

UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ELCHE

GRADO EN INGENIERÍA ELECTRÓNICA Y  
AUTOMÁTICA INDUSTRIAL



**UNIVERSITAS**  
*Miguel Hernández*

"ANÁLISIS DE LOS PATRONES DE  
CONSUMO ELÉCTRICO EN ESPAÑA"

TRABAJO FIN DE GRADO

Junio -2025

AUTOR: Pedro José Bellón Ibarra

DIRECTOR: Miguel López García

## ÍNDICE GENERAL

### Contenido

1. INTRODUCCIÓN .....	4
2. MATERIALES Y MÉTODOS .....	10
2.1. ENTORNO DE ANÁLISIS .....	10
2.2. DATOS.....	10
2.3. METODOLOGÍA.....	14
3. ANÁLISIS.....	16
3.1. ERROR DEL MODELO.....	17
3.2. NÚMERO DE DÍAS CON CIELO CUBIERTO .....	18
3.3. NÚMERO DE DÍAS NUBOSOS .....	23
3.4. RADIACIÓN GLOBAL .....	27
3.5. HORAS DE SOL MENSUALES.....	31
3.6. HORAS DE SOL DIARIAS .....	36
3.7. PRECIPITACIÓN .....	41
3.8. RADIACIÓN SOLAR NETA SOBRE LA SUPERFICIE .....	45
3.9. RADIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL.....	49
3.10. RADIACIÓN TOTAL.....	53
3.11. RADIACIÓN DIRECTA .....	61
3.12. RADIACIÓN DIFUSA .....	69
3.13. RADIACIÓN REFLEJADA .....	77
3.14. POTENCIA FOTOVOLTAICA ESTIMADA.....	85
3.15. GENERACION_FOTOVOLTAICA_ESIOS .....	93
4. SIMULACIÓN DEL MODELO CON NUEVAS VARIABLES.....	97
4.1. HORAS DE SOL DIARIAS .....	98
4.2. RADIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL.....	100
4.3. RADIACIÓN TOTAL.....	102
4.4. RADIACIÓN DIRECTA .....	104
4.5. RADIACIÓN DIFUSA .....	106
4.6. RADIACIÓN REFLEJADA .....	108
4.7. POTENCIA FOTOVOLTAICA ESTIMADA.....	110
4.8. RESULTADOS .....	112
5. SIMULACIÓN DEL MODELO CON NUEVAS VARIABLES A NIVEL HORARIO.....	113
5.1. RADIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL.....	114
5.2. POTENCIA FOTOVOLTAICA ESTIMADA.....	116

5.3. RADIACIÓN TOTAL.....	118
5.4. RESULTADOS .....	120
6. CONCLUSIONES .....	121
7. BIBLIOGRAFÍA.....	123



## 1. INTRODUCCIÓN

El pronóstico de cargas a corto plazo (STLF, short-term load forecasting) es una parte fundamental del funcionamiento de los sistemas de transmisión eléctricos para llevar a cabo una gestión eficiente de su operativa y garantizar el suministro eléctrico. Se trata de modelos que ofrecen previsiones en un rango que va desde minutos hasta varios días y que son utilizadas por los diferentes agentes que participan del sistema eléctrico ya sea como operadores de la red, comercializadores o productores de electricidad.

Los generadores de electricidad utilizan estos modelos para optimizar su producción incurriendo en menores costes de producción y mayor competitividad, lo que redonda en beneficios para el usuario. Su uso también reduce las emisiones de CO<sub>2</sub> a la atmósfera gracias a que permite una mayor utilización de energías renovables – especialmente solar y eólica - que por su naturaleza dependiente de las condiciones meteorológicas hacen necesario el uso de fuentes de respaldo, normalmente basadas en el uso de combustibles fósiles. Esto es especialmente relevante en Europa, que carece de fuentes propias de importancia de estos combustibles, lo que da lugar a debilidades políticas y económicas que se han materializado en los últimos años.

Por su parte, los operadores del sistema eléctrico (en el caso de España, Red Eléctrica Española, REE) también hacen uso de estos pronósticos para asegurar la estabilidad de la red. Una demanda superior a la generación de energía puede llevar a cortes de suministro y una sobrecarga de la red puede dañar infraestructuras y equipos y provocar apagones. Además, la previsión de la demanda permite también optimizar el mantenimiento de la red.

Debido a estas razones y aunque el STLF ha sido objeto de investigación durante años el interés por desarrollar métodos cada vez más sofisticados y precisos de previsión es persistente en el tiempo. Su relevancia sigue vigente debido a los continuos cambios tecnológicos, económicos, sociales y políticos. La extensión y el desarrollo de las fuentes de energía renovables, especialmente el abaratamiento de las placas solares fotovoltaicas, la relevancia del suministro de energía en el contexto internacional antes mencionado, la generalización del uso del aire acondicionado o el uso creciente del

coche eléctrico y los desafíos que plantea para la red eléctrica son algunos de los cambios que justifican este interés.

En cuanto a los modelos empleados para el STLF, estos han evolucionado desde métodos estadísticos clásicos hasta técnicas de inteligencia artificial. Los modelos estadísticos incluyen las regresiones lineales, métodos de suavizado exponencial o el análisis de series temporales. Estos modelos producen resultados precisos y se siguen utilizando hoy en día. Las técnicas de inteligencia artificial incluyen diferentes tipos de redes neuronales, algoritmos genéticos, máquinas de vector soporte, lógica difusa (fuzzy) o mapas autoorganizados. Finalmente tenemos también los sistemas híbridos que combinan los dos sistemas antes mencionados en diferentes etapas del proceso de generación de pronósticos [3].

No obstante, los métodos empleados no son el único aspecto que considerar a la hora de desarrollar sistemas de predicción. Técnicas de procesamiento de datos como la normalización, clusterización o filtrado de valores atípicos son también fundamentales. Además, las características de los datos de entrada y el sistema sobre el que queramos trabajar también se deben tener en cuenta a la hora de elegir los métodos a utilizar por lo que no es posible afirmar que haya una técnica mejor que el resto.

El punto de partida del presente trabajo es el modelo STLF que está en uso en Red Eléctrica Española, el operador que gestiona la red de transporte de energía eléctrica de alta tensión en España, y en cuyo desarrollo han participado investigadores de esta universidad. Se trata de un sistema híbrido que consiste en combinación lineal optimizada de una red neuronal con un modelo autorregresivo que produce pronósticos horarios de carga para cada una de las 18 regiones definidas por REE y un agregado a nivel nacional con 9 días de antelación [4].

Este modelo ha ido evolucionando a lo largo del tiempo. Las variables que alimentaban al modelo original son los datos de carga del sistema nacional de electricidad, información meteorológica y clasificación de días según calendario.

Los datos de carga son utilizados después de tratamiento para filtrar las anomalías. Inicialmente el modelo contenía dos variables para modelar las tendencias a largo plazo

de la carga como una función del tiempo, pero luego se vio que para periodos mayores de tres años esta aproximación perjudicaba la precisión del modelo por lo que fueron sustituidas por una media móvil de 52 semanas. Adicionalmente, el modelo incluía la última carga conocida.

La información meteorológica consiste en máximos y mínimos diarios de temperatura que se han preprocesado utilizando los conceptos de grados día de calefacción (HDD) y grados día de refrigeración (CDD) para tener en cuenta la relación no lineal entre temperatura y carga, ya que tanto temperaturas muy frías como muy cálidas incrementan el consumo de electricidad. Para modelar esta relación se definen dos umbrales y la temperatura se calcula como una función a trozos: una zona fría, otra neutral y otra caliente. Además, la temperatura de días pasados influye en la demanda de electricidad de los días siguientes, por lo que se incluyen valores de los últimos dos días. Por otra parte, no es deseable utilizar una media de temperaturas a nivel nacional ya que es una región que presenta diversidad climática y esta medida podría enmascarar eventos extremos a nivel local, por lo que se seleccionaron 5 estaciones meteorológicas (Barcelona, Bilbao, Madrid, Sevilla y Zaragoza) de las 59 disponibles. En total la información de temperatura añade 30 variables: el tratamiento temperaturas en HDD y CDD (3 valores, frío, neutral y caliente), el valor actual y dos valores pasados para cada una de las 5 localizaciones [5].

La clasificación de días según calendario es crítica a la hora de generar pronósticos de cargas. Dependiendo de si es entre semana o fin de semana o si es festivo o laborable la demanda de electricidad cambia. Las categorías (no exhaustivas) son: días festivos, días alrededor de los festivos (lunes antes de festivo, viernes después de festivo, etc.) o periodos especiales (semana antes de navidad, semana después de navidad, semana después de año nuevo, semanas de agosto, etc.). En el modelo los días se clasifican en 53 variables y se añaden otras 11 para asignar el mes [3].

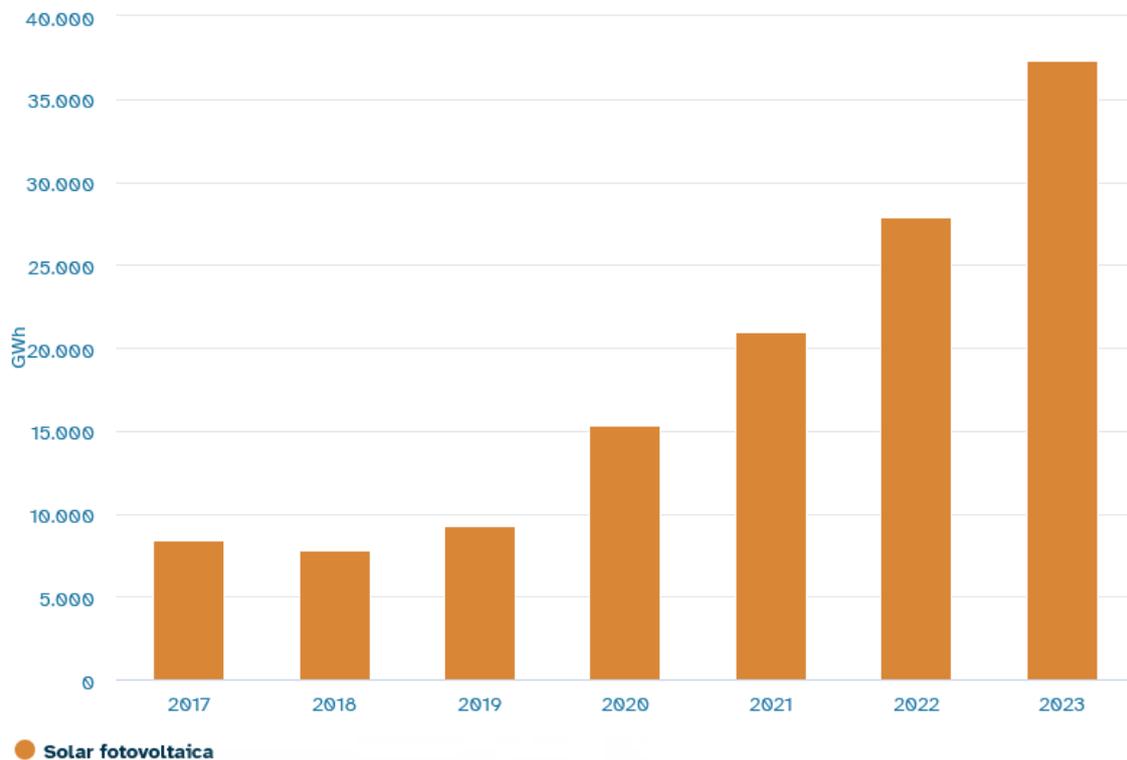
Se incluyen también en el modelo variables horarias relativas a la luz diurna ya que se vio que el error del modelo aumentaba alrededor de la salida y la puesta de sol, particularmente en las épocas del año en las que las horas de amanecer y el anochecer cambian rápido (alrededor de los equinoccios) o cuando se producen los cambios de

hora. Esto es especialmente importante en las horas cercanas la puesta de sol ya que estas horas coinciden con uno de los máximos de carga del día [5].

Como hemos mencionado, el modelo con las entradas de datos descritas es el que está actualmente en funcionamiento en REE y debido a que el error ha venido aumentando en los últimos años se hace necesario buscar las causas y plantear mejoras que permitan aumentar la precisión del modelo para que siga siendo útil.

Una de las posibles explicaciones es el aumento de instalaciones de placas fotovoltaicas en los hogares.

En los últimos años y especialmente desde 2018, la potencia solar fotovoltaica instalada en España prácticamente se ha cuadruplicado, pasando de los 8,7 MW en 2019 a 25,1 MW al cierre del año 2023 y 30,6 a finales del año 2024. En cuanto a la energía producida, a 5 de octubre del año 2024 ya se había superado la producción contabilizada el año anterior. En 2023 se generaron 37,4 GWh y en la citada fecha de 2024 se habían registrado 37,5 GWh, un 19% de aumento respecto del periodo anterior. Alcanzando también el 18,3% de la energía total producida en el *mix*, siendo la tercera fuente de electricidad del país [8]. Se prevé que esta tendencia continúe en el futuro.



Fuente: ree.es

Aunque las instalaciones de autoconsumo no están registradas por el operador ya que este contabiliza la energía producida mediante la generación en plantas solares, es razonable pensar que su número ha aumentado en paralelo a la potencia instalada en estas plantas gracias a su abaratamiento en los últimos años, la idoneidad del clima del país, el impulso político mediante ayudas económicas y regulaciones, la difusión mediática y la proliferación de empresas dedicadas a tal fin.

La producción de energía fotovoltaica mediante paneles solares está condicionada por diferentes factores, algunos ya mencionados. Entre ellos se encuentran los ambientales como la irradiancia solar (definida como la cantidad de energía procedente del Sol que llega a una superficie específica en un periodo determinado medida en  $W/m^2$ ), la temperatura, el clima, las condiciones meteorológicas o la contaminación. Tenemos también los factores técnicos como el tipo de paneles instalados, su mantenimiento, su degradación o la calidad y cantidad de equipos auxiliares, aunque estos son propios de cada instalación más allá de las regulaciones mientras que los ambientales son generales.

El objetivo del presente trabajo es investigar la posible relación entre el aumento del error del modelo y variables ambientales que puedan afectar a la generación de energía fotovoltaica como paso previo a nuevas iteraciones del modelo que tengan en cuenta estos efectos.



## 2. MATERIALES Y MÉTODOS

### 2.1. ENTORNO DE ANÁLISIS

El análisis se ha llevado a cabo en Google Colab, una plataforma en la nube basada en Jupyter que nos permite ejecutar código Python de forma interactiva y en la que además están disponibles las librerías necesarias.

En cuanto a las librerías de Python, se ha utilizado Pandas para manipulación, exploración y análisis de datos y Sklearn (Scikit-learn) para simular el funcionamiento del modelo STLF mediante una regresión lineal. Para generar los gráficos la herramienta de elección ha sido Matplotlib.

También se han ejecutado scripts en entorno local para descargar datos mediante las APIs de AEMET y ERA5.

### 2.2. DATOS

Se han utilizado diversos conjuntos de datos para analizar la relación entre el error del modelo STLF de REE y variables meteorológicas que condicionan la producción de energía fotovoltaica como son irradiancia solar y las relativas a condiciones atmosféricas. La temperatura ya se tiene en cuenta en el modelo y otros factores que puedan afectar a la generación fotovoltaica son propios de cada instalación.

Además de los datos del modelo proporcionados por el tutor del trabajo y participante en su desarrollo se han seleccionado fuentes de datos fiables de organismos reconocidos como como AEMET, PVGIS y ERA5.

#### 2.2.1. DATOS MODELO

Los datos del modelo comprenden tanto la entrada como la salida del mismo en un rango temporal que va desde el 1 de enero de 2013 hasta el 30 de abril de 2022.

Los datos de entrada consisten en un vector de 96 matrices, una para cada cuarto de hora del día. A su vez en cada matriz hay 102 variables y datos de 3407 días que comprenden el rango temporal descrito anteriormente. Entre estas variables se

encuentran la media móvil de 52 semanas de la carga y las 53 variables de clasificación de días y 30 de temperatura descritas junto con el modelo en la introducción.

Los datos de salida son otros 96 vectores con la demanda de energía en MW de cada cuarto de hora correspondiente para 3407 días.

Los ficheros que contenían estos datos se han procesado con Matlab para generar un fichero en formato CSV para cada cuarto de hora del día. Posteriormente se han filtrado todos los días especiales ya que para el análisis solo se utilizan días que no presenten características especiales.

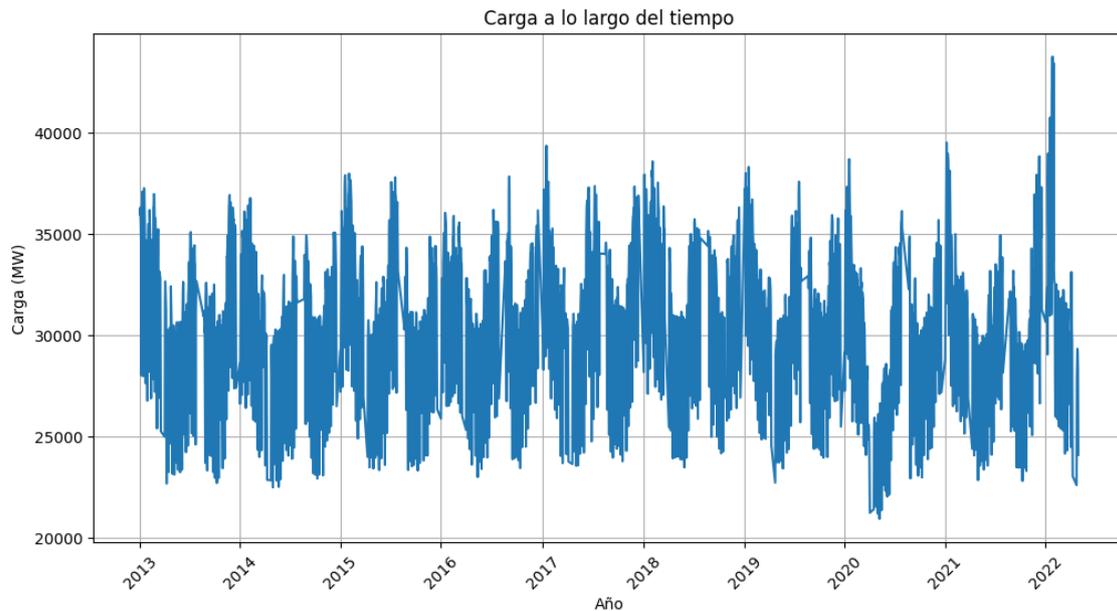
Para análisis diarios, en las primeras iteraciones se utilizaron datos entre las 11:00h y las 17:00h y en las siguientes se amplió el rango a entre las 08:00h y las 20:00h para abarcar todos los momentos de luz diurna. Para el análisis horario los datos se agregan por horas.

Una vez seleccionado el rango temporal diario se obtiene la media de cada variable agregando los datos por día.

El conjunto de datos final contiene:

- Año
- Mes
- Día
- Día de la semana
- Hora (si procede)
- Media móvil de 52 semanas de la carga
- 30 variables de temperatura
- Salida del modelo

En la siguiente figura se muestra la distribución de la carga a lo largo del tiempo en número de días desde el inicio del conjunto de datos:



### 2.2.2. DATOS AEMET

Se han empleado datos meteorológicos de la AEMET (Agencia Estatal de Meteorología) tanto diarios como agregados mensuales. Los datos diarios contienen valores de insolación (horas de sol con cielo despejado) y cantidad de precipitación y se han obtenido mediante la API de datos abiertos de la AEMET para las 5 localizaciones utilizadas en el modelo: Bilbao, Barcelona, Madrid, Sevilla y Zaragoza. Los agregados mensuales se han obtenido de la Estadística de las variables meteorofenológicas de la web de la AEMET para las mismas 5 localizaciones. Consisten en medias mensuales de insolación, radiación global, número de días cubiertos y número de días nublados.

### 2.2.3. PVGIS

PVGIS (Photovoltaic Geographical Information System) es una herramienta desarrollada por el JRC (Joint Research Centre) de la Comisión Europea que contiene información de radiación solar y rendimiento de sistemas fotovoltaicos para todo el mundo excepto los polos. Se han seleccionado las 5 localizaciones referidas anteriormente (Barcelona, Bilbao, Madrid, Sevilla y Zaragoza).

La configuración utilizada es la siguiente:

Se han utilizado las bases de datos SARAH3 y ERA5. La primera de ellas contiene observaciones satelitales y la segunda es un reanálisis. Un reanálisis es una combinación de observaciones meteorológicas históricas con modelos numéricos para generar representaciones climáticas.

Con ello se obtienen datos horarios de:

- P: Potencia del sistema fotovoltaico (W)
- Gb(i): Irradiancia directa (radiación directa) sobre el plano inclinado (plano del sistema) (W/m<sup>2</sup>)
- Gd(i): Irradiancia difusa sobre el plano inclinado (plano del sistema) (W/m<sup>2</sup>)
- Gr(i): Irradiancia reflejada sobre el plano inclinado (plano del sistema) (W/m<sup>2</sup>)

Que luego serán filtrados para reflejar las mismas horas que los datos del modelo (08:00h a 20:00h) y agregados a nivel diario.

#### 2.2.4. ERA5

ERA5 es un conjunto de datos globales de clima y meteorología de reanálisis producido por el ECMWF (Centro Europeo de Previsiones Meteorológicas a Medio Plazo) que forma parte del programa Copérnico de la Unión Europea.

Los datos se han obtenido a través de la API de Climate Data Store (CDS) de Copérnico. Los datos están disponibles por superficie por lo que se han delimitado

puntos de la mínima resolución posible para obtener datos horarios de las 5 localizaciones del trabajo.

Las variables utilizadas en el trabajo son:

- Radiación solar neta sobre la superficie (Surface net solar radiation) ( $\text{J/m}^2$ )
- Radiación solar sobre el plano horizontal (Surface solar radiation downwards) ( $\text{J/m}^2$ )

Estos datos han sido posteriormente convertidos a  $\text{W/m}^2$ , filtrados y agregados a nivel diario para reflejar los mismos intervalos temporales que los datos del modelo.

### 2.2.5. ESIOS

El Sistema de Información del Operador del Sistema (ESIOS) es una plataforma gestionada por REE que ofrece datos del mercado y del sistema eléctrico. Entre los datos disponibles de generación y consumo se han utilizado datos de generación de energía solar fotovoltaica medida en tiempo real.

### 2.3.METODOLOGÍA

Para el análisis se ha buscado replicar el modelo en funcionamiento en REE. Para ello se ha llevado a cabo el siguiente tratamiento de datos:

- Se descartan los datos de 2013 y se utilizan datos a partir de 2014 ya que la media móvil de la carga es de 52 semanas.
- Los datos en origen son horarios por lo que se filtra para obtener aquellos registros entre las 08h y las 20h, que son los de las horas de sol de todo el año, y posteriormente se agregan de forma diaria.
- Se filtra el periodo de la pandemia del COVID-19 y el confinamiento en España: entre 1 de marzo y el 1 de junio de 2020.
- Para evitar factores de confusión se han filtrado los días especiales

Una vez tratados los datos se utiliza un modelo de regresión lineal ya que con esta técnica conseguimos una precisión suficiente, un error relativamente pequeño, para el análisis.

Se predice el consumo de los años 2017, 2018, 2019, 2020, 2021 y 2022 (este último parcialmente pues solo tenemos datos hasta el 30 de abril) utilizando una división de los

conjuntos de datos de 3:1, es decir, tres años para entrenar el modelo y uno de predicción o test. Por ejemplo, para predecir el consumo del año 2017 utilizamos los datos de 2014, 2015 y 2016.

Una vez se obtiene la predicción de carga se compara con los datos reales y se calcula el error como la diferencia entre la realidad y la predicción, que posteriormente se pone en relación con las variables meteorológicas descritas en diagramas de dispersión. Queremos encontrar posibles relaciones lineales por lo que se compara el error del modelo en el eje de ordenadas con la magnitud de cada variable en el de abscisas y se calculan las pendientes de las rectas de regresión.

$$Error = P_{real} - P_{predicción}$$

Un error negativo quiere decir que la carga predicha es mayor que la real y que por lo tanto existe sobreestimación de la carga, que es el efecto cuya relación con las variables descritas queremos investigar.

En el caso opuesto, el de infraestimación, la carga predicha es menor que la real y el error es positivo.

En el caso de análisis en agregado para la comparación entre modelos la métrica utilizada es el RMSPE.

Con el fin de facilitar la interpretabilidad del análisis los puntos de datos se han codificado por colores:

- Enero, febrero, marzo y diciembre: azul
- Abril y noviembre: verde
- Junio, julio, agosto y septiembre: rojo
- Mayo y octubre: naranja

De esta forma se puede observar de forma visual la distribución de los diferentes puntos para identificar patrones, distinguir épocas del año, etcétera.

### 3. ANÁLISIS

Se procede en esta sección al análisis de las siguientes variables:

Variable	Descripción
n_cub	Número de días con el cielo cubierto por mes.
n_nub	Número de días nublados por mes.
glo	Radiación global mensual (MJ/m <sup>2</sup> ).
inso	Media mensual de horas de sol diarias.
sol	Horas de sol diarias.
prec	Cantidad de precipitación diaria (mm).
ssr_wm2	Radiación solar neta sobre la superficie (W/m <sup>2</sup> )
ssrd_wm2	Radiación solar sobre el plano horizontal (W/m <sup>2</sup> )
total_radiation_pvgis	Radiación solar total (W/m <sup>2</sup> ) (PVGIS)
total_radiation_pvgis_e5	Radiación solar total (W/m <sup>2</sup> ) (PVGIS E5)
Gb(i)_pvgis	Radiación directa (W/m <sup>2</sup> ) (PVGIS)
Gb(i)_pvgis_e5	Radiación directa (W/m <sup>2</sup> ) (PVGIS E5)
Gd(i)_pvgis	Radiación difusa (W/m <sup>2</sup> ) (PVGIS)
Gd(i)_pvgis_e5	Radiación difusa (W/m <sup>2</sup> ) (PVGIS E5)
Gr(i)_pvgis	Radiación reflejada (W/m <sup>2</sup> ) (PVGIS)
Gr(i)_pvgis_e5	Radiación reflejada (W/m <sup>2</sup> ) (PVGIS E5)
P_pvgis	Potencia estimada fotovoltaica (W) (PVGIS)
P_pvgis_e5	Potencia estimada fotovoltaica (W) (PVGIS E5)
generacion_fotovoltaica_esios	Energía generada por instalaciones fotovoltaicas (MWh) (ESIOS)

Su origen y las transformaciones de datos están descritos en la sección 2.2 y la metodología seguida en la sección 2.3 así como en sus secciones respectivas.

### 3.1. ERROR DEL MODELO

Año	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	636.362	761.512	2.078	2.401	1.156
2018	637.628	711.998	2.065	2.267	1.098
2019	609.785	1099.742	1.972	3.641	1.846
2020	626.509	1063.770	2.028	3.667	1.808
2021	684.635	1097.356	2.264	3.439	1.519
2022	765.115	2524.357	2.484	6.838	2.752

El error en la simulación del funcionamiento del modelo de REE presenta un error creciente en el tiempo con los datos disponibles.

Se puede observar que el error en los conjuntos de entrenamiento es estable y relativamente preciso salvo un pequeño incremento en los dos últimos años por lo que el modelo está aprendiendo bien los patrones de los datos de entrada.

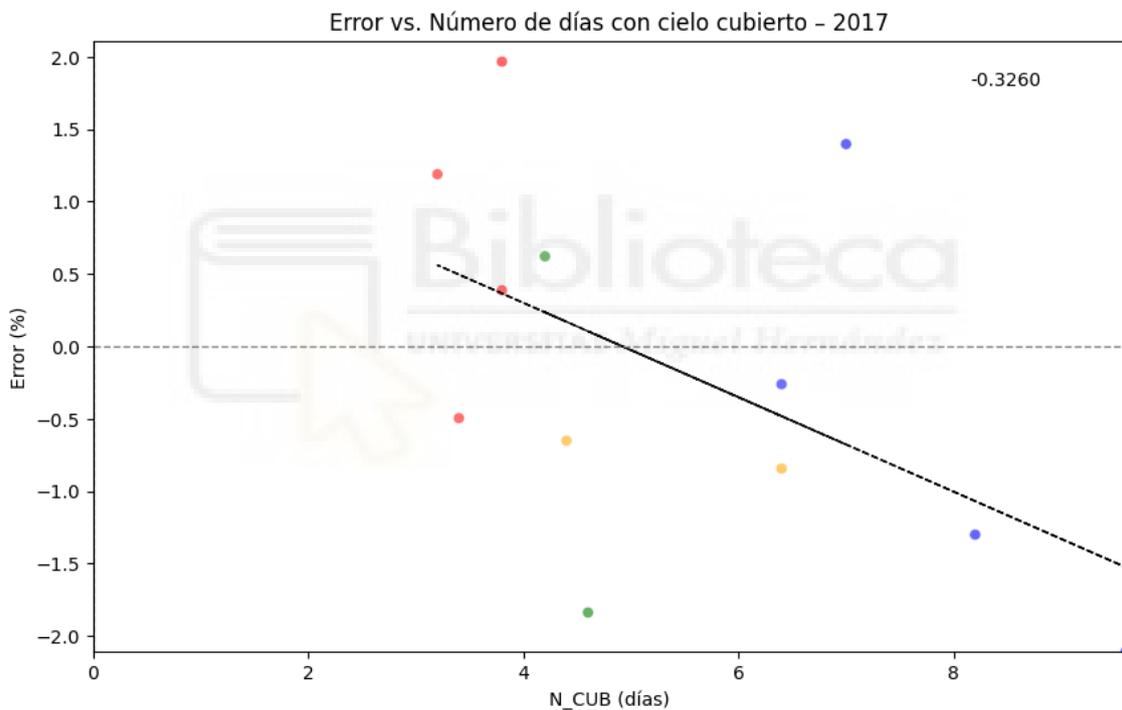
En las predicciones (test) el error es similar en 2017 y 2018 pero en 2019, 2020 y 2021 es mayor incrementándose en torno a un 40% respecto de los dos primeros años. En 2022 el error es significativamente mayor, aunque los datos para este año son incompletos.

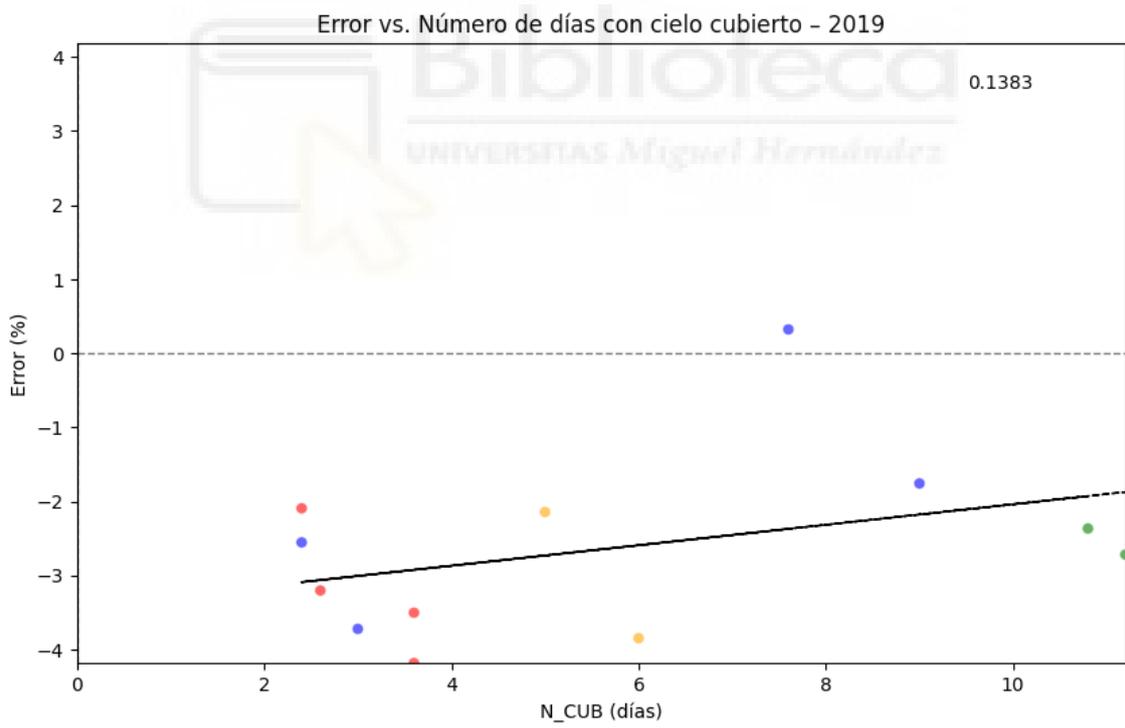
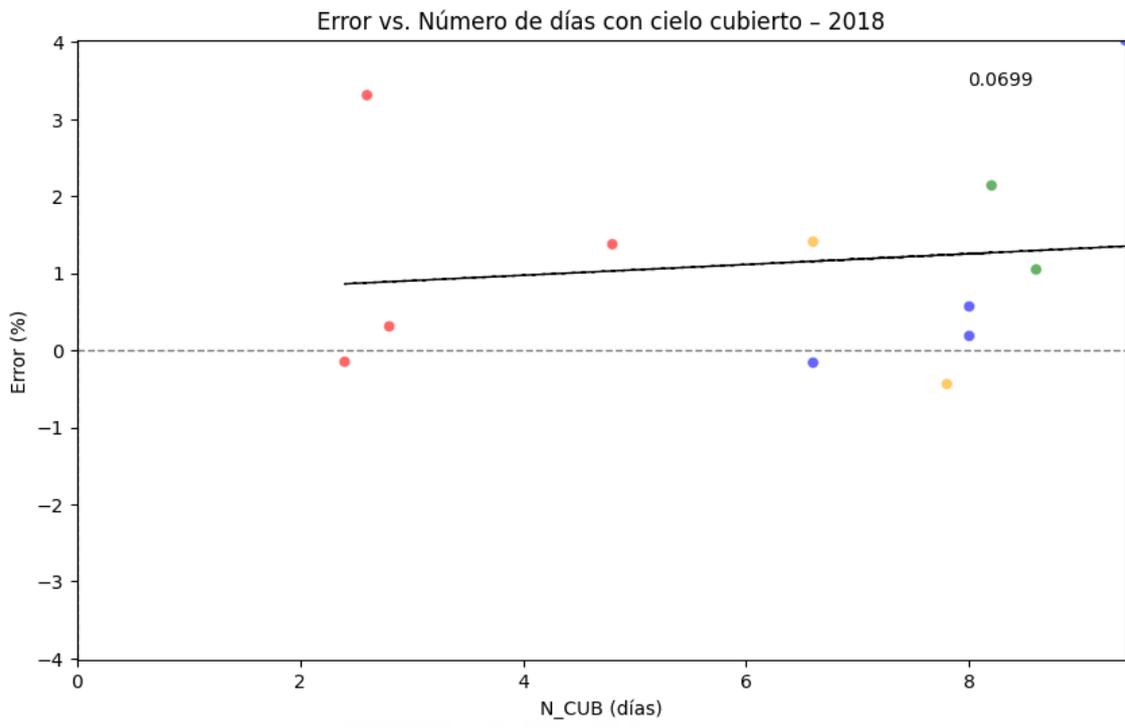
Esta tendencia creciente se refleja en la proporción entre el error del entrenamiento y el de predicción.

### 3.2. NÚMERO DE DÍAS CON CIELO CUBIERTO

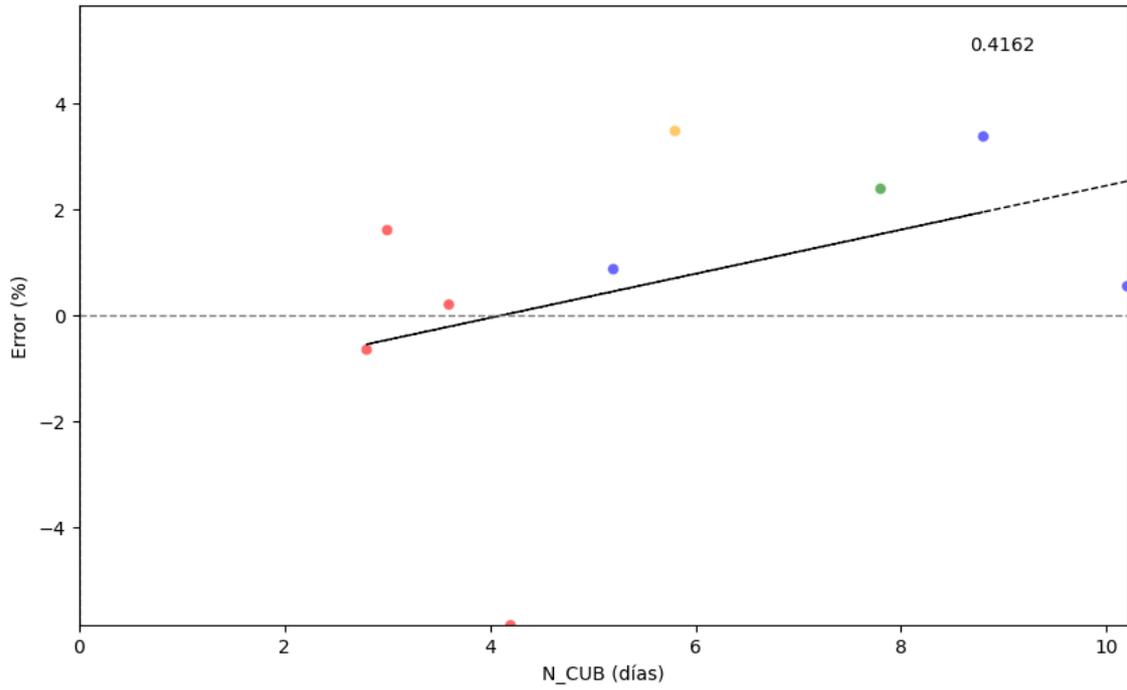
Esta variable (N\_CUB) hace referencia al número de días por mes con el cielo cubierto. En meteorología, para describir la nubosidad se observa el cielo, se divide imaginariamente en 8 partes (u octavas) y se cuenta el número de ellas en las que el cielo está tapado por las nubes. Si hay más de 6 se dice que el cielo está cubierto.

Los datos de esta variable son mensuales por lo que los resultados del modelo se han agregado a nivel mensual para poder realizar la comparación. Se procede luego a poner en relación el error con los valores de N\_CUB para cada año de cálculo en un gráfico de dispersión y se calcula la recta de regresión:

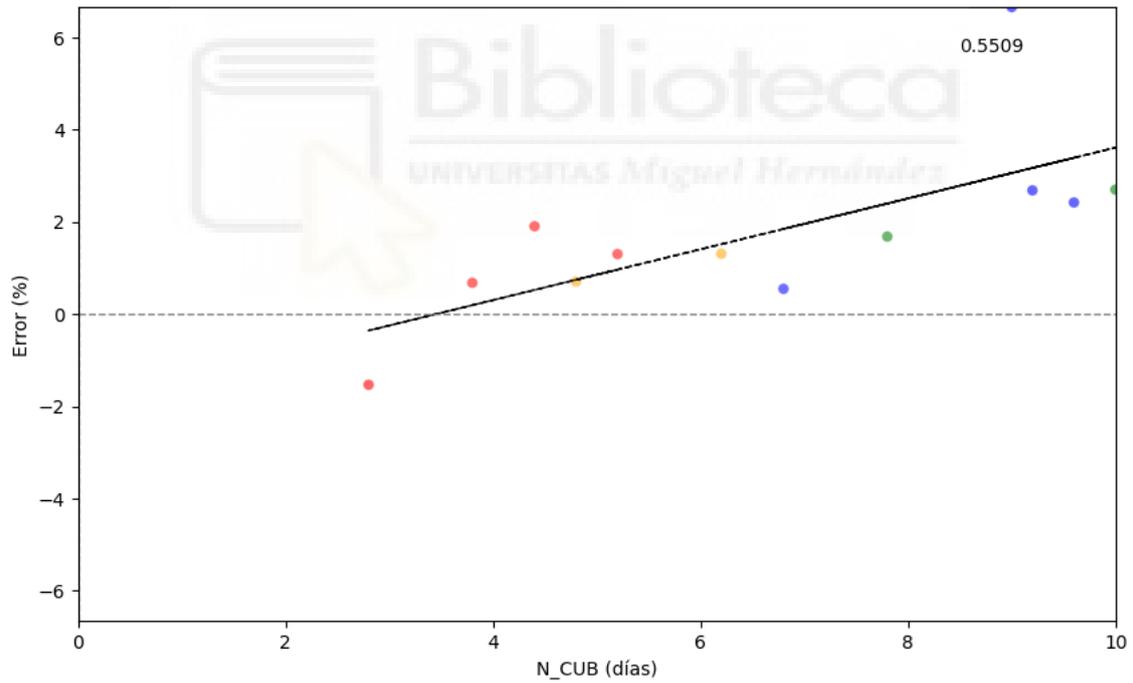


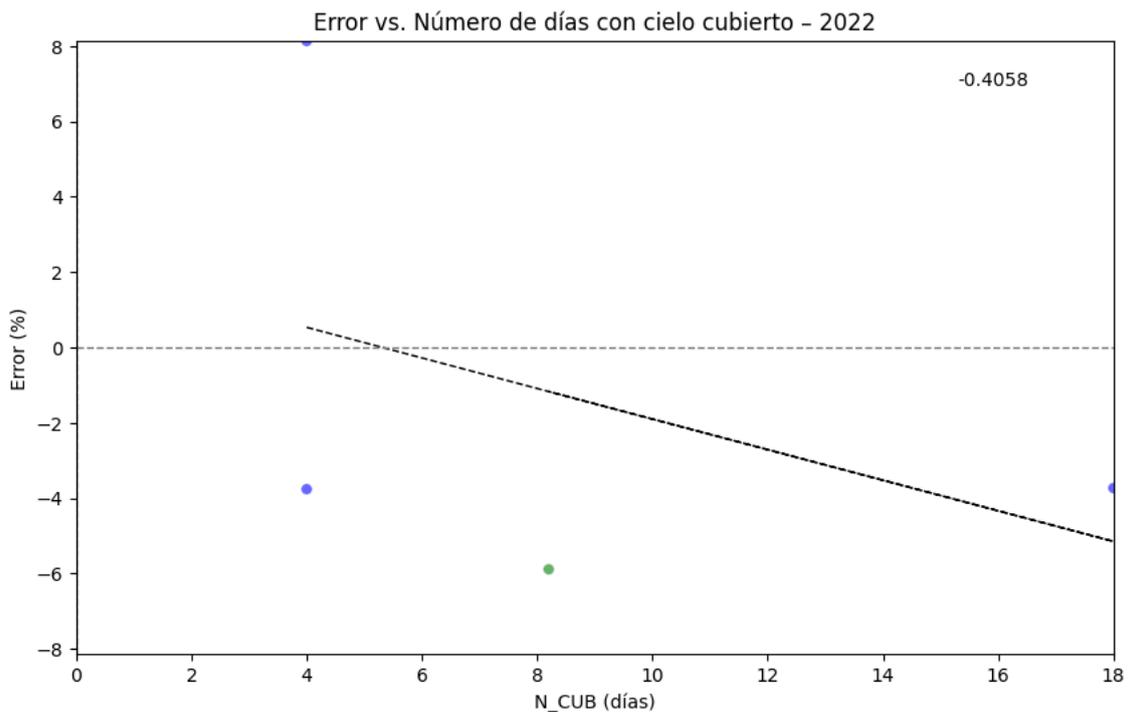


Error vs. Número de días con cielo cubierto - 2020



Error vs. Número de días con cielo cubierto - 2021





Año	Pendiente error vs. n_cub
2017	-0.326
2018	0.070
2019	0.138
2020	0.416
2021	0.551
2022	-0.406

En los gráficos podemos ver que la pendiente de la línea cambia pasando de ser negativa en 2017 a ser positiva desde 2018 hasta 2021 incrementándose la magnitud de esta. En 2022 solo tenemos datos de los meses de invierno (color azul) y la pendiente de la línea vuelve a ser negativa.

Entre 2018 y 2021 la relación entre el número de días nublados y el error es proporcional y hay una relación positiva entre el número de días nublados con una predicción de la carga del sistema eléctrico inferior a la que luego se produjo.

Los meses de verano se encuentran en la parte izquierda de las gráficas ya que en esta estación hay menos días nublados y el error está más cerca de la región negativa en esos meses. Es decir, unas condiciones favorables a la producción fotovoltaica son inversamente proporcionales a la sobreestimación de la carga y esta relación se hace más evidente en años más tardíos.

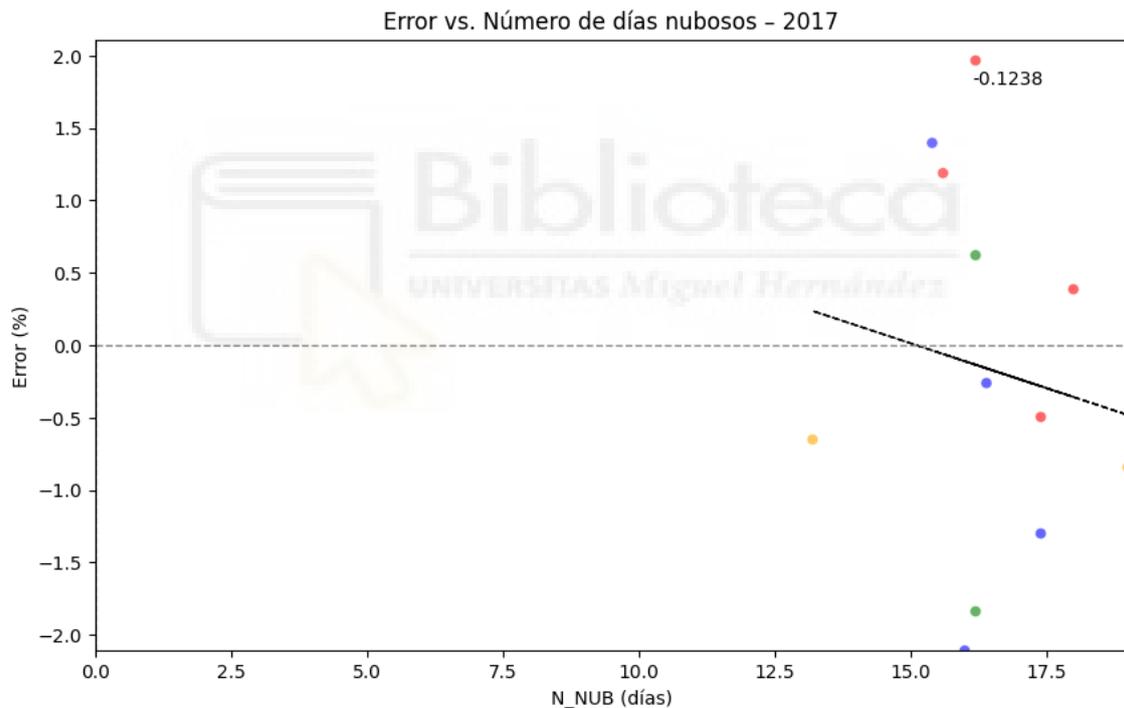
Es necesario apuntar que la correlación de la línea de regresión es baja dado que el número de puntos es reducido por ser esta variable de carácter mensual.

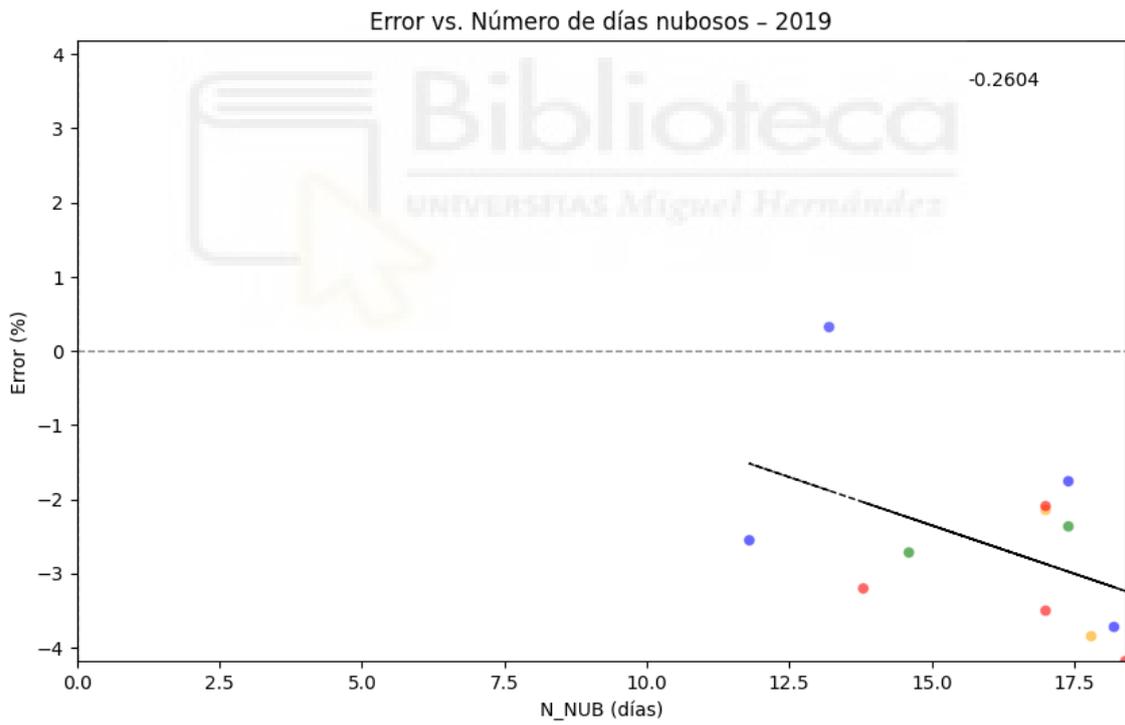
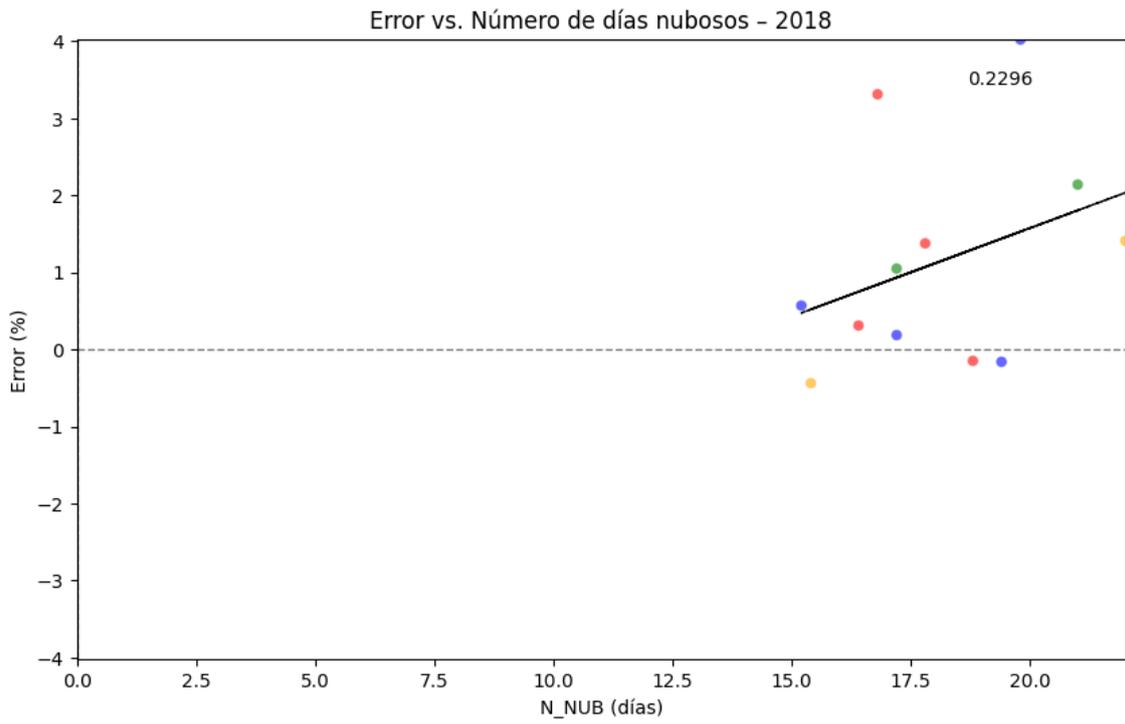


### 3.3. NÚMERO DE DÍAS NUBOSOS

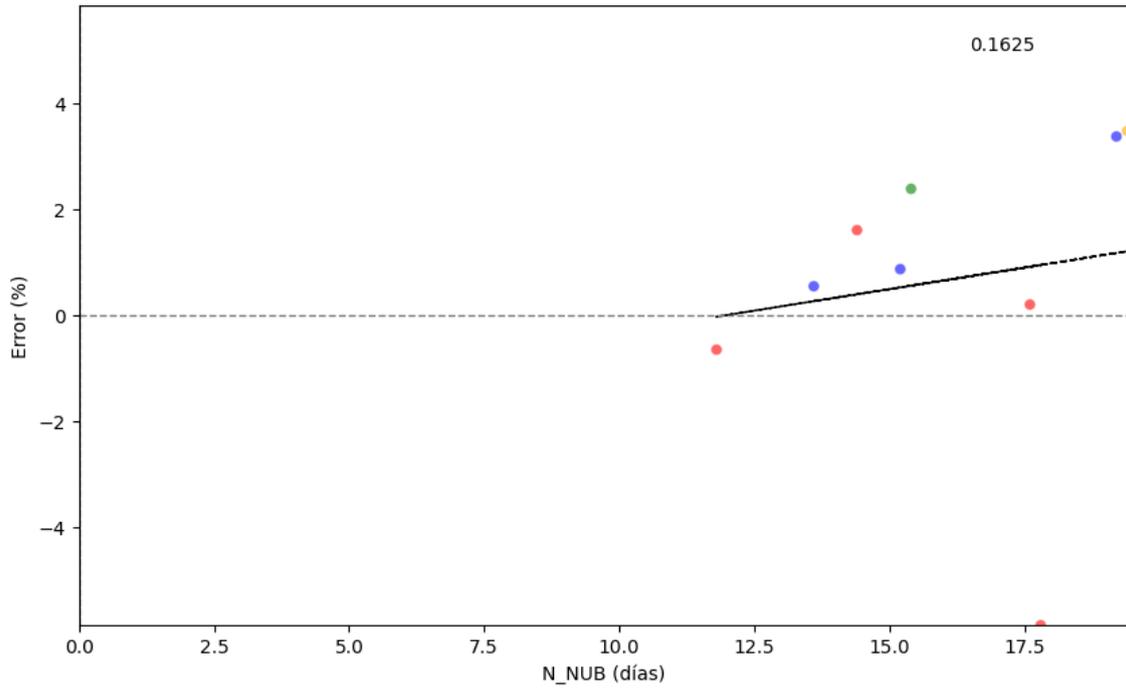
Esta variable (N\_NUB) hace referencia al número de días nublados cada mes. Utilizando la forma de medir la nubosidad de la variable N\_CUB (dividir el cielo imaginariamente en 8 partes) se dice que el cielo está nublado si entre 2 y 6 octavas del mismo están cubiertas por nubes.

Como en el caso anterior los datos de esta variable son mensuales por lo que se ha seguido la misma metodología de análisis: los resultados del modelo se han agregado a nivel mensual para poder realizar la comparación. Se procede luego a poner en relación el error con los valores de N\_NUB para cada año de cálculo en un gráfico de dispersión y se calcula la recta de regresión:

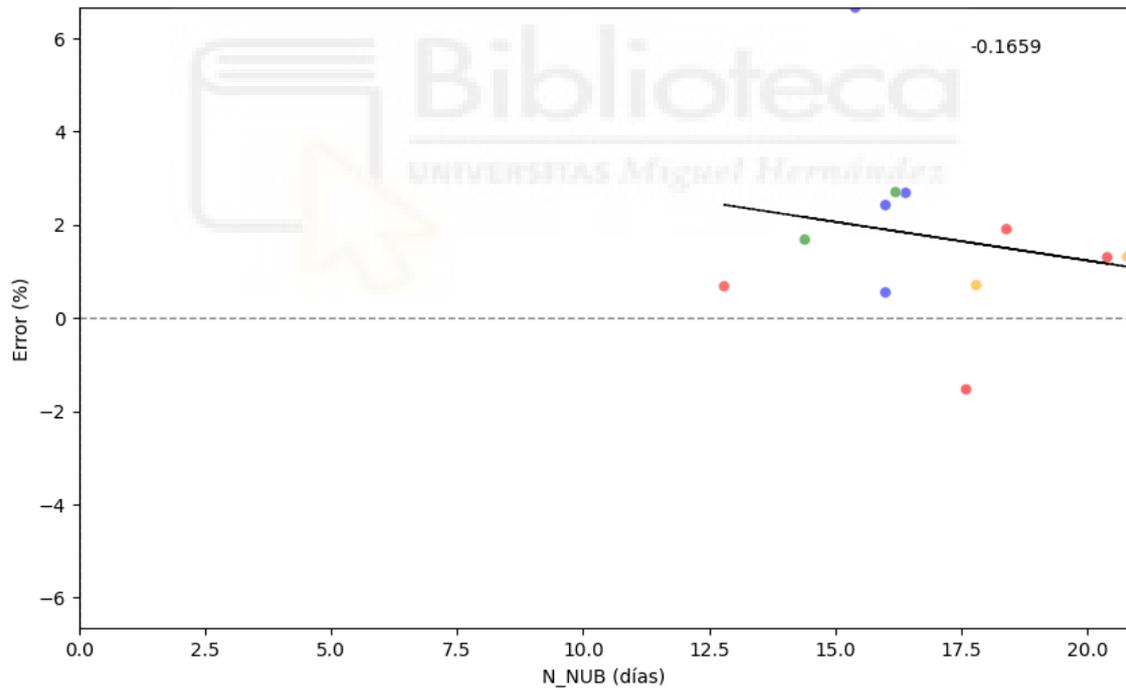


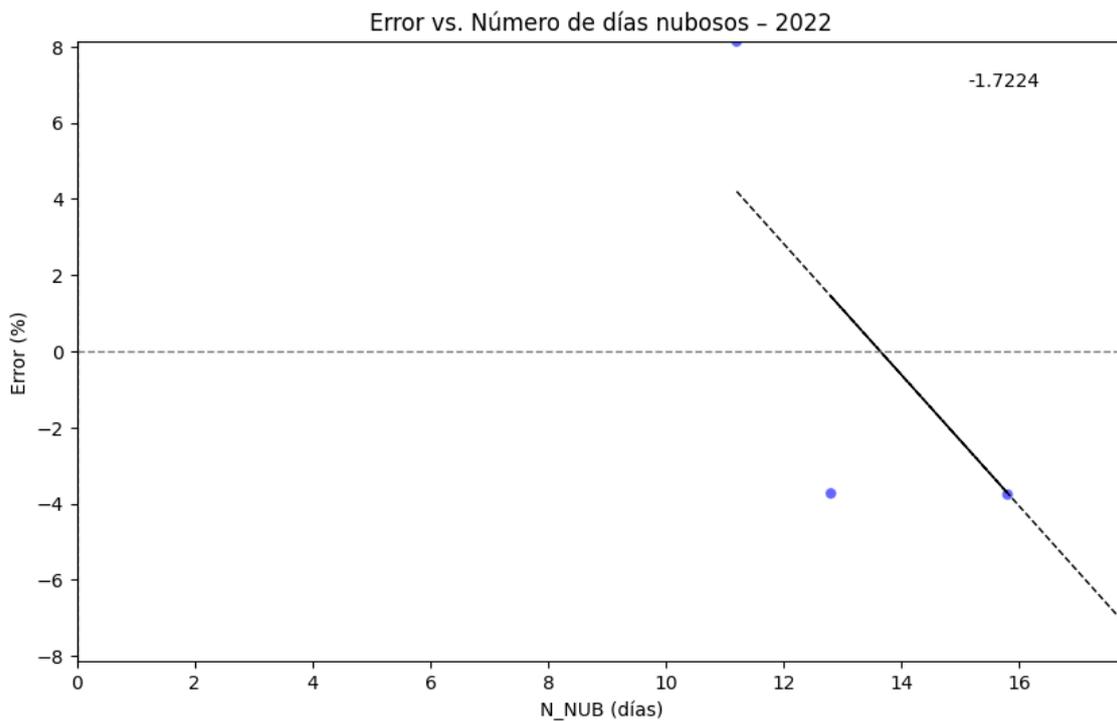


Error vs. Número de días nublados - 2020



Error vs. Número de días nublados - 2021





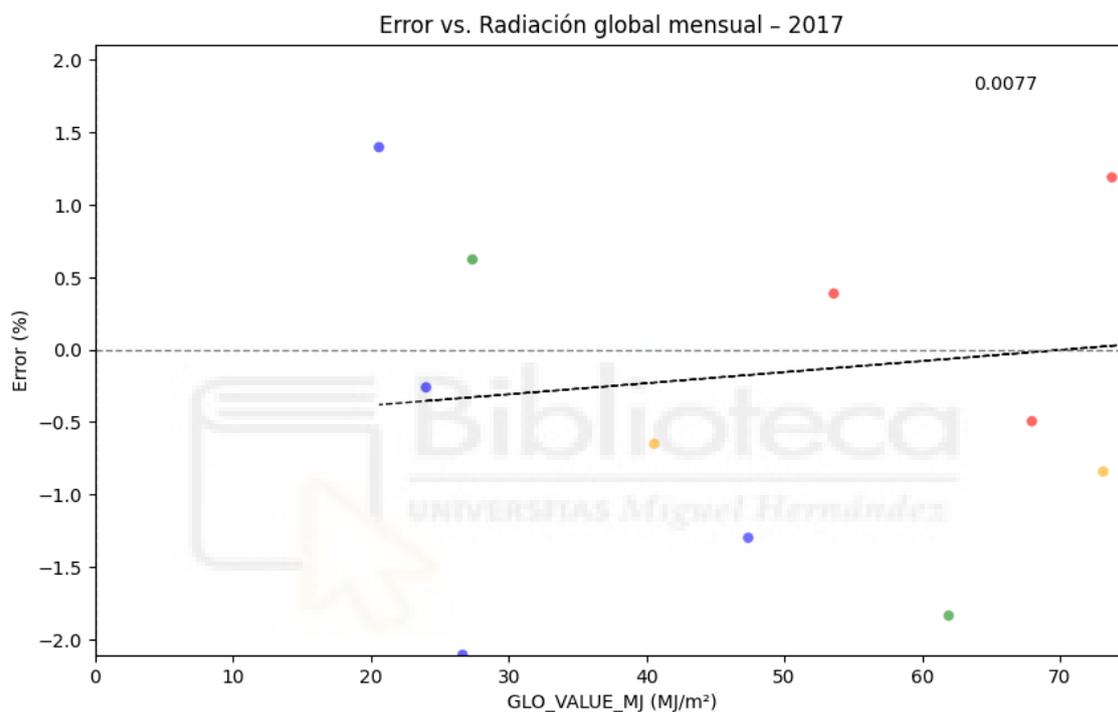
Año	Pendiente error vs. n_nub
2017	-0.124
2018	0.230
2019	-0.260
2020	0.163
2021	-0.166
2022	-1.722

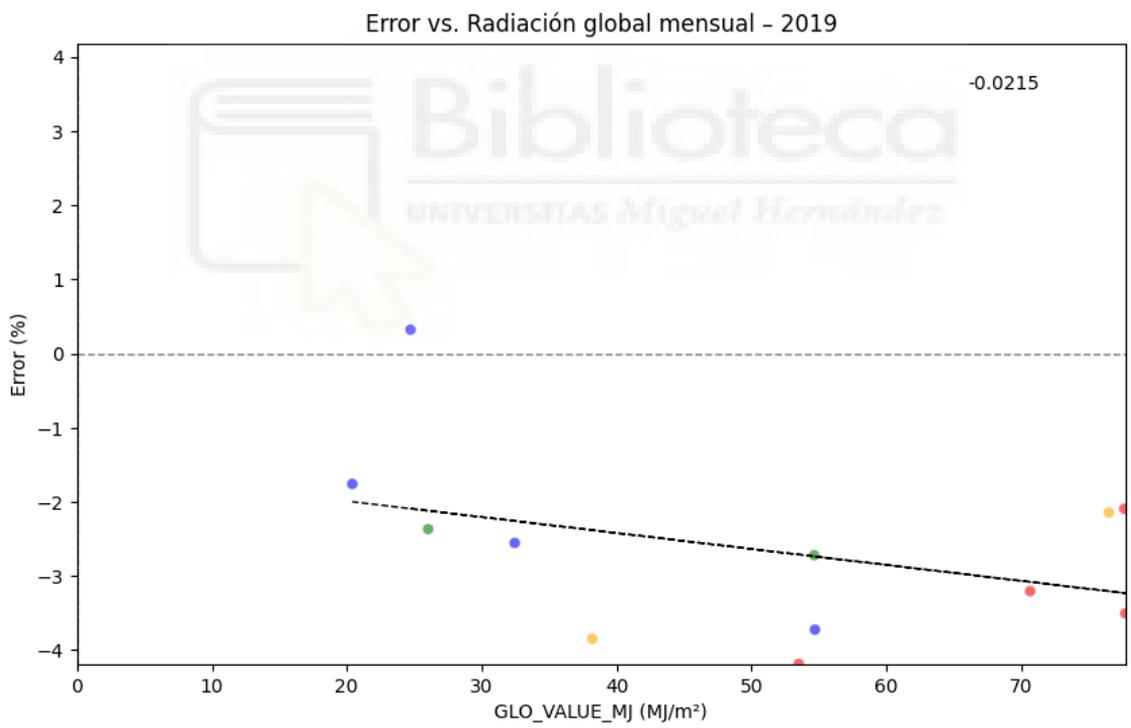
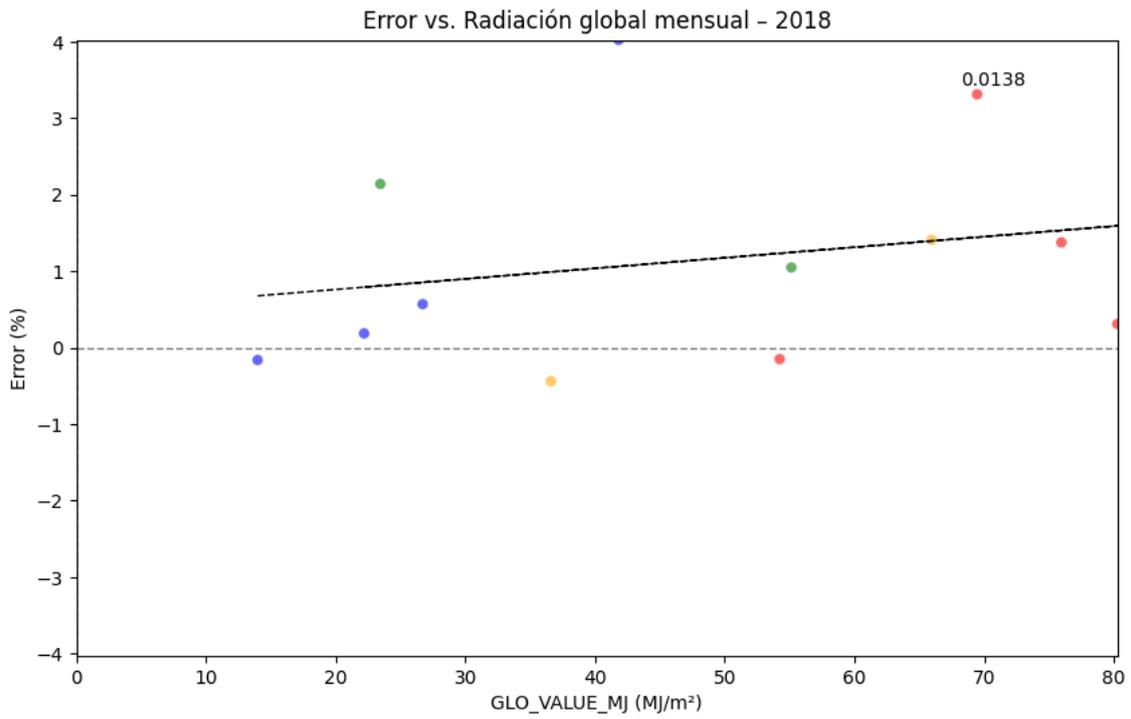
En el caso de esta variable no se observa ninguna correlación entre su magnitud y el error de acuerdo con la línea de regresión. Esto puede deberse a que esta variable abarca un abanico amplio de estados del cielo, lo cual es especialmente relevante en lo referente a la producción fotovoltaica. En el extremo inferior la cantidad de nubes puede que no sea suficiente para tener un efecto relevante en la misma y en el superior todavía permitiría el paso de la luz del sol.

### 3.4. RADIACIÓN GLOBAL

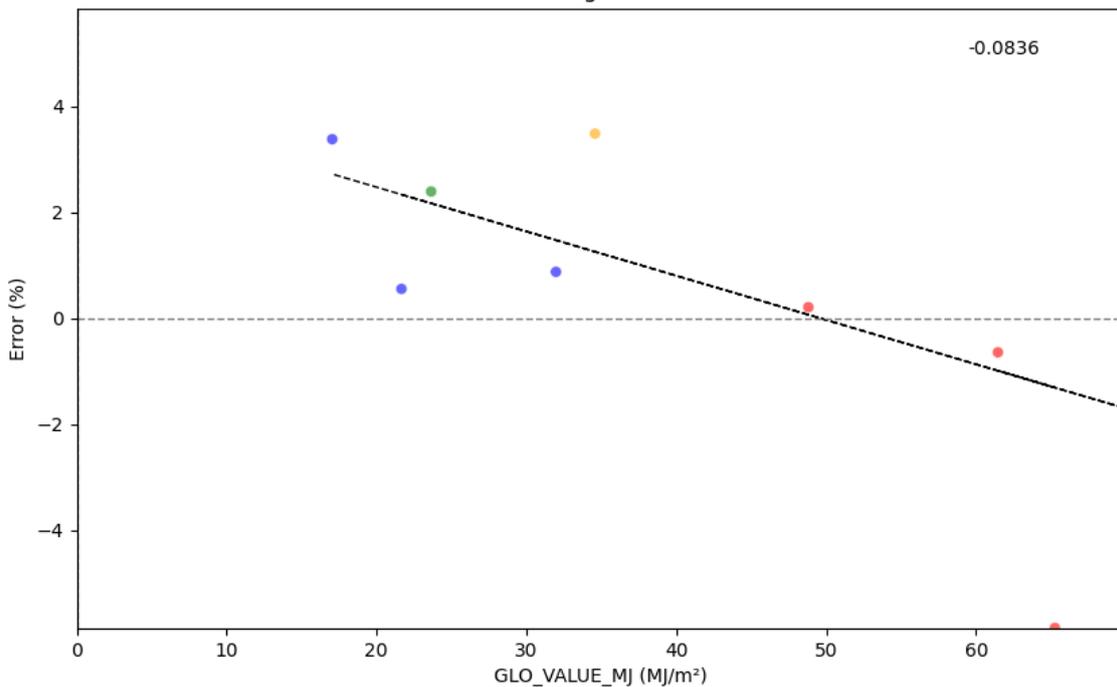
Esta variable (GLO) representa la radiación global mensual en MJ/m<sup>2</sup>.

Al igual que en los casos anteriores, esta variable presenta datos mensuales, por lo que se ha seguido la misma metodología de análisis. Los resultados del modelo se han agregado a nivel mensual y se examina la relación entre el error y los valores de esta variable para cada año en gráficos de dispersión.

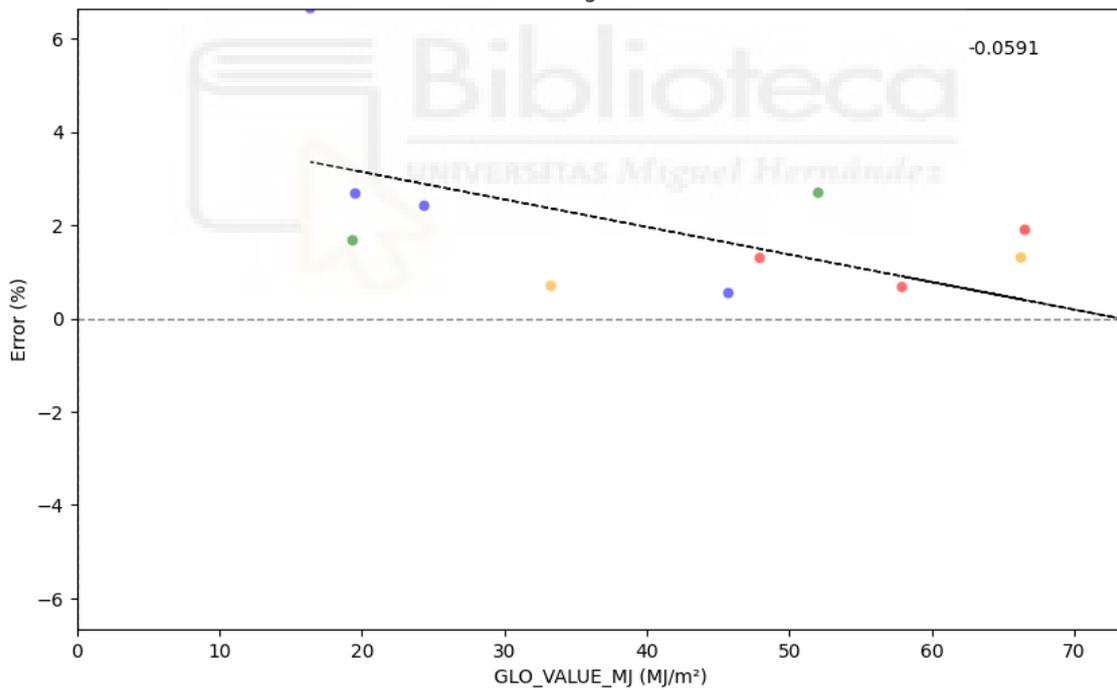


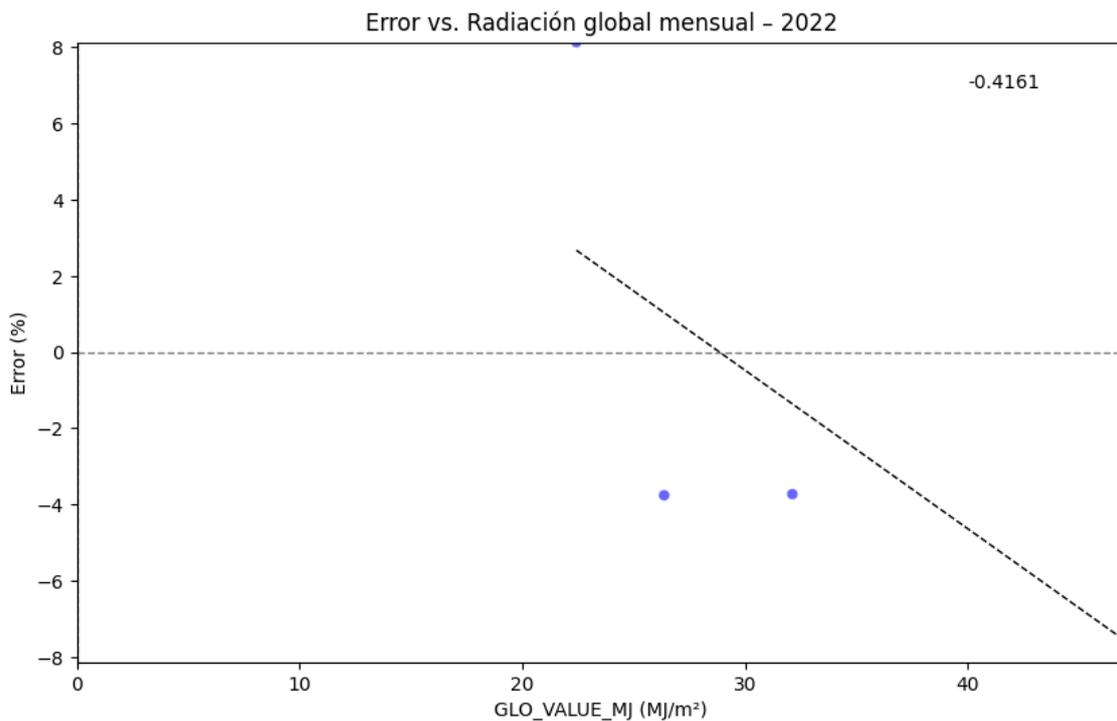


Error vs. Radiación global mensual - 2020



Error vs. Radiación global mensual - 2021





Año	Pendiente error vs. glo
2017	0.008
2018	0.014
2019	-0.021
2020	-0.084
2021	-0.059
2022	-0.416

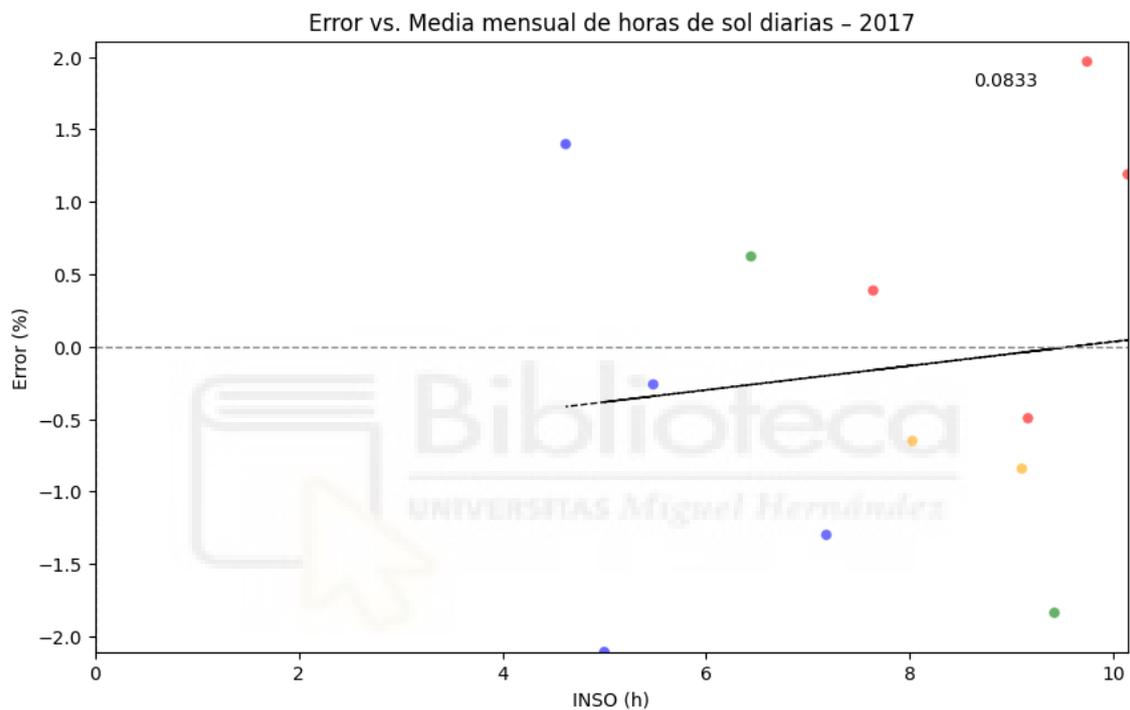
En las gráficas no se observa relación entre la magnitud de la variable en los años 2017 y 2018. Para los años 2019, 2020 y 2021 sí se aprecia una pendiente negativa de la línea de regresión entre la magnitud de la radiación y el error, siendo el error tendente al área negativa del diagrama (sobrestimación de la carga) en los meses de verano (color rojo) que coinciden con los de mayor radiación.

En el año 2022 la relación entre la radiación y el error también es inversamente proporcional pero solo contamos con datos de los 4 primeros meses del año.

### 3.5. HORAS DE SOL MENSUALES

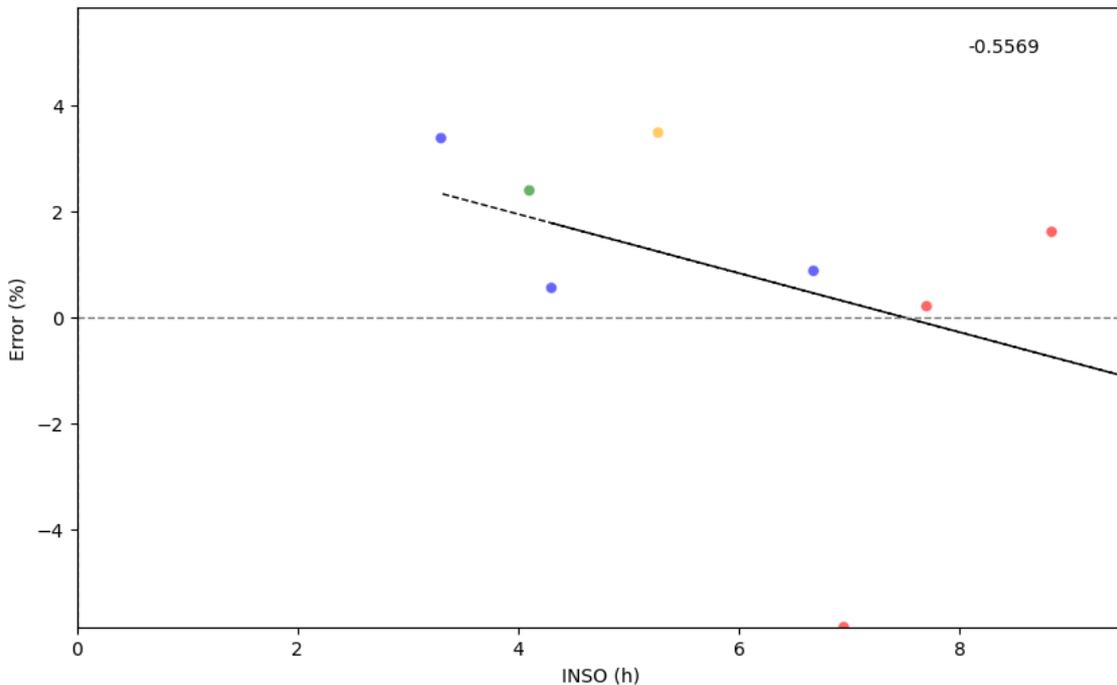
Esta variable (INSO) hace referencia a la media mensual de horas de sol diarias.

Como en las variables mensuales anteriores, se han agregado los datos a nivel mensual y se han puesto en relación al error del modelo para el año de cálculo en gráficos de dispersión.

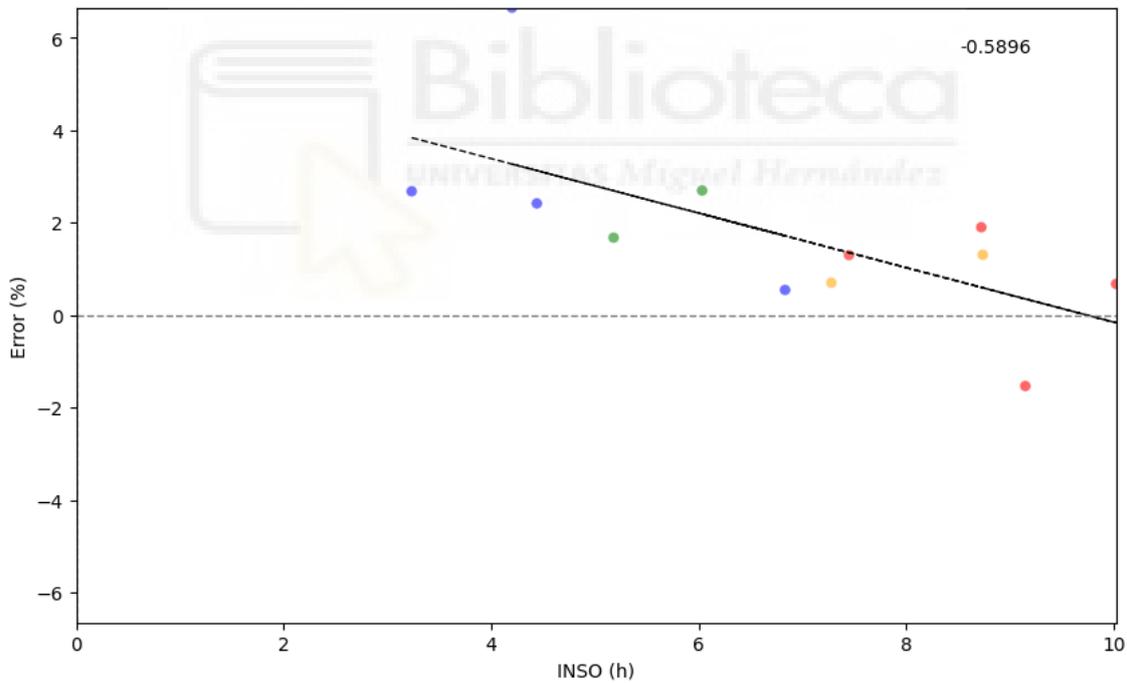


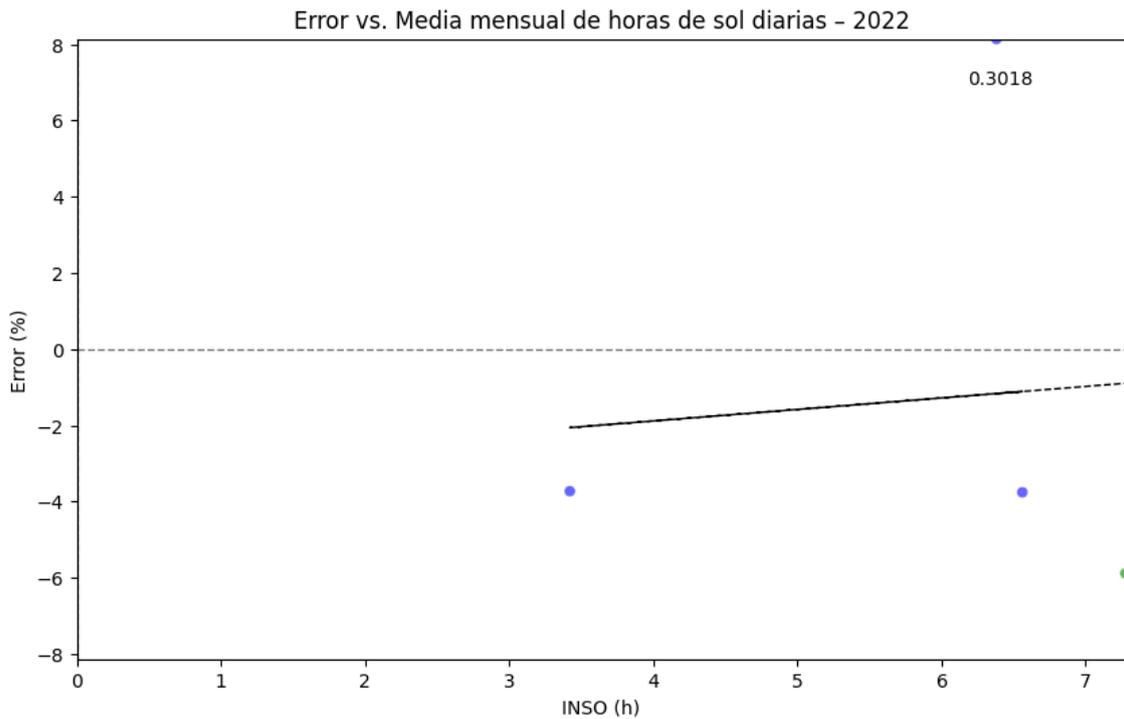


Error vs. Media mensual de horas de sol diarias - 2020



Error vs. Media mensual de horas de sol diarias - 2021





Año	Pendiente error vs. inso
2017	0.083
2018	0.045
2019	-0.247
2020	-0.557
2021	-0.590
2022	0.302

Como se puede observar las gráficas son similares a la variable GLO\_MJ. En 2017 y 2018 no hay relación evidente entre el error y la variable estudiada ya que la pendiente es pequeña. En los años 2019, 2020 y 2021 se puede apreciar que el error tiende a estar cerca de la zona negativa en los meses de verano. En 2022 la pendiente es similar a la de los primeros años pero, de nuevo, solo tenemos información para los meses de

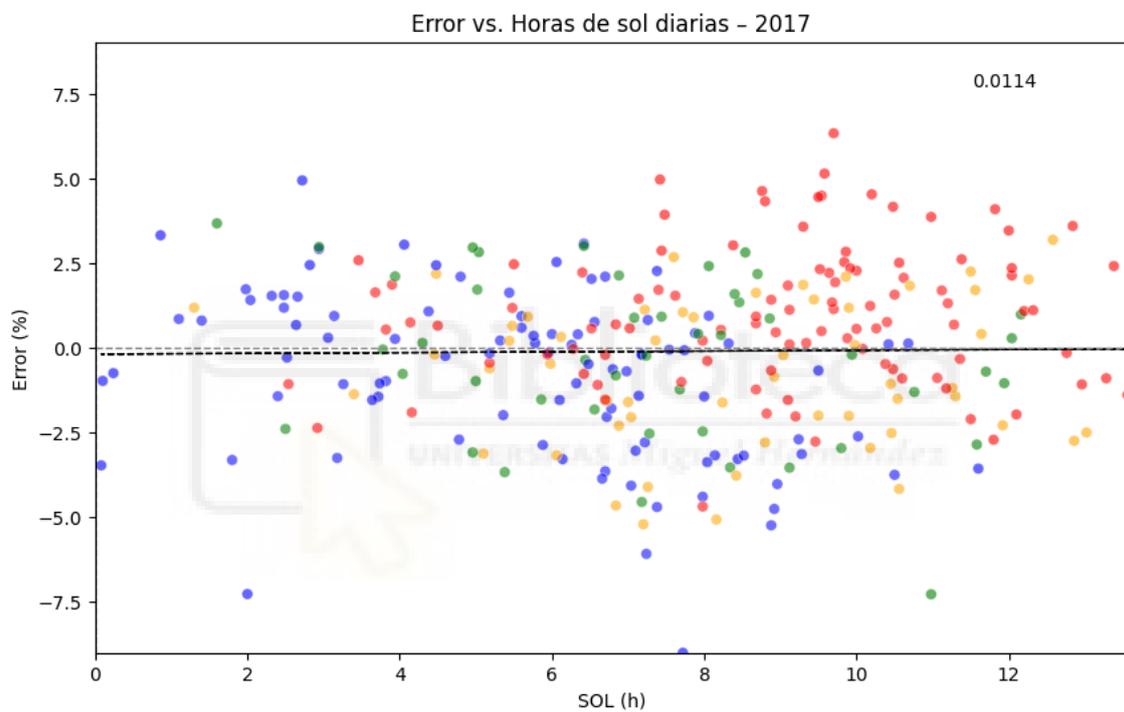
invierno y el punto correspondiente a los datos de primavera tiene una magnitud negativa mayor que los de invierno.



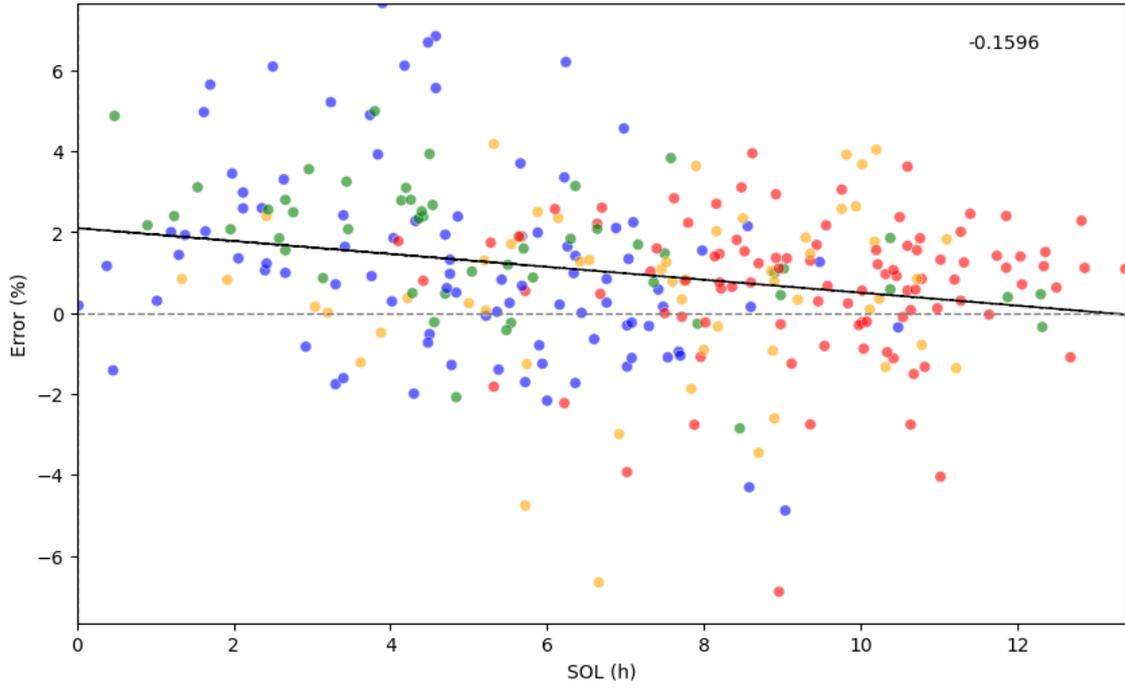
### 3.6. HORAS DE SOL DIARIAS

Esta variable (SOL) representa las horas de sol diarias.

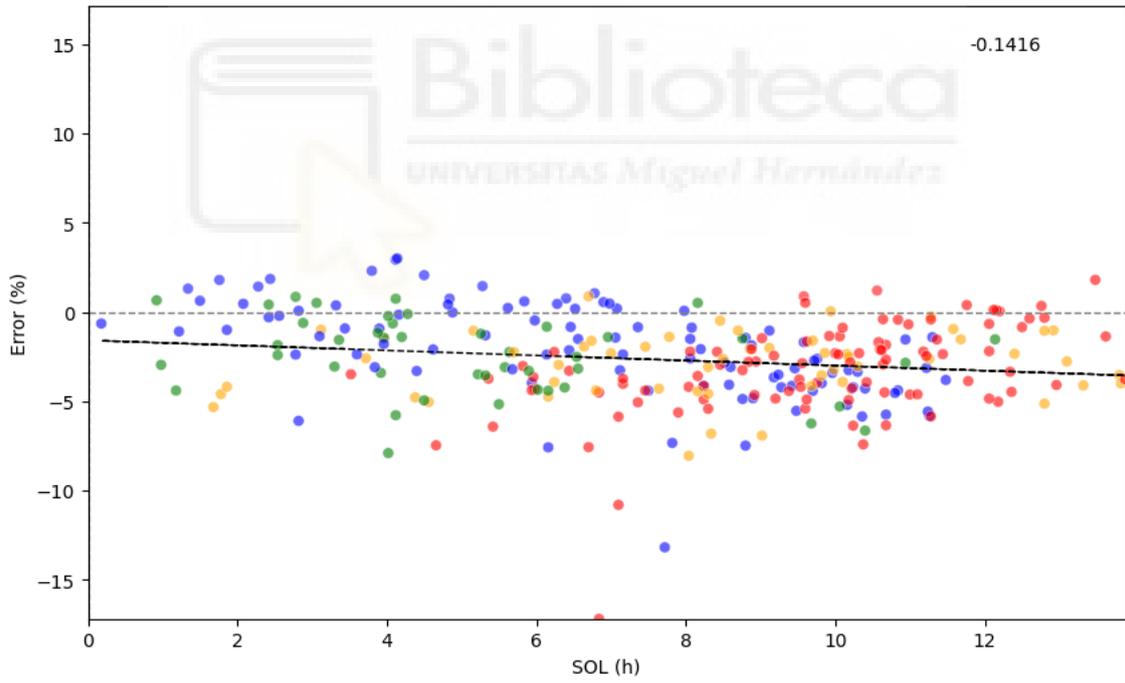
Los datos de esta variable son diarios por lo que se ha comparado su magnitud con el error del modelo a dicho nivel. Nuevamente el análisis se ha hecho a través de gráficos de dispersión para cada año de cálculo.



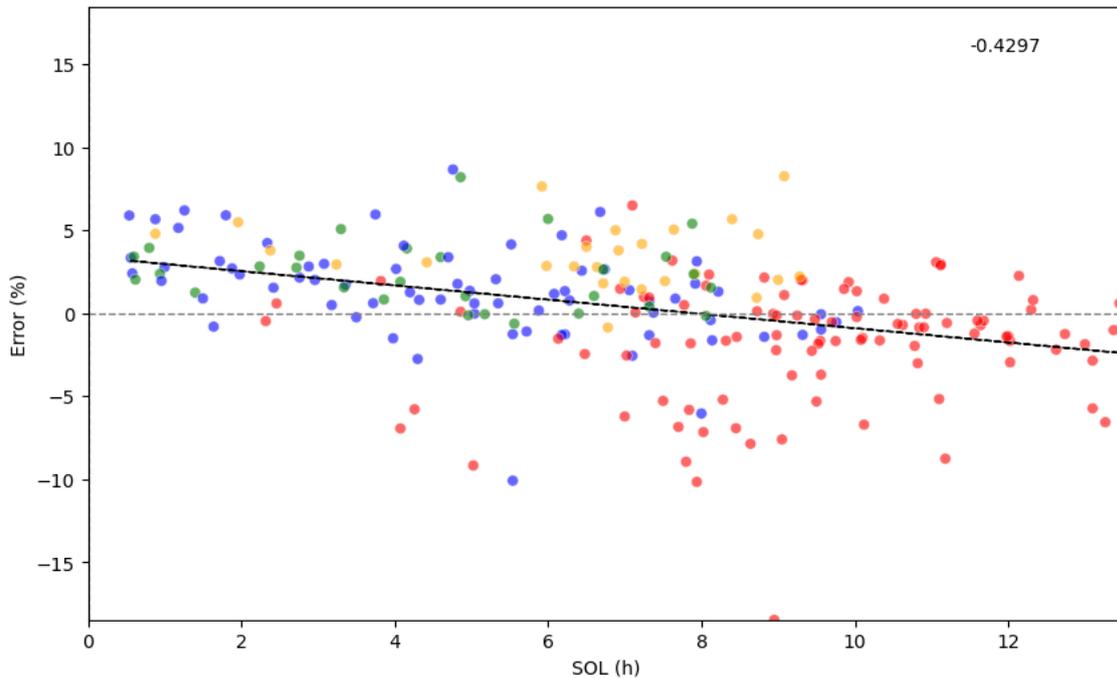
Error vs. Horas de sol diarias - 2018



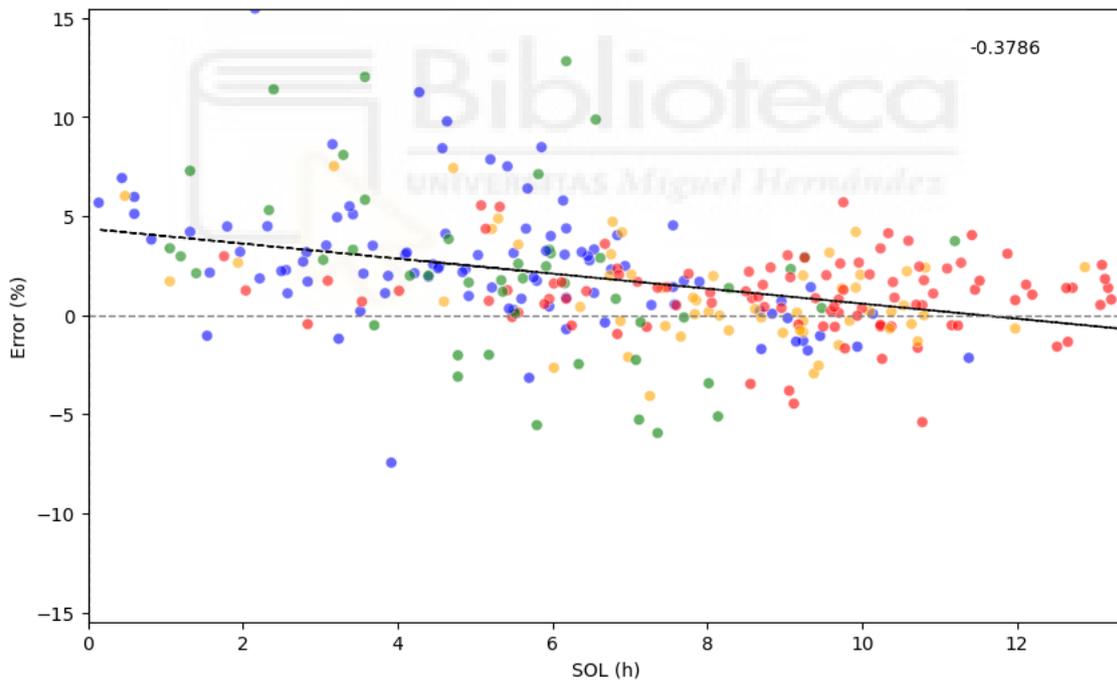
Error vs. Horas de sol diarias - 2019

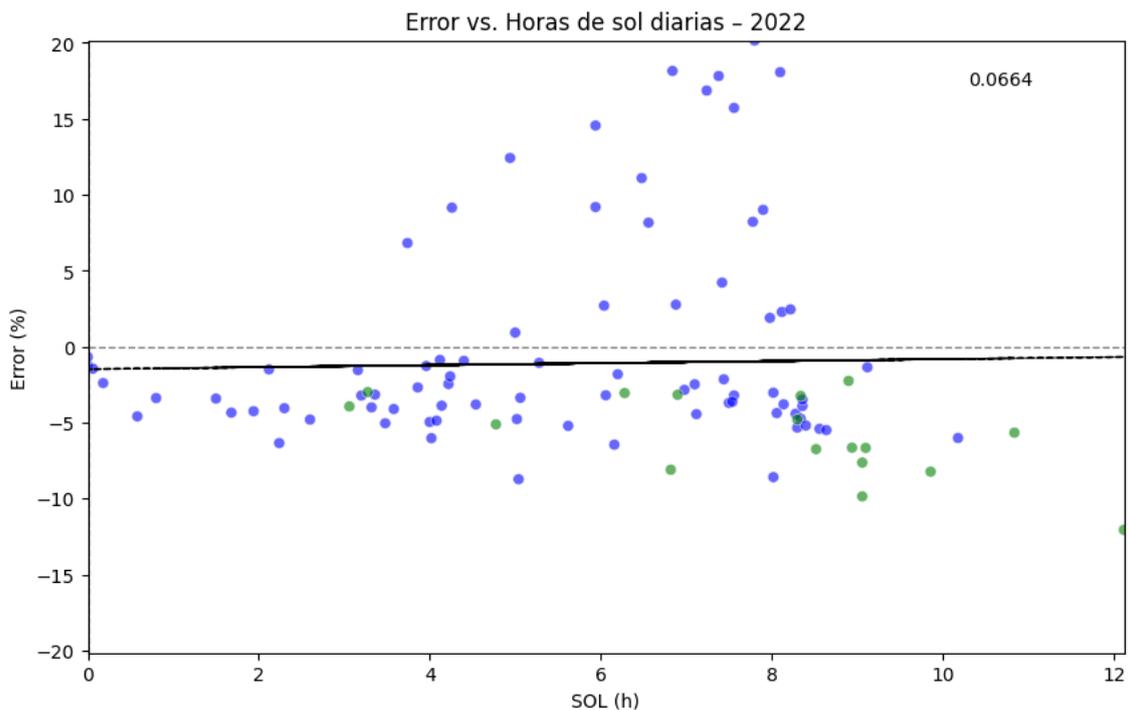


Error vs. Horas de sol diarias - 2020



Error vs. Horas de sol diarias - 2021





Año	Pendiente error vs. sol
2017	0.011
2018	-0.160
2019	-0.142
2020	-0.430
2021	-0.379
2022	0.066

El patrón de variables anteriores se repite en este caso, pero por tener más puntos de datos la correlación de la recta de regresión es mayor. Se puede observar que la pendiente de la curva se va haciendo negativa y la correlación se va haciendo aparente, especialmente en los dos últimos años para los que hay datos de verano (2020 y 2021), en los que además el error tiende a valores negativos en relación con la variable, indicando que aunque no siempre haya sobreestimación de la carga, la magnitud de la infraestimación es menor en estos meses. En 2022 la recta de regresión vuelve a ser

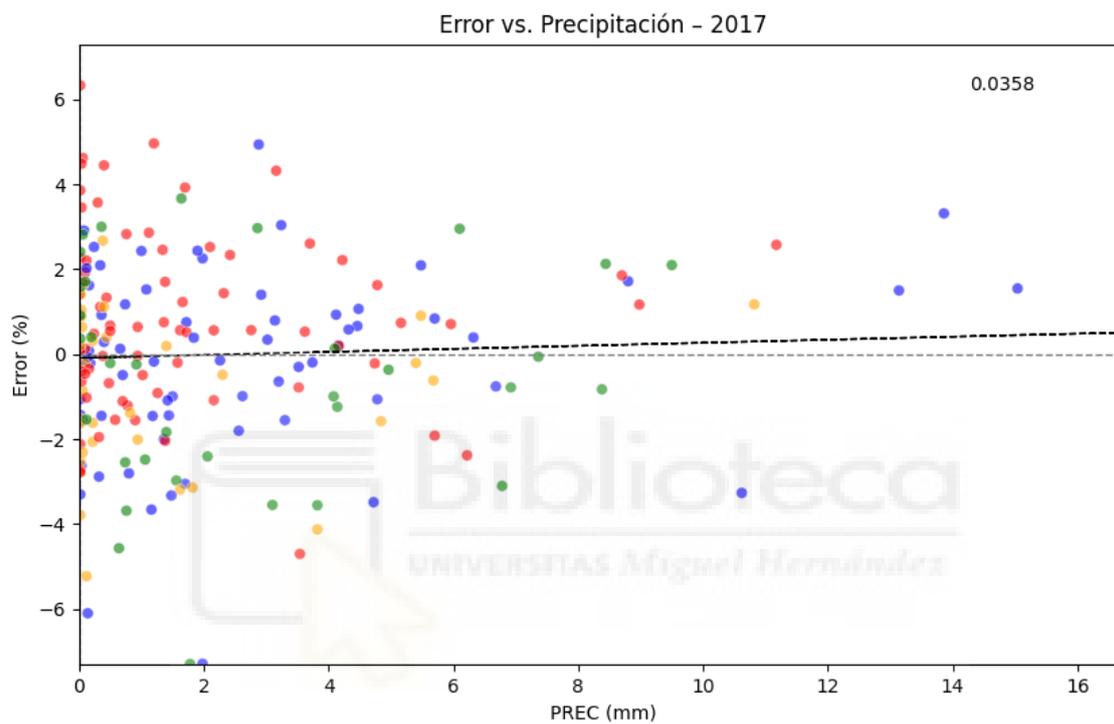
plana pero, de nuevo, solo tenemos datos de invierno. Aun así los puntos con una magnitud mayor en el eje de abscisas parecen tener un error mayor en sentido negativo.



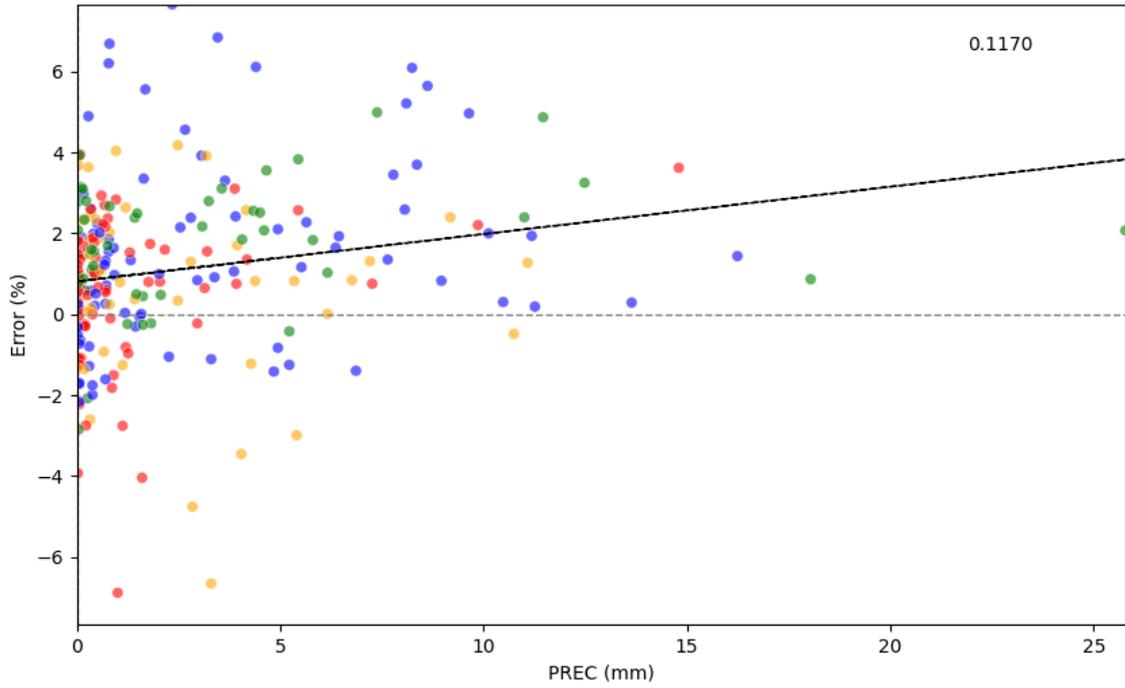
### 3.7. PRECIPITACIÓN

Esta variable (PREC) hace referencia a la cantidad de precipitación en mm.

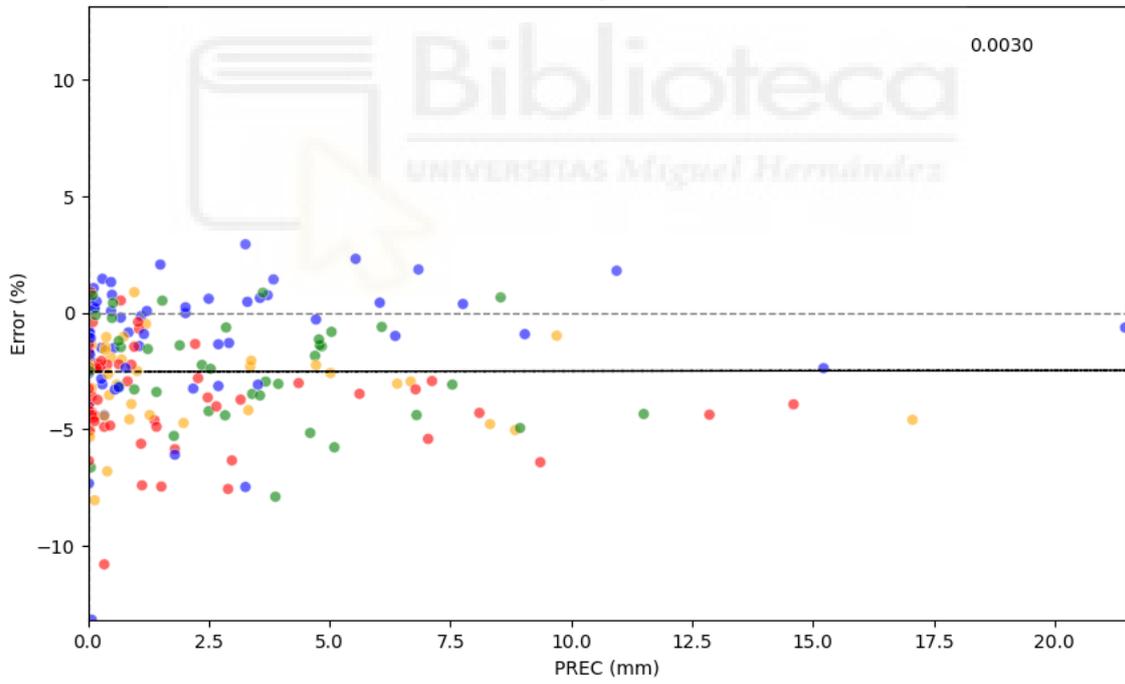
Para analizar esta variable se han filtrado aquellos días que tienen una precipitación mayor que 0 y se han puesto en relación con el error del modelo para cada año de cálculo en gráficos de dispersión.



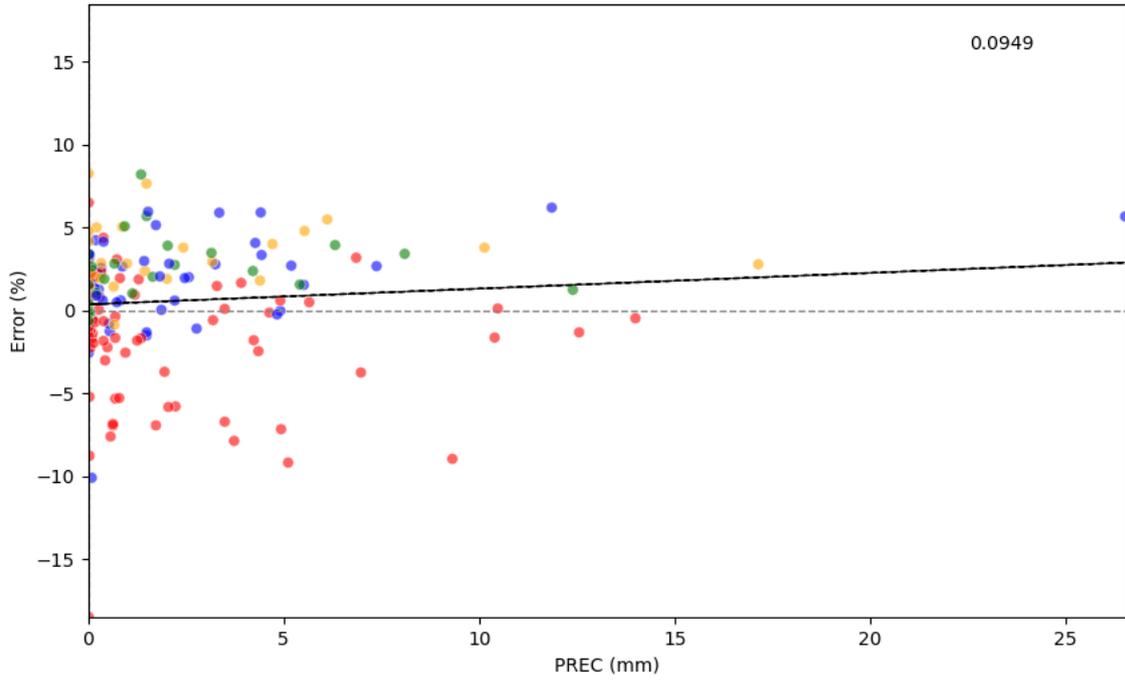
Error vs. Precipitación - 2018



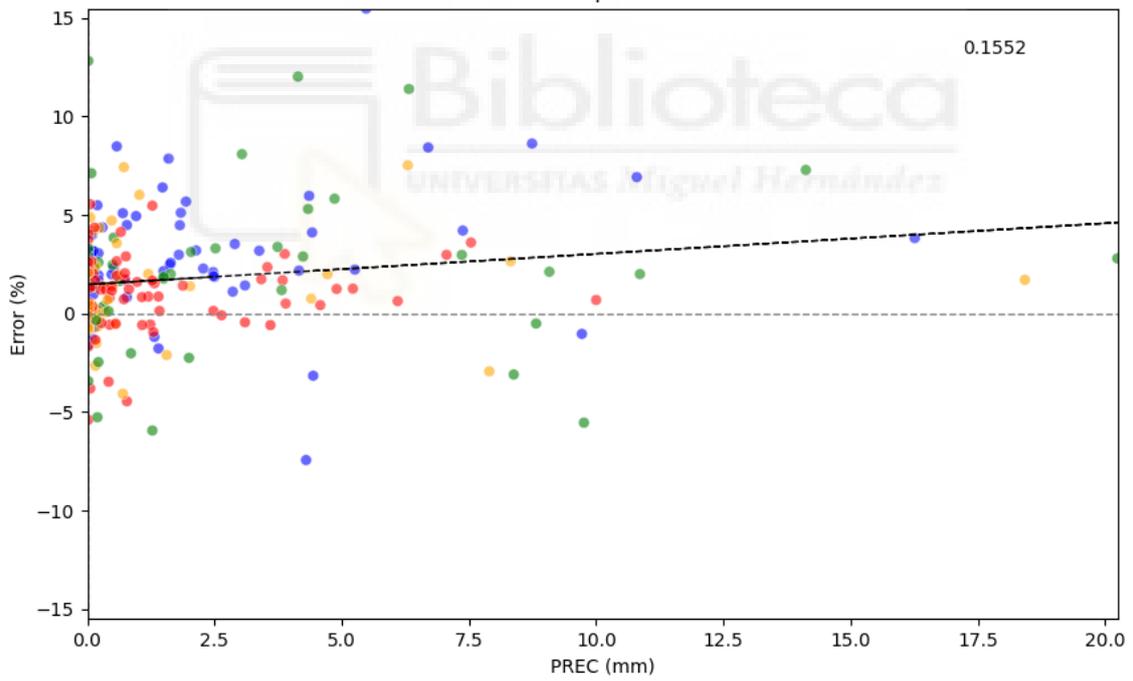
Error vs. Precipitación - 2019

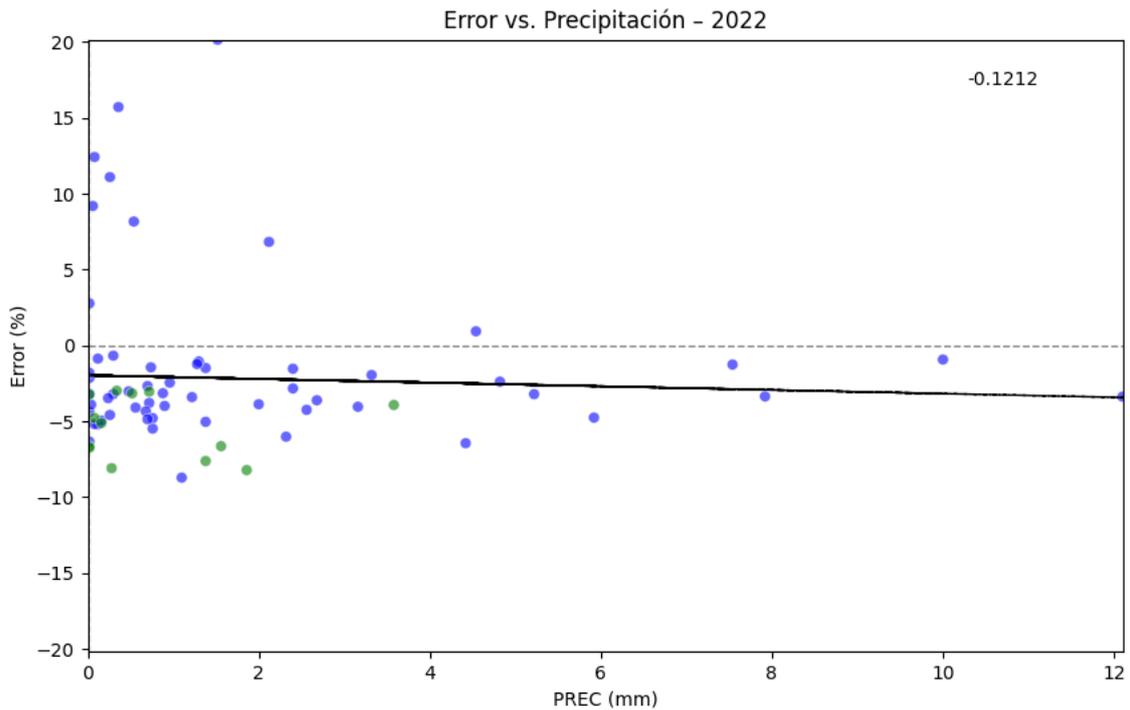


Error vs. Precipitación - 2020



Error vs. Precipitación - 2021





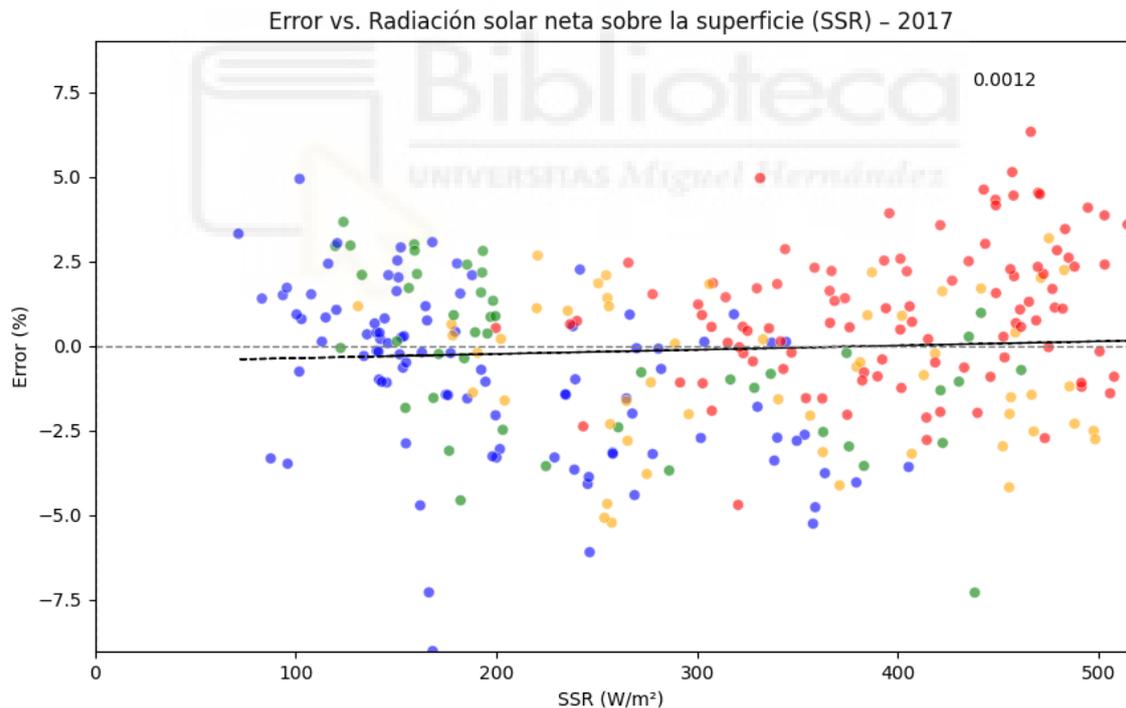
Año	Pendiente error vs. prec
2017	0.089
2018	0.103
2019	-0.003
2020	0.056
2021	0.098
2022	-0.043

Aunque no se vea reflejado en la pendiente de la recta de regresión sí se puede apreciar que a partir de 2019 en verano cuando hay precipitación el error tiende a ser negativo por lo que se produce una sobreestimación de energía.

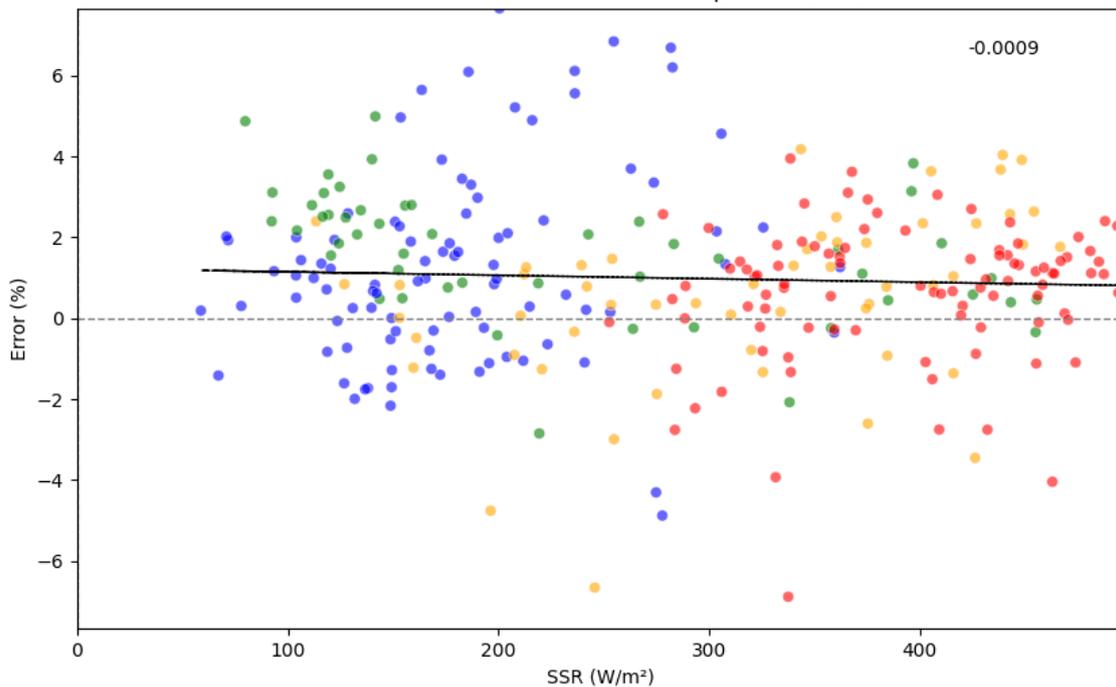
### 3.8. RADIACIÓN SOLAR NETA SOBRE LA SUPERFICIE

Esta variable (SSR\_WM2) representa la radiación solar neta sobre la superficie en W/m<sup>2</sup> del conjunto de datos ERA5. Se define como la cantidad de radiación solar que llega a la superficie de la Tierra después de tener en cuenta los efectos del albedo (% de radiación que cualquier superficie refleja respecto a la que incide sobre ella). Es importante para la producción fotovoltaica porque es una medida de la energía solar disponible para su conversión en electricidad.

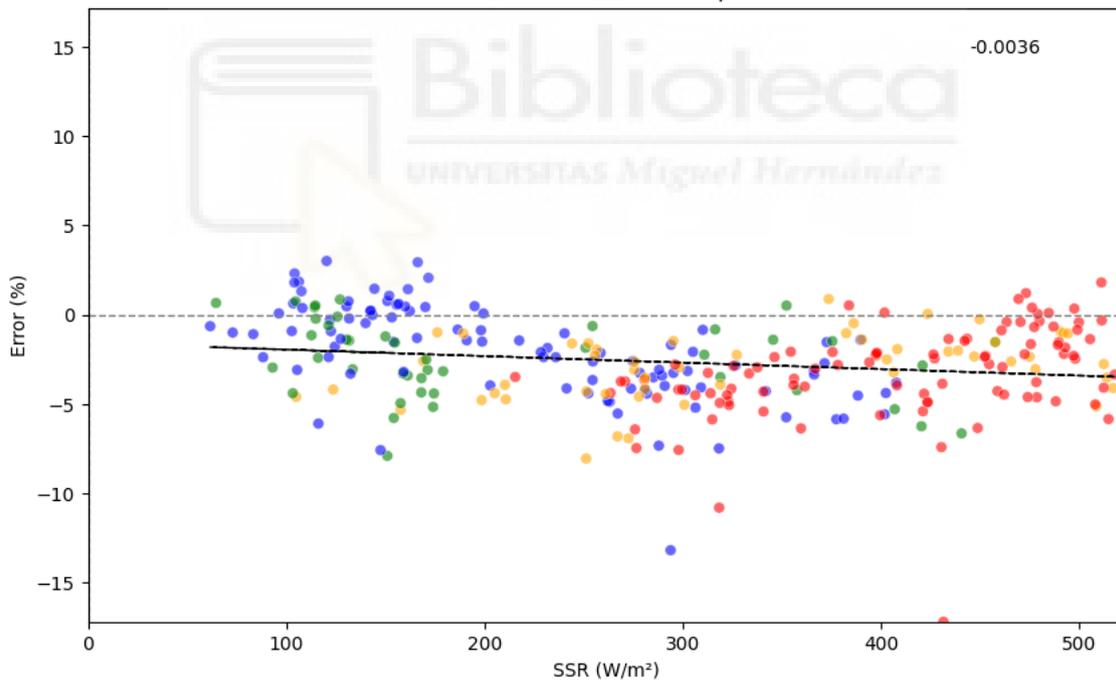
Los datos de esta variable se presentaban a nivel horario y para su análisis se han agregado a nivel diario en una media en el mismo rango temporal que los de entrada del modelo (08-20h). Posteriormente se ha representado la correspondencia del error y de la magnitud de la variable para los diferentes días en gráficos de dispersión y se ha calculado la recta de regresión.



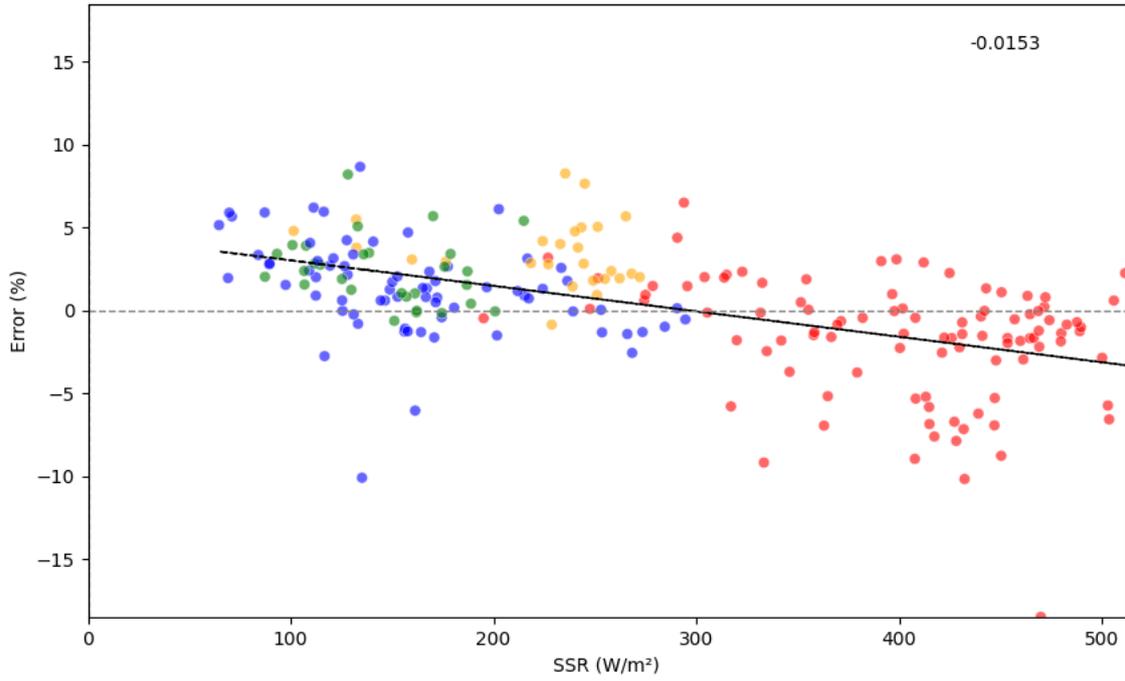
Error vs. Radiación solar neta sobre la superficie (SSR) - 2018



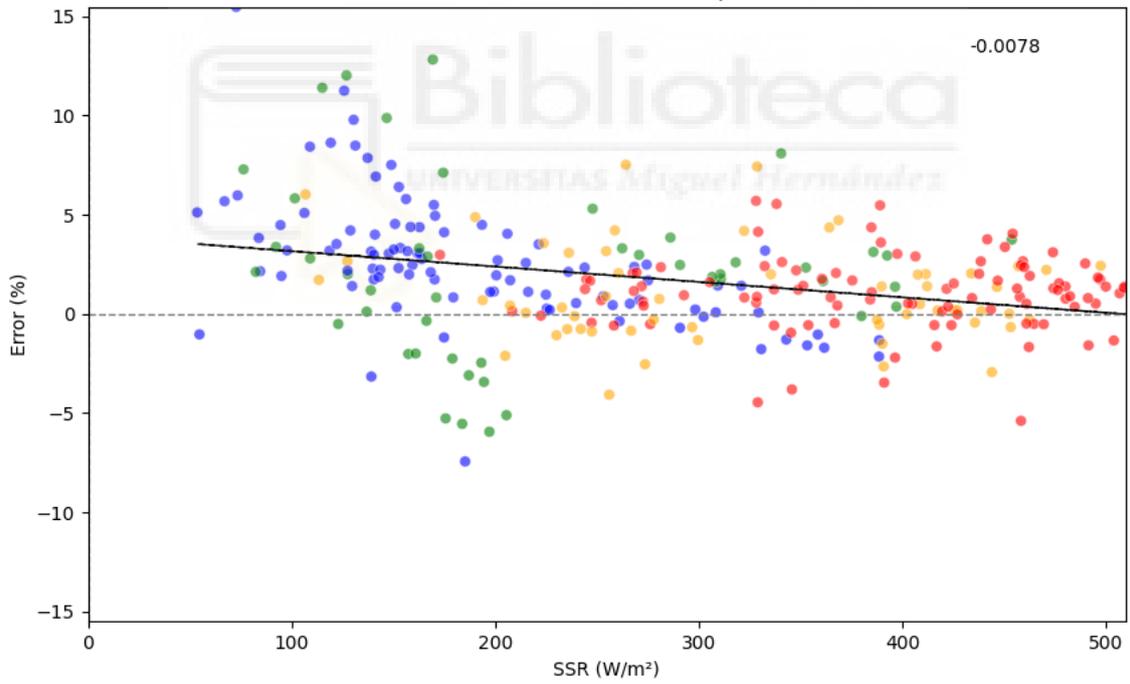
Error vs. Radiación solar neta sobre la superficie (SSR) - 2019

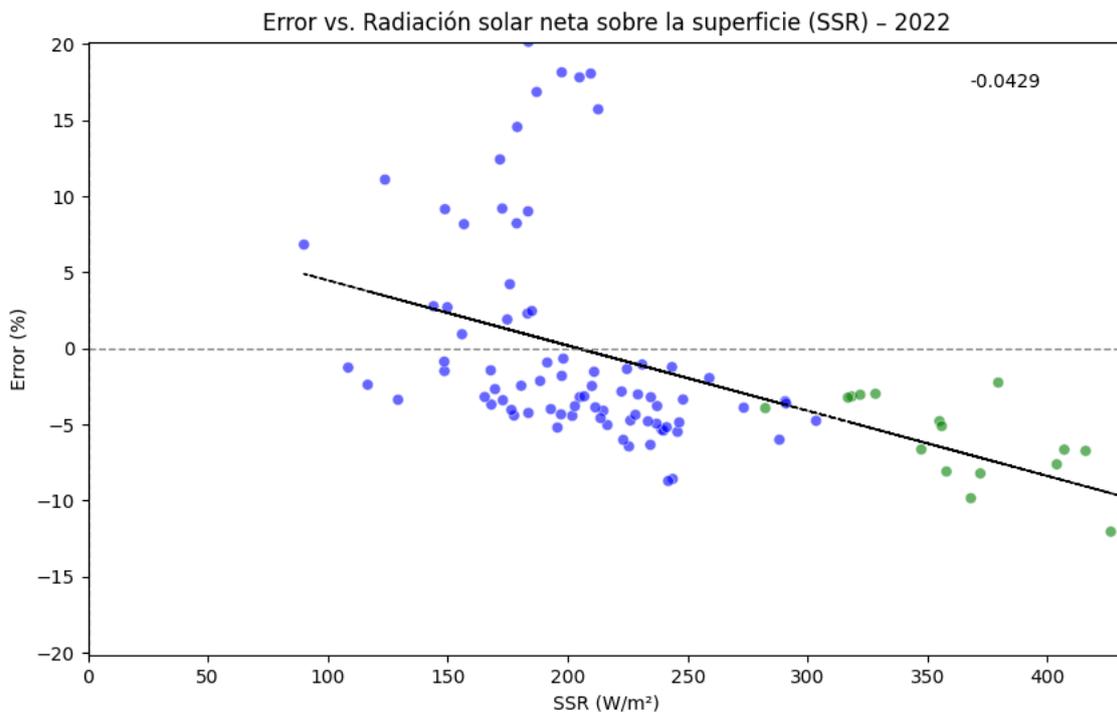


Error vs. Radiación solar neta sobre la superficie (SSR) - 2020



Error vs. Radiación solar neta sobre la superficie (SSR) - 2021





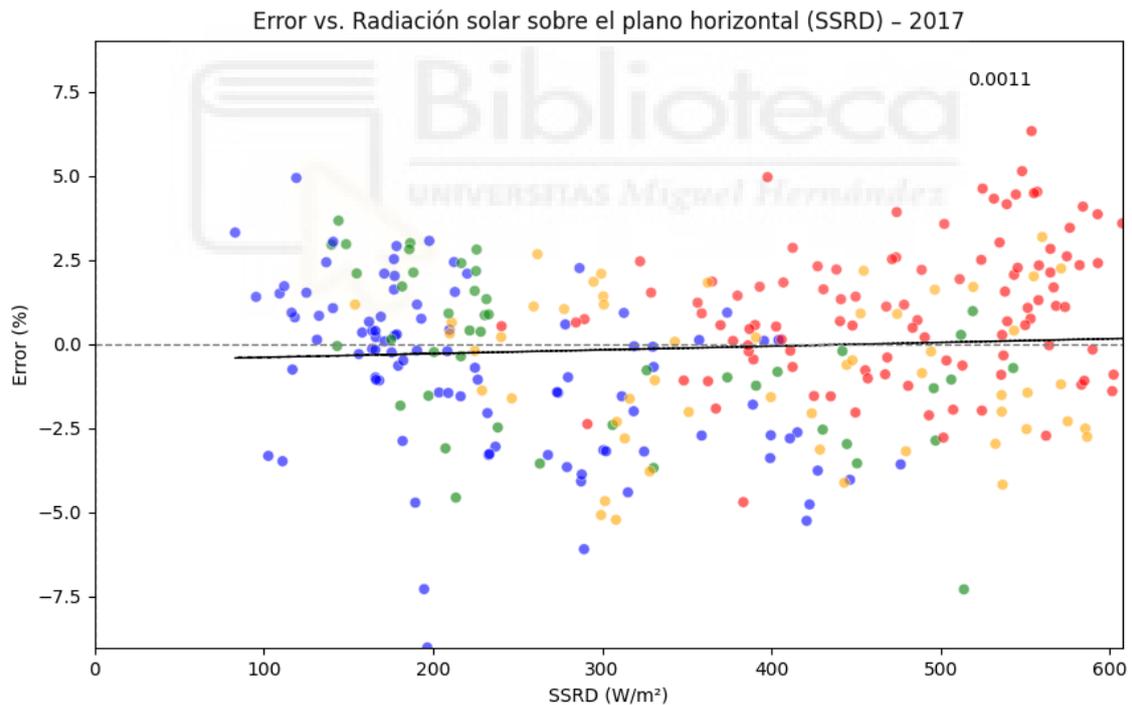
Año	Pendiente error vs. <i>ssr_wm2</i>
2017	0.001
2018	-0.001
2019	-0.004
2020	-0.015
2021	-0.008
2022	-0.043

Podemos ver en las gráficas que la relación entre el error del modelo y esta variable es plana durante los primeros años para los que tenemos datos, pero en 2020 y en 2021 es aparente una relación inversamente proporcional entre la radiación y el error. Esta relación se mantiene también en 2022 con los datos de primavera (en verde).

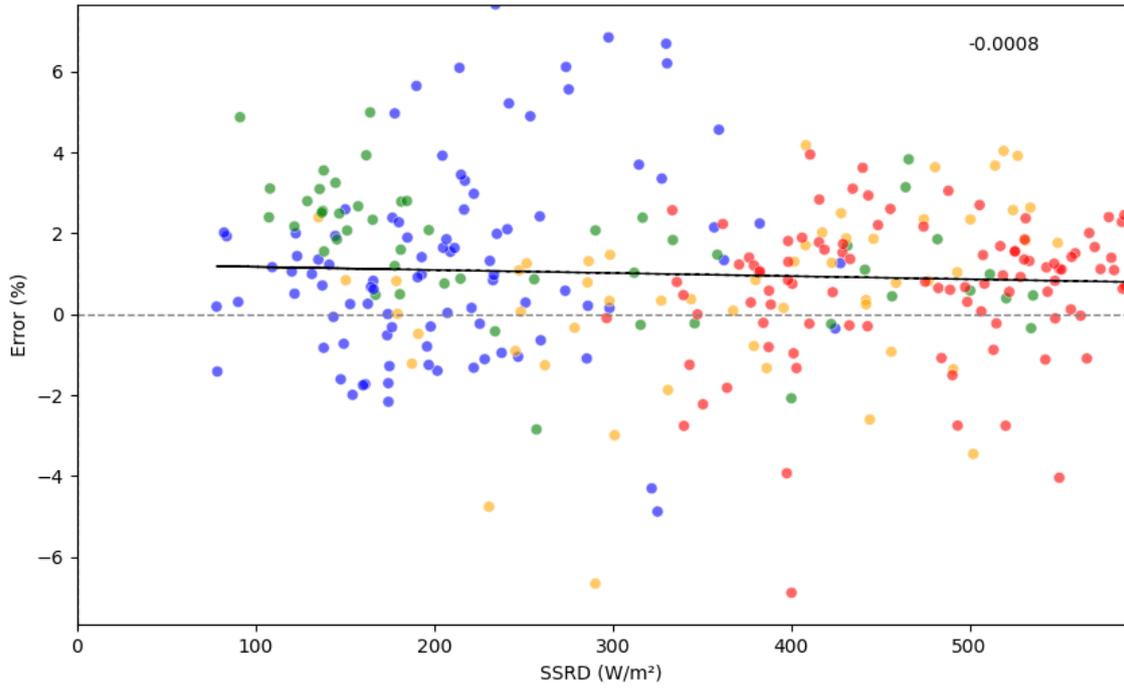
### 3.9. RADIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL

Esta variable (SSRD\_WM2) representa la radiación solar sobre el plano horizontal en W/m<sup>2</sup> del conjunto de datos ERA5. Se define como la radiación solar total que incide en un plazo horizontal de la superficie de la tierra incluyendo la radiación directa y la difusa. Proporciona información sobre la energía solar total disponible.

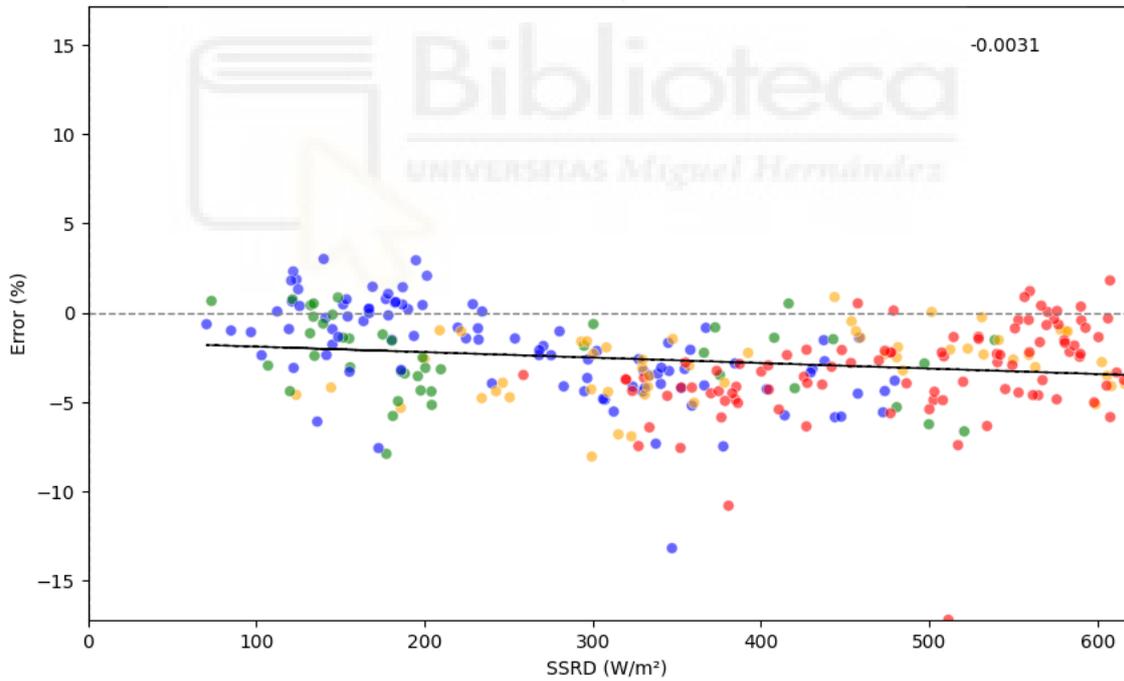
Como en el caso de la variable SSR\_WM2, los datos de SSRD\_WM2 tienen una granularidad diaria por lo que se han filtrado y agregado en una media para el mismo rango temporal (08-20h) para transformarlos en diarios. En este punto se han unido los datos a los del modelo por el día para obtener la correspondencia y esta se ha representado en gráficos de dispersión.



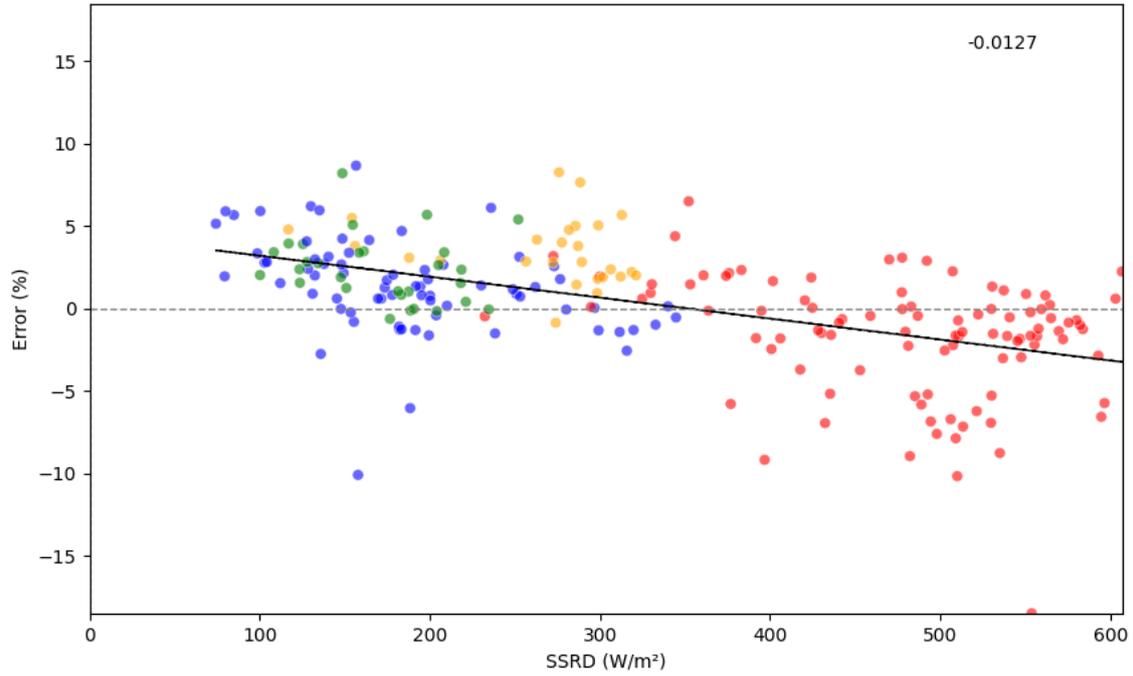
Error vs. Radiación solar sobre el plano horizontal (SSRD) - 2018



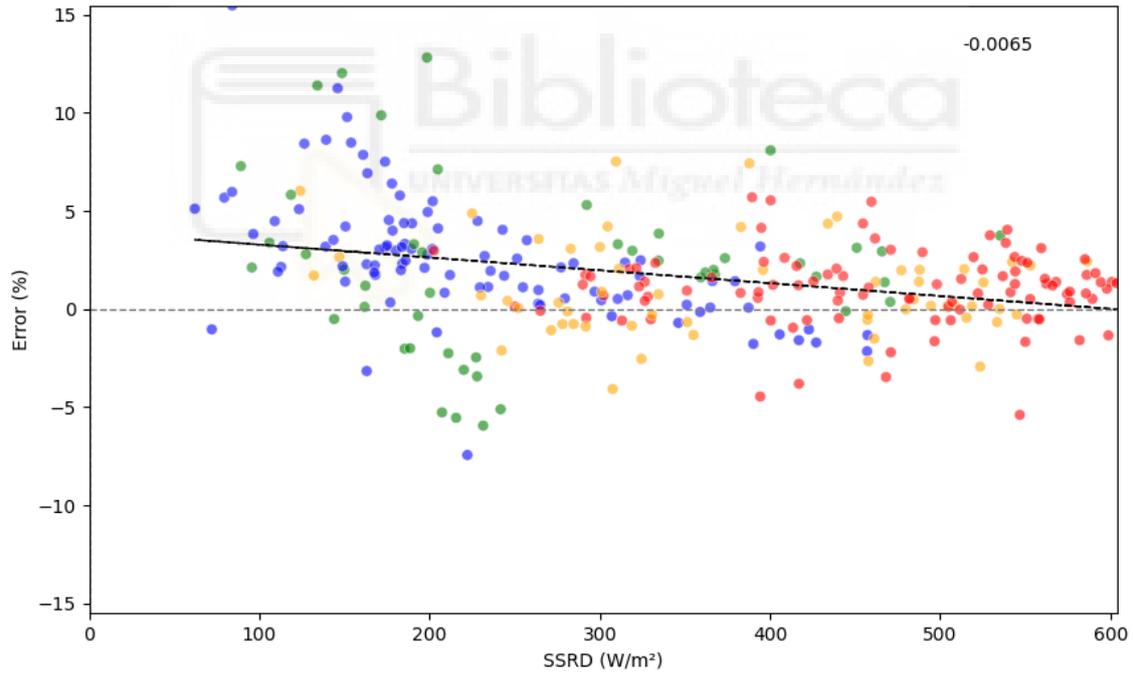
Error vs. Radiación solar sobre el plano horizontal (SSRD) - 2019

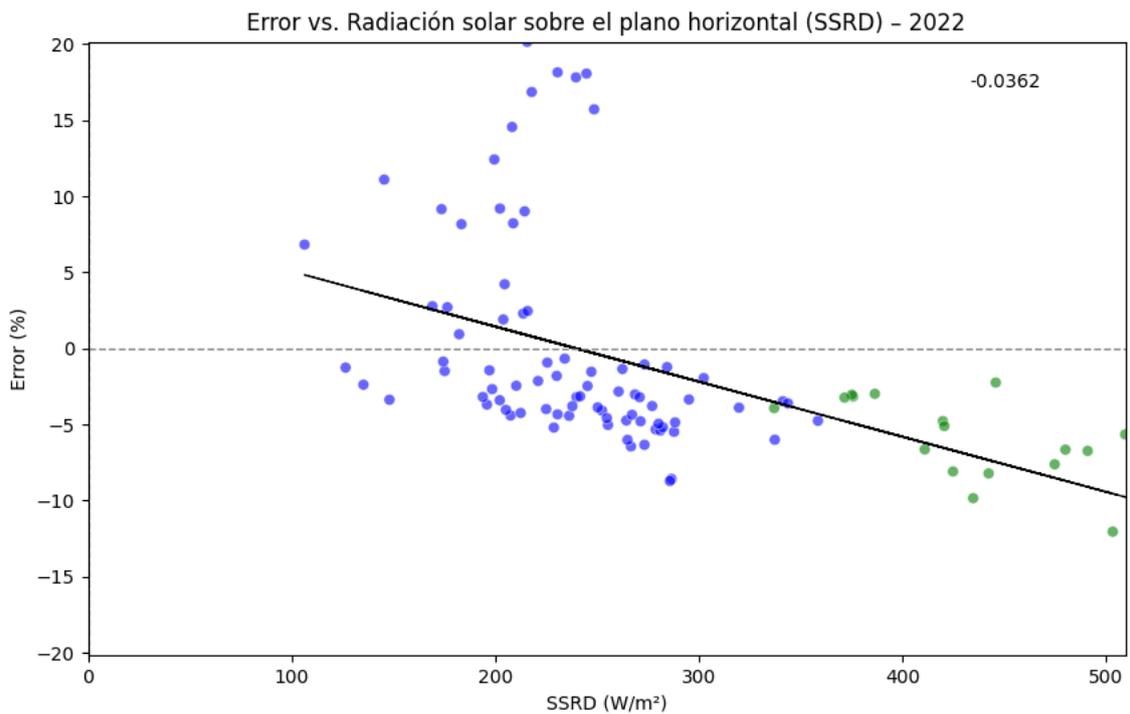


Error vs. Radiación solar sobre el plano horizontal (SSRD) - 2020



Error vs. Radiación solar sobre el plano horizontal (SSRD) - 2021





Año	Pendiente error vs. ssrd_wm2
2017	0.001
2018	-0.001
2019	-0.003
2020	-0.013
2021	-0.006
2022	-0.036

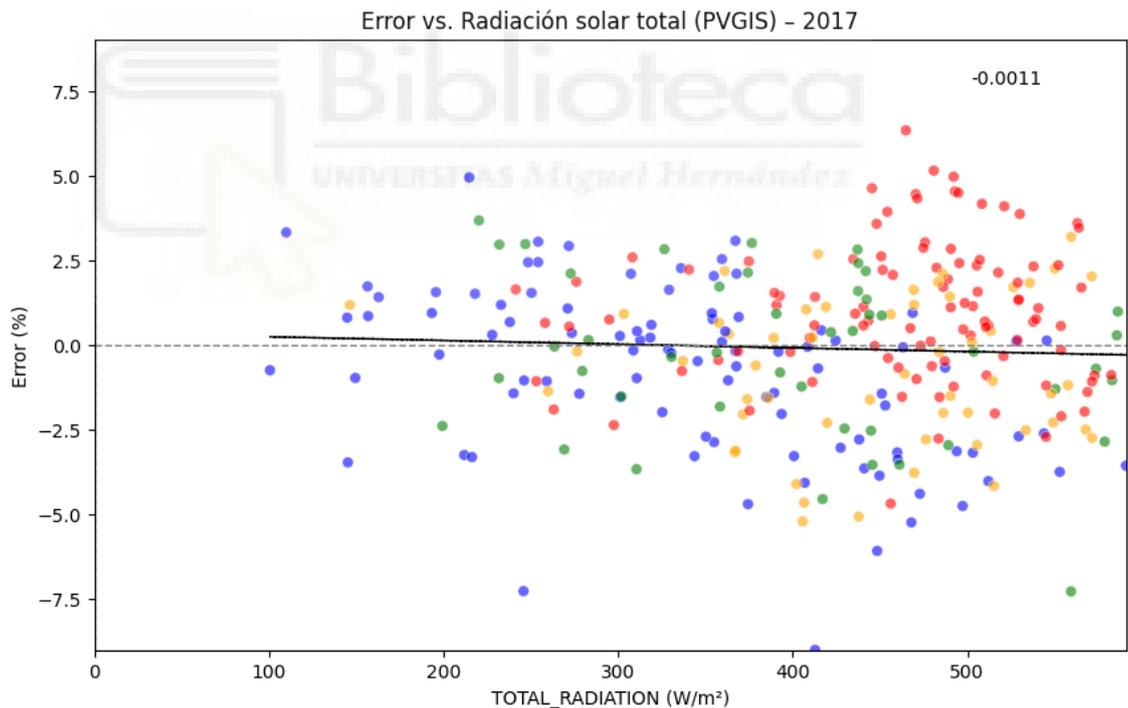
El comportamiento es similar al que observamos en la variable SSR\_WM2. En 2017, 2018 y 2019 no se aprecia relación entre esta variable y el error, pero en 2020 y 2021 los días con radiación más alta están más cerca de la región negativa del diagrama. Se puede observar esta tendencia también en 2022.

### 3.10. RADIACIÓN TOTAL

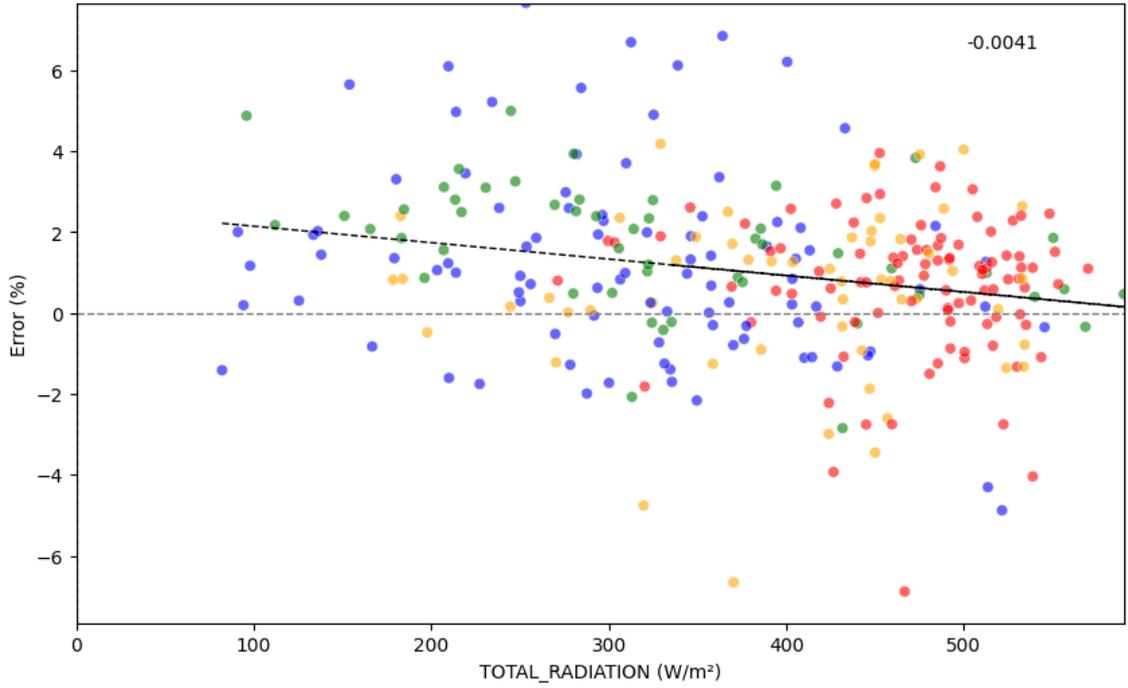
Esta variable (TOTAL\_RADIATION) representa la radiación total de los conjuntos de datos SARA3 y E5 de PVGIS. Representa la energía total disponible para la producción fotovoltaica y está compuesta por tres componentes: radiación directa (GBI, global beam irradiance), radiación difusa (GDI, global diffuse irradiance) y radiación reflejada (GRI, global reflected irradiance).

Los datos originales de esta variable estaban distribuidos horariamente por lo que se han agregado en el mismo rango temporal (08-20h) que los del modelo para obtener una media diaria que se ha puesto en relación con el error del mismo por día. Una vez obtenido, este conjunto de datos se ha representado en gráficas de dispersión, calculándose también la recta de regresión.

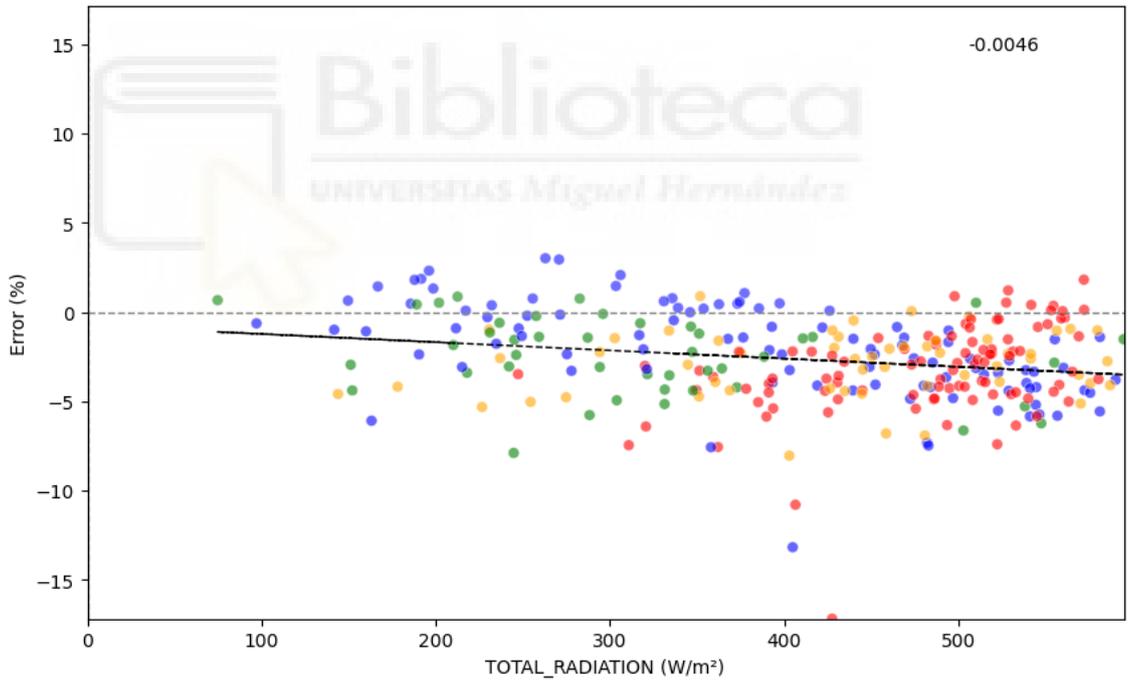
#### 3.10.1. SARA3



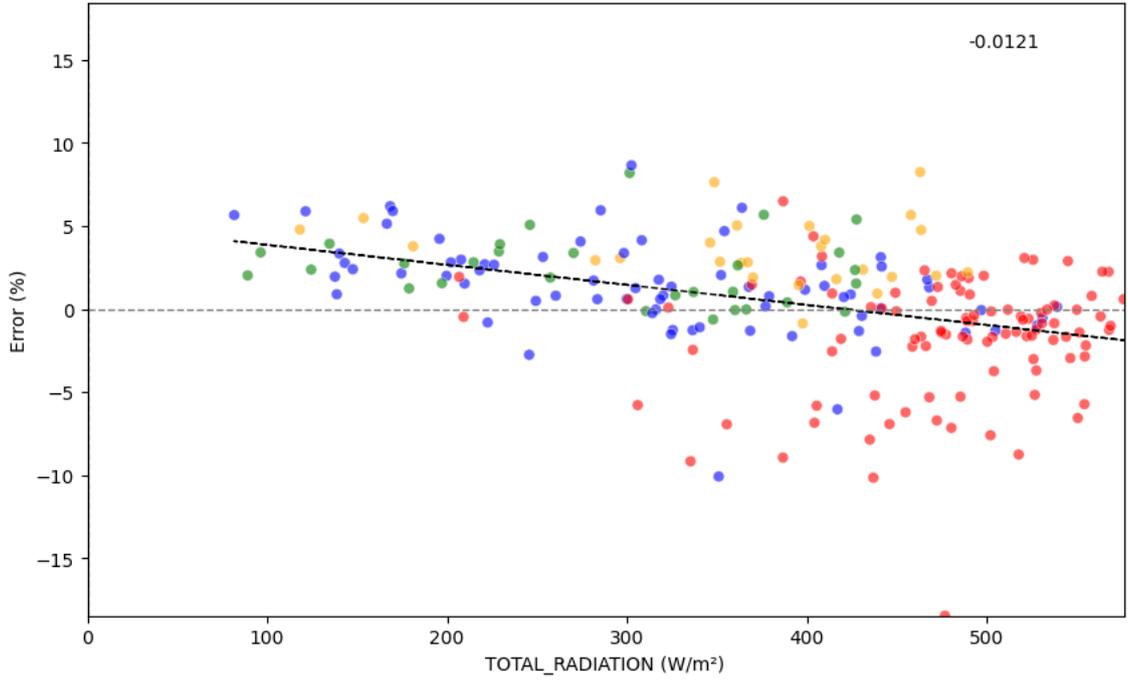
Error vs. Radiación solar total (PVGIS) - 2018



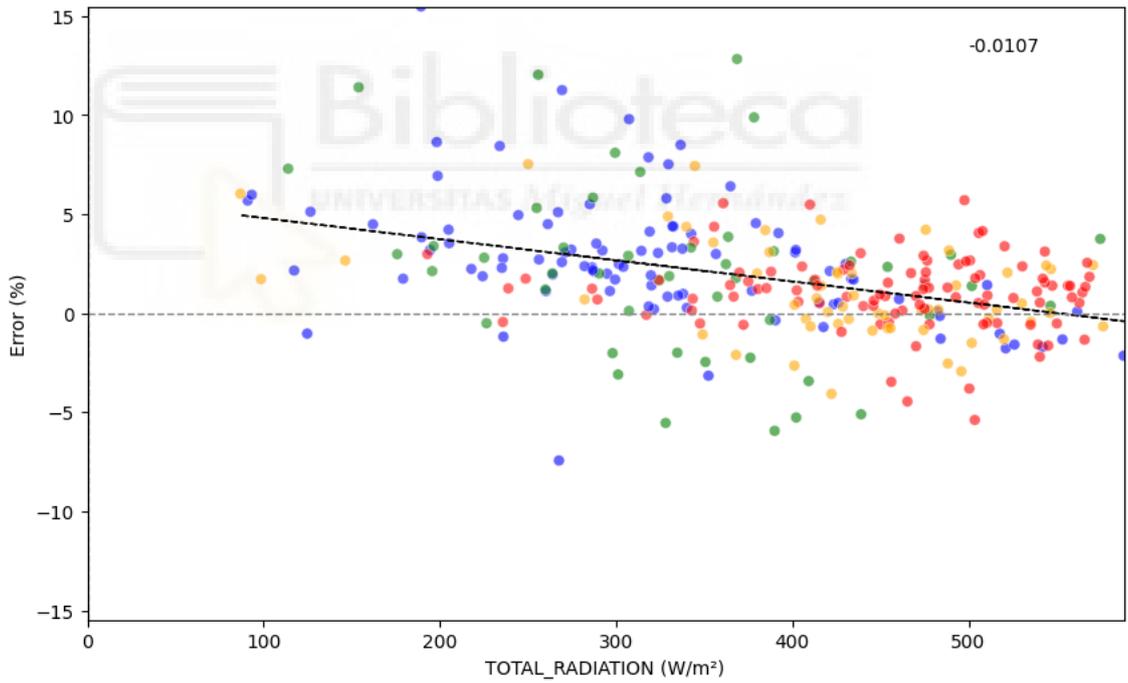
Error vs. Radiación solar total (PVGIS) - 2019

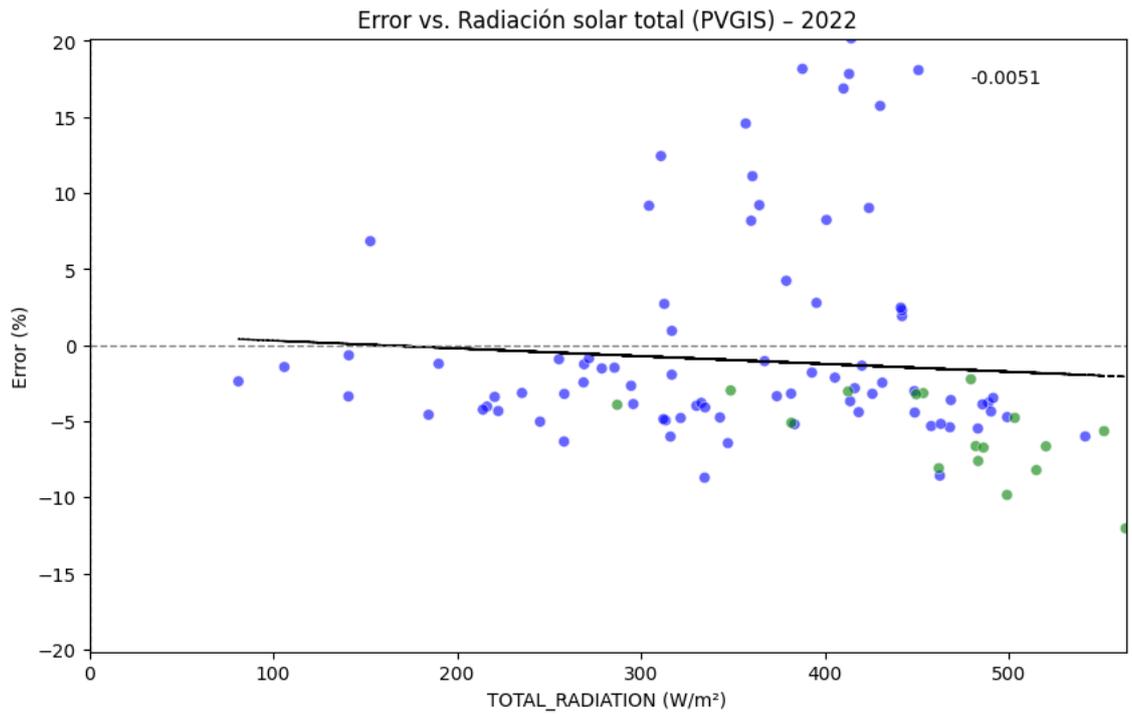


Error vs. Radiación solar total (PVGIS) - 2020

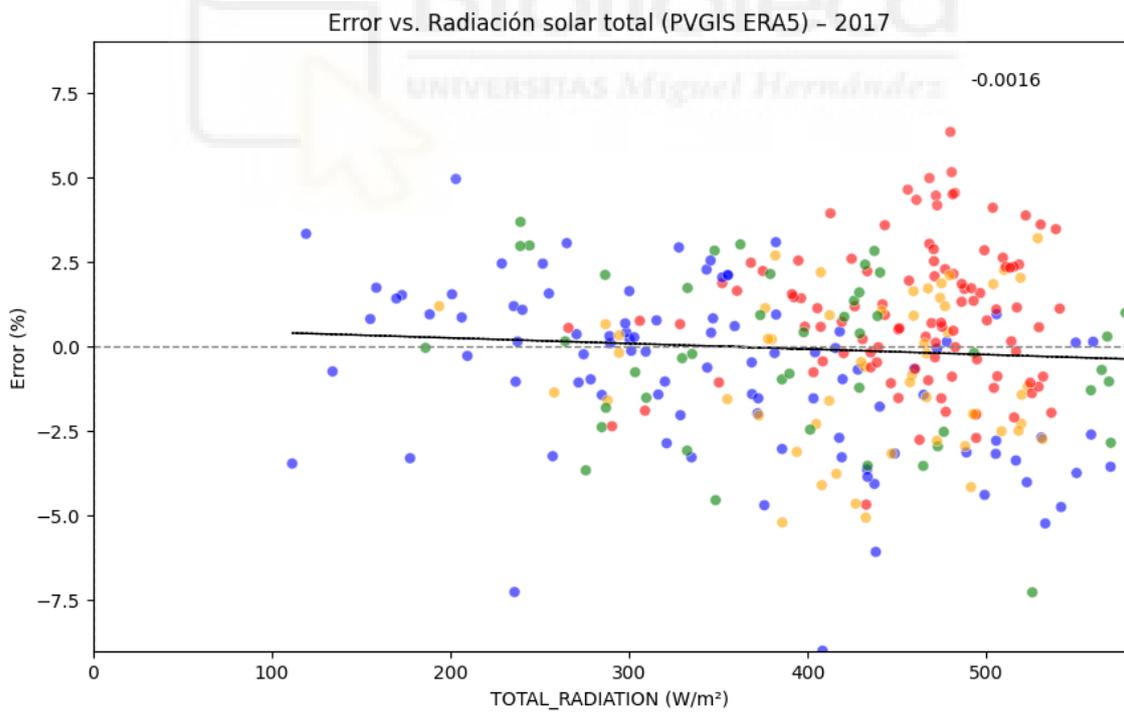


Error vs. Radiación solar total (PVGIS) - 2021

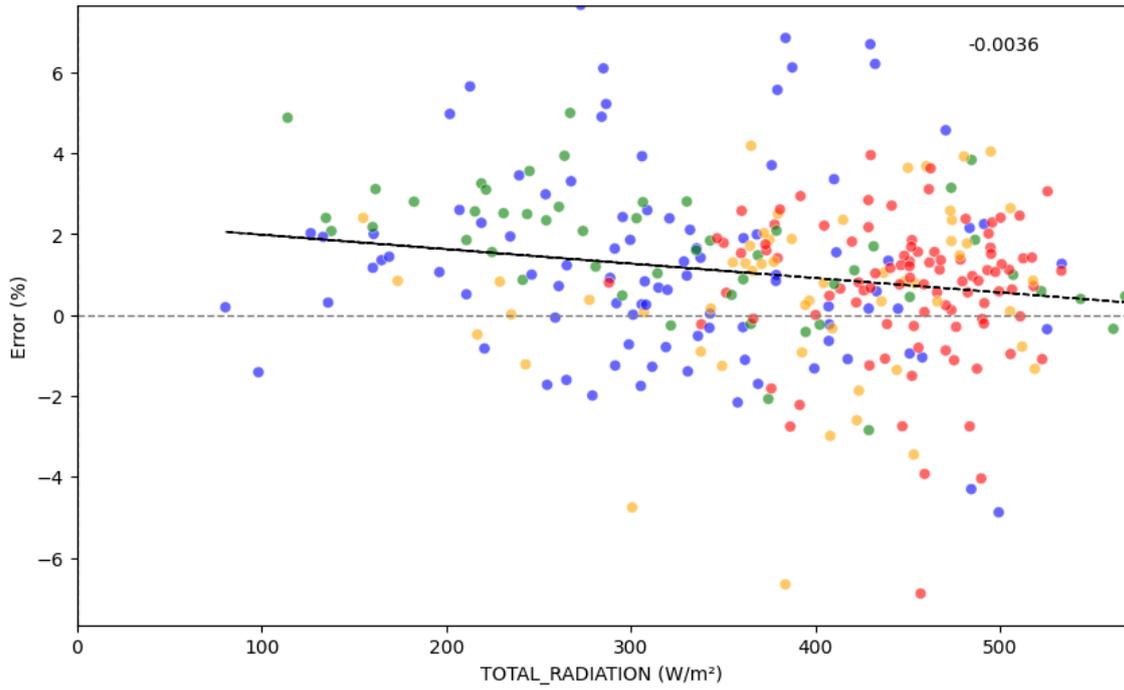




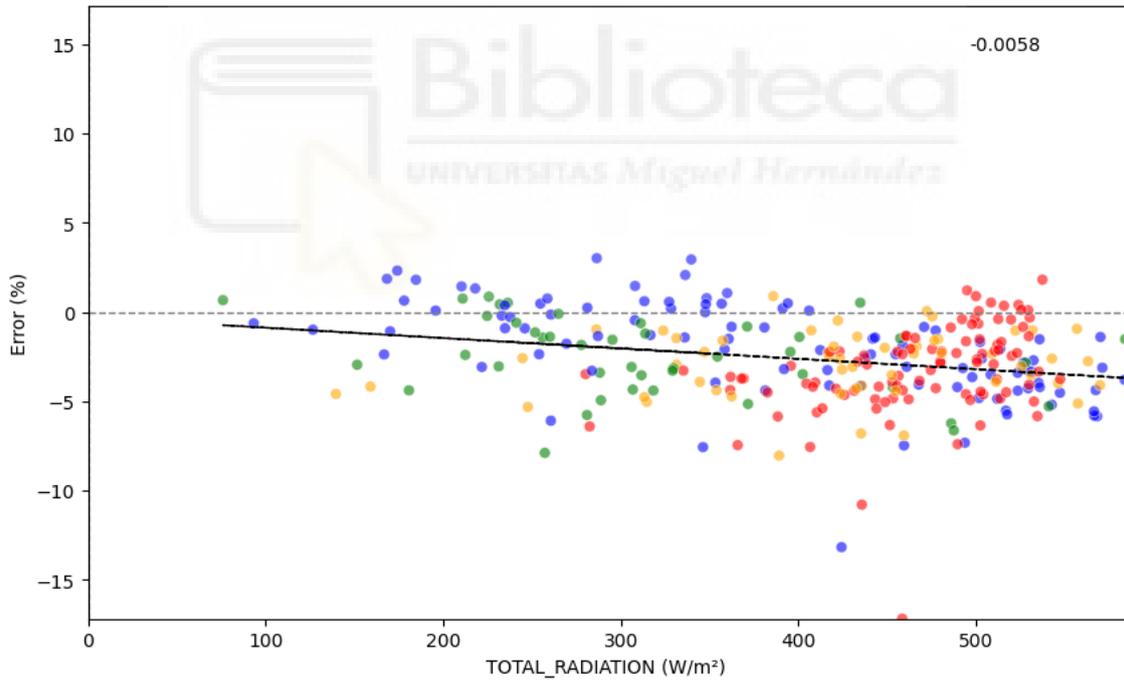
3.10.2. E5



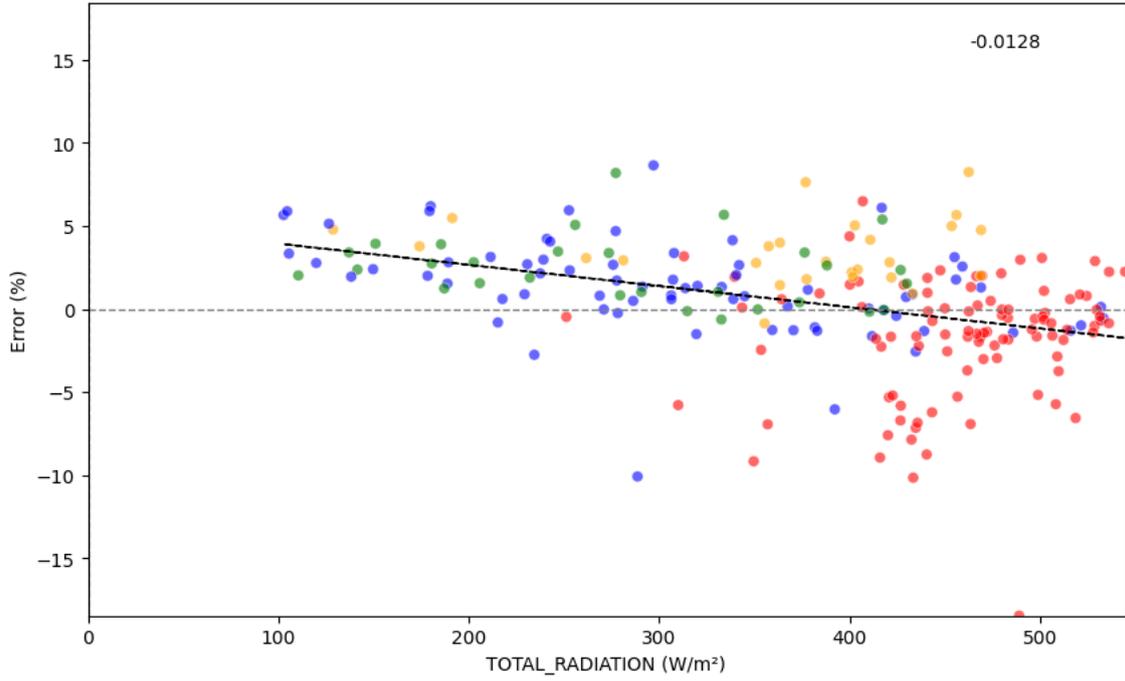
Error vs. Radiación solar total (PVGIS ERA5) - 2018



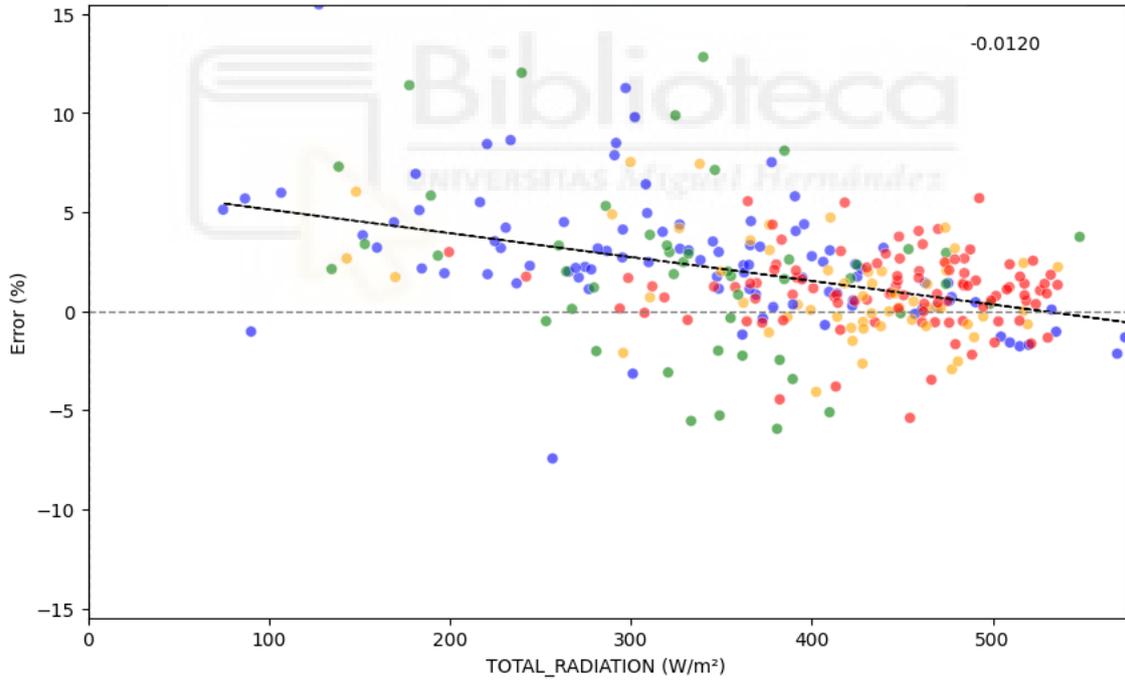
Error vs. Radiación solar total (PVGIS ERA5) - 2019

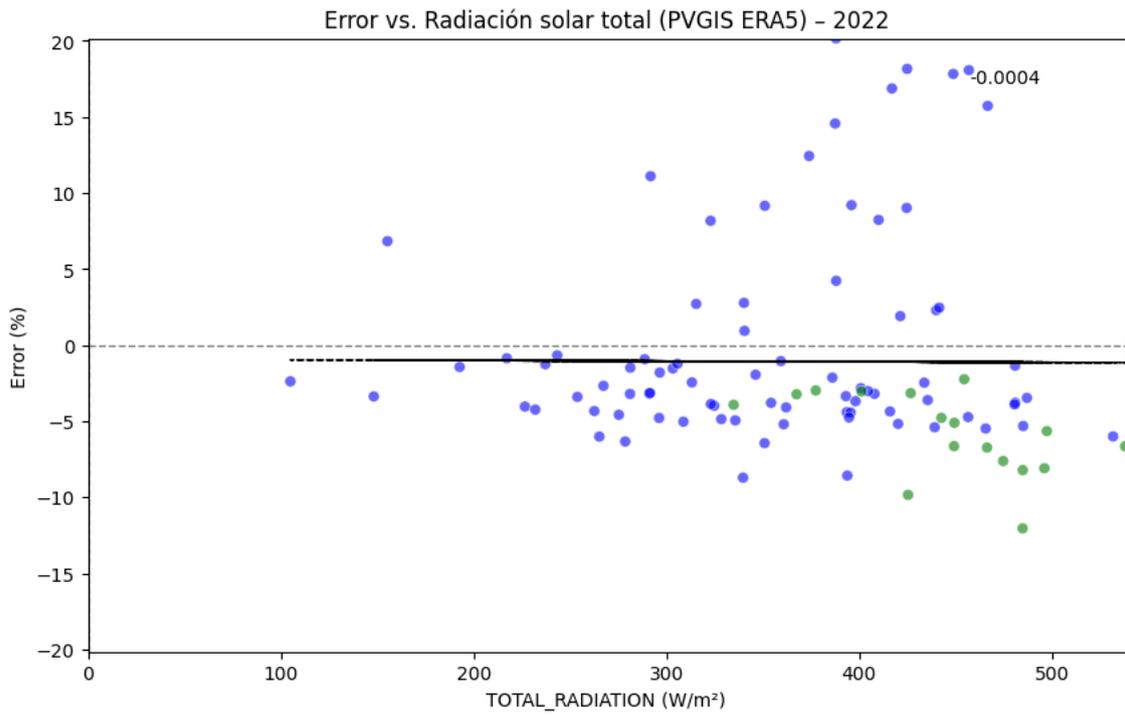


Error vs. Radiación solar total (PVGIS ERA5) - 2020



Error vs. Radiación solar total (PVGIS ERA5) - 2021





### 3.10.3. Análisis

Año	Pendiente error vs. total_radiation_pvgis
2017	-0.001
2018	-0.004
2019	-0.005
2020	-0.012
2021	-0.011
2022	-0.005

Año	Pendiente error vs. total_radiation_pvgis_e5
2017	-0.002
2018	-0.004
2019	-0.006
2020	-0.013
2021	-0.012
2022	-0.000

Se puede observar que durante 3 primeros años de la serie no hay relación aparente entre esta variable y el error pero en 2020 y 2021 sí se puede apreciar una relación inversamente proporcional con el error. También en los días de primavera (en color verde) de 2022.

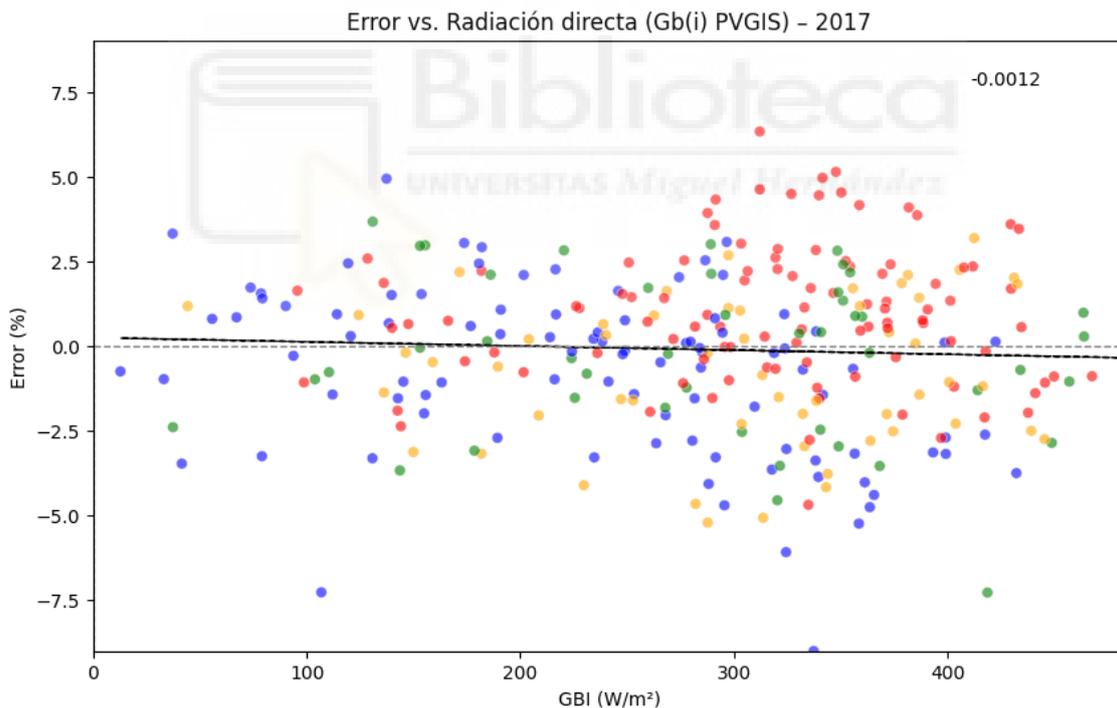
Esta relación se mantiene tanto en las observaciones satelitales (SARAH3 como en los reanálisis E5).

### 3.11. RADIACIÓN DIRECTA

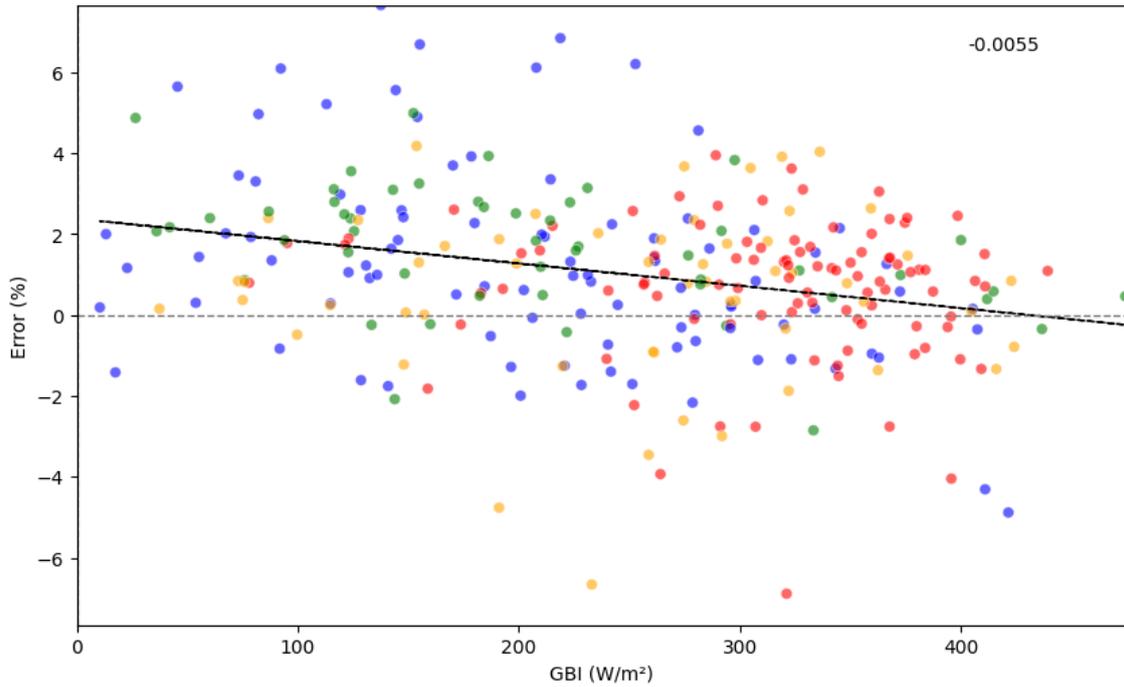
Esta variable (GBI) representa la radiación directa (beam) en  $W/m^2$  en los conjuntos de datos SARA3 y E5 de PVGIS. Se define como la energía directa que llega del sol a la superficie del planeta. La mayoría de la energía aprovechada por los paneles solares procede de este componente de la radiación total, especialmente en regiones soleadas.

Los datos originales se presentaban a nivel horario y se ha seguido la misma metodología que el resto de casos que compartían esta característica: filtrado (08-20h), agregación para obtener la media diaria y comparación con el error del modelo en gráficas de dispersión.

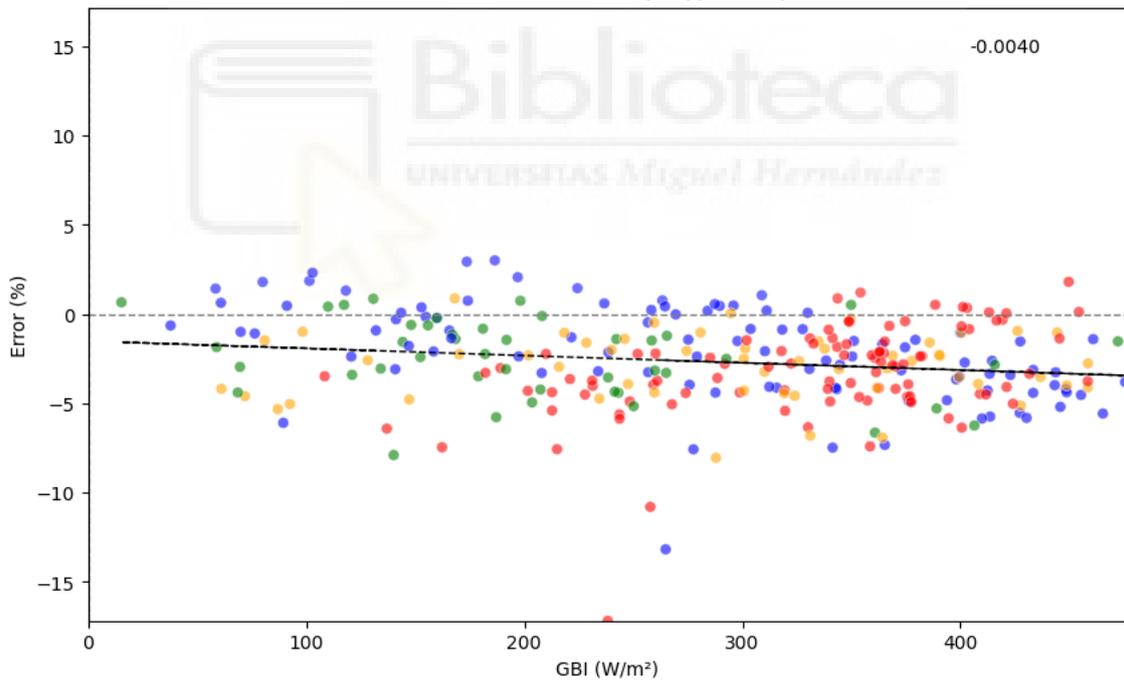
#### 3.11.1. SARA3



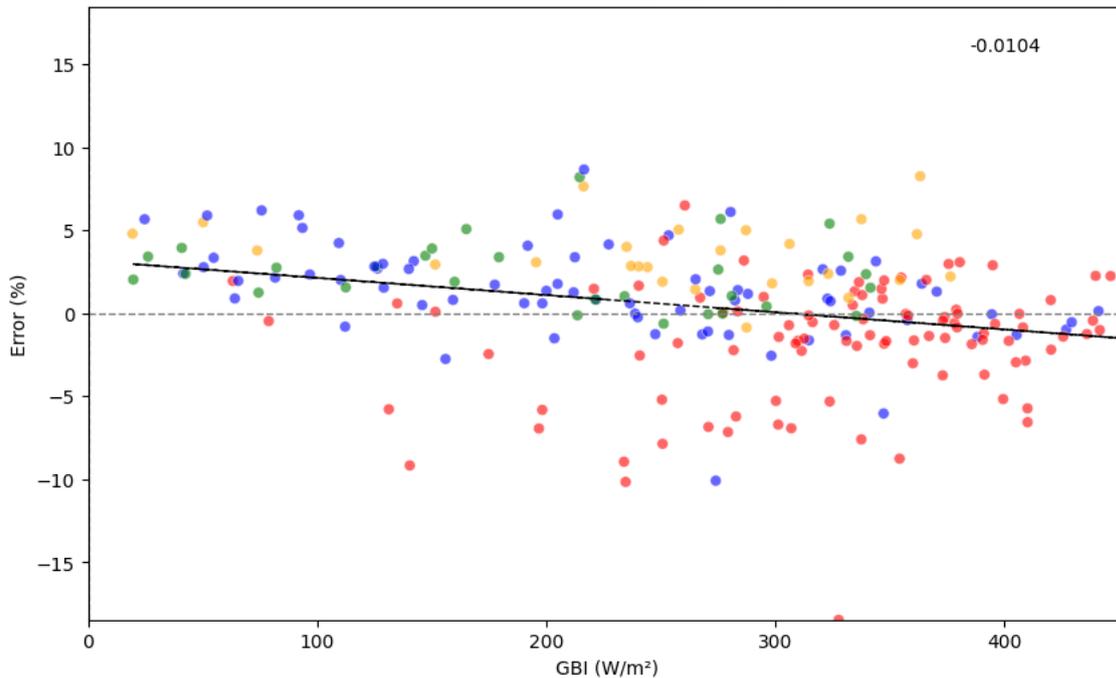
Error vs. Radiación directa (Gb(i) PVGIS) - 2018



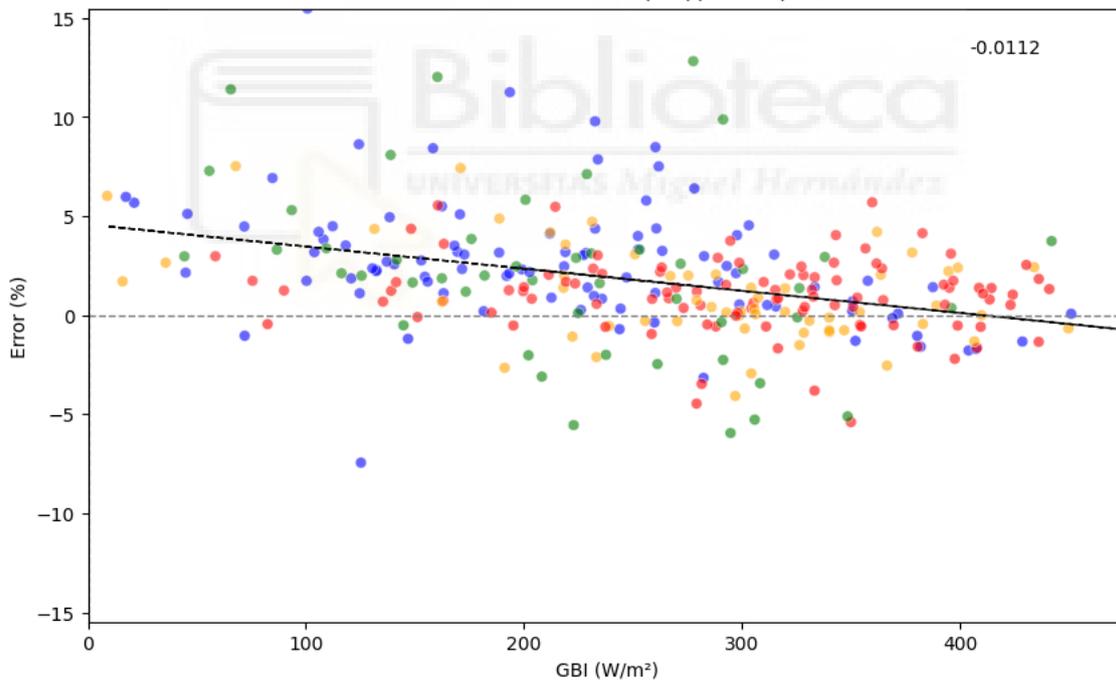
Error vs. Radiación directa (Gb(i) PVGIS) - 2019

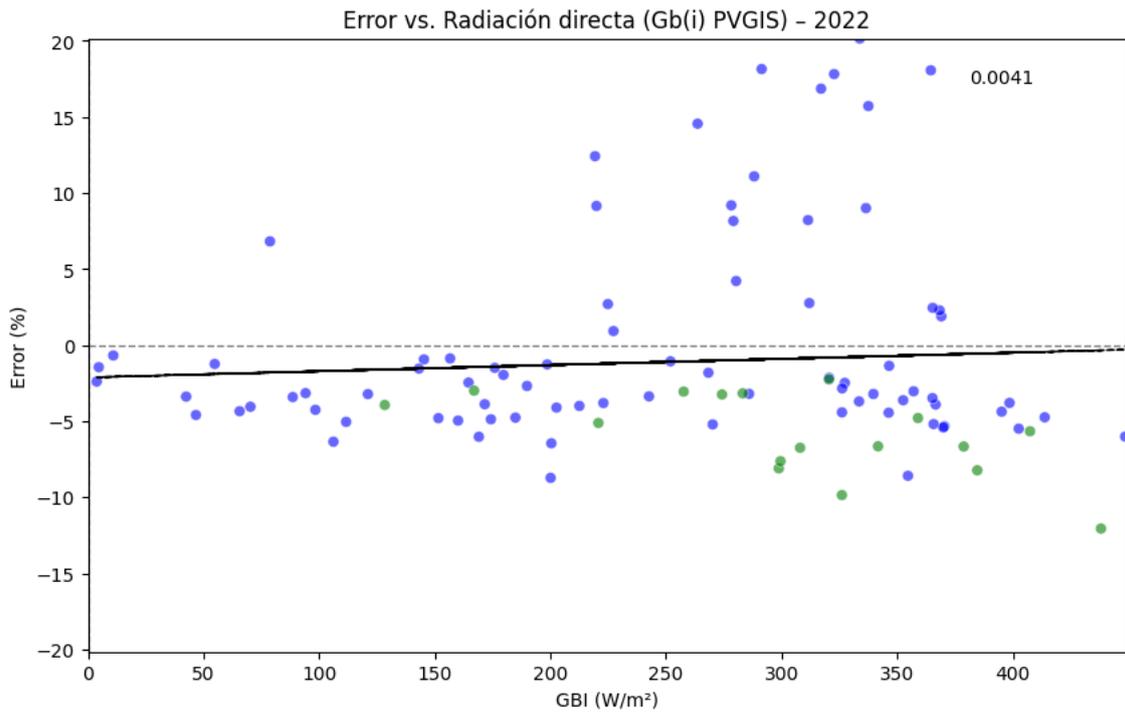


Error vs. Radiación directa (Gb(i) PVGIS) - 2020

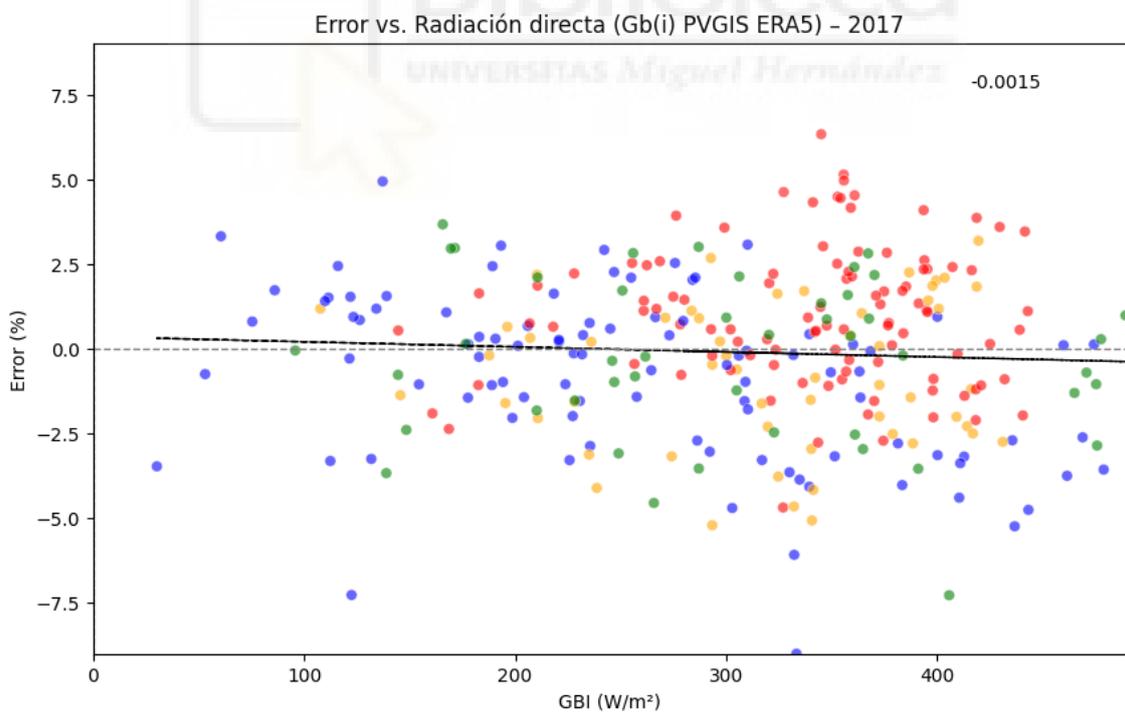


Error vs. Radiación directa (Gb(i) PVGIS) - 2021

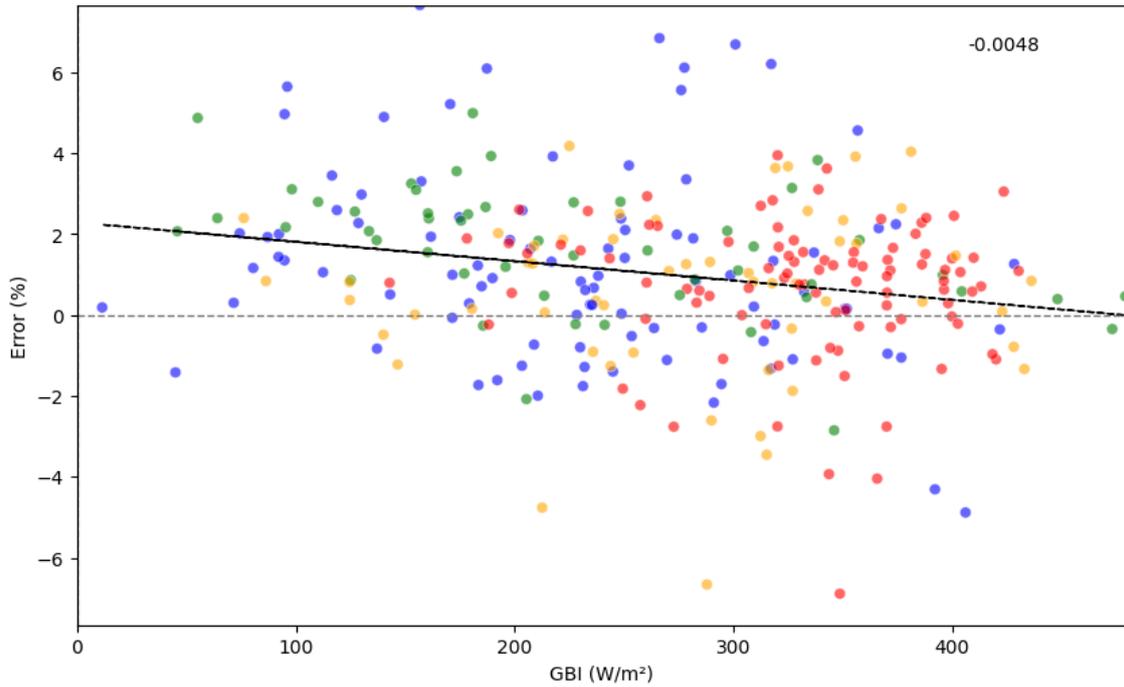




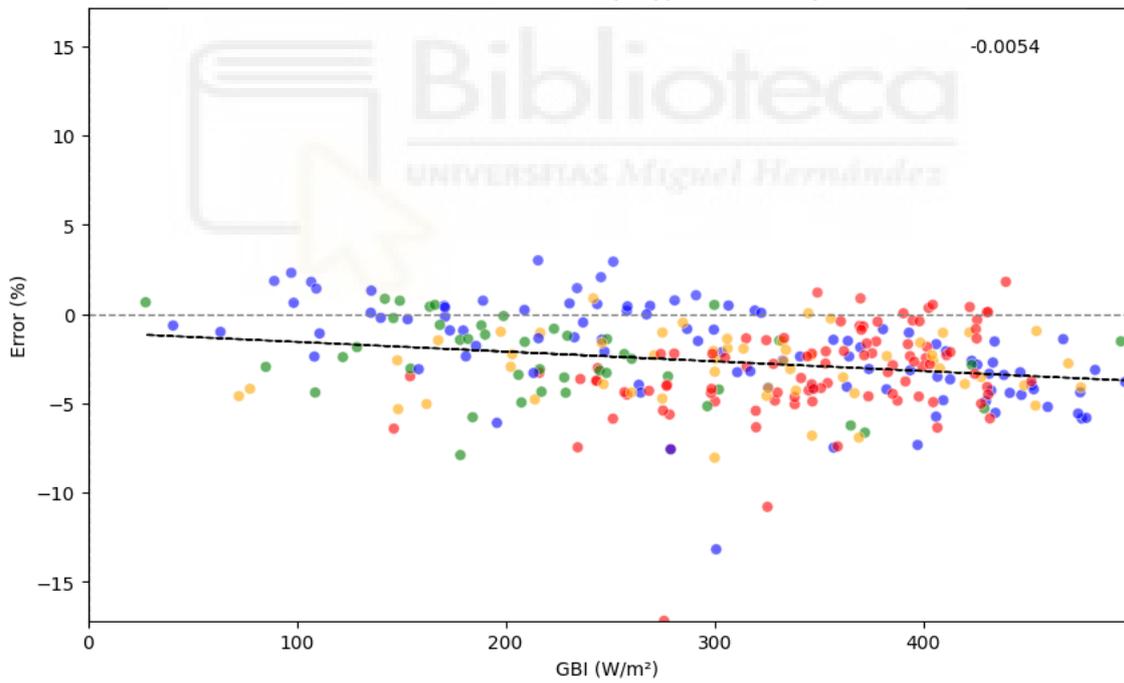
3.11.2. E5



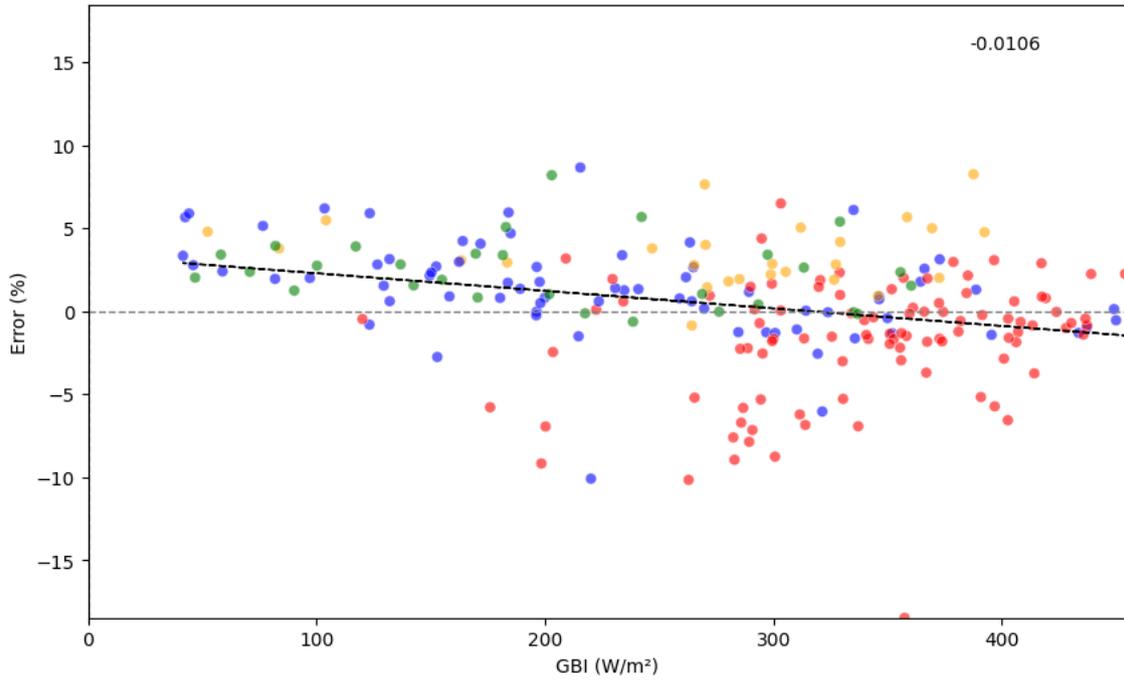
Error vs. Radiación directa (Gb(i) PVGIS ERA5) - 2018



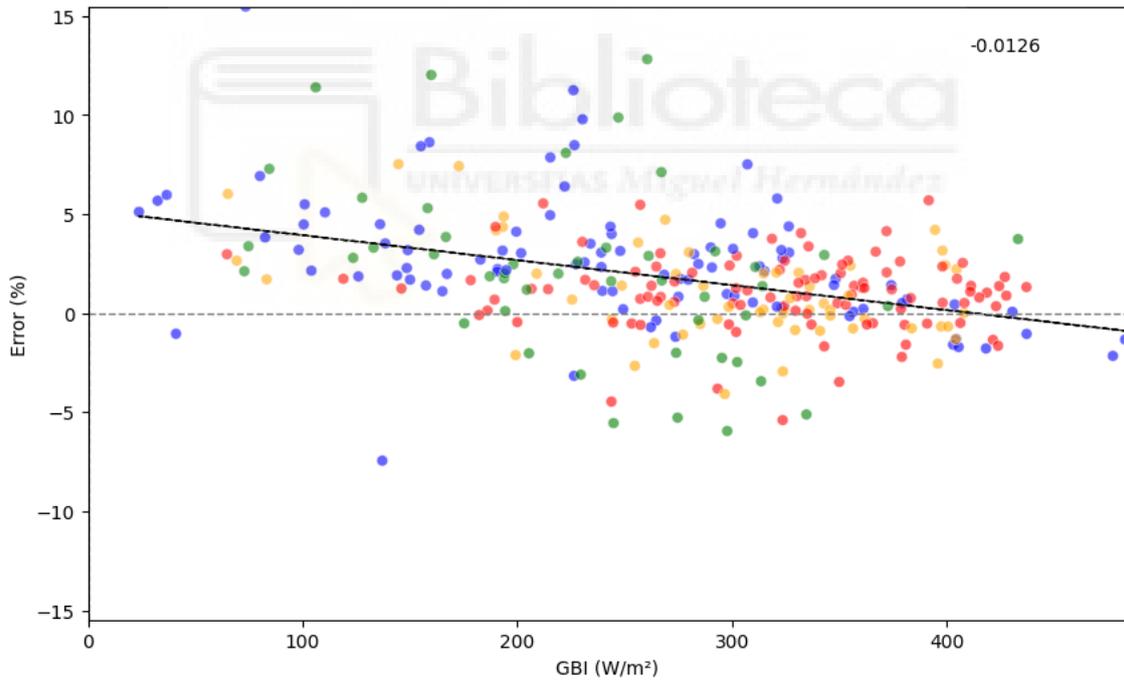
Error vs. Radiación directa (Gb(i) PVGIS ERA5) - 2019

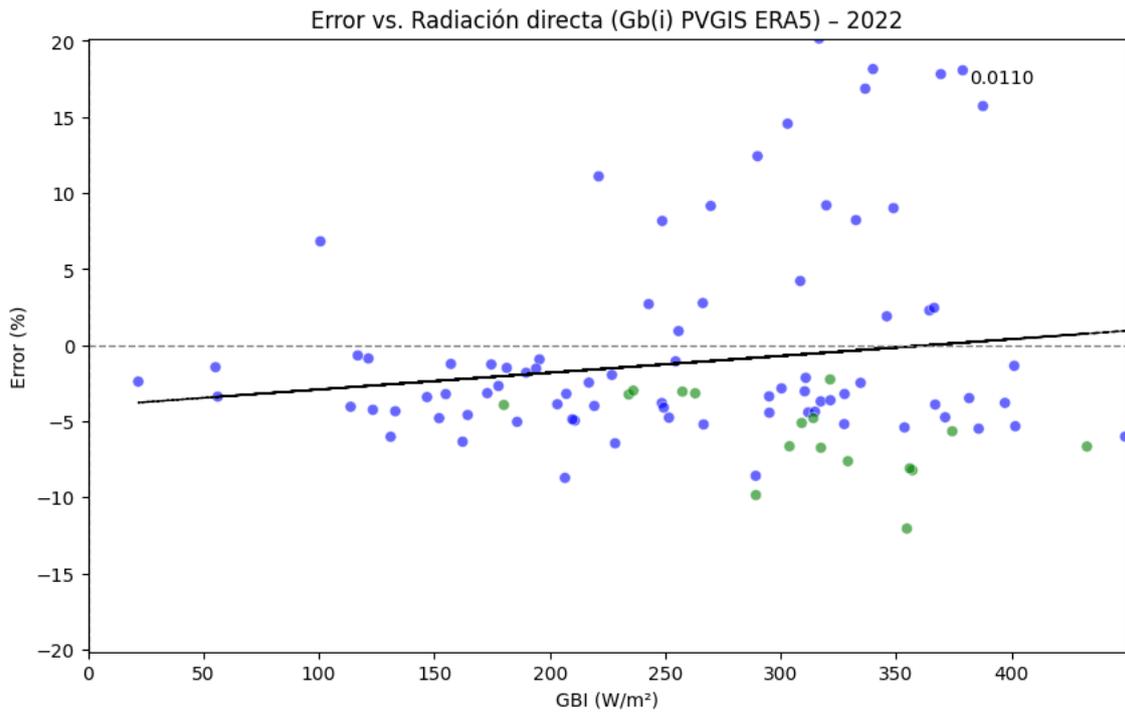


Error vs. Radiación directa (Gb(i) PVGIS ERA5) - 2020



Error vs. Radiación directa (Gb(i) PVGIS ERA5) - 2021





### 3.11.3. Análisis

Año	Pendiente error vs. Gb(i)_pvgis
2017	-0.001
2018	-0.005
2019	-0.004
2020	-0.010
2021	-0.011
2022	0.004

Año	Pendiente error vs. Gb(i)_pvgis_e5
2017	-0.002
2018	-0.005
2019	-0.005
2020	-0.011
2021	-0.013
2022	0.011

Se puede observar que durante 3 primeros años de la serie no hay relación aparente entre esta variable y el error pero en 2020 y 2021 sí se puede apreciar una relación inversamente proporcional con el error aunque las pendientes son suaves. También en los días de primavera (en color verde) de 2022.

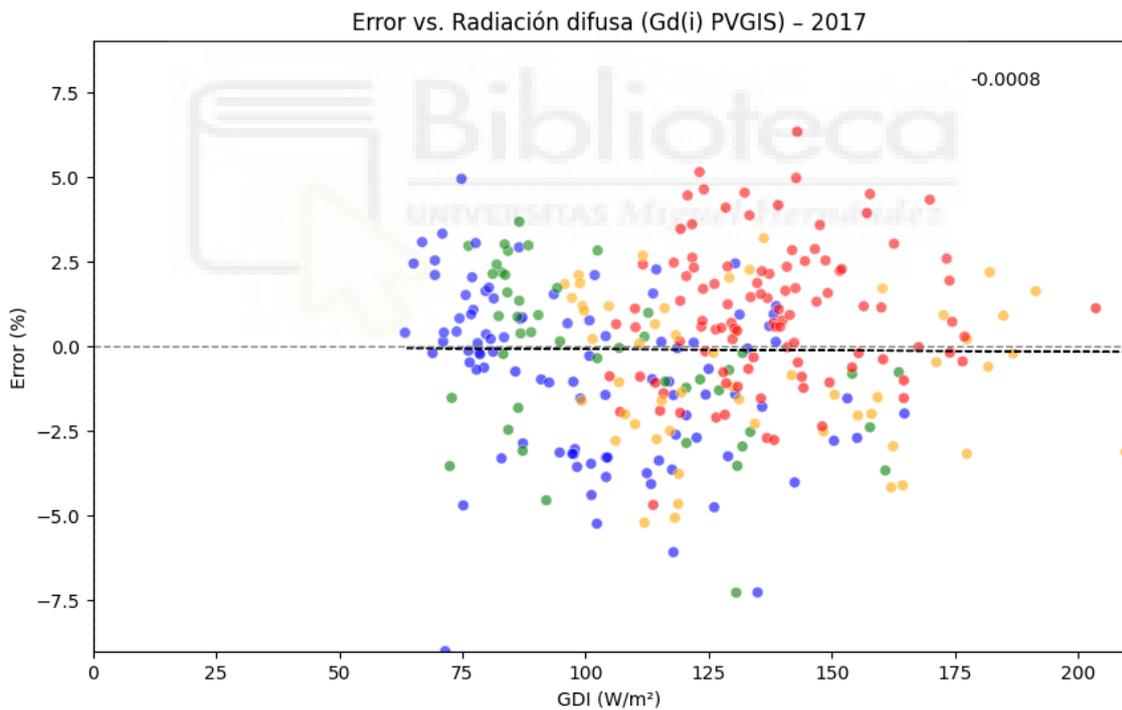
Esta relación se mantiene tanto en las observaciones satelitales (SARAH3) como en el reanálisis (E5).

### 3.12. RADIACIÓN DIFUSA

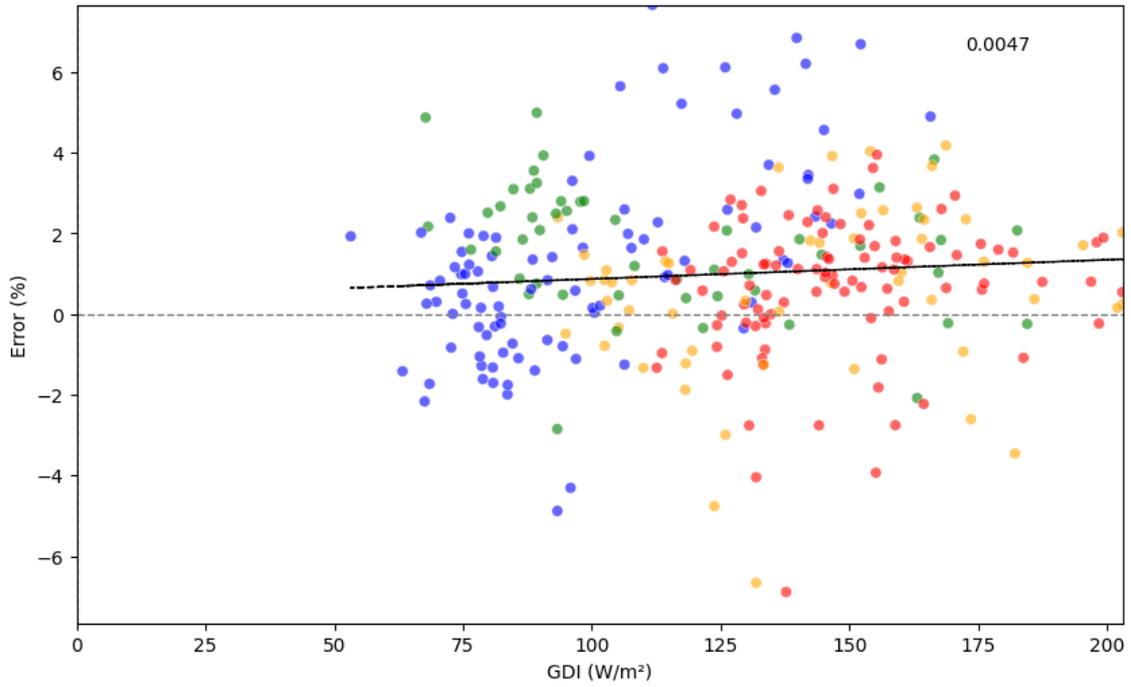
Esta variable (GDI) representa la radiación solar difusa (diffuse) en  $W/m^2$  de los conjuntos de datos SARA3 y ERA5 de PVGIS. Se define como la radiación solar dispersada por las partículas atmosféricas como aerosoles o nubes alcanzando la superficie indirectamente.

Los datos originales estaban distribuidos de forma horaria y se ha seguido la misma metodología que el resto de casos que presentaban esta característica: filtrado (08-20h), agregación para obtener la media diaria y comparación con el error del modelo en gráficas de dispersión.

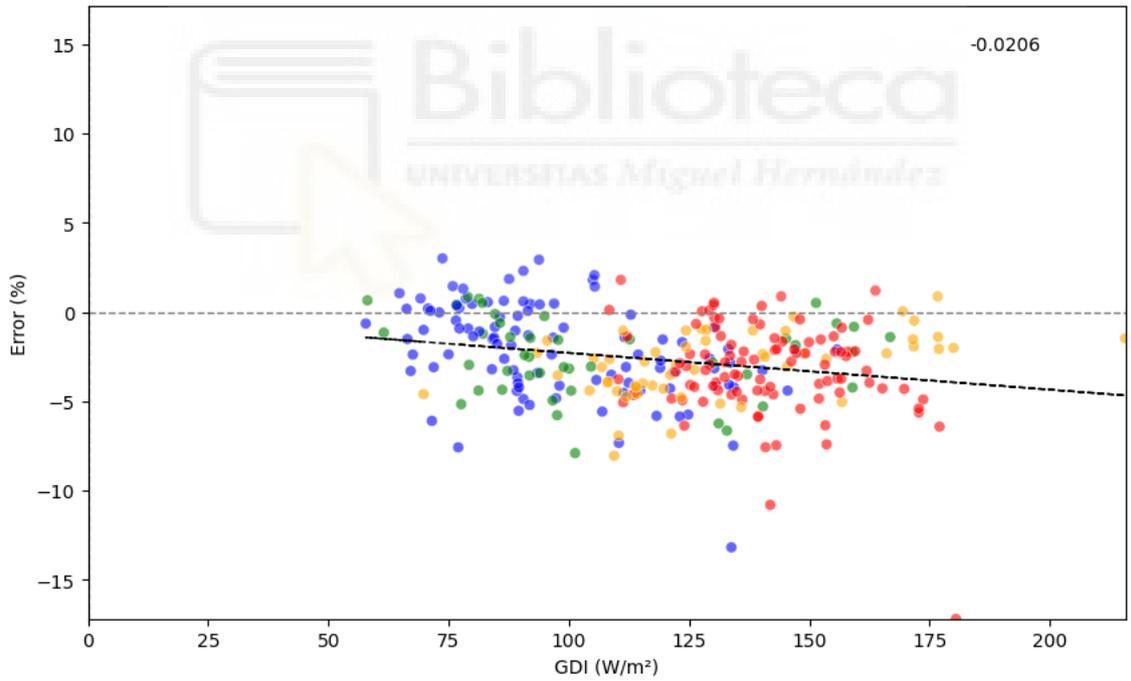
#### 3.12.1. SARA3



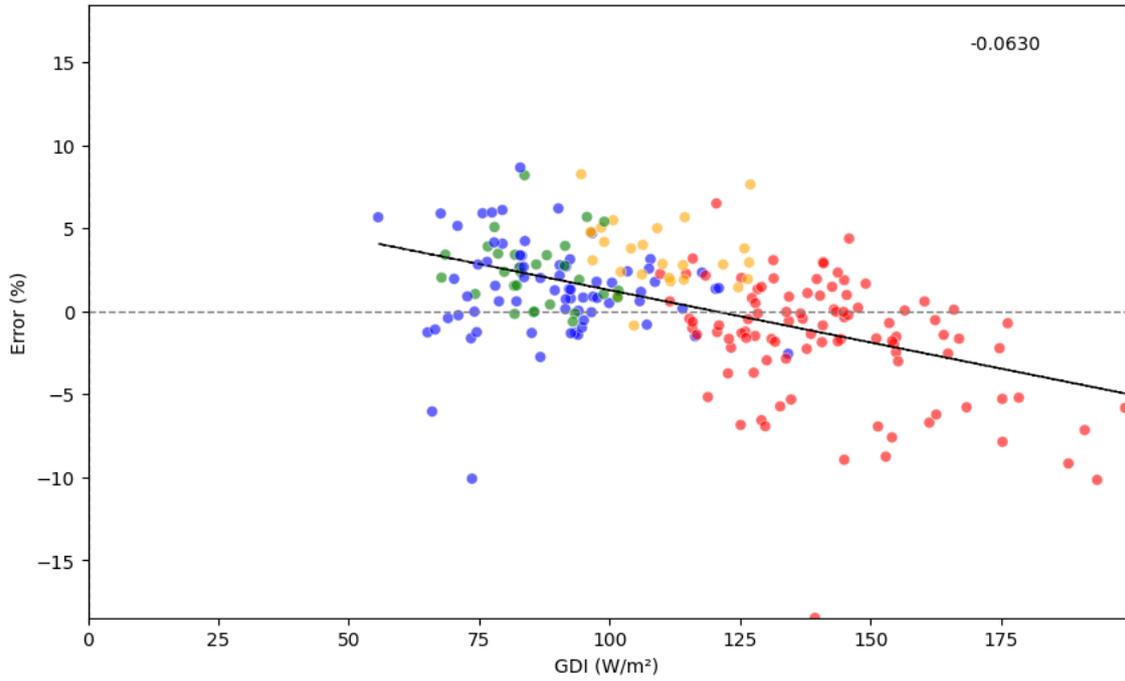
Error vs. Radiación difusa (Gd(i) PVGIS) - 2018



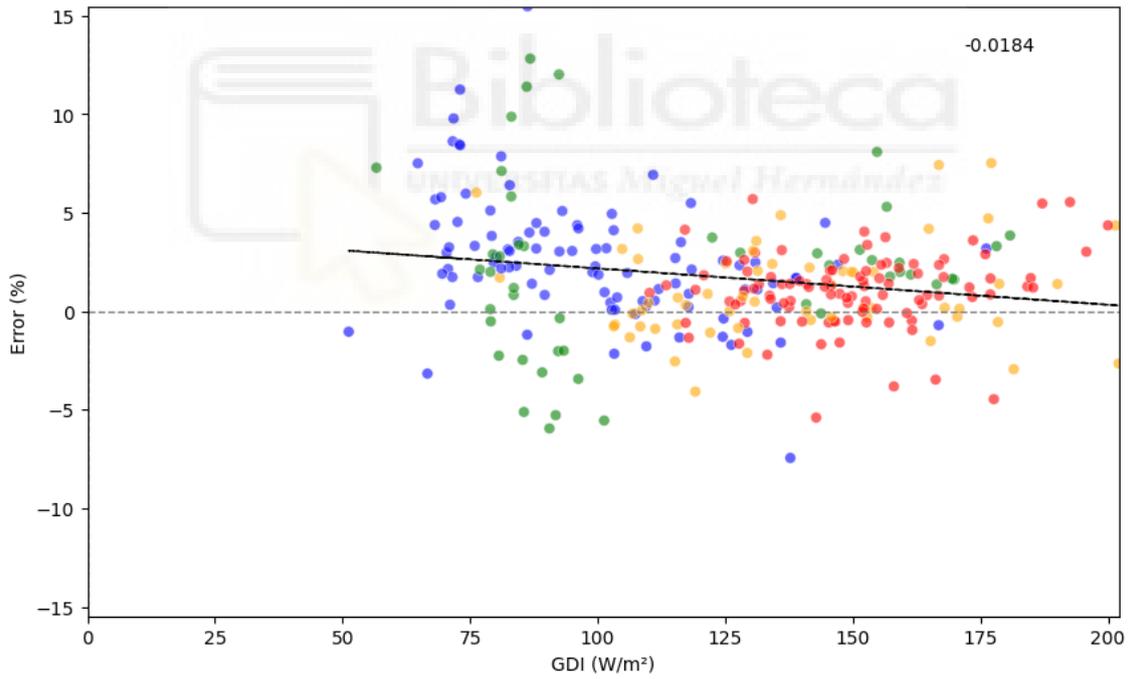
Error vs. Radiación difusa (Gd(i) PVGIS) - 2019

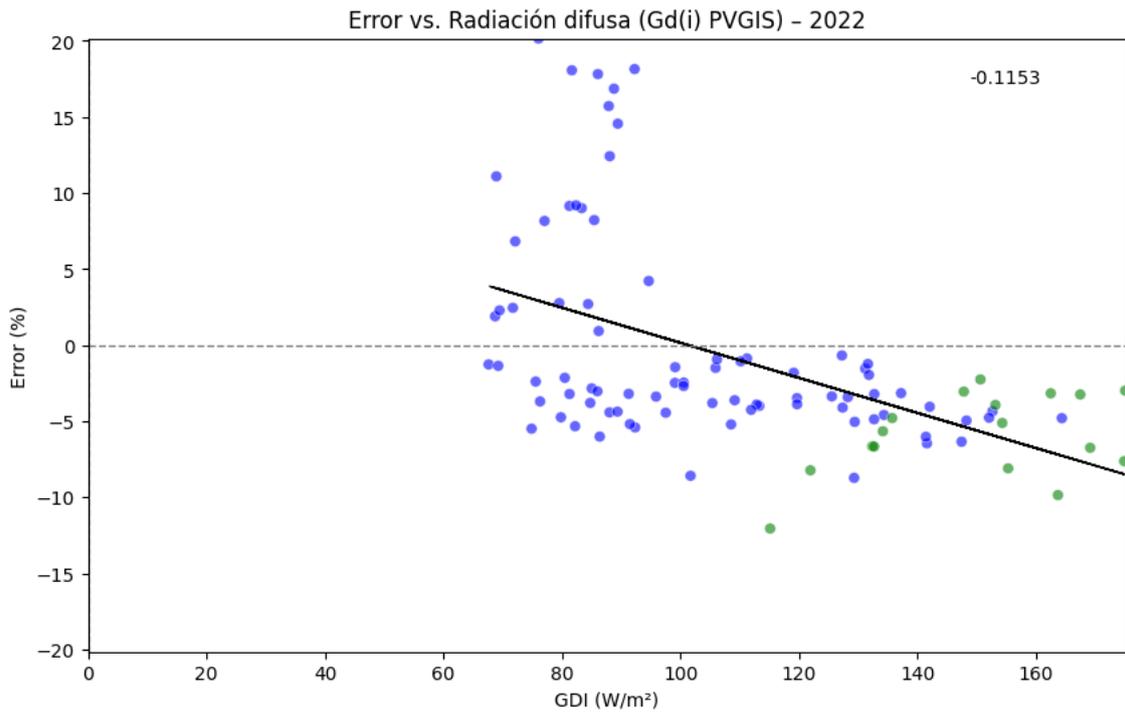


Error vs. Radiación difusa (Gd(i) PVGIS) - 2020

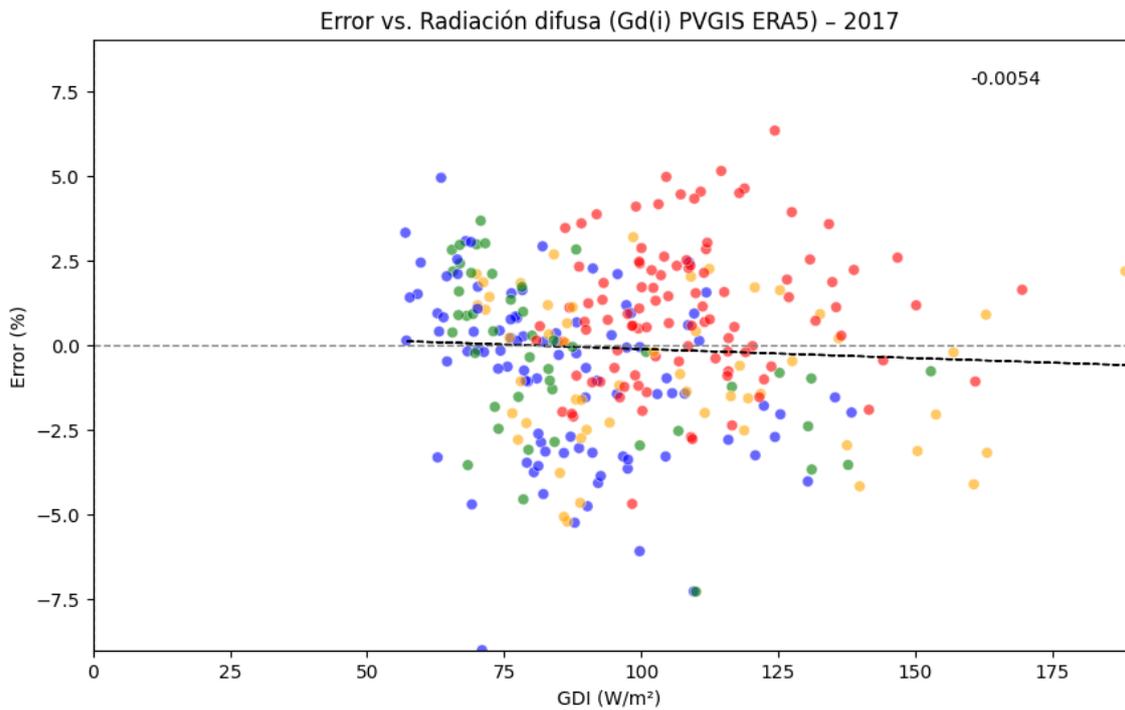


Error vs. Radiación difusa (Gd(i) PVGIS) - 2021

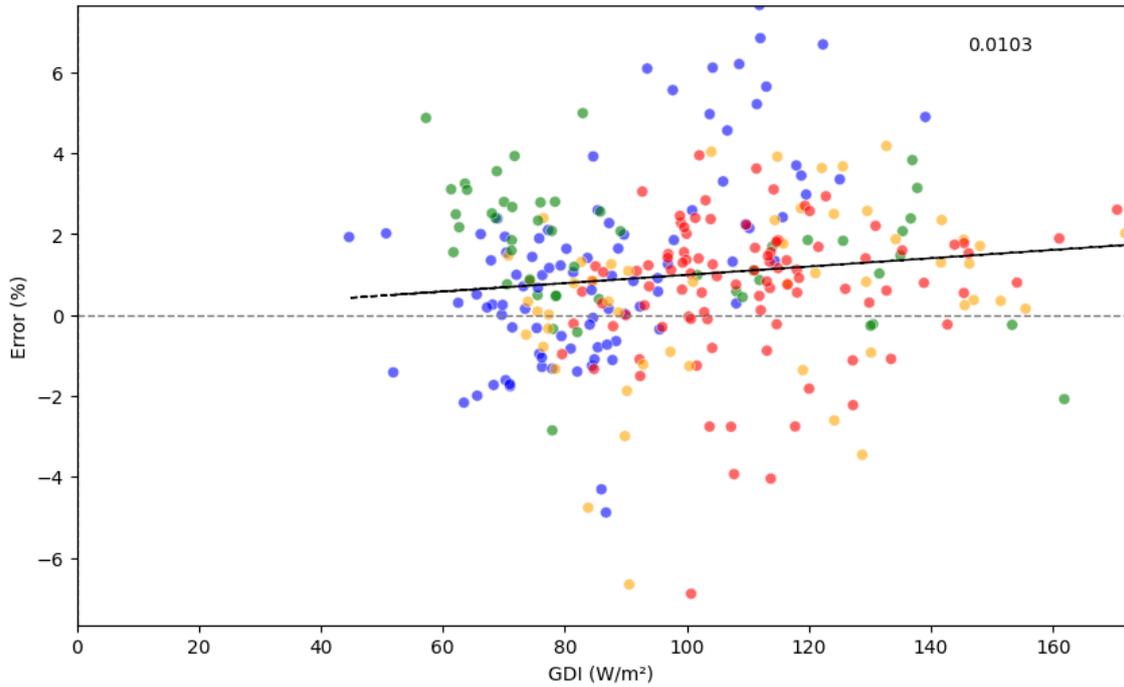




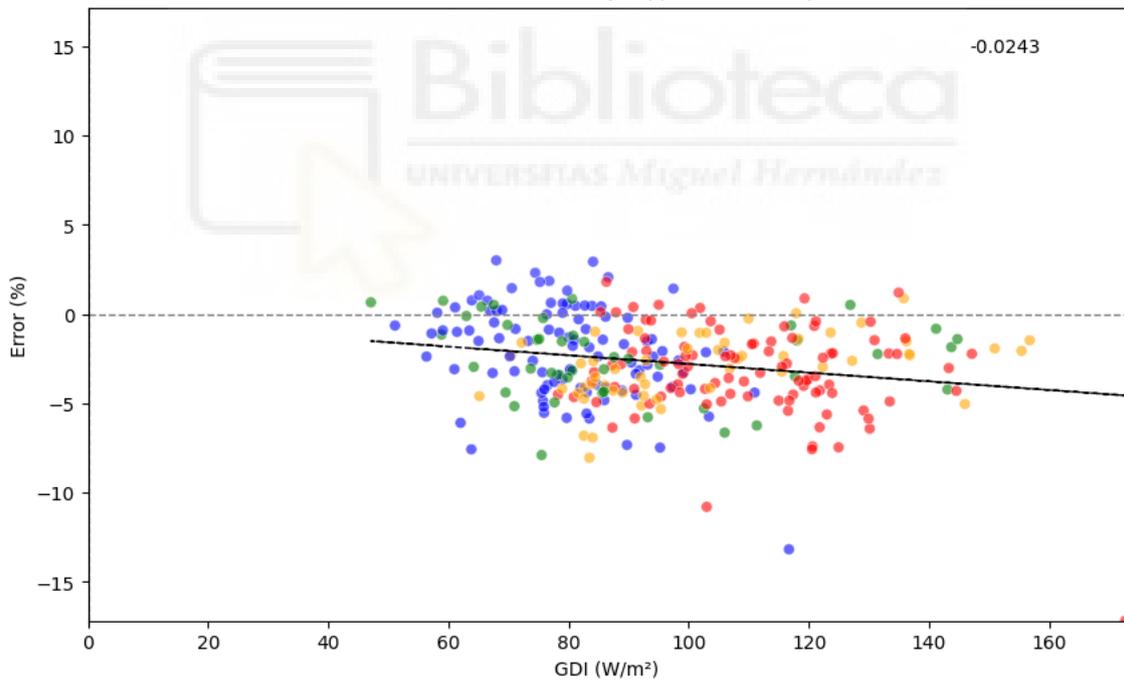
3.12.2. E5



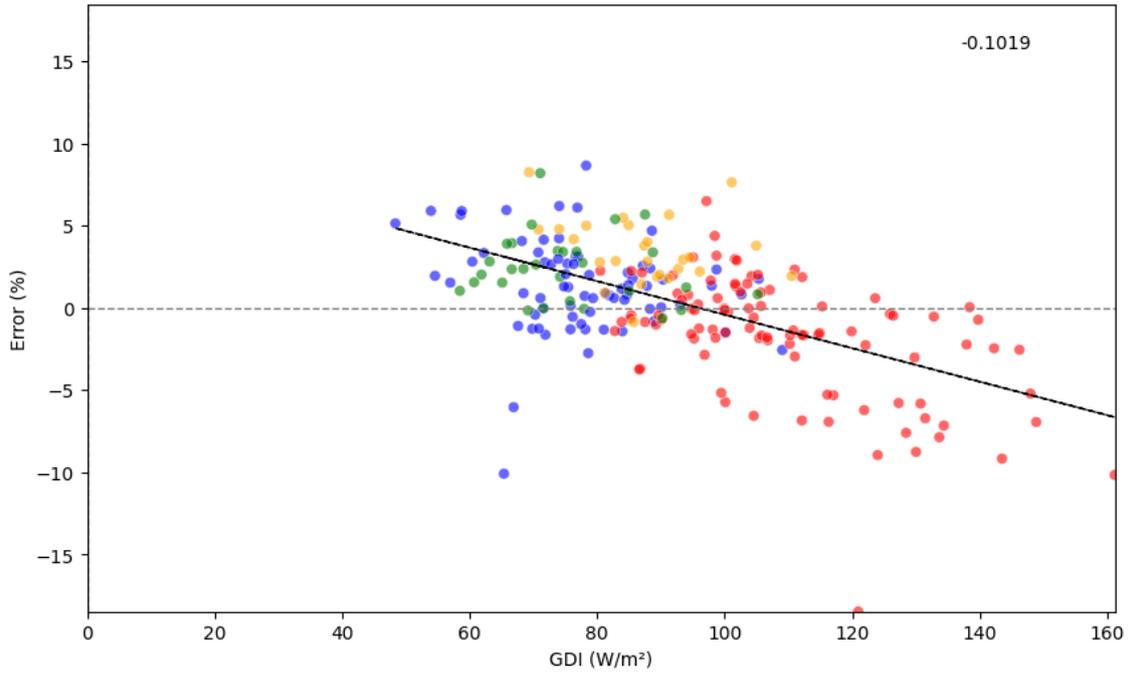
Error vs. Radiación difusa (Gd(i) PVGIS ERA5) - 2018



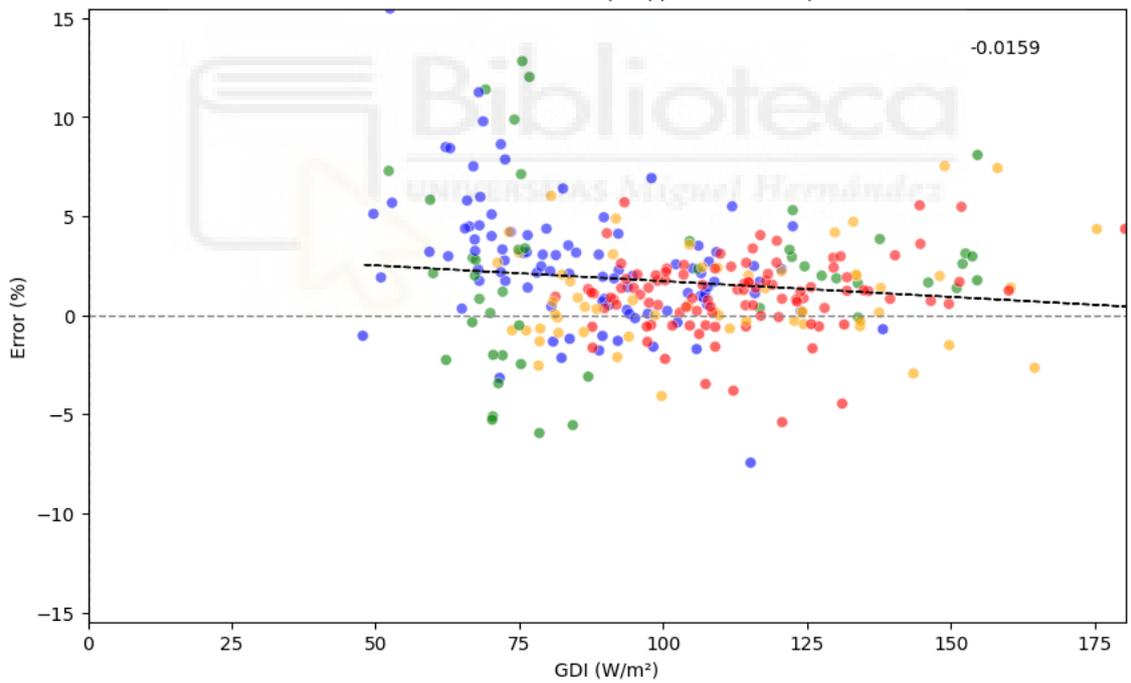
Error vs. Radiación difusa (Gd(i) PVGIS ERA5) - 2019

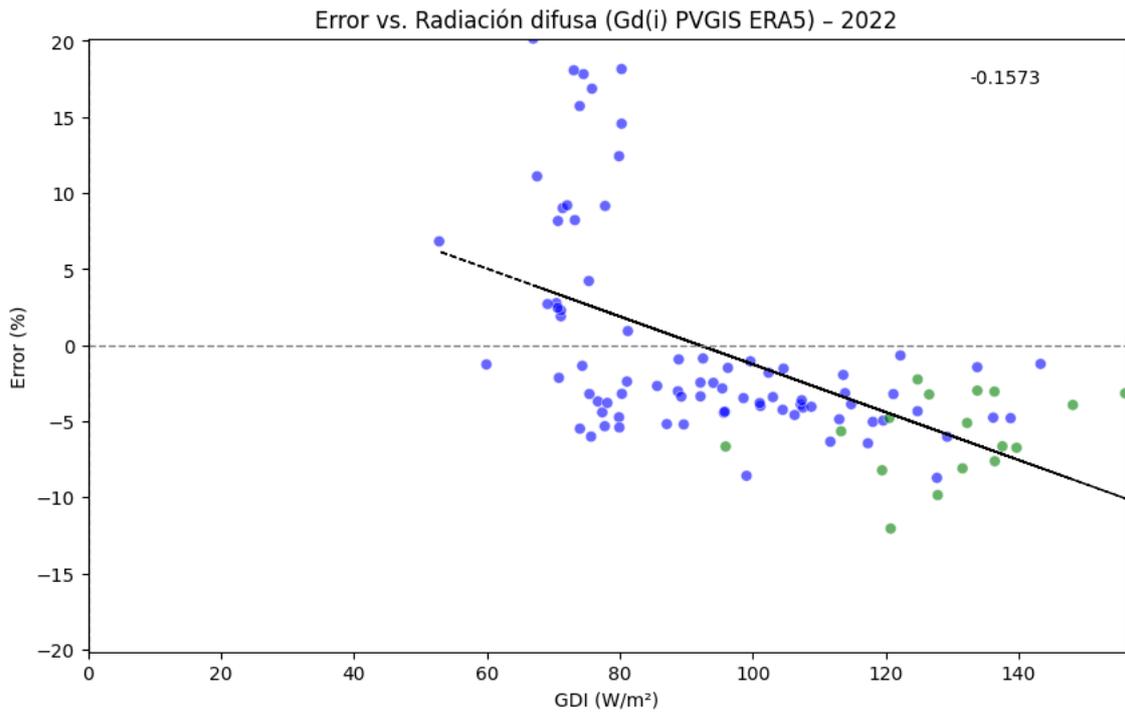


Error vs. Radiación difusa (Gd(i) PVGIS ERA5) - 2020



Error vs. Radiación difusa (Gd(i) PVGIS ERA5) - 2021





### 3.12.3. Análisis

Año	Pendiente error vs. Gd(i)_pvgis
2017	-0.001
2018	0.005
2019	-0.021
2020	-0.063
2021	-0.018
2022	-0.115

Año	Pendiente error vs. Gd(i)_pvgis_e5
2017	-0.005
2018	0.010
2019	-0.024
2020	-0.102
2021	-0.016
2022	-0.157

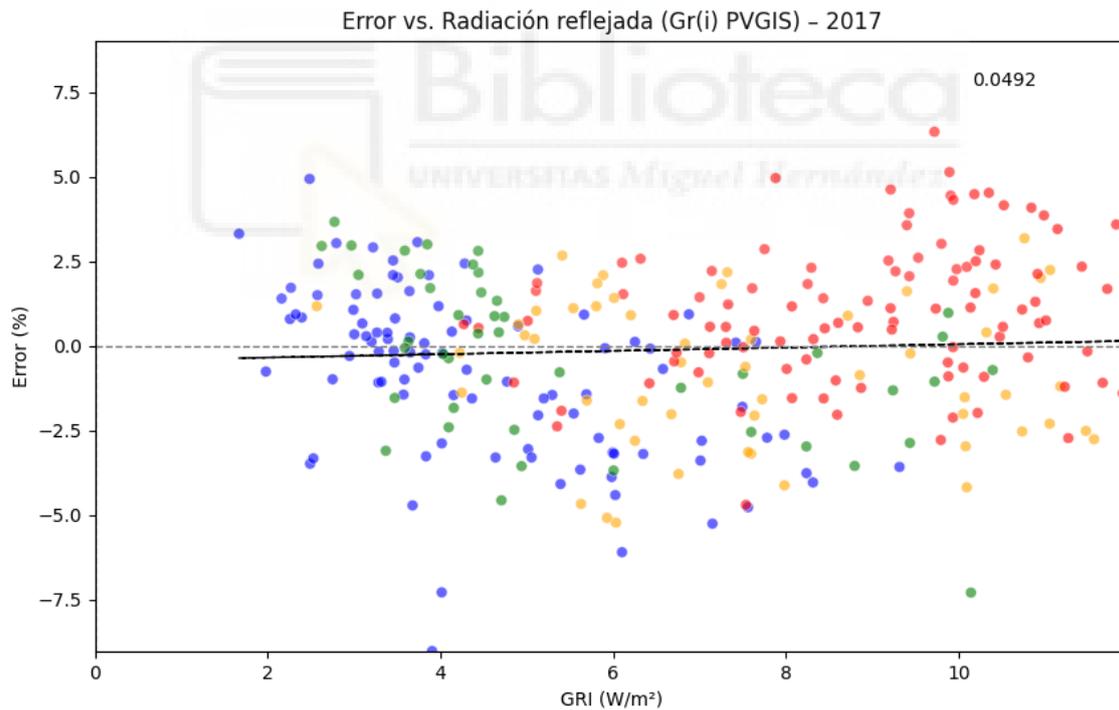
Se aprecia en los gráficos que en los primeros años de la serie de datos no hay una relación evidente entre el error del modelo y esta variable, pero en 2019 y 2020 la relación es inversamente proporcional. En 2021 esta relación se atenúa y la pendiente de la curva es más suave.

### 3.13. RADIACIÓN REFLEJADA

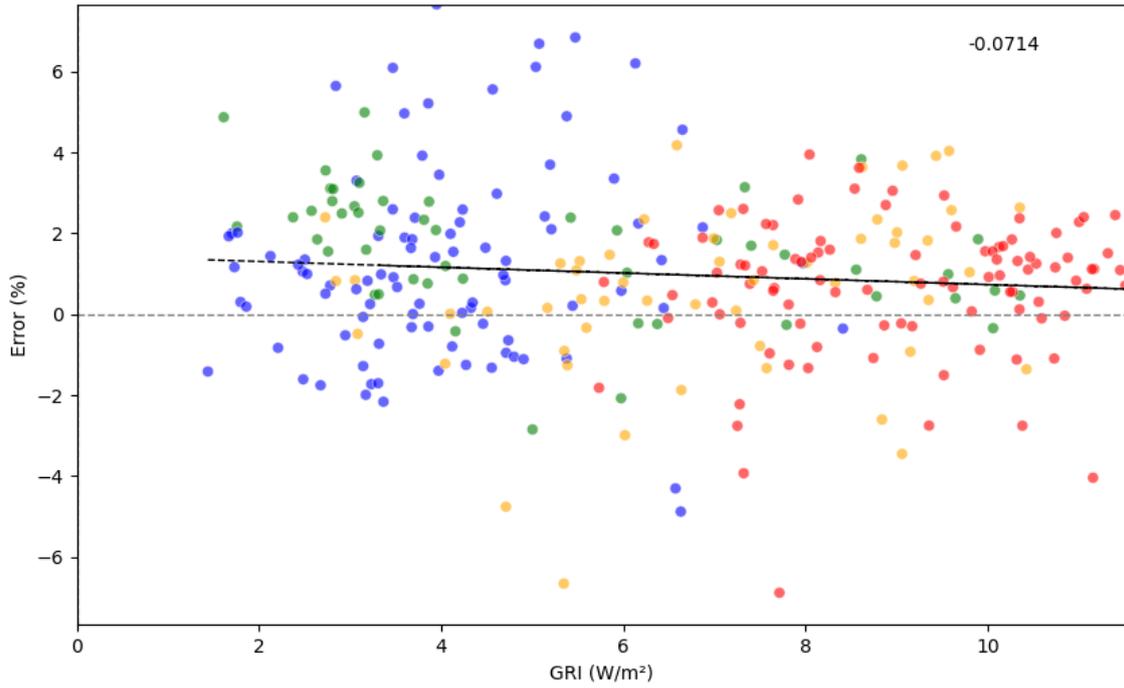
Esta variable (GRI) representa la radiación solar reflejada (reflected) en W/m<sup>2</sup> de los conjuntos de datos SARA3 y ERA5 de PVGIS. Se define como la fracción de la radiación solar que se refleja en el panel. La magnitud depende de la reflectividad de la superficie (albedo).

Los datos de esta variable presentaban una granularidad horaria y se ha seguido la misma metodología que en otras variables en esta misma situación: filtrado (08-20h), agregación para obtener la media diaria y comparación con el error del modelo en gráficas de dispersión.

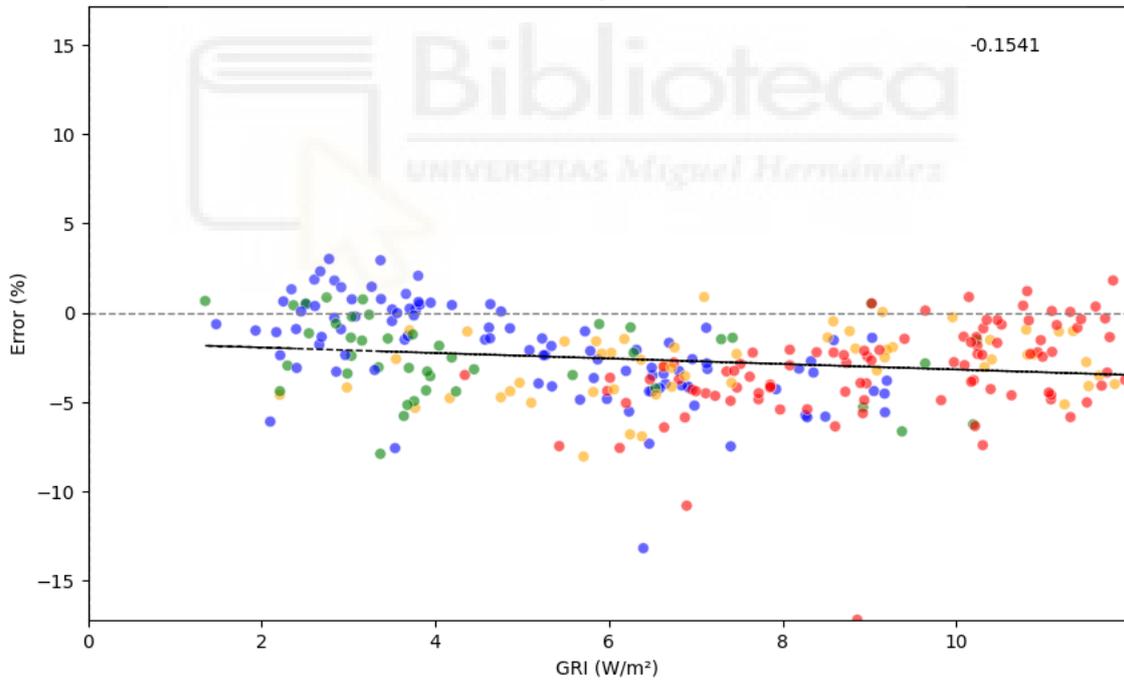
#### 3.13.1. SARA3



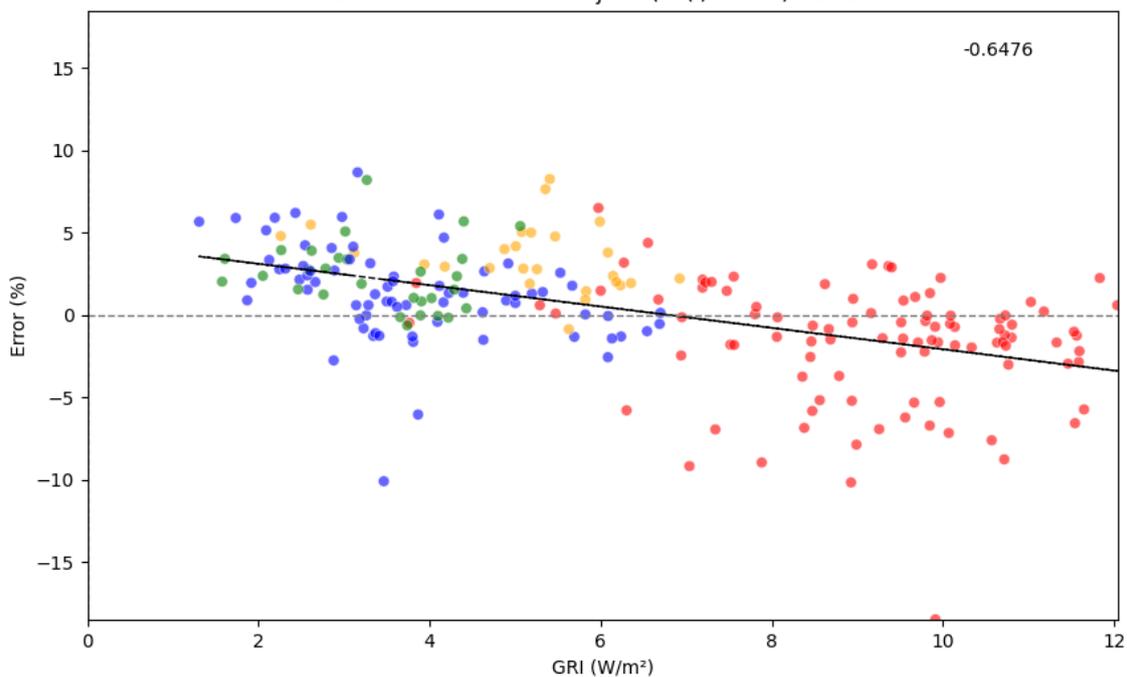
Error vs. Radiación reflejada (Gr(i) PVGIS) - 2018



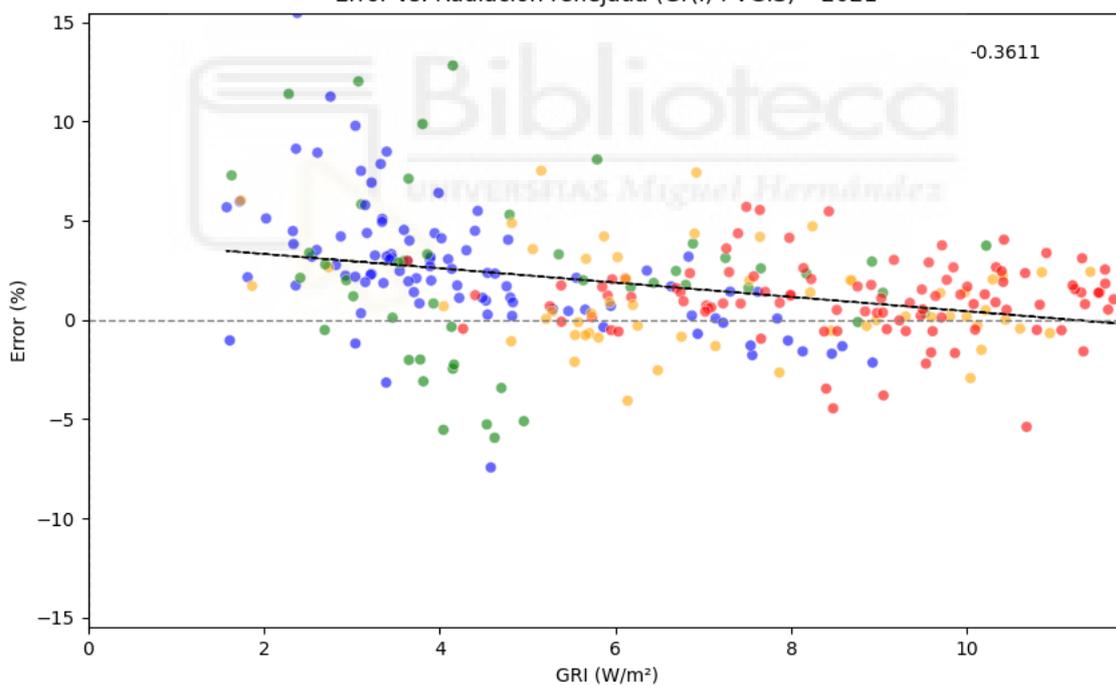
Error vs. Radiación reflejada (Gr(i) PVGIS) - 2019

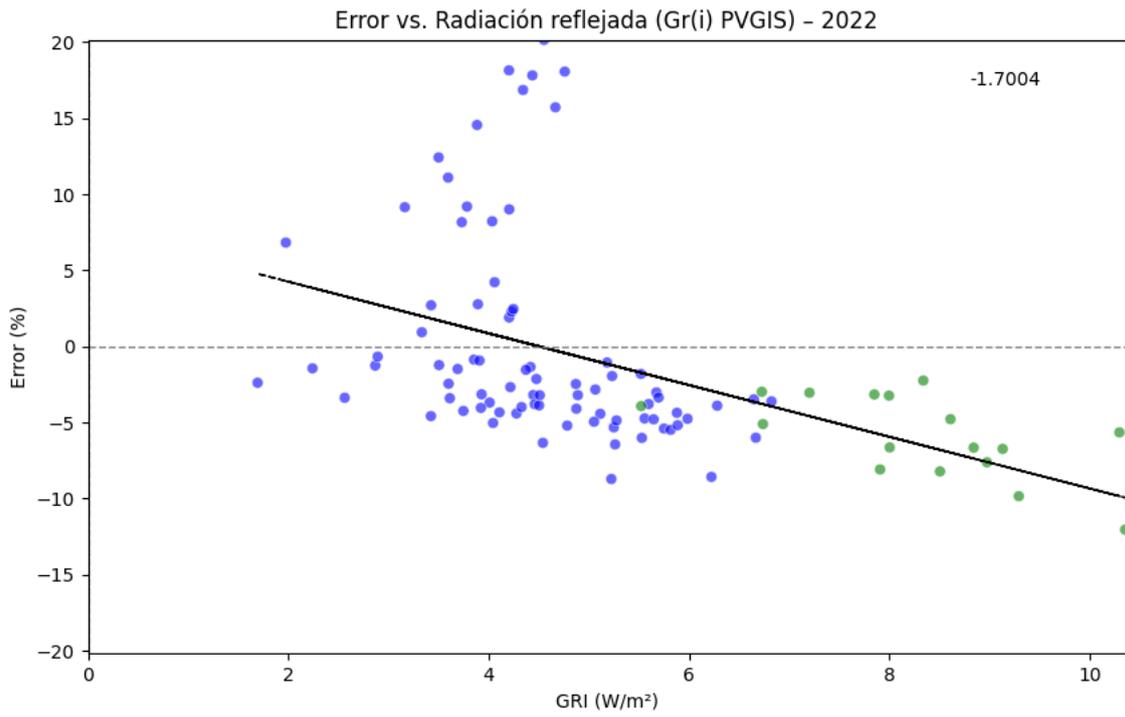


Error vs. Radiación reflejada (Gr(i) PVGIS) - 2020

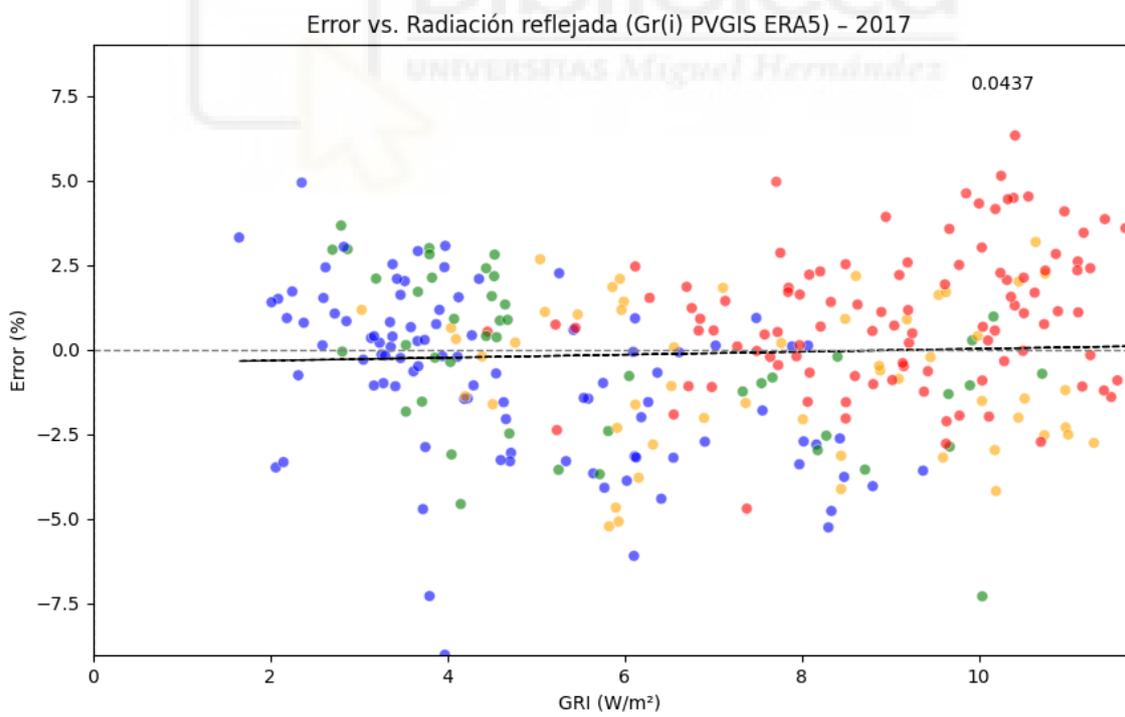


Error vs. Radiación reflejada (Gr(i) PVGIS) - 2021

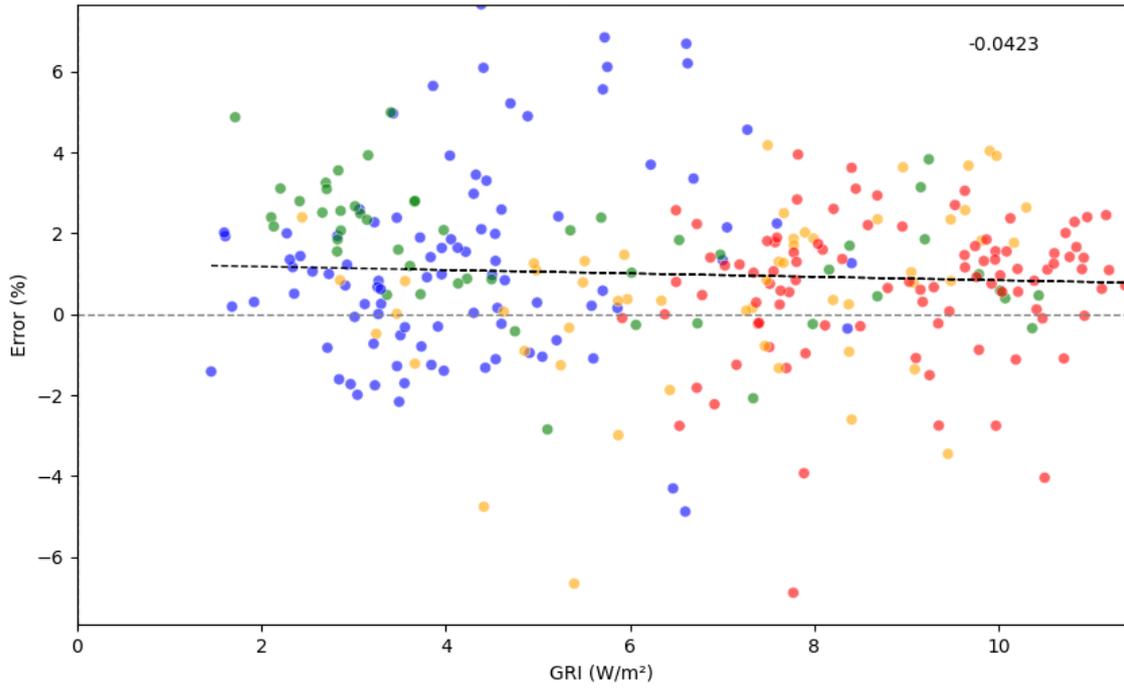




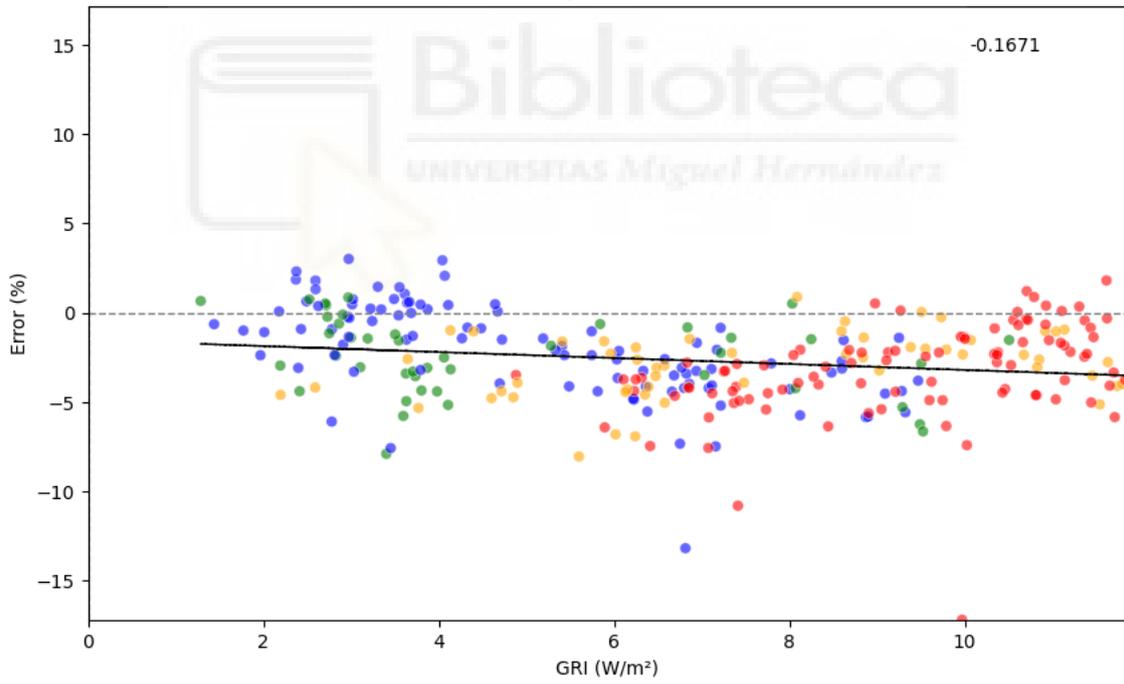
3.13.2. E5



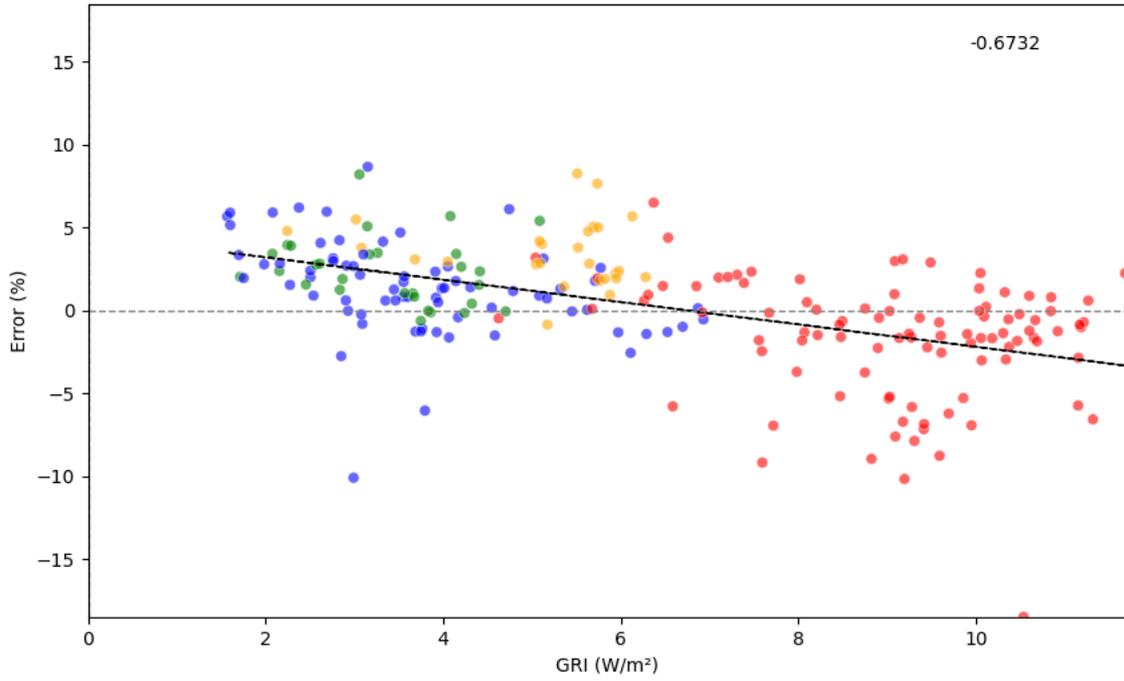
Error vs. Radiación reflejada (Gr(i) PVGIS ERA5) - 2018



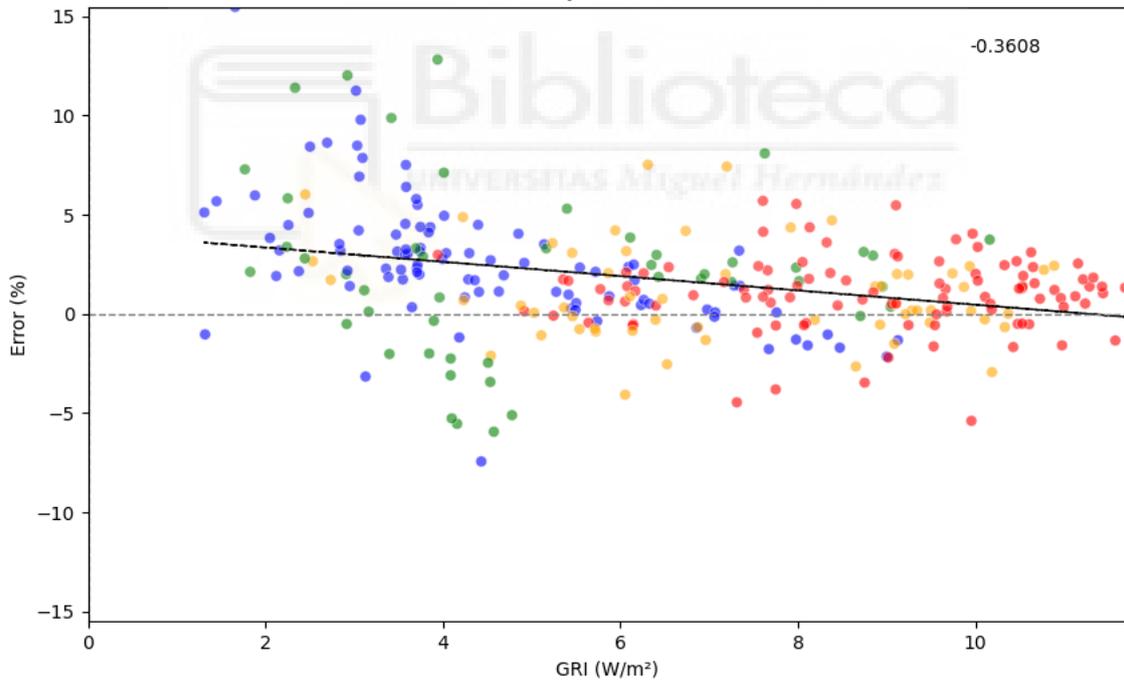
Error vs. Radiación reflejada (Gr(i) PVGIS ERA5) - 2019

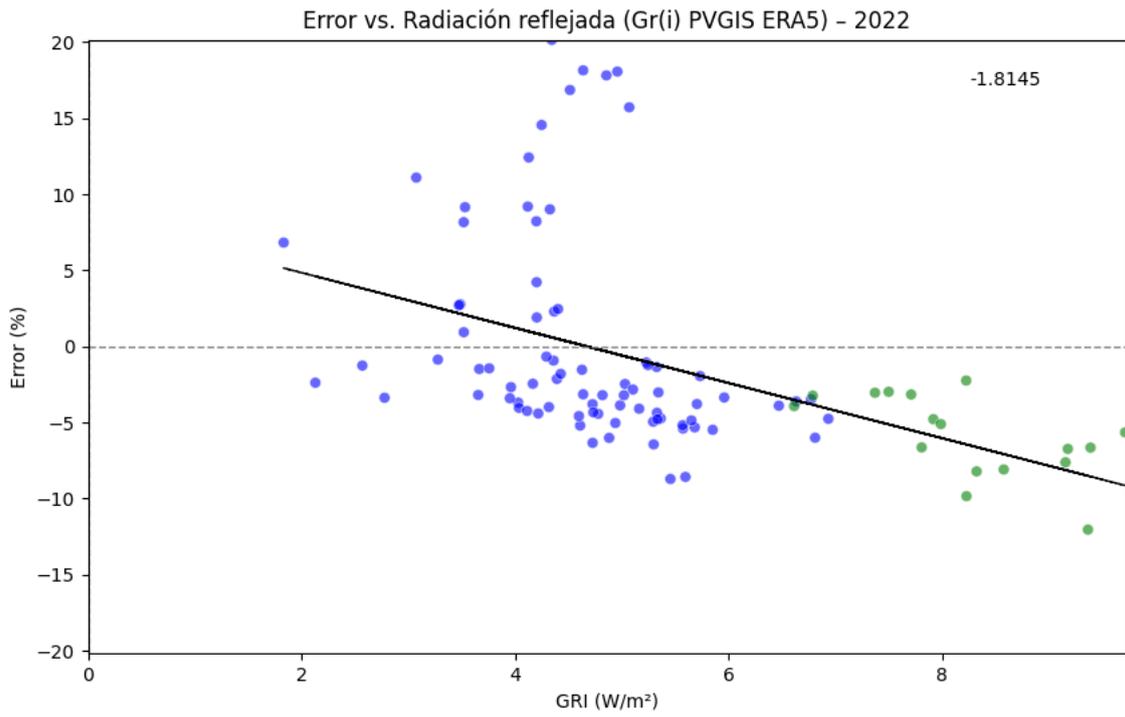


Error vs. Radiación reflejada (Gr(i) PVGIS ERA5) - 2020



Error vs. Radiación reflejada (Gr(i) PVGIS ERA5) - 2021





### 3.13.3. Análisis

Año	Pendiente error vs. Gr(i)_pvgis
2017	0.049
2018	-0.071
2019	-0.154
2020	-0.648
2021	-0.361
2022	-1.700

Año	Pendiente error vs. Gr(i)_pvgis_e5
2017	0.044
2018	-0.042
2019	-0.167
2020	-0.673
2021	-0.361
2022	-1.815

Se repite la dinámica observada en los otros componentes con la pendiente de la línea de regresión plana o muy suave los primeros años de la serie y negativa a partir de 2020.

### 3.14. POTENCIA FOTOVOLTAICA ESTIMADA

Esta variable (P) hace referencia a la potencia estimada (W) alcanzada por un sistema fotovoltaico teórico con las siguientes características:

**Mounting type:**\*

Fixed     Vertical axis     Inclined axis     Two axis

Slope [°]         Optimize slope

Azimuth [°]         Optimize slope and azimuth

PV power

PV technology\*   

Installed peak PV power [kWp]\*   

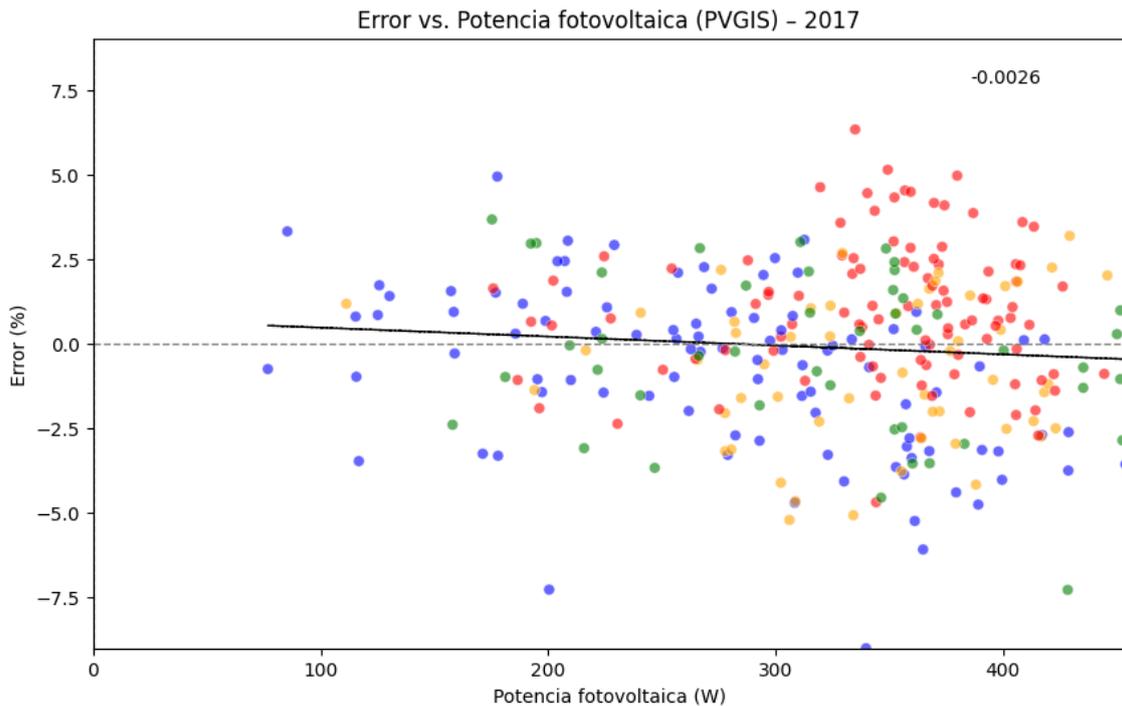
System loss [%]\*   

Radiation components

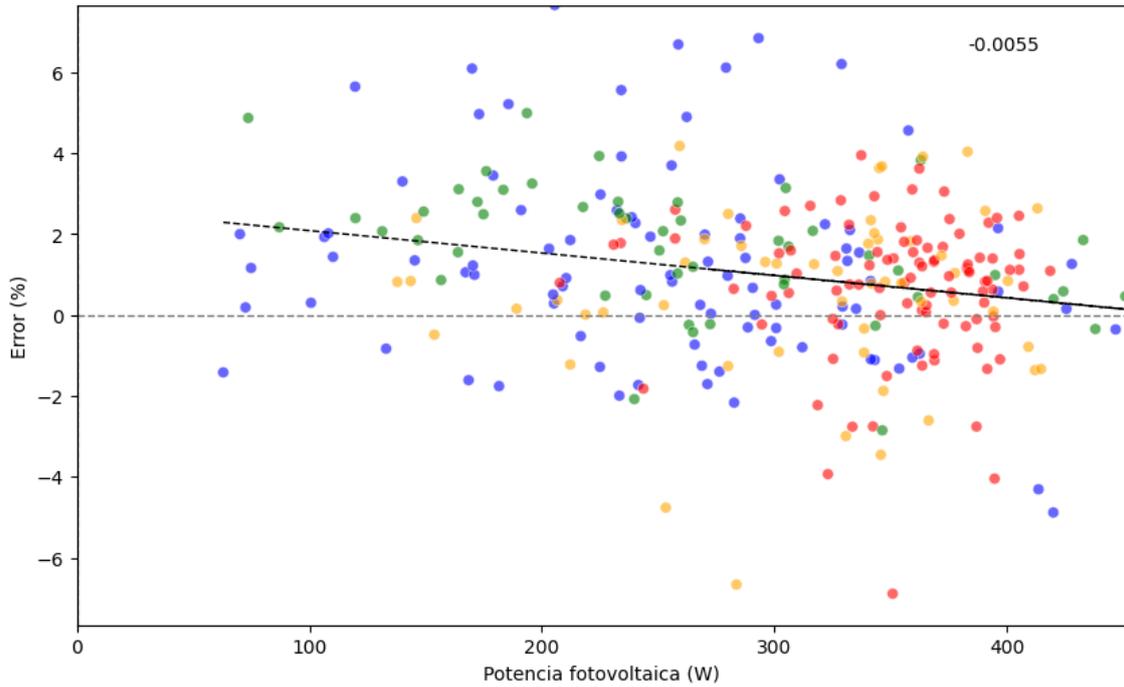
Utilizando los datos de PVGIS, tanto SARA3 como ERA5.

Los datos de esta variable estaban distribuidos de forma horaria y se ha seguido la misma metodología que en otras variables en esta misma situación: filtrado (08-20h), agregación para obtener la media diaria y comparación con el error del modelo en gráficas de dispersión.

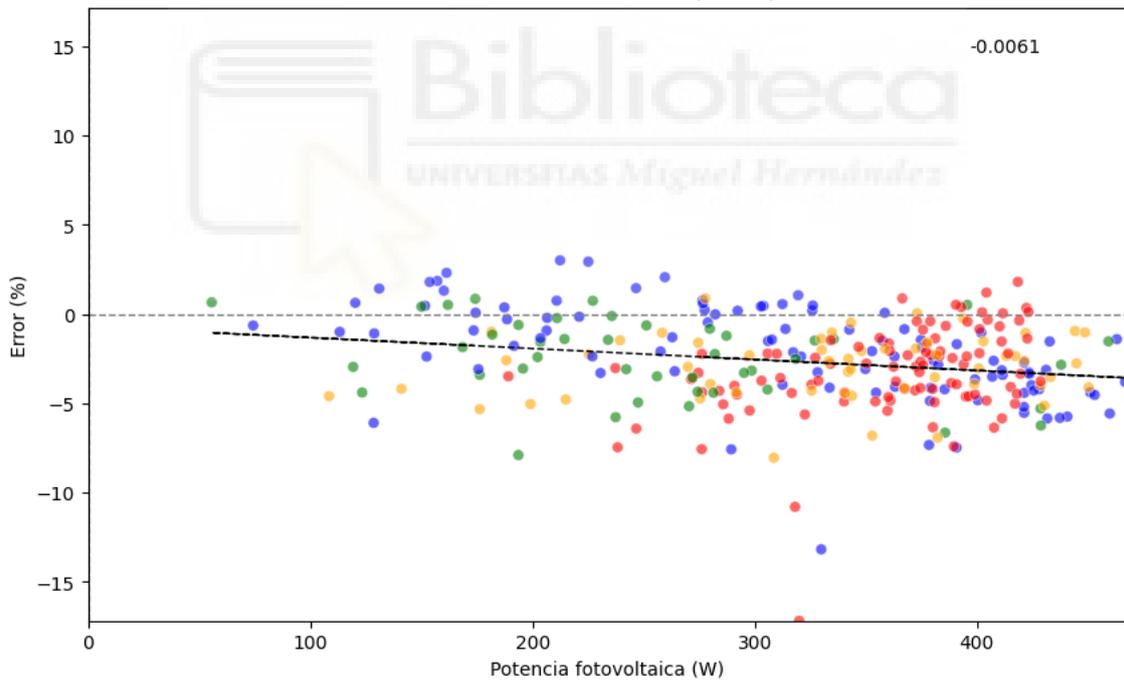
#### 3.14.1. SARA3



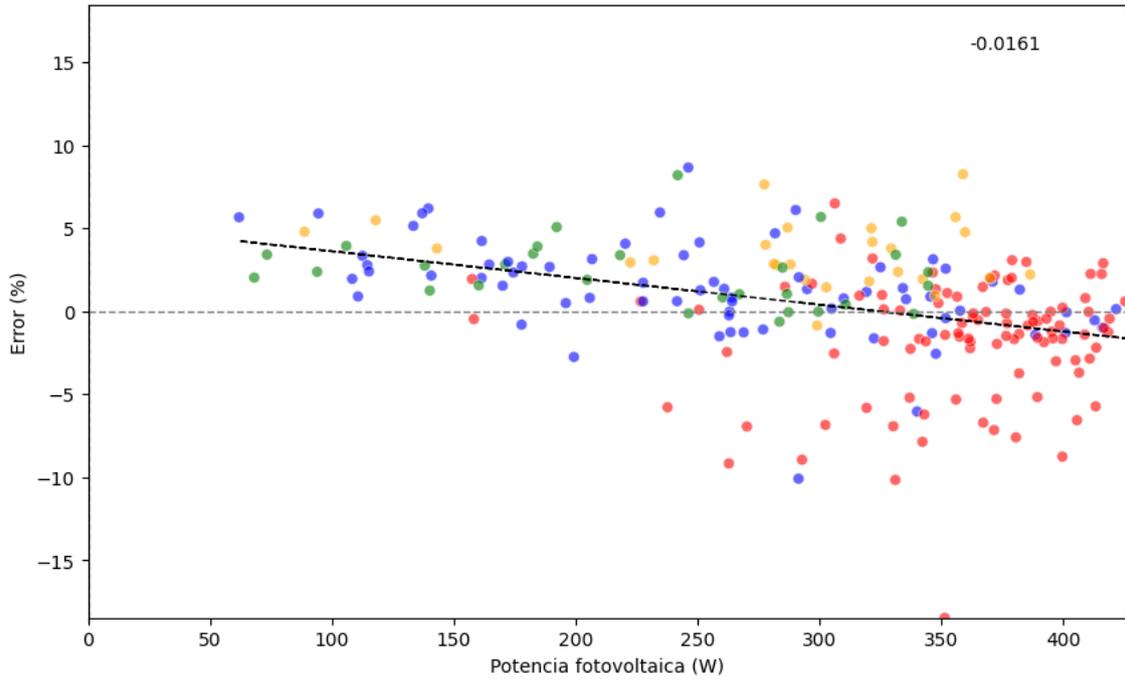
Error vs. Potencia fotovoltaica (PVGIS) - 2018



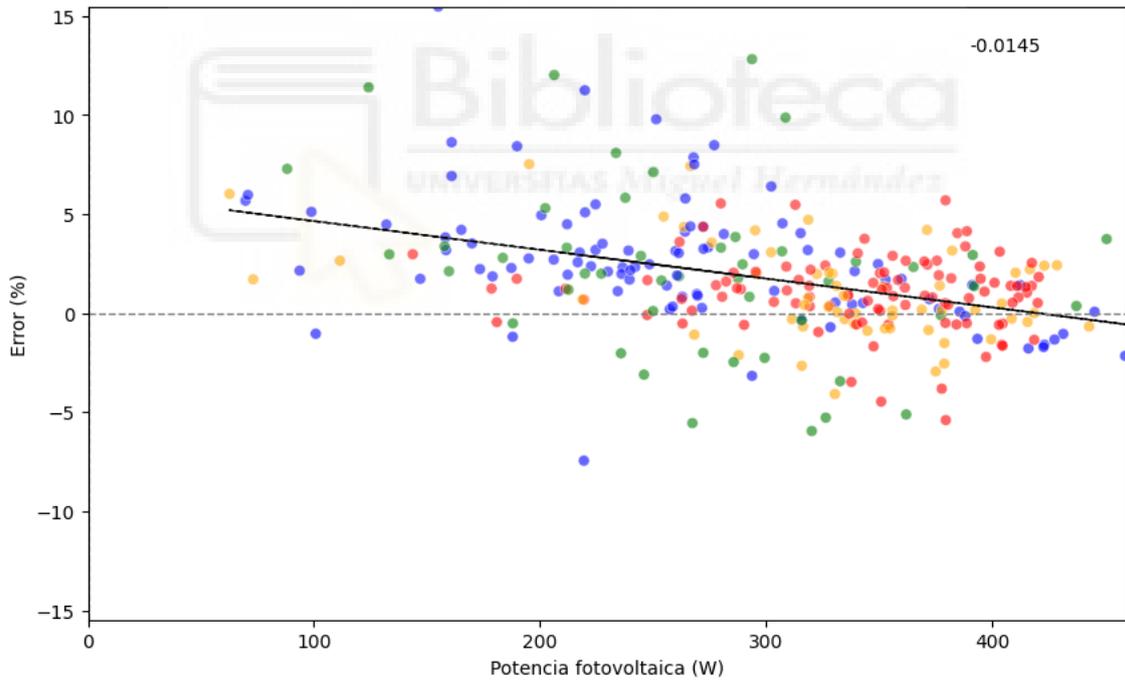
Error vs. Potencia fotovoltaica (PVGIS) - 2019

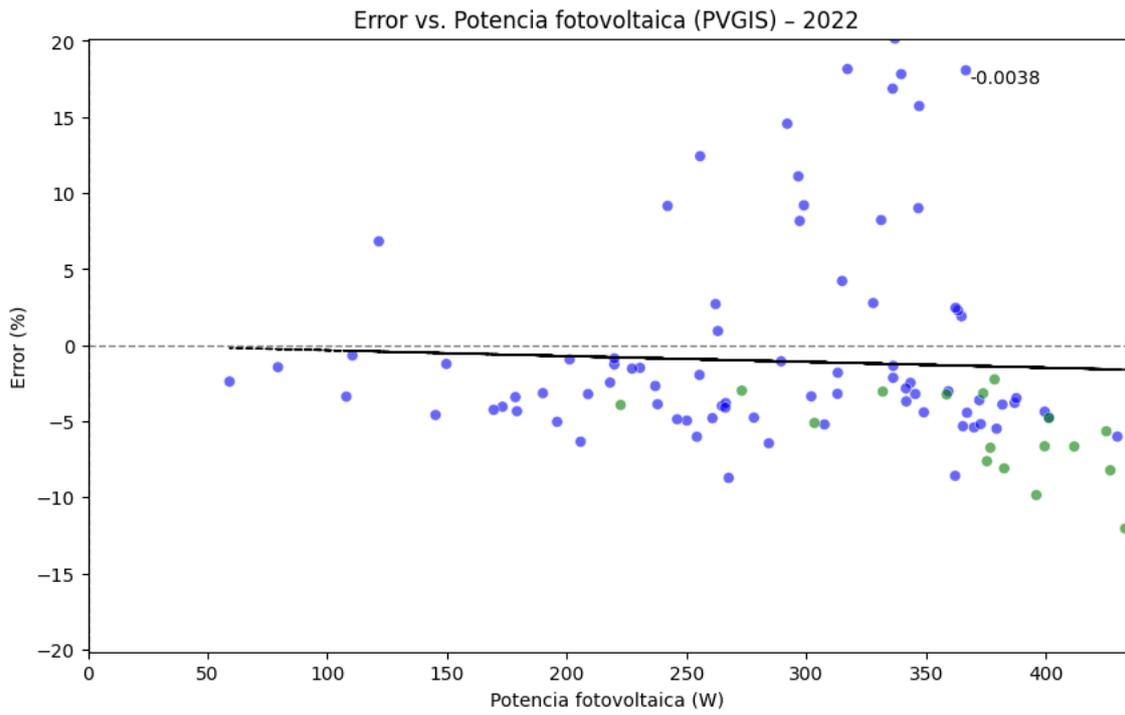


Error vs. Potencia fotovoltaica (PVGIS) - 2020

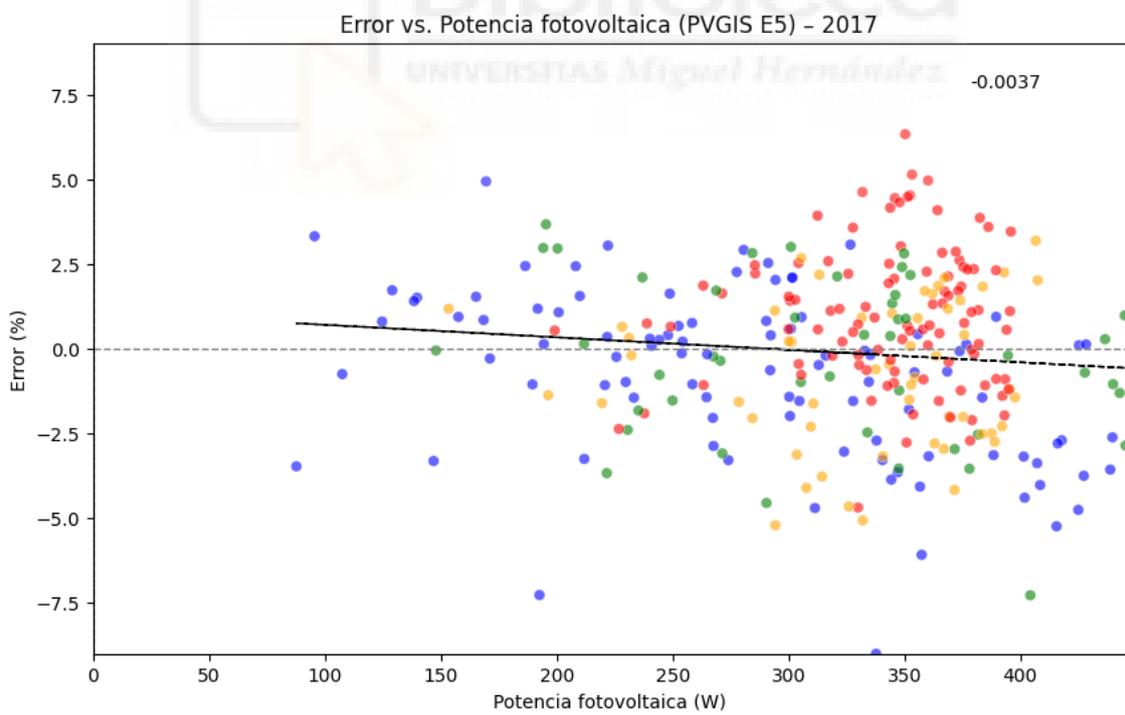


Error vs. Potencia fotovoltaica (PVGIS) - 2021

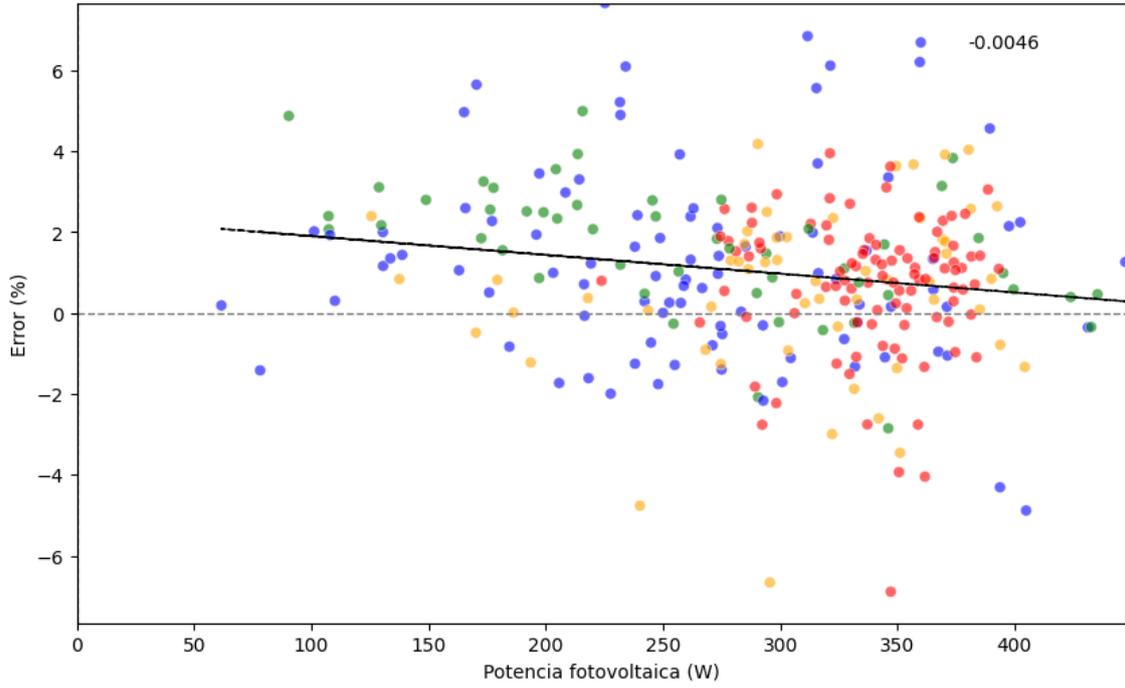




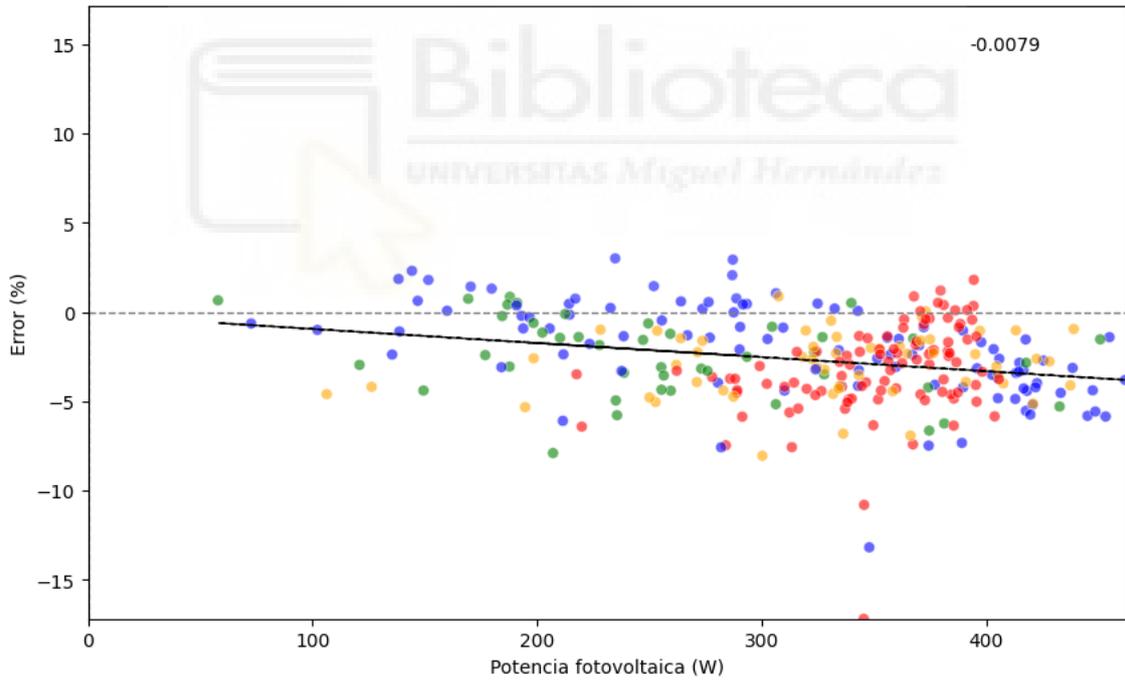
3.14.2. E5



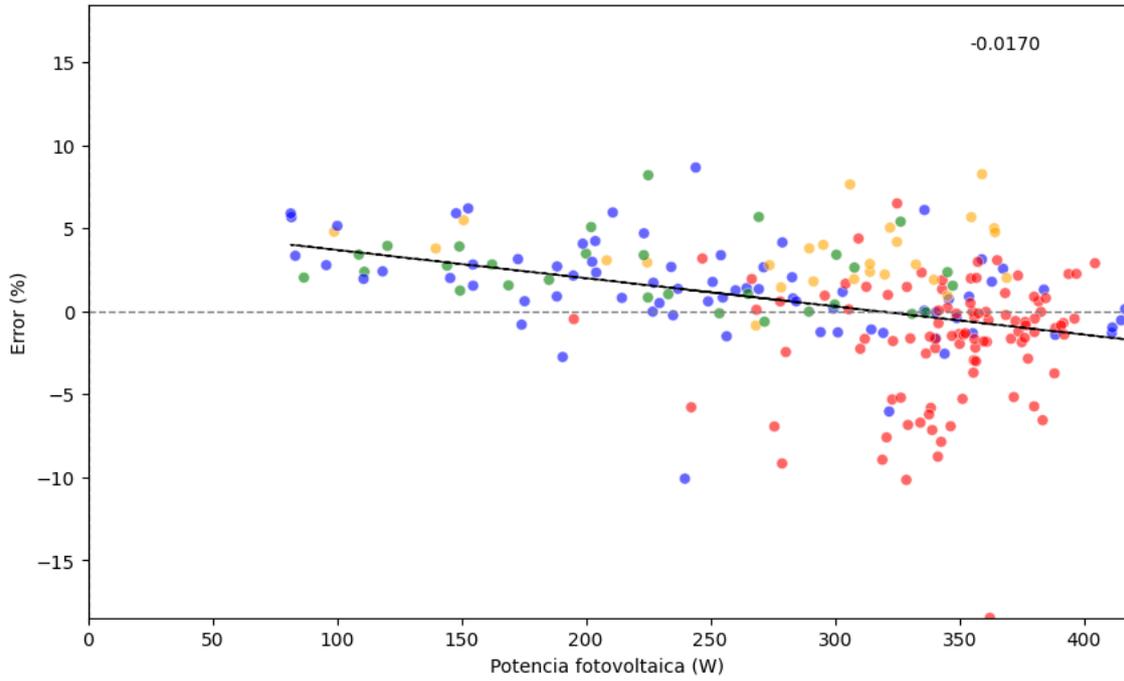
Error vs. Potencia fotovoltaica (PVGIS E5) - 2018



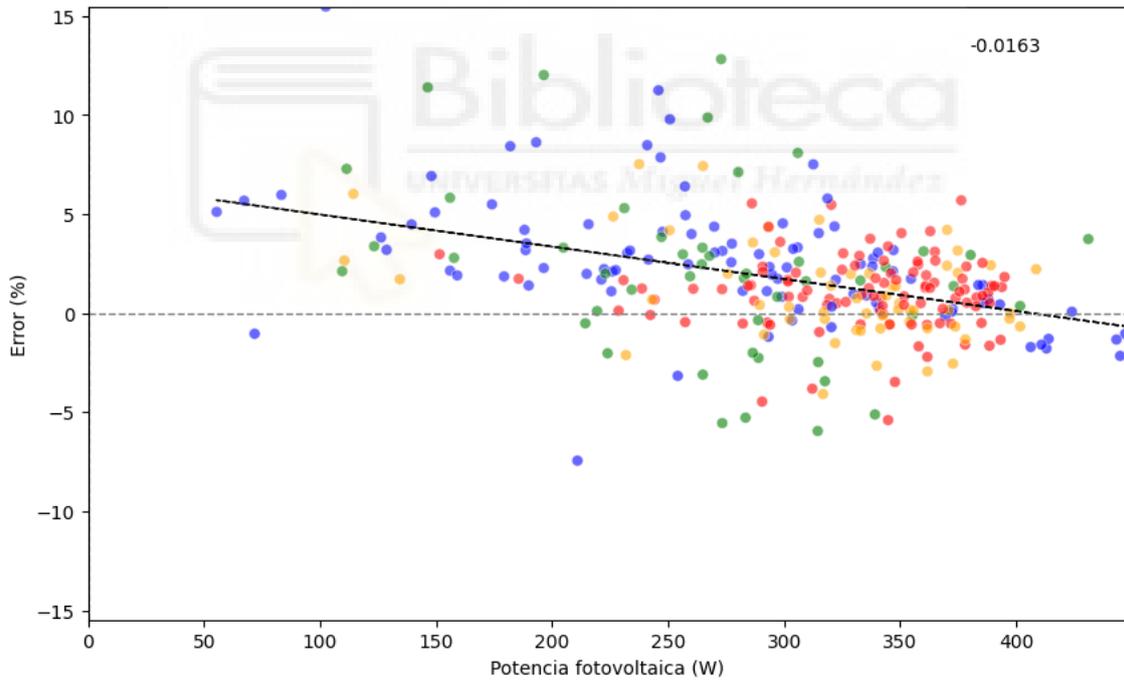
Error vs. Potencia fotovoltaica (PVGIS E5) - 2019

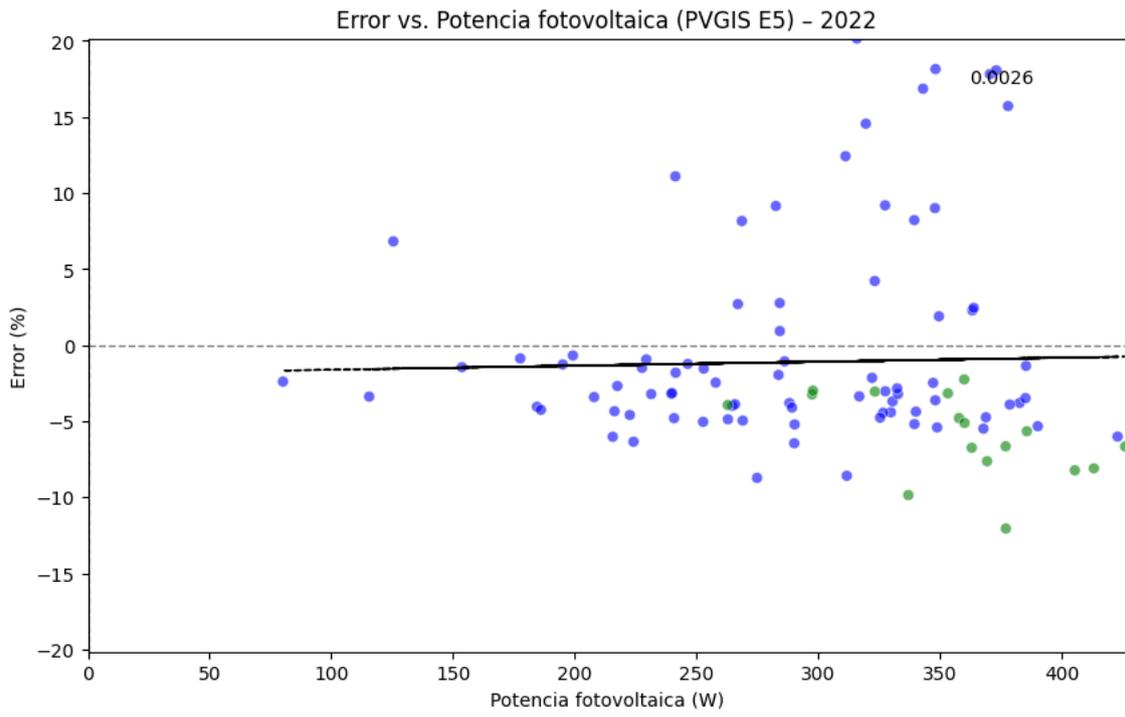


Error vs. Potencia fotovoltaica (PVGIS E5) - 2020



Error vs. Potencia fotovoltaica (PVGIS E5) - 2021





### 3.14.3. Análisis

Año	Pendiente error vs. P_pvgis
2017	-0.003
2018	-0.005
2019	-0.006
2020	-0.016
2021	-0.015
2022	-0.004

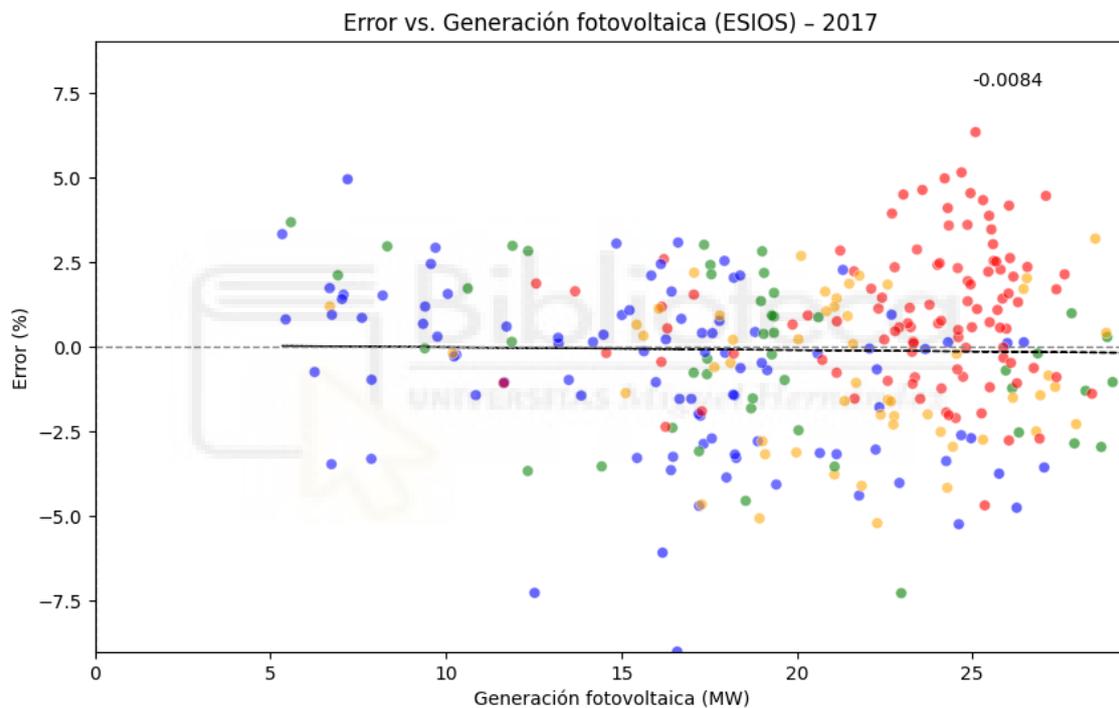
Año	Pendiente error vs. P_pvgis_e5
2017	-0.004
2018	-0.005
2019	-0.008
2020	-0.017
2021	-0.016
2022	0.003

Como se observa en las gráficas los primeros años no hay relación aparente entre el error y la magnitud de la potencia. Sin embargo, a partir de 2019 se puede apreciar una relación inversamente proporcional entre ambas magnitudes y podemos observar que una mayor potencia se asocia con una sobreestimación de la carga de energía.

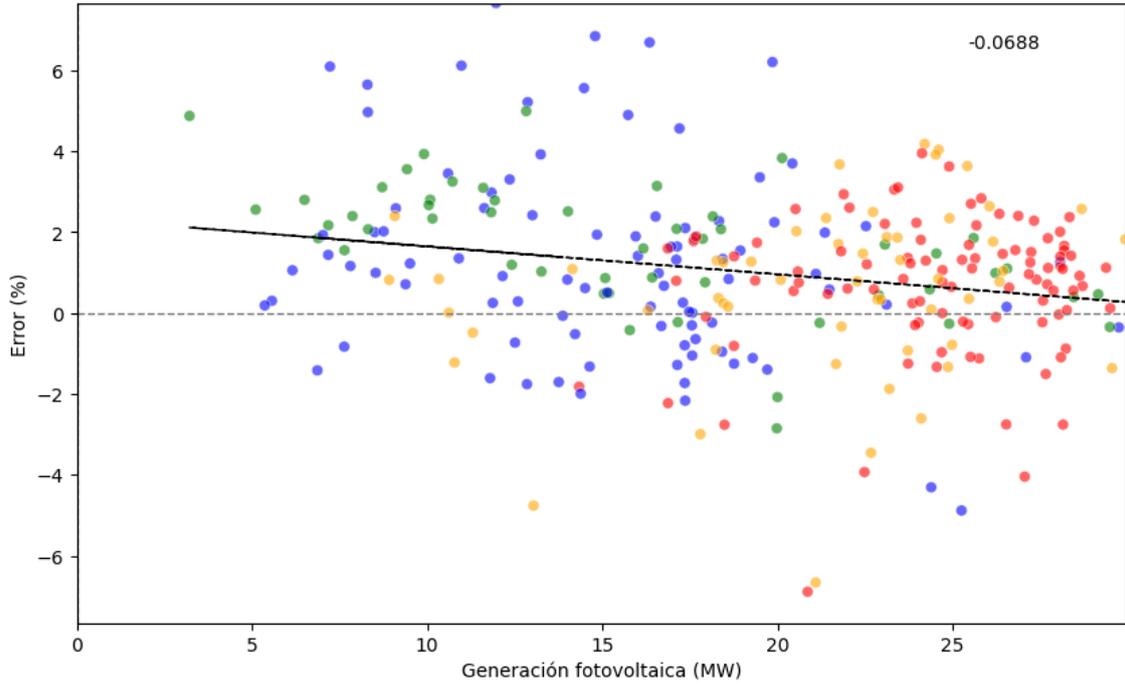
### 3.15. GENERACION\_FOTOVOLTAICA\_ESIOS

Esta variable (GENERACION\_FOTOVOLTAICA\_ESIOS) representa la generación de energía solar fotovoltaica medida en tiempo real en MW de la red ESIOS. Esta variable se refiere a la energía generada por los productores y puede servir como medida aproximada de la generada en instalaciones de autoconsumo.

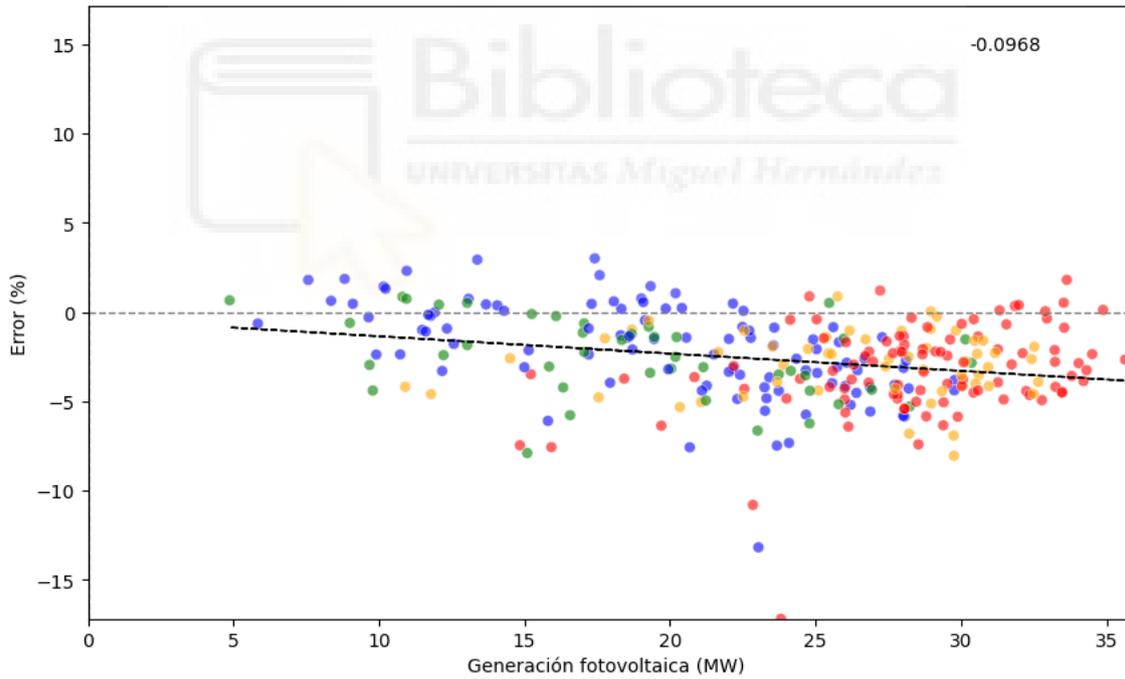
Los datos de esta variable en origen se han obtenido a nivel horario por lo que se ha seguido la misma metodología que el resto de casos de esta naturaleza: filtrado (08-20h), agregación para obtener la media diaria y comparación con el error del modelo en gráficas de dispersión.



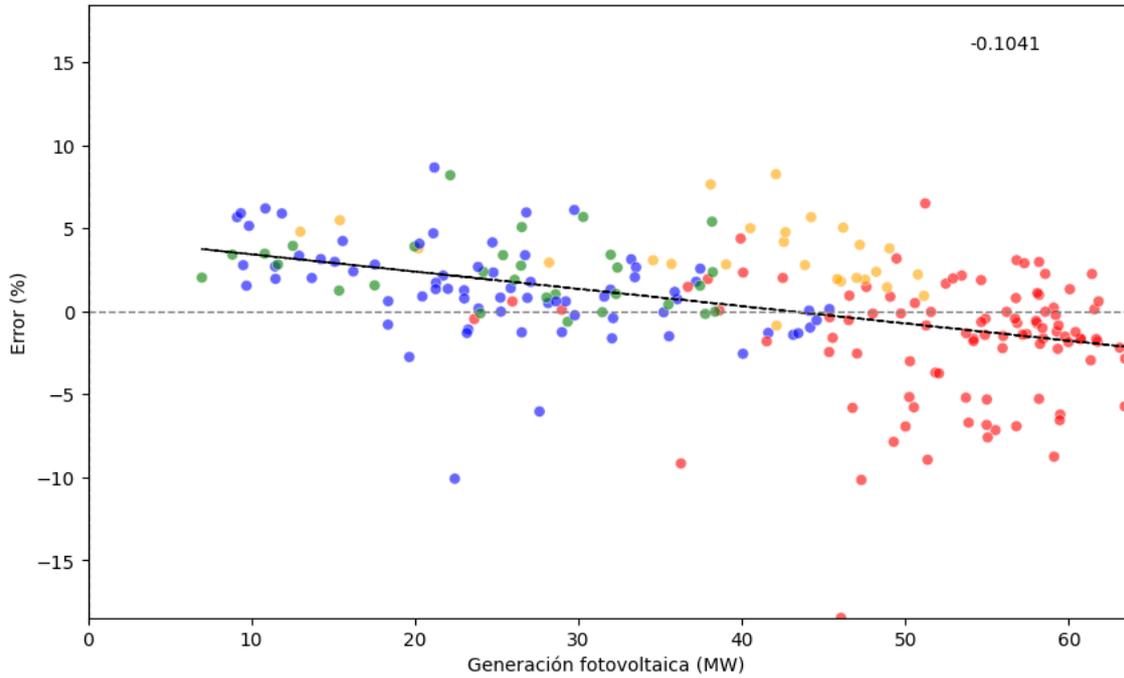
Error vs. Generación fotovoltaica (ESIOS) - 2018



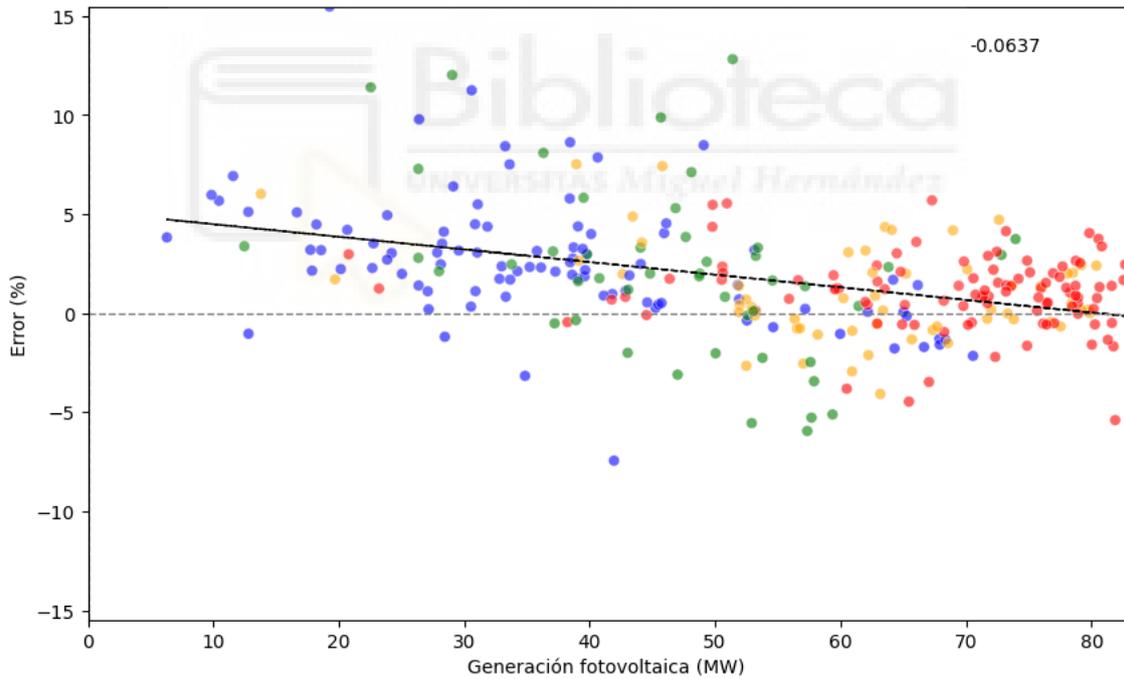
Error vs. Generación fotovoltaica (ESIOS) - 2019

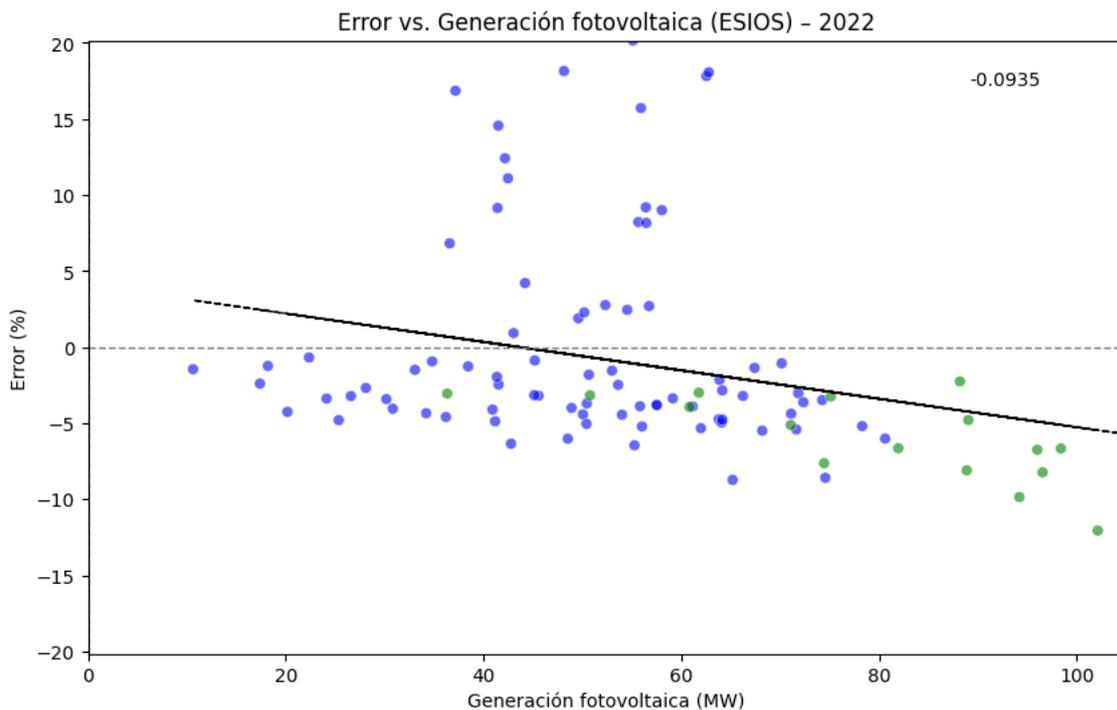


Error vs. Generación fotovoltaica (ESIOS) - 2020



Error vs. Generación fotovoltaica (ESIOS) - 2021





Año	Pendiente error vs. generacion_fotovoltaica_esios
2017	-0.008
2018	-0.069
2019	-0.097
2020	-0.104
2021	-0.064
2022	-0.093

En 2017 y 2018 no hay relación, pero a partir de 2019 se observa que la relación entre la potencia alcanzada por las instalaciones de energía fotovoltaica y el error del modelo de es inversamente proporcional y es aparente incluso en 2022 a pesar de que los datos llegan solo hasta abril, siendo precisamente este mes (puntos verdes) en el que se puede observar la relación.

#### 4. SIMULACIÓN DEL MODELO CON NUEVAS VARIABLES

Los resultados de los análisis del error del modelo en conexión con las variables meteorológicas estudiadas refuerzan la hipótesis de partida. Puede existir una relación entre estas variables y el error mediada por las instalaciones fotovoltaicas de autoconsumo.

El objetivo que se persigue al estudiar esta posible relación es la mejora del rendimiento del sistema original a través de la reducción del error. Para ello y como propuesta inicial se van a realizar nuevas simulaciones del modelo incluyendo algunas de las variables antes analizadas.

Las variables elegidas son:

- sol: horas de sol diarias
- ssrd\_wm2: radiación solar sobre el plano horizontal ( $W/m^2$ )
- total\_radiation\_pvgis: radiación solar total ( $W/m^2$ ) (PVGIS)
- Gb(i)\_pvgis: radiación directa ( $W/m^2$ ) (PVGIS)
- Gd(i)\_pvgis: radiación difusa ( $W/m^2$ ) (PVGIS)
- Gr(i)\_pvgis: radiación reflejada ( $W/m^2$ ) (PVGIS)
- P\_pvgis: potencia estimada fotovoltaica (W) (PVGIS)

Dado que los resultados para SARA3 y ERA5 eran muy similares para las variables procedentes de PVGIS, se ha optado por utilizar solo las de SARA3. Es el mismo caso de SSR\_WM2 y SSRD\_WM2, por lo que se ha escogido esta última.

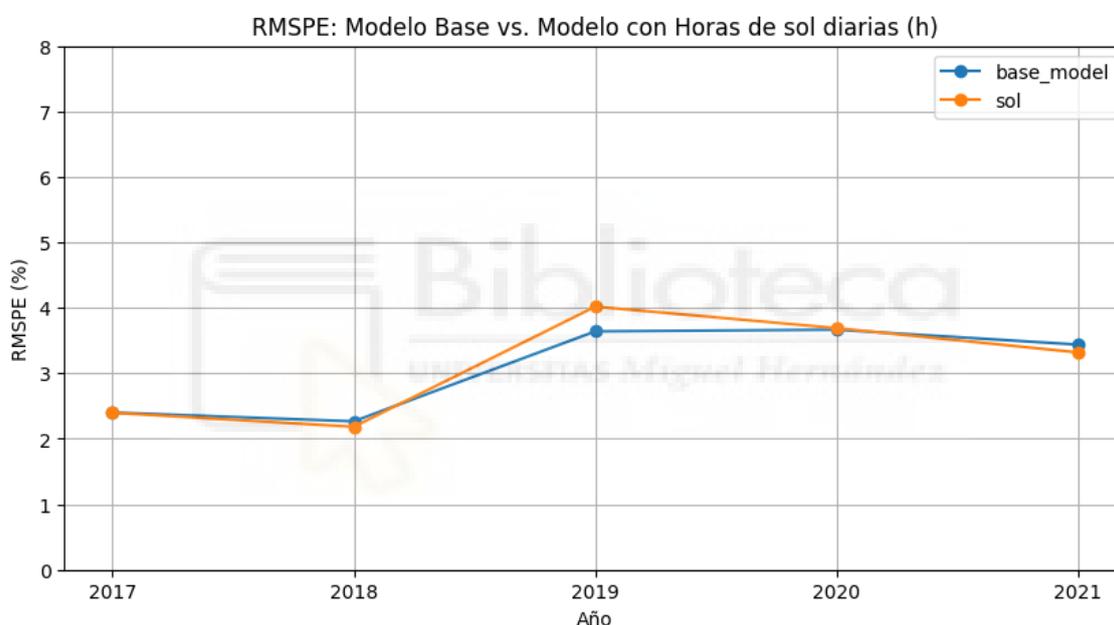
Esta nueva simulación consiste en comparar las simulaciones del modelo base del apartado 3 con la metodología descrita en el apartado 2.3 con esas mismas simulaciones pero añadiendo campos adicionales a los datos de entrada que contienen las variables a estudiar con el fin de comparar el error de ambas simulaciones respecto de la carga real del sistema.

#### 4.1.HORAS DE SOL DIARIAS

En el siguiente cuadro se muestran los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	sol	607.989	751.178	1.985	2.398	1.208
2018	sol	614.069	682.589	1.997	2.183	1.094
2019	sol	586.373	1209.716	1.911	4.019	2.103
2020	sol	607.884	1062.144	1.988	3.691	1.857
2021	sol	670.783	1070.324	2.236	3.320	1.485

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (sol):



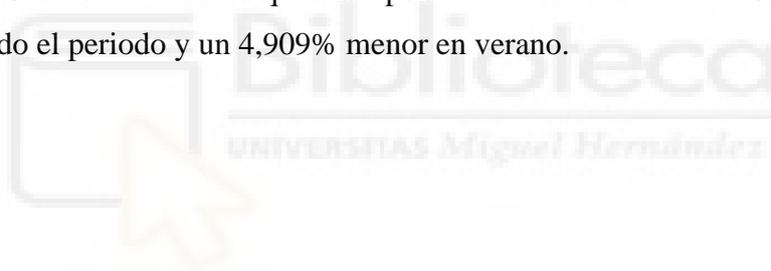
Se puede ver que, a partir de 2019, la tendencia es hacia la mejora del modelo con la variable respecto al original. Esto es lo que esperaríamos encontrar si este parámetro, relevante para la producción fotovoltaica, adquiriera mayor importancia a medida que se expanden este tipo de instalaciones.

En los siguientes gráficos se presentan los resultados del modelo comparados con el modelo base (columnas con “Base” en el nombre) para todo el periodo y para el verano respectivamente:

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	sol	2.398	2.401	-0.003	-0.124
2018	sol	2.183	2.267	-0.084	-3.707
2019	sol	4.019	3.641	0.378	10.377
2020	sol	3.691	3.667	0.024	0.664
2021	sol	3.320	3.439	-0.120	-3.475

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	sol	2.372	2.269	0.103	4.528
2018	sol	2.041	1.959	0.082	4.197
2019	sol	4.440	3.908	0.532	13.614
2020	sol	4.219	4.052	0.167	4.126
2021	sol	2.202	2.316	-0.114	-4.909

Se puede apreciar la misma tendencia que en el gráfico en los dos gráficos, mejorando progresivamente la diferencia expresada porcentualmente entre ambos modelos acabando en 2021 con el modelo que incorpora la variable con un RMSPE un 3,475% menor para todo el periodo y un 4,909% menor en verano.

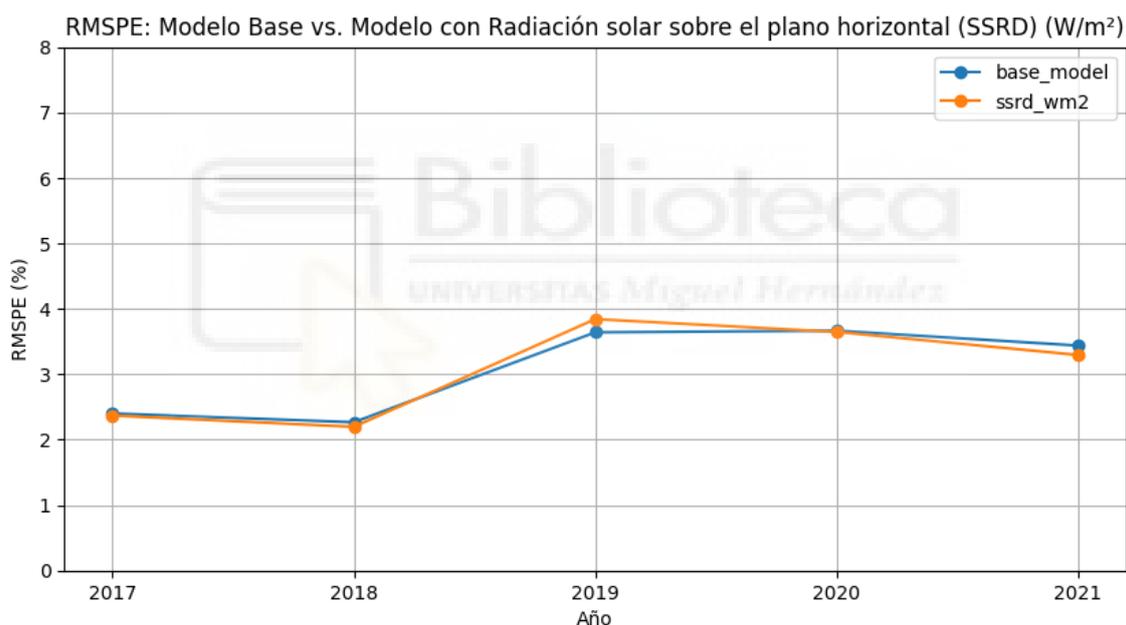


## 4.2.RADIACIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL

En el siguiente cuadro se muestran ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	ssrd_wm2	601.070	741.334	1.970	2.369	1.202
2018	ssrd_wm2	608.384	683.432	1.984	2.197	1.108
2019	ssrd_wm2	582.467	1156.134	1.901	3.844	2.022
2020	ssrd_wm2	604.860	1048.551	1.979	3.648	1.843
2021	ssrd_wm2	666.043	1063.540	2.223	3.295	1.483

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (ssrd\_wm2):



El comportamiento es similar al de la variable anterior, lo cual es esperable ya que ambas (y las siguientes) miden de forma directa o indirecta la energía disponible procedente del sol para su aprovechamiento por sistemas fotovoltaicos.

Se puede ver que entre 2019 y 2021 la tendencia es hacia la mejora del rendimiento de la SSRD\_WM2 respecto del modelo.

En las siguientes tablas se muestra la comparación del modelo que incorpora ssrd\_wm2 con el modelo base (columnas con “Base” en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	ssrd_wm2	2.369	2.401	-0.032	-1.351
2018	ssrd_wm2	2.197	2.267	-0.070	-3.108
2019	ssrd_wm2	3.844	3.641	0.202	5.558
2020	ssrd_wm2	3.648	3.667	-0.019	-0.528
2021	ssrd_wm2	3.295	3.439	-0.144	-4.186

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	ssrd_wm2	2.380	2.269	0.111	4.875
2018	ssrd_wm2	2.037	1.959	0.078	3.971
2019	ssrd_wm2	4.259	3.908	0.351	8.980
2020	ssrd_wm2	4.217	4.052	0.165	4.078
2021	ssrd_wm2	2.166	2.316	-0.150	-6.457

En la simulación del año 2021 el modelo que incorpora esta variable reduce el error un 4,186% para todo el periodo y un 6,457% en verano.

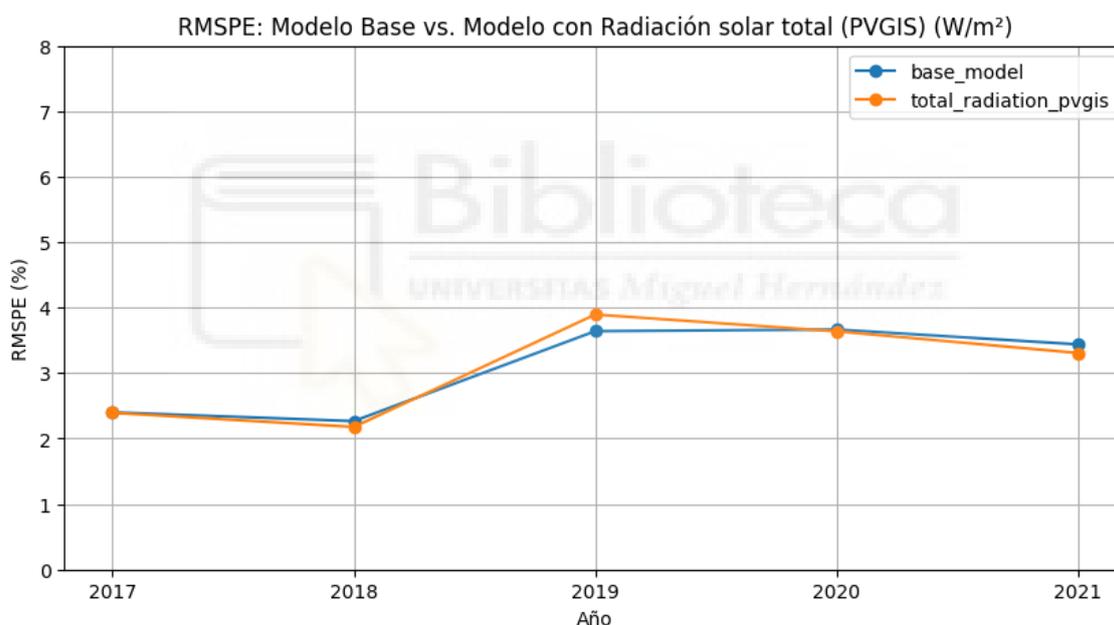


### 4.3.RADIACIÓN TOTAL

En el siguiente cuadro podemos ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	total_radiation_pvgis	604.547	748.326	1.975	2.394	1.212
2018	total_radiation_pvgis	613.996	680.506	1.998	2.178	1.090
2019	total_radiation_pvgis	585.508	1172.313	1.907	3.896	2.043
2020	total_radiation_pvgis	605.033	1046.544	1.979	3.639	1.838
2021	total_radiation_pvgis	661.083	1065.985	2.208	3.307	1.498

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (total\_radiation\_pvgis):



Se repite el patrón de las variables anteriores con una tendencia a la mejora del modelo que incorpora la variable respecto del modelo base entre los años 2019 y 2021.

En las siguientes tablas se muestra la comparación del modelo que incorpora total\_radiation\_pvgis con el modelo base (columnas con “Base” en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	total_radiation_pvgis	2.394	2.401	-0.007	-0.292
2018	total_radiation_pvgis	2.178	2.267	-0.090	-3.960
2019	total_radiation_pvgis	3.896	3.641	0.254	6.983
2020	total_radiation_pvgis	3.639	3.667	-0.028	-0.767
2021	total_radiation_pvgis	3.307	3.439	-0.132	-3.832

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	total_radiation_pvgis	2.371	2.269	0.102	4.512
2018	total_radiation_pvgis	2.025	1.959	0.066	3.349
2019	total_radiation_pvgis	4.313	3.908	0.404	10.348
2020	total_radiation_pvgis	4.126	4.052	0.075	1.847
2021	total_radiation_pvgis	2.212	2.316	-0.104	-4.475

El modelo que incorpora esta variable reduce el error en 2021 un 3,832% en para todo el periodo y un 4,475% en verano.

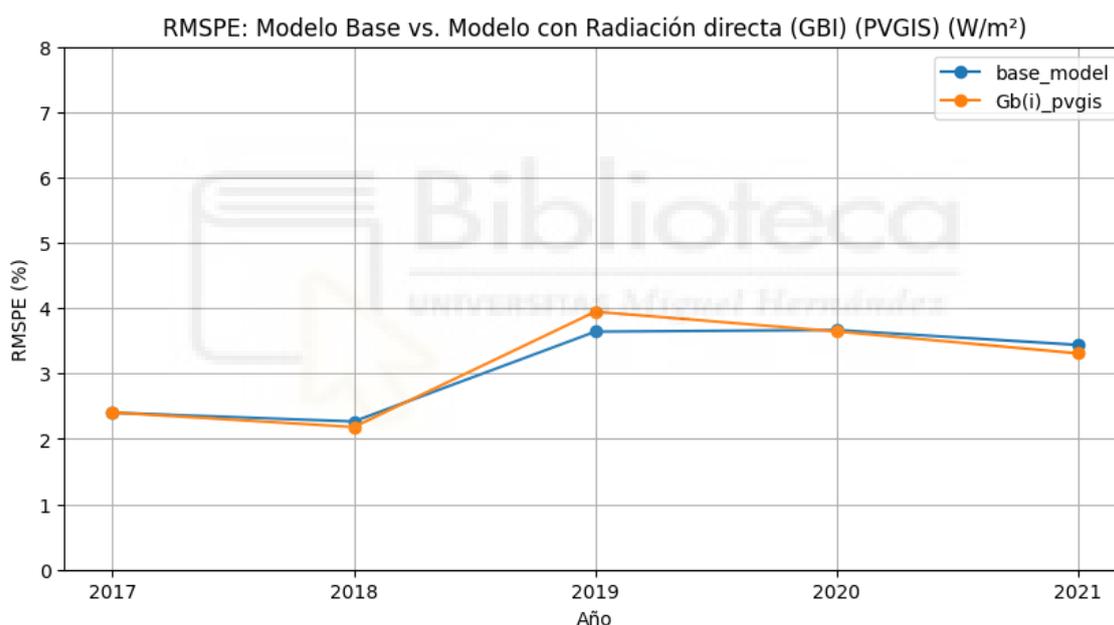


#### 4.4.RADIACIÓN DIRECTA

En el siguiente cuadro podemos ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	Gb(i)_pvgis	608.696	752.987	1.987	2.405	1.211
2018	Gb(i)_pvgis	616.895	681.807	2.005	2.182	1.088
2019	Gb(i)_pvgis	588.522	1187.806	1.915	3.947	2.061
2020	Gb(i)_pvgis	608.502	1049.109	1.989	3.645	1.833
2021	Gb(i)_pvgis	665.358	1066.428	2.221	3.309	1.490

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (Gb(i)\_pvgis):



Se repite el patrón de las variables anteriores con una tendencia a la mejora del modelo que incorpora la variable respecto del modelo base entre los años 2019 y 2021.

En las siguientes tablas se muestra la comparación del modelo que incorpora Gb(i)\_pvgis con el modelo base (columnas con “Base” en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	Gb(i)_pvgis	2.405	2.401	0.004	0.164
2018	Gb(i)_pvgis	2.182	2.267	-0.085	-3.762
2019	Gb(i)_pvgis	3.947	3.641	0.306	8.390
2020	Gb(i)_pvgis	3.645	3.667	-0.022	-0.592
2021	Gb(i)_pvgis	3.309	3.439	-0.130	-3.785

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	Gb(i)_pvgis	2.383	2.269	0.114	5.012
2018	Gb(i)_pvgis	2.028	1.959	0.069	3.529
2019	Gb(i)_pvgis	4.365	3.908	0.457	11.681
2020	Gb(i)_pvgis	4.150	4.052	0.098	2.430
2021	Gb(i)_pvgis	2.206	2.316	-0.110	-4.753

En la simulación del año 2021 el modelo que incorpora esta variable reduce el error un 3,785% para todo el periodo y un 4,753% en verano.

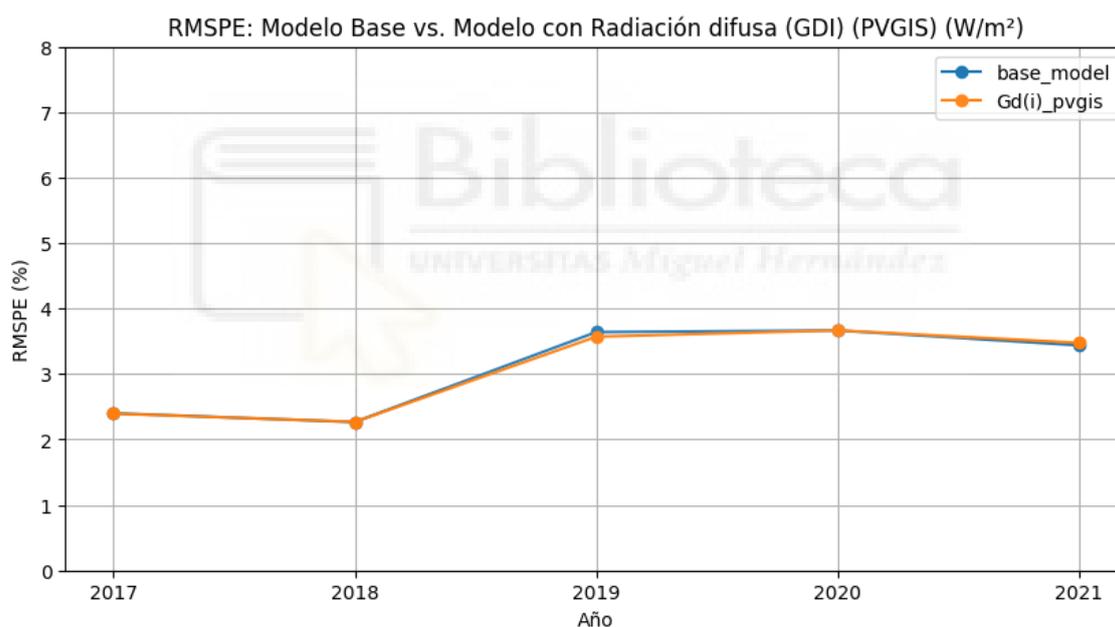


#### 4.5.RADIACIÓN DIFUSA

En el siguiente cuadro podemos ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	Gd(i)_pvgis	635.336	759.419	2.076	2.397	1.154
2018	Gd(i)_pvgis	637.391	712.580	2.065	2.269	1.099
2019	Gd(i)_pvgis	609.003	1078.835	1.969	3.571	1.813
2020	Gd(i)_pvgis	625.442	1064.981	2.024	3.668	1.812
2021	Gd(i)_pvgis	682.472	1105.855	2.255	3.477	1.542

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (Gd(i)\_pvgis):



Es aparente que la incorporación de esta variable no mejora el comportamiento del modelo base.

En las siguientes tablas se muestra la comparación del modelo que incorpora Gd(i)\_pvgis con el modelo base (columnas con "Base" en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	Gd(i)_pvgis	2.397	2.401	-0.004	-0.186
2018	Gd(i)_pvgis	2.269	2.267	0.001	0.063
2019	Gd(i)_pvgis	3.571	3.641	-0.070	-1.935
2020	Gd(i)_pvgis	3.668	3.667	0.001	0.031
2021	Gd(i)_pvgis	3.477	3.439	0.038	1.091

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	Gd(i)_pvgis	2.257	2.269	-0.012	-0.538
2018	Gd(i)_pvgis	1.956	1.959	-0.003	-0.175
2019	Gd(i)_pvgis	3.822	3.908	-0.086	-2.211
2020	Gd(i)_pvgis	4.009	4.052	-0.042	-1.042
2021	Gd(i)_pvgis	2.377	2.316	0.061	2.650

La comparación del desempeño de ambos modelos no comparte la misma tendencia que la del resto de variables. No hay una tendencia de mejora visible entre 2019 y 2021.

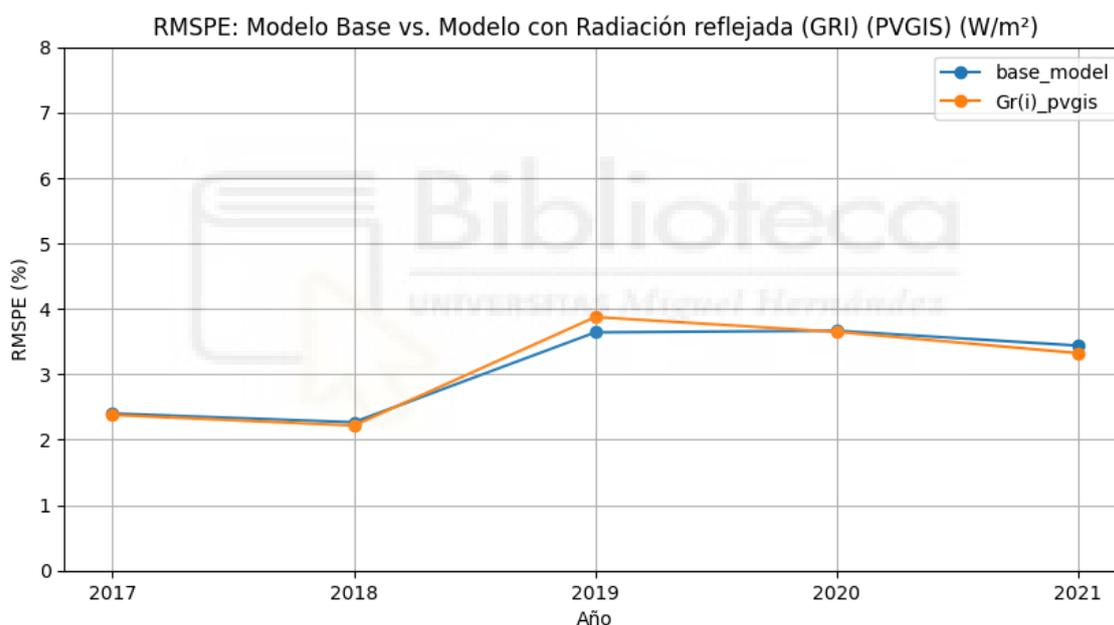


#### 4.6.RADIACIÓN REFLEJADA

En el siguiente cuadro podemos ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	Gr(i)_pvgis	609.857	744.326	1.995	2.380	1.193
2018	Gr(i)_pvgis	615.155	693.758	2.004	2.217	1.106
2019	Gr(i)_pvgis	588.218	1166.528	1.917	3.876	2.021
2020	Gr(i)_pvgis	610.008	1052.389	1.995	3.650	1.830
2021	Gr(i)_pvgis	670.327	1070.670	2.232	3.324	1.489

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (Gr(i)\_pvgis):



Se repite el patrón de las variables anteriores con una tendencia a la mejora del modelo que incorpora la variable respecto del modelo base entre los años 2019 y 2021.

En las siguientes tablas se muestra la comparación del modelo que incorpora Gr(i)\_pvgis con el modelo base (columnas con “Base” en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	Gr(i)_pvgis	2.380	2.401	-0.021	-0.883
2018	Gr(i)_pvgis	2.217	2.267	-0.051	-2.244
2019	Gr(i)_pvgis	3.876	3.641	0.234	6.438
2020	Gr(i)_pvgis	3.650	3.667	-0.017	-0.458
2021	Gr(i)_pvgis	3.324	3.439	-0.116	-3.359

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	Gr(i)_pvgis	2.378	2.269	0.109	4.805
2018	Gr(i)_pvgis	2.027	1.959	0.067	3.445
2019	Gr(i)_pvgis	4.277	3.908	0.368	9.427
2020	Gr(i)_pvgis	4.166	4.052	0.115	2.833
2021	Gr(i)_pvgis	2.206	2.316	-0.110	-4.735

En la simulación del año 2021 el modelo que incorpora esta variable reduce el error un 3,359% para todo el periodo y un 4,735% en verano.

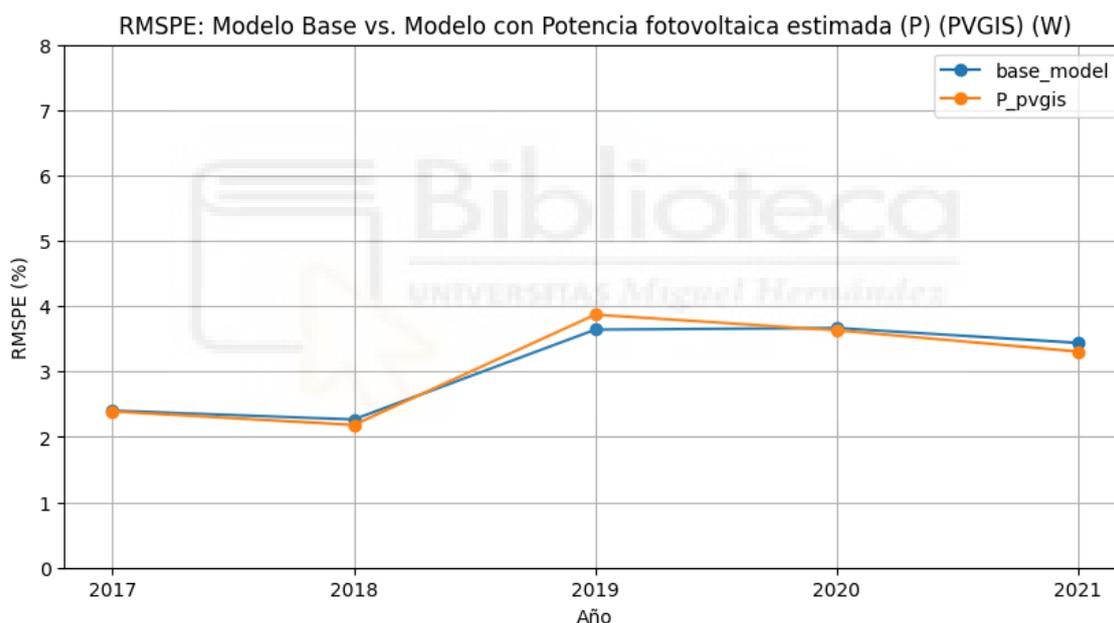


#### 4.7.POTENCIA FOTOVOLTAICA ESTIMADA

En el siguiente cuadro podemos ver los resultados de la simulación:

Año	Modelo	Train RMSE (MW)	Test RMSE (MW)	Train RMSPE (%)	Test RMSPE (%)	Test/Train RMSPE
2017	P_pvgis	604.586	747.156	1.975	2.390	1.210
2018	P_pvgis	613.963	682.661	1.998	2.183	1.093
2019	P_pvgis	585.852	1165.322	1.908	3.871	2.029
2020	P_pvgis	604.590	1044.224	1.978	3.632	1.837
2021	P_pvgis	660.065	1065.620	2.205	3.306	1.500

Y en el siguiente gráfico se puede ver el comportamiento tanto del modelo original (base\_model) como el que incorpora la variable (P\_pvgis):



Se repite el patrón de las variables anteriores con una tendencia a la mejora del modelo que incorpora la variable respecto del modelo base entre los años 2019 y 2021.

En las siguientes tablas se muestra la comparación del modelo que incorpora P\_pvgis con el modelo base (columnas con “Base” en el nombre):

Año	Modelo	Test RMSPE (%)	Base Test RMSPE (%)	Test RMSPE Dif	Test RMSPE Dif (%)
2017	P_pvgis	2.390	2.401	-0.011	-0.461
2018	P_pvgis	2.183	2.267	-0.084	-3.701
2019	P_pvgis	3.871	3.641	0.230	6.311
2020	P_pvgis	3.632	3.667	-0.035	-0.942
2021	P_pvgis	3.306	3.439	-0.133	-3.875

Año	Modelo	Test RMSPE (verano) (%)	Base Test RMSPE (verano) (%)	Test RMSPE verano Dif	Test RMSPE verano Dif (%)
2017	P_pvgis	2.358	2.269	0.089	3.908
2018	P_pvgis	2.017	1.959	0.057	2.928
2019	P_pvgis	4.282	3.908	0.374	9.568
2020	P_pvgis	4.116	4.052	0.065	1.603
2021	P_pvgis	2.206	2.316	-0.110	-4.732

En la simulación del año 2021 el modelo que incorpora esta variable reduce el error un 3,875% para todo el periodo y un 4,732% en verano.



#### 4.8.RESULTADOS

En el siguiente cuadro se muestra un resumen de los resultados de las simulaciones en relación con el modelo base:

Año	Modelo	Test RMSPE Dif (%)	Test RMSPE verano Dif (%)
2021	ssrd_wm2	-4.186	-6.457
2021	P_pvgis	-3.875	-4.732
2021	total_radiation_pvgis	-3.832	-4.475
2021	Gb(i)_pvgis	-3.785	-4.753
2021	sol	-3.475	-4.909
2021	Gr(i)_pvgis	-3.359	-4.735
2021	Gd(i)_pvgis	1.091	2.650

Las variables estudiadas mejoran su precisión respecto del modelo con la excepción de la radiación reflejada (Gd(i)\_pvgis) entre un 3,3% y un 4,2%.

Los resultados de las simulaciones siguen en general en línea con la hipótesis de partida y refuerzan su validez: las variables relacionadas con la disponibilidad de energía solar afectan al rendimiento del modelo en las simulaciones por lo que su inclusión en el modelo mejoraría la precisión del mismo.

## **5. SIMULACIÓN DEL MODELO CON NUEVAS VARIABLES A NIVEL HORARIO**

Tanto el análisis del error en relación con las variables meteorológicas como las simulaciones del modelo con nuevas variables avalan la validez de la hipótesis inicial.

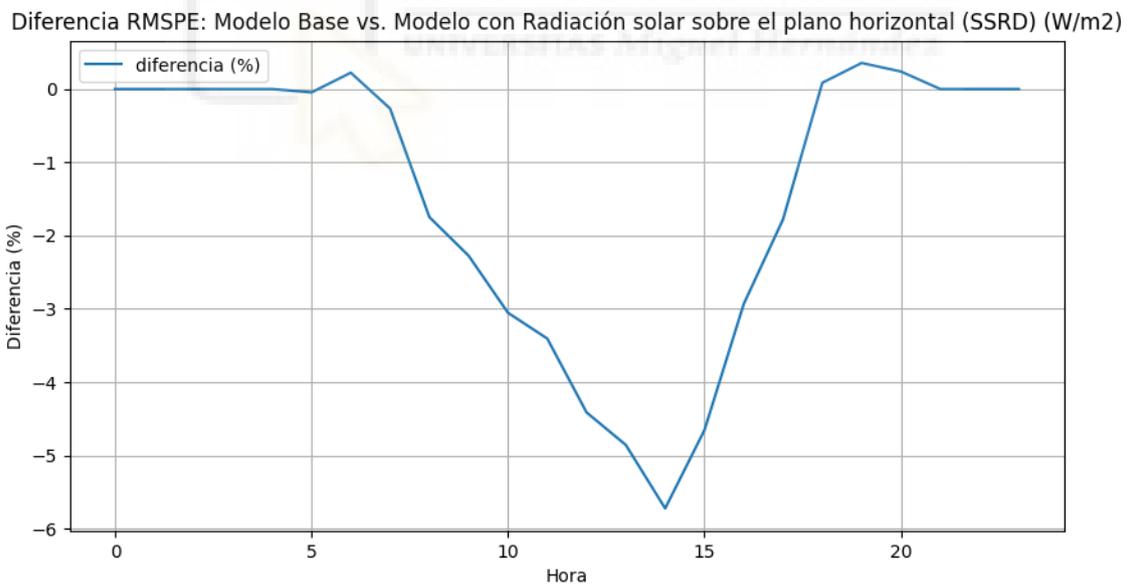
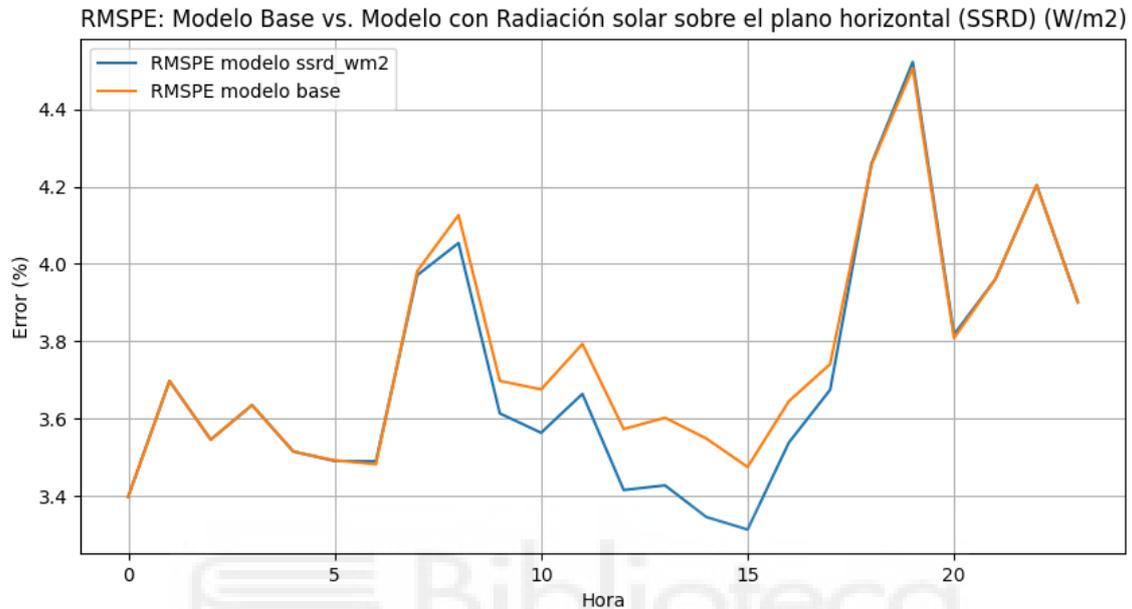
Para continuar con el análisis se cree conveniente realizar simulaciones con datos a nivel horario teniendo en cuenta los resultados del apartado 4. Para ello se van a utilizar las tres variables que mejoran más la precisión del modelo para el año 2021:

- ssrd\_wm2: radiación solar sobre el plano horizontal ( $W/m^2$ )
- total\_radiation\_pvgis: radiación solar total ( $W/m^2$ ) (PVGIS)
- P\_pvgis: potencia estimada fotovoltaica (W) (PVGIS)

La metodología seguida es similar a la seguida en análisis y simulaciones anteriores. Los datos del modelo junto con la variable a estudiar se agregan a nivel horario y por cada hora del día se lanza una regresión lineal para obtener la predicción. Posteriormente se lanza una simulación adicional con datos de entrada que no contienen dicha variable y la salida de ambos se compara con la carga real del sistema para calcular el error.

## 5.1. RADIACIÓN SOLAR SOBRE EL PLANO HORIZONTAL

A continuación se muestra la comparación entre error cuadrático medio porcentual entre la simulación del modelo horario que incluye la radiación solar sobre el plano horizontal y el modelo base.

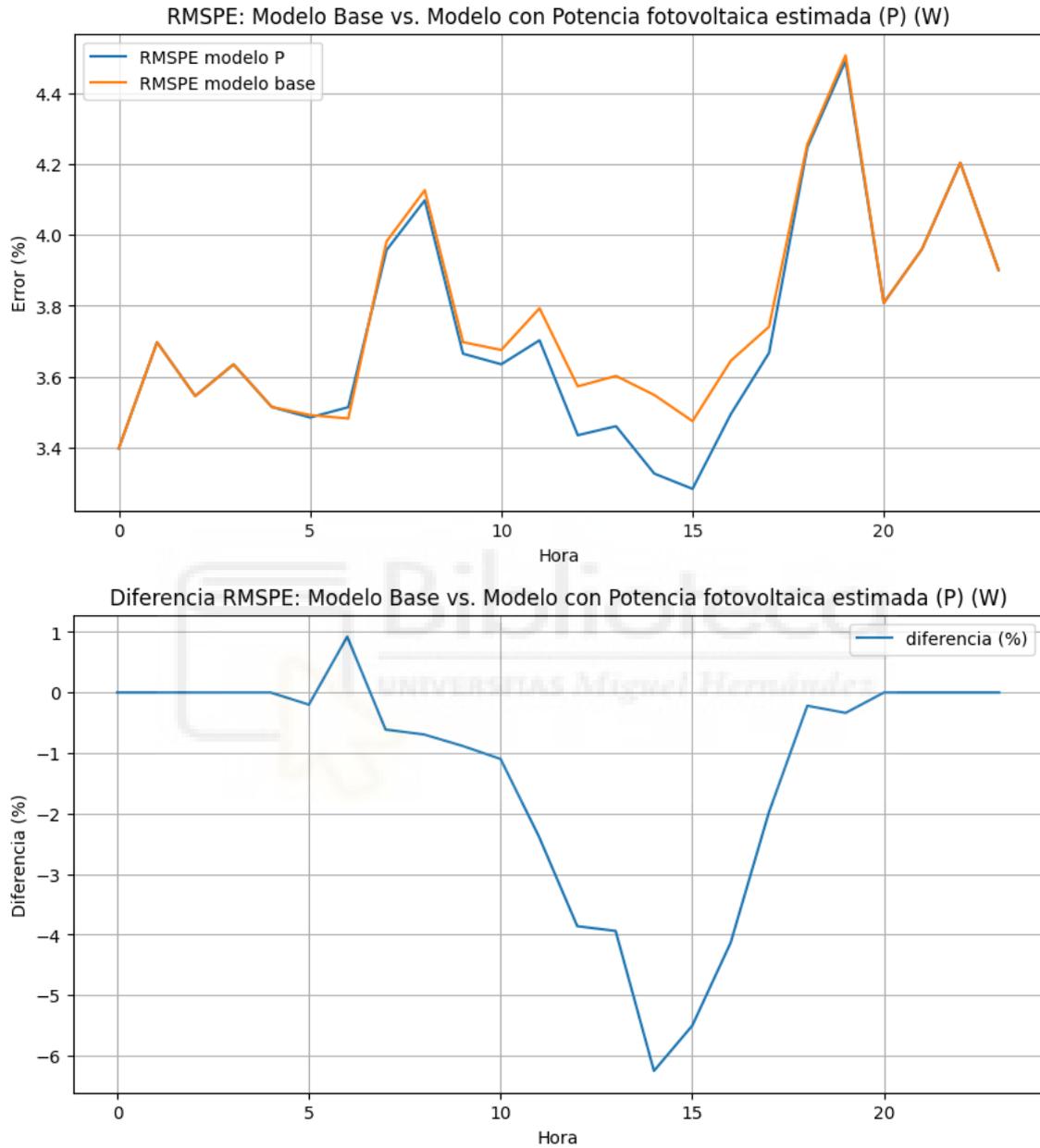


Se puede apreciar con claridad que se reduce el error a medida que progresa el día llegando al máximo de la reducción en las horas centrales del día. En concreto, el máximo de reducción es a las 14h alcanzando el 5,7%.

hora	RMSPE modelo ssrd_wm2	RMSPE modelo base	diferencia (%)
0	3.398	3.398	-0.000
1	3.697	3.697	-0.000
2	3.545	3.545	0.000
3	3.635	3.635	-0.000
4	3.515	3.515	0.000
5	3.490	3.492	-0.046
6	3.490	3.482	0.223
7	3.971	3.981	-0.266
8	4.054	4.126	-1.748
9	3.613	3.697	-2.276
10	3.563	3.675	-3.054
11	3.664	3.793	-3.406
12	3.415	3.573	-4.413
13	3.427	3.602	-4.858
14	3.345	3.548	-5.723
15	3.313	3.475	-4.655
16	3.538	3.645	-2.934
17	3.674	3.741	-1.779
18	4.259	4.256	0.085
19	4.523	4.507	0.356
20	3.817	3.808	0.239
21	3.960	3.960	0.000
22	4.204	4.204	0.000
23	3.901	3.901	0.000

## 5.2. POTENCIA FOTOVOLTAICA ESTIMADA

Simulando el modelo horario incluyendo la potencia fotovoltaica estimada y comparando con el modelo base obtenemos resultados similares al apartado anterior.

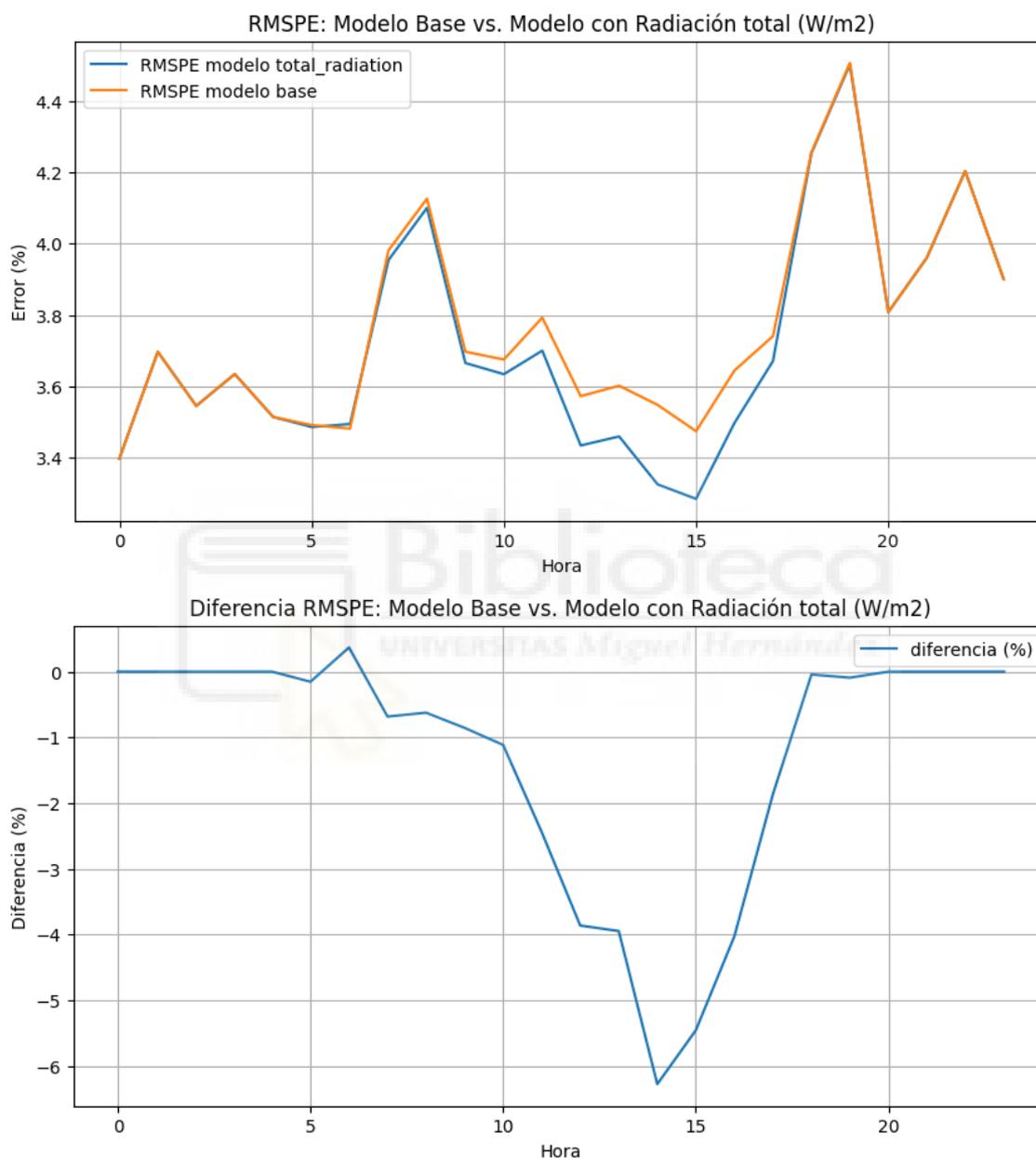


Nuevamente se reduce el error en las horas centrales del día alcanzando una reducción máxima a las 14h de un 6,2%.

hora	RMSPE modelo P	RMSPE modelo base	diferencia (%)
0	3.398	3.398	-0.000
1	3.697	3.697	-0.000
2	3.545	3.545	0.000
3	3.635	3.635	-0.000
4	3.515	3.515	0.000
5	3.485	3.492	-0.201
6	3.514	3.482	0.922
7	3.957	3.981	-0.612
8	4.098	4.126	-0.693
9	3.665	3.697	-0.881
10	3.635	3.675	-1.098
11	3.703	3.793	-2.379
12	3.435	3.573	-3.859
13	3.460	3.602	-3.938
14	3.327	3.548	-6.249
15	3.283	3.475	-5.501
16	3.494	3.645	-4.134
17	3.667	3.741	-1.975
18	4.246	4.256	-0.219
19	4.491	4.507	-0.337
20	3.808	3.808	-0.000
21	3.960	3.960	0.000
22	4.204	4.204	0.000
23	3.901	3.901	0.000

### 5.3. RADIACIÓN TOTAL

Realizando la simulación del modelo horario incluyendo la radiación total y comparando con el modelo base se obtienen los resultados que se presentan a continuación.



Como en los análisis anteriores de este apartado se reduce progresivamente el error en las horas de sol hasta alcanzar un pico de reducción a las 14h del 6,3%.

hora	RMSPE modelo total_radiation	RMSPE modelo base	diferencia (%)
0	3.398	3.398	-0.000
1	3.697	3.697	-0.000
2	3.545	3.545	0.000
3	3.635	3.635	-0.000
4	3.515	3.515	0.000
5	3.486	3.492	-0.154
6	3.495	3.482	0.367
7	3.954	3.981	-0.684
8	4.100	4.126	-0.626
9	3.666	3.697	-0.856
10	3.634	3.675	-1.116
11	3.700	3.793	-2.442
12	3.435	3.573	-3.865
13	3.460	3.602	-3.950
14	3.325	3.548	-6.280
15	3.285	3.475	-5.460
16	3.498	3.645	-4.030
17	3.671	3.741	-1.866
18	4.254	4.256	-0.045
19	4.502	4.507	-0.093
20	3.808	3.808	-0.000
21	3.960	3.960	0.000
22	4.204	4.204	0.000
23	3.901	3.901	0.000

## 5.4. RESULTADOS

En los siguientes cuadros se muestra un resumen de los resultados de la simulación horaria:

hora	RMSPE base	RMSPE P	dif base vs P (%)	RMSPE ssrd_wm2	dif base vs ssrd_wm2 (%)	RMSPE total_radiation	dif base vs total_radiation (%)
0	3.398	3.398	-0.000	3.398	-0.000	3.398	-0.000
1	3.697	3.697	-0.000	3.697	-0.000	3.697	-0.000
2	3.545	3.545	0.000	3.545	0.000	3.545	0.000
3	3.635	3.635	-0.000	3.635	-0.000	3.635	-0.000
4	3.515	3.515	0.000	3.515	0.000	3.515	0.000
5	3.492	3.485	-0.201	3.490	-0.046	3.486	-0.154
6	3.482	3.514	0.922	3.490	0.223	3.495	0.367
7	3.981	3.957	-0.612	3.971	-0.266	3.954	-0.684
8	4.126	4.098	-0.693	4.054	-1.748	4.100	-0.626
9	3.697	3.665	-0.881	3.613	-2.276	3.666	-0.856
10	3.675	3.635	-1.098	3.563	-3.054	3.634	-1.116
11	3.793	3.703	-2.379	3.664	-3.406	3.700	-2.442
12	3.573	3.435	-3.859	3.415	-4.413	3.435	-3.865
13	3.602	3.460	-3.938	3.427	-4.858	3.460	-3.950
14	3.548	3.327	-6.249	3.345	-5.723	3.325	-6.280
15	3.475	3.283	-5.501	3.313	-4.655	3.285	-5.460
16	3.645	3.494	-4.134	3.538	-2.934	3.498	-4.030
17	3.741	3.667	-1.975	3.674	-1.779	3.671	-1.866
18	4.256	4.246	-0.219	4.259	0.085	4.254	-0.045
19	4.507	4.491	-0.337	4.523	0.356	4.502	-0.093
20	3.808	3.808	-0.000	3.817	0.239	3.808	-0.000
21	3.960	3.960	0.000	3.960	0.000	3.960	0.000
22	4.204	4.204	0.000	4.204	0.000	4.204	0.000
23	3.901	3.901	0.000	3.901	0.000	3.901	0.000

Como se ha mencionado anteriormente, la reducción del error tiene un máximo en torno al 6 % en las horas centrales del día que, coinciden con el pico de generación de energía fotovoltaica.

## 6. CONCLUSIONES

El análisis de las distintas variables respecto al error de la emulación del modelo de predicción de cargas a corto plazo de Red Eléctrica Española permite concluir la importancia creciente de las variables meteorológicas relevantes en la producción de energía fotovoltaica en relación con el error. Es más, condiciones favorables a la generación de energía solar aparecen asociadas a errores negativos, lo que implica una sobreestimación de la carga del sistema.

Si estas relaciones no existieran esperaríamos una infraestimación de la energía demandada también en verano por el uso generalizado de aire acondicionado en un país como España y la distribución de puntos tendría forma de uve en variables relacionadas con la radiación.

Además, la simulación del modelo con la incorporación de nuevas variables relacionadas con la radiación solar ha permitido observar una reducción progresiva del error en el modelo entre 2019 y 2021. Estos resultados refuerzan la hipótesis de que la producción fotovoltaica tiene un efecto significativo en la precisión del modelo, sugiriendo que su integración puede mejorar las predicciones.

Como principal limitación encontramos la falta de datos para años más recientes ya que los datos solo llegan hasta abril de 2022. El número de instalaciones de autoconsumo ha continuado creciendo desde entonces y el entrenamiento de los modelos se ha realizado con periodos que no incorporaban suficiente información. Una serie más amplia de datos permitiría establecer relaciones más robustas y confirmar si la tendencia observada se mantiene en el tiempo.

La hipótesis de partida del presente trabajo es que puede existir una relación entre la proliferación de instalaciones de autoconsumo con placas solares y el error del modelo. A pesar de las limitaciones, tanto el análisis de las variables como las simulaciones del modelo incorporando algunas de las variables estudiadas apuntan a que esta relación es probable y es razonable pensar que a falta de futuros análisis que dispongan de más información el error del modelo podría reducirse incorporando variables meteorológicas relevantes para la producción de energía fotovoltaica como la radiación total o la radiación solar sobre el plano horizontal.

Futuros trabajos en esta línea pueden plantear análisis con mayor precisión aumentando la granularidad y la amplitud temporales de los datos, incorporar predicciones en tiempo

real, estudiar la relación de estas variables con la temperatura y el rendimiento de las placas fotovoltaicas o segregar geográficamente los modelos ya que la incidencia de la meteorología no es igual en todo el territorio.



## 7. BIBLIOGRAFÍA

1. AEMET (Agencia Estatal de Meteorología). (s.f.). *Agencia Estatal de Meteorología - AEMET*. Recuperado de <https://www.aemet.es>
2. ERA5 (ECMWF Reanalysis). (s.f.). *Copernicus Climate Data Store - ERA5*. Recuperado de <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels>
3. López, M., Sans, C., Valero, S., & Senabre, C. (2019). Classification of Special Days in Short-Term Load Forecasting: The Spanish Case Study. *Energies*, *12*(7), 1253. <https://doi.org/10.3390/en12071253>
4. López, M., Valero, S., Rodríguez, A., Veiras, I., & Senabre, C. (2018). New online load forecasting system for the Spanish Transport System Operator. *Electric Power Systems Research*, *154*, 401–412. <https://doi.org/10.1016/j.epsr.2017.09.003>
5. López, M., Valero, S., Sans, C., & Senabre, C. (2021). Use of Available Daylight to Improve Short-Term Load Forecasting Accuracy. *Energies*, *14*(1), 95. <https://doi.org/10.3390/en14010095>
6. PVGIS (Photovoltaic Geographical Information System). (s.f.). *Comisión Europea - PVGIS*. Recuperado de <https://ec.europa.eu/jrc/en/pvgis>
7. Red Eléctrica de España (ESIOS). (s.f.). *Sistema de información del operador del sistema - ESIOS*. Recuperado de <https://www.esios.ree.es>
8. Red Eléctrica de España. (2024). *Generación, solar fotovoltaica*. Recuperado de <https://www.sistemaelectrico-ree.es/informe-de-energias-renovables/sol/generacion/solar-fotovoltaica-solgeneracion>