

Trabajo Fin de Grado

Desarrollo de un gestor de energía de una
planta fotovoltaica con baterías de litio de
segunda vida

Development of an energy management system
for a photovoltaic plant with second life lithium
batteries

Autor

Sara Boira Lázaro

Directores

Marius Crisan
Pablo Pastor Flores

Departamento de Ingeniería de Tecnologías Industriales

Escuela de Ingeniería y Arquitectura

Universidad de Zaragoza

Junio 2023

Resumen

En el trabajo se desarrolla la gestión de los flujos de energía de un punto de consumo que cuenta con una instalación fotovoltaica y una batería de segunda vida. Para ello, se plantea un entorno de simulación compuesto por un modelo tanto de batería como de la predicción de generación fotovoltaica y los consumos. Para la predicción de la generación solar a futuro se ha diseñado y entrenado un modelo basado en redes neuronales. En cuanto a la estimación de consumos, se ha buscado una base de datos adecuada con la que trabajar. Tras su análisis, se ha ajustado un modelo estadístico SARIMAX. Ambos modelos se han validado haciendo uso de distintas métricas con las que evaluar su rendimiento. Mediante los modelos creados y la definición de condiciones de contorno para el gestor de energía, se ha entrenado un optimizador basado en un modelo “*Model Predictive Control*”, MPC. Este sistema prioriza el autoconsumo de la instalación, minimizando las pérdidas de energía o “*curtailment*”. Se analizan varios casos de estudio para validar el comportamiento en situaciones límite del sistema. Tras el análisis de resultados, se plantean distintas líneas de trabajo futuro, destacando las posibilidades que ofrece un aumento en la complejidad y número de las condiciones impuestas en el optimizador.

El código del trabajo se ha publicado y es consultable en el repositorio de Gitlab [1].

Índice

Índice de figuras.....	5
Índice de tablas.....	7
1. Introducción.....	1
1.1 Antecedentes y Contexto.....	1
1.2 Objetivos y alcance.....	2
1.3 Distribución de tareas y estructura del documento.....	3
2. Estudio del Estado del Arte	4
2.1 Modelos de Producción.....	4
2.1.1 Generadores fotovoltaicos	4
2.1.2 Geometría solar.....	7
2.1.3 Inversores.....	8
2.2.4 Baterías	8
2.2 Modelos de Predicción.....	10
2.2.1 Redes neuronales.....	12
2.2.2 Modelos ARIMA.....	14
2.2.3 Optimizador para el Sistema de Gestión de Energía (EMS).....	15
3. Metodología	16
3.1 Herramientas utilizadas	16
3.2 Modelo de producción de la instalación	17
3.2.1 Condiciones y especificaciones.....	17
3.2.2 Métricas.....	19
3.3 Modelo de consumo de la instalación.....	19
3.3.1 Condiciones y especificaciones.....	19
3.3.2 Métricas.....	21
3.4 Modelo de batería virtual	21
3.5 Sistema de gestión de energía (EMS).....	21
3.5.1 Condiciones de contorno	22
3.5.2 Función objetivo	22

4.	Resultados y discusión	24
4.1	Modelo de producción de la instalación	24
4.2	Modelo de consumo de la instalación	38
4.3	Sistema de gestión de energía (EMS)	43
5.	Conclusiones y trabajo futuro	47
6.	Bibliografía	48

Índice de figuras

Fig 1 Esquema generador fotovoltaico compuesto por dos ramas de tres módulos en serie. Fuente: Fig. 5.10, libro [7].....	5
Fig 2 Célula solar de silicio y su funcionamiento. Fuente: Fig. 3, artículo [8]	5
Fig 3 Efecto de la temperatura en la curva característica de una célula solar ($G = 800 \text{ W/m}^2$). Fuente: Fig. 4.11, libro[11].....	6
Fig 4 Efecto de radiación en panel 50 Wp EverExceed a 25°C Fuente: [12]	6
Fig 5 Esquema geometría solar y orientación del panel	7
Fig 6 Modelo eléctrico de batería Fuente: Fig. 7.2, libro [11].....	9
Fig 7 Esquema de Modelo P2D en batería de Li-ion. Fuente Fig. [3], artículo [26]	10
Fig 8 Estructura de una neurona artificial junto con su función de red. Fuente: Figs. 2a y b, [33].....	13
Fig 9 MLP con una capa oculta. Fuente: Fig 5.2, trabajo [31].....	13
Fig 10 Irradiancia durante el año 2018.....	17
Fig 11 Irradiancia solar durante los días 10/01/2018 y 10/07/2018.....	18
Fig 12 Irradiancia promedio por años del conjunto de datos	18
Fig 13 Gráficas de potencia media mensual consumida en distintos años de la base de datos de consumo.....	20
Fig 14 Media del consumo energético durante diferentes años.....	21
Fig 15 Esquema de la instalación con el criterio de signos utilizado	22
Fig 16 Criterio de parada de entrenamiento eligiendo métrica el error RMS. Fuente: Fig.4., artículo [43]	24
Fig 17 Gráfica de la función ReLU. Fuente: Fig 3.8 trabajo [45]	26
Fig 18 Predicción de la irradiancia en ventanas de 24 horas realizada por el segundo modelo	27
Fig 19 Análisis de sensibilidad del modelo 2.....	28
Fig 20 Vectores con los valores de las 24 horas predichas por las neuronas 4 y 8 del modelo 2	29
Fig 21 Análisis de sensibilidad del modelo 2 entrenado con una semilla de valor 0.....	29
Fig 22 Métricas de error de los conjuntos de entrenamiento y validación del modelo 73, correspondiente al modelo 2 entrenado con una semilla de valor 25999	30
Fig 23 Métricas de error de los conjuntos de entrenamiento y validación del modelo 92 , correspondiente al modelo 2 entrenado con una semilla de valor 900000	31

Fig 24 Métricas de error de los conjuntos de entrenamiento y validación del modelo 93, correspondiente al modelo 3 entrenado con una semilla de valor 900000	31
Fig 25 Métricas de error de los conjuntos de entrenamiento y validación del modelo 94, correspondiente al modelo 4 entrenado con una semilla de valor 900000	31
Fig 26 . Análisis de sensibilidad de las épocas 47, 48 y 49 del modelo 2 en las que se producen la bajada brusca en la gráfica de error	32
Fig 27 Análisis de sensibilidad de las distintas épocas en las que se producen bajadas bruscas en las gráficas de error de los diferentes modelos ejemplo	33
Fig 28 Predicción de la irradiancia en ventanas de 24 horas realizada por el segundo modelo entrenado con una semilla de valor 25999	33
Fig 29 Gráfica de la función ELU. Fuente: Fig 16 trabajo [44]	34
Fig 30 Fig. Gráfica de la función función de activación lineal	35
Fig 31 Gráfica de métricas de error del modelo escogido	36
Fig 32 Error MAE y RMSE horarios del modelo seleccionado	36
Fig 33 Predicciones realizadas por el modelo escogido	38
Fig 34 Gráfica de los valores máximos, mínimos y medios de consumo durante el año 2007	39
Fig 35 Gráficas resultado de la descomposición de la serie temporal mediante la función <code>seasonal_decompose()</code>	40
Fig 36 Gráfica con el componente estacional de la serie ampliado	40
Fig 37 Gráfica de autocorrelación de la serie temporal de estudio	41
Fig 38 Gráfica de autocorrelación parcial de la serie temporal de estudio	42
Fig 39 Predicción de la potencia en ventanas de 24 horas realizada por el modelo SARIMAX	43
Fig 40 Resultados obtenidos tras implementar el modelo MPC	44
Fig 41 Resultados obtenidos con el modelo MPC en un día en el que la batería satura	46

Índice de tablas

Tabla 1 Cronograma de distribución de tareas.....	3
Tabla 2 Características distintas químicas de batería.....	10
Tabla 3 Comparación de los principales métodos de predicción.....	11
Tabla 4 Distintas combinaciones de redes neuronales elegidas.....	25
Tabla 5 Información sobre la red neuronal escogida	35
Tabla 6 Métricas de los modelos con los que se ha obtenido mejores resultados junto con sus características	36
Tabla 7 Métricas de error del modelo SARIMAX implementado.....	42
Tabla 8 Características de la batería de estudio	44
Tabla 9 Energías obtenidas tras los balances realizados por el modelo MPC.....	45
Tabla 10 Características de la batería de estudio tras el redimensionamiento	45
Tabla 11 Energías obtenidas tras los balances realizados por el modelo MPC en un día en que la batería satura	46

1. Introducción

1.1 Antecedentes y Contexto

El interés por la integración de las energías renovables en el sistema eléctrico ha incrementado en los últimos años, esto se debe a factores como la creciente preocupación por la protección del medio ambiente, el agotamiento de las fuentes de energía convencionales o el aumento de los precios de estas últimas. Actualmente, el sector fotovoltaico es el que ha experimentado un mayor auge.

La utilización de energías renovables, además de reducir el impacto ambiental, permite complementar el servicio que ofrece la red eléctrica permitiendo así un mayor nivel de autosuficiencia energética. Sin embargo, muchas de estas nuevas fuentes de energía presentan una gran variabilidad al ser directamente dependientes del clima, provocando un desequilibrio entre el consumo y la producción de estas instalaciones.

Esta incertidumbre en la producción hace que sea necesario en muchas instalaciones el uso de tecnologías de almacenamiento de energía como elementos auxiliares en los momentos de menor generación. Para mejorar el uso de dicho almacenamiento, así como de la instalación completa, en términos de vida útil y amortización, resulta imprescindible realizar una buena gestión de los flujos de energía.

Es por ello que la búsqueda de métodos precisos de predicción solar, que permitan modelar tanto la demanda como la producción de las instalaciones de interés, se ha convertido en una prioridad, apareciendo distintas y prometedoras líneas de investigación dependiendo del tamaño de las ventanas temporales que se desean estudiar.

Este trabajo se enmarca dentro del proyecto “BFR50kW” de la plataforma interdisciplinar PTI+Transener del Consejo Superior de Investigaciones Científicas, CSIC. El proyecto consiste en la fabricación de una batería de flujo redox de vanadio y la integración de distintos tipos de tecnologías de almacenamiento y generación renovables. Con este trabajo se exploran varios de los conceptos transversales de la línea de investigación en el CSIC.

En este contexto, el presente trabajo se centra en diseñar e implementar modelos de predicción, en distintos horizontes temporales, tanto de producción como de consumo de la instalación. De esta manera, se podrá anticipar el déficit de energía solicitado, que se deberá suplir con la ayuda de una batería, o el exceso de potencia generada, que será vertida a la red. Respecto a la utilización de la mencionada batería complementaria, se realizará también un modelado de la misma que permitirá simular distintos escenarios de carga y descarga y conocer cómo se comportará la plataforma ante ellos.

Por último, para integrar los modelos mencionados y conseguir el máximo aprovechamiento de los recursos disponible se establecerán una serie de reglas y se creará un modelo de optimización que permita utilizar eficazmente la batería según nuestros intereses, gracias al conocimiento previo de la energía disponible en ésta, y las potencias demandadas y generadas por la instalación completa.

Para este propósito se realizará una revisión de las técnicas más utilizadas actualmente para la resolución de este tipo de problemas, destacando los modelos de Inteligencia Artificial (AI), como son las redes neuronales (ANNs), cuyo uso es cada vez más frecuente ya que permiten trabajar con gran cantidad de datos de entrada, aprendiendo patrones complejos y adaptándose automáticamente a los cambios de las condiciones meteorológicas. A su vez, también se hará hincapié en los modelos conocidos como “*AutoRegressive Integrated Moving Average*” (ARIMA) y los “*Model Predictive Control*” (MPC) dadas las ventajas que suponen y que serán expuestas más adelante.

1.2 Objetivos y alcance

Este trabajo busca obtener una visión general sobre los diferentes modelos de producción existentes en el ámbito de la generación de energía fotovoltaica, así como de las metodologías de predicción más utilizadas con el objetivo de conseguir implementar nuestros propios modelos matemáticos para aplicarlos en nuestra instalación de estudio.

De forma concreta los objetivos perseguidos son los siguientes:

- Investigación bibliográfica de los distintos sistemas de generación de energía fotovoltaica existentes, profundizando en sus componentes principales, así como una pequeña revisión acerca de los elementos de almacenamiento y de las métricas más utilizadas. Por otro lado, se estudiarán las técnicas de modelado más utilizadas en la predicción energética en casos similares al nuestro.
- Búsqueda y adecuación de bases de datos reales de consumo y producción en instalaciones fotovoltaicas.
- Diseño e implementación de los modelos predictivos de generación y consumo de energía solar a partir del entrenamiento con las bases de datos seleccionadas.
- Creación de un modelo que permita simular el comportamiento de una batería.
- Selección de métricas con las que validar los modelos mediante la comparación de las predicciones obtenidas con los datos reales disponibles.
- Establecer unas condiciones de contorno que nos permitan tener un orden de prioridad en la creación de un sistema de gestión de la energía para la optimización de los flujos de energía en la instalación fotovoltaica.
- Análisis de los resultados obtenidos y obtención de conclusiones.

1.3 Distribución de tareas y estructura del documento

Este documento se presenta de manera organizada, exponiendo en primer lugar los fundamentos básicos y los conceptos que permiten entender el contexto en el que se ha trabajado, y a continuación las herramientas y técnicas utilizadas.

De esta forma el trabajo ha sido dividido en las siguientes secciones:

- Estudio del estado del arte: apartado que comienza con la revisión de las instalaciones fotovoltaicas y sus principales componentes como sistemas de producción, y continuando con los diferentes modelos de predicción más habituales.
- Metodología: se presentan las herramientas de las que se ha hecho uso a lo largo del proyecto y a continuación, se explican detalladamente los modelos diseñados junto con las condiciones y especificaciones establecidas para ello y las métricas escogidas para evaluar su eficacia.
- Resultados y discusión: se describe el proceso de creación de los modelos y sistema de gestión de energía, exponiendo los resultados más relevantes, realizando comparativas entre las predicciones y los valores reales y realizando un comentario acerca de ello.
- Conclusión: a partir de los resultados mostrados en el apartado anterior y todo lo presentado se hace una valoración final del proyecto.

Para concluir este apartado, en el siguiente cronograma se plasma la distribución de tareas seguida:

Tabla 1 Cronograma de distribución de tareas

Tareas	Sept.	Oct.	Nov.	Dic.	Ene.	Feb.	Mar.	Abr.
Revisión bibliográfica								
Modelo de producción								
Modelo de consumo								
Modelo de batería								
EMS								
Redacción de memoria								

2. Estudio del Estado del Arte

2.1 Modelos de Producción

La energía solar se puede aprovechar de diversas maneras, entre las que se distinguen la utilización de métodos indirectos y directos. Para la aplicación de estos últimos se utilizan sistemas fotovoltaicos que se entienden como un conjunto de componentes mecánicos, eléctricos y electrónicos que captan la radiación solar y la transforman directamente en energía eléctrica.

Como se describe en [2], existen dos principales grupos de sistemas de generación fotovoltaica en función del modo de empleo de la energía producida:

- **Sistemas de generación autónomos, aislados u “*off-grid*”**, en los que la energía generada por módulos fotovoltaicos es almacenada en acumuladores (baterías) para luego ser empleada en instalaciones. Como se indica en [3], los principales componentes de este tipo de sistemas son un generador fotovoltaico, inversor, regulador y un sistema de baterías.

Dentro de este último grupo se encuentran los sistemas de generación mixtos o híbridos, que son un caso especial de los autónomos y cuya diferencia radica en que permiten el trabajo en conjunto con otra fuente de generación de energía eléctrica como por ejemplo la eólica, la hidráulica, etcétera.

- **Sistemas de generación fotovoltaica conectados a la red u “*on-grid*”**, que como se detalla en los artículos [4] y [5], inyectan la energía generada a la red de distribución. Los elementos que componen estos sistemas generalmente son un generador fotovoltaico, inversor, contadores para monitorear el flujo de energía, las cargas eléctricas, la red eléctrica y elementos de conexión a esta como pueden ser los convertidores DC-AC o el panel principal.

La energía generada por los módulos fotovoltaicos es directamente proporcional a la irradiancia solar que incide sobre su superficie, por lo tanto, muchas veces se integran baterías en los sistemas “*on-grid*” para compensar los desequilibrios entre consumo y generación. Más detalles se pueden encontrar en [6]. De este modo, la energía producida durante las horas de sol se puede utilizar en momentos en los que no se disponga de la suficiente radiación solar para generar la energía necesaria.

A continuación, se explican brevemente algunos parámetros necesarios para comprender el fundamento teórico de las instalaciones fotovoltaicas y se profundiza sobre varios de los componentes más esenciales de un sistema fotovoltaico.

2.1.1 Generadores fotovoltaicos

Un generador fotovoltaico es la asociación de varios módulos fotovoltaicos, también llamados paneles solares, que trabajan juntos para producir energía eléctrica a partir de

luz solar. A su vez, cada uno de los módulos se compone de un conjunto de células solares encapsuladas y montadas sobre un marco de soporte. Para obtener la potencia deseada, estas últimas pueden estar interconectadas en serie, incrementando así la tensión final en los extremos de la célula equivalente o en paralelo, para aumentar la intensidad total del conjunto.

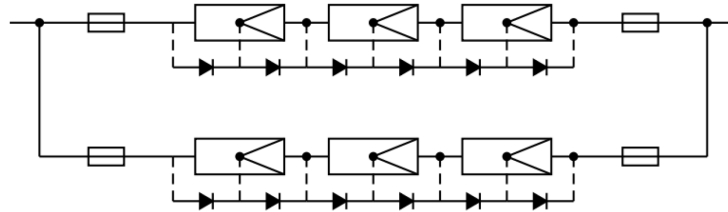


Fig 1 Esquema generador fotovoltaico compuesto por dos ramas de tres módulos en serie.

Fuente: Fig. 5.10, libro [7]

Por otro lado, el número de módulos por serie define la tensión del generador y el número de ramas define la corriente total del generador y así se nombra en [7].

El funcionamiento de los generadores se basa en el efecto fotoeléctrico que se produce en las células solares, las cuales actúan como diodos: cuando el material semiconductor que las forma (generalmente silicio) se excita ante la presencia de radiación, los electrones pueden fluir de la zona P a la N (expuesta a la radiación). Esto ocasiona un voltaje interno que ante la presencia de una resistencia producirá una corriente continua, como se menciona en [8].

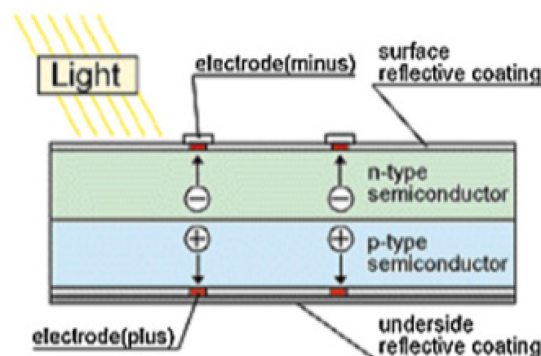


Fig 2 Célula solar de silicio y su funcionamiento. Fuente: Fig. 3, artículo [8]

Profundizando un poco más en los paneles, como aparece en los documentos [9] y [10], se debe conocer que existen distintos tipos, clasificados principalmente en tres generaciones según los materiales con los que se fabrican sus células: la primera y segunda generación en las que se utilizan distintas clases de silicio, tanto monocristalino como policristalino y amorfo. Por último, la tercera generación está todavía en proceso de desarrollo y busca conseguir con el uso de materiales innovadores de bajo coste de producción una mayor eficiencia en la conversión de la luz solar a energía eléctrica.

También es esencial conocer que existe una influencia de la temperatura y radiación en el funcionamiento de los módulos fotovoltaicos, como se muestra en las Fig 3 y Fig 4 respectivamente. A partir de dichas influencias se definen unas condiciones estándar denominadas STC ("Standard Testing Condition") con las que caracterizar las células y

utilizadas para la determinación de los parámetros eléctricos de los paneles. Estas corresponden a:

- Irradiancia: $G_{\text{stc}} = 1000 \text{ W/m}^2$ con incidencia normal.
- Temperatura de célula: $T_c = 25 \text{ }^\circ\text{C}$.
- Masa de aire: $AM = 1.5$

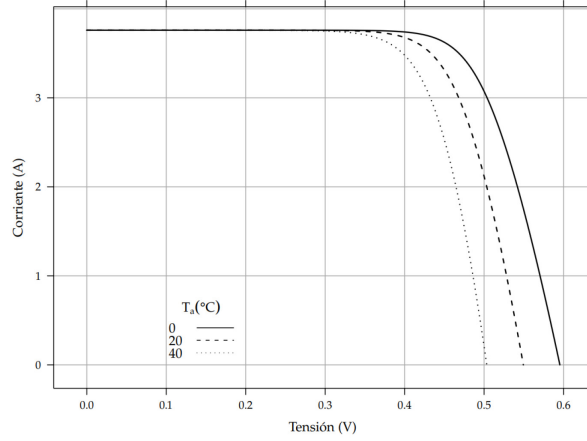


*Fig 3 Efecto de la temperatura en la curva característica de una célula solar ($G = 800 \text{ W/m}^2$).
Fuente: Fig. 4.11, libro[11]*

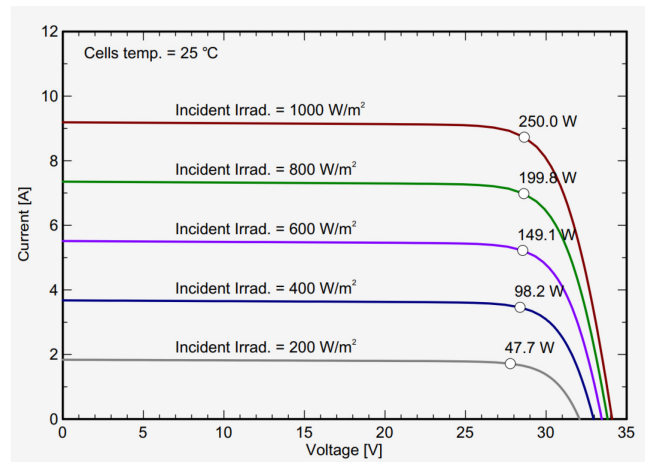


Fig 4 Efecto de radiación en panel 50 Wp EverExceed a 25°C Fuente: [12]

Algunos de los mencionados parámetros eléctricos que determinan los datos técnicos de los módulos y que se obtienen modelando las células como distintos circuitos, son el voltaje en vacío V_{ca} (entre los terminales del panel en circuito abierto), la corriente de cortocircuito I_{cc} (generada por el panel los terminales se encuentran cortocircuitados) o la potencia de pico, W_p (máxima potencia que es capaz de entregar el panel a una carga entre sus terminales) medida en Wp y para la que es necesario conocer el punto de máxima potencia (MPP), en el que se equilibran los valores máximos de corriente y tensión. Permanecer el máximo tiempo posible en este punto es el objetivo de cualquier instalación fotovoltaica. Esto lo podemos consultar en como se muestra en [11], [12].

2.1.2 Geometría solar

Para entender cómo trabajan los componentes explicados, resulta necesario tener unas nociones básicas sobre los ángulos que determinan la orientación de los paneles. Como se explica en el artículo [13], definir esta orientación correctamente es imprescindible para optimizar la eficiencia en la generación de energía eléctrica, así como para calcular la cantidad de luz solar que incide en los módulos fotovoltaicos.

La disposición de los paneles viene pues definida por los ángulos de inclinación (posición respecto a la superficie horizontal) y azimuth (posición respecto al plano de referencia vertical). Estos ángulos a su vez, como desarrolla el autor de [14], se determinan atendiendo a los ángulos cenital y azimuts solares, que corresponden al ángulo formado entre la dirección solar y el vector cenital, perpendicular a la superficie terrestre y el formado por el meridiano solar y el meridiano del lugar respectivamente. En la siguiente figura aparecen representados esquemáticamente estos ángulos, siendo θ_z el ángulo cenital, α_s el complementario de este último, llamado ángulo elevación solar, γ_s el ángulo azimutal solar, β la inclinación del panel y γ el ángulo azimut del mismo.

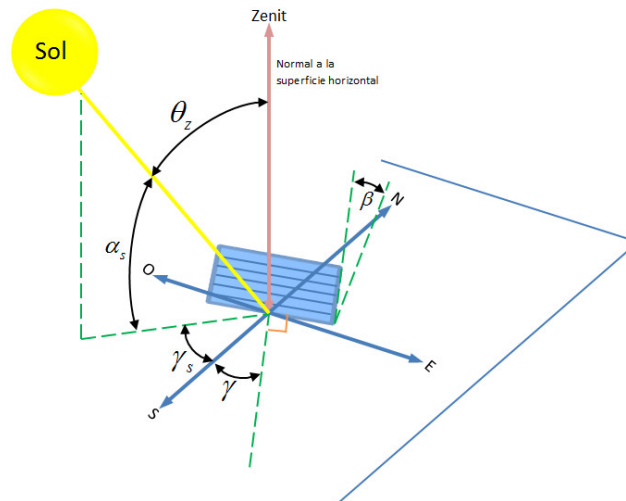


Fig 5 Esquema geometría solar y orientación del panel

La irradiancia solar es la cantidad de energía (radiación electromagnética solar) que llega a la superficie terrestre en un área determinada, expresada en vatios por metro cuadrado (W/m^2). Dependiendo de la geometría solar descrita y otros factores quedan definidas tres contribuciones diferentes de la irradiancia solar que incide sobre una superficie (en este caso los paneles de la instalación de estudio):

- Directa: hace referencia a la cantidad de energía solar que llega en línea recta del sol.
- Difusa: es la fracción de irradiancia que incide en la superficie después de haber sido dispersada por la atmósfera o por otros obstáculos.
- Reflejada: se trata de la cantidad de energía procedente de la reflexión en el suelo (supone una contribución muy pequeña e incluso despreciable).

2.1.3 Inversores

Son elementos imprescindibles en instalaciones conectadas a red ya que son los responsables de transformar la señal de potencia producida por el generador en tensión continua a tensión variable. Es decir, es el encargado de la conversión de corriente continua a alterna con las mismas características de la red eléctrica (220 voltios de valor eficaz y 50 Hz de frecuencia).

Se pueden clasificar en tipo “*on-grid*”, cuyo control se acopla a la red y “*off-grid*”, capaces de generar la onda sinusoidal. Además, también se puede distinguir entre monofásicos y trifásicos, dependiendo principalmente de la potencia.

2.2.4 Baterías

Las baterías convencionales, BES (“*Battery energy storage*”) o acumuladores son dispositivos compuestos por celdas galvánicas o electroquímicas, con la capacidad de acumular energía química, transformarla en energía eléctrica y producir así un voltaje entre sus terminales. Su uso es muy interesante ya que evitan, como se puede consultar en [15] que se produzcan grandes picos de consumo en breves periodos de tiempo al ser capaces de almacenar gran cantidad de energía que será utilizada en las horas de mayor demanda de consumo. Es decir, consiguen el recorte de picos o llenado de valles (“*peak shaving*”- “*load levelling*”).

Acerca de las características físicas de las baterías hay que considerar principalmente: la capacidad nominal (Ah) o la cantidad de energía que el acumulador es capaz de almacenar al comienzo de su vida útil; la tensión nominal (V) es decir, el voltaje en funcionamiento; la energía específica (Wh/kg), medida de la energía que el dispositivo puede almacenar por unidad de peso; potencia específica (W/kg) que relaciona la de la batería con su peso. Por último, las densidades de energía (Wh/L) y de potencia (W/L) son medidas que relacionan la energía y potencia con el volumen respectivamente, así se puede leer más extensamente en [16]–[18].

También es importante conocer los conceptos de “*State of Charge*” (SoC) o Estado de Carga [%], y “*State of Health*” (SoH) o Estado de Salud [%], siendo el primero una métrica indicadora de la cantidad de energía almacenada en el dispositivo de almacenamiento en el momento actual en relación con su capacidad máxima de almacenamiento, de modo que un SoC = 100% corresponde a la batería completamente cargada. El Estado de Salud por su parte, plasma el envejecimiento de la batería, siendo la diferencia entre la carga máxima de la batería y capacidad inicial de esta, por tanto, un estado de salud del 100% se refiere a la capacidad inicial del dispositivo e ira disminuyendo con el uso y degradación de la batería, como explican los autores de [19], [20].

El cálculo del SOC de la batería en el momento actual se puede efectuar mediante distintas técnicas, siendo el método de “*Coulomb Counting*”, como aparece en [21], uno de los más habituales debido a su sencillez. Este último se basa en la aplicación de la siguiente ecuación:

$$SOC_{initial} = SOC_{initial} - 100 \times \frac{\int_0^t Idt}{Q} \quad (1)$$

Donde $SOC_{initial}$ es la carga en porcentaje de la batería en el estado inicial, I la corriente de carga o descarga (de valor negativo en el estado de carga y positivo en el caso contrario) y Q la capacidad nominal de la batería.

Por otro lado, se pueden utilizar diversos modelos eléctricos equivalentes para definir el comportamiento de una batería. Uno de los más sencillos y aceptados, conocido como modelo Thévenin, considera una batería como una fuente de tensión, V_{BI} que será función de SOC, en serie con una resistencia, R_{BI} , como se muestra en la Fig 6:

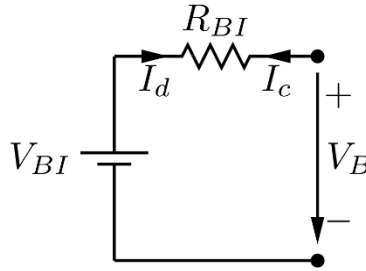


Fig 6 Modelo eléctrico de batería Fuente: Fig. 7.2, libro [11]

De este modo, la tensión de salida del acumulador durante la carga se puede calcular como:

$$V_B = V_{BI} + I_C \times R_{BI} \quad (2)$$

Y la misma tensión en el proceso de descarga será:

$$V_B = V_{BI} - I_C \times R_{BI} \quad (3)$$

En cuanto a la clasificación de las baterías, éstas pueden ser primarias en caso de no ser recargables, o secundarias, en caso contrario. Dentro de estas últimas se encuentran distintos dispositivos atendiendo a su naturaleza interna, como son los de plomo-ácido, níquel, ion sodio o ion litio. Por último, cabe destacar que existen nuevas líneas de investigación para el desarrollo de otros tipos de baterías como son las de flujo redox de Vanadio (VRB) o ZnBr o las baterías NaS.

Las baterías de litio, de las que consta la instalación de estudio de este trabajo, gracias a su alta capacidad energética, su ciclo de vida prolongado y su bajo ratio de autodescarga son las más idóneas para su uso en sistemas conectados a red y así se puede consultar en [22].

A continuación, se muestra una tabla comparativa de las características de algunos de los tipos de celdas mencionados, extraída de [23]:

Tabla 2 Características distintas químicas de batería

Características	Plomo-acido	Li- ion	VRB
Densidad de energía (Wh/kg)	25-50	75-200	43768
Densidad de potencia (W/kg)	75-300	500-2000	80-150
Ciclo de vida (Profundidad de descarga del 100%)	200-1000	1000-10000	>12000
Coste (\$/kWh)	100-300	300-2500	150-1000
Eficiencia de almacenamiento (%)	75-85	85-97	75-95
Auto descarga	Baja	Media	Despreciable

Con respecto al proceso químico que tiene lugar en el interior de las celdas durante la carga y descarga del dispositivo, existen numerosos modelos electroquímicos que lo explican basados en las reacciones que se producen. Así, en el caso de las baterías de litio, como se detalla en [24], durante la carga los iones se trasladan desde el electrodo negativo (ánodo) hacia el electrodo positivo (cátodo) a través del electrolito. Este movimiento de los iones de litio se produce gracias a que se aplica una diferencia de potencial entre los electrodos. Durante la descarga de una batería, se produce el proceso inverso.

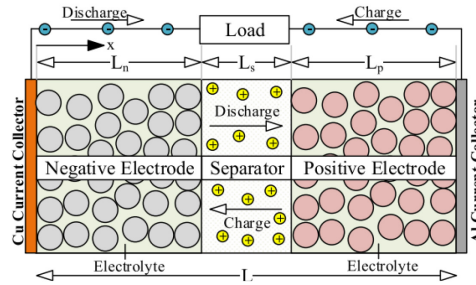


Fig 7 Esquema de Modelo P2D en batería de Li-ion. Fuente Fig. [3], artículo [26]

2.2 Modelos de Predicción

Como se ha mencionado anteriormente y exponen diversos autores, como en [25],[26] o [27], es necesario crear modelos de predicción en instalaciones fotovoltaicas debido a la gran dependencia de la energía producida y consumida a gran cantidad de factores externos, como la radiación solar, la temperatura ambiental, la humedad o la sombra, que pueden variar de manera impredecible. Revisando la bibliografía existente sobre el tema encontramos que, tanto en producción como en consumo, se utilizan infinidad de modelos de predicción, variando su aplicación en función del horizonte temporal que se desea anticipar (corto, medio o largo plazo). Así, se puede hacer una clasificación a grandes rasgos en tres grupos principales, cuyas ventajas e inconvenientes principales se plasman en la Tabla 3:

Tabla 3 Comparación de los principales métodos de predicción

MODELOS	CARACTERÍSTICAS PRINCIPALES	VENTAJAS	INCONVENIENTES	EJEMPLOS
ESTADÍSTICOS	Se basan en el análisis de datos históricos y en el uso de técnicas estadísticas para hacer predicciones sobre una variedad de fenómenos.	- No requieren información meteorológica adicional. - Realizan una buena estimación para horizontes de tiempo cortos.	-Dependen de la precisión de los datos históricos. - Poca flexibilidad al basarse en supuestos	- <i>Generalized linear model (GLM)</i> - <i>NWP (Numerical Weather Prediction)</i> - <i>ARIMA ("AutoRegressive Integrated Moving Average")</i>
BASADOS EN EL APRENDIZAJE AUTOMÁTICO	Utilizan algoritmos de aprendizaje automático, como redes neuronales, para aprender de los datos históricos de producción	Pueden trabajar con gran cantidad de datos y detectar patrones no lineales en ellos	-Resultan más difíciles de implementar -En muchos casos su interpretación resulta difícil ya que actúan como una caja negra - Necesitan mayores recursos computacionales -Necesitan una gran cantidad de datos para ser entrenados	- Redes neuronales - Árboles de decisión
BASADOS EN SISTEMAS FÍSICOS	Utiliza modelos matemáticos para describir el comportamiento de los componentes de la instalación fotovoltaica y simular su funcionamiento	-Resultan útiles para el diseño y la planificación de nuevas instalaciones fotovoltaicas -No necesitan una gran cantidad de datos históricos para su correcto funcionamiento	Necesitan información detallada y precisa sobre la instalación.	- Circuitos equivalentes de las celdas PV - Modelos térmicos de las celdas PV

Cabe destacar que muchos modelos se pueden considerar híbridos ya que combinan características de las técnicas ya mencionadas.

A continuación, se van a explicar en profundidad dos de los modelos cuyo uso está más extendido. Por un lado, aquellos basados en redes neuronales ya que, como se detalla en [28], actualmente se busca desarrollar algoritmos de predicción más sofisticados que sean capaces de predecir a largo plazo y no requieran de conocimientos del comportamiento del sistema PV ni se encuentren limitados por la no linealidad de los patrones de los datos. Esto resulta especialmente importante en la predicción de la producción fotovoltaica debido a la gran variabilidad de factores como la dirección e intensidad de la luz solar.

Por otro lado, los modelos ARIMA, cuya popularidad se debe entre otras cosas a su capacidad de capturar patrones de estacionalidad y tendencia en los datos, lo que es de gran ayuda en la predicción a corto plazo.

2.2.1 Redes neuronales

Estos modelos también llamados ANN (*“Artificial Neural Network”*), son algoritmos inspirados en el funcionamiento del cerebro humano de modo que, basándose en la capacidad de aprendizaje del sistema nervioso, la red aprende a identificar y generalizar los patrones asociados a los valores de entrada a partir de los cuales generará salidas.

La arquitectura de estas redes, como se puede consultar en los trabajos [29], [30], consiste en un conjunto de nodos o neuronas artificiales interconectadas entre sí mediante relaciones que tienen asignado un valor numérico modificable (peso). Así, cada neurona puede recibir una suma de entradas multiplicadas por sus respectivos pesos, como si se tratara de un “impulso nervioso”, y al mismo tiempo producir información en una salida que puede servir de entrada a otro nodo y así sucesivamente. Esto, se traduce en que en el interior de todas las neuronas los datos de entrada son procesados mediante funciones de red que comúnmente suelen ser combinaciones lineales de las entradas más un sesgo (*“bias”*).

Por otro lado, las neuronas además se encuentran agrupadas en diferentes niveles denominados capas, considerándose como “ocultas” aquellas que se encuentran entre la capa de entrada y la capa de salida. Se denominan así porque a diferencia de las anteriores, sus valores de salida no son observables. Estas capas tienen gran importancia ya que son las que permiten a una red modelar relaciones no lineales complejas entre los datos de entrada y salida, de modo que a medida que se agregan más capas de este tipo más aumenta esta capacidad de la red de modelar relaciones no lineales. Esto se produce mediante la aplicación en cada una de las neuronas de la capa oculta de una función de activación.

Existen diferentes tipos de funciones de activación, y su elección depende del problema que se está abordando y de la arquitectura de la red neuronal, siendo utilizada la misma función en todas las neuronas de una misma capa. Algunas de las más comunes

son la función sigmoide, la función tangente hiperbólica o la función ReLU (“*Rectified Linear Unit*”).

A continuación, se muestra un esquema de la estructura de una neurona:

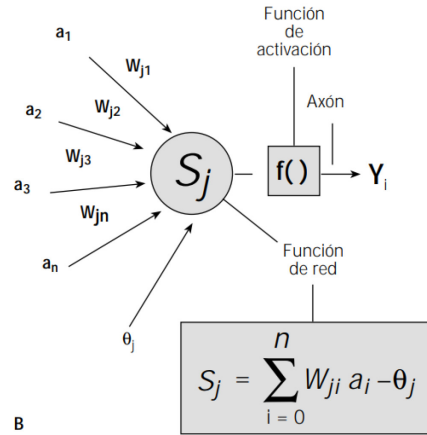


Fig 8 Estructura de una neurona artificial junto con su función de red. Fuente: Figs. 2a y b, [33]

Siendo S_i el valor de la salida de la neurona i , $[a_1, a_2, \dots, a_n]$ el vector de entradas, w_{ji} el peso de la conexión entre las neuronas i y j y θ_i el umbral o sesgo (“*bias*”) de la neurona i .

Por otro lado, se pueden encontrar distintos tipos de redes neuronales, como pueden ser la modular, la de función de base radial (RBF), o las redes multicapa o MLP (“*Multilayer Perceptron*”), como la que ha sido diseñada en este trabajo y que se representa a continuación:

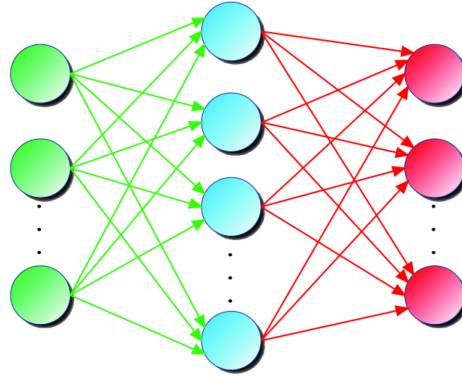


Fig 9 MLP con una capa oculta. Fuente: Fig 5.2, trabajo [31]

Esta clase de redes se denominan así porque pueden constar de varias capas ocultas. En el caso de la figura anterior, solo existe una, representada en color azul, mientras que las neuronas en color verde forman la capa de entrada, cuyo número depende de las variables de entrada. Por último, la capa roja representa la capa de salidas, que tendrá tantas neuronas como numero de variables se desea predecir.

2.2.2 Modelos ARIMA

Su uso es muy habitual en la predicción en el contexto de instalaciones fotovoltaicas, como el implementado en el trabajo [31], entre otros. Como, se explica en el artículo [32] estos modelos ARIMA (*“AutoRegressive Integrated Moving Average”*) son estadísticos y basan sus predicciones en datos históricos de series temporales, es decir, tienen en cuenta la dependencia entre las distintas observaciones a la hora de modelar.

Son la combinación de modelos AR (*“AutoRegressive”*) o autorregresivos y AM (*“Moving Average”*) o de medias móviles, los cuales asumen que el valor actual es función de los valores pasados. Estos modelos determinan los parámetros p (número de observaciones pasadas o rezagos necesarios para predecir el valor actual) y q (que indica el número de componentes aleatorios previos que configuran el valor presente en la serie temporal).

Al tratarse de modelos autorregresivos, similares a modelos lineales de regresión, el valor actual de la serie dependerá de los valores previos y se trabajará mejor si estos no están correlacionados entre ellos. Esto se traduce en que la serie de datos ha de ser estacionaria y no ser función del tiempo, es decir, la media y la varianza han de ser constantes a lo largo del tiempo.

El número de diferenciaciones que se precisen para conseguir la estacionariedad determinará el tercer parámetro necesario para definir el modelo: d. De este modo, la expresión algébrica de un modelo ARIMA queda definida como:

$$X_t^d = c + \phi_1 X_{t-1}^d + \dots + \phi_p X_{t-p}^d + \theta_1 \varepsilon_{t-1}^d + \theta_2 \varepsilon_{t-2}^d + \dots + \theta_q \varepsilon_{t-q}^d + \varepsilon_t^d \quad (4)$$

Donde X_t^d es la variable dependiente en el tiempo t (en el instante actual), c es una constante, $\phi_1, \dots, \phi_p$ son los coeficientes del modelo autorregresivo (AR) de orden p (representan el impacto que tienen los valores anteriores de la variable dependiente en el valor actual), $\theta_1, \dots, \theta_q$ son los coeficientes del modelo de media móvil (MA) de orden q (representan el impacto que tienen los errores anteriores en el valor actual) y ε_t es el error actual. De manera que cuánto más elevados sean los valores de ϕ_1 y θ_1 , mayor serán la influencia el valor anterior de la variable dependiente y del error anterior sobre el valor actual.

Sin embargo, aunque estos modelos pueden trabajar con datos que presentan tendencia, no admiten el componente estacional, es decir la existencia en la serie temporal de patrones regulares y predecibles que se repitan durante un período de tiempo específico. Esto puede resultar relevante en el modelado de consumo de una instalación fotovoltaica, donde un claro patrón es la variación en la cantidad de luz solar incidente. Así, ésta siempre será máxima a ciertas horas del día y mínima a otras, produciendo fluctuaciones con un periodo de 24 horas en el consumo. Aparece así, en la búsqueda de una mayor precisión, la necesidad de utilizar, como en el caso del presente trabajo y del autor de [33], extensiones del modelo ARIMA ya sea la denominada SARIMA (*“Seasonal AutoRegressive Integrated Moving Average”*) o la versión SARIMAX (*“Seasonal AutoRegressive Integrated Moving Average with exogenous variables”*).

Para definir este tipo de modelos se utilizan los mismos parámetros del modelo ARIMA, pero incorporando ahora además los parámetros P, D, Q (los órdenes del modelo autorregresivo, de integración y del modelo de media móvil estacionales respectivamente) y el periodo estacional (s). Además, los modelos SARIMAX permiten incluir variables exógenas, es decir, aquellas que afectan al resultado que estamos tratando de predecir, pero que no son parte del proceso que se está modelando, mejorando así la precisión en la predicción de la variable de interés. En el caso de este trabajo, serán todas las variables que influyen en el consumo, pero no están directamente relacionadas con la instalación, como pueden ser la temperatura exterior o el mes del año estudiado, que hará variar la cantidad de energía consumida.

2.2.3 Optimizador para el Sistema de Gestión de Energía (EMS)

Por último, en el caso de las instalaciones que cuentan con una batería, ésta resulta un componente crítico, ya que almacena la energía generada por los paneles solares y permite su uso posterior cuando no hay suficiente luz solar.

Por tanto, la optimización del uso de estos dispositivos es esencial para, por un lado, asegurar el suministro de energía requerido, y por otro, maximizar su vida útil y rendimiento. Así será necesario definir un EMS (*“Energy Management System”*) o sistema de gestión de energía.

En este contexto, con el propósito de conseguir una generación más constante en los sistemas fotovoltaicos, existen numerosos estudios sobre los diferentes algoritmos de control y las políticas de gestión para los sistemas de almacenamiento de energía. Algunos ejemplos de distintas metodologías utilizadas para optimizar el uso de los dispositivos de almacenamiento son: la usada en el trabajo [34] donde se propone una optimización mediante algoritmos genéticos utilizando un algoritmo de *“Rainflow counting”* y reduciendo el envejecimiento de la batería; el autor de [35] presenta un sistema de control predictivo considerando el efecto de envejecimiento de la batería y con el objetivo de conseguir *“peak shaving”* o por último los modelos MPC (*“Model Predictive Control”*) como el que se aplica en [36], en los que se va a centrar este trabajo.

Los MPC, como se menciona en [37], son modelos de control que, a partir del comportamiento futuro predicho del sistema por un modelo matemático previamente diseñado, actúan para optimizar los criterios establecidos en el horizonte de tiempo deseado. Es decir, el MPC calcula las acciones requeridas para cumplir con unos objetivos de control y restricciones en el sistema establecidos y tras la elección de una función real a minimizar (función objetivo o de coste) que es la que indica el criterio a optimizar.

En cuanto a las ventajas del uso de este modelo, destaca que permite optimizar en tiempo real y adaptarse a los cambios del sistema, establecer restricciones para garantizar la estabilidad del sistema y tratar con sistemas lineales y no lineales.

3. Metodología

3.1 Herramientas utilizadas

En este apartado se describen brevemente los recursos de los que se ha hecho uso para llevar a cabo la implementación y validación de los modelos desarrollados:

- En primer lugar, se ha utilizado el lenguaje de programación Python, cuyo uso resulta sencillo y está muy extendido. Posee la gran ventaja de que existen gran cantidad de bibliotecas especializadas. Así, algunas de las utilizadas han sido: “*TensorFlow*” (utilizada para construir nuestra red neuronal), “*Keras*” (usada complementariamente junto a “*TensorFlow*”, facilita el entrenamiento y diseño de la red), “*Pandas*” (empleada para leer y manipular el conjunto de datos), “*Numpy*” (para realizar cálculos), *statsmodels* (para el modelado de nuestro SARIMAX), “*Matplotlib*” y “*Plotly*” para graficar.
- Visual Studio: es el entorno de desarrollo integrado (IDE) que ha sido empleado para escribir el código en Python.
- Para la obtención de los datos necesarios para entrenar el modelo de predicción de producción (que más adelante se presentarán explicativamente) se ha hecho uso de la herramienta PVGIS (“*Photovoltaic Geographical Information System*”), referenciada en [38]. Se trata de un sistema de información geográfica fotovoltaica que almacena información precisa y actualizada sobre la producción solar en Europa, proveniente principalmente de datos satelitales y con la que se puede estimar la producción en una ubicación determinada.
- Bases de datos: para poder entrenar y validar el modelo de predicción de consumo fue necesaria la búsqueda de un “*Dataset*” adecuado cuyo contenido se va a exponer más adelante.
- Además, la implementación y validación de los modelos de este trabajo ha requerido procesar grandes cantidades de datos y su entrenamiento ha resultado un proceso muy exigente y complejo que llevaba demasiado tiempo para completarse en un ordenador personal. Para trabajar con este gran volumen de datos y el procesamiento de los modelos el uso de un servidor se volvió necesario. Para ello, se hizo uso del clúster de servidores Claudio de Madrid, propiedad del CSIC, al cual se accedía mediante acceso remoto vía SSH (“*Secure Shell*”). De esta manera, se ha podido agilizar el entrenamiento de los modelos notablemente y realizar muchas más pruebas, ya que además contaba con la gran ventaja de que permitía lanzar procesos paralelos con los que ejecutar diferentes configuraciones cambiando los distintos parámetros de los modelos.
- Por último, para llevar a cabo un control de las distintas versiones de los modelos que se iban realizando y poder revertir en caso de que fuera necesario los distintos cambios en el código, se utilizó la herramienta GitLab. De esta forma, además, los directores también tenían acceso y seguimiento de los avances del trabajo. El repositorio de código se ha publicado para la revisión pública del mismo y es consultable en [1].

3.2 Modelo de producción de la instalación

3.2.1 Condiciones y especificaciones

Como se ha comentado previamente, en el caso de la predicción de la producción se ha elegido un modelo de redes neuronales, para cuya implementación se han utilizado los datos obtenidos gracias a PVGIS, herramienta que permitió introducir las especificaciones de la instalación solar de estudio, proporcionadas por el CSIC. Estas fueron: la potencia de pico instalada (33.8 kWp), las pérdidas del sistema (14%) y los ángulos de azimut (0°) e inclinación (10°). de las placas. Con estas informaciones se descargó una base con datos horarios durante los años comprendidos entre 2005 y 2020.

El archivo CSV (“*comma-separated values*”; fichero cuyos valores están separados por coma) obtenido contiene distintas informaciones como la altura solar o la velocidad del viento, pero para este trabajo fueron necesarias únicamente las diferentes componentes de la irradiancia solar. En concreto, la irradiancia directa (Gb), proporcionada en W/m^2 , cuyos valores fueron los elegidos como datos de entrada para el entrenamiento de la red neuronal ya que está directamente relacionada con la potencia producida y los valores de esta última resultan muy fáciles de calcular a partir de ella.

A continuación, se presentan distintas gráficas para comprender más fácilmente los datos utilizados contenidos en el “*dataset*”.

En primer lugar, en la Fig 10 se muestra la irradiancia de un año completo de los presentes en el archivo, habiéndose elegido ejemplo 2018.

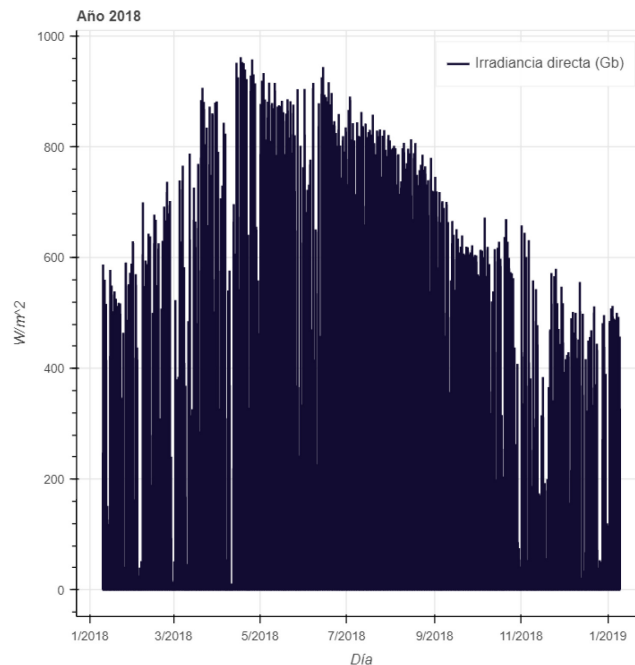


Fig 10 Irradiancia durante el año 2018

Se observa como el rango de valores representado oscila entre 0 y 900 W/m^2 y se aprecia el patrón anual que sigue la irradiancia solar, aumentando durante los meses de verano y reduciéndose durante el invierno. Para comprobarlo, se grafica en la Fig 11 la

irradiación directa durante un día de enero de 2018 y de julio de ese mismo año respectivamente. Así, se corrobora que las horas de sol son notablemente menores en invierno. Además, se pueden ver en ambas gráficas los ciclos diarios de día y de noche.

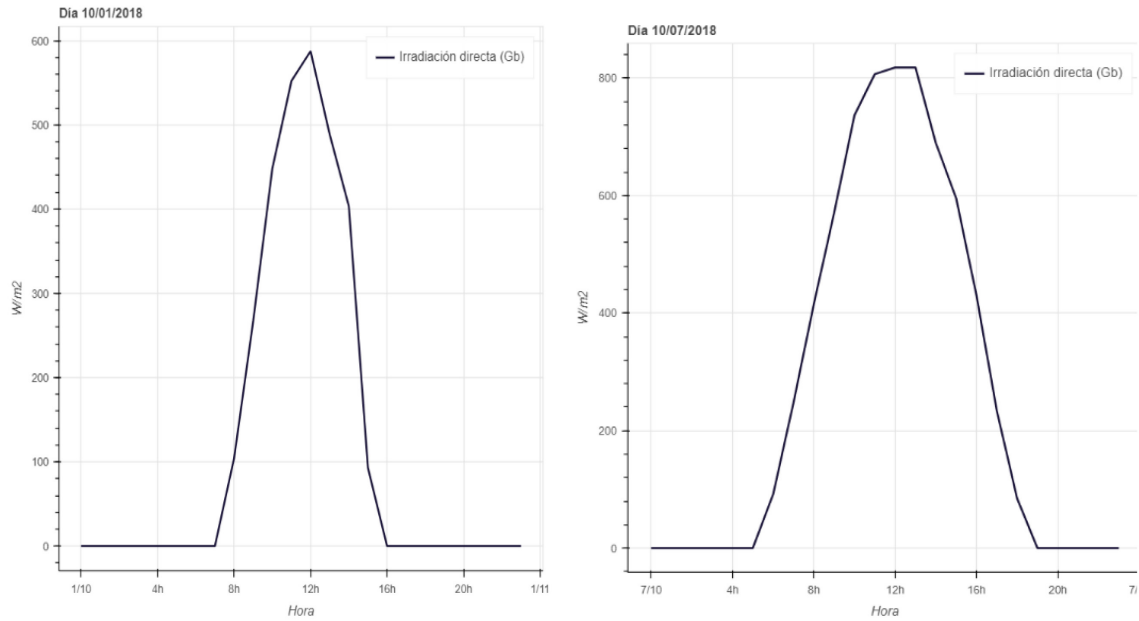


Fig 11 Irradiancia solar durante los días 10/01/2018 y 10/07/2018

Por último, se presentan en Fig 12 los valores de irradiancia medios de cada uno de los años presentes en el archivo CSV y expresados en unidades de W/m^2 . Se puede observar que el rango de valores oscila entre 0 y 160 W/m^2 , ya que como se ha comentado, durante muchas horas del día no incidirá luz solar en los paneles.

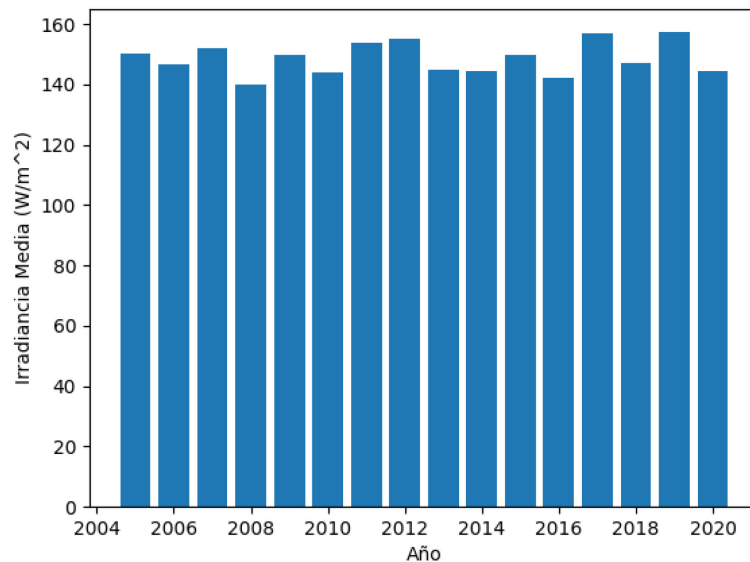


Fig 12 Irradiancia promedio por años del conjunto de datos

Por otro lado, para el entrenamiento de la red se estableció que se haría uso de 72 neuronas de entrada y 24 de salida (iguales al número de horas que se desea que sean predichas por el modelo). Esto se traduce en que, a partir de una fecha concreta

durante un periodo de tiempo determinado, introduciendo los datos en ventanas de 72 horas, el modelo predecirá las 24 horas futuras de irradiancia correspondientes a cada una de las horas del total de los días escogidos de entrada.

Además, para crear y entrenar el modelo, se decidió realizar el procedimiento descrito con los datos contenidos en 10 años comenzando desde el día 10/01/2005.

Sin embargo, no se fijó un número concreto de capas ocultas de la red ni sus respectivas neuronas, sino que se probaron modelos con distintas combinaciones. Lo mismo sucedió con el resto de los parámetros que se introducen al crear el modelo.

3.2.2 Métricas

Las métricas son las medidas utilizadas para evaluar la calidad del modelo. Existen múltiples maneras de evaluar los modelos, pero en este trabajo se han escogido dos de las más habituales, que se calculan a partir de un conjunto con N ejemplos con una salida real ($\mathcal{Y}$) y una salida estimada ($\hat{\mathcal{Y}}$).

- Error absoluto medio (“*Mean Absolute Error*”, MAE), se trata de una medida de la diferencia absoluta entre los valores predichos y los reales cuya fórmula es la siguiente:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (5)$$

Un valor bajo indica que el modelo está produciendo predicciones precisas.

- Raíz del error cuadrático medio (“*Root Mean Squared Error*”, RMSE): indica cuánto se desvían las predicciones del modelo de los valores reales en promedio. Cuanto menor sea su valor, mejor será la predicción del modelo:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (6)$$

3.3 Modelo de consumo de la instalación

3.3.1 Condiciones y especificaciones

Respecto al modelo de consumo se ha implementado un modelo SARIMAX, para cuyo entrenamiento se ha utilizado una base de datos, referenciada en [39], que contiene medidas minútales del consumo de energía eléctrica en un hogar durante un período de casi 4 años. En total, el archivo contiene 2075259 medidas tomadas entre diciembre de 2006 y noviembre de 2010, incluyendo la fecha, la potencia activa global expresada en kilovatios (kW) de la casa (es decir la potencia consumida), la potencia reactiva en kW, el voltaje en vatios (V), la intensidad global en amperios (A) y por último tres submediciones de energía expresada en vatios por horas y correspondientes a la cocina, una habitación de lavado, un calentador de agua eléctrico y un aire acondicionado. Estas tres últimas resultan importantes ya que como indica la persona que proporciona la base de datos la energía consumida total, que es de interés para crear el modelo de

consumo, será la medición de la potencia activa global dada, transformándola a kW/h, menos las tres submediciones nombradas.

Por otro lado, faltan algunas mediciones durante ciertos instantes, como por ejemplo en el día 28 de abril de 2007 y será necesario preprocesar el conjunto de datos para: por un lado, cambiar el valor de estos datos nulos y poder trabajar con ellos y, por otro, reagrupar la información del archivo de manera horaria para tener la misma frecuencia en los datos que en el modelo de producción.

Para ayudar a la comprensión de los datos del conjunto utilizados en el modelo se muestran ahora una serie de gráficas:

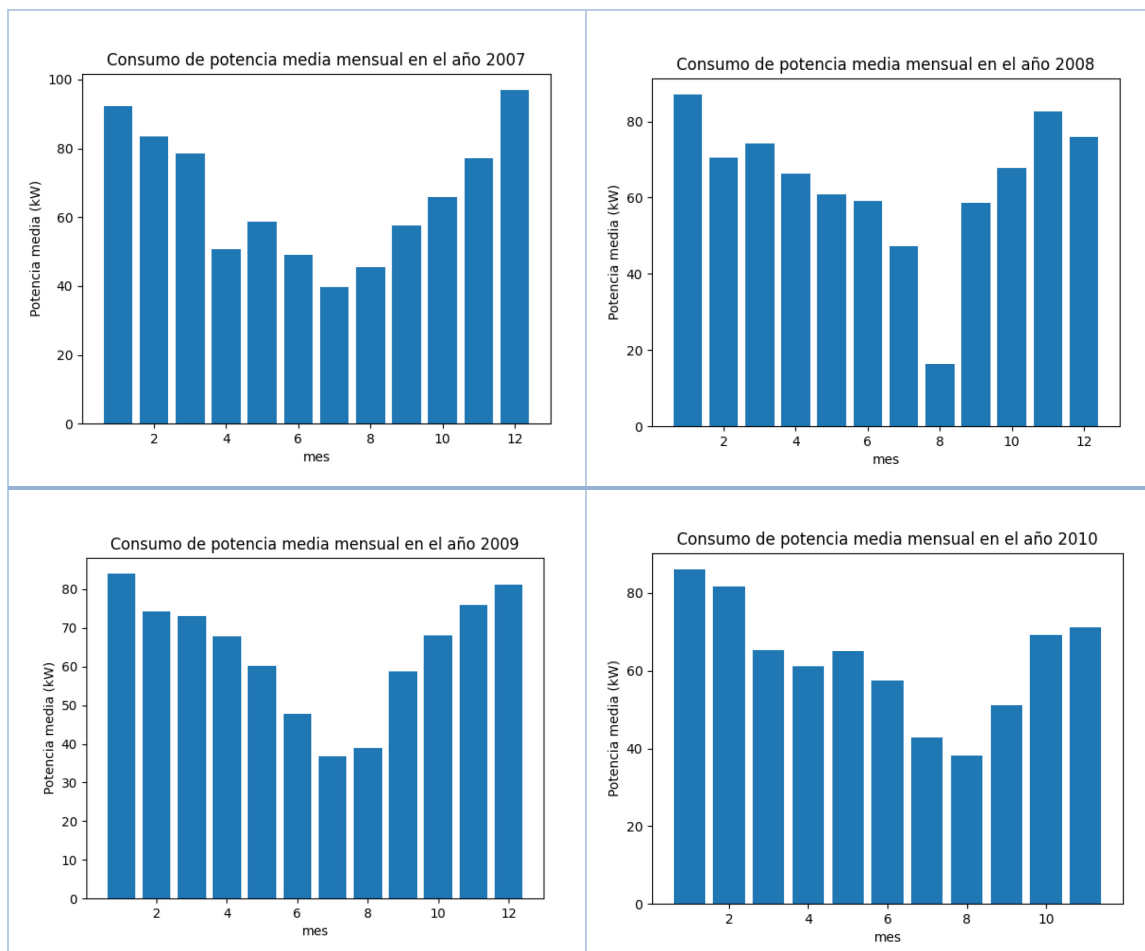


Fig 13 Gráficas de potencia media mensual consumida en distintos años de la base de datos de consumo

Primeramente, se muestra la comparativa de la media de consumo mensual durante los cuatro años de los que se dispone más información (solo se tienen valores del año 2006 disponibles a partir de diciembre del mismo año). En todos estos gráficos de barras se puede identificar una clara tendencia en el consumo, siendo este menor durante los meses de verano. Esto se debe a que en esas fechas factores como la existencia de más horas de luz durante el día, o mayores temperaturas son determinantes para que se produzca una disminución notable en el consumo, como se ha comprobado. Por otro lado, se puede observar cómo existen mediciones erróneas en la base de datos

proporcionada que pueden afectar en la predicción del modelo, como es el caso del mes de agosto de 2008 que presenta un consumo anómalamente bajo.

Por otro lado, se grafica la media de consumo anual de la base de datos al completo a excepción del año 2006 que apenas aporta información y se observa como los valores están comprendidos entre 0 y aproximadamente 65 kW.

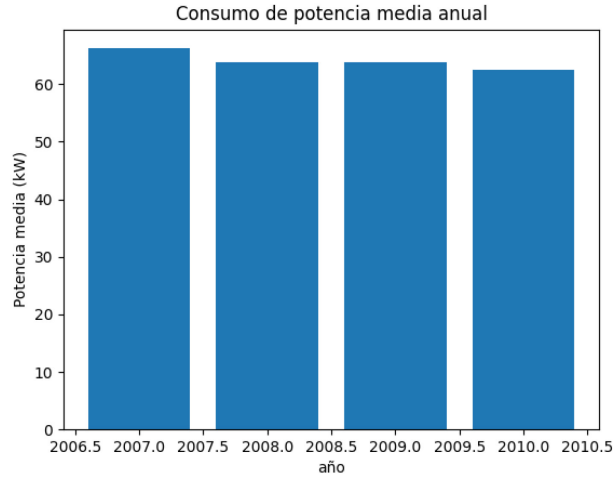


Fig 14 Media del consumo energético durante diferentes años

3.3.2 Métricas

En cuanto a las métricas de evaluación del modelo de consumo se decidió utilizar las ya explicadas en el caso de las redes neuronales y se añadió otra de las más comunes: el error porcentual absoluto medio (“*Mean Absolute Percentage*” Error, MAPE) que mide el porcentaje medio en que las predicciones difieren de los valores reales. Además, resulta muy útil debido a su fácil interpretación. Su expresión matemática es la siguiente:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (7)$$

Cuanto menor sea el porcentaje obtenido, mayor se la precisión de las predicciones.

3.4 Modelo de batería virtual

El modelo escogido se basa en la explicación dada anteriormente en el apartado 2.2.4 del estado del arte. Es decir, se trata de un modelo básico con “*Coulomb Counting*”, que se basa en la ecuación (1) seleccionado de esta manera al no ser el foco del trabajo.

3.5 Sistema de gestión de energía (EMS)

Como se ha mencionado antes, se ha elegido utilizar un modelo MPC, de modo que el primer paso para crearlo ha sido definir correctamente aquello que se quiere optimizar

en el sistema y las restricciones que se desea imponer, las cuales se detallarán en el siguiente subapartado y fueron:

- $P_{\text{cur}} \rightarrow 0$ (Potencia sobrante del sistema)
- $P_{\text{red}}^2 * 0.05$ (Potencia aportada por la red eléctrica o cedida a la misma)

3.5.1 Condiciones de contorno

En primer lugar, se establecieron las condiciones restrictivas y las funciones de coste necesarias:

- Por un lado, se busca que la energía sobrante sea cero, esto da lugar a la primera función de coste del modelo que hará que la potencia sobrante sea cero ($P_{\text{cur}} \rightarrow 0$).
- Por otro lado, primará el autoconsumo, es decir el uso de la batería antes que el consumo de la red eléctrica, lo que se traduce en una segunda función de coste que minimiza el cuadrado de la potencia entregada por la red al sistema en cada intervalo de tiempo ($P_{\text{red}}^2 * 0.05$).

3.5.2 Función objetivo

El modelo resuelve la ecuación del sumatorio de potencias que entra y sale al sistema, que será igual a cero, y así se garantiza la conservación de energía en la red eléctrica. Para esto, se ha definido un criterio de signos de modo que la potencia entrante a la red se establece como positiva y la saliente como negativa. Así, el balance de potencias resulta:

$$\sum P = 0 = \hat{P}_{\text{solar}} + \hat{P}_{\text{batería}} - \hat{P}_{\text{cons}} + \hat{P}_{\text{red}} - \hat{P}_{\text{cur}} \quad (8)$$

Siendo $\hat{P}_{\text{solar}}$ la potencia generada por las placas de la instalación, $\hat{P}_{\text{cons}}$ la que es consumida, $\hat{P}_{\text{red}}$ aquella que o se vierte a la red eléctrica (en este caso tendrá signo negativo) o se obtiene de la misma (siendo entonces positiva). Por último, $\hat{P}_{\text{batería}}$ es la potencia que se emplea para cargar o descargar la batería (siendo negativa en la primera situación y positiva en la segunda) y $\hat{P}_{\text{cur}}$ la potencia sobrante del sistema.

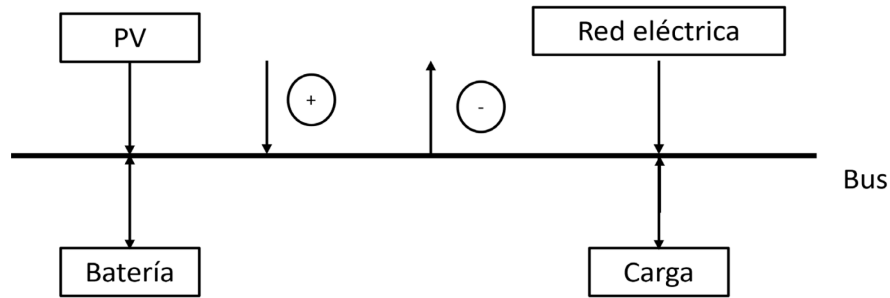


Fig 15 Esquema de la instalación con el criterio de signos utilizado

El modelo se diseñará utilizando la librería “*GEKKO*” y en él, ambas funciones objetivo serán sumadas y optimizadas, devolviendo como resultado una serie de valores de las potencias activas y reactivas que se deben establecer en cada instante de tiempo

para cumplir con el balance de potencias y las restricciones establecidas en el modelo. Sin embargo, estas potencias estimadas implican un error de predicción, por lo que como último paso se procederá a calcularlo a partir de esas primeras potencias estimadas y las potencias reales, de modo que:

$$E_{pred} = (P_{solar} - P_{cons}) - (\hat{P}_{solar} - \hat{P}_{cons}) \quad (9)$$

Siendo P_{solar} y P_{cons} los datos reales de potencia producida por las placas solares y potencia consumida por la instalación respectivamente, mientras que $\hat{P}_{solar}$ y $\hat{P}_{cons}$ son los valores estimados por el MPC de dichas potencias.

Si el resultado de este error resulta mayor o igual a cero significará que se necesita menor energía de lo estimado, mientras que si el error toma valores menores de cero se precisará más energía de lo estimado.

Finalmente, se distribuye el error calculado primero a la red, luego a la batería y por último a la potencia de “*curtailment*”, es decir la sobrante, de modo que la potencia de la red cumpla:

$$P_{min-red} \leq \hat{P}_{red} - E_{red} \leq P_{max-red} \quad (10)$$

Y si se supera alguno de estos límites se recalculará la potencia de la batería cumpliendo los rangos de la misma de la siguiente manera:

$$\hat{P}_{bateria} = \hat{P}_{solar} + \hat{P}_{cons} - \hat{P}_{red} + \hat{P}_{cur} - E_{red} \quad (11)$$

4. Resultados y discusión

4.1 Modelo de producción de la instalación

En este apartado se abordan los distintos pasos realizados para el diseño, implementación y análisis de las distintas redes neuronales.

4.1.1 Preprocesamiento de los datos

Para poder trabajar con los datos proporcionados por PVGIS fue necesario hacer algunas transformaciones en el conjunto. Por un lado, se modificó el formato de las fechas del archivo CSV a “*datetime*” que permite filtrar los datos en los periodos temporales deseados (necesario para crear las ventanas del modelo). Además, se referenciaron a cero minutos ya que en el archivo original todas las mediciones habían sido tomadas con 10 minutos de retraso respecto a las horas punta y así poder manipularlos más cómodamente.

Por otro lado, es necesario conocer la posibilidad de que aparezca sobreaprendizaje o sobreajuste (“*overfitting*” en inglés) en una red neuronal. Como se puede consultar en [40], se trata de un problema común que se produce cuando el modelo comienza a memorizar los datos de entrenamiento y, aunque para estos puede tener una gran precisión, no es capaz de generalizar y predecir correctamente ante nuevos datos de entrada. Para solucionar esto se decidió utilizar la validación cruzada (“*crossvalidation*”) que consiste en dividir los datos en tres conjuntos: entrenamiento (“*training set*”), test o verificación (“*test set*”) y validación (“*validation set*”). Así, el conjunto de entrenamiento será el empleado para ajustar los pesos del modelo mientras que el de validación se utiliza para decidir cuándo parar el proceso de entrenamiento, que será en el momento en el que el error de la validación (con la métrica escogida) sea mínimo. Superado este punto se producirá el nombrado sobreajuste de la red, como se muestra en la Fig 16:

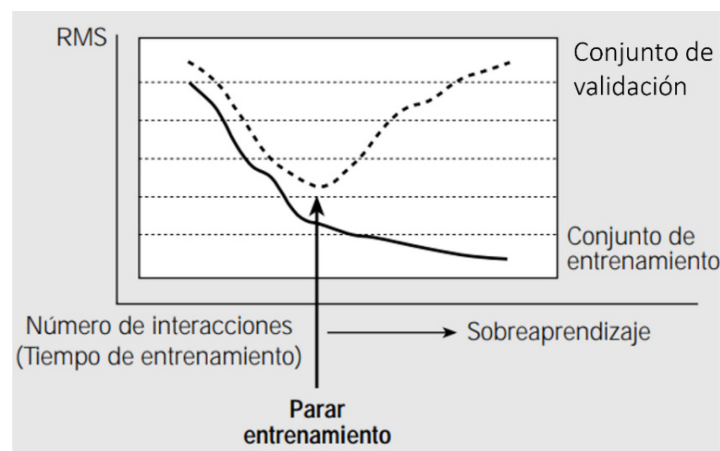


Fig 16 Criterio de parada de entrenamiento eligiendo métrica el error RMS. Fuente: Fig.4., artículo [43]

Una vez el modelo está entrenado se utilizará el conjunto de test para realizar la prueba final, prediciendo con los datos aún no utilizados y comprobando la precisión de la red.

Para este trabajo, se decidió que el conjunto de datos de validación estaría compuesto de dos años a partir de la fecha inicial 10/1/2015 mientras que el de test contendría un año desde el día 10/1/2018 y el de entrenamiento el descrito en el apartado 3.2.1. De esta forma, se dividieron los datos originales y se creó un archivo de “*inputs*” y otro de “*outputs*” de cada conjunto utilizando ventanas de 24 y 72 horas respectivamente. Esto se traduce en que se realizarán predicciones a un día futuro, utilizando para ello tres días previos.

Finalmente, se procedió a normalizar todos los grupos de datos creados para evitar que los valores más grandes dominaran a los pequeños, lo que podría alterar los pesos de las neuronas, como explica el autor de [41]. De esta manera, todos los valores de irradiancia se sitúan en el mismo rango de 0 a 1 y también se consigue que la interpretación de los resultados sea más sencilla. Para llevar esta normalización a cabo, se dividieron todos los valores entre la irradiación máxima que llega a la superficie terrestre a nivel del mar (1000 W/m^2).

4.1.2 Diseño de las redes neuronales

En este apartado se van a exponer las distintas combinaciones de los diferentes parámetros a introducir para crear el modelo con “*keras*”:

- **Número de capas y neuronas en cada una de ellas:** como se ha comentado anteriormente el número de neuronas en las capas de entrada y salida fue fijo. Sin embargo, como capas ocultas se probaron las que se muestran en la siguiente tabla:

Tabla 4 Distintas combinaciones de redes neuronales elegidas

Nº de modelo	Nº de capas ocultas	Neuronas por capa
1	2	[70,70,40,24]
2	3	[70, 60, 40,30, 24]
3	2	[70, 60, 40, 30]
4	5	[70, 75, 65, 50, 40, 30, 24]
5	3	[70, 60, 40,30, 24]
6	4	[70, 70, 40, 40, 30, 24]
7	2	[70, 60, 35, 24]
8	2	[70, 50, 40, 24]
9	4	[70, 70, 60, 50, 40, 24]

- **Funciones de activación:** se comenzó creando modelos con funciones ReLU (“*rectified linear unit*”) para todas las capas de la red. Estas funciones, como se puede consultar en [42], anulan los valores negativos de x y su expresión matemática es la siguiente:

$$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ x & \text{para } x > 0 \end{cases} \quad (12)$$

Esta función mejora el entrenamiento en las capas ocultas, pero puede generar “neuronas muertas”. Durante la implementación de los modelos se descubrió que estas aparecían en aquellas cuya entrada tenía valor negativo, ya que esta función provocaba que produjeran siempre una salida nula y que nunca se activaran durante el aprendizaje. Es decir, durante el proceso de entrenamiento no se ajustaban los pesos de estas neuronas.

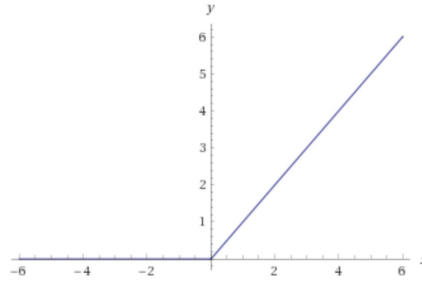


Fig 17 Gráfica de la función ReLU. Fuente: Fig 3.8 trabajo [45]

- **Función de coste** o pérdida (en inglés “*loss function*”): como se explica en [43] es una medida de la discrepancia entre los valores predichos por la red y los valores reales. Por tanto, durante el proceso de entrenamiento el objetivo es reducirla al máximo, ajustando para ello los pesos. Es muy común la aplicación de la función cuadrática MSE (“*Mean Squared Error*”) y es la que se ha elegido en todos los modelos, cuya expresión es:

$$MSE = \frac{1}{N} \sum_{i=1}^N (f(x) - y)^2 \quad (13)$$

Donde $f(x)$ es la predicción de la red neuronal, y el valor esperado y N el número de muestras totales del conjunto de datos.

- **Optimizadores:** se trata de algoritmos empleados para calcular los pesos de las redes neuronales con el objetivo de mejorar la predicción. En este trabajo se han entrenado los modelos con las capas descritas tanto con optimizadores ADAM (“*Adaptative Moment estimation*”) como RMSprop.
- **“Batch size”** o tamaño de lote: es la cantidad de observaciones del conjunto de entrenamiento que se utilizan en cada iteración. Para este trabajo se escogió para todos los modelos un tamaño de lote de valor 100.
- **Numero de épocas o “epochs”:** es el número total de interacciones completas que se realizan durante el entrenamiento, es decir, la cantidad de veces que se procesa todo el conjunto de datos de entrenamiento. Para la creación de los

primeros modelos se utilizaron 60 épocas, aunque luego se probaron otras combinaciones como se explicará más adelante.

Con estos primeros parámetros se obtuvieron 18 modelos de redes iniciales.

4.1.3 Desarrollo y aplicación de las redes neuronales

Una vez guardados los modelos mencionados se procedió a cargarlos y obtener las predicciones calculadas con los datos del conjunto de test. Al representar con todos ellos las distintas predicciones que estiman, se obtienen figuras similares a la Fig 18, que corresponde al segundo modelo en la lista. Así, las líneas de un mismo color representan las predicciones de irradiancia en ventanas de 24 horas que comienzan a la misma hora en los distintos días del año y, por tanto, habrá un total de 24 líneas por día representado.

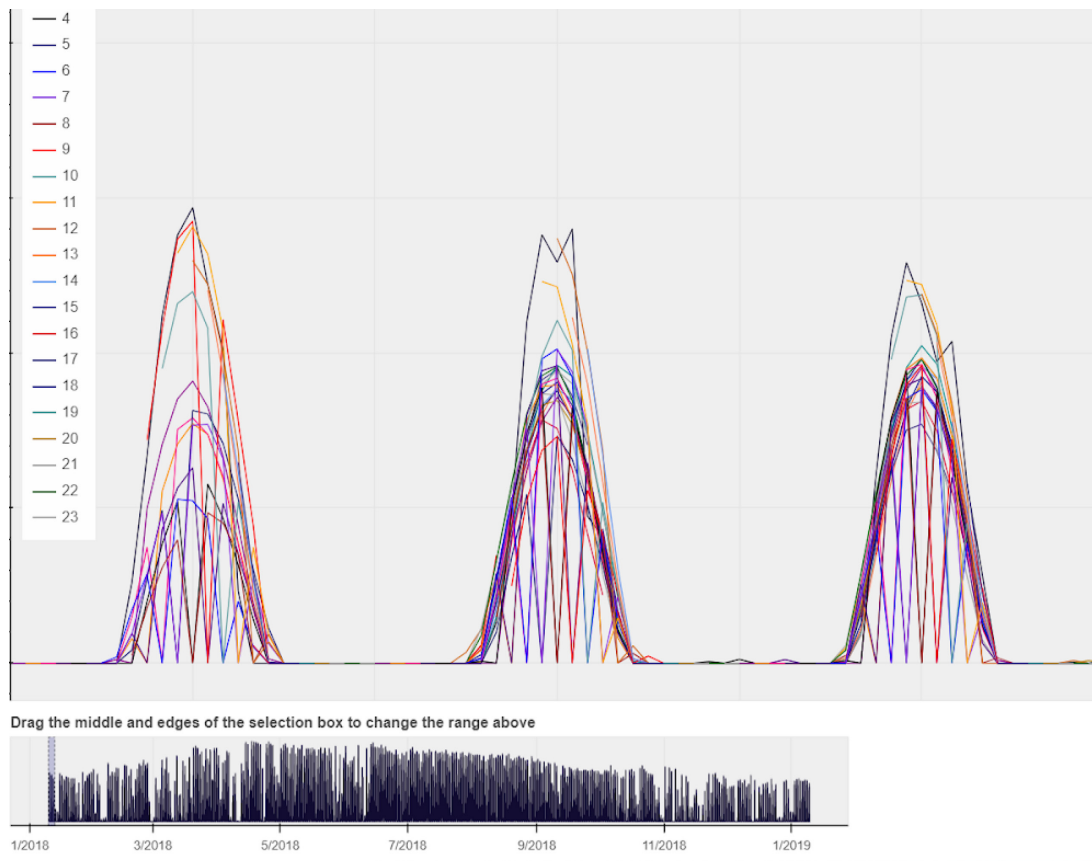


Fig 18 Predicción de la irradiancia en ventanas de 24 horas realizada por el segundo modelo

Al graficar se observan en todos los modelos varios picos en determinadas líneas, lo que indica que alguna de las neuronas no ha sido correctamente entrenada o tiene demasiada sensibilidad y da más peso a los valores de datos que corresponden a horas en la noche.

Algunas de las posibles soluciones podrían ser:

- Forzar a la red a comenzar el entrenamiento con una semilla que se elija, ya que es posible que los pesos iniciales que ha asignado aleatoriamente la red difieran

mucho del peso óptimo para algunas de las neuronas. Si se decidiera únicamente aumentar el número de “*epoch*” puede que la predicción mejorara, pero sin solucionar el problema ya que se puede producir un sobreajuste.

- Seleccionar otra métrica ya que puede que el error cuadrático medio no se ajuste adecuadamente al problema. Esto podría deberse a que con MSE se tiene en cuenta la diferencia entre resultados reales esperados y valores predichos por el modelo siendo muchos de los datos de ambos grupos idénticos ya que corresponden a horas de la noche (nulos).

Como parece haber un problema en la activación de ciertas neuronas, se realiza un análisis de sensibilidad para comprobar la sensibilidad de la salida a variaciones en la entrada. Este consiste en ir variando los valores de las 72 entradas de 0 a 1 para ver cómo afecta cada una de estas a las correspondientes 24 salidas. Para ello, se introducen en una matriz de identidad de dimensiones 72x72 como “Inputs” para la red y se grafica la matriz 72x24 correspondiente a las entradas y salidas con colores para que sea más visual e identificar fácilmente el fallo. En el caso del modelo 2 que se está mostrando de ejemplo se obtiene:

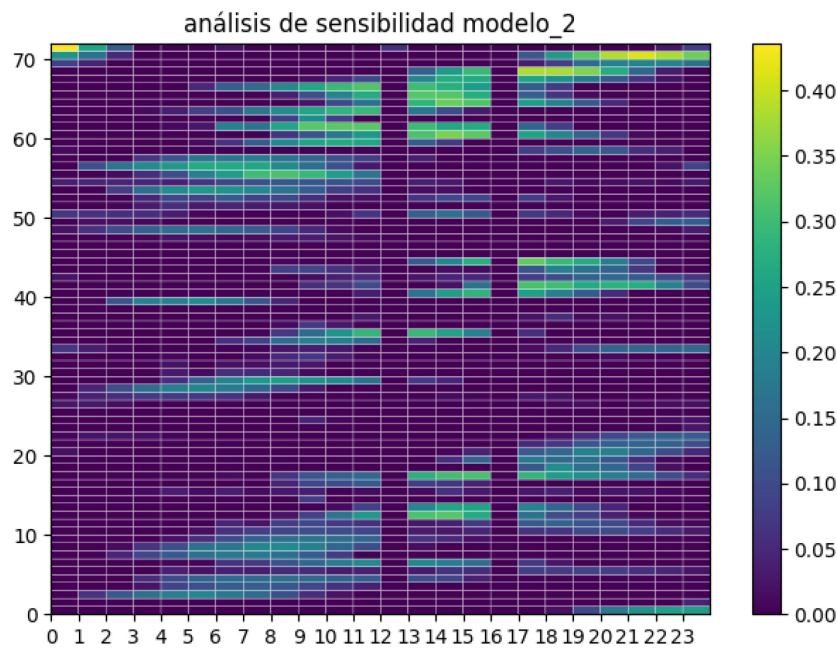


Fig 19 Análisis de sensibilidad del modelo 2

Se observa claramente que en este modelo son en concreto dos las neuronas con los pesos mal asignados que no trabajan adecuadamente: las correspondientes a las salidas 4 y 8. Se realiza una comprobación:

```

new_predictions[:,4]
<tf.Tensor: shape=(72,), dtype=float32, numpy=
array([0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0.], dtype=float32)>
new_predictions[:,8]
<tf.Tensor: shape=(72,), dtype=float32, numpy=
array([0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 0.], dtype=float32)>

```

Fig 20 Vectores con los valores de las 24 horas predichas por las neuronas 4 y 8 del modelo 2

Se puede ver que todas las salidas de las neuronas 4 y 8 son nulas, es decir se trata de las ya nombradas “neuronas muertas”, incapaces de predecir. Este mismo error se da en distintas neuronas de todos los modelos entrenados.

En la búsqueda de una solución, se opta por aplicar una semilla inicial con valor cero y reentrenar todos los modelos, obligando así a las redes a asignar los pesos a partir de ésta. Se repite el análisis de sensibilidad con todos los modelos y esta semilla para ver si se han producido mejoras. El resultado para el modelo que estaba siendo estudiado es el siguiente:

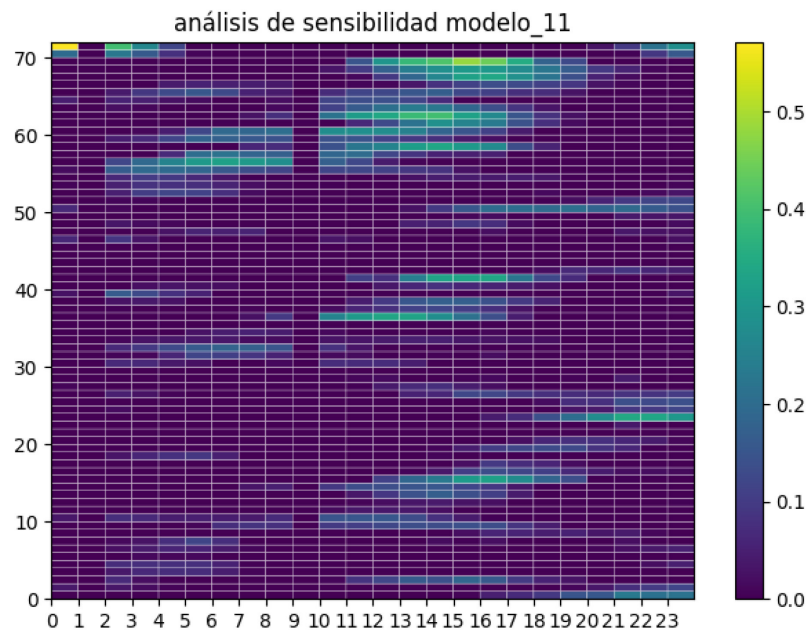


Fig 21 Análisis de sensibilidad del modelo 2 entrenado con una semilla de valor 0

Se observa que se ha producido una mejora y ahora solamente es una la neurona que no funciona por lo que se trata de un problema de asignación de los pesos en la creación de los modelos y que sigue persistiendo.

Se continua con la creación de nuevos modelos variando el parámetro de la semilla y buscando qué valor es el que ofrece un mejor resultado.

Paralelamente se procede a graficar las métricas de error de los modelos con las distintas semillas y para el caso de 60 “*epoch*”, tanto con los datos de validación como con los de entrenamiento. De esta forma se puede observar los problemas que se están dando y si el modelo se ha sobreajustado o si la métrica escogida (MSE) es o no adecuada. Así, continuando con el mismo modelo presentado antes, se obtiene:

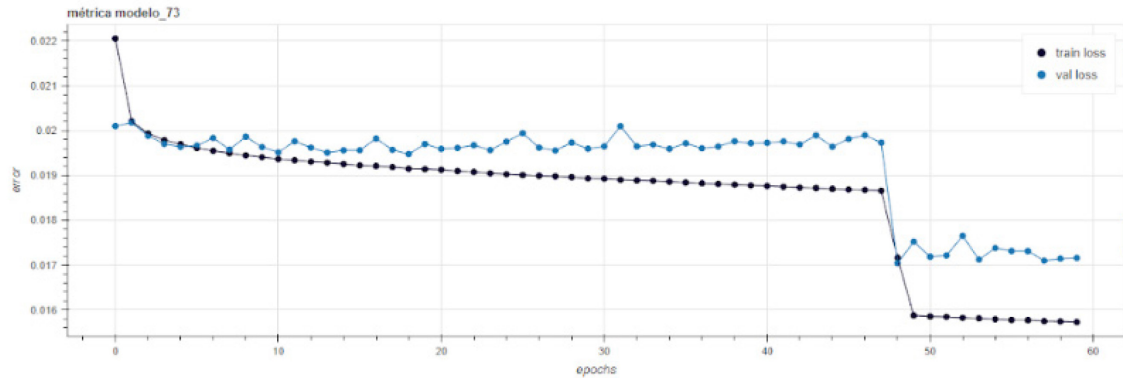


Fig 22 Métricas de error de los conjuntos de entrenamiento y validación del modelo 73, correspondiente al modelo 2 entrenado con una semilla de valor 25999

Se puede ver una bajada progresiva de los errores en la red con ambos grupos de datos hasta llegar a la época 47 donde se produce una bajada abrupta e inesperada que puede deberse a que se acaba de producir una actualización en los pesos de un mínimo local, es decir, la red creía haber alcanzado el mínimo global y al encontrar en la siguiente época uno mayor ha cambiado los pesos para que se ajuste el modelo adecuadamente.

Si se realiza esta misma gráfica en diferentes modelos, se observa el mismo fenómeno en distintas épocas, como se aprecia en estos dos ejemplos:

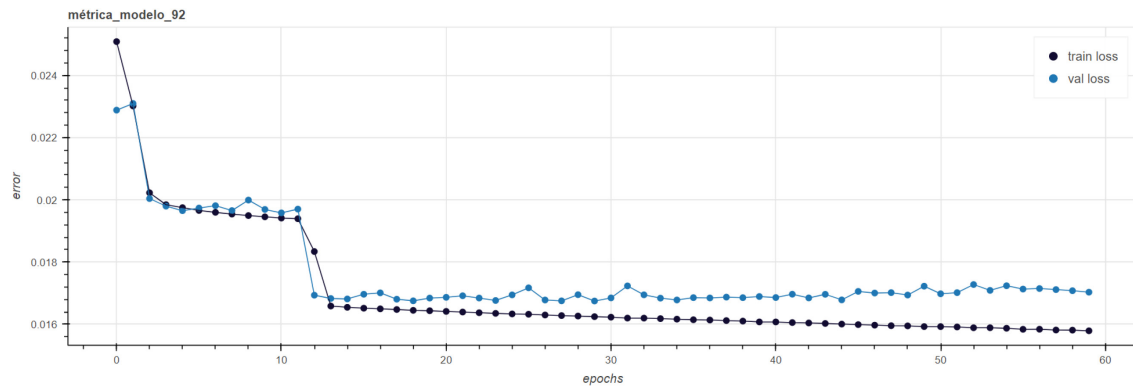


Fig 23 Métricas de error de los conjuntos de entrenamiento y validación del modelo 92 ,
correspondiente al modelo 2 entrenado con una semilla de valor 900000

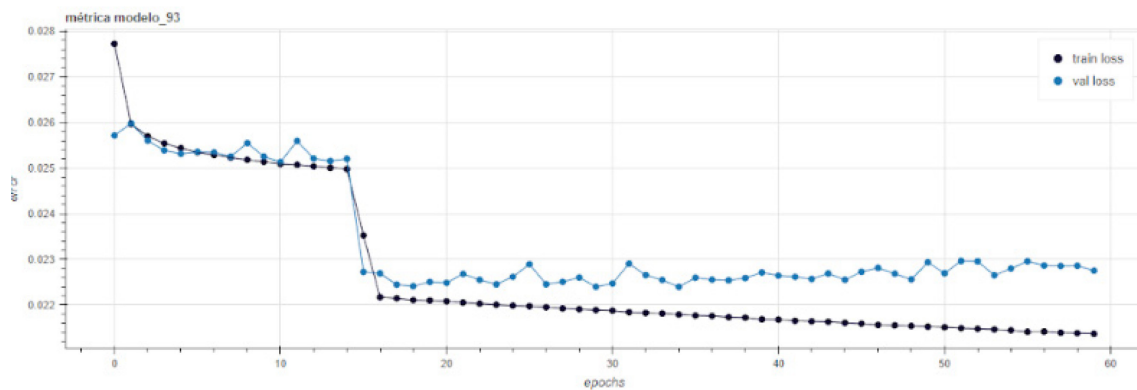


Fig 24 Métricas de error de los conjuntos de entrenamiento y validación del modelo 93,
correspondiente al modelo 3 entrenado con una semilla de valor 900000

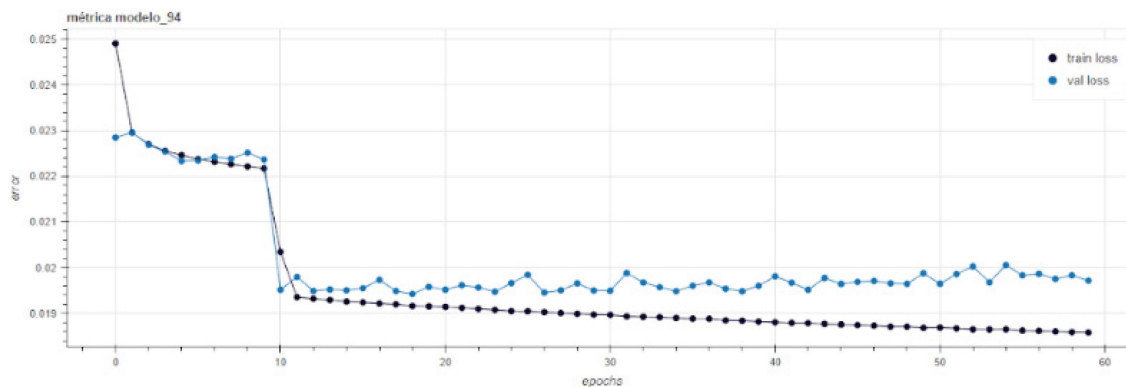


Fig 25 Métricas de error de los conjuntos de entrenamiento y validación del modelo 94,
correspondiente al modelo 4 entrenado con una semilla de valor 900000

Para comprobar si la hipótesis es correcta, y si de la época 47 a la 48 del modelo 74 se ha producido un ajuste en los pesos que ha hecho que alguna de las neuronas que previamente no funcionaban ahora sí lo haga, se generan nuevos modelos donde se guarden los pesos que se asignan a cada neurona época por época.

A continuación, se carga el modelo creado con los pesos correspondientes a cada época y se realizan los respectivos análisis de sensibilidad en todas ellas, atendiendo especialmente a aquellas donde se han producido los cambios bruscos de las anteriores

gráficas, para ver si han sucedido en el momento en que han empezado a funcionar las neuronas:

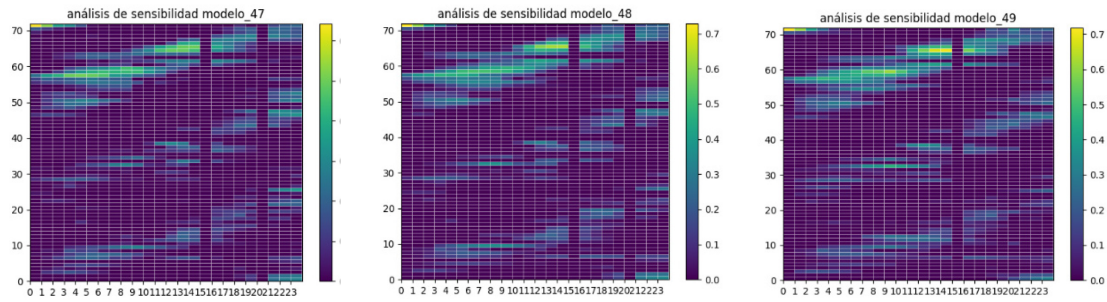
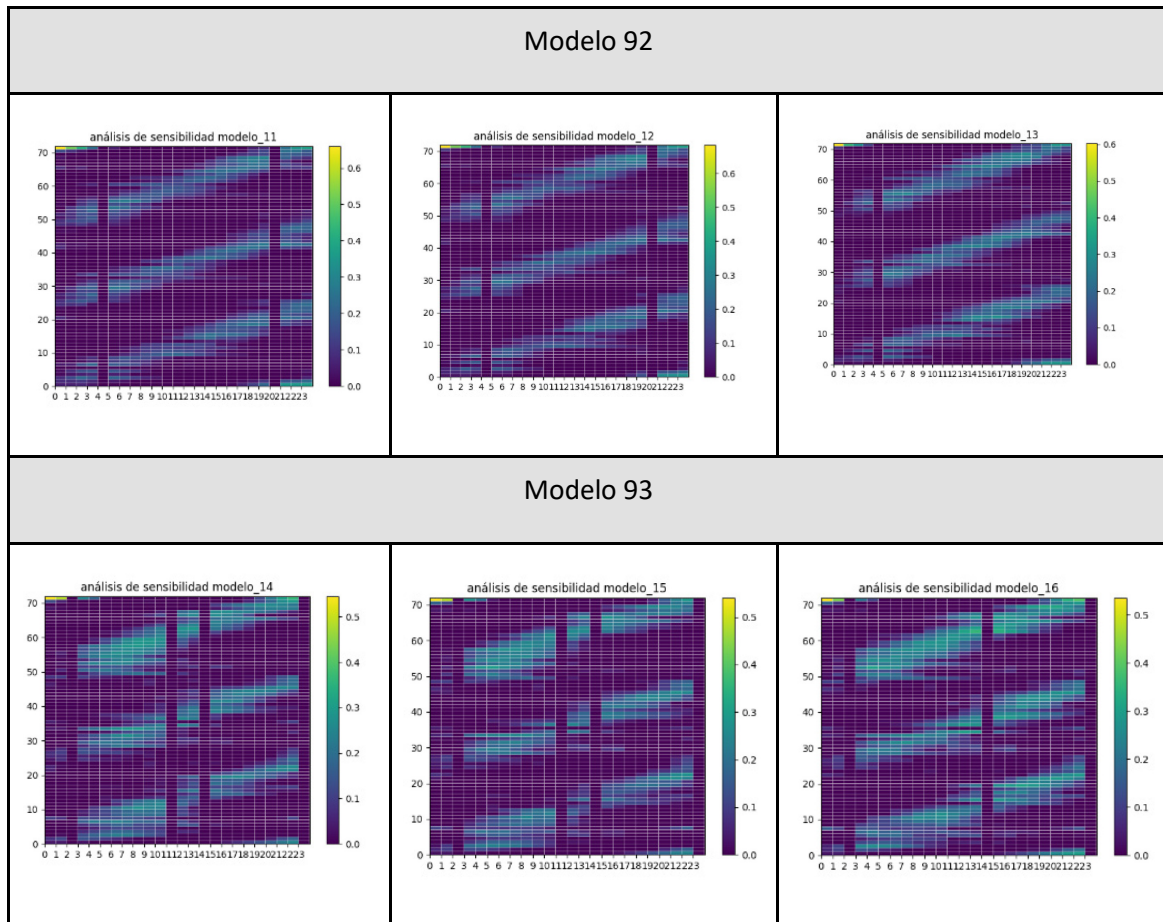


Fig 26 . *Análisis de sensibilidad de las épocas 47, 48 y 49 del modelo 2 en las que se producen la bajada brusca en la gráfica de error*

Se comprueba que en este caso la hipótesis es correcta y en las mencionadas “*epochs*” las neuronas han comenzado a funcionar.

Para asegurar que esta teoría es correcta, se realiza el mismo procedimiento en todos los modelos dónde se producía este efecto. A continuación, se muestran los análisis de sensibilidad de las épocas de los mismos modelos mostrados antes como ejemplo donde se producían los saltos en las gráficas de error:



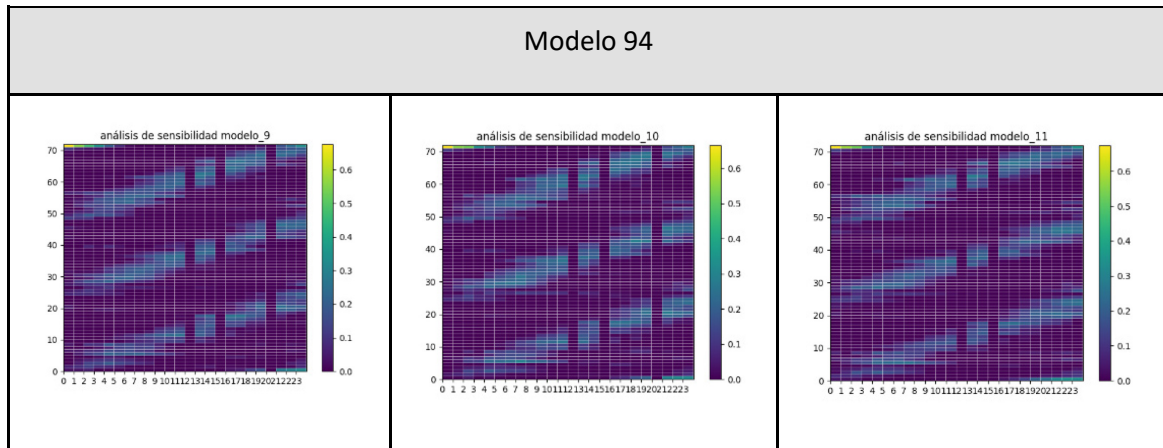


Fig 27 Análisis de sensibilidad de las distintas épocas en las que se producen bajadas bruscas en las gráficas de error de los diferentes modelos ejemplo

Se observa que en todos los casos es cierto que los cambios abruptos en las gráficas de “loss” se debían a un ajuste de los pesos que provoca que una neurona que no se activaba ahora si lo haga.

Si se repite la gráfica de interés inicial de las predicciones de la irradiancia diaria, se comprueba que incluso en los casos en los que ya no existen neuronas muertas las predicciones no se ajustan a los datos reales:

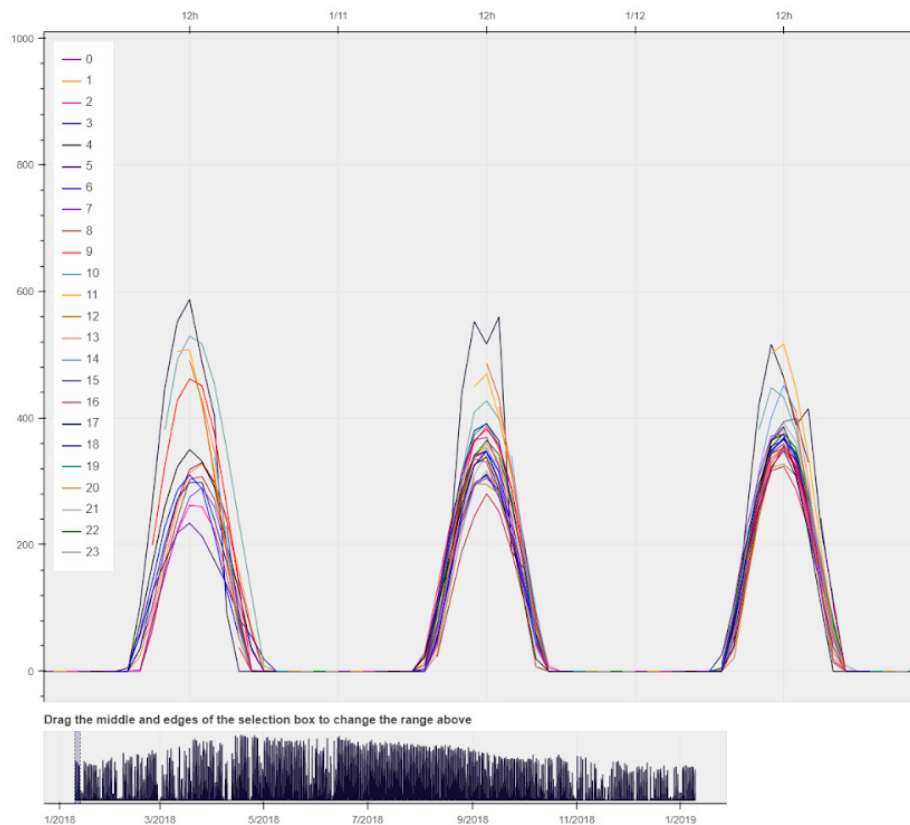


Fig 28 Predicción de la irradiancia en ventanas de 24 horas realizada por el segundo modelo entrenado con una semilla de valor 25999

Se procede a buscar una solución para resolver la nombrada mala asignación de los pesos que es la causante que las neuronas carezcan de precisión. Como se ha explicado

anteriormente las funciones de activación ReLU favorecen la aparición de neuronas muertas. Por tanto, se opta por entrenar modelos con las mismas características en cuanto a número de capas y neuronas que los anteriores y con una semilla de valor 25999, con la que se han obtenido mejores resultados, pero aplicando ahora una función de activación ELU (“*Exponential Linear Unit*”) tanto en la capa de entrada como en todas las capas ocultas. Estas funciones son una modificación de la función ReLU que incorporan una cola exponencial para suavizar la pendiente en el semieje x negativo y cuya expresión matemática y gráfica son las siguientes:

$$f(x) = \begin{cases} \alpha(e^x - 1) & x < 0 \\ x & x > 0 \end{cases} \quad (14)$$

Donde x es la entrada a la función y α es una constante positiva.

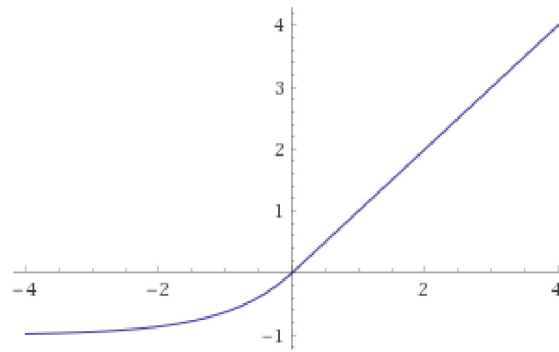


Fig 29 Gráfica de la función ELU. Fuente: Fig 16 trabajo [44]

Como se puede observar en la gráfica esta función a diferencia de la ReLU no se satura en la región negativa, lo que consigue que las neuronas tengan una salida diferente a cero incluso cuando su entrada es menor o igual a cero.

Además, en las capas de salida se decidió utilizar funciones de activación lineares de la forma:

$$f(x) = x \quad (15)$$

Donde x es la entrada a la función. A continuación, se muestra la gráfica correspondiente a dicha expresión algebraica:

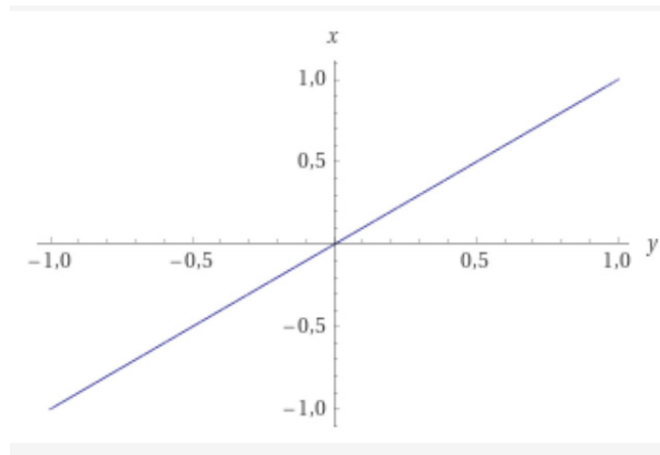


Fig 30 Fig. Gráfica de la función función de activación lineal.

Esta función tiene la ventaja de ofrecer unos resultados más fáciles de interpretar.

Por otro lado, en la búsqueda de la mejora del rendimiento de los modelos se decidió añadir el hiperparámetro “*Learning rate*” o tasa de aprendizaje, que como se expone en [41], [43], es una medida de cómo de rápido se desea que aprenda la red y suele tomar valores entre 0,1 y 0,0001 en una escala logarítmica. Para ajustarlo hay que tener en cuenta que valores muy elevados pueden provocar problemas en la convergencia para alcanzar un mínimo. Por otro lado, valores demasiado bajos harán que el algoritmo de optimización converja muy lentamente para alcanzar el mínimo de la función de coste y se necesiten muchas iteraciones. Además, pueden provocar que, dependiendo del algoritmo de optimización empleado, algunos pesos converjan a mínimos locales en lugar de globales. Así, se decide crear modelos con tasas de aprendizaje de valores de 0.1, 0.01, 0.02 y 0.0001.

Se crearon más de 200 modelos con diversas combinaciones de épocas y de los mencionados parámetros, lo que fue posible gracias a la utilización del servidor del que se disponía. Entre ellos, se escogió el siguiente modelo, que ofreció los mejores resultados y cuyas características aparecen en la Tabla 5:

Tabla 5 Información sobre la red neuronal escogida

Modelo escogido	
Nº de neuronas por capa	[70, 70, 40, 40, 30, 24]
“ <i>Batch size</i> ”	100
Nº de “ <i>epochs</i> ”	60
Función de coste	MSE
Optimizador	ADAM
Valor de la semilla	25999
“ <i>Learning rate</i> ”	0.01

Al modelo seleccionado le corresponde la siguiente gráfica de métricas, donde ya no se producen los anteriores cambios abruptos y no existe sobreentrenamiento:

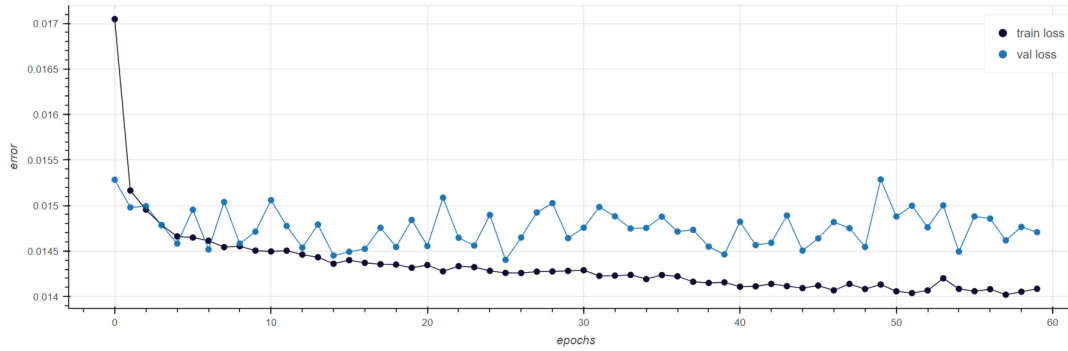


Fig 31 Gráfica de métricas de error del modelo escogido

En cuanto a las métricas de error, se grafican a continuación los errores MAE y MAPE de la red seleccionada en cada hora predicha, es decir, el eje x está expresado en horas relativas y la hora 1 implica la predicción en dos horas y así sucesivamente:

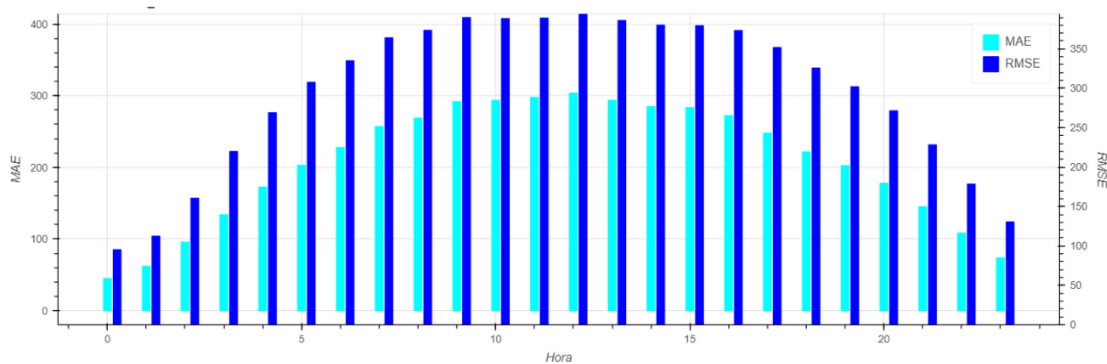


Fig 32 Error MAE y RMSE horarios del modelo seleccionado

En ella se puede observar que los errores de las primeras y las últimas horas predichas son notablemente menores que cuando se predice en las horas intermedias. Una hipótesis que se plantea es que esto se deba a que los valores cada 24 horas se asemejan más a los predichos inicialmente y a la red le resulta más fácil aproximarse al valor real.

Por otro lado, se plasman en la siguiente tabla los errores obtenidos con la red seleccionada junto con las de los otros cuatro modelos, y sus parámetros, que han demostrado dar los mejores resultados después de la selección:

Tabla 6 Métricas de los modelos con los que se ha obtenido mejores resultados junto con sus características

Modelo escogido	
MAE	205.9 W/m ²
RMSE	120.20 W/m ²
Modelo 110	
Nº de neuronas por capa	[70, 60, 35, 24]
"Batch size"	100
Nº de "epochs"	60
Función de coste	MSE

Optimizador	ADAM
Valor de la semilla	25999
“ <i>Learning rate</i> ”	0.01
MAE	205.15 W/m ²
<i>RMSE</i> ”	291.77 W/m ²
Modelo 146	
Nº de neuronas por capa	[70, 60, 35, 24]
“ <i>Batch size</i> ”	100
Nº de “ <i>epochs</i> ”	10000
Función de coste	MSE
Optimizador	ADAM
Valor de la semilla	25999
“ <i>Learning rate</i> ”	0.01
MAE	209.39 W/m ²
<i>RMSE</i> ”	297.01 W/m ²
Modelo 148	
Nº de neuronas por capa	[70, 70, 60, 50, 40, 24]
“ <i>Batch size</i> ”	100
Nº de “ <i>epochs</i> ”	10000
Función de coste	MSE
Optimizador	ADAM
Valor de la semilla	25999
“ <i>Learning rate</i> ”	0.01
MAE	206.73 W/m ²
<i>RMSE</i> ”	296.55 W/m ²
Modelo 147	
Nº de neuronas por capa	[70, 50, 40, 24]
“ <i>Batch size</i> ”	100
Nº de “ <i>epochs</i> ”	10000
Función de coste	MSE
Optimizador	ADAM
Valor de la semilla	25999
“ <i>Learning rate</i> ”	0.01
MAE	207.07 W/m ²
<i>RMSE</i> ”	299.57 W/m ²

Por último, se presentan a continuación algunos días como muestra de la predicción hecha por esta red junto a los datos reales para observar los resultados de manera cualitativa:

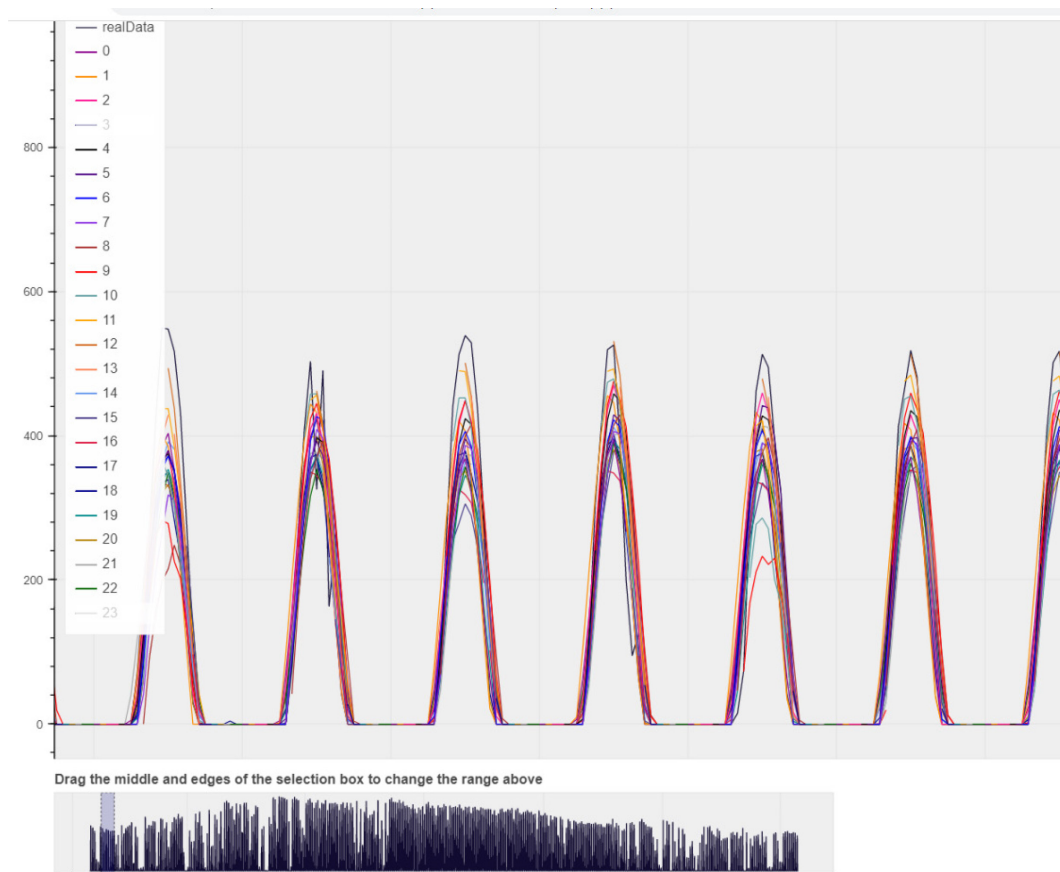


Fig 33 Predicciones realizadas por el modelo escogido

4.2 Modelo de consumo de la instalación

A continuación, se expone el proceso realizado para el diseño de nuestro modelo ARIMA

4.2.1 Preprocesamiento de los datos

Como se explicó en el apartado 3.3.1 fue necesario crear una nueva columna en el CSV dado que contuviera los valores de energía total consumida medidos en kW/h. Estos datos fueron agrupados con frecuencia horaria igual a la del modelo de producción mediante la función “*sum()*” del paquete “*Numpy*” de Python, por lo que se obtuvieron los datos en potencia (kW).

Además, antes de crear el modelo, se dividió el conjunto de datos en dos bloques: el primero, de entrenamiento o “*train*” se utilizó para crear el modelo y el segundo de test fue necesario para comprobar si las predicciones estimadas por el modelo con los datos de “*train*” son precisas o no. Se eligieron como datos de entrenamiento los referentes a los años 2006 y 2008 y los del 2009 como test.

Por otro lado, en el apartado 2.2.2 se ha explicado la necesidad de que la serie de datos sea estacionaria en los modelos ARIMA, por lo que antes de comenzar con la creación del modelo se comprueba si lo es. En caso contrario, se deberá conseguir que lo sea (mediante, por ejemplo, transformaciones logarítmicas o exponenciales). Para realizar

dicha comprobación gráficamente se toma un día ejemplo de la base de datos, seleccionando los valores de consumo del año 2007 y calculando la media entre ellos y los valores máximos y mínimos:

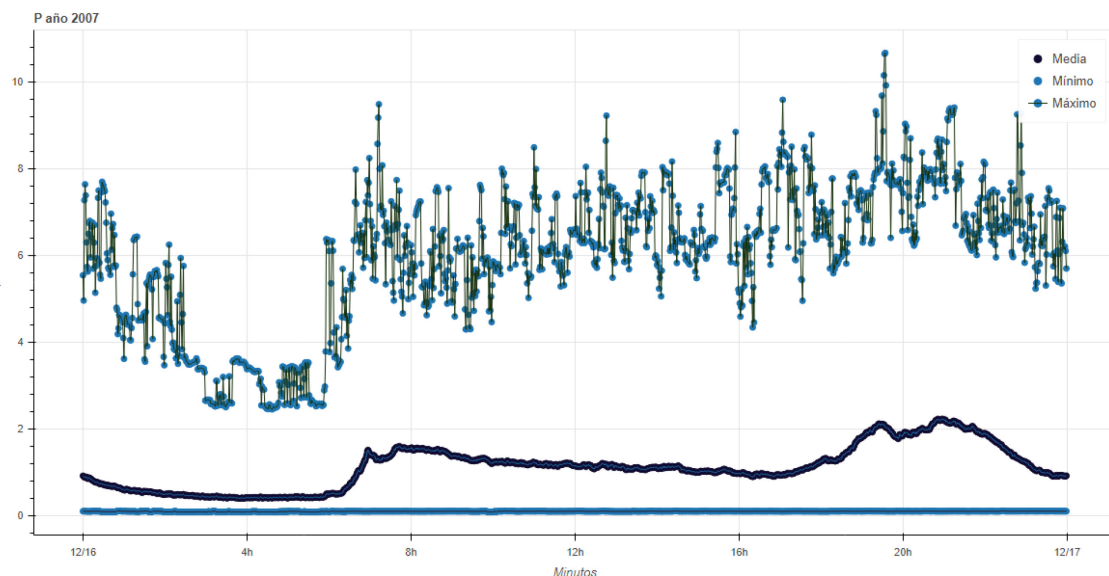


Fig 34 Gráfica de los valores máximos, mínimos y medios de consumo durante el año 2007

Se observa que no existe ninguna tendencia a largo plazo en los datos ni una gran dispersión entre ellos y que la media de la serie fluctúa en torno a una media estable a lo largo del tiempo, por lo que podemos concluir que nuestra serie temporal es estacionaria.

Por otro lado, es importante observar si existe estacionalidad o no en la serie temporal de datos para decidir si en lugar de elegir un modelo ARIMA se obtendrán resultados más precisos escogiendo alguna de sus variantes. Para ello, realizamos una descomposición de la serie utilizando la función “*seasonal_descompose()*” del paquete de “*statsmodels*” de Python. Como resultado se muestran cuatro gráficas: la serie original, una serie temporal con la tendencia de la misma, otra con su estacionalidad y finalmente una con los residuos de la serie original.

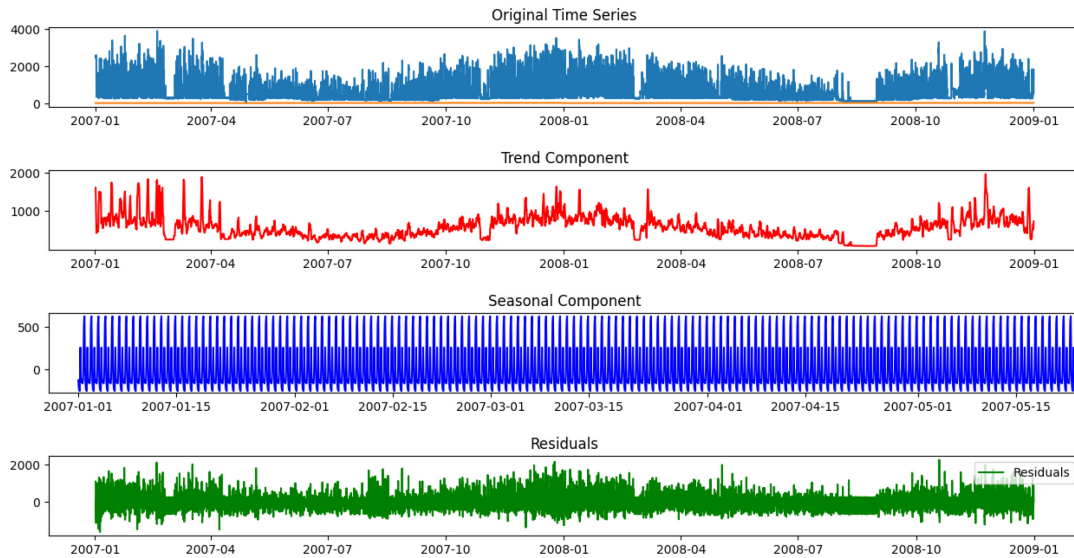


Fig 35 Gráficas resultado de la descomposición de la serie temporal mediante la función `seasonal_decompose()`

De esta manera, atendiendo a la tendencia, representada en color rojo, se aprecia que se mantiene entorno a los mismos valores constantes a lo largo del tiempo y se reafirma que la serie original es estacionaria. Observando la tercera gráfica obtenida, se ve que existe claramente una estacionalidad puesto que se produce una variación repetitiva en intervalos de tiempo regulares. Para averiguar el periodo de la estacionalidad (cada cuánto aparece el mismo patrón), se amplía la gráfica:

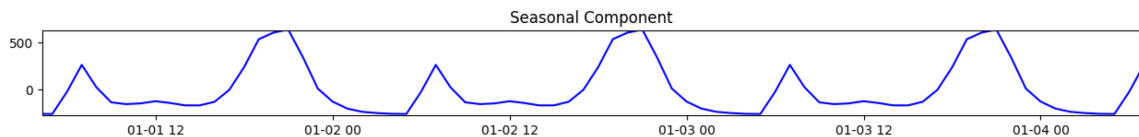


Fig 36 Gráfica con el componente estacional de la serie ampliado

Se puede comprobar que el periodo de repetición del mismo patrón es de 24 horas, es decir, diariamente se consume más energía durante las mismas horas punta del día y menos durante los mismos momentos de la noche.

Con esta información se decide implementar un modelo SARIMAX ya que es más completo que un modelo SARIMA y el modelo ARIMA no recoge la estacionalidad de los datos.

4.2.2 Diseño del modelo y elección de sus parámetros

Previamente se ha visto que para el diseño de los modelos SARIMAX es necesario determinar los parámetros p , d , q , P , D , Q y s . En cuanto al parámetro s correspondiente al periodo de estacionalidad se puede concluir, observando las gráficas anteriores, que tiene un valor de 24.

El parámetro d corresponde al número de diferenciaciones realizadas para hacer la serie estacionaria, y como se ha podido observar la serie de la que se partía ya lo era por lo

que tendrá un valor de cero. No obstante, para confirmarlo se realiza complementariamente el test ADF (*Augmented Dickey-Fuller*) o prueba de raíz unitaria, que como se explica en [45]. Es una prueba de hipótesis, donde la hipótesis nula es que la serie temporal es no estacionaria (tiene una raíz unitaria), y la hipótesis alternativa es que la serie temporal es estacionaria (no tiene una raíz unitaria). Para descartar o no la hipótesis se calcula un valor-p y si este es menor que el nivel de significancia establecido (normalmente 0.05), se rechaza la hipótesis nula y se concluye que la serie temporal es estacionaria. Esta prueba se lleva a cabo mediante la función *adfuller()* de la librería *statsmodels* de Python, obteniendo un valor $p\text{-value} = 7.58e-08$, muy por debajo de 0.05, demostrando así que la serie de datos es estacionaria.

En cuanto al parámetro q , como se puede consultar en [44], se puede determinar mediante la gráfica de autocorrelación o ACF, que representa la influencia que tienen las observaciones pasadas en la actual. Los valores con mayor correlación serán los que más contribuyan y con ellos será suficiente para predecir. Por otro lado, la zona azul es la de significación y muestra qué coeficientes son distintos de cero (fuera del área colorida) y cuáles no (dentro del área). Así, en el caso estudiado se observa un gran pico en el primer desfase y por tanto se puede establecer el parámetro q con valor 1.

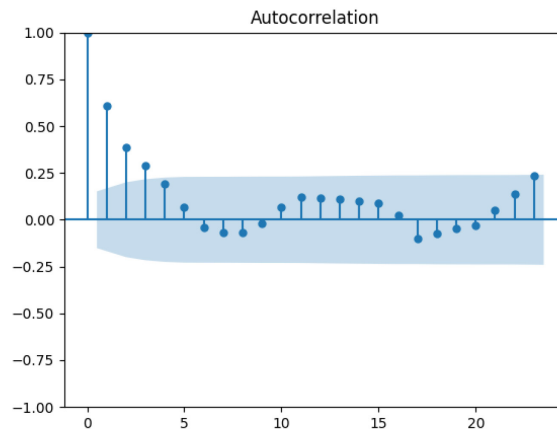


Fig 37 Gráfica de autocorrelación de la serie temporal de estudio

El siguiente paso es encontrar el orden del término AR (p), analizando para ello la gráfica de correlación parcial (PACF), como se explica en el trabajo citado anteriormente, que mide la correlación entre la serie temporal y un retardo o *“lag”* concreto. Se aprecia que la correlación más significativa también se da con el primer retardo, y por tanto se puede definir el parámetro p con valor 1.

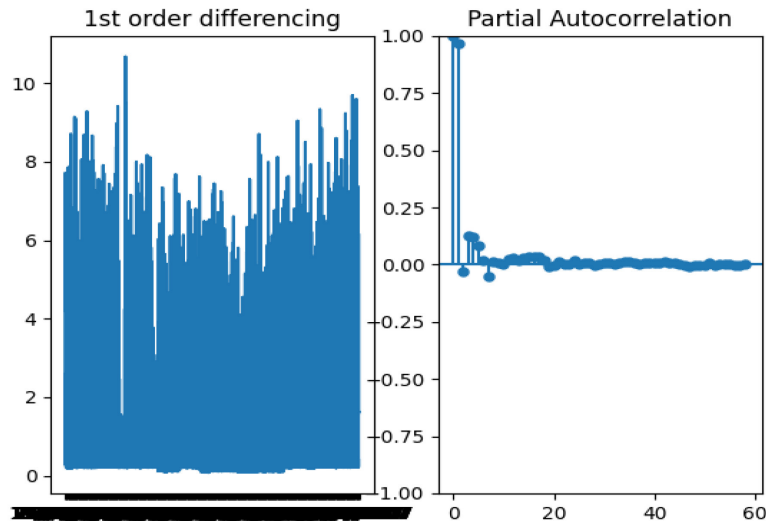


Fig 38 Gráfica de autocorrelación parcial de la serie temporal de estudio

Por tanto, el modelo final de ARIMA queda definido como $(p=1, d=0, q=1)$. En cuanto a los componentes estacionales P, D y Q, se utilizó la función “*auto.arima()*” del paquete “*pmdarima*” de Python, obteniendo valores de 1, 0 y 1 respectivamente. Finalmente, se añadió una columna con los meses del año como variable exógena y se crearon este y otros modelos.

4.2.3 Resultados obtenidos

Tras el diseño de los modelos se obtuvieron con ellos las potencias futuras, expresadas en kW, de varios días en ventanas de 24 horas, y se calcularon las métricas de error de las predicciones realizadas. Tras un barrido paramétrico se seleccionó el primero de los modelos, cuya elección de parámetros ha sido descrita en 4.2.2 y cuyas métricas son las mostradas en la Tabla 7:

Tabla 7 Métricas de error del modelo SARIMAX implementado

Métricas de error del modelo	
MAE	0.216 kW
RMSE	0.317 kW
MAPE	24.6%

Así, atendiendo al valor del error MAE, significa que en promedio las predicciones estarán 200 W por encima o por debajo de los valores reales de consumo de la instalación, con un valor nominal de 2 kW. Por tanto, se puede concluir que el modelo realizará predicciones suficientemente precisas.

Una vez elegido el modelo, se procedió a graficar los resultados. Así, en el caso de once días predichos con respecto al último de los valores del conjunto de entrenamiento la figura obtenida fue la siguiente:

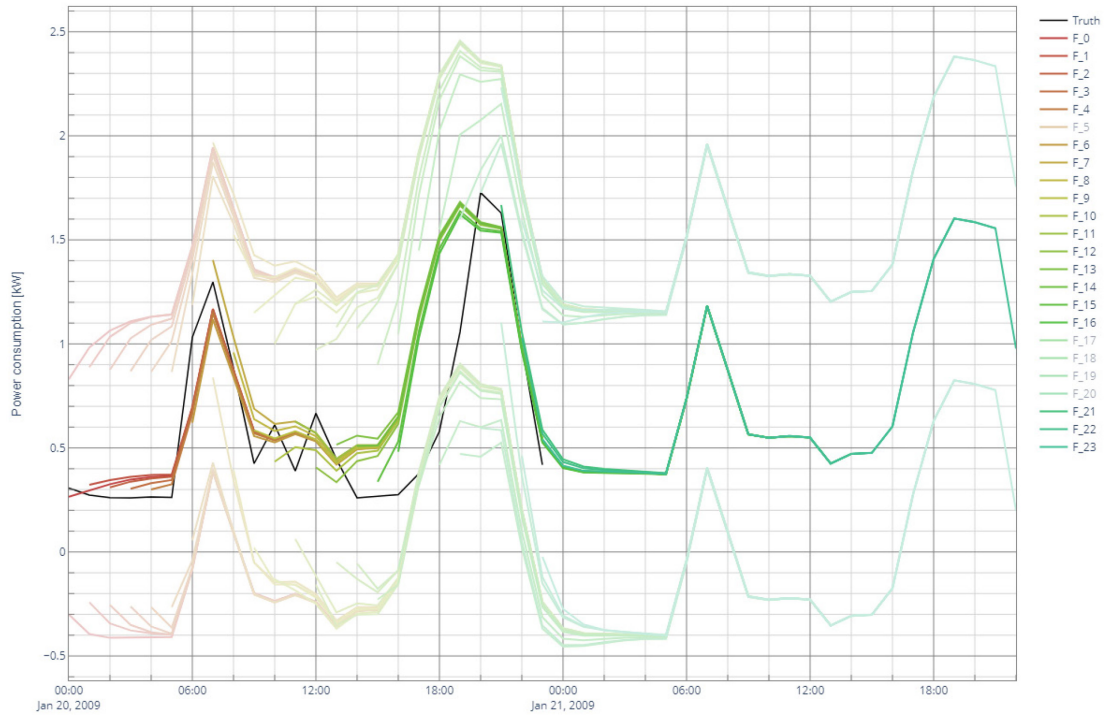


Fig 39 Predicción de la potencia en ventanas de 24 horas realizada por el modelo SARIMAX

En ella cada una de las líneas coloridas representa la predicción de las 24 horas futuras respecto a una de las horas del día del conjunto de test, al igual que se hacía en el caso de la producción, mientras que la línea más oscura contiene los valores reales procedentes de los datos de test. Así mismo, las líneas sombreadas representan el margen del intervalo de confianza del 95% con el que se ha graficado.

Se puede observar además que al comienzo el modelo es más sensible a las entradas y existe una mayor variación entre las distintas predicciones mientras que a largo plazo los valores predichos se asemejan más los unos a los otros, lo que puede deberse a que el modelo encuentra algún patrón estadístico repetitivo.

4.3 Sistema de gestión de energía (EMS)

Una vez obtenidas las predicciones tanto del modelo de predicción como de consumo, se introducen los resultados en el modelo MPC. Junto con estas, se añaden además los valores reales para poder obtener las potencias, tanto activas como reactivas, que han de establecerse. Como el modelo trabaja con potencias, se requiere introducir los valores en kW y en el caso del modelo de producción habrá que transformarlos ya que inicialmente se trata de irradiancia expresada en W/m^2 . Para ello se multiplican los valores por el área de estudio provista con placas solares ($200 m^2$), el rendimiento de los paneles indicado por el fabricante (21,2%) y se divide entre mil.

Además, se introducen los datos característicos de la batería de segunda vida que se encuentra en el centro en el que se centra este trabajo que se plasman en:

Tabla 8 Características de la batería de estudio

Características de la batería

SOC máximo	95%
SOC mínimo	15%
Capacidad nominal	20 kWh
Potencia máxima	6 kW

De esta manera, al utilizar el modelo MPC y representar los resultados con las predicciones de las próximas 24 horas se obtiene la siguiente figura:



Fig 40 Resultados obtenidos tras implementar el modelo MPC

En la gráfica superior se representa el balance de potencia consumida y producida tanto con los datos reales como con los predichos, es decir la diferencia entre el consumo y la generación en la instalación. Simultáneamente, se observa el SOC de la batería la gráfica que está en medio (*"Battery SOC"*). Se puede apreciar que, durante los momentos en los que no hay generación de energía la batería se descarga mientras que conforme avanza el día y comienza la producción procede a cargarse. Al terminar el día de nuevo comienza a descargarse ya que ya no incide luz sobre los paneles y con ello cesa la producción

Por otro lado, en la gráfica titulada *"Powers"*, quedan representadas la energía almacenada en la batería en cada instante, la potencia que se vierte o se obtiene de la red eléctrica y la potencia sobrante, que corresponden a las líneas de *"Battery_Power,"*

“Grid_Power” y “Cur_Power”, respectivamente. Se aprecia que durante el día la potencia sobrante es muy elevada, lo que se debe a que el consumo es mucho menor que la producción durante el día (momentos en los que se produce la carga de la batería). Además, con los balances realizados por el modelo, se obtiene los siguientes valores numéricos en unidades de energía que confirman lo mostrado en las figuras:

Tabla 9 Energías obtenidas en 24h tras los balances realizados por el modelo MPC

Resultados obtenidos modelo MPC	
Energía total sobrante	51.21 kWh
Energía total comprada	7.32 kWh
Energía total vendida	22.89 kWh

Tras varios resultados se observa una diferencia muy significativa entre la producción y el consumo estudiados. Existe una gran cantidad de energía sobrante. Esto se debe a que no ha sido posible obtener datos suficientes de consumo procedentes del CSIC y se ha recurrido a una base de datos de consumo de una vivienda, con consumos mucho más pequeños. Por tanto, en este caso de aplicación, no se exprime toda la capacidad de optimización del MPC. Para conseguir resultados donde las reglas del modelo se apliquen, se decide redimensionar el tamaño del acumulador. Tras una revisión de las baterías de segunda vida de uso doméstico existentes se selecciona una batería como la descrita en [45] cuya capacidad y potencia máximas se adaptan más a los datos disponibles. Así, se introducen los siguientes datos en el modelo MPC:

Tabla 10 Características de la batería de estudio tras el redimensionamiento

Características de la batería	
SOC máximo	95%
SOC mínimo	15%
Capacidad nominal	6 kWh
Potencia máxima	3 kW

A su vez, se reducen las dimensiones de la instalación fotovoltaica hasta 100 m², ya que las iniciales son demasiado elevadas para el consumo introducido y provoca que siempre haya un exceso de energía generada. Una vez redimensionada la instalación se repiten las simulaciones.

A continuación, la Fig 41 muestra un ejemplo de un día en que la batería ha saturado:

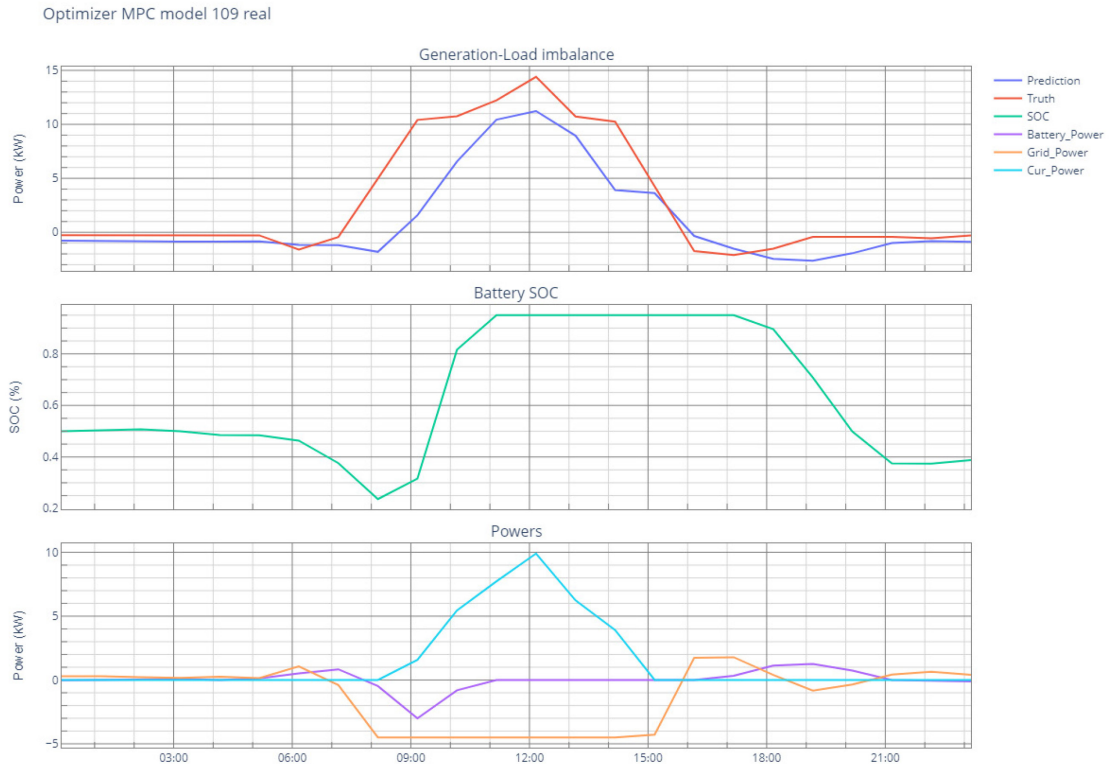


Fig 41 Resultados obtenidos con el modelo MPC en un día en el que la batería satura

Se observa en la gráfica cómo la batería inicialmente comienza a descargarse mientras no existe generación de energía. Sin embargo, una vez comienza la producción, el dispositivo ha alcanzado su capacidad máxima entorno a las 11 de la mañana. A partir de ese punto ya no acepta más carga, por tanto, toda la energía que se continúe generando deberá ser vertida a la red eléctrica, como corroboran los resultados numéricos obtenidos con el modelo que aparecen en la Tabla 11. Por otro lado, al final del día cuando termina la producción se empieza a hacer uso de la batería y esta comienza a descargarse, priorizando el autoconsumo.

Tabla 11 Energías obtenidas tras los balances realizados por el modelo MPC en un día en que la batería satura

Resultados obtenidos modelo MPC	
Energía total sobrante	34,81 kWh
Energía total comprada	7.82 kWh
Energía total vendida	37.36 kWh

5. Conclusiones y trabajo futuro

En el presente trabajo se han alcanzado con éxito los objetivos planteados. Se ha creado un modelo basado en redes neuronales que predice de manera anticipada la producción de una planta fotovoltaica. Complementariamente, se ha diseñado un modelo estadístico SARIMAX para la estimación de consumos. Ambos modelos han sido optimizados definiendo varias métricas de error que los evalúan. Además, tras una revisión bibliográfica, se ha seleccionado como optimizador del sistema de gestión de energía un Modelo de Control Predictivo, MPC. Combinándolo con los resultados obtenidos con los modelos mencionados, se ha realizado una gestión eficiente del flujo de energía haciendo uso de la batería.

Durante el desarrollo de los modelos basados en redes neuronales, se ha comprobado la importancia de la selección de hiperparámetros. Se han estudiado especialmente los parámetros de “*learning rate*”, “*epoch*” y función de activación. Para este último, se ha evaluado el uso de funciones de activación novedosas, como la “*leaky ReLU*”, con la que se han obtenido mejores resultados.

Tras el análisis exploratorio de los datos de consumo seleccionados, se ha seleccionado el uso del modelo de predicción estadístico SARIMAX, dada la estacionalidad de las muestras. Cabe destacar la influencia de los datos históricos en las predicciones a corto plazo; sin embargo, las predicciones más lejanas convergen al patrón medio de consumo. Además, el modelo final obtenido tiene capacidad de predicción de consumo con un error promedio de 200 W.

Para la optimización de flujos de energía mediante MPC se han definido distintas condiciones de contorno, priorizando el autoconsumo de la instalación. Se ha validado el funcionamiento del optimizador estudiado distintas situaciones: días con perfiles de irradiación solar alta, nublados y distintos perfiles de consumo.

Dentro de las líneas de trabajo futuro, se plantea la validación experimental de los algoritmos, dada la imposibilidad de la utilización de las instalaciones del CSIC durante el desarrollo del trabajo. Se propone también utilizar un modelo de batería más complejo, ya que no era el foco principal del trabajo.

Otra línea de trabajo futuro que ya se está explorando en el grupo de investigación es el aumento del número y complejidad de las condiciones impuestas en el modelo MPC. Esto podría permitir realizar cálculos de costes económicos, amortización de la instalación y el correcto dimensionado de instalaciones solares con almacenamiento.

6. Bibliografía

- [1] “pablop / TFG Sara · GitLab.” <https://gitlab.com/pastorpflores/tfg-sara> (accessed May 12, 2023).
- [2] C. Aarthi Vigneshwari, S. Siva Sakthi Velan, M. Venkateshwaran, M. Adam Mydeen, and V. Kirubakaran, “Performance and economic study of on-grid and off-grid solar photovoltaic system,” *2016 International Conference on Energy Efficient Technologies for Sustainability, ICEETS 2016*, pp. 239–244, Oct. 2016, doi: 10.1109/ICEETS.2016.7582933.
- [3] A. Ghafoor and A. Munir, “Design and economics analysis of an off-grid PV system for household electrification,” 2014, doi: 10.1016/j.rser.2014.10.012.
- [4] O. P. Mahela and A. G. Shaik, “Comprehensive overview of grid interfaced solar photovoltaic systems,” *Renewable and Sustainable Energy Reviews*, vol. 68, pp. 316–332, Feb. 2017, doi: 10.1016/J.RSER.2016.09.096.
- [5] “IEEE Xplore Full-Text PDF:” <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8442403&tag=1> (accessed Oct. 19, 2022).
- [6] Y. Zhang, A. Lundblad, P. E. Campana, and J. Yan, “Employing battery storage to increase photovoltaic self-sufficiency in a residential building of Sweden,” *Energy Procedia*, vol. 88, pp. 455–461, Jun. 2016, doi: 10.1016/J.EGYPRO.2016.06.025.
- [7] R. Chenni, M. Makhoulf, T. Kerbach, and A. Bouzid, “A detailed modeling method for photovoltaic cells,” *Energy*, vol. 32, pp. 1724–1730, 2007, doi: 10.1016/j.energy.2006.12.006.
- [8] A. Ndiaye, A. Charki, A. Kobi, C. M. F. Kébé, P. A. Ndiaye, and V. Sambou, “Degradations of silicon photovoltaic modules: A literature review,” *Solar Energy*, vol. 96, pp. 140–151, 2013, doi: 10.1016/J.SOLENER.2013.07.005.
- [9] V. Muteri *et al.*, “Review on life cycle assessment of solar photovoltaic panels,” *Energies (Basel)*, vol. 13, no. 1, Jan. 2020, doi: 10.3390/EN13010252.
- [10] J. Pastuszak and P. Węgierek, “Photovoltaic Cell Generations and Current Research Directions for Their Development,” *Materials 2022, Vol. 15, Page 5542*, vol. 15, no. 16, p. 5542, Aug. 2022, doi: 10.3390/MA15165542.

- [11] V. J. Chin, Z. Salam, and K. Ishaque, "Cell modelling and model parameters estimation techniques for photovoltaic simulator application: A review," *Appl Energy*, vol. 154, pp. 500–519, Sep. 2015, doi: 10.1016/J.APENERGY.2015.05.035.
- [12] A. R. Jordehi, "Parameter estimation of solar photovoltaic (PV) cells: A review," *Renewable and Sustainable Energy Reviews*, vol. 61, pp. 354–371, Aug. 2016, doi: 10.1016/J.RSER.2016.03.049.
- [13] R. Gardashov, M. Eminov, G. Kara, E. G. Emecen Kara, T. Mammadov, and X. Huseynova, "The optimum daily direction of solar panels in the highlands, derived by an analytical method," *Renewable and Sustainable Energy Reviews*, vol. 120, p. 109668, Mar. 2020, doi: 10.1016/J.RSER.2019.109668.
- [14] O. P. Lamigueiro, "Energía Solar Fotovoltaica", Accessed: Oct. 19, 2022. [Online]. Available: <https://www.researchgate.net/publication/249012821>
- [15] C. Rahmann, B. Mac-Clure, V. Vittal, and F. Valencia, "Break-Even Points of Battery Energy Storage Systems for Peak Shaving Applications," *Energies 2017, Vol. 10, Page 833*, vol. 10, no. 7, p. 833, Jun. 2017, doi: 10.3390/EN10070833.
- [16] P. A. Nelson, S. Ahmed, K. G. Gallagher, and D. W. Dees, "Modeling the Performance and Cost of Lithium-Ion Batteries for Electric-Drive Vehicles, Third Edition," Mar. 2019, doi: 10.2172/1503280.
- [17] M. Ceraolo, "New dynamical models of lead-acid batteries," *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1184–1190, 2000, doi: 10.1109/59.898088.
- [18] P. V Braun, J. Cho, J. H. Pikul, W. P. King, and H. Zhang, "High power rechargeable batteries," 2012, doi: 10.1016/j.cossms.2012.05.002.
- [19] B. S. Bhangu, P. Bentley, D. A. Stone, and C. M. Bingham, "State-of-charge and state-of-health prediction of lead-acid batteries for hybrid electric vehicles using non-linear observers," *2005 European Conference on Power Electronics and Applications*, vol. 2005, 2005, doi: 10.1109/EPE.2005.219601.
- [20] W. Pan, T. Xu, Q. Chen, and M. Zhu, "An integration and selection scheme for capacity estimation of Li-ion battery based on different state-of-charge intervals," *J Energy Storage*, vol. 53, p. 105073, 2022, doi: 10.1016/j.est.2022.105073.

- [21] C. De Electricidad Caracterización, S. U. Carga, and Y. Descarga, “UNIVERSIDAD POLITÉCNICA SALESIANA SEDE QUITO”.
- [22] T. M. I. Mahlia, T. J. Saktisahdan, A. Jannifar, M. H. Hasan, and H. S. C. Matseelar, “A review of available methods and development on energy storage; technology update,” *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 532–545, May 2014, doi: 10.1016/J.RSER.2014.01.068.
- [23] Á. Ibáñez Casao, P. Pastor-Flores, and F. Barreras, “Jornada de Jóvenes Investigadores del I3A,” *Jornada de Jóvenes Investigadores del I3A*, vol. 10, p. 100060, Jul. 2012, doi: 10.26754/JJII3A.20227132.
- [24] A. Jokar, B. Rajabloo, M. D. Esilets, and M. Lacroix, “Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries,” 2016, doi: 10.1016/j.jpowsour.2016.07.036.
- [25] L. Alfredo Fernandez-Jimenez and E. Zorzano Alba, “MODELOS DE PREDICCIÓN A CORTO PLAZO PARA SISTEMAS FOTOVOLTAICOS,” 2015, Accessed: Mar. 27, 2023. [Online]. Available: <https://www.researchgate.net/publication/264870223>
- [26] C. Voyant *et al.*, “Machine learning methods for solar radiation forecasting: A review,” *Renew Energy*, vol. 105, pp. 569–582, May 2017, doi: 10.1016/J.RENENE.2016.12.095.
- [27] T. -González *et al.*, “ESTIMACIÓN DE LA PRODUCCIÓN ELÉCTRICA DE UNA PLANTA FOTOVOLTAICA UTILIZANDO MODELOS ESTADÍSTICOS Y REDES NEURONALES ARTIFICIALES: COMPARACIÓN DE LOS RESULTADOS”.
- [28] S. Theocharides, G. Makrides, A. Kyprianou, V. Venizelou, P. Kaimakis, and G. E. Georghiou, “PV PRODUCTION FORECASTING MODEL BASED ON ARTIFICIAL NEURAL NETWORKS (ANN) Smart Control Systems for Energy Management: New Master Degree View project Erasmus+ Project MOST-Advanced Master’s Education based on Smart Grid Technology (smartgridsmaster.eu) View project PV PRODUCTION FORECASTING MODEL BASED ON ARTIFICIAL NEURAL NETWORKS (ANN),” 2017, Accessed: Mar. 29, 2023. [Online]. Available: <https://www.researchgate.net/publication/320166784>
- [29] L. Alfredo Fernández Jiménez and C. Académico, “Modelos de predicción a corto plazo de la generación eléctrica en instalaciones fotovoltaicas,” *Modelos de predicción a corto plazo de la generación eléctrica en instalaciones fotovoltaicas*, 2014.
- [30] T. -González *et al.*, “ESTIMACIÓN DE LA PRODUCCIÓN ELÉCTRICA DE UNA PLANTA FOTOVOLTAICA UTILIZANDO

MODELOS ESTADÍSTICOS Y REDES NEURONALES ARTIFICIALES: COMPARACIÓN DE LOS RESULTADOS”.

- [31] C. Hernández, F. Rodríguez, J. C. Moreno, P. R. da Costa Mendes, and J. E. Normey-Rico, “Modelado y predicción a corto plazo del consumo y producción de energía eléctrica en una micro-red utilizando métodos basados en series temporales y redes neuronales artificiales,” pp. 1010–1017, Feb. 2022, doi: 10.17979/SPUDC.9788497498081.1010.
- [32] D. Alberg and · Mark Last, “Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms,” *Vietnam Journal of Computer Science*, vol. 5, pp. 241–249, 2018, doi: 10.1007/s40595-018-0119-7.
- [33] J. Arias Martín, “Predicción de consumos eléctricos en Inglaterra, Gales y Escocia a través de datos de medidores inteligentes,” Jun. 2019, Accessed: Mar. 31, 2023. [Online]. Available: <https://openaccess.uoc.edu/handle/10609/102266>
- [34] E. Pérez, H. Beltrán, C. Ariño, and J. C. Alfonso-Gil, “Optimización de la vida de baterías en plantas fotovoltaicas mediante algoritmos genéticos,” pp. 949–956, Feb. 2022, doi: 10.17979/SPUDC.9788497498081.0949.
- [35] Y. Riffonneau, S. Bacha, F. Barruel, and S. Ploix, “Optimal power flow management for grid connected PV systems with batteries,” *IEEE Trans Sustain Energy*, vol. 2, no. 3, pp. 309–320, Jul. 2011, doi: 10.1109/TSTE.2011.2114901.
- [36] B. Liu, Z. Lu, K. Yao, and F. Gao, “A MPC Operation Method for A Photovoltaic System with Batteries,” *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 807–812, Jan. 2015, doi: 10.1016/J.IFACOL.2015.09.068.
- [37] D. L. Marruedo, “Control predictivo de sistemas no lineales con restricciones: estabilidad y robustez”.
- [38] “JRC Photovoltaic Geographical Information System (PVGIS) - European Commission.” https://re.jrc.ec.europa.eu/pvg_tools/en/ (accessed May 12, 2023).
- [39] “Household Electric Power Consumption | Kaggle.” <https://www.kaggle.com/datasets/uciml/electric-power-consumption-data-set> (accessed Jun. 03, 2023).
- [40] J. Trujillano, J. March, A. Sorribas, and J. Trujillano, “Aproximación metodológica al uso de redes neuronales artificiales para la predicción de resultados en medicina,” *Med Clin (Barc)*, vol. 122, p. 59, 2004.

- [41] J. C. Colás Herrera, “Aplicación de modelos de redes neuronales a la predicción de la fiebre,” 2019.
- [42] L. M. Cabral Guzmán, “Detección de automóviles usando un teléfono inteligente Android con redes neuronales convolucionales,” Mar. 2020, Accessed: May 03, 2023. [Online]. Available: <http://tesis.ipn.mx/xmlui/handle/123456789/28258>
- [43] P. Heras, A. Tutor, M. Luisa, and C. Cabeza, “Predicción de la Radiación mediante el Uso de una Red Neuronal Recurrente”.
- [44] D. Alfonso, L. Sarmiento, C. Andrés, and M. Alayón, “Modeling video transmission losses using SARIMA and ARIMA models,” *Revista*.
- [45] “Almacenamiento sostenible (low) Home ESS - BeePlanet factory.” <https://beeplanetfactory.com/beebatteryhome/> (accessed May 29, 2023).