

UNIVERSIDAD DE LA LAGUNA

FACULTAD DE CIENCIAS

SECCIÓN DE FÍSICA



TRABAJO DE FIN DE GRADO

**Técnicas de inteligencia artificial para la predicción
de energía renovable.**

Aitor L. Ramírez Cabrera

Supervisado por:

Dr. Albano José González Fernández

D. Luis Alonso Siverio

Grado en Física 2020-2021

Abstract

A greater social concern about climate change and the future of our planet together with efficiency improvements in producing renewable energies have made this a rising sector in the last decade. Thus, to estimate the daily produced energy is important for its integration into the electric system, especially for isolated systems as the Canary Islands. The main problem in the forecasting of the energy that will be generated in the next days lies in the dependence of renewable energy on atmospheric conditions, which makes these have a highly variable nature.

In this work, we propose to combine time series of the produced renewable energy, synoptic meteorological conditions and machine learning techniques, specifically, artificial neural networks (ANN) to forecast the daily energy generated in each island.

Some statistical metrics (RMSE, R^2 and BIAS) are used to compare the results of three different types of ANN trained with different data inputs:

- DA: where the inputs are just the time series of generated energy.
- SP: where the inputs are just the atmospheric sea level pressure in an area around the Canary Islands.
- DAP: which is a combination of the other two, and the inputs are the atmospheric sea level pressure and the time series of generated energy.

Those results are also compared with those obtained for the simplest persistence method (ZOH) and with a traditional method for the prediction of time series (ARMA).

Finally, very varied results were obtained, with some cases in which the ANNs significantly improve the ZOH results, and other cases in which there is no significant improvement. In general, for solar energy, DA and DAP are the most reliable, being in all cases better than the ZOH. For wind energy, DA and DAP are also the most reliable, excepting for the cases of La Palma and Gran Canaria in which the best model is SP. Unlike for solar energy, these models do not provide a significant improvement with respect to the ZOH.

Resumen

Una mayor preocupación social sobre el cambio climático y el futuro del planeta junto a mejoras de rendimiento en la obtención de energías renovables han hecho del mismo un sector en auge durante la última década. Sin embargo, la predictibilidad de estas es un gran problema ya que, al contrario que una planta térmica, la disponibilidad de la energía solar y eólica no es continua, ya que poseen una naturaleza altamente variable. Por tanto, hacer una estimación previa de las energías renovables generadas cada día es de gran importancia para la integración de estas en el sistema eléctrico, sobre todo para sistemas aislados como los de Canarias.

En este trabajo se propone el uso de herramientas de aprendizaje automático. En concreto, se implementan redes neuronales artificiales (ANN) para realizar las predicciones de energía diaria generada en cada isla.

Se han utilizado algunas métricas estadísticas (RMSE, R^2 y BIAS) para comparar los resultados obtenidos con tres modelos de ANNs, entrenadas con distintas entradas:

- DA: la entrada para esta red es sólo la serie temporal de energía generada.
- SP: las entradas de esta red son sólo las presiones atmosféricas a nivel del mar en un entorno de Canarias.
- DAP: esta red es una combinación de las dos anteriores, las entradas son las presiones atmosféricas a nivel del mar y la serie temporal de energía generada.

Esos resultados se han comparado, además, con los obtenidos por el método de persistencia más sencillo (ZOH) y por un método tradicional de predicción de series temporales (ARMA).

Hecho esto, se obtuvieron unos resultados muy diversos, con algunos casos en los que las ANN superan en gran medida al ZOH y otros en los que la mejora no es notable (esto se puede observar en los cuadros 1 - 11). A modo general, para la energía solar, DA y DAP suelen ser los más eficientes, superando en todos los casos al ZOH. Para la energía eólica, DA y DAP son también los más eficientes, salvo en los casos de La Palma y Gran Canaria en los que el mejor modelo es SP. A diferencia del caso de energía solar, la mejora de estos con respecto al ZOH no suele ser muy grande.

Índice

1. Introducción	1
1.1. Evolución de las renovables	1
1.2. Métodos tradicionales de predicción	3
1.2.1. Métodos de persistencia	3
1.2.2. Métodos físicos	3
1.2.3. Métodos estadísticos	4
1.3. Inteligencia artificial	4
1.4. Estructura de la memoria	6
2. Objetivos del trabajo	7
3. Datos	8
3.1. Potencia generada	8
3.2. Presión atmosférica	9
4. Metodología	10
4.1. Descarga de los datos	10
4.2. Análisis y pretratamiento de los datos	11
4.3. Tipos de redes usadas	16
4.4. Comparación con métricas estadísticas	16
5. Resultados y discusión	18
5.1. La palma	19
5.2. El Hierro	21
5.3. Tenerife	22
5.4. Gran Canaria	24
5.5. Fuerteventura	26
5.6. Lanzarote	28
6. Resumen y conclusiones	30
7. Agradecimientos	33
A. Nomenclatura	37

Índice de figuras

1.	Evolución de la generación de energía en España (fuente: REE)	2
2.	Energía eólica en Tenerife entre 2016 y 2021	3
3.	Distribución de una red neuronal	5
4.	Ejemplo de la tabla de datos de la REE	8
5.	Puntos de presión tomados en el estudio	9
6.	Energía solar diaria	12
7.	Energía eólica diaria	13
8.	Energía eólica y solar en Tenerife	14
9.	Descomposición de la serie temporal de datos de Tenerife. El eje vertical está en MWh y el horizontal en años.	15
10.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en La Palma.	19
11.	Diagrama de dispersión para las predicciones realizadas por SP sobre energía eólica en La Palma	20
12.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía eólica en El Hierro.	21
13.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Tenerife.	22
14.	Diagrama de dispersión para las predicciones realizadas por DA sobre energía eólica en Tenerife.	23
15.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Gran Canaria.	24
16.	Diagrama de dispersión para las predicciones realizadas por SP sobre energía eólica en Gran Canaria.	25
17.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Fuerteventura.	26
18.	Diagrama de dispersión para las predicciones realizadas por DA sobre energía eólica en Fuerteventura.	27
19.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Lanzarote.	28
20.	Diagrama de dispersión para las predicciones realizadas por DAP sobre energía eólica en Lanzarote.	29

1. Introducción

Summary

In this section a quick introduction to the renewable energies and machine learning fields is made. Firstly, some of the energy plans proposed by the European Union and the Spanish government with the objective of the decarbonisation that led to a boom in the renewable energy field are mentioned. Secondly, some of the traditional methods used for the forecasting of the energy production are shown. Thirdly, a brief explanation of what is an artificial neural network and how it works is given.

1.1. Evolución de las renovables

Con el objetivo de reducir las emisiones de carbono durante la producción de energía en centrales térmicas, en las últimas décadas se han propuesto diversos planes energéticos para favorecer la transición a fuentes de energía más limpias.

A nivel europeo, en 2008 se aprobó un plan energético con el que se proporcionó un marco de apoyo a las renovables a través de la propuesta de unos objetivos claros: para 2020 las emisiones de dióxido de carbono se deberían haber recortado en un 20 % respecto a los niveles de 1990 y las energías renovables deberían representar el 20 % del consumo total de la Unión Europea^[1]. Teniendo en cuenta que para 2006 solo el 7 % de la energía que se consumía en la UE era renovable^[2], observamos que este último era un objetivo bastante ambicioso con el que se favoreció en gran manera el auge de las renovables a nivel europeo. Este objetivo no solo se cumplió con creces sino que además en 2020 las energías renovables superaron a los combustibles fósiles a nivel europeo, alcanzando un 38 % de la energía producida^[3]. Para las próximas décadas se prevé continuar con este crecimiento, con el objetivo de alcanzar la descarbonización total en 2050.

A nivel nacional, el panorama con respecto a las energías renovables ha ido cambiando a lo largo de los años en función de las diversas leyes y recortes que se han hecho a estas. Así en 2007, se publicó en el BOE el Real Decreto 661/2007^[4] en el que se propuso una retribución a la producción de energía procedente de fuentes renovables. Esto tuvo bastante éxito, logrando alcanzar una potencia fotovoltaica casi 10 veces superior a la

que se había fijado inicialmente. De esta manera, en 2009, España se convirtió en uno de los líderes mundiales en inversiones en energías renovables. Posteriormente, entre 2010 y 2012, se aprobaron duros recortes en la retribuciones y se estableció un impuesto a la generación de electricidad del 7%, las consecuencias que estos recortes tuvieron en las renovables se pueden observar en la **Figura 1**. Tras 2015, la capacidad de generación de energía renovable fue creciendo poco a poco, superando a los combustibles fósiles.

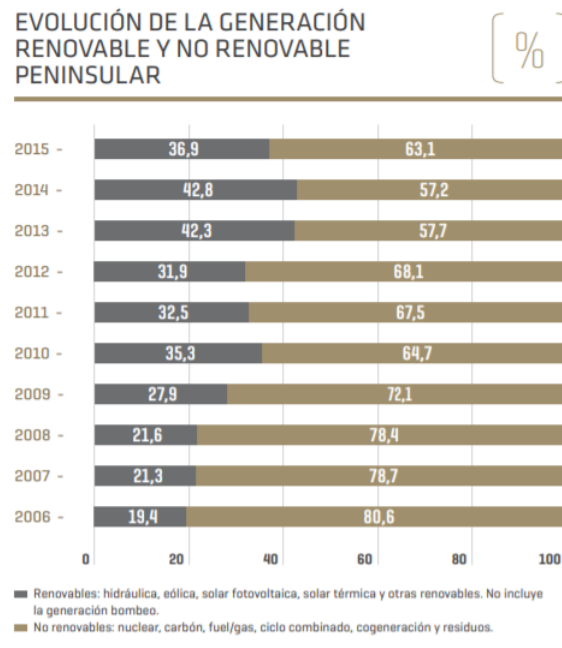


Figura 1: Evolución de la generación de energía en España (fuente: REE)

Mientras tanto, en Canarias, según datos de Red Eléctrica de España, en 2018 solo el 7.7% de la energía consumida fue procedente de fuentes renovables^[5]. Tras esto, se propuso el objetivo de alcanzar para 2025 una producción de energía eléctrica de fuentes renovables de un 45% del total. Los efectos de este objetivo han sido que en 2020 las renovables generaron un 17.5% del total de la energía eléctrica producida en canarias^[6]. Este aumento de producción de energías renovables se puede observar en la **Figura 2**.

Gracias a estos objetivos para conseguir la descarbonización en las próximas décadas, el sector de las energías renovables ha ganado bastante peso, lo que ha hecho posible las mejoras en la producción de estas. A lo largo de la última década (2010-2019), el coste en la producción de energía solar y eólica ha ido decayendo. En concreto, para la energía solar fotovoltaica el coste se ha reducido en un 82% a lo largo de la década, para centrales termosolares ha disminuido un 47%, para centrales eólicas en la costa un 39% mientras

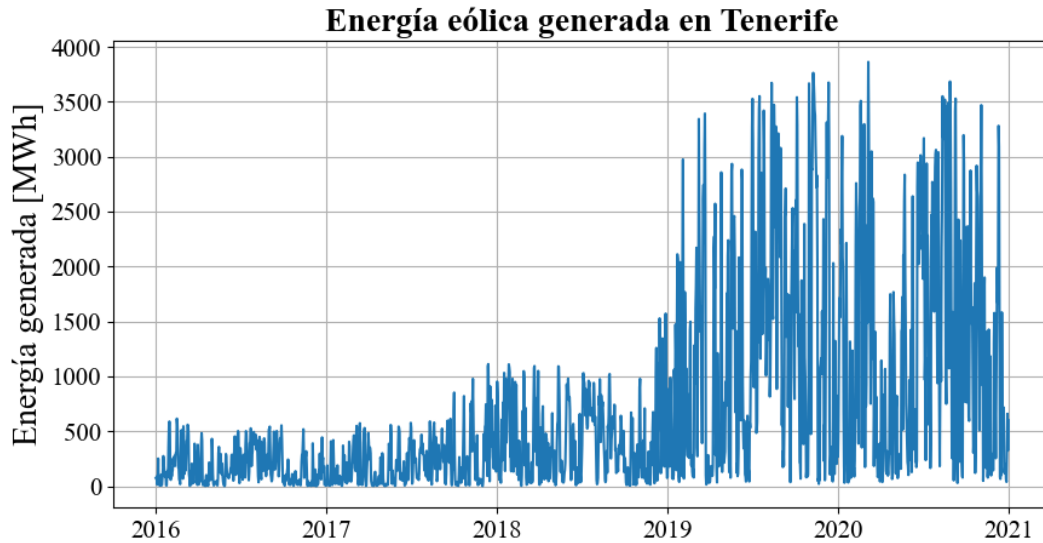


Figura 2: Energía eólica en Tenerife entre 2016 y 2021

que para eólicas fuera de la costa ha disminuido un 29%.^[7]

1.2. Métodos tradicionales de predicción

Los métodos que se usan tradicionalmente para la predicción de energías renovables pueden clasificarse en tres grupos: métodos de persistencia, métodos físicos y métodos estadísticos.^[8]

1.2.1. Métodos de persistencia

Estos son los que se usan normalmente como referencia y se basan en que la energía generada en el futuro será igual a la generada en el presente. A lo largo de este trabajo, por ejemplo, se toma como referencia el retenedor de orden cero (ZOH) en el que las predicciones son las observaciones del día anterior.

1.2.2. Métodos físicos

Son métodos en los que para la predicción no se usan los valores de energía directamente sino que se usan modelos numéricos mesoescalares para predecir el recurso eólico o solar y, posteriormente, relacionarlos con las curvas de producción de los diferentes parques para poder modelizar la generación de energía. Estos son especialmente buenos para

la predicción a medio y largo plazo. Un ejemplo de su aplicación es el de Focken et al.^[9], quienes crearon un modelo capaz de hacer predicciones para las 48 horas siguientes.

1.2.3. Métodos estadísticos

A diferencia de los anteriores, en estos se hace la predicción en base a las series temporales de energías generadas o, en los modelos más sofisticados, combinando estas con predicciones de las condiciones meteorológicas. Cabe destacar que estos solo son buenos para predicciones a corto plazo ya que a medida que aumenta el tiempo en el que queremos predecir, más error tendrá la predicción.^[10]

Concretamente, dentro de modelos estadísticos nos encontramos con modelos de series temporales tales como los autorregresivos de media móvil (ARMA) o los autorregresivos integrados de media móvil (ARIMA). Los primeros han sido usados, por ejemplo, en el estudio realizado por Q.Chen y K.A.Folly^[11] con el que obtuvieron un RMSE de 11,3% y un MAE del 8,8% para la predicción de energía eólica a corto plazo. Con respecto a los ARIMA, estos han sido usados, por ejemplo, por S.Atique et al.^[12] para la predicción de energía generada diariamente por una planta solar, obteniendo, para una predicción de 30 días, un MAE del 17,7%.

1.3. Inteligencia artificial

La inteligencia artificial es un campo que, a pesar de su antigüedad (fue en 1956 cuando el término inteligencia artificial recibió su primer significado: hacer que una máquina se comporte como lo haría un ser humano, de tal manera que se la podría llamar inteligente^[13]), está sufriendo un gran crecimiento en la actualidad gracias a la mejora en la velocidad y capacidad de cálculo de los ordenadores comerciales.

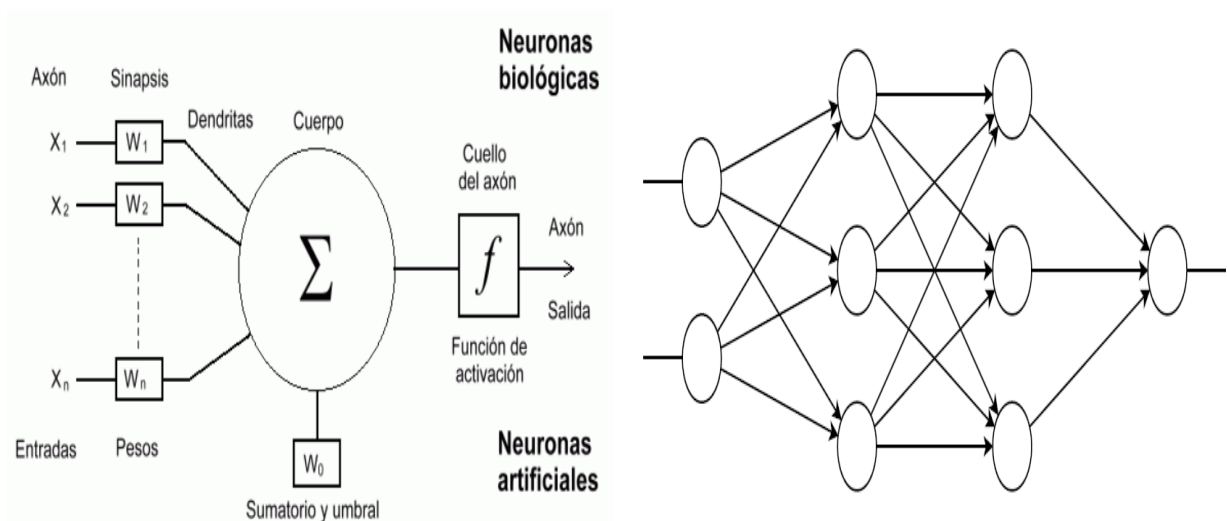
El campo de la inteligencia artificial es un campo muy amplio que podemos dividir en diferentes categorías en función del trabajo a realizar. En el caso de estudio, para la resolución de problemas y reconocimiento de patrones, trabajaremos con sistemas que imitan el pensamiento humano, en concreto, con las llamadas redes neuronales artificiales por su facilidad a la hora de reconocer patrones y aproximar problemas de regresión complejos.

Las redes neuronales artificiales (ANN por sus siglas en inglés) son modelos matemáticos que tratan de emular el comportamiento de una red neuronal biológica (BNN). Así,

una BNN está formada por unidades más simples llamadas neuronas mientras que en una ANN a estas se les denomina perceptrones. Las similitudes entre estas unidades simples se muestran en la **Figura 3a**. En una neurona real, la dendrita recibe señales eléctricas de los axones de otras neuronas, en un perceptrón, estas señales son los valores numéricos x_1, x_2, \dots, x_n . Además, mientras que en una neurona real en la sinapsis entre axón y dendrita la señal eléctrica se modula, en una neurona artificial esto se emula multiplicando los valores de entrada por unos pesos W_1, W_2, \dots, W_n . Finalmente, se modela la señal de salida que emite cada neurona real haciendo la suma ponderada de las entradas y aplicándole una función de activación. La salida de la neurona tendrá la forma:

$$Y = f \left(\sum_{i=1}^n w_i x_i \right) \quad (1)$$

Como se mencionó anteriormente, una ANN no es más que un conjunto de perceptrones, en concreto, en este trabajo usaremos las denominadas perceptrón multicapa en las que los perceptrones están ordenados por capas de manera que la salida de uno de ellos alimenta al resto de perceptrones de la capa siguiente. Un esquema más visual de cómo se ordenan los perceptrones en una ANN se puede observar en la **Figura 3b**.



(a) Comparación entre una neurona biológica y una artificial (Fuente: [14])

(b) Ejemplo de perceptrón multicapa (Fuente: [15])

Figura 3: Distribución de una red neuronal

Para trabajar con una ANN hace falta entrenarla. El proceso de entrenamiento consiste en proporcionarle unos valores en la entrada e intentar que en la salida se obtengan los valores deseados, a través de la modificación de los pesos^[16].

En el caso de estudio, debido a que tenemos un *set* de entrenamiento del que conocemos las salidas, usaremos algoritmos de aprendizaje supervisados. La manera de proceder en el entrenamiento de este tipo de algoritmos será proporcionarle unos valores de entrada (x_{train}) de los cuales conocemos la salida (y_{train}) y dejar que la red calcule sus valores de salida (y_{pred}). A continuación, se define una función de coste (por ejemplo, el error cuadrático medio) con la que calcular la bondad de la red en función del error entre y_{train} e y_{pred} . Una vez hecho este cálculo, se modifican los pesos haciendo uso de un algoritmo de *Backpropagation* que se basa en hacer una propagación del error hacia atrás entre las distintas capas de la red y optimizar el error usando un optimizador (por ejemplo, el descenso del gradiente estocástico, o SGD por sus siglas en inglés). Tras la optimización de los pesos, se repite este procedimiento hasta obtener una red que obtenga buenos resultados.

A diferencia de los métodos tradicionales de predicción de series temporales, en los que se hace una predicción de un parámetro en función de los anteriores, para las redes neuronales en ningún momento se ha limitado el número de atributos ni el de muestras que se le pueden proporcionar como entrada. Esto quiere decir que, si tenemos un problema regresivo, en el que varios atributos están relacionadas pero no conocemos a priori cómo, podemos pasárselos a la red neuronal y dejar que sea esta la que busque la relación entre estos atributos.

1.4. Estructura de la memoria

Tras esta breve introducción, se procede a hacer un resumen sobre la estructura que adopta la memoria. En la Sección 2 se exponen los principales objetivos del trabajo. A continuación, en la Sección 3, se muestran los datos con los que se va a trabajar durante el proyecto y las bases de datos de las que se han extraído. En la Sección 4, se enumeran los pasos seguidos durante el proyecto para la descarga de los datos, su tratamiento y el posterior entrenamiento de las ANN. En la Sección 5 se exponen los resultados obtenidos para cada caso estudiado. Finalmente, en la Sección 6 se hace un breve resumen del trabajo hecho, se exponen unas conclusiones sobre los resultados obtenidos y se proponen algunos métodos para continuar con este trabajo en el futuro.

2. Objetivos del trabajo

Summary

The main goal of this project is to make use of artificial neural networks to forecast the daily renewable energy production in the Canary Islands by using the time series of energy production and the synoptic meteorological conditions. Then, we can compare these predictions with the ones obtained with traditional statistical methods for time series analysis to conclude if ANNs are more reliable.

Con el auge de las energías renovables y la previsión del crecimiento que estas tendrán a lo largo de la próxima década, es importante obtener un buen modelo matemático con el que hacer una predicción fiable de la potencia que se generará por fuentes renovables en los días posteriores. Este estudio cobra mayor importancia en Canarias, no solo por el ambicioso plan de crecimiento que se ha propuesto (mencionado en **1.1**) sino, además, por tratarse de sistemas pequeños y aislados en los que la estimación de la potencia generada es primordial para la estabilidad de los sistemas.

En este trabajo se propone que los modelos matemáticos encargados de la predicción sean las ya mencionadas redes neuronales. La idea es aprovechar su capacidad en el reconocimiento de patrones para analizar las series temporales de energía generada y dar una predicción de la energía que se generará en el día siguiente. Además dada la posibilidad de hacer predicciones a partir de varias variables, se propone no solo usar la serie temporal de energía generada sino también la de condiciones sinópticas de la atmósfera de Canarias ya que, a priori, estas están relacionadas con la nubosidad y el viento y, por lo tanto, podrían estar relacionadas con la generación de energía solar y eólica.

Por lo tanto, de este trabajo se espera comprobar que los modelos regresivos basados en redes neuronales son más fiables que los tradicionales y, más importante aún, comprobar si el uso de las condiciones sinópticas de la atmósfera mejoran la predicción.

3. Datos

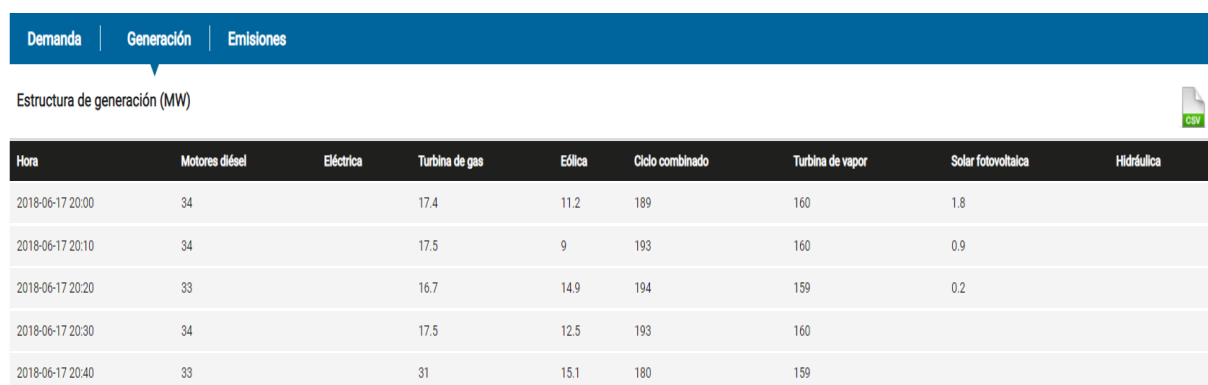
Summary

For the analysis, two databases were used: one from REE with the data of the power generated every 10 minutes in each island, and the other from ECMWF with the daily mean atmospheric pressure at sea level in 20 different points in an area around the Canary Islands.

Para este trabajo usaremos datos de potencia generada por fuentes eólicas y solar fotovoltaicas en cada una de las islas, así como datos de presiones a nivel del mar en un entorno de Canarias.

3.1. Potencia generada

Las series temporales de potencia generada en las islas se extrajeron de las bases de datos de Red Eléctrica de España (REE), unas bases de datos públicas que se pueden consultar en su página web^[16]. En esta página web se proporcionan unas tablas por día y por isla en las que se muestra la potencia generada (en MW) por cada una de las fuentes cada diez minutos. En la **Figura 4** se muestra un recorte de esta tabla para el día 17 de junio de 2018 en Tenerife.



Hora	Motores diésel	Eléctrica	Turbina de gas	Eólica	Ciclo combinado	Turbina de vapor	Solar fotovoltaica	Hidráulica
2018-06-17 20:00	34		17.4	11.2	189	160	1.8	
2018-06-17 20:10	34		17.5	9	193	160	0.9	
2018-06-17 20:20	33		16.7	14.9	194	159	0.2	
2018-06-17 20:30	34		17.5	12.5	193	160		
2018-06-17 20:40	33		31	15.1	180	159		

Figura 4: Ejemplo de la tabla de datos de la REE

En concreto, para el caso de estudio, se extrajeron de estas tablas los datos entre el 1 de enero de 2013 y el 31 de diciembre de 2020.

3.2. Presión atmosférica

Para las condiciones sinópticas de la atmósfera se usó la base de datos del Centro Europeo de Previsiones Meteorológicas a Plazo Medio (ECMWF por sus siglas en inglés). En concreto, se trabajó con los datos de presión a nivel del mar proporcionados por ERA5, el último reanálisis atmosférico realizado por ECMWF, que proporciona datos horarios con una resolución $0,25^\circ \times 0,25^\circ$.^[17]

Para el caso de estudio, se tomó la media diaria y se seleccionaron los puntos de presión que se muestran en la **Figura 5**, siendo los rojos los más cercanos al área de estudio y los azules aquellos más lejanos y, presumiblemente, de menor influencia.

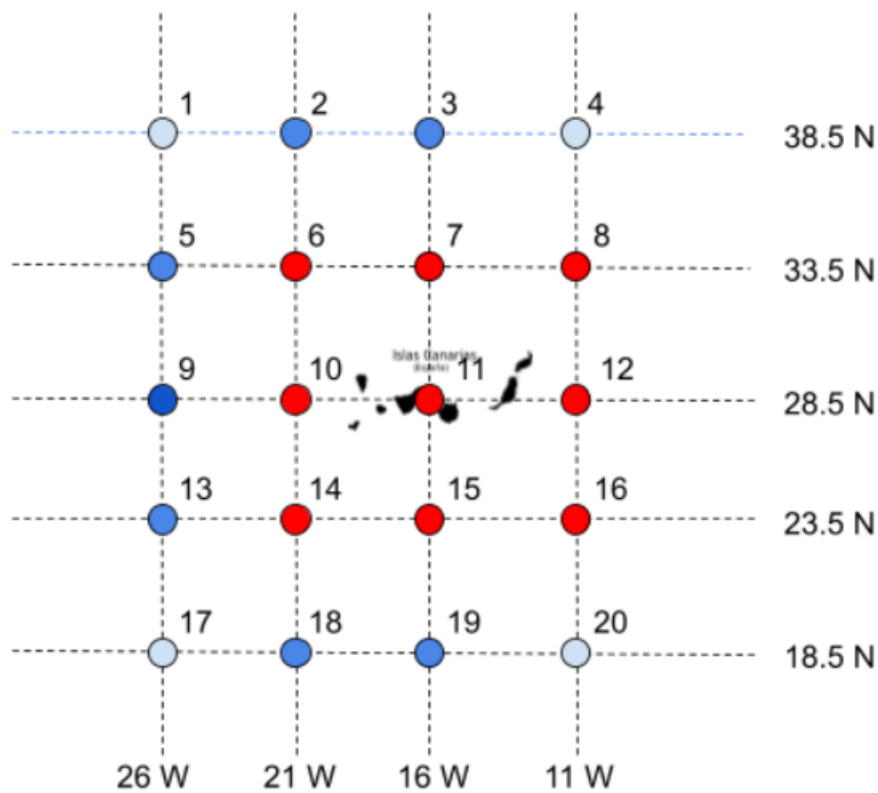


Figura 5: Puntos de presión tomados en el estudio

4. Metodología

Summary

The methodology followed in this work was the general one in the study of time series. First, the data was downloaded from the two databases. Then it was analyzed and some transformations, like a daily integration, the detrend of the series and a scaling in the range [0,1], were applied to it. The third step was to define three neural networks, one that only analyzes the time series of energy (DA), another that only analyzes the atmospheric pressure (SP) and a third one that analyzes the time series of energy and the atmospheric pressure (DAP). After that, the last step was to apply some statistical metrics (RMSE, R^2 and BIAS) to evaluate the models and compare them with two of the traditional prediction models such as ZOH and ARMA.

En general, el método para el desarrollo de modelos predictivos comienza con la descarga de los datos y su análisis para obtener conocimientos previos sobre el tipo de datos con el que estamos trabajando. A continuación, se pasa al tratamiento de estos haciendo las transformaciones correspondientes para asegurar que el modelo converge y que ningún parámetro externo afecta a la predicción. Hecho esto, se define el tipo de algoritmos a usar para la predicción y se evalúan a través del uso de métricas como el error cuadrático medio (MSE) o el coeficiente de determinación (R^2). Por último, se estudian distintos ajustes de estos algoritmos con los que se pueda mejorar la predicción.

4.1. Descarga de los datos

La primera tarea a realizar en este estudio sería la descarga de los datos de potencia generada de la página web de REE^[18]. Pero, dado que las tablas que proporciona REE son generadas de forma dinámica tras la solicitud sobre una fecha concreta, la descarga a mano de estos datos se vuelve una tarea imposible. Para solucionar esto se optó por desarrollar un código en *Python* que, a través de la librería *Selenium*^[19], permitiese el *web scrapping* automático de estas páginas web, es decir, que permitiese leer el código *html* de las páginas generadas para así descargar los datos de las tablas.

4.2. Análisis y pretratamiento de los datos

Tras la descarga de los datos, dada la gran cantidad de estos para cada isla, el primer paso fue convertir la potencia generada cada diez minutos en energía generada cada hora. Para ello se integraron los datos de potencia haciendo uso de la regla del trapecio:

$$\int_a^b f(x)dx \approx \sum_{k=1}^N \frac{f(x_{k-1}) + f(x_k)}{2} \Delta x_k \quad (2)$$

A continuación, para empezar con un breve análisis de los datos, se representó, en diagramas de cajas, la energía media generada cada día durante los ocho años de estudio. Estas gráficas se muestran en la **Figura 6** y la **Figura 7**.

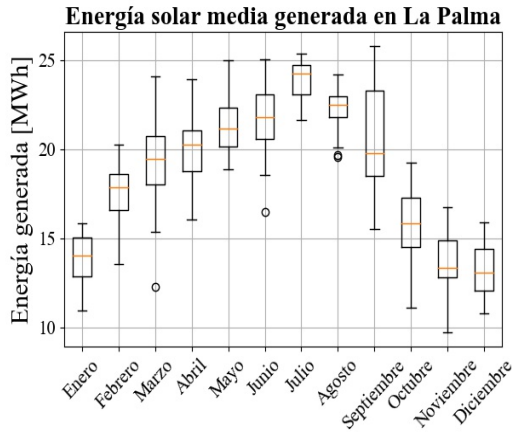
En las gráficas de la **Figura 6** es fácil observar que, como era de esperar, hay un comportamiento estacional por el que aumenta la generación de energía solar en los meses de verano y disminuye en los meses de invierno. Además, en las islas más orientales (Fuerteventura y Lanzarote) se puede distinguir una curva más suave, principalmente por la falta de nubosidad que hay en estas islas, y que alcanza el máximo antes que el resto.

En las gráficas de la **Figura 7** es interesante observar el aumento que tiene la energía eólica durante los meses de Julio y Agosto. Este aumento es debido principalmente a los vientos alisios que, como se comprobó en [20], en los meses de verano afectan más a las zonas en las que se encuentran las instalaciones actuales de parques eólicos.

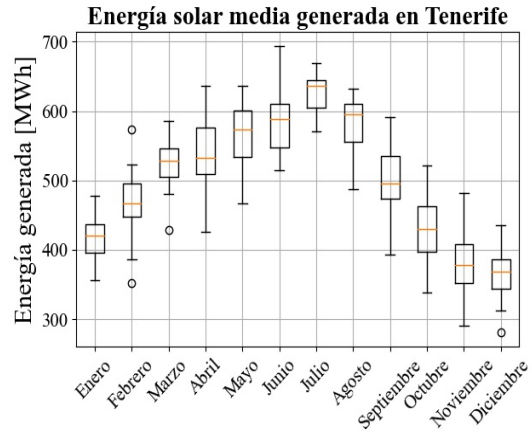
El siguiente paso fue analizar las series temporales de la energía diaria para facilitar el tratamiento de los datos. En la **Figura 8** se muestra un ejemplo de estas gráficas para la energía generada en Tenerife.

En estas se pudo observar que en los últimos años hay una tendencia al alza en la generación de energía debido a la instalación de nuevos parques eólicos en las islas. Pero, como en el caso de estudio nos interesa hacer predicciones para un sistema aislado en el que no se añaden nuevas infraestructuras a la red eléctrica, el siguiente paso en el tratamiento fue quitarle la tendencia a estos datos. Para ello, primero tenemos que comprobar si las series temporales que se están tratando se pueden modelar como series aditivas, esto es, que los valores se pueden descomponer en la suma de sus componentes de tendencia, estacionalidad y unos residuos o, si por el contrario, se pueden modelar como series multiplicativas en las que los valores se pueden descomponer como la multiplicación de sus componentes.

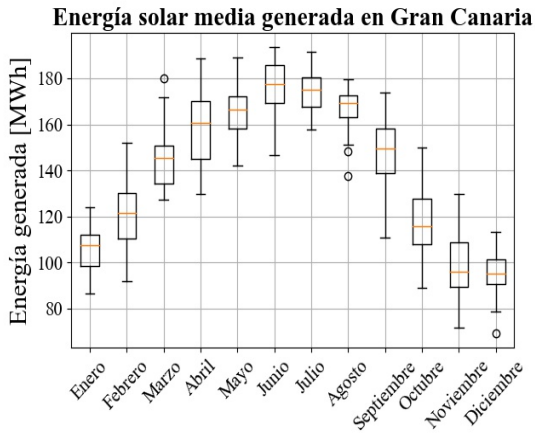
Para determinar cuál de estos dos modelados es el apropiado para la serie temporal,



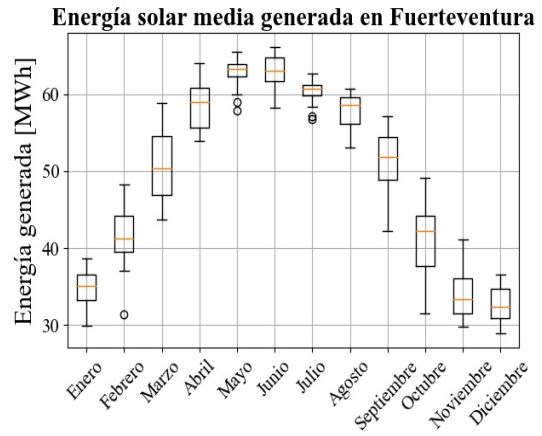
(a) La Palma



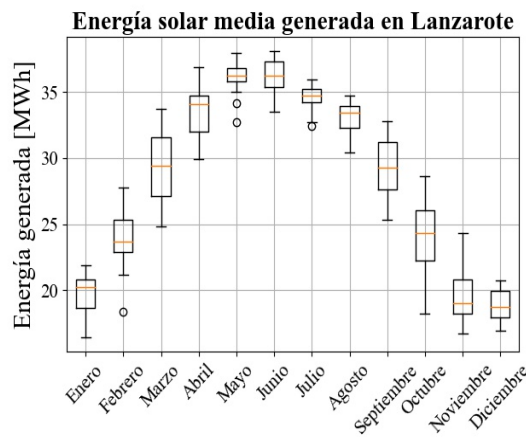
(b) Tenerife



(c) Gran Canaria

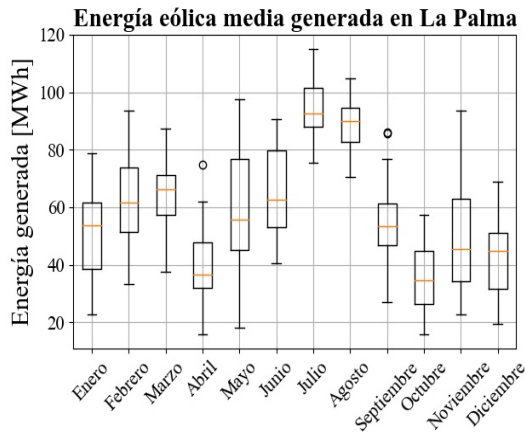


(d) Fuerteventura

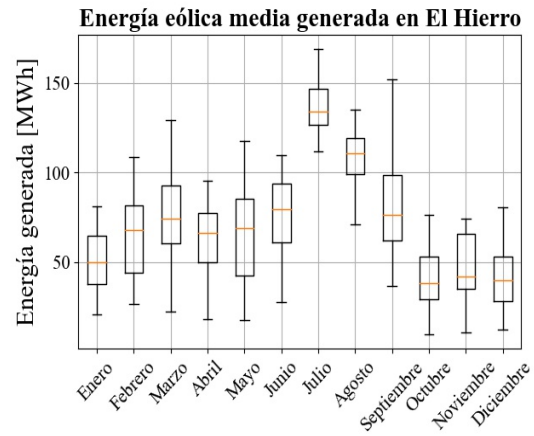


(e) Lanzarote

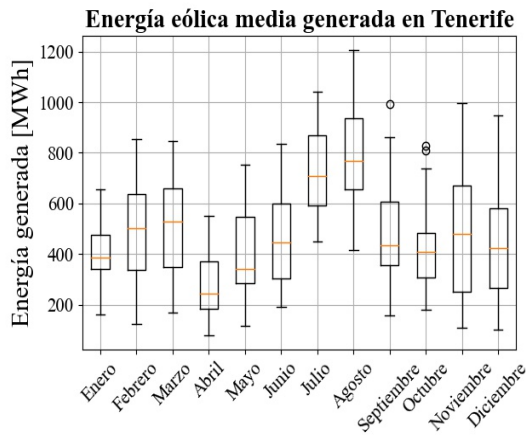
Figura 6: Energía solar diaria



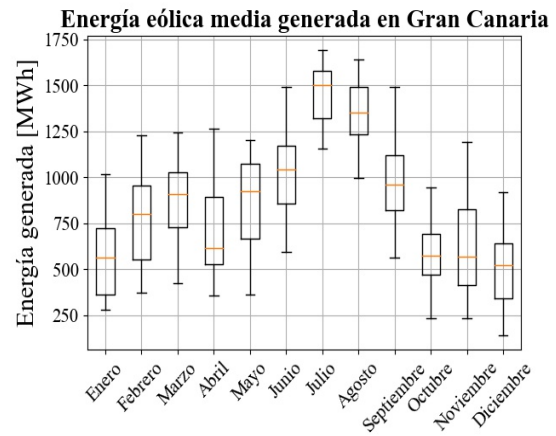
(a) La Palma



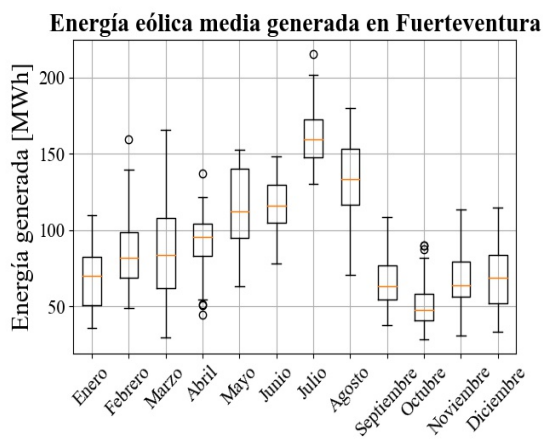
(b) El Hierro



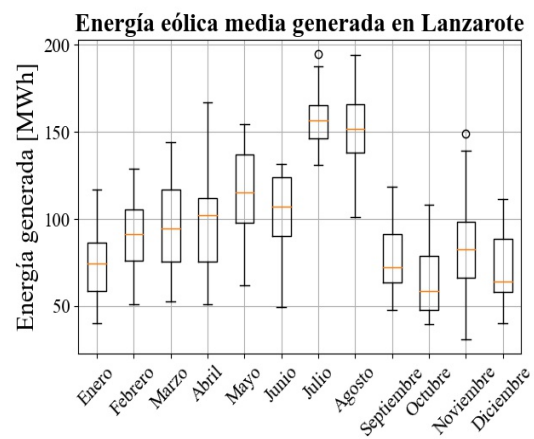
(c) Tenerife



(d) Gran Canaria



(e) Fuerteventura



(f) Lanzarote

Figura 7: Energía eólica diaria

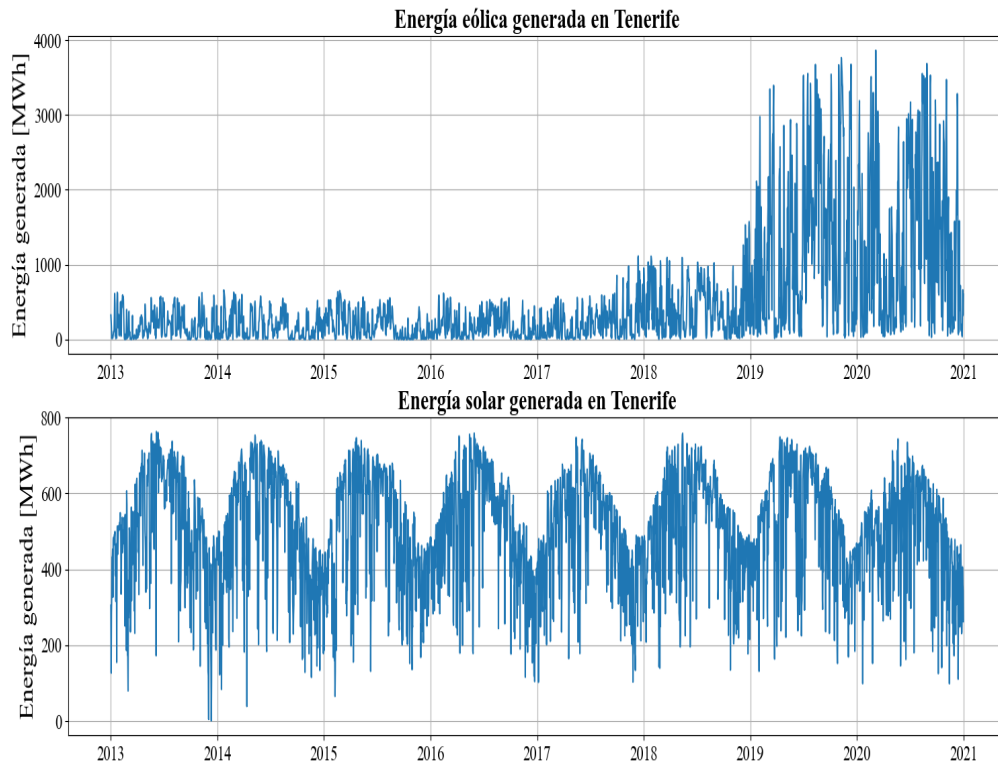


Figura 8: Energía eólica y solar en Tenerife

se hace una descomposición de esta en sus componentes estacionales y se comprueba con cual se obtienen unos residuos menores y con una distribución estocástica, que no presente patrones o pueda aproximarse de forma funcional. En el caso tratado, se hizo la descomposición estacional usando el módulo *seasonal_decompose* de la librería *statsmodel* de *Python*^[21] y se obtuvo un mejor resultado con el modelado como series multiplicativas.

Un ejemplo de la descomposición como serie multiplicativa para los datos de energía eólica en Tenerife se puede observar en la **Figura 9**, en la primera fila de esta se muestran los datos reales. En la segunda se muestra la tendencia de estos. Como se había comentado anteriormente, se observa el aumento de la tendencia en los últimos años. En la tercera fila se representa la estacionalidad de los datos y se observa que es bastante regular a lo largo de los años. Por último, en la cuarta fila se muestran unos residuos bastante planos con una distribución bastante uniforme, dejando claro que la modelización como serie multiplicativa es la correcta, ya que representa un comportamiento mucho más acorde a lo esperado en la aditiva.

Descomposición estacional de energía eólica en Tenerife

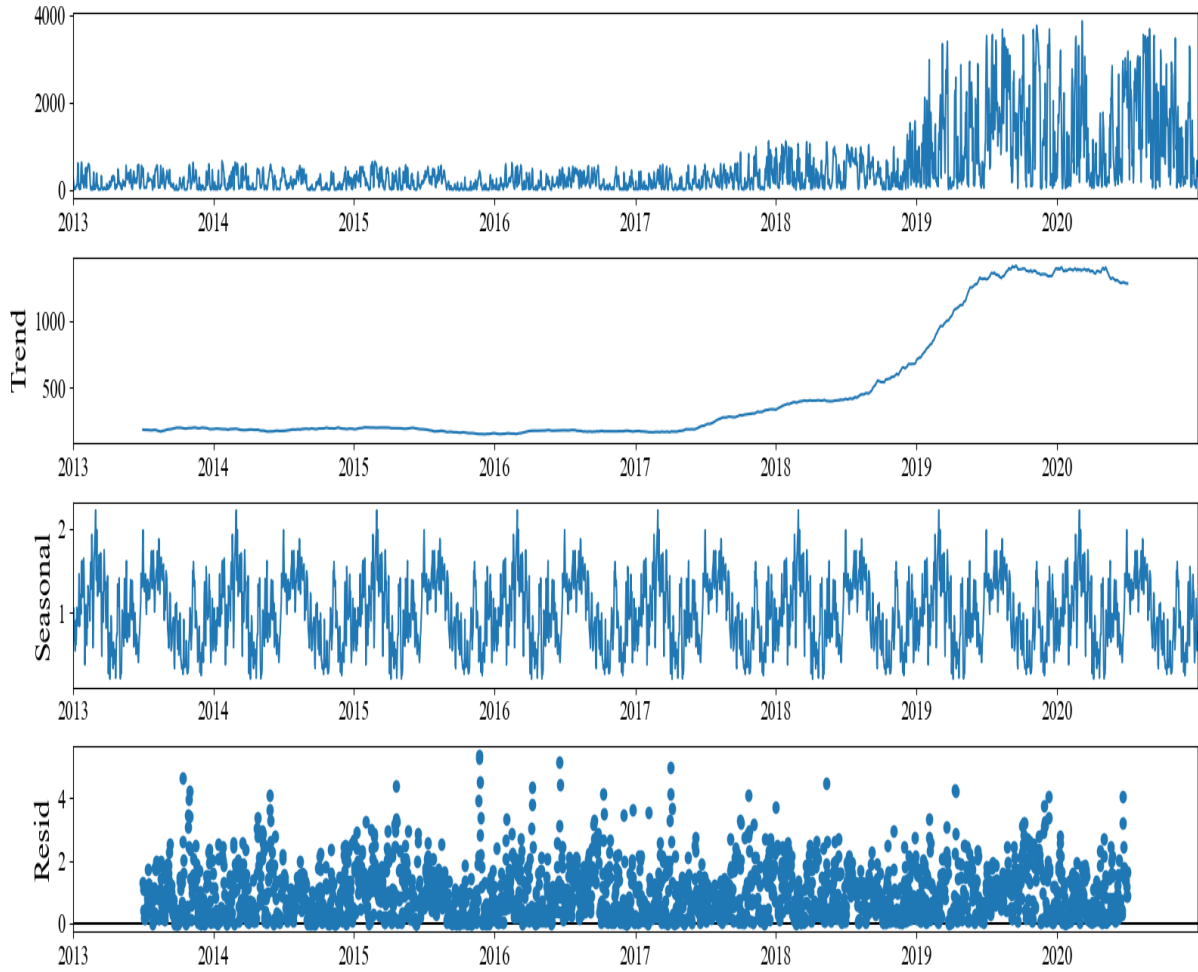


Figura 9: Descomposición de la serie temporal de datos de Tenerife. El eje vertical está en MWh y el horizontal en años.

Otro paso importante para tratar de obtener mejores resultados es el escalado de los datos, en el caso de estudio los datos se escalaron en el intervalo $[0, 1]$. Este escalado es importante por varios motivos. Para empezar, hace que el entrenamiento sea más rápido y reduce las posibilidades de que el error se quede atascado en un mínimo local obteniendo un resultado peor. Otro motivo por el que escalar los datos es la diferencia de escala entre las energías y las presiones, lo que puede degenerar en que la red dé más importancia al atributo con mayor escala y se obtengan unas predicciones imprecisas.

Por último, para poder evaluar la red, es importante dividir aleatoriamente los datos en un conjunto de entrenamiento y un conjunto de test (tal y como se explicó en **(1.3)**). En el caso de estudio se decidió que esta división fuera de un 80 % para entrenamiento y un 20 % para test.

4.3. Tipos de redes usadas

Como se ha venido adelantando en la introducción, las redes neuronales elegidas para el estudio son las denominadas perceptrón multicapa (**Figura 3**). En concreto, se usó el *MLPRegressor* (perteneciente a librería *Scikit learn* de *Python*). Este modelo usa como función de coste el error cuadrático medio y la optimiza haciendo uso del descenso del gradiente estocástico o a través del optimizador adam (*Adaptive Moment Estimation*), una mejora del descenso del gradiente estocástico.

Durante el trabajo se estudió como de buenas eran las predicciones de estas redes en tres casos distintos:

- DA: Tomando como valores de entrada los datos de energía generada los tres días anteriores y como salida la energía que se generará el día siguiente.
- SP: Tomando como valores de entrada sólo la presión atmosférica a nivel del mar y como salida la energía generada el mismo día.
- DAP: Haciendo una combinación de las dos anteriores y tomando como valores de entrada las energías generadas los tres días anteriores y la presión atmosférica del mismo día y como salida la energía generada ese día.

Para cada uno de los casos, los parámetros de red óptimos para la predicción (esto es, el número de capas, número de neuronas por capa, función de activación y optimizador) se obtuvieron a partir de una búsqueda exhaustiva entre diversas combinaciones de los mismos haciendo uso de las funciones *GridSearchCV*^[22] y *RandomizedSearchCV*^[23] del módulo *Scikit learn*.

4.4. Comparación con métricas estadísticas

Por último, para contrastar la bondad de las redes se calcularon tres métricas de ajuste y se compararon con las obtenidas por métodos tradicionales de predicción de series temporales.

Las métricas usadas fueron:

- Error cuadrático medio (MSE) es la media de los errores al cuadrado y ofrece una magnitud del error.

$$MSE = \frac{1}{n} \sum_{i=1}^n (SR_p(i) - SR_o(i))^2 \quad (3)$$

- Raíz del error cuadrático medio (RMSE) con la que convertimos el MSE a las unidades reales de la variable de salida obteniendo una descripción más significativa del resultado.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (SR_p(i) - SR_o(i))^2} \quad (4)$$

- Coeficiente de determinación (R^2) que nos indica la bondad del ajuste de un conjunto de predicciones a los valores reales.

$$R^2 = 1 - \frac{\sum_{i=1}^n (SR_o(i) - SR_p(i))^2}{\sum_{i=1}^n (SR_o(i) - \frac{1}{n} \sum_{i=1}^n (SR_o(i)))^2} \quad (5)$$

siendo SR_o y SR_p los datos reales y las predicciones respectivamente.

- Sesgo (BIAS) que proporciona una media de la diferencia que se produce entre la predicción y el valor real, cuanto menor sea el BIAS en valor absoluto mejor será el modelo.

$$BIAS = \frac{1}{n} \sum_{i=1}^n (SR_p - SR_o) \quad (6)$$

Los métodos tradicionales de predicción de series temporales con los que se contrastó la bondad de las ANN fueron:

- Retenedor de orden cero (ZOH): este modelo se basa en suponer que la energía que se generará un día es igual a la que se generó el día anterior, es decir, que la salida del modelo sea igual a la entrada desplazada un día.
- Modelo autorregresivo de media móvil (ARMA): es un modelo amplio y flexible en el que se combinan los procesos autorregresivos AR(p) y los procesos de medias móviles MA(q). La forma general de los ARMA(p,q) puede ser escrita como:

$$X_t = \varepsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (7)$$

donde ϕ_1, \dots, ϕ_p son los parámetros del modelo autorregresivo, $\theta_1, \dots, \theta_q$ son los parámetros del modelo de medias móviles y $\varepsilon_t, \varepsilon_{t-1}, \dots$ son los términos de error.

5. Resultados y discusión

Summary

In this section the metrics obtained for each method are shown. For a better understanding, these are grouped according to the island, showing for each island two different tables, corresponding to the two renewable energies sources (solar and wind energy). For each table, the best metrics are highlighted in yellow.

Additionally, for the best model of each case, a scatter plot is shown to facilitate the interpretation of the statistical metrics.

A continuación, se muestran las métricas obtenidas por los modelos sencillos y por las redes neuronales. Estas están calculadas en tanto por uno para poder hacer comparaciones entre las distintas islas y fuentes. Con el fin de facilitar la visualización de estos resultados se ha decidido agrupar los mismos por isla, mostrando para cada una dos tablas con las métricas obtenidas por los modelos para las dos fuentes de energía renovable existentes en la isla. En cada tabla está resaltado en amarillo el modelo con el que se ha obtenido la mejor métrica.

Además, para tener una visión más gráfica de la bondad del elegido como mejor modelo, se adjuntan diagramas de dispersión entre las observaciones y las predicciones. En estos se muestra también una recta (en color verde), que corresponde a la bisectriz del primer y tercer cuadrante, y representa la posición en la que estarían los puntos si la predicción fuera perfecta.

5.1. La palma

Solar			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,28	0,24	0,00
ARMA(6,8)	0,20	-0,48	0,11
DA	0,25	0,40	0,02
SP	0,26	0,34	-0,00
DAP	0,24	0,46	0,03

Cuadro 1: Métricas para energía solar en La Palma

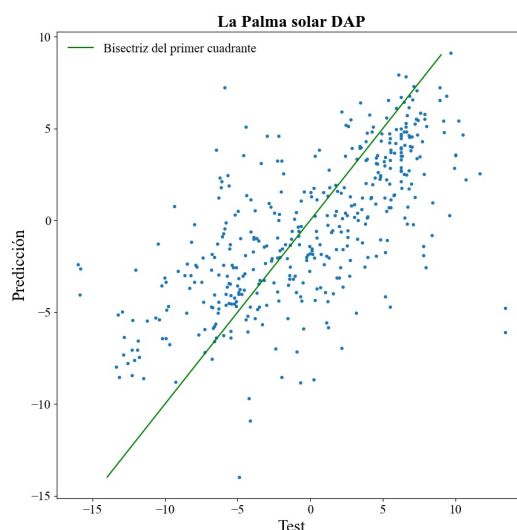


Figura 10: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en La Palma.

Como podemos observar en el **Cuadro 1** el ZOH es el que menor parámetro de BIAS tiene aunque tiene un RMSE alto con respecto al resto de modelos y un R² muy bajo. Con el ARMA(6,8) se obtiene un RMSE bastante bajo pero el parámetro de BIAS de este es alto con respecto a los otros modelos. Con respecto a las redes neuronales, estas están igualadas aunque si hubiera que elegir alguna se tomaría la que combina las energías de días anteriores y las presiones (DAP) por ser la que menor RMSE y mayor R² tiene aunque, como se puede observar en el diagrama de dispersión de las predicciones hechas

con este modelo (**Figura 10**), algunos puntos se alejan de la bisectriz, por lo tanto, las predicciones no mejoran mucho las del ZOH, mostrando la dificultad del problema estudiado. Por último, destacar que para ninguno de estos modelos $R^2 \geq 0,5$.

Eólica			
	RMSE relativo	R^2	BIAS relativo
ZOH	0,48	0,60	0,00
ARMA(8,7)	0,57	-1,20	0,40
DA	0,43	0,66	0,06
SP	0,37	0,77	-0,12
DAP	0,43	0,66	-0,03

Cuadro 2: Métricas para energía eólica en La Palma

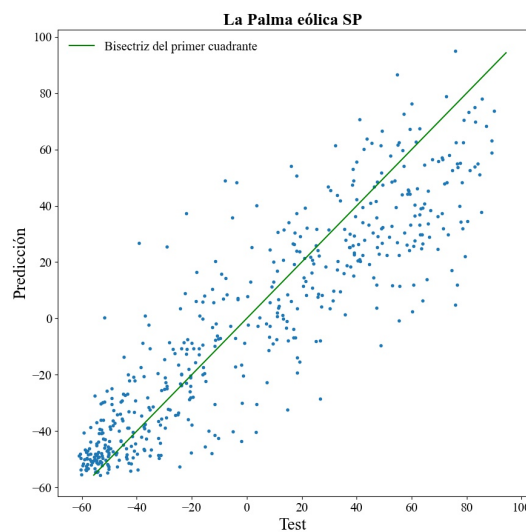


Figura 11: Diagrama de dispersión para las predicciones realizadas por SP sobre energía eólica en La Palma

Para energía eólica, vemos en el **Cuadro 2** que el mejor modelo en cuanto a RMSE y R^2 es el SP, la red neuronal en la que las entradas son solo las presiones, aunque para este se obtiene un parámetro de BIAS bastante grande. Para este modelo, la gráfica de dispersión (**Figura 11**) muestra una mayor concentración de puntos en torno a la bisectriz para las energías más bajas, además, como era de esperar debido al alto R^2 en este caso se observan menos *outliers* que para el caso de energía solar (**Figura 10**).

5.2. El Hierro

Eólica			
	RMSE relativo	R^2	BIAS relativo
ZOH	0,57	0,57	0,00
ARMA(4,4)	0,65	-0,14	0,13
DA	0,52	0,60	0,10
SP	0,62	0,51	0,16
DAP	0,51	0,61	0,01

Cuadro 3: Métricas para energía eólica en El Hierro

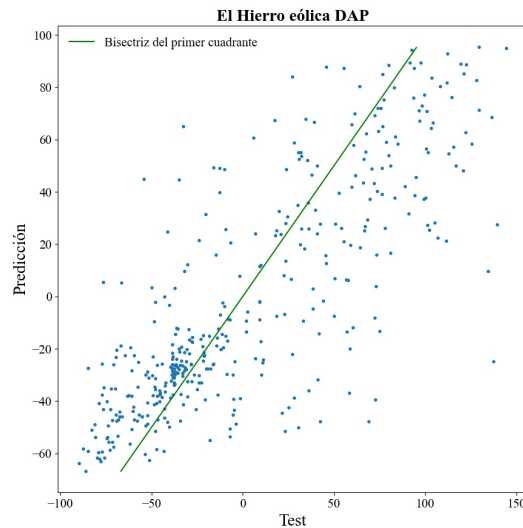


Figura 12: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía eólica en El Hierro.

Respecto a la energía eólica en El Hierro, vemos en el **Cuadro 3** que el mejor modelo para la predicción es, sin ninguna duda, el DAP. En la **Figura 12** se muestra el diagrama de dispersión de las predicciones hechas por este frente a los valores reales. En esta se puede explicar el alto R^2 obtenido para este modelo ya que hay muchos puntos en torno a la bisectriz pero, además, se pueden observar bastantes *outliers* que generan el alto RMSE.

5.3. Tenerife

Solar			
	RMSE relativo	R^2	BIAS relativo
ZOH	0,26	0,21	0,00
ARMA(1,7)	0,19	-0,19	0,08
DA	0,23	0,39	-0,05
SP	0,22	0,38	0,01
DAP	0,23	0,40	-0,03

Cuadro 4: Métricas para energía solar en Tenerife

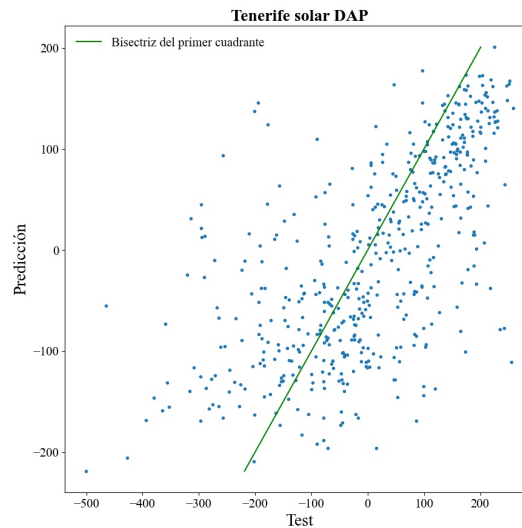


Figura 13: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Tenerife.

En el **Cuadro 4**, vemos que no hay un claro vencedor para la predicción de energía solar en Tenerife ya que todos tienen un $R^2 \leq 0,5$ y por lo tanto son considerados modelos predictivos menos eficientes. Aún así, vemos que, como era de esperar, los modelos basados en ANN superan al ZOH.

En la **Figura 13** se muestra el diagrama de dispersión entre las predicciones y las observaciones obtenidas con el modelo con un R^2 mayor, es decir, el DAP. En esta, vemos una gran cantidad de *outliers* que provocan el bajo valor de R^2 obtenido para el modelo.

Eólica			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,84	0,55	0,00
ARMA(7,3)	1,79	-0,72	0,26
DA	0,77	0,59	0,03
SP	0,90	0,27	0,12
DAP	0,78	0,59	0,04

Cuadro 5: Métricas para energía eólica en Tenerife

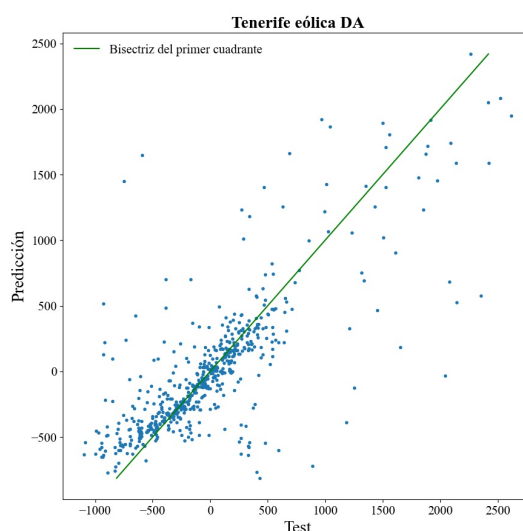


Figura 14: Diagrama de dispersión para las predicciones realizadas por DA sobre energía eólica en Tenerife.

Con respecto a la energía eólica en Tenerife, en el **Cuadro 5** vemos que el modelo que aporta la mejor predicción es el DA, es decir, la red neuronal a la que solo se le ha aportado información de la energía generada los tres días anteriores, aunque con resultados muy similares a los obtenidos por DAP. La gráfica de dispersión obtenida para este modelo se muestra en la **Figura 14**. En esta, los puntos correspondientes a energías más bajas se encuentran en torno a la bisectriz mientras que los correspondientes a energías más altas se alejan de la misma. Esto puede ser debido al gran aumento que tuvo la energía eólica en Tenerife en 2019 (tal y como se muestran en la **Figura 8**), ya que, a pesar de haberle

quitado la tendencia a los datos, la red no sería capaz de predecir un aumento tan brusco de energía generada.

5.4. Gran Canaria

Solar			
	RMSE relativo	R^2	BIAS relativo
ZOH	0,27	0,35	0,00
ARMA(1,8)	0,12	-0,05	0,03
DA	0,24	0,51	0,01
SP	0,28	0,31	-0,02
DAP	0,24	0,51	0,00

Cuadro 6: Métricas para energía solar en Gran Canaria

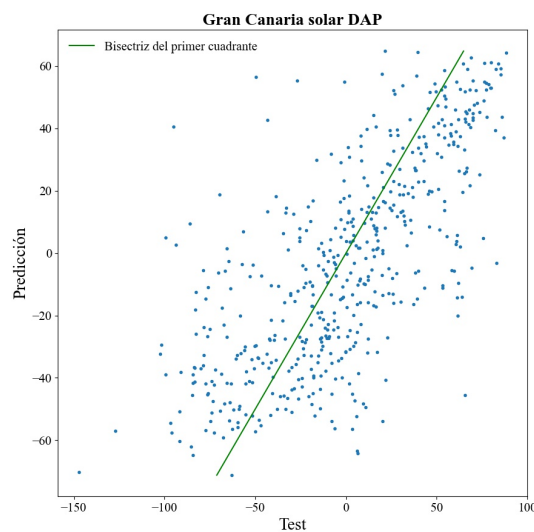


Figura 15: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Gran Canaria.

Para energía solar en Gran Canaria, en el **Cuadro 6** podemos observar que, de entre los modelos comparados, el mejor para la predicción es el DAP, pero sin ser un modelo eficiente debido a su bajo R^2 . En la **Figura 15**, se muestra el diagrama de dispersión de

predicciones frente a observaciones. Tal y como era de esperar, debido al bajo R^2 obtenido por el modelo, lo puntos se encuentran bastante dispersos con respecto a la bisectriz.

Eólica			
	RMSE relativo	R^2	BIAS relativo
ZOH	0,51	0,58	0,00
ARMA(3,2)	0,81	-0,09	0,12
DA	0,49	0,59	-0,02
SP	0,41	0,67	0,13
DAP	0,49	0,60	-0,05

Cuadro 7: Métricas para energía eólica en Gran Canaria

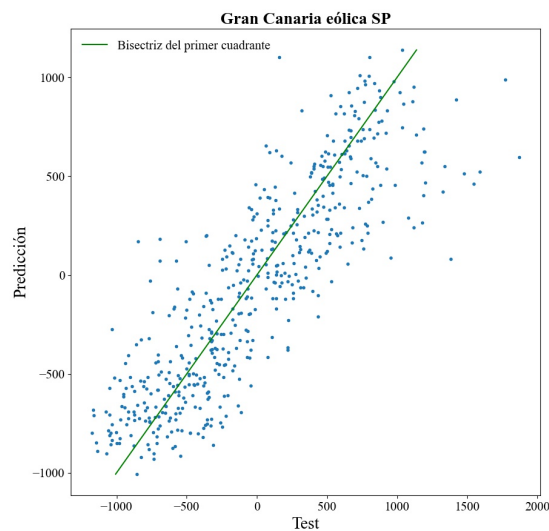


Figura 16: Diagrama de dispersión para las predicciones realizadas por SP sobre energía eólica en Gran Canaria.

Con respecto a la energía eólica en Gran canaria, vemos en el **Cuadro 7** que el mejor modelo para la predicción es el SP. Aunque este tiene un BIAS bastante grande, para salvar este problema sería mejor tomar la red a la que se le da información sobre la energía generada los días anteriores.

5.5. Fuerteventura

Solar			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,14	0,74	0,00
ARMA(3,3)	0,10	-0,07	0,02
DA	0,12	0,81	0,03
SP	0,20	0,47	-0,01
DAP	0,12	0,81	0,02

Cuadro 8: Métricas para energía solar en Fuerteventura

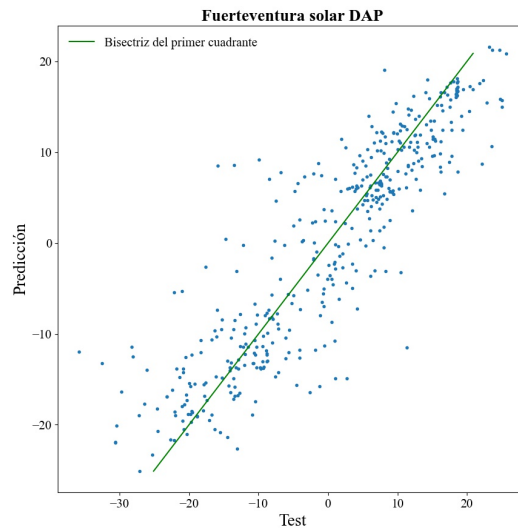


Figura 17: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Fuerteventura.

En Fuerteventura, para energía solar, vemos en el **Cuadro 8** que los mejores modelos para la predicción son el DA y DAP. Se observa, además, que estos tienen un R² de 0,81 lo cual quiere decir que es un buen modelo. Esto último lo podemos observar en el diagrama de dispersión (**Figura 17**) ya que hay muchos puntos en torno a la bisectriz. Este resultado, es decir, el buen comportamiento de los modelos para la predicción de energía solar en esta isla, no es de extrañar, ya que, como se había comentado al inicio, el efecto de las nubes en Lanzarote y Fuerteventura es menos relevante que en el resto.

Dependiendo de forma más directa de la variación anual debida a la posición de la Tierra con respecto al Sol.

Eólica			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,57	0,56	0,00
ARMA	1,01	-0,12	0,10
DA	0,54	0,62	-0,07
SP	0,60	0,46	-0,01
DAP	0,54	0,60	-0,13

Cuadro 9: Métricas para energía eólica en Fuerteventura

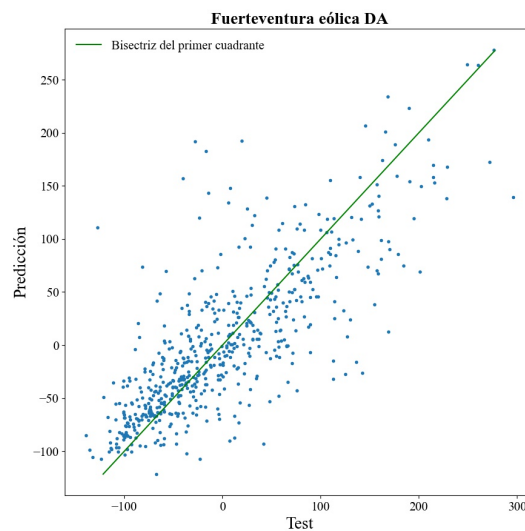


Figura 18: Diagrama de dispersión para las predicciones realizadas por DA sobre energía eólica en Fuerteventura.

Con respecto a la energía eólica en Fuerteventura, en el **Cuadro 9** podemos observar que el mejor modelo para la predicción es el DA. En la **Figura 18** se muestra el diagrama de dispersión para este modelo, al igual que en los casos anteriores para energía eólica, vemos que para energías bajas los puntos están más cerca de la bisectriz que para energías altas, como se explicó para el caso de Tenerife, esto debe ser debido al brusco aumento en la generación de energía eólica en 2019.

5.6. Lanzarote

Solar			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,13	0,15	0,00
ARMA(8,1)	0,08	0,05	0,02
DA	0,12	0,82	0,01
SP	0,20	0,46	0,01
DAP	0,12	0,82	0,01

Cuadro 10: Métricas para energía solar en Lanzarote

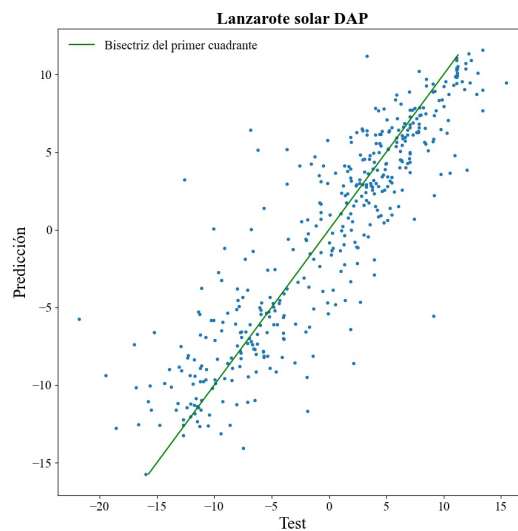


Figura 19: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía solar en Lanzarote.

Para Lanzarote, podemos observar en el **Cuadro 10** que los mejores modelos obtenidos son DA y DAP. Además, al igual que en el caso de Fuerteventura, se observa que el R² de los mejores modelos es muy alto lo que implica que es muy buen modelo, como podemos contrastar en la **Figura 19**.

Eólica			
	RMSE relativo	R ²	BIAS relativo
ZOH	0,56	0,53	0,00
ARMA(6,6)	0,84	-0,34	0,29
DA	0,54	0,61	0,09
SP	0,54	0,48	-0,05
DAP	0,53	0,62	0,07

Cuadro 11: Métricas para energía eólica en Lanzarote

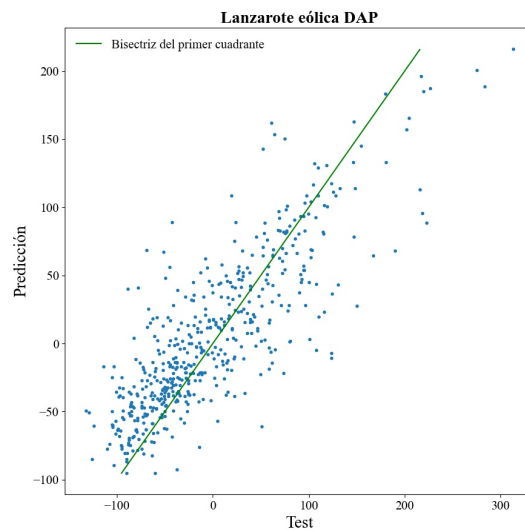


Figura 20: Diagrama de dispersión para las predicciones realizadas por DAP sobre energía eólica en Lanzarote.

Con respecto a la energía eólica en Lanzarote, en el **Cuadro 11** vemos que el mejor modelo para la predicción es el DAP, la red neuronal a la que se le da información sobre la energía generada los días anteriores y la presión atmosférica a nivel del mar. En la **Figura 20** se muestra el diagrama de dispersión obtenido para las predicciones hechas por este modelo, en esta observamos como los puntos correspondientes a energías bajas se encuentran más concentrados cerca de la bisectriz que aquellos que tienen una energía más alta, un comportamiento similar al de los casos mencionados anteriormente para energía eólica.

6. Resumen y conclusiones

Summary

In this chapter a quick recap of what has been done during the project is exposed. Then, from the results obtained in Section 5, some general conclusions are made: for the case of solar energy, ARMA models are in all cases the ones with the lowest RMSE but with a very big BIAS, on the other hand, in terms of R^2 the best model in all cases is the DAP model. For the case of wind energy, all models based on neural networks are better than the ZOH and ARMA but the best one between these three depends on the island. Finally, some ideas to continue with this project and to try to obtain better results for the neural networks are given.

A lo largo de este trabajo se ha realizado un estudio sobre el uso de redes neuronales artificiales para la predicción diaria de energías renovables en las islas Canarias, para ello se han comparado los resultados obtenidos por estas con los obtenidos por el retenedor de orden cero y por modelos ARMA, que son los métodos tradicionales de predicción de series temporales que se usan actualmente en la industria^[11].

Para empezar, dado que en el momento de la realización de este trabajo REE no disponía de una base de datos pública para cada una de las islas en la que descargar los datos, sino que todos los datos se muestran en tablas dinámicas en su página web^[18], se desarrolló un código en *Python* con el que obtener estos datos haciendo *web scraping*.

Con los datos descargados se pasó a tratarlos, calculando primero la energía generada por día en lugar de la potencia cada diez minutos que proporciona REE. A continuación, para tener una idea de como varían a lo largo del año se representó la media de las energías de cada mes en diagramas de cajas. El siguiente paso en el tratamiento de los datos fue hacer una descomposición estacional y eliminar la tendencia para así poder hacer predicciones a un caso aislado en el que se minimiza el efecto de la variación en el número de parques solares o eólicos.

Tras esto, se pasó a aplicar un algoritmo denominado *gridsearch* para relizar una búsqueda exhaustiva entre distintas combinaciones de parámetros para las redes neuronales y así determinar con cuales se obtenía una red neuronal que hiciera buenas predicciones

de energía generada. Este procedimiento se llevó a cabo para redes en las que los valores de entrada eran las energías generadas en los tres días anteriores, otras en las que las entradas eran los valores de presión atmosférica a nivel del mar del mismo día en un entorno de Canarias (condiciones sinópticas generales) y otras que combinaran las energías de los tres días anteriores y las presiones del mismo día.

Una vez determinadas los mejores parámetros de las redes para la distintas predicciones, se pasó a comparar las métricas de los resultados obtenidos con ellas con las de los métodos sencillos. A modo de resumen general, en estas se puede observar que en todos los casos el parámetro de BIAS relativo para el retenedor de orden cero (ZOH) es muy bajo o nulo. Esto era algo de esperar ya que este es solo la media de las diferencias entre predicción y valor real y como las predicciones del ZOH son solo las entradas desplazadas un día, estas diferencias se compensan unas con otras de manera que la media es casi nula. Centrando la atención en las fuentes por separado podemos obtener las siguientes conclusiones generales:

- **Energía solar:** Los modelos ARMA suelen ser los que menor RMSE tienen aunque suelen tener un parámetro de BIAS muy grande, es decir, sobrestiman la producción. Para esta se observa también que las redes cuyas entradas son solo la presión atmosférica a nivel del mar no mejoran mucho con respecto al ZOH. Por último, en lo que al parámetro de correlación R^2 respecta, el modelo vencedor en todos los casos es la red neuronal a la que se le proporciona información sobre la energía generada los días anteriores y sobre la presión atmosférica del propio día, se observa además que este es mucho mayor para las islas más orientales. Esto era de esperar ya que estas son las islas más bajas^[24] y, por lo tanto, las que menos nubosidad tienen, es decir, en estas la energía solar producida a lo largo del año depende en mayor medida del movimiento de la Tierra, un movimiento periódico a lo largo de los años. Las diferencias entre los resultados de DAP y DA son realmente bajas, lo que indica que la adición de más información a las entradas de las redes neuronales, en este caso las presiones a nivel del mar, no mejoran significativamente los resultados frente al uso exclusivo de la serie temporal de la producción de energía solar.
- **Energía eólica:** Por lo general, los modelos ARMA no aportan mejores resultados que el ZOH. Con respecto a los modelos basados en redes neuronales, estos son mejores que los sencillos, aunque la elección de cual es el mejor de entre los tres

depende de la isla elegida. Así, en casi todas las islas las redes neuronales que toman como entradas las energías producidas los días anteriores y las presiones a nivel del mar (DAP) obtienen mejores resultados, aunque, de nuevo, muy similares a las redes que sólo dependen de la serie temporal de la energía eólica (DA). Sólo en dos islas, La Palma y Gran Canaria, los mejores resultados corresponden a las redes que toman como entrada sólo las presiones a nivel del mar (SP), es decir, se basan sólo en las condiciones sinópticas generales en un entorno de la región. Sea cual sea la red elegida, los resultados no son superiores a los obtenidos para la producción solar, excepto para las dos islas más orientales.

Por último, como planes de trabajo futuro, se podría tratar de mejorar los resultados obtenidos con modelos basados en redes neuronales artificiales, sobre todo para los casos de las islas occidentales en las que los resultados no son tan buenos como uno esperaría. Para ello, se podría continuar con los análisis ya realizados y aplicar algoritmos de selección de parámetros (*feature selection*) que tratan de reducir el número de parámetros de entrada determinando cuales de estos son los que más información aportan a la predicción^[25].

También, se podría probar a realizar el estudio con modelos de redes neuronales artificiales más sofisticados como por ejemplo las redes *LSTM* (*Long short-term memory*) que ya fueron probadas en un estudio para la predicción de energía solar en Vietnam^[26] u otras técnicas que permitan realizar análisis de regresión complejos, como *random forest*, etc. Asimismo, se podría explorar la utilización de otras entradas a los métodos de aprendizaje automático, como variables meteorológicas de modelos globales de predicción numérica más directamente relacionadas con la producción de energía renovable, como la velocidad y dirección del viento o estimación de cobertura nubosa.

7. Agradecimientos

Quisiera comenzar agradeciendo a mis tutores, Albano González y Luis Alonso, por haber aceptado guiarme durante la realización de este proyecto y resolver todas las dudas lo antes posible a pesar de la gran carga de trabajo que han tenido.

Deseo también expresar el agradecimiento a mi familia, especialmente a mis padres, que han estado al pie del cañón día a día, celebrando las cosas buenas y ayudando en las malas. También a mis hermanos, ellos han sido mi referente, por ellos me he marcado estas metas y gracias a ellos las he ido superando.

Y, por último, pero no por ello menos importante, a mis compañeros del grupo C.M. y, cómo no, a mis amigos y compañeros de piso, David y Lucía, por haberme acompañado en este corto pero intenso viaje de cuatro años. Probablemente, sin su apoyo incondicional, no habría llegado tan lejos.

Referencias

- [1] Comisión europea. Paquete de medidas sobre clima y energía hasta 2020. https://ec.europa.eu/clima/policies/strategies/2020_es.
- [2] Renewable energy in the european union. https://en.wikipedia.org/wiki/Renewable_energy_in_the_European_Union.
- [3] Agora energiewende and ember (2021): The european power sector in 2020: Up-to-date analysis on the electricity transition, january 2021. <https://ember-climate.org/project/eu-power-sector-2020/>.
- [4] Boletín oficial del estado, 126, de 25 de mayo de 2007, real decreto 661/2007 por el que se regula la actividad de producción de energía eléctrica en régimen especial. recuperado de <https://www.boe.es/buscar/act.php?id=BOE-A-2007-10556&p=20130713&tn=1>.
- [5] Canarias, fuente de energía limpias (28/12/18). recuperado de: <https://www.laprovincia.es/economia/2018/12/29/canarias-fuente-energia-limpias-9394244.html>.
- [6] Red eléctrica de españa, nota de prensa: Las renovables generan en 2020 el 17,5 % del total de la energía eléctrica canaria, la mayor cuota desde que hay registros. recuperado de https://www.ree.es/sites/default/files/07_SALA_PRENSA/Documentos/2021/1203_NP_Avance_Canarias.pdf.
- [7] International Renewable Energy Agency. Renewable power generation costs in 2019. Abu Dhabi. Capítulo 1 pag.21. ISBN: 978-92-9260-244-4.
- [8] Shahram Hanifi, Xiaolei Liu, Zi Lin, and Saeid Lotfian. A critical review of wind power forecasting methods—past, present and future. *Energies*, 13(15), 2020.
- [9] U. Focken, M. Lange, and H.-P. Waldl. Previento - a wind power prediction system with an innovative upscaling algorithm. In *Proceedings of the European Wind Energy Conference (EWEC)*, pages 1–4, 2001.
- [10] W. Chang. A literature review of wind forecasting methods. In *Journal of Power and Energy Engineering*, pages 161–168, 2014.

- [11] Q. Chen and K.A. Folly. Wind power forecasting. *IFAC-PapersOnLine*, 51(28):414–419, 2018. 10th IFAC Symposium on Control of Power and Energy Systems CPES 2018.
- [12] Sharif Atique, Subrina Noureen, Vishwajit Roy, Vinitha Subburaj, Stephen Bayne, and Joshua Macfie. Forecasting of total daily solar energy generation using arima: A case study. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0114–0119, 2019.
- [13] Alexei Serna, Eder Acevedo, and Edgar Serna. Principios de la inteligencia artificial en las ciencias computacionales. *Desarrollo e Innovación en Ingeniería*, page 167, 2017.
- [14] Fuente de la imagen 3a. <http://www.cs.us.es/~fsancho/?e=72>.
- [15] Fuente de la imagen 3b. <https://censorcosmico.blogspot.com/2017/11/creacion-de-una-estructura-de-red.html>.
- [16] R.M. Aguilar, J.M. Torres, and C.A. Martín. Aprendizaje automático en la identificación de sistemas. un caso de estudio en la predicción de la generación eléctrica de un parque eólico. *Revista Iberoamericana de Automática e Informática industrial*, 16(1):114–127, 2019.
- [17] Copernicus Climate Change Service. Climate reanalysis. <https://climate.copernicus.eu/climate-reanalysis>.
- [18] Red Eléctrica de España. Seguimiento de la demanda de la energía eléctrica. <https://demanda.ree.es/visiona/seleccionar-sistema>.
- [19] Selenium documentation. <https://www.selenium.dev/selenium/docs/api/py/api.html>.
- [20] Albano González, Juan C. Pérez, Juan P. Díaz, and Francisco J. Expósito. Future projections of wind resource in a mountainous archipelago, canary islands. *Renewable Energy*, 104:120–128, 2017.
- [21] Prashant Banerjee. Complete guide on time series analysis in python, October 2020. <https://www.kaggle.com/prashant111/complete-guide-on-time-series-analysis-in-python>.

- [22] Gridsearchcv documentation. https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html.
- [23] Randomizedsearchcv documentation. https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html.
- [24] F.J. Pérez Torrado y E. Rodríguez Badiola J.C. Carracedo. Canarias: islas volcánicas intraplaca.
- [25] S. Khalid, T. Khalil, and S. Nasreen. A survey of feature selection and feature extraction techniques in machine learning. In *Proceedings of 2014 Science and Information Conference, SAI 2014*, pages 372–378, 2014. Cited By :284.
- [26] N. Q. Nguyen, L. D. Bui, B. V. Doan, E. R. Sanseverino, D. D. Cara, and Q. D. Nguyen. A new method for forecasting energy output of a large-scale solar power plant based on long short-term memory networks a case study in vietnam. *Electric Power Systems Research*, 199, 2021.

A. Nomenclatura

ANN:	Red Neuronal Artificial
ARIMA:	Modelos autorregresivos integrados de media móvil
ARMA:	Modelos autorregresivos de media móvil
BNN:	Red Neuronal Biológica
BOE:	Boletín Oficial del Estado
DA:	Red neuronal entrenada con la serie temporal de energía generada
DAP:	Red neuronal entrenada con la serie temporal de energía generada y las presiones atmosféricas a nivel del mar
ECMWF:	Centro Europeo de Previsiones Meteorológicas a Plazo Medio
LSTM:	<i>Long short-term memory</i>
REE:	Red Eléctrica de España
RMSE:	Raíz del error cuadrático medio
SP:	Red entrenada con las presiones atmosféricas a nivel del mar
UE:	Unión Europea
ZOH:	Retenedor de orden cero