

Redes Neuronales Artificiales. Aplicación a la regionalización de la precipitación y temperaturas diarias

Nota técnica 34 de AEMET



Juan Andrés García Valero

AUTOR:

JUAN ANDRÉS GARCÍA VALERO

AGENCIA ESTATAL DE METEOROLOGÍA
DELEGACIÓN TERRITORIAL EN LA REGIÓN DE MURCIA



Aviso Legal: los contenidos de esta publicación podrán ser reutilizados, citando la fuente y la fecha, en su caso, de la última actualización


Edita:

© Ministerio para la Transición Ecológica y el Reto Demográfico
Agencia Estatal de Meteorología
Delegación Territorial en la Región de Murcia
Murcia, 2021

Catálogo de Publicaciones de la Administración General del Estado:
<https://cpage.mpr.gob.es>

NIPO: 666-20-028-5
<https://doi.org/10.31978/666-20-028-5>

Agencia Estatal de Meteorología (AEMET)
C/ Leonardo Prieto Castro, 8
28040 Madrid
<http://www.aemet.es/>

 @Aemet_Esp

 <https://www.facebook.com/AgenciaEstataldeMeteorologia>



ÍNDICE GENERAL

1. Introducción	2
2. Redes neuronales artificiales	4
2.1. Estructura de una RNA	4
2.2. Señal de entrada a una neurona	5
2.3. Salida de una neurona. Función de activación	6
2.4. Entrenamiento RNA. Ajuste de pesos	7
3. Datos empleados	11
4. Calibración de la RNA	12
4.1. Arquitectura de red considerada	12
4.2. Validación	13
4.2.1. Precipitación	13
4.2.2. Temperatura	14
5. Conclusiones y discusión	19

1 INTRODUCCIÓN

Las proyecciones de cambio climático constituyen el punto de partida de un gran número de estudios que investigan el impacto que los cambios futuros pueden tener en los ecosistemas naturales y en multitud de sectores socioeconómicos. Las proyecciones derivadas directamente de los modelos globales de clima son en muchas ocasiones insuficientes para los estudios de impacto debido a su baja resolución espacial, siendo necesario disponer de proyecciones de mayor resolución. El aumento de resolución se obtiene básicamente mediante dos tipos de técnicas de regionalización distintas, unas de tipo dinámico y otras de tipo estadístico. Las primeras emplean modelos climáticos regionales de mayor resolución que los modelos globales sobre una región determinada tomando como condiciones de frontera los resultados de los modelos globales. En cambio, las segundas buscan relaciones de tipo estadístico entre variables proporcionadas por los modelos globales, denominadas predictores, y variables de tipo puntual, normalmente observaciones, denominadas predictandos, como puede ser la precipitación, temperatura o el viento sobre una determinada localidad. En general, las técnicas dinámicas requieren de numerosos recursos de computación, siendo su obtención muy costosa, mientras que las técnicas estadísticas presentan una exigencia de cálculo muy inferior. Por contra, el problema de las técnicas estadísticas es que las relaciones estadísticas se calibran durante un período de tiempo muy distinto para el que luego se obtienen las proyecciones de cambio climático, manteniéndose como hipótesis el que dichas relaciones se mantienen en un contexto de cambio climático.

Las redes neuronales artificiales (RNA) son un tipo de técnica estadística capaz de reproducir de manera empírica posibles relaciones no lineales entre un conjunto de variables de entrada y salida, las cuáles podrían asimilarse a los campos predictores y predictandos. Por ello, las RNA han sido empleadas como un método más de regionalización estadística dentro de la familia de los métodos de regresión que componen este tipo de regionalización (Amblar y otros. [2017]). Están basadas en la fisiología del cerebro, de forma que están constituidas por una serie de unidades de cómputo o nodos conectadas entre sí, simulando lo que serían las relaciones sinápticas entre las neuronas cerebrales. Su método de aprendizaje consiste en calibrar las conexiones entre las neuronas mediante un proceso de ajuste iterativo. La duración de este ajuste se extenderá más o menos en el tiempo en función de la complejidad de las relaciones entre predictores y predictando, del algoritmo de aprendizaje adoptado, así como de la

tipología o estructura de red neuronal considerada. De forma general, las RNA exigen un mayor coste computacional que otro tipo de técnicas de regresión estadísticas, siendo este factor el que probablemente haya condicionado su menor aplicación a estudios de este tipo. Por otro lado, su potencial para detectar relaciones no lineales las hace muy atractivas para su aplicación a variables más complejas como puede ser la precipitación.

En la actualidad, solo un número reducido de trabajos ha empleado las RNA como método de regionalización de la temperatura y precipitación a escala diaria. Así, en Dibike y Coulibaly [2006] se muestra la eficiencia de las RNA para su aplicación en el downscaling de las temperaturas extremas y precipitaciones diarias en la región de Quebec (Canadá). En Trigo y Palutikof [1999] se presenta un estudio comparativo del funcionamiento del uso de esta técnica frente a otros métodos de regresión en relación a las temperaturas extremas diarias en Portugal. Los resultados de este trabajo evidenciaron que las RNA entrenadas eran capaces de obtener relaciones entre predictores y predictandos de mayor consistencia física que otro tipo de técnicas de regresión empleadas de forma más habitual como es el caso de la regresión múltiple. Recientemente, las RNA se han usado también para el aprendizaje de las relaciones entre modelos de clima globales y regionales (Gómez-Navarro y otros. [2019]), de modo que es posible entrenar una RNA empleando como predictando la salida del modelo regional y como predictores los campos procedentes del modelo global que sirvieron para alimentar al modelo regional. Entrenando una RNA de esta forma es posible aplicarla a simulaciones centenarias obtenidas con el modelo global, permitiendo así una regionalización de dichas simulaciones con un coste computacional muy inferior al que sería necesario mediante la simulación dinámica regional. Hay tres factores fundamentales que justifican la necesidad de trabajos como el que se presenta aquí: la escasez de proyecciones regionalizadas mediante RNA en el territorio español, las ventajas que las RNA pueden ofrecer en relación a otro tipo de técnicas, y la mayor capacidad de cálculo que hoy en día tienen los procesadores.

El objetivo de este trabajo consiste en presentar el método mediante el cuál se ha calibrado un elevado número de RNA para su uso en la generación de escenarios regionalizados de cambio climático. Dicho entrenamiento se ha efectuado a escala diaria sobre un importante número de estaciones distribuidas por todo el territorio español. Con este trabajo se pretende engrosar el catálogo de escenarios regionalizados que ya dispone la Agencia Estatal de Meteorología (AEMET) obtenido mediante otras dos técnicas de regionalización estadísticas: análogos y regresión lineal. Para ello, en la Sección 2 se explica en términos generales el funcionamiento de una RNA. En la sección 3 se describen los datos empleados. En el apartado 4 se explica la estructura de red utilizada y se muestran una serie de resultados de validación de las RNA calibradas con dicha estructura. Por último, en la sección 5 se presentan las principales conclusiones.

2 REDES NEURONALES ARTIFICIALES

2.1 ESTRUCTURA DE UNA RNA

El perceptrón es la unidad básica que alimenta lo que hoy en día se conoce como aprendizaje profundo, denominándosele también neurona artificial. Un perceptrón es una máquina de aprendizaje muy simple. Puede incluir algunas entradas, cada una de las cuales tiene un peso para indicar su grado de importancia, y generar una decisión de salida de tipo binario, 0 o 1 (Figura 2.1). Cuando el perceptrón se combina con muchos otros forma lo que se conoce como red neuronal artificial (RNA), la cuál, teóricamente, puede responder a cualquier pregunta siempre y cuando cuente con suficientes datos de entrenamiento y capacidad de cálculo.

La estructura o topología de una RNA se define en función de la forma en que se encuentran distribuidas las neuronas. Una de las estructuras más empleadas es la distribución en capas, denominándose a este tipo multicapa (multilayer perceptrón en su denominación anglosajona, en adelante MLP). En la Figura 2.2 se muestra un ejemplo de RNA formada por 3 capas. La primera capa (izquierda) es la de entrada donde cada neurona gestiona la información de una variable de entrada a la red. En el caso de la aplicación de una RNA como método estadístico de regionalización, en esta capa se recibiría la información de los campos predictores. La última capa (derecha) constituye la capa

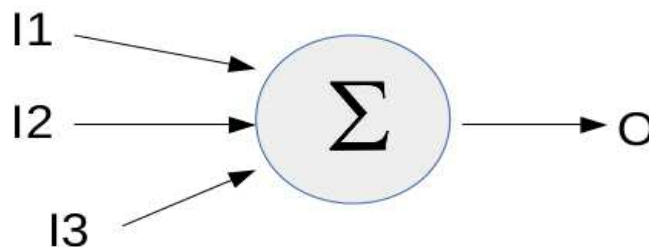


Figura 2.1: Esquema de una simple neurona (perceptrón) con tres entradas y una salida.

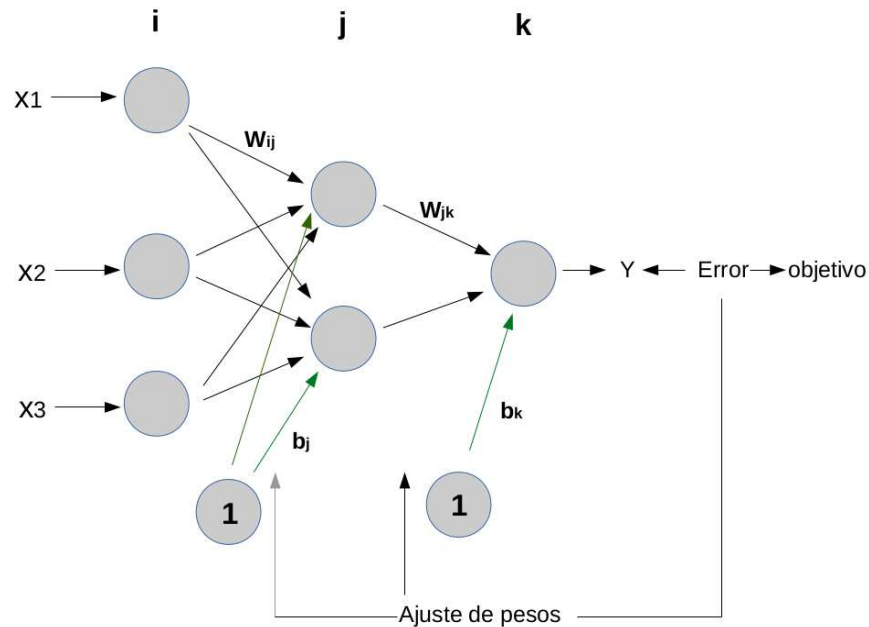


Figura 2.2: Esquema de una RNA de 3 capas

de salida de la red neuronal, que puede estar constituida por una o varias neuronas que gestionan la salida. En el caso que se quisiera entrenar la RNA para determinar la relación entre una variable predictando con sus campos predictores esta capa de salida estaría formada por una sola neurona. La capa intermedia, también denominada capa oculta, recibe la información de las capas de entrada, gestionándola y redirigiéndola hacia la capa de salida. Así, la información fluye desde la entrada hacia la salida por lo que a este tipo de MLP se le denomina con el término anglosajón *feedforward*.

En este trabajo se ha utilizado la estructura MLP. Dicha topología es la que tradicionalmente se ha empleado en la mayoría de trabajos desarrollados hasta la fecha para fines muy similares al que pretendemos aquí (Trigo y Palutikof [1999], Dibike y Coulibaly [2006], entre otros). Esta estructura parece suficiente para que esta pueda aprender de las relaciones entre predictores y predictandos [Hornik y otros., 1989], aunque nuevas investigaciones apuntan a la necesidad de considerar RNA más complejas, con un mayor número de capas ocultas, que permitan un aprendizaje más profundo de las relaciones [Baño-Medina y Gutiérrez, 2018], especialmente en variables como la precipitación.

2.2 SEÑAL DE ENTRADA A UNA NEURONA

Definida la topología de la RNA, número de capas y neuronas en cada una de ellas, ésta debe ser entrenada para la resolución de un problema determinado. En una RNA tipo MLP las neuronas de las diferentes capas están conectadas con todas las neuronas de sus

capas adyacentes, definiéndose estas conexiones mediante pesos. Así, el entrenamiento de la red consiste en el ajuste de dichos pesos representados en la Figura 2.1 por W_{ij} . Antes de describir cómo se lleva a cabo el proceso de ajuste hay que tener en cuenta algunas consideraciones del funcionamiento de una simple neurona.

La información que le llega a una determinada neurona ubicada en una capa determinada j , es la señal integrada de la salida devuelta por todas las neuronas de su capa anterior i . La señal de entrada definida en la ecuación 2.1 constituye una combinación lineal de las salidas de las neuronas anteriores ponderadas por los pesos que definen las relaciones entre las neuronas.

$$I_j = \sum_{i=1}^{N_i} (W_{ij}A_i) + b_j \quad (2.1)$$

donde N_i es el número de neuronas de la capa anterior; W_{ij} el peso de una determinada neurona en la capa anterior i con dicha neurona localizada en la capa j ; A_i es la salida de una neurona determinada en la capa anterior; por último b_j es el peso de la neurona bias con la neurona j que recibe su señal.

2.3 SALIDA DE UNA NEURONA. FUNCIÓN DE ACTIVACIÓN

A la señal integrada de entrada a una determinada neurona, ésta le aplica una función conocida como función de activación, por ello, a su salida suele denominársele también *activación*. Dicha salida constituirá una de las entradas a las neuronas de la siguiente capa. Una característica necesaria que debe tener la función de activación es que sea continua y diferenciable. Las funciones de activación pueden ser de tipo escalón, lineales o no lineales, siendo éstas últimas las que se utilizan para entrenar RNA que resuelven problemas complejos. De entre las no lineales una de las más empleadas es la función logística (ecuación 2.2). Esta función cumple las características descritas (Figura 2.3) y tiene una sencilla derivada (ecuación 2.3) que la hace muy útil para el proceso de ajuste de los pesos entre las neuronas. Además de la función logística, hay otras que también suelen emplearse como funciones de activación (tangente hiperbólica, unidad lineal rectificadora conocida como RELU, otras funciones con modificaciones de RELU, etc.). Para este trabajo se ha empleado la función logística. El uso de neuronas bias es habitual para desplazar hacia la izquierda o hacia la derecha del eje de abscisas la función de activación. Este desplazamiento puede permitir que valores de entrada, por ejemplo un valor 5 tome un valor de activación 0, lo cuál puede resultar muy útil cuando se ajusta la señal de salida de la RNA a un resultado conocido. Así, es frecuente incluir en cada capa una neurona bias que almacena el valor de 1 multiplicado por el peso de dicha neurona con cada una de las neuronas de la capa sobre la que actúa dicha neurona bias (término b_j de la ecuación 2.1). Dicho peso también se ajusta durante la etapa de entrenamiento.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

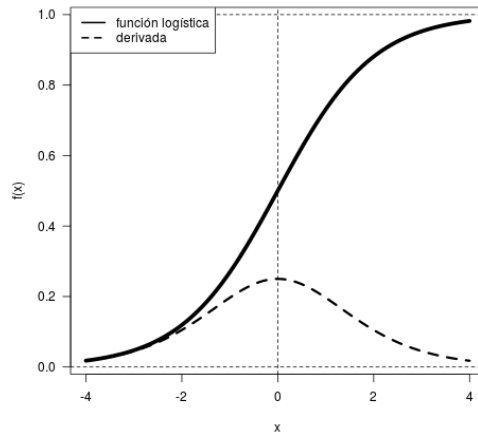


Figura 2.3: Funcion logística (línea continua) y su derivada (línea discontinua)

$$\frac{df}{dx} = f(x)(1 - f(x)) \quad (2.3)$$

2.4 ENTRENAMIENTO RNA. AJUSTE DE PESOS

Una forma de entrenar una RNA es mediante el procedimiento conocido como supervisión, es decir, empleando un conjunto de valores de entrada y salida conocidos. De este modo, las salidas de la red neuronal se van comparando con los valores conocidos determinándose así un error del valor de la salida. Generalmente, este error se define a partir del error cuadrático medio (ecuación 2.4).

$$E = \frac{1}{n} \sum_{t=1}^n (T_t - A_t)^2 \quad (2.4)$$

donde n es el número de etapas temporales que se emplearán en la calibración de los pesos; A_t es la activación o salida proporcionada por la RNA para un determinado instante de tiempo t ; T_t es la respuesta conocida en ese mismo instante de tiempo t .

El error de la RNA depende del valor de los pesos, de la muestra de entrenamiento, así como de la función de activación. El objetivo del método de entrenamiento es minimizar la función error de salida, necesitándose para ello un ajuste fino de los distintos pesos. Uno de los métodos habitualmente empleados para ello es el de propagación hacia atrás, más conocido por su término anglosajón *backpropagation* [Rumelhart y otros., 1986]. Su nombre proviene del hecho de que este método ajusta primero los pesos de las neuronas más próximas a la capa de salida, y a partir de ellos va procediendo al ajuste de los pesos de las capas superiores, y así hasta llegar a los pesos entre las neuronas de entrada y de la primera capa oculta (Figura 2.2). Para este proceso de ajuste el método utiliza la técnica del gradiente descendente, denominada así porque

es la dirección negativa del gradiente de la función error aquella se sigue para alcanzar su valor mínimo. Así, la función error dependerá de cada uno de los pesos que definen las distintas conexiones entre las neuronas. Siguiendo el esquema representado en la Figura 2.2 esta dependencia puede expresarse de la forma $E = f(W_{ij}, W_{jk}, b_j, b_k)$, de modo que el mínimo de la función error se alcanzará cuando se alcancen los valores mínimos de sus diferentes derivadas parciales (ecuación 2.5).

$$\frac{dE}{dW} = \sum_{j=1}^{N_j} \sum_{i=1}^{N_i} \frac{\partial E}{\partial W_{ij}} + \sum_{k=1}^{N_k} \sum_{j=1}^{N_j} \frac{\partial E}{\partial W_{jk}} + \sum_{j=1}^{N_j} \frac{\partial E}{\partial b_j} + \sum_{k=1}^{N_k} \frac{\partial E}{\partial b_k} \quad (2.5)$$

donde $N_{i,j,k}$ es el número de neuronas de las capas, i, j y k, respectivamente; W_{ij} son los pesos entre una determinada neurona en la capa i y otra neurona de la capa j; $b_{j,k}$ son los pesos bias aplicado a una determinada neurona en la capa j y k, respectivamente. De este modo el método de backpropagation va calculando los diferentes gradientes definidos en la ecuación anterior. Para ello emplea la regla de la cadena, considerando que el peso entre dos neuronas tiene una dependencia respecto de la señal de entrada que recibe la neurona en la capa inferior.

A continuación, se muestra un ejemplo del cálculo del gradiente de la función error respecto al peso W_{jk} entre una neurona de la capa j y otra neurona de la capa de salida k, tal y como aparece en el esquema de la Figura 2.2. Teniendo en cuenta que este peso va a depender de la entrada a la neurona k (I_k) aplicando la regla de la cadena el gradiente puede expresarse de la forma:

$$\frac{\partial E}{\partial W_{jk}} = \frac{\partial E}{\partial I_k} \frac{\partial I_k}{\partial W_{jk}} \quad (2.6)$$

Considerando la ecuación 2.1, podemos expresar $\frac{\partial I_k}{\partial W_{jk}} = A_j$. Sustituyendo ésto en la ecuación 2.6 obtendremos la ecuación 2.7.

$$\frac{\partial E}{\partial W_{jk}} = A_j \frac{\partial E}{\partial I_k} \quad (2.7)$$

Considerando además que la señal de activación de la neurona k, A_k , depende de la función de activación escogida, $A_k = g(I_k)$, y aplicando de nuevo la regla de la cadena, la ecuación anterior se puede expresar de la siguiente forma:

$$\frac{\partial E}{\partial W_{jk}} = A_j \frac{\partial E}{\partial A_k} \frac{\partial A_k}{\partial I_k} \quad (2.8)$$

La condición de la naturaleza diferenciable de la función de activación se justifica ahora para resolver el término $\frac{\partial A_k}{\partial I_k}$ de la ecuación 2.8. Por otro lado, el cálculo sencillo que presenta esta derivada cuando se usa la función logística (ver ecuación 2.3) justifica su uso en el entrenamiento de una RNA, y especialmente cuando el número de pesos a

ajustar es muy alto. Con estas consideraciones la ecuación 2.8 se puede expresar de la siguiente manera:

$$\frac{\partial E}{\partial W_{jk}} = A_j A_k (1 - A_k) \frac{\partial E}{\partial A_k} \quad (2.9)$$

donde los términos A_j y A_k son conocidos, y en el caso de la neurona de salida k , el término $\frac{\partial E}{\partial A_k}$ también es conocido puesto que es la diferencia entre el valor de salida esperado y la salida de esta neurona A_k .

Determinados los gradientes entre las neuronas de las capas j y k el método de entrenamiento ajusta los nuevos pesos mediante la siguiente ecuación:

$$W_{jk}^{t+1} = W_{jk}^t - \eta \frac{\partial E}{\partial W_{jk}} \quad (2.10)$$

donde η es la constante de aprendizaje y debe definirse previamente. La elección de esta constante es importante pues un valor muy pequeño provocaría una convergencia muy lenta del algoritmo de entrenamiento, mientras que un valor elevado podría provocar que no se alcanzase nunca la convergencia.

Una vez modificados los pesos entre las neuronas de las capas j y k , se determinarían los gradientes entre las neuronas de las capas superiores, i y j , en el caso del esquema representado en la Figura 2.2. De este modo, el gradiente para un determinado peso entre dos neuronas ubicadas en la capa i y j se expresaría de forma similar al de la ecuación 2.9 reemplazando el índice k por j y el índice j por i . Así, el último término de esta ecuación puede expresarse considerando la dependencia I_k respecto de A_j , quedando éste de la forma:

$$\frac{\partial E}{\partial A_j} = \sum_{k=1}^{N_k} \frac{\partial E}{\partial I_k} \frac{\partial I_k}{\partial A_j} \quad (2.11)$$

Teniendo en cuenta la ecuación 2.1 para una neurona de la capa k , el término $\frac{\partial I_k}{\partial A_j}$ se puede expresar de la forma:

$$\frac{\partial I_k}{\partial A_j} = W_{jk} \quad (2.12)$$

Mientras que el término $\frac{\partial E}{\partial I_k}$ considerando la dependencia de A_k con respecto a I_k y la derivada de la función de activación, el término se puede expresar como:

$$\frac{\partial E}{\partial I_k} = \frac{\partial E}{\partial A_k} \frac{\partial A_k}{\partial I_k} = \frac{\partial E}{\partial A_k} A_k (1 - A_k) \quad (2.13)$$

De modo que reemplazando las ecuaciones 2.13 y 2.12 en la ecuación 2.9 y considerando esta última ecuación aplicada a las capas i y j , se tiene que el gradiente es:

$$\frac{\partial E}{\partial W_{ij}} = A_i A_j (1 - A_j) \sum_{k=1}^{N_k} A_k (1 - A_k) W_{jk} \frac{\partial E}{\partial A_k} \quad (2.14)$$

Donde todos los términos son conocidos. A partir de aquí, y aplicando de nuevo la ecuación 2.11 se obtendrían los nuevos pesos W_{ij} .

El ajuste de pesos o entrenamiento de la RNA supone llevar a cabo un proceso iterativo como el que acaba de describirse. Para la primera iteración es necesario inicializar los pesos, no conocidos *a priori*, siendo habitual inicializarlos de forma aleatoria. El proceso de ajuste de estos pesos se irá repitiendo una y otra vez hasta alcanzar un mínimo de la función error, el cuál se producirá cuando el gradiente general de la función error respecto a todos los pesos (ecuación 2.5) durante dos iteraciones consecutivas no experimente un cambio significativo.

3 DATOS EMPLEADOS

Las RNA se han entrenado con datos observados de precipitación y temperaturas máximas y mínimas diarias procedentes de estaciones de la Agencia Estatal de Meteorología. Así, se dispone de un total de 2323 estaciones de precipitación y 374 de temperatura. Las estaciones se encuentran distribuidas por toda el territorio español y fueron seleccionadas de entre otras muchas tras superar unos controles de calidad y completud (Brunet y otros. [2008]). Dichas estaciones son las mismas que se han empleado en las dos técnicas de regionalización estadísticas desarrolladas previamente en AEMET ([Amblar y otros., 2017]). Como predictores se ha usado un conjunto de variables de reanálisis de NCEP/NCAR con una resolución de 2.5° (Kalnay y otros. [1996]). Dichos predictores son los mismos que se emplearon en el método de regresión múltiple para la obtención de los escenarios regionalizados. La relación de predictores para los diferentes predictandos puede verse en la Tabla 3.1.

El entrenamiento supervisado de las RNA requiere el uso de variables de entrada y salida conocidas durante un tiempo suficientemente largo (Sec. 2.4). Posteriormente, el funcionamiento de las RNA calibradas debe ser evaluado en un período distinto al de calibración con el fin de determinar sus limitaciones. En nuestro caso, tanto para la precipitación como para la temperatura se ha empleado para la calibración la información de campos predictores y predictandos disponibles en el período 1961-1990, mientras que para la validación se ha considerado el período 1991-2000. Previamente, todos los predictores han sido estandarizados considerando los valores de su media y desviación estandar correspondientes al período de calibración.

Predictandos	Predictores superficie	Predictores altura (500, 700 y 850 hPa)
Precipitación	. Viento meridional . Viento zonal	. Viento meridional . Viento zonal
Temperatura	. Presión a nivel del mar	. Temperatura

Tabla 3.1: Predictores empleados para la temperatura y la precipitación.

4 CALIBRACIÓN DE LA RNA

4.1 ARQUITECTURA DE RED CONSIDERADA

En este trabajo la topología que se ha considerado para las RNA es aquella que responde a un esquema MLP de 3 capas con dos neuronas bias (Figura 2.2). En la primera y segunda capa se ha empleado para cada neurona la función logística como función de activación, mientras que en la capa de salida se ha considerado una función lineal. En el caso de la precipitación se han calibrado dos redes neuronales para cada punto, una para determinar la ocurrencia y la otra para estimar la cantidad de precipitación.

Como algoritmo de entrenamiento se ha considerado el de *resilient backpropagation* (en adelante RPROP) que es una variante del método *backpropagation* explicado en la Sección 2.4. Este se explica con detalle en Riedmiller y Braun [1993] siendo su principal diferencia con respecto al de *backpropagation* que no requiere la definición de una constante de aprendizaje igual para todos los pesos (ecuación 2.11). A diferencia del método *backpropagation*, RPROP usa una constante de aprendizaje variable y dependiente de cada peso que no necesita ser inicializada. De este modo los pesos se ajustan atendiendo al signo de la derivada parcial de la función error (ecuación 4.1) en lugar de ajustarse considerando el valor de dicha derivada como sucede en *backpropagation*.

$$W_{jk}^{t+1} = W_{jk}^t - \eta_{jk}^t \left(\text{sign} \frac{\partial E}{\partial W_{jk}} \right) \quad (4.1)$$

Resilient backpropagation permite una convergencia más rápida del algoritmo de entrenamiento ya que en regiones donde el gradiente de la función error es muy pequeño la constante de aprendizaje aumenta su valor siempre que la derivada parcial mantenga su signo. Por el contrario, si la derivada cambia de signo, la constante disminuirá de valor indicando así que el mínimo de la función error se ha sobrepasado al emplear un valor para la constante demasiado elevado.

Un elemento necesario para definir la arquitectura de la RNA es el número de neuronas a considerar en la capa oculta, siendo este número *a priori* desconocido. Para inferir su valor se efectuaron previamente una serie de pruebas con estaciones pertenecientes a diferentes regiones climáticas, entrenando arquitecturas entre 1 y 25 neuronas en la capa oculta. Posteriormente, las RNA entrenadas se validaron frente a observaciones. Para

la validación se consideraron diferentes métricas: la correlación para la precipitación, y el error cuadrático medio para la temperatura. Los resultados mostraron que el uso de más de 10 neuronas no mejoraba significativamente los resultados obtenidos con un número menor de ellas. Si a esto se añade que el tiempo para calibrar la RNA aumenta considerablemente a medida que lo hace el número de neuronas empleadas, y que además se pretende calibrar un número muy elevado de RNA, se optó por entrenar para cada punto de observación 10 redes, cada una de ellas con un número de neuronas distinto entre 1 y 10. Finalmente, se seleccionó como RNA para cada punto observacional a aquella que mostró el mejor resultado de validación.

En cuanto a la red neuronal entrenada para determinar la ocurrencia de precipitación, ésta emplea como entradas los predictores de la Tabla 3.1 y su estructura está compuesta por una neurona en su capa oculta y otra en su capa de salida. Dado que el valor de la salida de la red está comprendido entre 0 (no ocurrencia) y 1 (ocurrencia), ésta salida se interpreta como una probabilidad de modo que valores próximos a la unidad se corresponden con aquellos días donde la probabilidad es mayor de que sean lluviosos. Para definir el umbral de salida de la red neuronal a partir del cuál se clasificaría un día como lluvioso o no, se ha generado una curva ROC correspondiente al período de calibración. Esta curva se ha obtenido considerando 1000 valores entre 0 y 1. De este modo se seleccionó el umbral cuyo ratio entre el porcentaje de aciertos y de falsas alarmas resultó mayor.

Para el calibrado de las RNA y su posterior uso como método predictivo se utilizaron las funciones *neuralnet* y *compute*, respectivamente, las cuáles están implementadas en la librería *neuralnet* (Fritsch y Guenther [2016]) del proyecto de software libre **R**.

4.2 VALIDACIÓN

4.2.1. Precipitación

La precipitación se ha validado considerando diferentes aspectos. El primero de ellos evalúa la ocurrencia del evento de precipitación mediante la tasa de aciertos y falsas alarmas. Otro aspecto analizado ha sido la desviación porcentual de la precipitación estimada frente a la observada. También se ha evaluado la proporción de la varianza observada que se explica mediante los valores estimados. El último aspecto considerado ha consistido en el análisis de la similitud temporal de las series observadas y estimadas, usando para ello el coeficiente de correlación obtenido entre ambas series. La ocurrencia de evento y desviación porcentual media se han evaluado para la escala diaria, mientras que la evaluación del resto de aspectos se ha llevado a cabo también para la escala mensual.

Respecto a la ocurrencia, se ha obtenido un valor mediana de tasa de aciertos y falsas alarmas alrededor del 80 y 20 %, respectivamente (Figura 4.1). Los patrones espaciales de ambos parámetros son complementarios, observándose en términos generales una mejor estimación de la ocurrencia en la mitad suroeste del país.

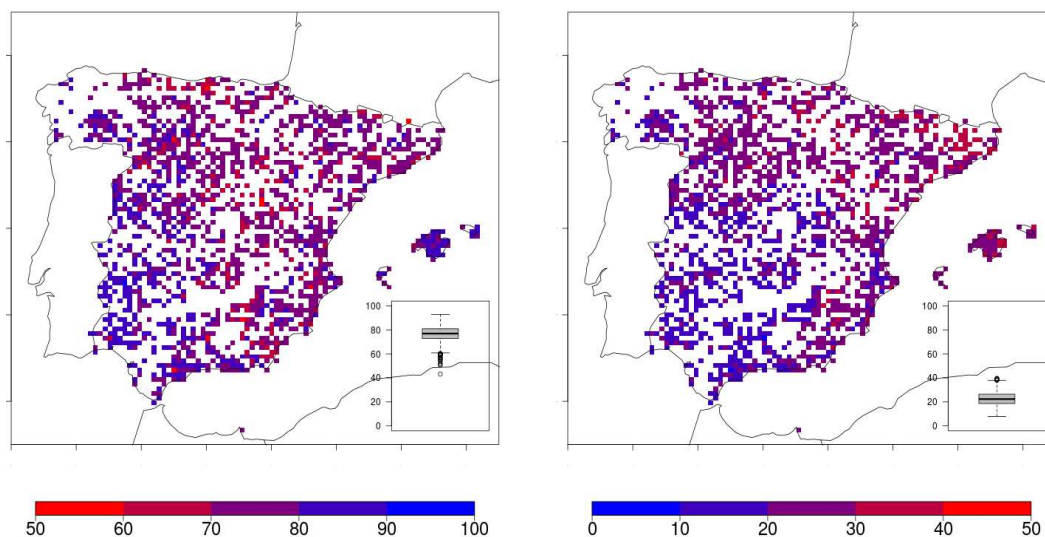


Figura 4.1: Tasa de aciertos (izquierda) y falsas alarmas (derecha), en %, sobre la ocurrencia de evento de precipitación diaria. En ambas gráficas se representa en el cuadro inferior derecho un gráfico de tipo boxplot que resume estadísticamente los resultados obtenidos para el conjunto de todas las estaciones.

En relación a la desviación porcentual de las cantidades de precipitación estimadas frente a las observadas, se aprecia una subestimación del orden de un 20 %, representando este dato el valor de la mediana obtenido para todo el conjunto de puntos (Figura 4.2).

En la Figura 4.3 se muestra una composición de imágenes con los patrones de validación de la correlación (gráficos en la columna de la izquierda) y de la varianza explicada (derecha), en la escala diaria (fila superior) y mensual (fila inferior). En general, a escala diaria los resultados de validación son significativamente peores que en la escala mensual. Así, a escala diaria la correlación mediana es de 0.55, mientras que la varianza está en torno al 25 %. Por contra, en la escala mensual la correlación y la varianza son del orden de 0.75 y 60 %, respectivamente. En ambas escalas temporales, los mejores resultados se observan en las regiones donde la influencia atlántica es mayor, así como en las regiones montañosas; en cambio, los resultados empeoran en las regiones de la vertiente mediterránea.

4.2.2. Temperatura

Para la validación de la temperatura se han empleado métricas similares a la precipitación como la correlación y la varianza explicada. No obstante, en este caso se han evaluado también el bias y la raíz cuadrada del error cuadrático medio (rmse). Al igual que en la precipitación, se ha considerado tanto la escala diaria como la mensual, a excepción del bias pues según su definición, error medio de las desviaciones, es igual en

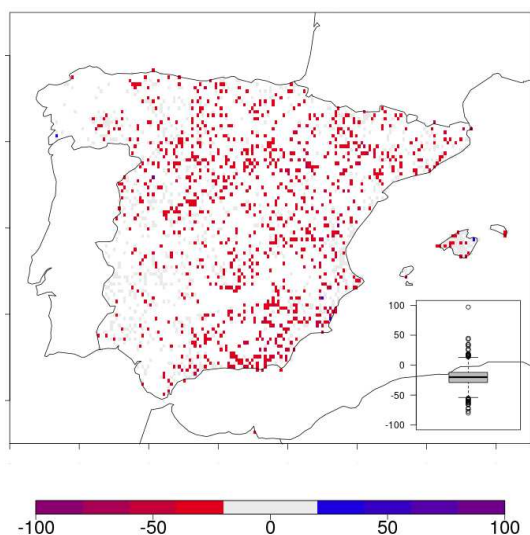


Figura 4.2: Desviación porcentual de la precipitación estimada frente a la observada.

ambas escalas puesto que la temperatura media mensual se obtiene como promedio de las series diarias.

Los resultados evidencian un bias prácticamente nulo para la temperatura máxima ($0.07\text{ }^{\circ}\text{C}$), y algo superior para la temperatura mínima ($0.29\text{ }^{\circ}\text{C}$). En cuanto a los patrones espaciales representados en la Figura 4.4 no se aprecia ningún patrón físico reconocible relacionado con la orografía o la continentalidad.

En relación al rmse, correlación y varianza explicada, los resultados son ligeramente mejores para la temperatura máxima, especialmente en las últimas dos métricas. Así, a escala diaria el rmse es de 2.7 y $2.5\text{ }^{\circ}\text{C}$ para las máximas y mínimas, respectivamente. La correlación toma valores de 0.93 y 0.90 , respectivamente, y la varianza explicada del 86.4 y 81.0% . A escala mensual todas las métricas mejoran, tanto para las temperaturas máximas como para las mínimas, aunque en esta escala son ligeramente mejores para la temperatura mínima. Así, en el caso de las máximas y las mínimas, el rmse es de 1.4 y $1.1\text{ }^{\circ}\text{C}$, la correlación de 0.98 y 0.99 , y la varianza del 96.3 y 96.8% , respectivamente. Todos estos resultados se pueden ver de forma detallada en las Figuras 4.5 y 4.6 donde se representan los patrones obtenidos, en general muy homogéneos, para las distintas variables, métricas (columnas) y escalas temporales (filas).

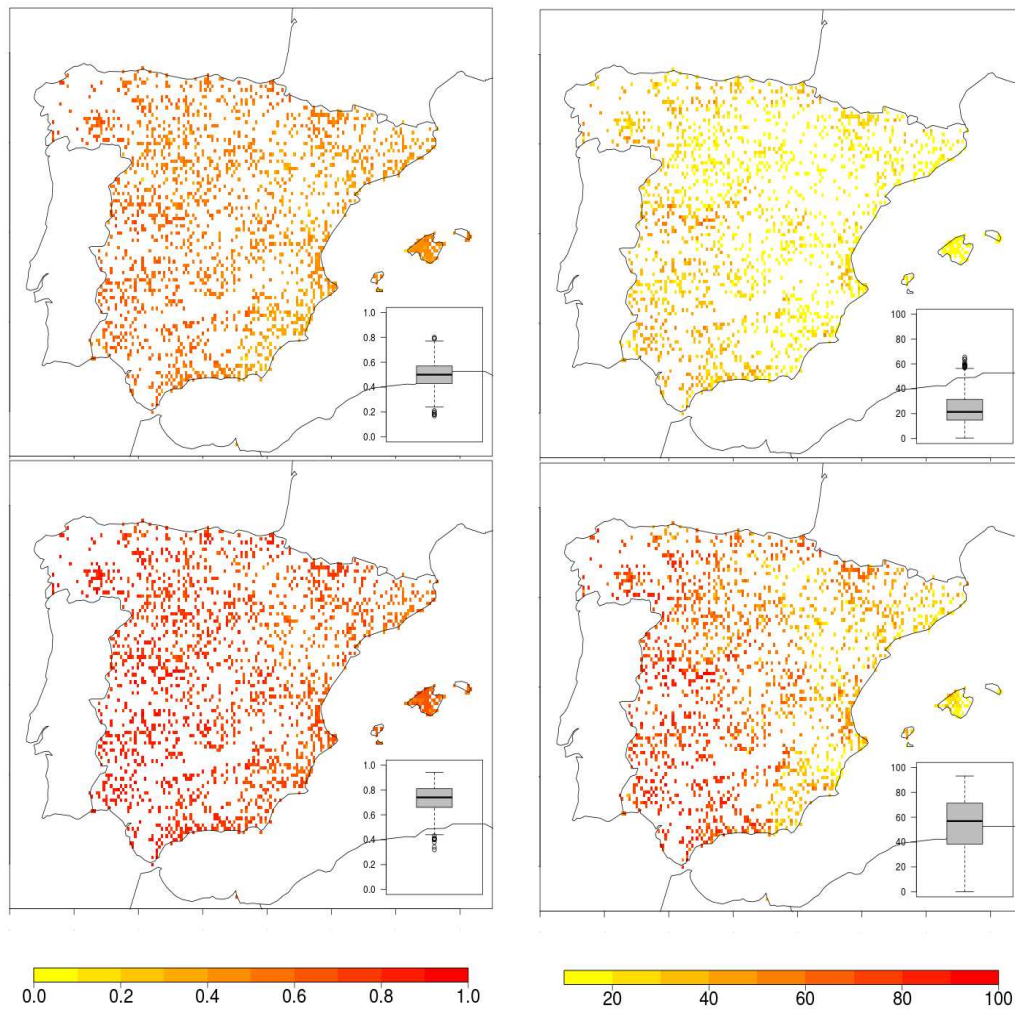


Figura 4.3: Correlación (columna izquierda) y varianza explicada (columna derecha) de la precipitación. La fila superior representa la escala diaria mientras que la inferior se refiere a la escala mensual.

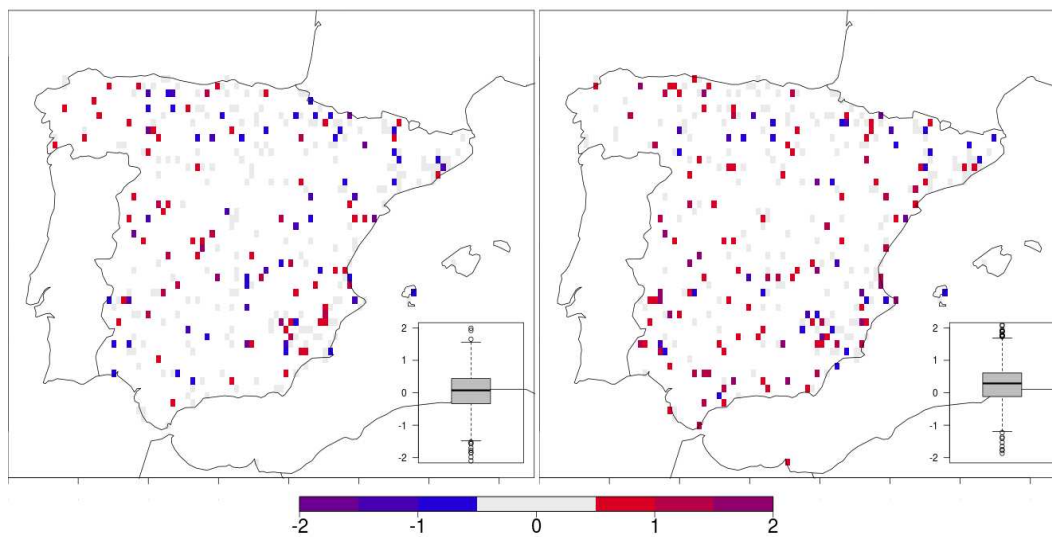


Figura 4.4: Bias (°C) obtenidos para la temperatura máxima (izquierda) y mínima (derecha).

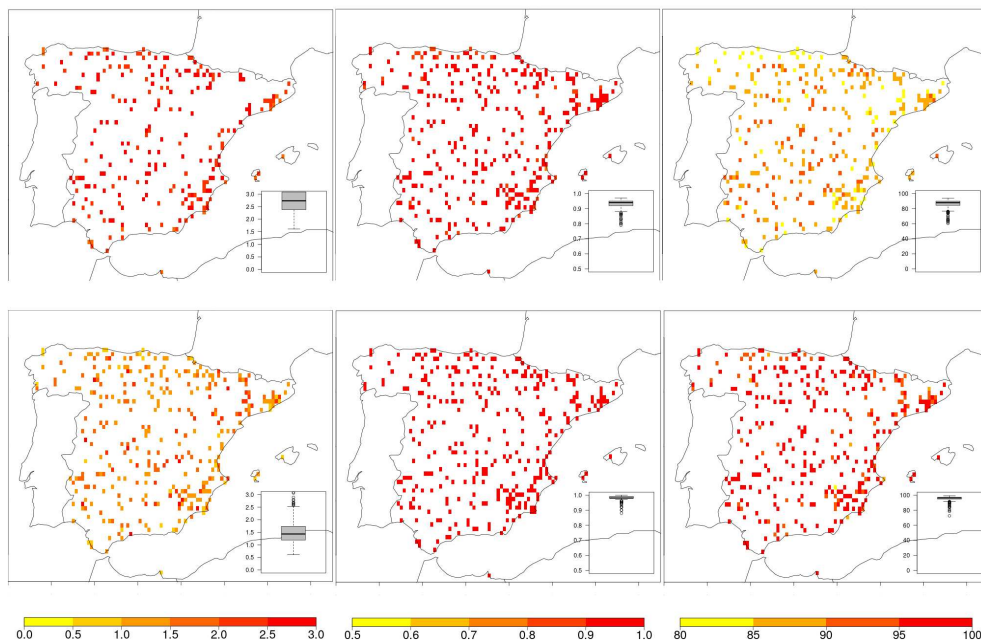


Figura 4.5: Patrones de validación del rmse (columna izquierda), correlación (columna central) y varianza explicada (columna derecha) obtenidos para la temperatura máxima. La fila superior representa la escala diaria mientras que la inferior la escala mensual.

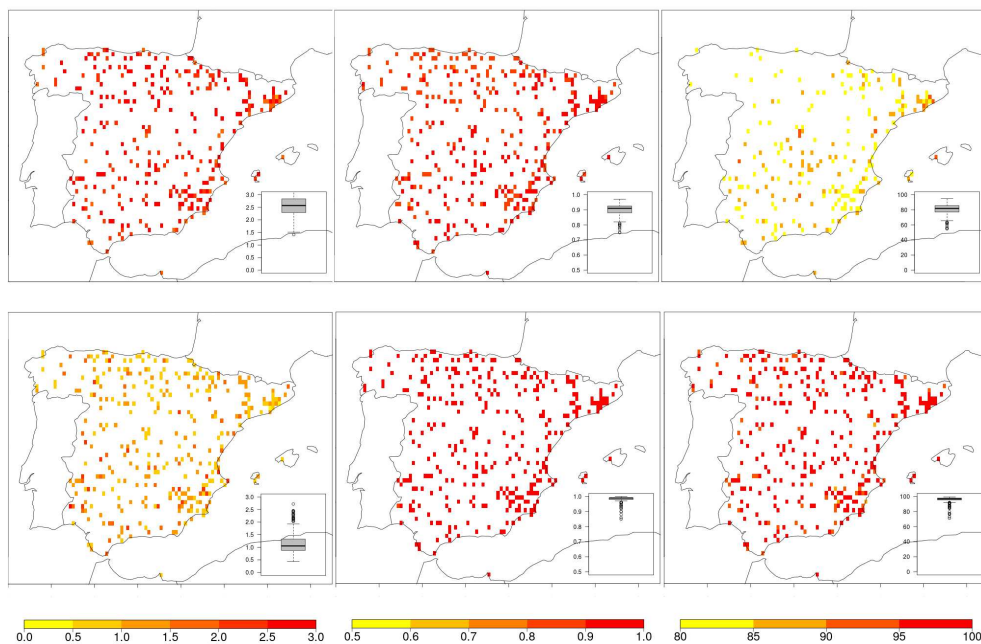


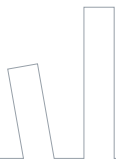
Figura 4.6: Idem que Figura 4.5 pero para la temperatura mínima.

5 CONCLUSIONES Y DISCUSIÓN

En este trabajo se presenta la metodología seguida para el entrenamiento de redes neuronales artificiales llevado a cabo sobre un número elevado de puntos observacionales sobre el territorio español. La calibración se ha aplicado a las variables precipitación y temperaturas máxima y mínima diaria.

Los resultados de validación obtenidos muestran mejores resultados para las temperaturas que para la precipitación. Así, mientras que para las temperaturas la varianza explicada a escala diaria se sitúa entre el 80 y 90 %, en el caso de la precipitación ésta apenas alcanza el 25 %. Este resultado evidencia la dificultad de modelizar la precipitación a escala diaria, especialmente en aquellas regiones cuya pluviometría presenta un régimen muy irregular y que en gran medida depende de factores de pequeña escala, tal y como sucede en amplias regiones de nuestro país, especialmente de la fachada mediterránea. Por otro lado, y a pesar de la mayor complejidad que supone la modelización de la precipitación diaria en comparación con la temperatura, el valor relativamente bajo que presenta en promedio la varianza explicada puede estar manifestando ciertas limitaciones de la metodología. Una de ellas podría estar relacionada con los predictores empleados. La inclusión de algún otro predictor relacionado con la humedad podría tal vez mejorar los resultados. No obstante, su uso podría resultar desaconsejable en la obtención de proyecciones climáticas a muy largo plazo debido al hecho de que la humedad es una variable altamente parametrizable en los modelos globales de clima. Otra limitación podría deberse a que la estructura multicapa escogida para las RNA resulte insuficiente en cuanto al número de capas ocultas. Aumentar su complejidad puede ayudar a que las RNA adquieran un aprendizaje más profundo de las relaciones físicas entre predictores y predictandos (Baño-Medina y Gutiérrez [2018]). Sin embargo, el uso de este tipo de redes podría dificultar la operatividad de trabajos como éste, cuyo fin es su aplicación a un número ingente de puntos observacionales. Otra causa relacionada con la estructura de la RNA empleada podría estar en el número de neuronas utilizadas. Este hecho fue analizado mediante validación cruzada probando distinto número de neuronas evidenciándose que un número superior a 10 neuronas tendía al sobreajuste en la estructura tricapa considerada.

Las RNA entrenadas en este trabajo se pretende sirvan de base para la generación de escenarios regionalizados de cambio climático. A pesar de las debilidades observadas



en los resultados de validación a escala diaria, éstos mejoran considerablemente en la escala mensual, manifestando que de algún modo las redes entrenadas son capaces de reproducir razonablemente bien la variabilidad de baja frecuencia, cualidad importante cuando se trata de aplicarlas a estudios de variabilidad y cambio climático.

AGRADECIMIENTOS

Al proyecto R de software libre (<https://www.r-project.org/>) y en particular a los desarrolladores de la librería neuralnet con la que se ha llevado a cabo este trabajo.

BIBLIOGRAFÍA

- MP Amblar, MJ Casado, MA Pastor, P Ramos, y E Rodríguez. *Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR5*. AEMET, 2017. ISBN 978-84-7837-096-2.
- J Baño-Medina y JM Gutiérrez. Deep convolutional networks for feature selection in statistical downscaling. In *8th International Workshop on Climate Informatics*, 2018.
- M Brunet, MJ Casado, M De_Castro, P Galán, JA López, JM Martín, A Pastor, E Petisco, P Ramos, J Ribalaygua, E Rodríguez, I Sanz, y L Torres. *Generación de escenarios regionalizados de cambio climático para España*. Agencia Estatal de Meteorología, 2008.
- YB Dibike y P Coulibaly. Temporal neural networks for downscaling climate variability and extremes. *Neural Networks*, 19(2):135–144, 2006.
- Stefan Fritsch y Frauke Guenther. *neuralnet: Training of Neural Networks*, 2016. URL <https://CRAN.R-project.org/package=neuralnet>. R package version 1.33.
- JJ Gómez-Navarro, CC Raible, JA García-Valero, M Messmer, JP Montávez, y O Martius. Event selection for dynamical downscaling: a neural network approach for physically-constrained precipitation events. *Climate Dynamics*, May 2019. ISSN 1432-0894. doi: 10.1007/s00382-019-04818-w. URL <https://doi.org/10.1007/s00382-019-04818-w>.
- K Hornik, M Stinchcombe, y H White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.
- E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, y otros. The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471, 1996.
- M Riedmiller y H Braun. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In *IEEE International Conference on neural network*, pages 586–591. IEEE, 1993.
- DE Rumelhart, GE Hinton, y RJ Williams. Learning representations by back-propagating errors. *Nature*, 323(9):533–536, 1986.



RM Trigo y JP Palutikof. Simulation of daily temperatures for climate change scenarios over portugal: a neural network model approach. *Climate Research*, 13(1):45–59, 1999.