

VALIDACIÓN DE REDES NEURONALES PARA SU APLICACIÓN EN PROYECCIONES DE CAMBIO CLIMÁTICO SOBRE ESPAÑA.

Juan Andrés GARCÍA VALERO¹, Juan Pedro MONTÁVEZ GÓMEZ², Juan José GÓMEZ NAVARRO²

¹ *Delegación de AEMET en la Región de Murcia. AEMET.*

² *Departamento de Física de la Tierra. Universidad de Murcia.*

jugarciav@aemet.es

RESUMEN

Una red neuronal es una herramienta que puede aplicarse al desarrollo de proyecciones climáticas regionalizadas. Su problema es el alto coste computacional que requieren comparado con otro tipo de técnicas estadísticas. Sin embargo, su ventaja principal es que pueden aprender relaciones de tipo no lineal entre predictores y predictando. Hasta la fecha, y en particular sobre el territorio español, este método apenas se ha aplicado a la obtención de proyecciones climáticas, y menos aún a un número alto de puntos observacionales.

En este trabajo se presentan los resultados de validación obtenidos tras entrenar un conjunto de redes neuronales para 220 y 1204 series diarias de temperatura máxima/mínima y precipitación, respectivamente. Los resultados demuestran la gran capacidad de las redes neuronales en la estimación de la temperatura en general en todas las zonas del territorio español. Sin embargo, los resultados obtenidos para la precipitación son peores, siendo algo mejores en aquellas zonas con un régimen más regular de las precipitaciones.

Palabras clave: red neuronal, validación, proyección climática.

ABSTRACT

Neural networks (NNs) can be applied for obtaining regional climate projections. The main problem is the high computational cost they need compared to other statistical methods. However, the main advantage of this tool is the non linear relationships they can learn between predictors and predictand. Until now, NNs have not hardly been applied for obtaining climate projections over the Spanish territory, and even less to a large number of observations.

This study shows the validation results obtained from training NNs for 220 and 1204 observational daily series of temperature and precipitation, respectively. Results show a high capacity of NNs in estimating temperature in many parts of the Spanish regions. Nevertheless, results for precipitation are worse in general, although they are bit better in regions having a more regular precipitation regime.

Key words: neural networks, validation, climate projection.

1. INTRODUCCIÓN

El cambio climático es sin duda uno de los grandes retos de hoy día. El uso de proyecciones climáticas de alta resolución espacio-temporal es imprescindible a la hora de plantear estudios y políticas de adaptación y mitigación a los posibles impactos derivados. Las proyecciones regionalizadas son la base para este tipo de estudios. Para ello, AEMET pone a disposición en su página web (http://www.aemet.es/es/serviciosclimaticos/cambio_climat) diferentes proyecciones obtenidas a partir de los modelos climáticos globales (GCM) del último informe del IPCC (AR5) mediante distintos métodos de regionalización. Por un lado, se han utilizado dos técnicas estadísticas (regresión múltiple y análogos) y por otro, otra dinámica basada en los resultados del proyecto CORDEX.

Con el objetivo de aumentar el número de proyecciones disponibles, AEMET lleva trabajando desde hace un tiempo con un nuevo método de regionalización estadístico basado en redes neuronales (RNs). Esta herramienta tiene la capacidad de reproducir relaciones de tipo no lineal entre un conjunto de variables de entrada y otra de salida. Están formadas por un conjunto de nodos capaces de transferir información entre ellos de forma parecida a como lo hacen las neuronas del cerebro. Su entrenamiento se basa principalmente en el ajuste de dichas conexiones, para lo cual hay un amplio abanico de métodos de aprendizaje. Hay muchos tipos de arquitecturas de RN posibles, entre ellas está la de multicapa (Multilayer Perceptron, MLP) que hasta ahora ha sido la más usada en este tipo de aplicaciones, concretamente aquella que presenta tres capas (Hornick et al., 1985): una de entrada, otra intermedia (u oculta), y otra de salida. En ésta arquitectura de RN las neuronas se disponen en cada capa y la información va fluyendo desde la capa de entrada hacia la de salida (ver figura 1). En comparación con otros métodos estadísticos de downscaling las RNs se han aplicado a un número de trabajos muy reducido, a pesar de los mejores resultados de validación obtenidos por ellas en algunos estudios como el de Trigo y Palutikof., 1999. Su menor uso posiblemente se deba a que exigen un coste computacional más elevado.

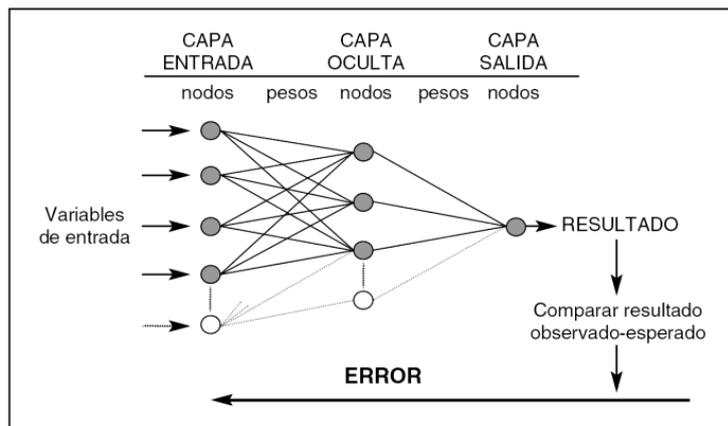


Fig. 1 Arquitectura de red neuronal con tres capas. Fuente: Cabello et al., 2005.

El objetivo principal de este trabajo es presentar los principales resultados de validación obtenidos de la calibración de un conjunto de RNs para un amplio número de estaciones de precipitación, temperatura máxima y mínima sobre el territorio español. En la sección 2 se muestra la metodología seguida, mientras que en las secciones 3 y 4 se presentan los principales resultados y su discusión, respectivamente.

2. MÉTODOS

Se ha entrenado una red neuronal para cada punto procedente de una base de datos observacional de AEMET. Así, se han considerado 222 series diarias de temperatura máxima y mínima (en adelante THX y TMN, respectivamente) y 1204 series de precipitación (en adelante, pcp). Además de la calidad contrastada de estas series, todas ellas presentan menos de un 10% de valores faltantes en el período 1979-2000. La arquitectura de RN usada fue la MLP de tres capas, empleándose como método de aprendizaje el de “resilient backpropagation” (Riedmiller and Braun, 1993) y como función de transferencia de salida de cada nodo la función logística.

Predictando	Predictores superficie	Predictores altura niveles: 850, 700 y 500 hPa
Precipitación	Presión a nivel del mar Divergencia del viento	Espesores referenciados al nivel de 850 hPa Humedad específica Humedad relativa Temperatura Viento zonal Viento meridional Vorticidad potencial referenciada a 850 hPa Índice de inestabilidad K Índice de inestabilidad TT Divergencia vector Q en 500
Temperatura	Presión a nivel del mar Temperatura máxima Temperatura mínima Viento zonal Viento meridional	Espesores referenciados a 850 Temperatura en 850 y 700 Humedad específica en 850 y 700 Viento zonal en 850 y 700 Viento meridional en 850 y 700

Tabla 1: Predictores (valores medios diarios) empleados en el downscaling de las diferentes variables predictando.

Para el entrenamiento de las RNs se han considerado como predictores (inputs) a una serie de campos atmosféricos (tabla 1) procedentes de reanálisis de NCAR (resolución 2.5°), empleándose para el calibrado de cada una el punto de reanálisis más próximo al de la observación. No obstante, dependiendo de la variable predictora, en cada

punto de observación se usó para el entrenamiento una submuestra de los campos predictores iniciales. Así, los campos predictores finalmente seleccionados fueron aquellos que dieron unos mejores resultados de validación. Dicha validación se basó en la correlación (en el caso de la precipitación) y en el error cuadrático medio (para la temperatura) entre las series diarias observadas, y las estimadas tras un proceso iterativo de inclusión de campos predictores. De este modo tras la inclusión de nuevos campos predictores se entrenaron 10 RNs para esa combinación de predictores, cada RN con un número distinto de neuronas, entre 1 y 10, en la capa oculta. Así, para cada combinación posible de predictores se obtuvieron 10 RNs, escogiendo finalmente como RN para el punto de observación a aquella que ofreció, entre todas las RNs calibradas, los mejores resultados de validación.

Los campos atmosféricos usados inicialmente para el downscaling de cada variable predictando se muestran en la tabla 1. No obstante, para cada punto de observación no se utilizaron todos ellos sino una submuestra que como acaba de explicarse en el párrafo anterior dio lugar a la mejor validación. Previamente al entrenamiento de las RNs los campos de reanálisis y las variables predictando se normalizaron, restando a cada dato su valor máximo y dividiéndolo por su rango. Por último, los períodos escogidos para el entrenamiento y validación fueron 1979-1995 y 1996-2000, respectivamente.

3. RESULTADOS

3.1 Patrones espaciales de Validación

En las Figuras 2 a 4 se presentan los patrones espaciales de las tres métricas de validación utilizadas, varianza explicada (EV), correlación (COR) y error cuadrático medio (ECM), obtenido para cada una de las variables consideradas.

En la Figura 2 se muestra el porcentaje de la varianza observada explicada por las series estimadas. Como resultado esperable, se aprecia que la validación de la precipitación es considerablemente peor que la obtenida para las otras dos variables. La EV es relativamente baja en precipitación siendo la mediana de la población del orden del 33%, aunque en zonas del Levante, este de Extremadura, oeste de Castilla La Mancha y en el Cantábrico, supera el 50% (Figura 3 izquierda). Tanto para THX como TMN la mediana de EV es del 100 y 93%, respectivamente.

En cuanto a la correlación (figura 4), la mediana en precipitación muestra un valor de 0.5, siendo las zonas peninsulares de clima atlántico las que presentan valores de correlación más altos entre 0.6 y 0.8 (Figura 3 derecha). Al igual que sucedía con la ECV, la correlación es muy alta para las variables THX y TMN, siendo la mediana de 0.96 y 0.94, respectivamente.

La última métrica considerada es el ECM (figura 5). La mediana en el caso de la precipitación es de 4.7 mm, siendo las zonas del interior del país las que presentan un error menor (inferior a 4 mm), mientras que las zonas más lluviosas del Cantábrico, litoral andaluz, especialmente Huelva, así como el litoral mediterráneo, desde el Golfo de Valencia hasta Gerona, presentan los valores más altos de ECM (superiores a 6 mm). Por último, la mediana del ECM para THX y TMN es 2.2 y 2.1 °C, respectivamente.

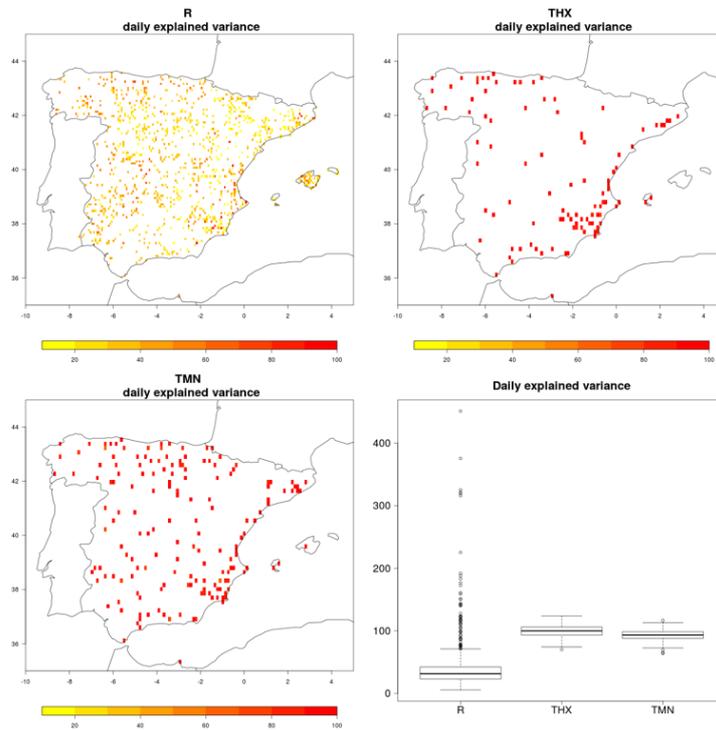


Fig. 2: Patrones espaciales de validación de las variables de precipitación (R), temperatura máxima (THX) y temperatura mínima (TMN) relativos a la métrica varianza explicada. En la gráfica inferior derecha se muestra el diagrama boxplot resumen de los mapas anteriores, representándose en la línea central de la caja el valor de la mediana.

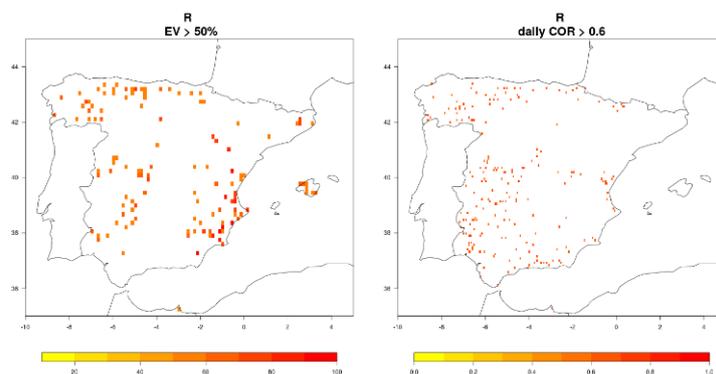


Fig. 3: Localización de las estaciones donde las series estimadas mediante RNs explican más del 50% de la varianza observada (izquierda), y presentan una correlación superior a 0.6 con sus series observadas (derecha).

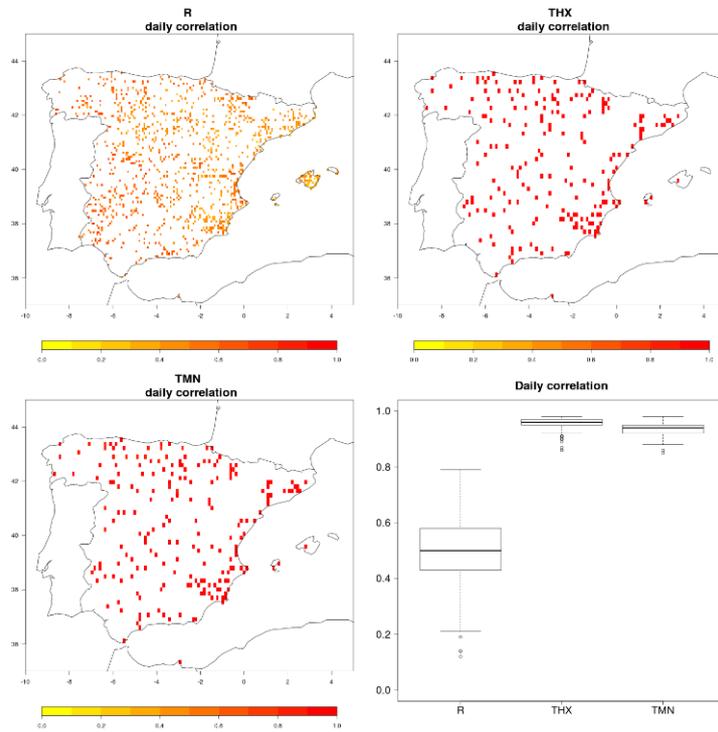


Fig. 4: Igual que la figura 2 pero en este caso se representa la correlación.

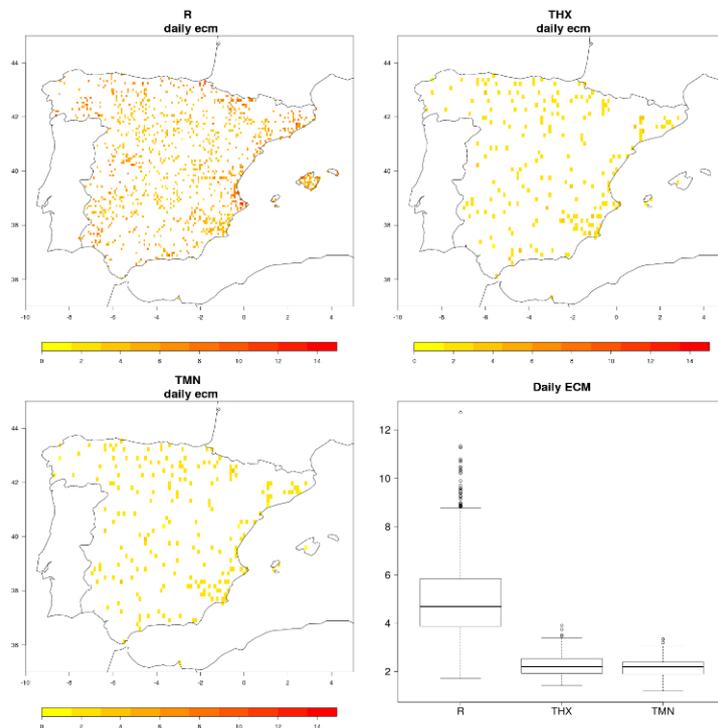


Fig. 5: Igual que la figura 2 pero en este caso se representa el ECM.

3.2 Validación puntual

La figura 6 muestra a modo de ejemplo la validación de la precipitación para la serie 3303E (estación: Carpio de Tajo, Toledo). En la parte superior de la figura (series diarias) puede verse en el diagrama qqplot que la RN sobrestima ligeramente las precipitaciones de los días menos lluviosos, por debajo de 5mm. A partir de ese umbral, la RN subestima los valores de precipitación. A pesar de ello, la RN es capaz de reproducir eventos diarios de precipitación cercanos a 50 mm.

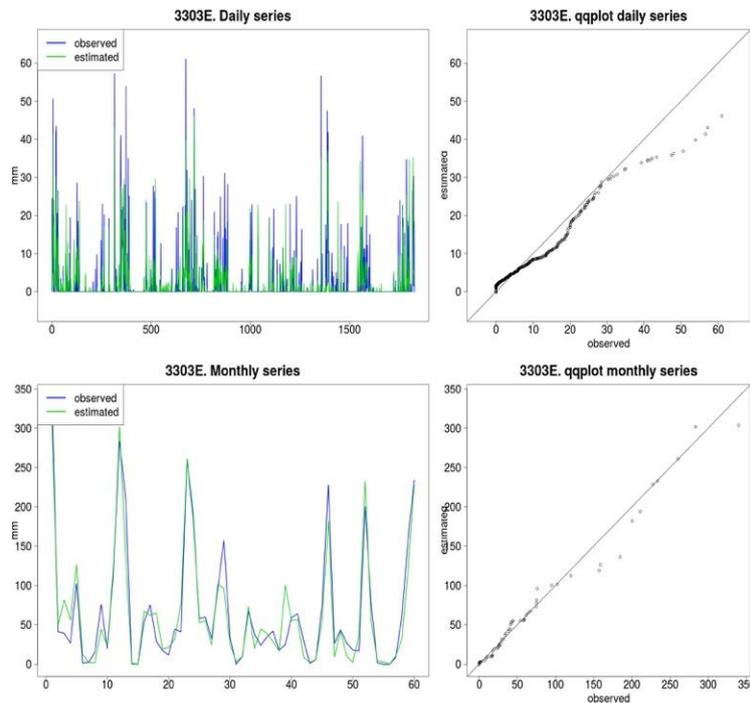


Fig. 6: En la parte izquierda se representan las series diarias estimadas (verde) y observadas (azul) de la precipitación para la estación 3303E. En la parte derecha se muestran los diagramas qq-plot de la estimación (ordenadas) vs observación (abcisas). Las gráficas de la parte superior representan la escala diaria mientras que las de la parte inferior están referidas a la escala mensual.

4. DISCUSIÓN

Los buenos resultados de validación derivados de la aplicación de RNs al downscaling de la temperatura diaria demuestran que esta herramienta presenta un gran potencial para ser empleada en la obtención de proyecciones de cambio climático. No obstante, los resultados relativamente peores obtenidos con la precipitación diaria ponen de manifiesto la necesidad de continuar investigando en esta variable, especialmente en aquellas zonas de la Península Ibérica que presentan un régimen más irregular.

Algunos de los resultados obtenidos con la precipitación demuestran la capacidad de las RNs de simular episodios intensos, los cuáles difícilmente pueden ser simulados con otro tipo de técnicas de downscaling. Aún así, los episodios más extremos en

general son subestimados, observándose este hecho tanto en precipitación como en temperatura. Este resultado evidencia la cautela con que debieran usarse las proyecciones de cambio climático en el caso que fueran a aplicarse al análisis de eventos extremos, así como la necesidad de disponer de todos estos parámetros de validación antes de emprender cualquier estudio relacionado con la aplicación de tales proyecciones.

Como continuidad a este estudio se pretende emplear las RNs entrenadas para obtener nuevas proyecciones de cambio climático para España, lo cuál contribuirá a ampliar el catálogo de las ya existentes. Por último, y como aspecto de mejora, se llevará a cabo el entrenamiento de nuevas RNs de forma estacional, hecho que podría mejorar los resultados de validación aquí presentados.

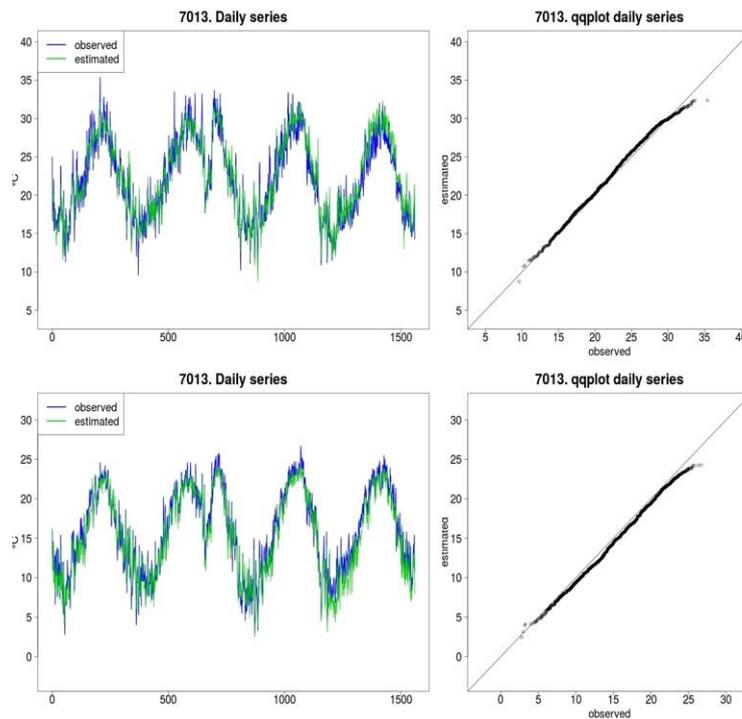


Fig. 7: Idem que la figura 6, pero aquí sólo se muestran los resultados para la escala diaria de la temperatura máxima (parte superior) y la temperatura mínima (parte inferior). Los gráficos están referidos a las series estimadas y observadas en la estación 7013 (Cartagena Puerto).

AGRADECIMIENTOS

Al proyecto R de software libre (<https://www.r-project.org/>) y en particular a los desarrolladores de la librería neuralnet (Günther y Fritsch, 2010) con la que se ha llevado a cabo este estudio.

REFERENCIAS

- Cabello, J. Castelló, M. Llanes, J. Pozo, A. Goixart, L. Tello, A. 2005. Redes neuronales artificiales en Medicina Intensiva. Ejemplo de aplicación con las variables del MPM II. *Medicina intensiva*, volumen 28, 13-20.
- Günther, F. Fritsch, S. 2010. Neuralnet: Training of neural networks. *The R journal*, 2, 20-38.
- Hornick, K. Stinchcombe, M. White, H. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2, 359–366,.
- Riedmiller, M. Braun, H. 1993. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *IEEE International Conference on neural network*. 586-591.
- Trigo, R. Palutikof, J. 1999. Simulation of daily temperatures for climate change scenarios over Portugal: a neural network model approach. *Climate Research*, volumen 13, 45-59.