

ESTRUCTURAS ESPACIALES COHERENTES DE PRECIPITACIÓN EN LA ESPAÑA PENINSULAR: RÉGIMEN ATLÁNTICO

M.Y. Luna, A. Morata, E. Avello, B. Peraza
Servicio de Desarrollos Climatológicos, Instituto Nacional de Meteorología.
Leonardo Prieto Castro, 8, 28040 Madrid.
Tfno.: 91 5819 703
Fax: 91 5819 767
Correo electrónico: yluna@inm.es

Resumen

Con el fin de estudiar el comportamiento climático de la precipitación en la España peninsular, es necesario identificar las regiones en las que existe una evolución similar. Para agrupar de una manera objetiva y matemáticamente rigurosa las series de precipitación, y puesto que este campo no sigue un comportamiento lineal, se ha empleado un método de clasificación no lineal para la detección de estructuras espaciales coherentes de la precipitación. Los datos observacionales se extraen de la totalidad de los datos contenidos en la base de datos histórica del Instituto Nacional de Meteorología (INM) para el estado climático 1961-1990. El número de observatorios es variable dependiendo del instante temporal y su distribución espacial es irregular. Con estas observaciones, el Servicio de Desarrollos Climatológicos del INM ha elaborado una base de datos de precipitación diaria mediante interpolación espacial en una rejilla de 25 km utilizando la técnica IDW (Inverse Distance Weighted). Esta base de datos posee una distribución espacial y temporal homogénea, en la que no existen datos faltantes. La metodología empleada en la determinación de estructuras espaciales coherentes de la precipitación consta de dos fases. En la primera de ellas, se realiza un filtrado de la señal de precipitación mediante una Descomposición Wavelet discreta. En la segunda fase, y a partir de las series de coeficientes wavelet obtenidas en la descomposición, se forman conglomerados mediante una red neuronal no supervisada de Kohonen. De este modo las series originales son agrupadas en función de su comportamiento similar en la banda de frecuencia seleccionada, encontrándose un número reducido de estructuras coherentes o mapas autosimilares independientes, identificándose de forma totalmente objetiva la zona de influencia meramente atlántica.

Palabras clave

Precipitación, Península Ibérica, análisis wavelet, red neuronal de Kohonen.

Abstract

In order to study the climatic behaviour of Spanish precipitation it is necessary to identify regions characterised by similar evolution. The coherent spatial structures are determined by a classification methodology of precipitation time series that should be objective, mathematically rigorous and non-linear. The observational data are extracted from the whole data contained in the historical database of the Spanish Meteorological Centre (INM) for the climatic state 1961-1990. The spatial distribution of the observatories is irregular and their number depends on the time period. From these data, the Servicio de Desarrollos Climatológicos of INM has built a database of daily precipitation by means of spatial interpolation in a grid of 25 km using an IDW (Inverse Distance Weighted) method. This database is spatially and temporally homogeneous without missing data. The methodology used for determining the precipitation spatial coherent structures combines both wavelet

theory and a neuronal network. First, the precipitation signal is filtered by discrete wavelet decomposition. Second, conglomerates are obtained from a Kohonen neural network applied to the wavelet coefficients. As a result, the original precipitation time series are grouped in terms of their similar behaviour at the selected frequency. Independent spatial coherent structures or self-organised maps are isolated identifying the zone of Atlantic regime in an objectively way.

Keywords

Precipitation, Iberian Peninsula, wavelet analysis, Kohonen neural network.

1. INTRODUCCIÓN

El estudio del comportamiento climático de la precipitación en la España peninsular necesita en primer lugar de una identificación rigurosa de las regiones en las que se puede detectar una evolución similar. Algunos autores identifican de un modo cualitativo dos regiones fundamentales de comportamiento climático claramente diferente en la Península Ibérica, una regida por un régimen Atlántico y la otra por un régimen Mediterráneo (Jansá, 1969; Font, 2000). Análogamente, Rodó et al. (1997) identifica un área de influencia claramente Atlántica regida por el patrón de la NAO y una zona que se ajusta al litoral Mediterráneo, cuyo comportamiento relaciona con el ENSO. Por otra parte, diversos autores han abordado este tema aplicando diferentes criterios en la segregación de los observatorios que presentan características comunes. González-Rouco (1997) utiliza la isolínea de correlación de 0.4 del primer modo del análisis de componentes principales de la precipitación mensual invernal como criterio de delimitación obteniendo dos zonas; una región que denomina de régimen de influencia Atlántica, y otra región de influencia Mediterránea. Sáenz (2000) tiene en cuenta un criterio similar para determinar una tercera zona de comportamiento climático que abarca la Cornisa Cantábrica; Esteban-Parra et al. (1998) siguiendo el mismo criterio que González-Rouco divide la Península Ibérica en tres regiones: interior y sur de la Península, costa Mediterránea y costa Cantábrica. Serrano et al., (1999) aplica un análisis *cluster* en función de un valor umbral del coeficiente de congruencia de un análisis de componentes principales rotadas fijado en 0.82. Fernández-Mills (1995) aplica el procedimiento de *clustering* a las series temporales de un análisis de componentes principales de las series de precipitación originales utilizando el *método de la distancia media* con el que identifica 10 regiones de comportamiento climático diferente en la Península Ibérica.

Todos los métodos mencionados resultan en mayor o menor medida subjetivos o, por razones de linealidad, inapropiados para la regionalización de la precipitación. Ésta variable no sigue un comportamiento lineal y, por tanto, no deben aplicarse métodos lineales para su regionalización. El objetivo de este trabajo consiste en agrupar de manera objetiva y matemáticamente rigurosa las series de las que se dispone; para ello, se ha empleado un método de clasificación no lineal que detecta estructuras coherentes de la precipitación mensual en baja frecuencia. En primer lugar, es necesario realizar un filtrado de las series de anomalías originales de manera que desenmascaren la señal de baja frecuencia; posteriormente se procede a su clasificación mediante una *red neuronal no supervisada de Kohonen*.

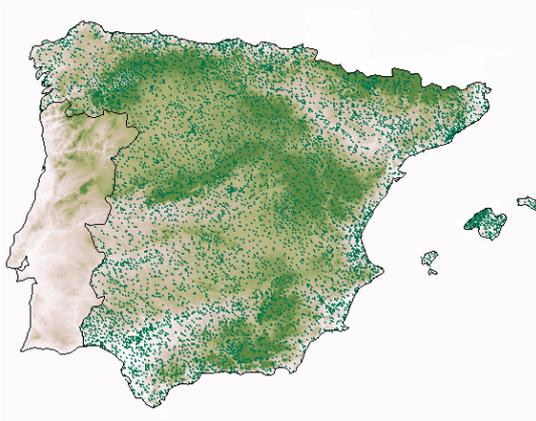
2. DATOS

Los datos observacionales de precipitación diaria acumulada se extraen de la totalidad de los datos contenidos en la base de datos histórica del Instituto Nacional de Meteorología (INM) para el estado climático 1961-1990. El número de observatorios es variable dependiendo del instante temporal y su distribución espacial es irregular. Concretamente, se contaba con 4617

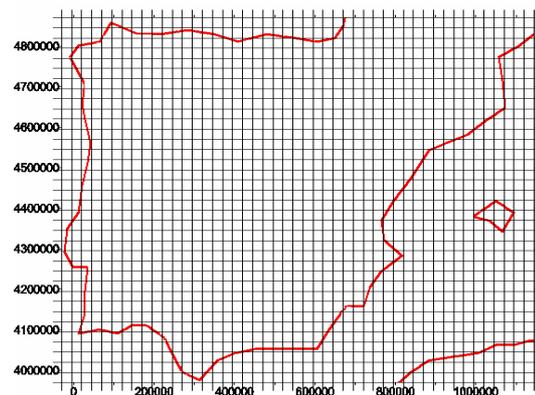
estaciones pluviométricas y 164 completas en el último instante temporal considerado, mientras que en el primer instante temporal sólo se tenían 644 estaciones. En la Figura 1 se muestra el mapa de la Península Ibérica y las Islas Baleares con los observatorios disponibles en el último instante temporal (31 de diciembre de 1990). Con estas observaciones, el Servicio de Desarrollos Climatológicos del INM ha elaborado una base de datos de precipitación diaria mediante interpolación espacial en una rejilla de 25 km utilizando la técnica IDW (Inverse Distance Weighted). La disposición de los puntos de la rejilla donde se han obtenido datos interpolados espacialmente se muestra también en La Figura 1. El algoritmo IDW proporciona resultados razonables incluso con datos tan irregulares como los de precipitación. Los métodos de interpolación espacial basados en algoritmos Kriging y Spline no ofrecen resultados adecuados debido a que suavizan fuertemente los campos resultantes perdiéndose la información de sucesos extremos de precipitación (Dittmann et al., 1999). La base de datos así construida posee una distribución espacial y temporal homogénea, en la que no existen datos faltantes, constituyendo un conjunto de datos de gran utilidad para todo tipo de estudio climático.

Figura 1.(a) Mapa del área en estudio con indicación de las observaciones disponible el 31 de diciembre de 1990; (b) Disposición de la rejilla utilizada en la interpolación espacial.

(a)



(b)



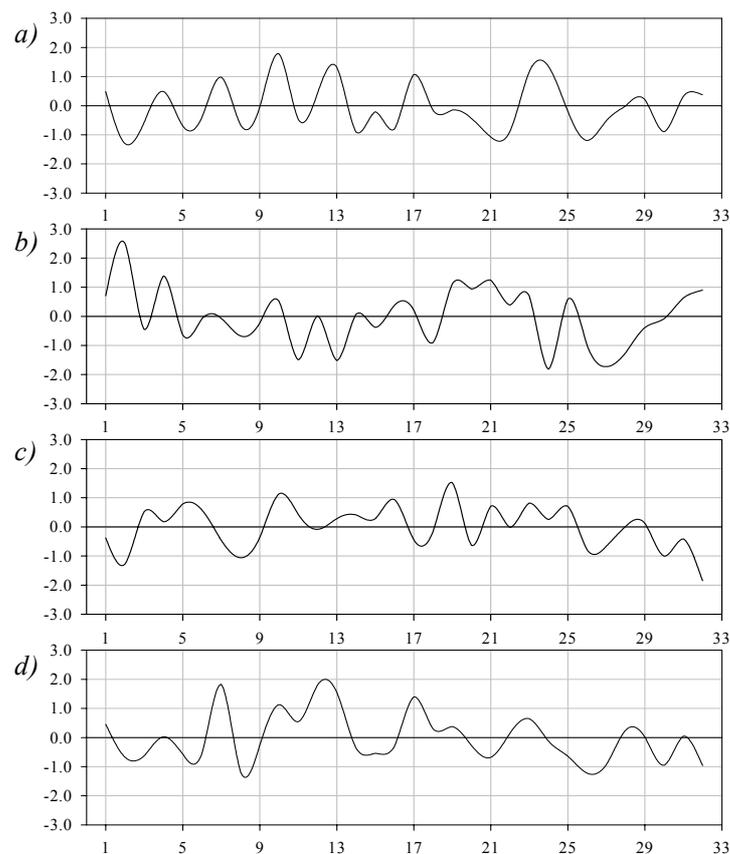
3. APROXIMACIÓN DE LA BAJA FRECUENCIA A LA SEÑAL DE PRECIPITACIÓN.

La base de datos de precipitación ha sido sometida a un Análisis Multirresolución *Wavelet* (AMW) con objeto de conseguir una representación de la señal como suma de sus componentes esenciales en distintas frecuencias. El AMW fue introducido por Morlet et al. (1982) y Grossman y Morlet (1984) para el análisis tiempo-frecuencia de señales no estacionarias. En el análisis *wavelet* se utilizan combinaciones lineales de funciones llamadas *wavelet* para representar una señal $f(t)$ mediante dilataciones y traslaciones de una misma función. Las funciones *wavelet* son funciones matemáticas de cuadrado integrable ($f(t) \in L^2(R)$), en general ortogonales, de soporte compacto y media cero, que permiten descomponer una señal no estacionaria en componentes multirresolución a distintas

frecuencias, capturando tanto los detalles finos como los gruesos de cada componente así como las características fundamentales de cada escala. Para la identificación de bajas frecuencias en las series de precipitación se ha realizado un análisis de Transformada *Wavelet* Discreta (TWD) especialmente apropiada para identificar procesos de baja frecuencia y larga memoria. Se ha utilizado para ello la familia de *wavelets Symmlet* a un nivel de multirresolución correspondiente a un período característico de baja frecuencia de 5-7 años.

La Figura 2 muestra para 4 puntos de la red, a modo de ejemplo, las series de coeficientes *wavelet* correspondientes a una variabilidad de 5-7 años. Las series cercanas a Badajoz (Figura 2a) y a San Fernando (Figura 2d) muestran un patrón de variabilidad similar mientras que las series cercanas a Murcia y Bilbao (Figura 2b y c) tienen un comportamiento en baja frecuencia diferente a las dos anteriores y también diferente entre sí. Para agrupar series de precipitación de comportamiento similar de manera objetiva, las 1775 series de coeficientes *wavelet*, obtenidas de la descomposición, se han sometido a una red neuronal para la formación de conglomerados.

Figura 2. Series de coeficientes *wavelet* de baja frecuencia de las series de precipitación en los puntos de rejilla más próximos a: (a)Badajoz; (b) Murcia; (c) Bilbao y (d) San Fernando. El eje de abcisas representa el coeficiente y el eje de ordenadas su valor adimensional.

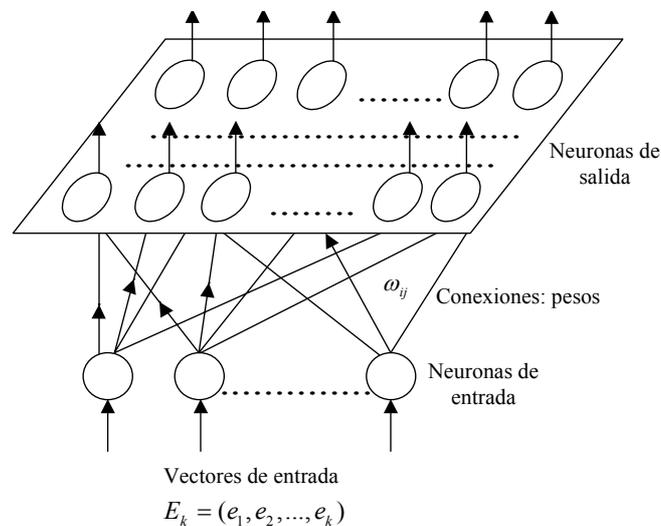


4. CREACIÓN DE CONGLOMERADOS: RED NEURONAL DE KOHONEN.

La segunda fase consiste, pues, en la formación de conglomerados con las series de coeficientes *wavelet* obtenidas en la descomposición *wavelet* llevada a cabo en el apartado anterior. De este modo, se han caracterizado las series de precipitación originales por la semejanza en baja frecuencia de las señales, es decir, las series originales van a ser agrupadas en función de su comportamiento en baja frecuencia. Para organizar consistentemente la información de las series de coeficientes *wavelet* de las series de precipitación se ha utilizado la red neuronal no supervisada de Kohonen (Kohonen, 1982). Se trata de una técnica de clasificación no lineal que identifica, basándose en el cálculo de distancias euclídeas, características comunes en señales para la creación de mapas topológicos autosimilares.

El modelo de Kohonen tiene dos variantes denominadas *Learning Vector Quantization* (LVQ) y *Topology Preserving Map* (TPM) o *Self-Organizing Map* (SOM) basadas en el principio de formación de mapas topológicos para establecer características comunes entre los vectores de entrada a la red (Simpson, 1990). La diferencia entre las dos variantes radica en las dimensiones de los vectores de entrada; mientras que el modelo LVQ únicamente procesa vectores unidimensionales, el modelo SOM puede procesar vectores bidimensionales e incluso tridimensionales. La red utilizada para la segregación de series de precipitación en esta segunda fase del proceso ha sido, pues, la variante SOM de malla rectangular (Figura 3).

Figura 3. Arquitectura de la red SOM (Self-Organizing Map) de Kohonen



Se trata de una red que utiliza un *aprendizaje no supervisado de tipo competitivo* (Kohonen, 2001). Las neuronas de la capa de salida compiten por activarse y sólo una de ellas permanece activa ante una determinada información de entrada a la red ajustándose los pesos, w_{ij} , de las conexiones en función de la neurona clave. Durante la etapa de entrenamiento se presenta a la red un conjunto de informaciones de entrada denominado *conjunto de vectores de entrenamiento*, $E_k = (e_1^{(k)}, \dots, e_n^{(k)})$. Estos vectores de entrenamiento forman la base en función de la cual se establece la semejanza entre los datos constituyendo así las diferentes categorías, una por cada neurona de salida, que servirán durante la fase de funcionamiento para realizar clasificaciones de los datos que se presentan a la red. Por otro lado, las neuronas

de salida también reciben las correspondientes entradas debidas a las conexiones laterales con el resto de las neuronas de salida y cuya influencia dependerá de la distancia a la que se encuentren. La salida generada por una neurona de salida j ante un vector de entrada E_k será:

$$s_j(t+1) = f \left(\sum_{i=1}^N w_{ij} e_i^{(k)} + \sum_{p=1}^M Int_{pj} s_p(t) \right) \quad (1)$$

donde Int_{pj} es una función del tipo *sombrero mejicano* que representa la influencia lateral de la neurona p sobre la neurona j , s_p la salida generada por una neurona p que a su vez es información de entrada de la neurona j y s_j es la salida generada por la neurona j .

Los valores finales de los pesos de las conexiones entre cada neurona de la capa de salida con las de entrada corresponden con los valores de las componentes del vector de aprendizaje que consigue activar la neurona correspondiente. Si existe mayor número de patrones de entrenamiento que de neuronas de salida, los patrones deben asociarse con una misma neurona perteneciendo, por tanto, a la misma clase. En tal caso, los pesos se obtienen como promedio de dichos patrones. De manera simplificada, el funcionamiento de la red responde a la expresión:

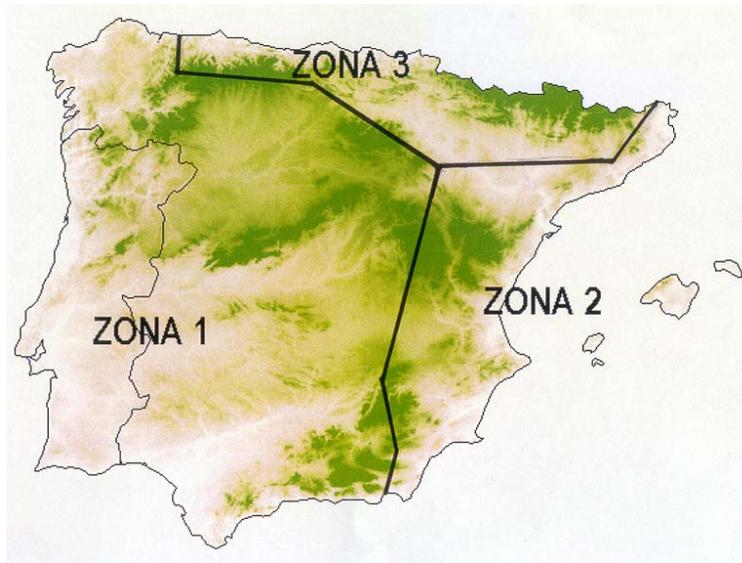
$$s_j = \begin{cases} 1 & \text{MIN} \|E_k - W_j\| = \text{MIN} \left(\sqrt{\sum_{i=1}^N (e_i^{(k)} - w_{ij})^2} \right) \\ 0 & \text{resto} \end{cases} \quad (2)$$

donde $\|E_k - W_j\|$ constituye la distancia euclídea entre el vector de entrada y el vector de peso de las conexiones entre cada una de las neuronas de entrada y la neurona de salida. La red de Kohonen, en definitiva, realiza una tarea de clasificación, ya que la neurona de salida activada ante una entrada representa la clase a la que pertenece la información de entrada. Así pues, ante entradas parecidas se activa una misma neurona de salida u otra cercana, debido a la semejanza de clases, garantizando que neuronas topológicamente próximas sean sensibles a entradas físicamente similares. Por lo tanto, la red de Kohonen resulta especialmente útil para establecer relaciones, *a priori* desconocidas, entre conjuntos de datos.

De la fase anterior se han obtenido 1776 series de coeficientes *wavelet* mediante la aplicación del análisis TWD. Estas series constituyen el conjunto de vectores de entrada que se introducen en la red neuronal mediante la aplicación completa de la metodología anteriormente expuesta. Se han podido identificar tres regiones de comportamiento de la precipitación de baja frecuencia que se muestran en la Figura 4. Estas regiones han sido denominadas como sigue:

- Zona 1: Meseta Central, Galicia y suroeste de la Península Ibérica
- Zona 2: Litoral Mediterráneo y la zona suroriental del Valle del Ebro
- Zona 3: Cornisa Cantábrica y la zona noroccidental del Valle del Ebro

Figura 4. Mapa con las regiones autosimilares de precipitación en baja frecuencia.



5. CONCLUSIONES

Se ha desarrollado un método de detección de mapas autosimilares de la precipitación en baja frecuencia en la Península Ibérica y Baleares que introduce de forma combinada dos tipos de análisis: wavelets y redes neuronales no supervisadas. El propósito es agrupar de una manera objetiva las series de precipitación por sus características comunes en baja frecuencia y para ello se emplea un método de clasificación que es no lineal y es capaz de detectar estructuras espaciales coherentes. Se dispone de una base de datos de precipitación en una rejilla de 25 km que abarca la Península Ibérica y las Islas Baleares. Las series originales son agrupadas en función de su comportamiento similar en la banda de frecuencia seleccionada, encontrándose un número reducido de estructuras coherentes o mapas autosimilares independientes. El método ha permitido detectar 3 estructuras en baja frecuencia de precipitación asociadas a distintos regímenes de flujo atmosférico sobre la zona. Así, la zona de influencia meramente atlántica (zona 1) queda identificada de forma totalmente objetiva abarcando la Meseta Central, Galicia y el suroeste de la Península Ibérica. Esta metodología propuesta será utilizada con otras variables climáticas y a otras escalas temporales para mejorar la clasificación de las zonas que encontradas en este trabajo.

REFERENCIAS

- Dittmann, E., G. Muller-Westermeier, A. Kreis, P. Fuchs, S. Szalai, Z. Bihari, A. Bussay, T. Szentimrey and C. Szinell (1999). *Spatial interpolation of climatological parameters*. Deutscher Wetterdienst Forschung und Entwicklung. Arbeitsergebnisse, Nr. 57.
- Esteban-Parra, M. J., F. S. Rodrigo, Y. Castro-Diez (1998). Spatial and temporal patterns of precipitation in Spain for the period 1880-1992. *Int. J. Climatol.* **18**: 1557-1574.
- Font, I. (2000). *Climatología de España y Portugal*, 2nd edn. Ediciones Universidad de Salamanca, Spain. pp422.
- Fernández-Mills G., X. Lana and C. Serra (1994). Catalanian precipitation patterns: Principal component analysis and automated regionalization. *Theor. Appl. Climatol.*, **49**, 201-212.

- González-Rouco J. F. (1997). *Modelo de predicción de la precipitación peninsular en climas perturbados*. Tesis Doctoral. UCM. 237 pp.
- Grossmann A. and J. Morlet (1984). Decomposition of Hardy functions into square integrable *wavelets* of constant shape, *SIAM. J. Math. Anal.*, **15**, 723-736.
- Jansá, J. M. (1969). *Curso de climatología*. Ed. INM Serie B. 445 pp.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59-69.
- Kohonen, T. (2001). *Self-Organizing maps*. 3rd ed, Springer-Verlag.
- Morlet, G. A., I. Fourgeau and D. Giard (1982). Wave propagation and sampling theory. *Geophysics*, **47**, 203-236.
- Rodó, X., E. Baert and A. Comin (1997). Variations in seasonal rainfall in Southern Europe during the present century: relationships with the North Atlantic Oscillation and El Niño-Southern Oscillation. *Clim. Dyn.*, **13**, 275-284.
- Sáenz J. (2000). *Estudio de las características regionales del clima invernal en el País Vasco and sus relaciones con el sistema climático global*. Tesis doctoral. Universidad del País Vasco.
- Serrano A., J. E. García, V. L. Mateos, M. L. Cancillo, and J. Garrido (1999). Monthly modes of variation of precipitation over the Iberian Peninsula. *J. Climate*, **12**, 2894-2919.
- Simpson, P. K. (1990). *Artificial Neural Systems*. Ed. Pergamon Press.