

ESTUDIO DE LA VARIABILIDAD CLIMÁTICA EN EL SUR DE ESPAÑA DURANTE EL PERIODO 1961-90 MEDIANTE ANÁLISIS DE CORRELACIÓN CANÓNICA

José Luis UCEDA, Juan Pedro MONTÁVEZ y Juan Ignacio JIMÉNEZ
Grupo de Climatología Urbana y Cambio Climático
Dpto. Física Aplicada, Facultad de Ciencias, Universidad de Granada

RESUMEN

Diferentes variables meteorológicas fueron medidas en siete estaciones en el sur de España cada enero en el periodo 1961-90. La presión media a nivel del mar fue medida en una rejilla que se extiende desde 50°N20°W hasta 30°N10°E. Mediante análisis en componentes principales se prepararon los datos, que fueron estudiados mediante análisis de correlación canónica. El periodo elegido se caracteriza por situaciones anómalas en la mayoría de las variables, lo cual queda reflejado en este estudio.

Palabras clave: Análisis de Correlación Canónica, Análisis en Componentes Principales, España.

ABSTRACT

Different meteorological variables were measured every January in seven stations in the south of Spain for the period 1961-90. Mean-sea-level pressure was measured in a grid-base point from 50°N20°W to 30°N10°E. The database was prepared through principal components analysis in order to be studied by canonical correlation analysis. The chosen period is characterized by anomalous situations in the most of the variables, which is reflected in this work.

Key words: Canonical correlation analysis, principal components analysis, Spain

1. INTRODUCCIÓN

Pretendemos en este trabajo hacer un Análisis de Correlación Canónica, técnica estadística del análisis multivariante que nos permite encontrar ciertos mecanismos causales entre diferentes variables físicas. Nos ayudaremos de otra técnica conocida como Análisis en Componentes Principales, la cual nos permitirá reducir las dimensiones de nuestra base de datos, estableciendo un cierto número de patrones y encontrando los vínculos entre las variables del problema y los patrones que mejor las describen. Además habremos conseguido eliminar gran parte del ruido que contaminaba a las señales climáticas originales y trabajaremos con series que son ahora funciones ortogonales con lo que se facilita el estudio de la estructura de covarianza.

El análisis de correlación canónica se basa en el estudio de la correlación que existe entre dos conjuntos de variables. Se encuentran ciertas variables (llamadas canónicas) que son combinaciones lineales de los datos originales y se construyen de tal manera que se maximice la correlación entre ellas. Pretendemos encontrar, por tanto, ciertas relaciones causales y vínculos entre diferentes variables físicas. Más adelante hacemos una breve introducción teórica a esta técnica. (Más detalles además en PREISENDORFER (1988); WERNER y VON STORCH (1993).)

2. DATOS

Siete estaciones fueron elegidas en el sur de España. En cada una de ellas se midieron diferentes variables, cuyas anomalías extremas, en la estación de invierno, pretendemos estudiar; para ello elegimos sólo las medidas del mes de enero, en vez de la media entre los meses de diciembre, enero y febrero. Tenemos, por tanto, series temporales anuales para el periodo 1961-90. Las variables de los registros climatológicos son: presión atmosférica, temperatura mínima y máxima, humedad relativa, precipitación e insolación. Se resumen a continuación las principales características de dichas estaciones (Tabla I, Fig. 1). Los datos fueron suministrados por el Instituto Nacional de Meteorología (INM).

La misma base de datos ha sido ya estudiada en varios trabajos en nuestro grupo de investigación, como el presentado en Cádiz en el First International Conference on Applied Sciences and the Environment (ASE98); SÁNCHEZ *et al.* (1988).

Estación	Iniciales	Altura (m)	Latitud	Longitud
Albacete	AL	698,56	38° 56'	01° 51'
Badajoz	BA	185	38° 53'	06° 49'
Ciudad Real	CR	627	38° 59'	03° 55'
Granada	GR	687	37° 08'	03° 37'
Málaga	MA	6,65	36° 39'	04° 28'
San Javier	SJ	2,6	37° 47'	00° 48'
Sevilla	SE	27	37° 27'	05° 53'

Tabla I. Principales características de las siete estaciones.

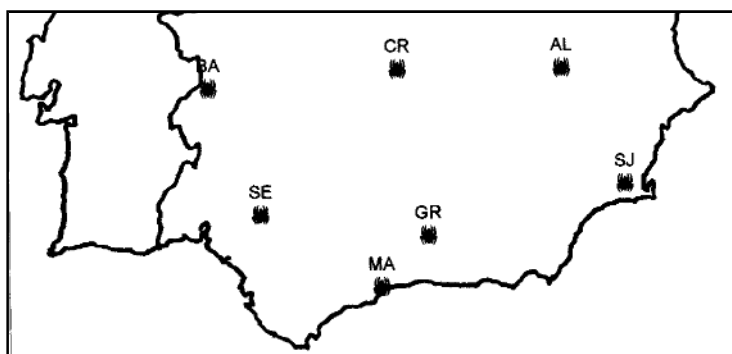


Figura 1. Distribución de estaciones.

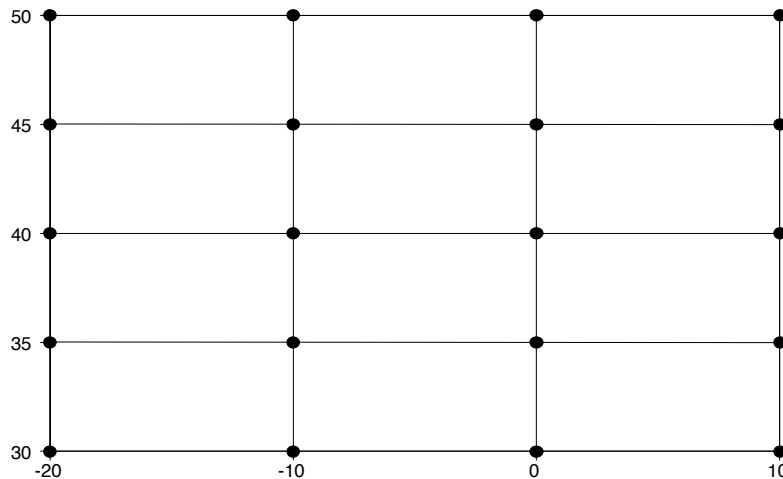


Figura 2: Distribución de la rejilla.

Para el análisis de correlación canónica, se utilizó una rejilla, que se extiende desde los 50°N20°W hasta los 30°N10°E, con datos de presión media a nivel del mar cada 5° de latitud y 10° de longitud (Fig. 2). La fuente de estos datos está en JONES, 1987.

Hacemos un estudio, por tanto, entre las presiones en esta rejilla (que será el patrón sinóptico) y diferentes variables físicas en las siete estaciones de la región de estudio.

3. METODOLOGÍA

En esta sección se describen brevemente las técnicas utilizadas.

3.1 Análisis en componentes principales

Las Componentes Principales se obtienen proyectando los datos originales en la base de los vectores característicos de la matriz de covarianza y son también series temporales. De manera compacta,

$$A = ETX \quad X = AE \quad (1)$$

donde A es la matriz de componentes principales, E la matriz de los vectores característicos y X la matriz de los datos originales. Cada nueva componente principal explica el máximo posible de varianza del proceso, estando incorrelada con las precedentes e incorpora una nueva característica del proceso. Los valores característicos, λ_k , autovalores de la matriz de covarianza, son la varianza explicada por cada componente y son tales que la primera componente principal da cuenta de la mayor cantidad de varianza explicada; seguida por la segunda y así sucesivamente. Las componentes que expliquen sólo una porción muy pequeña de varianza, son consideradas como ruido que contamina la señal climática original; se toman sólo las significativas según ciertas reglas de selección. Para su mejor interpretación, se suelen rotar ortogonalmente estos vectores característicos, mediante rotación Varimax.

3.2. Análisis de correlación canónica

El análisis de correlación canónica, básicamente, examina las relaciones lineales que hay entre dos conjuntos de variables, encontrando algunas combinaciones lineales entre las variables del primer conjunto y el mismo número de combinaciones en el segundo de tal manera que expresen, de la mejor manera posible, las correlaciones entre los dos conjuntos. En el análisis de correlación canónica se parte de la estructura de covarianza y covarianza cruzada entre los dos conjuntos de datos. De productos de estas matrices, se derivan los patrones adjuntos como autovectores. Encontramos, por tanto, una combinación lineal de las y (variables de un conjunto) y una combinación lineal de las x (variables del otro conjunto)

$$\begin{aligned}c_1(y) &= a_1y_1 + a_2y_2 + \dots + a_ky_k \\c_1(x) &= b_1x_1 + b_2x_2 + \dots + b_kx_k\end{aligned}\tag{2}$$

donde los coeficientes a y b son seleccionados de tal manera que se maximice la correlación entre $c_1(y)$ y $c_1(x)$, llamadas primeras variables canónicas para los y y x , respectivamente. La correlación entre ellas se llama primera correlación canónica. Derivamos, ahora, sucesivos pares de variables canónicas, con máxima correlación, de tal modo que estén incorrelados con los pares precedentes. Seguimos derivando pares hasta que haya el mismo número de pares que el número de variables en el conjunto más pequeño. Estas variables canónicas son también series temporales.

Finalmente ponemos nuestra base de datos como combinación lineal de estas variables canónicas que acabamos de obtener.

Si se hace un previo análisis en componentes principales, dentro de cada conjunto las variables son ahora funciones ortogonales, con la consiguiente simplificación de la estructura de covarianza (la matriz de covarianza es diagonal). Además al reducir las dimensiones del sistema hemos eliminado gran parte del ruido en la señal climática original.

Una vez acabado el análisis, se vuelve al espacio físico, obteniendo los patrones espaciales mediante una transformación lineal.

4. RESULTADOS

Un primer examen de los datos nos muestra que en el periodo de estudio las variables se comportan de manera anormalmente extrema respecto a la media del siglo. Por ejemplo, hay un gran déficit de precipitaciones y en el caso de las presiones sus medias son anormalmente altas; en algunos casos incluso superiores a 1020 mb, sobre todo al oeste de la Península Ibérica; sin embargo, durante la estación invernal, es usual encontrar actividad ciclónica en esta zona.

Previamente al análisis de correlación canónica se realizó un análisis en componentes principales entre las estaciones con cada una de las variables meteorológicas. En todos los casos se rotaron las componentes. Se tomaron cuatro componentes, superando en todos los casos el 97% de varianza explicada acumulada.

Los coeficientes de correlación entre estas componentes y las variables originales nos permiten agruparlas en subregiones (no se muestran las figuras). En general encontramos una región occidental

que suele incluir las estaciones de Badajoz y Sevilla y una oriental que suele incluir a las de Almería y San Javier. Otras dos componentes describen principalmente las características individuales de las estaciones de Málaga y Granada. La cercanía de estas dos últimas estaciones mencionadas nos sugiere, para el análisis posterior, la aparición de cierto gradiente. Con la rejilla, de un total de veinte puntos, bastó con sólo tres componentes (que dieron cuenta de un 97% de varianza explicada acumulada) que procedimos también a rotar.

El análisis de correlación canónica fue llevado a cabo entre la salida del análisis en componentes principales de la rejilla y la salida del análisis en componentes principales de distintas variables en nuestra región. En algunos casos fue también calculado el ritmo de varianza explicada $\xi_i(x)$, parámetro que mide la intensidad del vínculo entre un patrón de circulación y la distribución de cierta magnitud física (WERNER *et al*, 1993). En el caso de las precipitaciones, el primer par de correlación canónica resultó tener una correlación del 89% (Fig. 3). En el mapa de anomalías del patrón sinóptico apreciamos actividad anticiclónica al oeste de la Península Ibérica. Nuestra región de estudio, está influenciada, por tanto, por altas presiones, lo que provoca un déficit pluviométrico generalizado en toda la región; pero con más fuerza hacia el oeste. Sólo la estación de San Javier presenta anomalías positivas. El ritmo de varianza explicada, supera en todas las estaciones el 50%, excepto en la de San Javier (sólo un 3%), mostrando el vínculo que hay entre la situación sinóptica y el déficit en las precipitaciones. El segundo par, con una correlación del 39% (Fig. 4), nos muestra actividad ciclónica en el suroeste; en todas las estaciones el campo de anomalías es positivo, decreciendo hacia el norte.

En el caso de las temperaturas máximas, el primer par resultó tener una correlación del 85% (Fig. 5). En la zona central-occidental el campo de temperaturas tiene cierto grado de homogeneidad con anomalías positivas. Aunque la región está influenciada por anomalías positivas de presión (Fig. 5, arriba), hay anomalías negativas en el suroeste y sabemos que la presencia de lluvias y cielos cubiertos reducen el enfriamiento por radiación. Ambos factores pueden explicar dicha situación anómala.

Fue calculado el coeficiente de correlación entre algunas variables canónicas y la serie del índice NAO, encontrándose altos valores (algunos por encima de 0.7).

5. CONCLUSIONES

La localización e intensidad del anticiclón de las Azores es determinante para el régimen pluviométrico en nuestra zona de estudio. Hemos visto la alta correlación existente entre el índice NAO y las variables canónicas. El debilitamiento de su centro de altas presiones en invierno permite la entrada de flujos del oeste, húmedos y templados, que reducen el enfriamiento por radiación (por la presencia de lluvias y cielos cubiertos). Sabemos que periodos de descenso en las precipitaciones están relacionados con periodos de altas presiones en el anticiclón de las Azores o con su desplazamiento hacia el este. Los patrones obtenidos están, por tanto, conectados con la situación sinóptica real sobre la zona sur de España cada enero en el periodo 1961-90.

El análisis de correlación canónica se revela, pues, como una herramienta útil, capaz de mostrar conexiones y relaciones causales entre diferentes variables físicas.

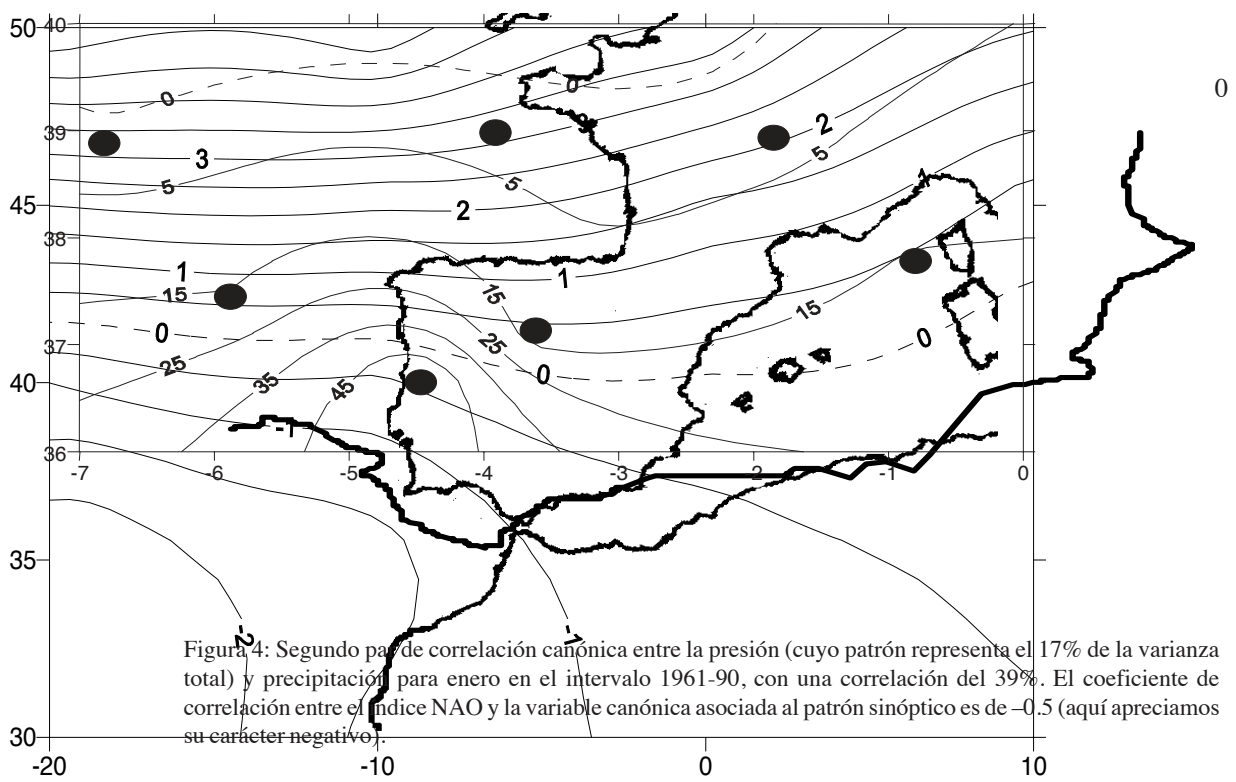




Figura 5: Primer par de correlación canónica entre la presión y temperatura máxima para enero en el intervalo 1961-90, con una correlación del 85%.

REFERENCIAS BIBLIOGRÁFICAS

JONES, P.D. (1987): "The Early Twentieth Century Arctic High – Fact or Fiction?", *Climate Dynamics* **1**. 63-75.

PREISENDORFER, R. W. (1988): *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, Amsterdam.

WERNER, C. y VON STORCH, H. (1993): "Interannual Variability of Central European Mean Temperature in January-February and Its Relation to Large-Scale Circulation", *Clim. Res.* **3**. 195-207.

SÁNCHEZ, E., MONTÁVEZ, J.P., RODRÍGUEZ, A. y JIMÉNEZ, J.I. (1998): "Statistical Methods in Climatology: A Case Study", p. 133. *Applied Sciences and the Environment*, Wit Press, Southampton.

