

3

**Fuentes y métodos para la caracterización de ambientes
meteorológicos o 'tipos de tiempo'**

Domingo F. Rasilla Álvarez

UNIVERSIDAD DE CANTABRIA

1. INTRODUCCIÓN: ESTADO ACTUAL DE LA CLIMATOLOGÍA SINÓPTICA.

Nacida como disciplina científica en la Segunda Guerra Mundial (JACOBS, 1946), la Climatología Sinóptica ha experimentado una larga evolución. Concebida inicialmente para caracterizar los climas regionales a través del análisis retrospectivo de las variaciones de la circulación atmosférica (HARE, 1955; BARRY Y PERRY, 1973), en las 2 últimas décadas ha experimentado una profunda renovación metodológica (YARNAL, 1993), protagonizada por la introducción de la informática. Esta progresión se sostiene igualmente por una diversificación de su objeto de estudio, por la estrecha vinculación de muchos problemas ambientales con la dinámica atmosférica (p.e. la calidad del aire, los recursos hídricos o el cambio climático).

Esta actualización de los procedimientos ha mejorado su habitual método deductivo de análisis, cuyos denominadores comunes son:

- La OBSERVACIÓN DE LA ATMÓSFERA, que implica la clasificación de sus distintos estados, siguiendo dos criterios:
 - ⇒ La *similitud en sus propiedades* (p.e. una combinación de variables meteorológicas –temperatura, humedad etc...– características de una masa de aire).
 - ⇒ Sus *relaciones comunes* (localización de los centros de acción en un mapa de presión).
- Un ANÁLISIS ESTADÍSTICO, que cuantifica las variaciones de la frecuencia y/o propiedades internas de esas clases.
- La EXPLICACIÓN de la variabilidad espacial y temporal de un fenómeno, natural o antrópico, a través de los cambios que experimenta la circulación atmosférica.

La clasificación ordena y simplifica grandes volúmenes de información, que servirán para identificar los rasgos esenciales de la circulación atmosférica sobre el ámbito de estudio. Inicialmente, esta tarea se llevaba a cabo manualmente, siendo el investigador el que definía los criterios. Ello requería mucho tiempo y esfuerzo (lo cual, salvo excepciones, limitó su aplicación a períodos no superiores a 10 ó 15 años); además, la subjetividad aumentaba su inconsistencia interna, dificultando su duplicación en otro momento o su reproducción por otros autores. Las clasificaciones “automáticas”, es decir, las realizadas con la ayuda de los ordenadores, se han popularizado a partir de los años setenta, permitiendo los complejos cálculos numéricos asociados a la aplicación de diversos algoritmos de clasificación (LUND, 1963; KIR-



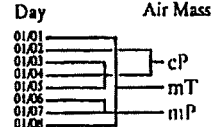

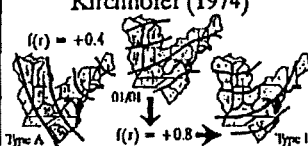
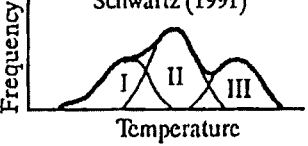



ESCALA ESPACIAL	FENÓMENO CLIMÁTICO	
	TIPOS DE CIRCULACIÓN	TIPOS DE TIEMPO
LOCALIDAD 	Muller (1977) Pielke et al. (1987) 	Christensen y Bryson (1966) Kalkstein y Corrigan (1986) 
REGIÓN 	Lund (1963) Kirchhofer (1974) 	Bryson (1966) Schwartz (1991) 
CONTINENTE 	Barnston y Livezey (1987) 	Davis y Kalkstein (1990) 

Figura 1: Tipo de enfoques sinópticos

CHOFFER, 1973), de técnicas estadísticas multivariadas (GALAMBOSI *et al*, 1995) o de inteligencia artificial (BARDOSSY *et al*, 1995).

Existen dos opciones a la hora de estudiar la variabilidad de la atmosfera (figura 1):

- A través de los TIPOS DE CIRCULACIÓN analizamos las diferentes configuraciones de un campo meteorológico, como la presión a nivel del mar, la altura geopotencial o los vientos.

- En los TIPOS DE TIEMPO, identificados como “el conjunto relativamente homogéneo de estados o ambientes atmosféricos producto de la interacción de los distintos factores dinámicos y geográficos”, los criterios de análisis son las propiedades termodinámicas de una masa de aire, plasmadas en una serie de parámetros meteorológicos como la temperatura, la humedad, la nubosidad etc....

Esta dualidad deriva de la no completa correspondencia entre el tipo de circulación y el tipo de tiempo. No siempre la misma situación sinóptica es responsable de las mismas condiciones atmosféricas; en principio, por la dificultad de acotar precisamente todas las estructuras de la circulación atmosférica; en segundo lugar, porque las características de las masas de aire asociadas a una misma situación sinóptica varían estacionalmente. Pero además, por la variedad de masas de aire involucradas en una misma situación sinóptica; en el modelo noruego de perturbación pueden sucederse distintos tipos de tiempo asociados al paso de sucesivas masas de aire y discontinuidades frontales. La primera opción ha sido utilizada para la descripción climática regional, su variabilidad temporal o los mecanismos de transporte en la atmósfera; la segunda es aconsejable cuando existan procesos vinculados a esas propiedades termodinámicas (contaminación, salud etc...). Sin embargo, ambos enfoques no son excluyentes, sino complementarios; tal y como sugiere Durand Dastés (1973), la climatología regional necesita contrastar los resultados procedentes de las dos opciones.

El concepto de tipo de tiempo se remonta a comienzos de siglo, cuando se relacionó el clima con los seres vivos (FEDEROV, 1927). Para resaltar la respuesta fisiológica de los segundos, las primeras clasificaciones de tipos de tiempo priorizaban, bien un parámetro meteorológico aislado, bien una combinación específica de variables, cuyo efecto era determinante sobre la naturaleza o las actividades humanas. El tiempo o ambiente atmosférico de cada día era identificado con un código de cuatro letras que permitía la definición de sus principales atributos. Más tarde, este concepto fue reformulado por Christensen y Bryson (1996) que lo consideraron como “el conjunto de observaciones meteorológicas que exhibían un elevado grado de consistencia interna”.

La complejidad de las relaciones entre la atmósfera, los seres vivos o los distintos fenómenos naturales o antrópicos, unido a la imposibilidad de manejar simultáneamente el gran número de parámetros atmosféricos disponibles en las bases de datos meteorológicas, dificultó la definición de estos ambientes atmosféricos (COMELLAS, 1964); de hecho, el análisis de tipos de circulación ha sido tradicionalmente, el enfoque elegido por la mayoría de los investigadores, particularmente en la Península Ibérica (OLCINA, 1994). En otros casos, se recurrió a simplificar el número de parámetros meteorológicos empleados para la caracterización sinóptica (presión, tempera-

tura y precipitación), cuyos umbrales, elegidos de manera arbitraria servían para definir los tipos de tiempo (CREUS, 1977; RUIZ URRESTARAZU, 1982; PEJENAUTE GOÑI, 1990). Esta aproximación es insuficiente, no sólo por la rigidez impuesta por dichos umbrales, sino porque prioriza una serie de variables cuya incidencia sobre los procesos ambientales puede ser nula por sí mismos. Éstos últimos responden al efecto combinado de diversas variables meteorológicas; por ejemplo, la irradiación asociada a la ausencia de nubes acentúa el efecto de las altas temperaturas, mientras el viento ejerce un efecto desecante que alivia el rigor de las jornadas estivales, si bien en invierno acentúa la sensación de frío.

En consecuencia, un tipo de tiempo debería definirse por la combinación de un número elevado de parámetros que reproduzcan con la mayor fidelidad posible la variedad de estados de la atmósfera sobre un lugar. Uno de los métodos más extendidos para la obtención de tipos de tiempo siguiendo esta hipótesis es el Índice Sinóptico Temporal (TSI; Temporal Synoptic Index; KALKSTEIN *et al*, 1988), diseñado por el Synoptic Lab del Departamento de Geografía de la Universidad de Delaware (EEUU). Este método implica el uso de dos procedimientos estadísticos multivariados, el análisis en componentes principales y el análisis de conglomerados (“cluster analysis”), habiendo sido usado para evaluar distintos problemas ambientales como:

- Calidad del aire (DAVIS, 1991; KALKSTEIN *et al*, 1990; KALKSTEIN y CORRIGAN, 1986; EDER *et al*, 1994; MCGREGOR y BAMZELIS, 1995; SHAHGEDANOVA *et al*, 1998).
- Distribución de la precipitación (MCCABE *et al*, 1989; GREENE, 1996).
- Salud (KALKSTEIN, 1993; GREENE, y KALKSTEIN, 1996; MCGREGOR *et al*, 1999; SMOYER *et al*, 2000)
- Cambio climático (KALKSTEIN *et al*, 1990; 1993).
- Productividad agrícola (DILLEY, 1992; JONES y DAVIS, 2000).
- Acumulación y ablación glacial (BRAZELET *et al*, 1992).
- Fenómenos extremos (DAVIS y ROGERS, 1992; ELLIS y LEATHERS, 1996).

A pesar de gozar de un sólida reputación, su empleo en España no ha sido frecuente (LANA y FERNÁNDEZ MILLS, 1994; SERRA *et al*, 1999), si bien en los últimos años se asiste a su creciente aplicación, sobre todo aplicado a problemas ambientales muy diferentes como la contaminación atmosférica (FERNÁNDEZ GARCÏA y RASILLA, 1999; FERNÁNDEZ GARCÏA *et al*, 2001)

o los incendios forestales (DIEGO *et al.*, 2001). En la misma línea se podría incluir otros estudios que, sin seguir estrictamente los pasos que enunciaremos a continuación, pueden integrarse perfectamente por sus objetivos y metodología (SÁNCHEZ *et al.*, 1990; TUDURI y RAMIS, 1997).

2. PAUTAS PARA LA ELABORACIÓN DEL ÍNDICE SINÓPTICO TEMPORAL.

Los procedimientos multivariados requeridos para la elaboración del Índice Sinóptico Temporal están incluidos en la mayoría de los paquetes estadísticos de uso común (SPSS, SAS, STATGRAPHICS etc...), por lo que se remite al lector a consultar los numerosos manuales existentes para conocer su implementación. Este trabajo está pensado exclusivamente como una guía práctica para resolver los problemas que surgirán al reproducir esta metodología, ya que, aunque existen publicaciones centradas en el uso de esas técnicas multivariadas en Climatología (PREISENDORFER, 1988; JOLIFFE, 1986; WILKS, 1996; von STORCH y ZWIERS, 1999), el investigador poco acostumbrado al manejo de una terminología anglosajona puede encontrar difícil su comprensión.

2.1. PREPARACIÓN INICIAL DE LOS DATOS CLIMÁTICOS

Los datos básicos son una serie de parámetros meteorológicos usuales en las estaciones de la red sinóptica mundial, donde se llevan a cabo entre 3 y 4 observaciones diarias (00, 07, 13 y 18 TMG). Estos datos, procedentes de un único observatorio meteorológico, deben abarcar un periodo temporal más o menos dilatado (serie climática). Aunque su número y naturaleza puede variar, los más frecuentes son los siguientes:

- Temperatura del termómetro seco.
- Temperatura del termómetro húmedo/temperatura del punto de rocío.
- Dirección y velocidad del viento.
- Nubosidad e insolación.
- Precipitación.
- Presión atmosférica.
- Visibilidad.

A ellos se añaden otros parámetros con periodicidad diaria, como las temperaturas máxima y mínima, la evaporación y la insolación. En algunos casos se han agregado datos de temperatura, humedad, viento y altura de la topografía en distintos niveles atmosféricos, suministrados por radiosondas,

si bien la experiencia personal permite afirmar que no mejoran esencialmente los resultados obtenidos con datos de superficie. Su disposición debe acomodarse a una matriz P (RICHMAN, 1985); cada parámetro constituye una variable, y cada miembro de la serie temporal cada uno de los casos.

Un problema frecuente en las series meteorológicas es la ausencia de datos. Si el número de lagunas es reducido y se distribuyen aleatoriamente durante el intervalo temporal disponible, puede optarse por alguno de los métodos de relleno incluidos en los paquetes estadísticos (sustitución por la media, promedio del día anterior y posterior, tendencia en ese punto de la serie, imputación por regresión etc...). A escala diaria, este relleno es facilitado por la alta autocorrelación temporal de algunas variables meteorológicas (p.e. temperatura, humedad y presión etc...). Sin embargo, su abundancia y su concentración durante un periodo concreto de la serie puede comprometer cualquier análisis posterior, siendo aconsejable prescindir de ellos.

Algunos de los citados parámetros meteorológicos muestran un ciclo anual, resultado de la incidencia de factores ajenos a la circulación atmosférica (p.e. radiación solar). Esta distorsión puede resolverse subdividiendo la matriz original según las estaciones astronómicas o meses, y repitiendo los cálculos tantas veces como sea necesario. Además de multiplicar el trabajo, los resultados no son directamente comparables, debido a la inclusión de estaciones de gran variabilidad (primavera y otoño) entre los relativamente estables promedios estivales e invernales. Otra solución es resaltar las oscilaciones de escala sinóptica, centrando las series originales respecto a un "estado medio de la atmósfera". A la hora de elaborar ese promedio aparecen varias soluciones en la literatura científica:

- Si se dispone de series largas (p.e. 30 años) se pueden calcular según días julianos (desde el 1 de enero -1- al 31 de diciembre -365-). El riesgo de inhomogeneidades (cambios de emplazamiento, de instrumental o de métodos de observación etc...), por no hablar de la inclusión de periodos climáticos con características diferentes, obliga a tomar con precaución esta posibilidad.
- Disponer de series climáticas diarias con una duración menor (p.e. 10 años) es lo más frecuente, pero el centrado según días julianos implica una especial sensibilidad a eventos excepcionales. Esta circunstancia se puede soslayar aplicando un nuevo filtrado consistente en suavizar esos promedios según días julianos. El más sencillo de los procedimientos es usar medias

móviles de 31 días; algo más complejo es ajustar funciones de seno y coseno de periodos de un año y medio año (Galambosi *et al*, 1996). También puede obtenerse la longitud óptima de ese filtro a través del grado de autocorrelación temporal y de las periodicidades fundamentales (análisis espectral) de las variables meteorológicas (HEWITSON y CRANE, 1992). Una vez obtenida esa periodicidad (que en latitudes medias suele oscilar entre 11 y 15 días), los valores originales se centran con respecto al promedio móvil de los n días en torno al día en cuestión (p.e. en caso de aplicar el de 11 días, del día $t-5$ al $t+5$ alrededor del día $t=0$; figura 2). Además de su sencillez, otras ventajas evidentes de este filtro móvil es su utilidad para bases de datos pequeñas, y la atenuación del efecto de las inhomogeneidades sobre los promedios (figura 2).

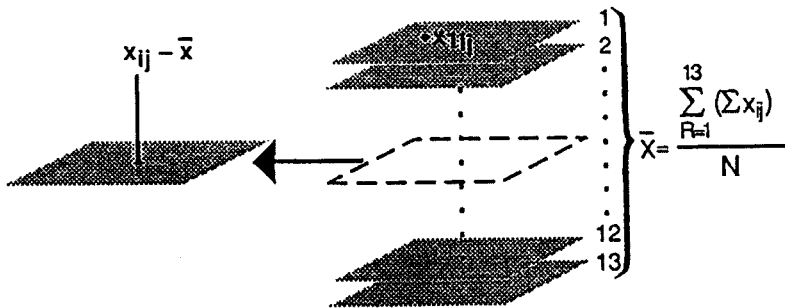


Figura 2: Desestacionalización con el método de medias móviles (en este caso, de 13 días)

La métrica original de algunas variables meteorológicas no es apropiada para su inclusión en procedimientos multivariados, al no cumplir los supuestos estadísticos que subyacen en este tipo de análisis (por ejemplo, el de normalidad). Este es el caso del viento, magnitud vectorial que posee una dirección medida en grados sexagesimales. Desde el punto de vista exclusivamente matemático, un viento NNW (dirección de 340°) está más separado de un NNE (20°) que de un SW (220°). Por ello debe proyectarse en sus componentes escalares zonal y meridiano, mediante la siguiente fórmula:

- componente zonal (E-W) $u = -V_h \sin \varnothing$
- componente meridiana (N-S) $v = -V_h \cos \varnothing$

siendo V_h la velocidad del viento en km/h y \varnothing el azimut de la dirección del viento. Igualmente, la asimetría que caracteriza la distribución de

las precipitaciones también perturba esos supuestos. La literatura sugiere diversas transformaciones para encoger la cola de la distribución de frecuencias. Entre ellos destaca el cálculo de su logaritmo neperiano, o de sus raíces cuadrada o cúbicas.

Queda finalmente como última precaución identificar los casos anómalos, frecuentes en las bases de datos meteorológicos a causa de fenómenos como olas de calor o de frío, temporales de viento o precipitaciones torrenciales. Salvo que presenten valores aberrantes, estos casos no deberían eliminarse, pues precisamente suelen estar asociados a crisis ambientales, como altas tasas

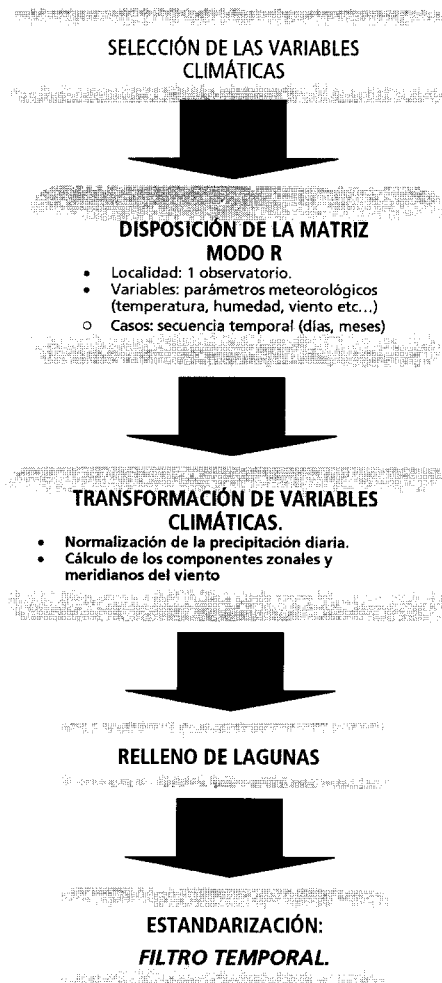


Figura 3: secuencia de procedimientos para la preparación inicial de los datos

de mortalidad, elevadas concentraciones de contaminantes etc... El procedimiento de filtrado móvil tiene además la virtud adicional de amortiguar estas situaciones excepcionales; por otro lado, el análisis en componentes principales constituye un segundo filtro que suaviza adicionalmente los valores climáticos. Una vez efectuadas estas modificaciones (figura 3), la base de datos inicial es sometida al primer procedimiento multivariado, el análisis en componentes principales.

2.2. REDUCCIÓN DE LA INFORMACIÓN: ANÁLISIS EN COMPONENTES PRINCIPALES

La fase de reducción de la información comienza sometiendo la matriz inicial, centrada y transformada, a un análisis en componentes principales. Este método multivariado es, posiblemente, el más extendido en el campo de las ciencias atmosféricas (LORENZ, 1967; KUZBACH, 1967). Su empleo se justifica por la necesidad de reflejar la actuación simultánea y combinada de los diversos parámetros meteorológicos sobre los fenómenos que constituyen el mundo real. Sin embargo, su evolución temporal puede estar a su vez condicionada entre sí; esta multicolinealidad aporta información redundante que complica los cálculos estadísticos sin aportar beneficios. Gracias al análisis en componentes principales podemos eliminar este exceso de información, condensando la más relevante. En consecuencia, esta técnica simplifica un conjunto de variables iniciales en una serie más reducida de nuevas variables o *componentes principales*, combinación de las anteriores. Como las nuevas variables son ortogonales (no están correlacionadas entre sí), se convierten en índices climáticos simples que pueden trabajarse por sí mismo (p.e. como variables independientes en un análisis de regresión múltiple), o mediante otros procedimientos estadísticos.

Esta reducción de la información se logra si las variables iniciales están correlacionadas entre sí; para confirmar esta condición existen algunos test incluidos en la mayoría de los paquetes estadísticos (p.e. el test de Bartlett). Debe tenerse en cuenta que alguno de ellos es sensible a la amplitud de la muestra, de tal manera que en grandes bases de datos pequeñas correlaciones son estadísticamente significativas pero climáticamente prescindibles.

Antes de proceder a los cálculos estadísticos requeridos para la obtención de los componentes principales es necesario elegir el tipo de matriz de dispersión (describe cómo las observaciones se dispersan alrededor de su vector medio en el espacio de K -dimensiones definido por las K -variables). Para este

tipo de análisis es recomendable emplear una matriz de correlación, ya que la matriz original combina variables cuantificadas en unidades diferentes (metros, °C, octas etc...). Una matriz de covarianza proporcionaría más peso estadísticos a los parámetros climáticos con mayor variabilidad, sesgando los resultados.

Una vez efectuados los cálculos pertinentes, el ordenador proporciona una serie de matrices que deben ser interpretadas desde el punto de vista climático (figura 4).

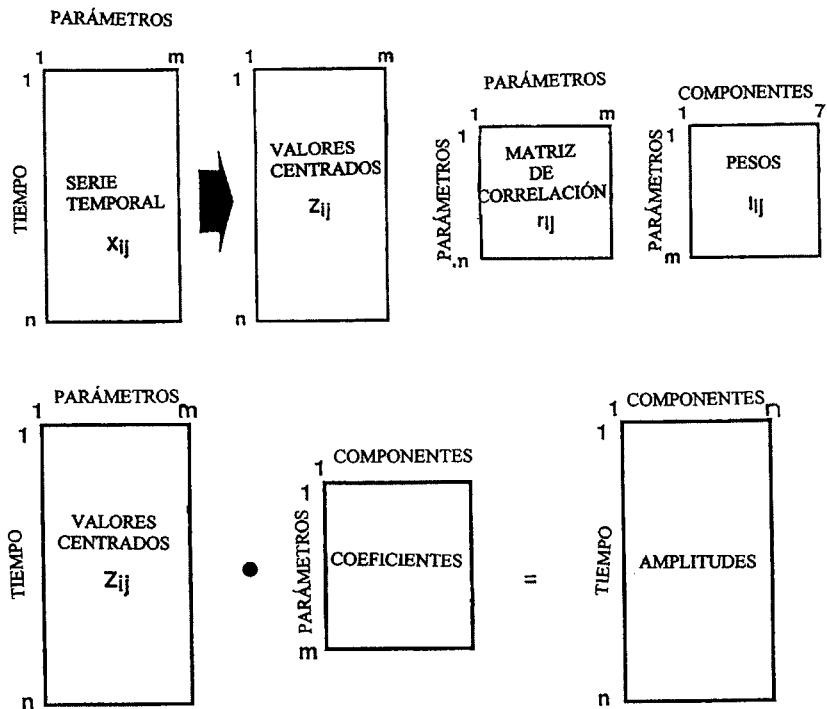


Figura 4: Matrices producto de un análisis en componentes principales

- La matriz de componentes o vectores propios (autovectores), que numera las nuevas variables en sentido decreciente y en función de su grado de explicación de la varianza original.
- La matriz de pesos o ponderaciones (*loadings*) describe la importancia de cada variable original en la conformación de los componentes principales:

cuanto más alto es su peso, más influencia tiene en la formación de ese componente. El valor de cada elemento de esta matriz equivale al coeficiente de correlación entre la variable original y el componente principal. No está muy claro umbral mínimo necesario para considerar significativa la influencia de un parámetro original; como regla de compromiso se plantea que es indispensable un valor superior a 0,4 ó 0,5. Desde el punto de vista climático esta matriz reproduce los mecanismos de variabilidad atmosférica presentes en los datos originales; normalmente, el primer componente principal es un índice térmico al que acompañan variables secundarias, cuya naturaleza es diferente en cada observatorio. Otros componentes equivalen a índices de nubosidad, insolación y visibilidad, o reflejan el grado de estabilidad de la atmósfera a través del efecto combinado de la presión y el viento.

- La matriz de puntuaciones o amplitudes (*scores*) es resultado de la multiplicación de la matriz de dispersión por la matriz de ponderaciones. Refleja la importancia que los mecanismos citados tienen en un momento determinado, y por lo tanto, nos permiten caracterizarlo con respecto a los días anteriores y posteriores. Normalmente, la mayoría de los paquetes estadísticos suministran estas puntuaciones estandarizadas, con media 0 y desviación típica 1.

Aunque este procedimiento extrae tantos componentes principales como variables iniciales, sólo un reducido número, que describen la mayor parte de la varianza original, proporciona información relevante (“señal climática”). El resto, que reproduce procesos mesoescalares o bien simples errores, es considerado “ruido climático”, y como tal, debe ser eliminado. Este proceso o “truncación de los componentes” constituye un momento delicado, sobre todo si viene seguido de una rotación, ya que los resultados finales dependerán en buena medida del acierto en esa selección. Para elaborar el ITS no es necesaria la rotación, por lo que la selección del número de componentes no constituyen un paso tan delicado como en otras aplicaciones climáticas, como en las regionalizaciones. En cualquier caso, establecer cuantos componentes principales resumen la parte substancial de la información descansa muy a menudo en el propio juicio del investigador, siendo recomendable que la decisión se fundamente en la interpretabilidad de los mecanismos climáticos reproducidos por los componentes principales. No obstante, para ayudar en la toma de decisiones se han diseñado una serie de tests (figura 5):

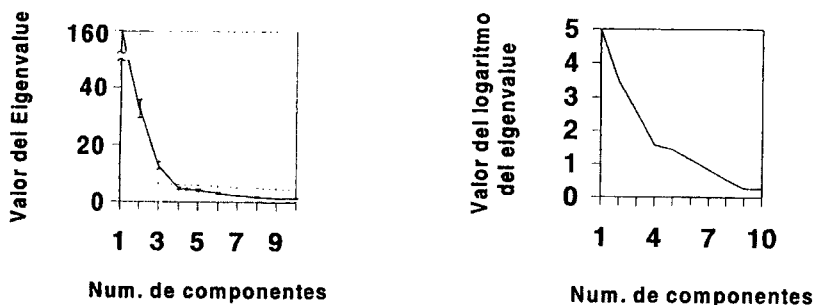


FIGURA 5: a) Gráfico de sedimentación con el valor de cada vector, el resultado de la aplicación de la regla N (trazo de puntos) y error típico según el test de separación de cada vector (barras) b) gráfico de sedimentación con el logaritmo de los autovectores (LEV)

- Los más sencillos aconsejan mantener aquellos componentes cuyos autovalores son superiores a la unidad o al 5 % de la varianza original (criterio de raíz latente, JOLLIFFE, 1986), opción por defecto en la mayoría de los paquetes estadísticos. Según ellos, el grado de explicación proporcionado por un componente debería ser, como mínimo, igual que la información suministrada por cualquiera de las variables originales. La cifra total del 80 % de varianza explicada es considerada por otros autores como deseable (criterio de porcentaje de varianza, o regla de Kaiser), aunque en análisis climáticos es difícil conseguir esta cifra, dado que los parámetros meteorológicos empleados en este método incluyen mucho “ruido”. Aunque fáciles de aplicar, establecen umbrales rígidos y arbitrarios de difícil justificación.
- Muy populares son los test de naturaleza gráfica (figura 4), como el de sedimentación o “Scree test” (CATTELL, 1966), también habitual en todos los programas estadísticos. Consiste en la representación gráfica del valor de cada componente en un eje de ordenadas y de su orden decreciente en la abscisa. El número de componentes a eliminar viene dado por un trazado plano del gráfico, tras una ruptura de la pendiente. También con lleva una gran subjetividad, dado que descansa en una inspección visual: en algunos casos, en vez de una nítida ruptura de pendiente se obtiene una curva. El gráfico LEV (logaritmo de los autovalores) es una variante en la que el valor original del autovector se sustituye por su logaritmo. Este último test plantea que cuando los componentes reproducen ruido, las magnitudes de sus autovalores deberían disminuir exponencialmente, lo que se refleja en una línea recta en el diagrama.

- Los test basados en técnicas de muestreo son cada vez más frecuentes. Los más utilizados son:

⇒ La regla N (OVERLAND y PREISENDOFER, 1982) consiste en la aplicación de sucesivos análisis a 100 matrices elaboradas aleatoriamente a partir de una distribución normal, compuestas del mismo número de casos y variables que la real. Se consideran significativos todos los autovalores reales superiores al percentil 95 de los simulados; un valor inferior implica que no podrían distinguirse de los causados por procesos aleatorios.

⇒ En el test de separación de autovalores (NORTH *et al*, 1982) subyace la idea de que los componentes principales están elaborados a partir de un número finito de muestras (N) que reproducen el estado instantáneo de uno o varios campos meteorológicos. La independencia estadística de cada componente es determinada representando el error típico de cada autovalor, dado por la fórmula $(2/N)^{0.5}$. Si las barras de error de dos autovalores consecutivos se superponen, es posible que, con una muestra diferente del mismo campo meteorológico, los componentes principales tendrían una estructura diferente.

Diversos estudios han mostrado que ninguno de los anteriores tests está libre de inconvenientes, de ahí que la selección de los componentes debe atender además a criterios como el propósito de nuestro estudio, el tipo de datos, y el equilibrio entre parsimonia e información.

Para mejorar la interpretación climática de los componentes es habitual recurrir a una transformación matemática conocida como rotación. Esta busca la obtención de una “estructura simple”, que se alcanza cuando el valor de la mayoría de las ponderaciones de un componente se aproximan a 0, mientras los restantes alcanzan valores muy altos. El resultado es que los vectores rotados enfatizan el impacto de algunas variables concretas, las que poseen ponderaciones más altas, sin contaminación de las restantes. Sin embargo, como la aplicación del análisis en componentes principales en este método busca simplemente la máxima reducción de la información original y la comprensión de las interrelaciones entre los parámetros meteorológicos iniciales, es recomendable no efectuar esta transformación matemática.

El último paso es el cálculo de las puntuaciones o amplitudes de los componentes que, al cuantificar la magnitud e importancia de los mecanismos dinámicos responsables del estado de la atmósfera, se convierten en las variables sobre las que se aplica la segunda fase (figura 6).

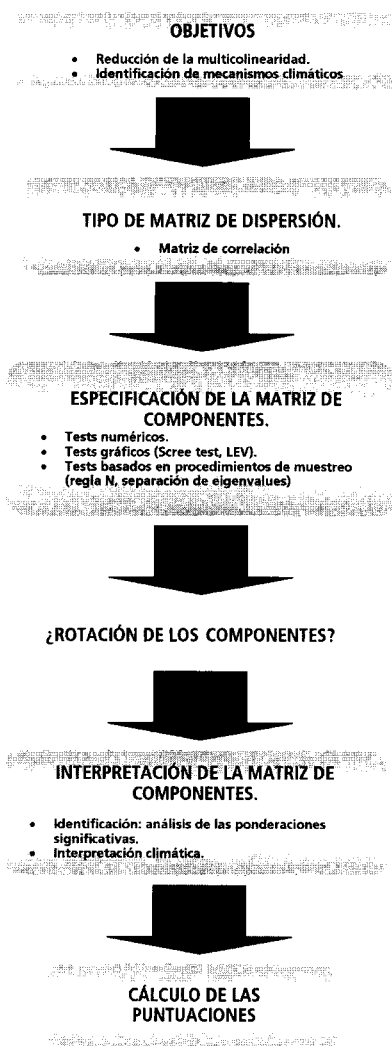


Figura 6: Secuencia de la fase de reducción de la información. Análisis en componentes principales.

2.3 CLASIFICACIÓN DE LA INFORMACIÓN (ANÁLISIS DE CONGLOMERADOS)

La siguiente fase involucra la aplicación de una técnica multivariada conocida como análisis de conglomerados (*cluster analysis*). Esta es la denominación de un conjunto de técnicas multivariadas cuyo propósito es separar objetos en grupos cuya composición y número son desconocidos de antemano, a través del grado de similitud o diferencia que poseen esos objetos según sus atributos o características. Estos grupos deben cumplir dos condiciones:

- Homogeneidad interna: las observaciones pertenecientes al mismo grupo deben ser similares entre sí (minimización de las diferencias entre sus miembros).
- Sustanciales diferencias entre los distintos grupos: las observaciones de una clase deben ser diferentes de las observaciones de los restantes (maximización de las diferencias).

En el caso que ocupa, los objetos son los días en los que se han efectuado unas observaciones meteorológicas, cuyos atributos son las puntuaciones de los componentes principales; cada clase, grupo o conjunto que se obtenga del análisis de conglomerados constituirá un tipo de tiempo.

Siendo la atmósfera dinámica y la transición de unos estados a otros paulatina, los tipos de tiempo son, en realidad, una construcción teórica para simplificar una realidad compleja, y generalizar una respuesta a las variaciones de la circulación atmosférica. El análisis de conglomerados es una técnica exploratoria excelente, por su capacidad para estratificar el número casi infinito de posibles combinaciones entre las citadas puntuaciones. Desde esta óptica es preferible a otras técnicas multivariadas similares, como el análisis discriminante, pues no suministra regla alguna para determinar la pertenencia de nuevos casos a grupos ya definidos, pero sí permite establecer diferentes niveles de agregación.

El primer paso es la selección de una métrica que reproduzca el grado de similitud entre los casos. Ya que no se trata de evaluar el parecido en las formas, sino la magnitud del espacio multidimensional en el que se asientan los casos, las distancias y pseudodistancias son más indicadas que los índices de correlación. Ente las primeras destaca la euclidiana, aunque se han empleado otras, como la cuadrática o el ángulo entre pares de vectores (ROMESBERG, 1984).

Otra importante decisión es la selección del tipo de algoritmo que evaluará las similitudes y diferencias entre las observaciones y agrupará los casos. En los algoritmos jerárquicos el proceso de agregación se inicia considerando la ausencia de vínculos entre las observaciones (cada una constituye un grupo). En el siguiente paso, las dos más próximas se combinan en un primer grupo, apareciendo $n-1$ grupos, uno de los cuales tiene ahora dos miembros. En pasos sucesivos, a este primer grupo de dos miembros se le pueden agregar nuevos casos, o constituirse nuevos grupos a partir de las observaciones libres. Este proceso continuará hasta que, al final, todas las for-

men un único grupo. Cada algoritmo jerárquico calcula de manera diferente la distancia entre los grupos, y esta formulación matemática pueden distorsionar la clasificación a través de dos efectos:

- El efecto de “ENCADENAMIENTO” o asignación de nuevas observaciones a grupos ya existentes en vez de formar nuevos grupos (si bien puede utilizarse para identificar soluciones con grupos heterogéneos).
- El efecto “BOLA DE NIEVE” consiste en la generación de un grupo compuesto por numerosos casos y una constelación de grupos de mucho menor tamaño, generados a partir de casos atípicos. Este efecto no es del todo perjudicial, ya que detecta casos anómalos –cuando su frecuencia no supera el 5 % del total–.

Dos algoritmos presentes en la mayoría de los paquetes estadísticos son recomendados en la literatura climática (KALKSTEIN *et al*, 1987; YARNAL, 1993): el de *Ward* (pero que tiende a crear clusters compactos y casi de igual tamaño) y el *encadenamiento medio* (*average linkage*, puede generar bolas de nieve). Sin embargo, antes de proceder a una elección definitiva, es recomendable un buen conocimiento del grado de dispersión de los datos, y el contraste de los resultados obtenidos con diferentes métodos, ya que las coincidencias existentes reflejarán procesos reales y no artificios matemáticos

Los algoritmos jerárquicos no requieren un conocimiento previo del número de grupos o de la partición inicial, pero presentan algunas limitaciones. Por ejemplo, combinaciones claramente mejorables pueden persistir a lo largo de todo el proceso, ya que una observación incluida inicialmente dentro de un grupo no puede ser reasignada posteriormente a otro, generando resultados erróneos, como ocurre frecuentemente con los casos atípicos. Finalmente, su implementación es costosa, pues a medida que aumenta el tamaño de la matriz las necesidades de memoria del ordenador crecen exponencialmente, limitando su aplicabilidad sobre bases de datos de gran tamaño. La fragmentación de la matriz inicial en submuestras aleatorias como alternativa no suele ser empleada, pues debilita la estabilidad de la clasificación final.

Los algoritmos aglomerativos, entre los que *k-means* ha alcanzado gran preeminencia, clasifican las observaciones en un número predefinido de grupos. El proceso de clasificación comienza con la selección de los centroides (“semillas”) de los grupos iniciales; más adelante, se asigna cada observación al grupo cuyo centroide se encuentra más próximo y se recalcula de nuevo

ese centroide. Este proceso se repite hasta que no sea posible reasignar más observaciones, o sea, cuando todos los casos se incorporan al grupo cuyo centroide está más próximo, o finalmente cuando se satisfaga algún otro tipo de criterio adicional. Esta naturaleza iterativa optimiza la composición de los grupos, evitando las consecuencias derivadas de la inclusión de casos anómalos y de las diferentes medidas de distancia, y puede conseguir una partición robusta si las “semillas” se eligen adecuadamente (MILLIGAN y COOPER, 1980), pero requiere un conocimiento previo del número de grupos a extraer y son muy sensibles a la bondad de la partición inicial.

Dadas las ventajas e inconvenientes de ambos tipos de algoritmos, se recomienda su uso combinado. El número de grupos y sus centroides se obtienen a partir de un análisis jerárquico, repitiendo el procedimiento para reducir la incidencia de los casos atípicos. Una vez obtenida una solución suficientemente estable, los centroides servirán como semillas en un procedimiento aglomerativo que ajustará de manera óptima la composición de los grupos. En definitiva, los métodos jerárquicos y aglomerativos son complementarios, adquiriendo los primeros carácter exploratorio y los segundos confirmatorio.

La elección del número de grupos es la última decisión de importancia en el proceso de clasificación. El carácter continuo de la atmósfera impide que la estructura de estas clasificaciones sea completamente estable, ya que los límites entre los tipos de tiempo serán arbitrarios. La elección del análisis de conglomerados implica la búsqueda de la mejor clasificación posible a través de diferentes niveles de agregación. Ni la partición inicial (cada caso constituye una clase) ni la final (todos los casos están englobados en el mismo grupo) constituyen la respuesta al problema. Su aplicación presupone que entre esas dos opciones extremas existen uno o varios niveles de agrupamiento, que proporcionan una buena capacidad explicativa del fenómeno natural o antrópico estudiado, y acordes a unas hipótesis climáticas de partida (tipos de tiempo “naturales”).

Frente a otros procedimientos multivariados, el análisis de conglomerados no puede utilizar los mismos criterios estadísticos para la inferencia. Para ayudar en el proceso de clasificación y solventar este inconveniente, se han diseñado otros métodos, conocidos como “reglas de parada”. Algunos utilizan alguna medida de similitud o distancia entre los conglomerados en cada paso sucesivo (figura 7). La progresión del agrupamiento y sus pasos

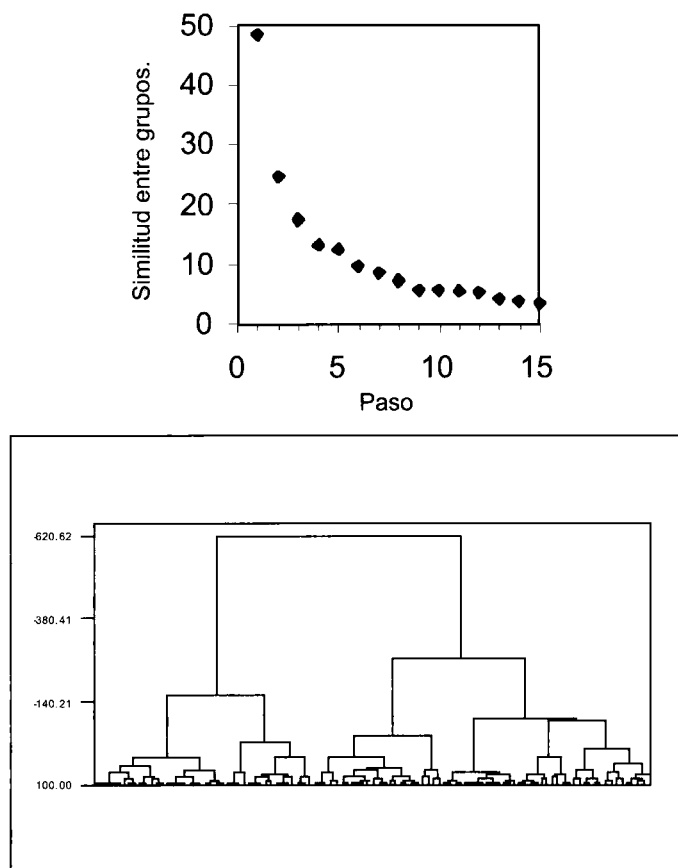


Figura 7: Dendrograma y evolución del índice de similitud entre grupos en sucesivos estadios de agrupación

intermedios se reproducen en el dendrograma, o diagrama arborescente, compuesto de una serie de ramas que se unen a medida que disminuye el número de grupos, proporcionando además información de las distancias entre los grupos. La detención del proceso de agrupamiento deberá ocurrir cuando esta medida exceda un valor prefijado, o cuando este valor experimente una brusca variación, seleccionando la solución previa a ese destacado incremento, lo cual reflejaría una reducción importante en la homogeneidad interna de los grupos. Otras reglas de parada convierten conocidos test de significación en pseudotests estadísticos (caso de los coeficientes de fusión, el R^2 , el Pseudo-Fy el Pseudo- t^2). No obstante, y al igual que ocurre con el análisis en componentes principales, las soluciones teóricas proporcionadas por estas regla

de parada deben contrastarse con las que proporciona el sentido común, basado en el conocimiento que el investigador posee de la dinámica atmosférica regional y con el grado de explicación que proporcionan acerca del fenómeno estudiado, teniendo siempre en cuenta la búsqueda de la parsimonia científica.

Dado el grado de subjetividad que rodea al análisis de conglomerados, es esencial que la solución elegida sea lo más estable y representativa posible, pudiendo con ello lograrse su generalización a nuevos casos; para ello se suelen emplear distintos procedimientos de validación. Uno de ellos es reservar una submuestra aleatoria y repetir en ella todos los pasos efectuados hasta ahora; también puede lograrse clasificando esa submuestra a partir de los centroides de la muestra original. Finalmente, cada vez es más frecuente la aplicación de técnicas de muestreo aleatorio, al igual que procedimientos de validación cruzada y validación externa.

Finalmente, el IST concluye con la interpretación climática de los grupos. Esta interpretación puede llevarse a cabo sobre sus centroides, pero debe recordarse que están calculados a partir de las puntuaciones o amplitudes de los componentes principales. Su interpretación, en consecuencia, es difícil, ya que constituyen unos índices sintéticos de variabilidad temporal respecto a unos promedios filtrados. Por ello, es más sencillo recurrir a los datos originales, calculando los perfiles medios de cada grupo, proporcionándoles finalmente una etiqueta que reproduzca sus características esenciales. Esta interpretación proporcionará además un medio para evaluar el grado de correspondencia de la solución propuesta y los tipos de tiempo definidos en otras clasificaciones, incluso de carácter manual, al igual que con los tipos de circulación más frecuentes en la zona estudiada (figura 8).

3. NUEVAS ORIENTACIONES METODOLÓGICAS.

El método anteriormente expuesto, pese a haber alcanzado un grado de madurez considerable, presenta algunos inconvenientes. Algunos están ligados al planteamiento matemático que subyace en cada técnica estadística. Otros descansan en la propia naturaleza dinámica del fenómeno estudiado. Su pretendida objetividad se convierte, finalmente, en un serio inconveniente, pues el investigador controla marginalmente el proceso de clasificación, pudiendo el agrupamiento reflejar únicamente relaciones matemáticas definidas por medio de medidas de similitud.

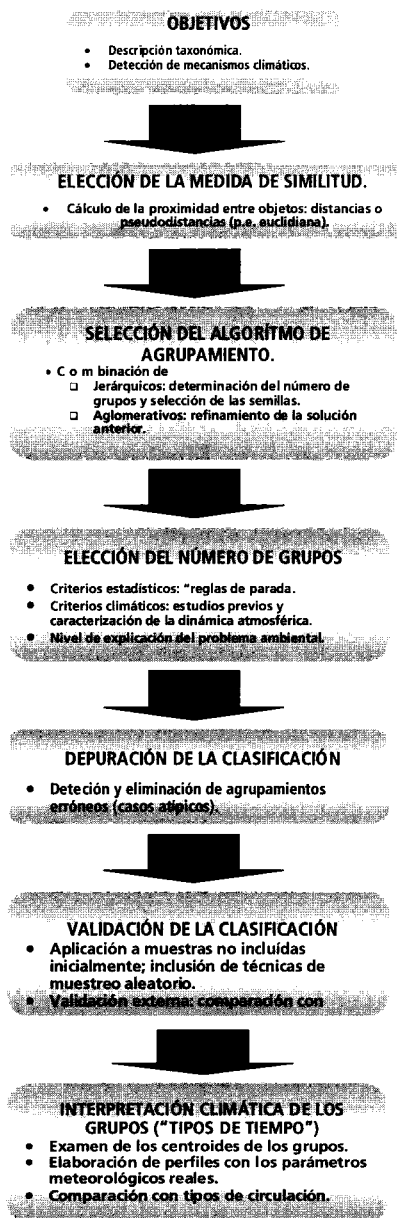


Figura 8: Secuencia de procedimientos para la fase de clasificación: análisis conglomerados

En las clasificaciones manuales, por el contrario, el investigador toma en consideración información muy variada que ayuda en el proceso de toma de decisiones.

Con objeto de mejorar los resultados se han planteado una serie de soluciones orientadas en dos líneas de investigación. La primera es la aplicación de técnicas que permiten aprehender la naturaleza no lineal de la dinámica atmosférica; este es el caso de técnicas como la lógica difusa (FUZZY LOGIC), las redes neuronales artificiales (ARTIFICIAL NEURAL NETWORKS) y los mapas autoorganizados (SELF ORGANIZING MAPS).

La segunda es el desarrollo de métodos “híbridos” combinando las ventajas de los métodos manuales y de las clasificaciones automáticas (FRANKS y YARNAL, 1997). El número inicial de grupos y sus características son establecidos a priori por el investigador sobre una muestra reducida de datos (entre el 10 y el 15 % de la población sería recomendable), basándose en su conocimiento de la dinámica atmosférica regional; posteriormente, los procedimientos automáticos antes citados se aplicarían al resto de la muestra. Uno de los ejemplos más claros de esta nueva orientación es el Spatial Synoptic Classification (KALKSTEIN *et al*, 1996). A diferencia de la propuesta original del TSI, el SSC requiere la identificación previa de las principales masas de aire que atraviesan una región, así como sus características meteorológicas más relevantes. La práctica climatológica juega un papel fundamental, pues determina los límites y las características de esas masas de aire a través de una cuidadosa selección de los días representativos de cada clase. El número de días representativos debería ser significativo, pero excluyendo aquellos casos que supongan una merma de su homogeneidad. Estas semillas serán usadas posteriormente para clasificar todos los restantes días a través del análisis discriminante, una técnica de agrupamiento estadística y conceptualmente diferente del análisis de conglomerados (CAÑADA, 1989; GALÁN, 1989; RASILLA, 1998).

BIBLIOGRAFÍA

- ANDEBERG, M. R. (1973). *Cluster Analysis for Applications*. New York, Academic Press.
- BARDOSSY, A.; DUCKSTEIN, L. Y BOGARDI, Y. (1995). "Fuzzy rule-based classification of atmospheric circulation patterns". *Int. J. Climatol*, 15, pp. 1087-1097.
- BARRY, R. G. y A. H. PERRY (1973). *Synoptic Climatology Methods and Applications*, Methuen & Co. Ltd., London, 555 pp.
- BRAZEL, A. J.; CHAMBERS, F.B. Y KALKSTEIN, L.S. (1992). "Summer energy balance on west Gulkana glacier, Alaska, and linkages to a temporal synoptic index". *Z. Geomorph. N.F.*, Suppl.Bd. 86, pp. 15-34.
- BRYSON, R. A. (1966). "Air masses, streamlines and the Boreal forest". *Geographical Bulletin*, 8, pp. 228-269.
- CAÑADA TORRECILLA, R. (1989) *El clima de Extremadura*. Tesis Doctoral, Dep. de Geografía, Universidad Autónoma de Madrid, Madrid.
- CATELL, R.B. (1966). «The scree test for the number of factors». *Multivar. Behav. Res.*, I, pp. 245-276.
- CHRISTENSEN, W. I. y BRYSON, R.A. (1966). "An investigation of the potential of component analysis for weather classification". *Mon. Wea. Rev.*, 94(12), 697-709.
- COMELLAS, S. (1964). «Los estados del tiempo en la Cuenca de Pamplona». *Geographica*, X-XI, pp. 3-34.
- CROWE, P.R., (1971). *Concepts in Climatology*, St. Martin's press, New York, 589 pp.
- CREUS, J. (1981). *El clima del Alto Aragón Occidental*. Instituto de Estudios Altoaragoneses, Jaca.
- DAVIS, R. E. (1991). "A synoptic climatological analysis of winter visibility trends in the mideastern United States". *Atmos. Environ.*, 25B, pp. 165-175.
- DAVIS, R. E. y KALKSTEIN, L.S. (1990a). "Using a spatial synoptic climatological classification to assess changes in atmospheric pollution concentrations". *Phys. Geogr.*, 11(4), pp. 320-342.
- DAVIS, R. E., y L. S. KALKSTEIN (1990b) "Development of an automated spatial synoptic climatological classification". *Int. J. Climatol*, 10, pp. 769-794.

- DAVIS, R.E. y ROGERS, R.F. (1992). "A synoptic climatology of severe storms in Virginia". *Profess. Geographer*, 44 (3), pp. 319-332.
- DAVIS, R.E. y WALKER, C.T. (1992). "An upper air synoptic climatology of the western United States". *J. Clim.*, 5, pp. 1449-1467.
- DILLEY, F. B. (1992). "The statistical relationship between weather-type frequencies and corn (maize) yields in southwestern Pennsylvania, USA". *Agr. Forest. Meteorol*, 59, pp 149-164.
- DIEGO, C.; GARCÍA CODRÓN, J.C.; PACHECO, S. y RASILLA, D. (2001). "Le climat et les incendies á Cantabria (Espagne): le feu, un outil intégré du système d'exploitation agrarie". *Climat et environnement. L'information climatique au service de la gestion de l'environnement*, (en prensa).
- DURAND-DASTÉS, F. (1973). "A propos des notions de type de temps et de type de circulation". *Journées de Climatologie de Rennes*.
- EDER, B.K.; DAVIS, J.M. Y BLOOMFIELD, P. (1994). "An automated classification scheme designed to better elucidate the dependence of ozono on meteorology". *J. App. Meteorol*, 33, pp. 1182-1198.
- ELLIS, A.W. y LEATHERS, D. J. (1996). "A synoptic climatological approach to the analysis of lake-effect snowfall: potential forecasting applications". *Wea. Forecast.* 11, pp. 216-229.
- FEDEROV, E.E. (1927). "Da klima as Wettergesamtheit". *Mon. Wea. Rev.*, 55, pp. 401-403.
- FERNANDEZ GARCIA, F. y RASILLA, D. (1999). "Concentraciones de ozono y tipos de tiempo en el Área Metropolitana de Madrid. VIII Coloquio Ibérico de Geografía, vol 1, Lisboa, pp. 31-36.
- FERNANDEZ GARCIA, F.; RASILLA, D.; PLAZA, G. y HORCAS, R. (2001). "Situations meteorologiques et pollution atmospherique a Madrid, Espagne". *Climat et environnement. L'information climatique au service de la gestion de l'environnement*, (en prensa).
- FRAKES, B. y YARNAL, B. (1997). "A procedure for blending manual and correlation-based synoptic classifications". *Int. J. Climatol.*, 17, pp. 1381-1396.
- GALAN GALLEGU, E.(1989): *Tipos de tiempo anticiclónicos invernales en la España Peninsular y Baleares. Ensayo metodológico*. Tesis Doctoral, 3 vol, Madrid, 1405 pp.

- GEENE, J.S y KALKSTEIN, L.S. (1996). "Quantitative analysis of summer air masses in the eastern United States and application to human mortality". *Clim. Res*, 7, pp. 43-53.
- GREENE, J.S. (1996). "A synoptic climatological analysis of summertime precipitation intensity in the Eastern United States". *Phys. Geogr.* 17(5), pp. 401-418.
- HARE, F.K. (1955). "Dynamic and Synoptic Climatology". *Ann. Assoc. Am. Geogr.*, 2-3, pp. 152-162.
- HARMAN, J.R. y WINKLER, J. A. (1991). "Synoptic climatology: themes, applications and prospects". *Phys. Geogr.* 12 (3), pp. 220-230.
- HEWITSON, B. y CRANE, R.G. (1992). "Regional climates in the GISS Global Circulation Model and Synoptic-scale Circulation". *J. Climate*, 5, pp. 1002-1011.
- HUFTY, A. (1976). "Situations synoptiques et analyse multivariée des temps á Qubec". *Cahier de géographie du Qubec*, 20, pp. 69-92.
- HUFTY, A. (1981). "Les types de temps á Argel (Analyse Multivariée)". *La Meteorologie*, 26, pp. 93-109.
- HUTH, R. y NEMESOVA, Y. (1995). "Estimation of missing daily temperatures: can a weather categorization improve its accuracy". *J. Climate*, 8, pp. 1901-1916.
- HUTH, R.; NEMESOVA, Y. y KLIMPEROVA, N. (1993). "Weather categorization based on the average linkage clustering technique: an application to european mid-latitudes". *Int. J. Climatol.* 13, pp. 817-835.
- JACOBS, W.C. (1946). "Synoptic Climatology". *Bull. Am. Met. Soc.*, 27, pp. 306-311.
- JAMES, M. (1985). *Classification Algorithms*. John Wiley and Sons, New York, 210 pp.
- JOLLIFFE, I.T. (1986). *Principal component analysis*. Springer, New York, 271 pp.
- JONES, G.V. y DAVIS, R.E. (2000). "Using a synoptic climatological approach to understand climate-viticulture relationships". *Int. J. Climatol.*, 20, pp. 813-837.
- KALKSTEIN, L. S. (1991). "A new approach to evaluate the impact of climate on human mortality". *Environ. Health Pers.*, 96, pp. 145-150.

- KALKSTEIN, L. S. y CORRIGAN, P. (1986). "A synoptic climatological approach for geographical analysis: Assessment of sulfur dioxide concentrations". *Ann. Ass. American Geograph.*, 76(3), pp. 381-395.
- KALKSTEIN, L. S.; DUNNE, P. y VOSE, D.R. (1990). "Detection of climatic change in the western North American Arctic using a synoptic climatological approach". *J. Climate*, 3(10), pp. 1153-1167.
- KALKSTEIN, L. S.; TAN, G. y SKINDLOV, J. (1987). "An evaluation of objective clustering procedures for use in Synoptic climatological classification". *J. Clim App. Meteorol*, 26(6), pp. 717-730.
- KALKSTEIN, L.S. y VALIMONT, K.M. (1986). "An Evaluation of Summer Discomfort in the United States Using Relative Climatological Index". *Bull. Amer. Meteor Soc*, 67(7), pp. 842-848.
- KALKSTEIN, L.S.; JAMESON, P.F.; GREENE, J.S.; LIBBY, J. y ROBINSON, L. (1996). "The Philadelphia Hot Weather Health Watch/Warning system: development and application, summer 1995". *Bull. Amer. Meteor Soc*, 77(7), pp. 1519-1528.
- KALKSTEIN, L.S.; NICHOLS, M.; BARHTEL, C. y GREENE, J. (1996). "A new spatial synoptic classifications: applications to air-mass analysis". *Int. J. Climatol.*, 16, pp. 983-1004.
- KALKSTEIN, L.S.; SHERIDAN, S.C y GRAYBEAL, D.Y. (1998). "A determination of character and frequency changes in air masses using a spatial synoptic classification". *Int. J. Climatol.*, 18, pp. 1223-1236.
- KUZBACH, J.E. (1967). "Empirical eigenvectors of sea level pressure, surface temperature and precipitation complexes over North America". *J. Appl. Meteorol.*, 6, pp. 792-802.
- LANA, X. y FERNÁNDEZ MILLS, G. (1994). "Minimum sample size for synoptic weather type classification. Application to winter period data recorded on the catalan coast (north-east Spain)". *Int. J. Climatol*, 15, pp. 1051-1060.
- LEATHERS, D. J. Y ELLIS, A.W. (1996). "Synoptic mechanisms associated with snowfall increases to the lee of the lakes Erie and Ontario". *Int. J. Climatol*, 16, pp. 1117-1135.
- LORENZ, E. N. (1967). *Empirical orthogonal functions and statistical weather prediction*. Scientific Report 1, Statistical Forecasting Project, Department of Meteorology, Massachusetts Institute of Technology, 48 pp.

- KIRCHOFFER, W. (1973). "Classification of 500 mb patterns". *Arbeitsbericht der Schweizerischen Zentralanstalt.*, 45, 16 pp.
- LUND, L. (1963). "Map pattern classification by statistical methods". *J. Appl. Meteorology*, 2, pp. 56-65.
- MCCABE, G. J.; HAY, L.E.; KALKSTEIN, L.S.; AYERS, M.A: y D. M. WOLOCK (1989). "Simulation of precipitation by weather-type analysis" *Hydraulic Engineering '89 Proceedings National Conf. on Hydraulic Engineering/HY Div/ASCE*, New Orleans, pp. 679-684.
- MCGREGOR, G. R.; WALTERS, S. y WORDLEY, J. (1999). "Daily hospital respiratory admissions and winter air mass types, Birmingham, U.K.". *Int. J. Biometeorol.*, 43, pp. 21-30.
- MCGREGOR, G.R. (1999). "Winter ischaemic heart disease deaths in Birmingham, U.K.: a synoptic climatological analysis". *Clim. Res*, 13, pp. 17-31.
- MCGREGOR, G.R. y BAMZELIS, D. (1995). "Synoptic typing and its application to the investigation of weather air pollution relationships, Birmingham, U.K.". *Theor. Appl. Climatol*, 51, pp. 223-236.
- MILLIGAN, G.W. y COOPER, M.C. (1985). "An examination of the effect of six types of error perturbation on fifteen clustering algorithms". *Psychometrica*, 45, pp. 325-342.
- OLCINA CANTOS, J. (1994). "Métodos de clasificación sinóptica en España. Revisión y propuesta". *Estudios Geográficos*, 215, pp. 357-387.
- OLIVER, J.E. (1991). "The history, status and futures of climatic classification". *Phys. Geogr*, 12 (3), pp. 231-251.
- OVERLAND, J.E. y PREISENDORFER, R.W. (1982). "A significance test for principal components applied to cyclone climatology". *Mon. Wea. Rev*, pp. 110, 1-8.
- ÖZELKAN, E. O.; NI, F. Y DUCKSTEIN, L. (1996). "Relationships between monthly atmospheric circulation patterns and precipitation: Fuzzy logic and regression approaches". *Water Res. Research*, pp. 2097-2103.
- PEJENAUTE GOÑI, J.M. (1996). *Tipos de tiempo y clima de las comarcas navarras*. Fondo de Publicaciones del Gobierno de Navarra, Serie Geográfica, nº7, 545 pp.
- PREISENDORFER, W.C. (1988). *Principal component analysis in Meteorology and Oceanography*, Elsevier, Amsterdam, 425 pp.

- RASILLA, D. (1996). *Viento Sur y efecto föhn en la vertiente septentrional de la Cordillera Cantábrica*. Universidad Autónoma de Madrid, Madrid.
- RICHMAN, M.B. (1985). "Rotation of principal components". *J. Climatol*, 6, pp. 293-335.
- ROMESBERG, H. C. (1984). *Cluster Analysis for Researchers*. Lifetime Learning Publications, Belmont, California, 334 pp.
- RUIZ URRESTARAZU, E.: *La transición climática del Cantábrico Oriental al Valle Medio del Ebro*. Diputación Foral de Alava, Servicio de Publicaciones, Vitoria, 1982, 651 pp.
- SÁNCHEZ, M.L.; PASCUAL, D.; RAMOS, C. y PÉREZ, Y. (1990). "Forecasting particulate pollutant concentrations in a city from meteorological variables and regional weather patterns". *Atm. Environm*, 6, pp. 1509-1519.
- SERRA, C; FERNÁNDEZ MILLS, G.; PERIAGO, M.C. y LANA, X. (1999). "Winter synoptic weather types in Catalonia (NE Spain) and their linkage with minimum temperature anomalies". *Int. J. Climatol.*, 19, pp. 1675-1695.
- SHAHGEDANOVA, M.; BURT, T.P. y DAVIES, T.D. (1998). "Synoptic climatology of air pollution in Moscow". *Theor. Appl. Climatol*, 61, pp. 85-102.
- SMOYER, K.E.; KALKSTEIN, L.S.; GREENE, J.C. y YE, H. (2000). "The impacts of wather and pollution on human mortality in Birmingham, Alabama, and Philadelphia, Pennsylvania". *Int. J. Climatol.*, 20, pp. 881-897.
- STONE, R.G. (1989). "Weather types at Brisbane, Queensland: an example of the use of principal components and cluster analysis". *Int. J. Climatol*, 9, pp. 3-32.
- TUDURI, E. Y RAMIS, C. (1997). "The environments of significative convective events in the Western Mediterranean". *Wea. and Forecast.*, 12, pp. 294-306.
- YARNAL, B. (1993). *Synoptic climatology in Environmental Analysis*. Studies in Climatology Series, Belhaven Press, London, 195 pp.
- VON STORCH, H. y ZWIERS, F. W. (1999). *Statistical Analysis in Climate Research*, Cambridge University Press, Londres, 484 pp.
- WILKS, S. *Statistical Methods in the Atmospheric Sciencies*, Academic Press, vol. 59, Nueva York, 461 pp.
- PERRY, A. (1991). "Growing points in synoptic climatology". *Progr. Phys. Geogr*, 17, pp. 91-96.