

# INSTITUTO NACIONAL DE METEOROLOGIA

SUB. GRAL. DE CLIMATOLOGIA Y APLICACIONES

## CLASIFICACION DE PATRONES ESPACIALES DE PRECIPITACION DIARIA SOBRE LA ESPAÑA PENINSULAR Y BALEARICA

por

Jaime Ribalaygua y Rafael Borén\*

REGISTRADO

\* Centro de Investigación Forestal (CIFOR), Instituto Nacional de Investigación y Tecnología Agraria y Agroalimentaria (INIA)

SERVICIO DE ANALISIS E  
INVESTIGACION DEL CLIMA

INFORME Nº 3

Septiembre 1995

AEMET-BIBLIOTECA



1003547



R<sup>o</sup>. - ~~14.106~~  
CB 1003547

Sig.: M83: M77(460)

# INSTITUTO NACIONAL DE METEOROLOGIA

SUB. GRAL. DE CLIMATOLOGIA Y APLICACIONES

## CLASIFICACION DE PATRONES ESPACIALES DE PRECIPITACION DIARIA SOBRE LA ESPAÑA PENINSULAR Y BALEARICA

Septiembre 1995

06 OCT. 1995

**Jaime Ribalaygua Batalla**

**Rafael Borén Iglesias\***

\* Centro de Investigación Forestal (CIFOR), Instituto Nacional de Investigación y Tecnología Agraria y Agroalimentaria (INIA).  
Carretera de la Coruña, Km 7  
28040 MADRID



# ÍNDICE

<b>1. INTRODUCCIÓN</b> .....	3
<b>2. ANTECEDENTES</b> .....	4
<b>3. DESCRIPCIÓN Y PREPARACIÓN DE LOS DATOS</b> .....	5
3.1. Datos Atmosféricos.....	5
3.B. Datos de Superficie.....	5
<b>4. METODOLOGÍA DE CLASIFICACIÓN</b> .....	7
4.1. Transformaciones del campo de datos.....	7
4.2. Medidas de disimilitud.....	9
4.3. Métodos de agregación.....	10
4.4. Selección del método de agregación.....	11
4.5. Método de "clustering" utilizado: Método de Petisco Modificado.....	11
4.5.1. Selección de semillas.....	12
4.5.2. Asignación de días a las semillas.....	14
4.6. Selección de las variables y la medida de disimilitud.....	14
4.7. Proceso final de clasificación.....	15
<b>5. RESULTADOS</b> .....	16
5.1. Tipos de reparto de precipitación.....	16
5.2. Aplicación de la clasificación al período 1941-1990.....	19
<b>6. AGRADECIMIENTOS</b> .....	20
<b>7. BIBLIOGRAFÍA</b> .....	21
<b>ANEXO: FIGURAS</b> .....	23



# 1. INTRODUCCIÓN

El presente estudio se enmarca en el contexto de una serie de trabajos que el Servicio de Análisis e Investigación del Clima (SAIC) del Instituto Nacional de Meteorología desarrolla, con el objetivo común de incrementar el conocimiento de las relaciones entre las configuraciones atmosféricas a escala sinóptica del área peninsular, y los efectos en superficie (principalmente precipitación y temperatura) a escala local o comarcal.

Aunque la resolución temporal de los análisis es diaria, la predictividad de las relaciones detectadas se contrastará siempre a nivel climatológico (compendios de centenas o millares de días). La resolución diaria viene determinada por la escala temporal propia de los fenómenos analizados (las perturbaciones sinópticas tienen una vida media del orden de días), y por la imposibilidad de contar con un banco de datos de fenómenos en superficie de resolución superior a la diaria.

El estudio de los factores de superficie se realiza a escala local o comarcal, ya que es en esta escala donde una fluctuación o cambio climático es percibida y sufrida por el hombre y los ecosistemas, y donde es planteable una respuesta técnica o de planificación territorial.

Para el análisis de las configuraciones atmosféricas, la escala es sinóptica. Hubiera sido deseable una mayor resolución, pues las estructuras mesoescalares pueden tener una influencia importante en determinadas épocas y zonas de nuestra geografía. Sin embargo, tanto los datos del pasado como las proyecciones del futuro climático ofrecidas por los Modelos de Circulación General (M.C.G.s), tienen y tendrán a medio plazo unas resoluciones para las que las formaciones de mesoescala son cuasi "sub-grid".

Al ser distintas las escalas espaciales de trabajo en atmósfera y superficie, las relaciones entre ellas deben necesariamente determinarse de forma estadística.

El conocimiento de dichas relaciones debe permitir alcanzar dos objetivos ulteriores:

- a) Estimar, para un determinado período, los campos climatológicos, en función de las configuraciones atmosféricas diarias del mismo período. Si esta estima es suficientemente precisa, se podrán obtener escenarios climáticos futuros a escala comarcal, realizando lo que se denomina un "downscaling" estadístico. Esta técnica tiene como objetivo soslayar las limitaciones de los M.C.G.s, derivadas de su baja resolución espacial. Puede considerarse la parte "predictiva" de este campo de investigación del SAIC.
- b) Explicar fluctuaciones del pasado en dichos campos climatológicos, en función de cambios en la frecuencia de aparición de determinados tipos sinópticos. Los trabajos que abordan este objetivo constituyen la aplicación "analítica" de las investigaciones.

El área de predicción es objeto de otros estudios llevados a cabo por este Servicio (Borén *et al.*, 1995). El presente estudio describe, enmarcado en la búsqueda de las relaciones atmósfera-superficie antes comentadas, la generación de una clasificación de efectos en superficie. Obtenida esta clasificación, se estudian brevemente las relaciones entre cada tipo de efectos en superficie y las configuraciones sinópticas que lo generan, avanzando en el campo analítico de las investigaciones del SAIC.

## 2. ANTECEDENTES

Para abordar el estudio de las interacciones entre configuraciones atmosféricas y efectos en superficie, cabe sondear estadísticamente las relaciones de una forma directa mediante la utilización de técnicas como: correlación canónica (Von Storch *et al.*, 1993), predicción analógica (Zorita *et al.*, 1993), regresión estadística entre cierta variable de superficie en un punto, y otras de altura (Wigley *et al.*, 1990), etc. Sin embargo en la mayoría de estos trabajos de hace difícil, cuando no imposible, encontrar una explicación conceptual a las relaciones detectadas, de forma que pueden conseguir buenos resultados en predicciones, pero no permiten analizar qué cambios en la circulación atmosférica son los que generan los de las variables de superficie. Es por ello que, sin desestimar una utilización posterior de algunos de dichos métodos, se ha preferido comenzar con una aproximación de mayor potencial analítico.

Cualquier metodología analítica implica una partición del material. En este caso, el material de estudio lo constituyen conjuntos o poblaciones de configuraciones atmosféricas, y de campos de efectos en superficie. Una partición en tipos de dichas poblaciones lleva a generar una clasificación.

La clasificación de las configuraciones atmosféricas es lo que tradicionalmente se ha denominado clasificación sinóptica. Las clasificaciones sinópticas existentes para España son mayoritariamente "subjetivas" (tipos elegidos por la experiencia de un meteorólogo, y días asignados a cada tipo por el propio investigador). Entre estos trabajos destacan Font (1983), Linés (1981), Sánchez Rodríguez (1993)... No existen publicaciones de períodos largos clasificados siguiendo alguno de estos esquemas, como ocurre, por ejemplo, en el Reino Unido (Lamb, 1972). Aunque son aportaciones de indudable valor, al no estar automatizadas, el análisis de períodos pasados, o de futuros generados por M.C.G.s, requeriría la inspección de las configuraciones de cada día por parte del experto. En el actual contexto de trabajo, es necesaria la utilización de clasificaciones "automatizables", es decir, que la asignación de cada día a los diferentes tipos sea realizable por un ordenador.

Recientemente se han concluido dos estudios que desarrollan clasificaciones automatizables (Calvo, 1993), (Petisco y Martín, 1995). La determinación de los tipos sinópticos se realiza mediante técnicas de "clustering", utilizando algoritmos de similitud entre campos de presión en superficie o geopotenciales. Una vez determinadas las configuraciones tipo, los días se asignan a los tipos que les son más parecidos. Un problema característico de esta aproximación clásica radica en la propia naturaleza de los campos a clasificar. La ausencia de fronteras espaciales nítidas para determinar las ubicaciones preferentes de las perturbaciones sinópticas (no hay "clusters" naturales nítidos), genera una gran sensibilidad de los resultados a pequeñas variaciones en los parámetros que rigen el proceso de agrupamiento. Esto hace que los resultados del proceso dependan mucho de la experiencia y la intuición meteorológica del investigador.

Para intentar disminuir parcialmente este problema, y para buscar la significación meteorológico-climatológica de los tipos (diferentes tipos, diferentes efectos en superficie), se ha abordado el problema desde una perspectiva novedosa. El proceso de cluster se aplicará, no a las configuraciones sinópticas directamente, sino a los campos de efectos en superficie. Dichos campos, debido a la existencia de discontinuidades espaciales claras (cordilleras, cuencas geográficas, etc.), pueden presentar más fácilmente "clusters" naturales. Los distintos tipos sinópticos se determinarán, en primera aproximación, como los campos atmosféricos medios de los días pertenecientes a cada reparto-tipo de la/s variable/s de superficie.

### 3. DESCRIPCIÓN Y PREPARACIÓN DE LOS DATOS

#### 3.1. DATOS ATMOSFÉRICOS

Los datos atmosféricos provienen del National Meteorological Center (NMC) Northern Hemisphere Grid Data Set. Una parte de este banco de datos es accesible en CD-ROM (Mass *et al.*, 1987). Se ha utilizado la segunda versión de este CD-ROM, generada por los autores antes citados.

Son análisis finales del NMC. La red es octogonal, con 1977 puntos igualmente espaciados en proyección polar estereográfica, centrada en el Polo Norte. Para nuestras latitudes, esta resolución equivale aproximadamente a 3 o 4 grados de espaciamiento internudos. El NMC proporciona también una subrutina de interpolación (Esquema Bessel) para convertir dicha red en una de latitud/longitud. La red finalmente utilizada en el estudio puede contemplarse en la figura 1. El paso es de 3° en latitud y de 4° en longitud, y los 120 puntos que la componen se extienden por un área centrada en el norte de Iberia, que se considera permite captar bien las perturbaciones que afectan nuestro territorio. Las variables disponibles son:

- Alturas de geopotencial de 850mb, 700mb, 500mb, 200mb, que desde ahora se nombrarán Z850, Z700, Z500 y Z200 respectivamente.
- Presión reducida a nivel del mar (SLP).
- Viento en 850 y 250mb.
- Temperatura en 850 y 700mb (T850, T700).

La mayoría de las variables tienen dato casi continuo cada 12 horas desde 1962 a 1989. Z500 y SLP a 12hUTC tienen además dato en el período 1945-1989.

#### 3.2. DATOS DE SUPERFICIE

Las variables de superficie con dato diario y densidad adecuada al estudio (más de 150-200 observatorios homogéneamente repartidos por la Península y Baleares), sólo son tres: temperaturas máxima y mínima, y precipitación acumulada.

En esta primera aproximación se ha decidido trabajar con la precipitación. Al partir de esta variable, la clasificación sinóptica finalmente obtenida tenderá a discriminar bien los repartos espaciales de la lluvia, y a desatender los de temperatura. Se considera sin embargo más conveniente tratar de buscar en lo posible una buena determinación de los campos de precipitación, ya que es una variable inherentemente más complicada de predecir que la temperatura.

Se ha dispuesto de todo el banco de datos de precipitación acumulada diaria del Instituto entre los años 1900 y 1994. Lógicamente, la densidad de observatorios con dato no es igual en todo el período. Antes del año 1941, la frecuente existencia de días sin dato en regiones extensas puede dificultar una correcta caracterización de los repartos. A partir de Junio de 1991, faltan los datos de Cataluña y de gran parte de los observatorios de la zona centro. La clasificación de repartos de precipitación se ha generado con los datos de 1961-1990. En la figura 2 se representa por un punto cada uno de los 7138 observatorios con algún dato en este período. La clasificación final se aplicará al período 1941-1990.

Para permitir la comparación entre repartos diarios de precipitación, se requiere un número fijo de datos homogéneamente repartidos por la superficie nacional. La existencia de lagunas (falta de dato) en casi todos los observatorios puede subsanarse de una de estas formas:

- Eligiendo un número determinado de observatorios bien repartidos territorialmente, y rellenando las lagunas que presenten mediante regresión múltiple con los datos síncronos existentes en estaciones del entorno.
- Creando un banco de datos en rejilla. Los valores de cada punto del grid se obtienen por interpolación espacial con los datos de su entorno.

Se eligió finalmente la segunda opción, debido a las mayores facilidades que ofrece para extender el análisis a períodos diferentes, ya que no depende de la existencia de datos de observatorios concretos en los nuevos períodos. Además permite utilizar más información (todos los datos disponibles), disminuyendo simultáneamente el ruido producido por fenómenos locales (de naturaleza microclimática o derivados de incorrectas mediciones).

La rejilla de trabajo (figura 3), con 203 puntos, permite una aproximación al fenómeno a escala comarcal (un punto cada 2500km<sup>2</sup>).

En la estima de los valores diarios del campo en los puntos de rejilla se utilizó la siguiente fórmula:

$$P_j = \frac{\sum_{i=1}^n \frac{p_i}{d_{ij}}}{\sum_{i=1}^n \frac{1}{d_{ij}}}$$

donde  $P_j$  es la precipitación estimada en el punto  $j$ ,  $p_i$  es la precipitación en el observatorio correspondiente, y  $d_{ij}$  es la distancia entre ese observatorio y el punto  $j$  (se pondera por la inversa de la distancia). El parámetro  $n$ , número de observatorios con los que estimar, es diferente para cada punto, e incluso para cada día. Para estimar  $P_j$  se utilizan todos los observatorios con dato del día en cuestión en un radio de 57km entorno al punto. Este límite se seleccionó subjetivamente mediante pruebas con distintos radios. Valores superiores tienden a comprometer la precisión comarcal buscada al producir un excesivo suavizamiento de las formas; radios inferiores generan campos a nuestro juicio demasiado ruidosos. Sólo cuando hay menos de tres datos en ese radio, se considera que la media no sería representativa, y se buscan los datos que faltan para completar esos tres fuera del entorno mencionado, por orden de proximidad al punto.

El algoritmo así descrito funciona como un interpolador exacto (si hay un observatorio  $i$  coincidente con el punto de rejilla  $j$ ,  $P_j = p_i$ ). Para evitar esto, y que por ello el punto pueda presentar el comportamiento ruidoso de una precipitación local, se asigna una distancia teórica de 10Km, a todos los observatorios situados a una distancia del punto inferior a dicho valor. Todos estos valores numéricos se seleccionaron tras las pruebas pertinentes.

Se desestimó el procedimiento de cálculo Krigging por su deficiente funcionamiento sobre estos datos (creaba tendencias exageradas) y por el mayor coste de computación.

Para dar una idea de la precisión de estos campos en puntos de rejilla, se ofrecen las figuras 4 y 5. La figura 4 representa la precipitación media anual del período 1961-1990, estimada como compendio de los 10957 días del mismo, en puntos de rejilla. La figura 5 muestra la precipitación media anual del mismo período, estimada en el Atlas Nacional del Clima (MOPT y IGN, 1992).

## 4. METODOLOGÍA DE CLASIFICACIÓN

Toda técnica de clasificación precisa la definición de una medida de disimilitud y de una estrategia de agregación entre individuos y/o grupos (Júdez, 1989). Es también posible "jugar" con transformaciones del campo de datos que se va a clasificar. La selección de la metodología de clasificación, se realizó entre las posibles combinaciones que surgen al utilizar determinadas opciones de estos tres factores.

### 4.1. TRANSFORMACIONES DEL CAMPO DE DATOS

El campo de precipitación diario puede representarse por un vector cuyas componentes sean:

- a) **Los valores de precipitación en los 203 puntos de rejilla.**
- b) **Estos mismos valores pero transformados.** La fuerte asimetría que caracteriza a la distribución frecuencial de la variable precipitación acumulada diaria, claramente no normal, con muchos valores nulos y otros pocos muy extremos, puede hacer su comportamiento inestable para las medidas de disimilitud. La literatura de análisis climático sugiere diversas transformaciones que buscan reducir la asimetría positiva de la variable, "retrayendo" la cola derecha de la distribución: raíces cuadradas (Richman and Lamb, 1985); raíces cúbicas (Douglas and Englehart, 1983); logaritmos neperianos (Willmott, 1977)... Debido a la alta sensibilidad del coeficiente de correlación de Pearson frente a las asimetrías en la distribución, y teniendo en cuenta que participa en alguna de las medidas de disimilitud ensayadas, se analizaron las correlaciones entre las series de los observatorios más completos dos a dos, utilizando los valores sin transformar, y transformados de diversas formas. La transformación más eficiente resultó ser la de raíz cúbica, y por ello fue esta la utilizada finalmente.
- c) **Los coeficientes de un número determinado de componentes principales (rotadas o no) de los valores (transformados o no) en puntos de rejilla.** Algunos autores (Lanzante and Harnack, 1982; Von Storch *et al.*, 1993; Richman and Lamb, 1985) proponen la utilización de transformaciones lineales por autovalores para reducir el número de variables y separar la señal meteorológica del ruido. Entre los modelos de transformación lineal, los denominados componentes principales (CPs) son más convenientes que las "Empirical Orthogonal Functions" (EOFs) para garantizar la significación física de los autovalores (Richman and Lamb, 1985).

En la determinación de las CPs, se optó por utilizar la matriz de correlación como matriz de dispersión, para evitar dar más peso a los puntos de rejilla de mayor precipitación.

El número de CPs a retener se determinó siguiendo diversos criterios: la regla N de Preisendorfer y la regla Q (Preisendorfer, 1988), el criterio de Anderson, el de retener los CPs que expliquen más varianza que una variable original (en este caso, que expliquen más del  $[100/203]\%$ )... Los criterios no coincidían en un determinado número de CPs a retener, por lo que hubo de llegarse a un compromiso entre ellos. Al final se optó por

retener 22 CPs del campo de precipitaciones sin transformar y 15 del de raíz cúbica de precipitación, que explicaban el 87% y el 88% de la varianza de sus datos respectivos.

Determinado el número de CPs a retener, además de probar las CPs sin rotar (CPSs), se estudiaron dos tipos de transformaciones ulteriores de las mismas, en un ensayo similar al propuesto por White *et al.* (1991): CPs rotadas ortogonalmente por el método VARIMAX (CPRs), y rotadas oblicuamente por el método D-QUART (CPOs). El motivo de estas pruebas es que, según los citados autores, estas transformaciones pueden mejorar la significación física de las CPs.

En efecto: Se denominan "factor loadings" de la componente principal  $i$  a los  $c_{i,1}, c_{i,2}, \dots, c_{i,j}, \dots, c_{i,203}$  de

$$\begin{aligned} C_1 &= c_{1,1}P_1 + c_{1,2}P_2 + \dots + c_{1,j}P_j + \dots + c_{1,203}P_{203} \\ C_2 &= c_{2,1}P_1 + c_{2,2}P_2 + \dots + c_{2,j}P_j + \dots + c_{2,203}P_{203} \\ &\dots\dots\dots \\ C_i &= c_{i,1}P_1 + c_{i,2}P_2 + \dots + c_{i,j}P_j + \dots + c_{i,203}P_{203} \\ &\dots\dots\dots \\ C_m &= c_{m,1}P_1 + c_{m,2}P_2 + \dots + c_{m,j}P_j + \dots + c_{m,203}P_{203} \end{aligned}$$

donde  $P_1, P_2, \dots, P_j, \dots, P_{203}$ , son los valores de la precipitación (transformados o no, según el caso) de cierto día en los 203 puntos de rejilla, y  $C_1, C_2, \dots, C_i, \dots, C_m$ , son los coeficientes de ese mismo día para las  $m$  componentes principales. Si se representan los factor loadings de cada componente, asignando a cada punto de rejilla el factor correspondiente, se obtienen mapas como los de la figura 6, que corresponden a las CPOs, denominados "factor loadings patterns". En el caso de CPSs, aparecen los denominados patrones de Buell (Buell, 1979), donde la "topografía" de los CPs parece ser función de la forma geométrica del dominio espacial representado. Esto indica que las CPSs responden, más que a las relaciones meteorológicas entre la precipitación en los diferentes puntos, a la configuración geométrica del territorio (en este caso, aproximada a un triángulo invertido). Además, las diferencias de explicación de la varianza entre ellas son muy grandes, lo que daría enormes diferencias en la utilización de medidas de disimilitud ponderadas o sin ponderar (ver apartado 4.2.), y excesivo peso a las primeras CPSs si se ponderan.

Las CPRs captan mucho mejor las relaciones entre los puntos, y sus factor loadings generan regionalizaciones con claro sentido meteorológico. Además, las diferencias de varianza explicada son mucho menores, equilibrándose así el peso que se asigna a cada CPR. Sin embargo, la exigencia de que las nuevas variables sean ortogonales entre sí introduce ciertas limitaciones.

Estas limitaciones desaparecen en la rotación oblicua, reduciéndose el solapamiento de los patrones de los CPs y aumentando la exclusividad de la regionalización producida (cada punto de rejilla tiende a estar relacionado con una sola CP) (ver figura 6). Así mismo, las diferencias de varianza explicada son, con mucho, las menores.

Los resultados obtenidos para nuestro banco de datos coinciden con las anteriores consideraciones teóricas, y, a la vez, con los obtenidos por White *et al.* (1991) para precipitaciones mensuales en el estado de Pensilvania, EEUU. Se pretende estudiar si esta mejora en la regionalización de las CPs al rotarlos, redundan en un mejor funcionamiento de la clasificación.

En definitiva, las variables ensayadas para representar los campos diarios de precipitación han sido:

- 1) Valores de precipitación en 203 puntos de rejilla, sin transformar:  $P_1, P_2, \dots, P_{203}$ .
- 2) Valores de raíz cúbica de precipitación en los mismos puntos:  $^3P_1, ^3P_2, \dots, ^3P_{203}$ .
- 3) Coeficientes de 22 CPSs del campo de precipitación en 203 puntos de rejilla sin transformar:  $Css_1, Ccss_2, \dots, Ccss_{22}$
- 4) Coeficientes de 22 CPRs del campo de precipitación en 203 puntos de rejilla sin transformar:  $Crs_1, Crs_2, \dots, Crs_{22}$
- 5) Coeficientes de 22 CPOs del campo de precipitación en 203 puntos de rejilla sin transformar:  $Cos_1, Cos_2, \dots, Cos_{22}$
- 6) Coeficientes de 15 CPSs del campo de raíz cúbica de precipitación en 203 puntos de rejilla:  $Cst_1, Cst_2, \dots, Cst_{15}$
- 7) Coeficientes de 15 CPRs del campo de raíz cúbica de precipitación en 203 puntos de rejilla:  $Crt_1, Crt_2, \dots, Crt_{15}$
- 8) Coeficientes de 15 CPOs del campo de raíz cúbica de precipitación en 203 puntos de rejilla:  $Cot_1, Cot_2, \dots, Cot_{15}$

#### 4.2. MEDIDAS DE DISIMILITUD

Se han analizado las siguientes:

- a) **Distancia euclídea**, si se aplica sobre variables ortogonales (CPSs y CPRs), o **pseudodistancia**, para variables no ortogonales (resto de campos), según:

$$D_{xy} = \sum_{i=1}^n (x_i - y_i)^2$$

donde  $\mathbf{X} = (x_1, x_2, \dots, x_n)$ ,  $\mathbf{Y} = (y_1, y_2, \dots, y_n)$ , son los campos de precipitación de los días X e Y, representados por las n (203, 22 o 15) variables correspondientes, según el caso.

Para los campos de componentes principales, además, se utilizó una distancia ponderando cada componente principal por la varianza explicada por ella, según:

$$D_{xy} = \sum_{i=1}^n w_i (x_i - y_i)^2$$

donde  $w_i$  es la varianza explicada por la componente principal i.

- b) **Índice de correlación (uno menos el coeficiente de correlación de Pearson) entre los campos**, según:

$$D_{xy} = 1 - \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{Y})^2}}$$

$$\text{donde } X = \frac{\sum_{i=1}^n x_i}{n} \quad Y = \frac{\sum_{i=1}^n y_i}{n}$$

Del mismo modo que antes, para los campos de componentes principales se utilizó también un coeficiente de correlación ponderado por la varianza explicada por cada componente, usando

$$D_{xy} = 1 - \frac{\sum_{i=1}^n w_i(x_i - X)(y_i - Y)}{\sqrt{\sum_{i=1}^n w_i(x_i - X)^2 \sum_{i=1}^n w_i(y_i - Y)^2}}$$

$$\text{donde } X = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad Y = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i}$$

Por tanto, las medidas de disimilitud ensayadas han sido:

- 1) Distancia euclídea o pseudodistancia, sin ponderar (para todos los campos).
- 2) Distancia euclídea o pseudodistancia, ponderada por la varianza explicada por cada CP(para CPs).
- 3) Índice de correlación (uno menos coeficiente de correlación de Pearson), sin ponderar (para todos los campos).
- 4) Índice de correlación (uno menos coeficiente de correlación de Pearson), ponderada por la varianza explicada por cada CP(para CPs).

### 4.3. MÉTODOS DE AGREGACIÓN

Sobre diferentes combinaciones de campos-medidas de disimilitud, se ensayan los métodos de agregación disponibles. Se hicieron intentos de clasificación con los siguientes métodos:

- Average method (SAS, 1990)
- Centroid method (SAS, 1990)
- Complete linkage method (SAS, 1990)
- Density linkage method:
  - K-nearest neighbour method(SAS, 1990)
  - Uniform-Kernel's method(SAS, 1990)
  - Wong's Hybrid method (SAS, 1990)
- Flexible-beta method (SAS, 1990)
- Single linkage method (SAS, 1990)
- Ward's method (SAS, 1990), (SPAD, 1993)
- MacQuitty's similarity analisis (SAS, 1990)
- Median method (SAS, 1990)

- Método de Petisco (Petisco y Martín, 1995).

El método de Petisco se utilizó modificando ligeramente los programas informáticos por él suministrados. Todos los demás se utilizaron con los paquetes estadísticos SAS, SPAD y BMDP.

#### **4.4. SELECCIÓN DEL MÉTODO DE AGREGACIÓN**

Se realizaron diversas pruebas, con varias combinaciones de campos-medidas, y, cuando el método ofrecía la posibilidad de modificar parámetros, con diferentes valores de éstos. En términos generales, se obtuvieron los siguientes resultados:

- Los métodos single y de densidad generaron "bolas de nieve": tendían a asignar la mayoría de los días a un hipergrupo, y un sólo día al resto de grupos.
- El resto de los utilizados con SAS, tendían a detectar algunos clusters naturales nítidos (ver apartado 4.5.1.), pero en general, al no permitir la reasignación de los días (tras una primera fase de clasificación, se calculan los centroides de cada grupo formado, como media de los elementos (días) que lo constituyen, y se reasignan de nuevo todos los elementos por similitud a esos centroides) producen "clusters" claramente no "hiperesféricos", con días dentro de un mismo grupo muy diferentes entre sí (extremos muy "alejados" o disímiles).
- Los resultados más satisfactorios fueron los del programa SPAD, que utiliza el método de Ward, seguramente porque permite la reasignación de elementos. Sin embargo, este método presenta como deficiencia la rigidez de la versión utilizada, que sólo permite usar como medida de disimilitud la distancia euclídea sobre componentes principales rotadas ortogonalmente.
- Sin duda un estudio más exhaustivo hubiera permitido mejores resultados, combinando métodos en diferentes fases del proceso. Sin embargo, tras las primeras pruebas del mismo, se seleccionó el método de Petisco, un método de "clustering" de naturaleza diferente, que desde el principio parecía no sufrir ninguna de las limitaciones antes comentadas. Se pasa ahora a comentar con detalle este método y las modificaciones sobre él realizadas.

#### **4.5. MÉTODO DE "CLUSTERING" UTILIZADO: MÉTODO DE PETISCO MODIFICADO**

Consta de dos fases bien diferenciadas. En la primera fase, se trata de detectar conjuntos de elementos (días) que presenten gran similitud entre sí. El centroide de cada conjunto, que en realidad es un día imaginario, media de los días del conjunto, constituye lo que denominamos una semilla, que representará a un determinado tipo de reparto de precipitación. Establecido un número adecuado de semillas, la segunda fase consiste en asignar cada día del período a clasificar a la semilla con la que presente mayor similitud, quedando el día inclasificado si no supera un umbral de similitud con ninguna.

#### 4.5.1. Selección de semillas

En esta fase se ha utilizado un programa informático basado en el desarrollado por D. Eduardo Petisco, del Servicio de Análisis e Investigación del Clima del Instituto Nacional de Meteorología. Dicho programa detecta las zonas de "alta densidad" del campo de datos, es decir, los días imaginarios que superan cierto umbral de similitud con un mayor número de elementos del banco de datos (días reales). Su funcionamiento, a grandes rasgos, es el siguiente:

- 1) Construye, para cada día, un grupo, siguiendo el proceso:
  - a) asigna a ese día todos aquellos otros cuya similitud con él supere un cierto umbral UMB1,
  - b) calcula el centroide del grupo así formado,
  - c) reasigna al centroide todos los días cuya similitud con él supere UMB1,
  - d) repite b) y c) hasta que el grupo formado sea estable.

La reasignación consigue que los centroides se vayan desplazando hacia las zonas de "alta densidad", hasta detectarse el centroide de máxima "densidad" relativa. Esto hace que muchos días tengan asociado un mismo grupo, pues sus centroides se han ido desplazando hacia un mismo máximo relativo. Y este desplazamiento de centroides permite que muchas veces un día no pertenezca al grupo que tiene asignado, pues su centroide se ha desplazado, en busca de un máximo relativo, llegando tener una similitud con el día original menor de UMB1. Esta idea de la reasignación, y el consiguiente desplazamiento de centroides, es lo que hace de este método una herramienta inestimable para detectar tipos de reparto de precipitación.

- 2) Definidos los grupos asociados a cada día, ahora debe seleccionarse un número de grupos adecuado, que constituyan los tipos de reparto de precipitación. En primer lugar se ordenan todos los grupos por número de días asignados, es decir, por número de días con similitud al centroide mayor que UMB1. En caso de igualdad de número de días asignados, se ordena por la media de similitud entre el centroide y los días que constituyen el grupo.

El primer grupo seleccionado es el que encabeza esa ordenación. Posteriormente se va descendiendo en la misma, y se van seleccionando grupos siempre que sus centroides no sean excesivamente similares a los ya seleccionados, es decir, si la similitud del centroide del grupo estudiado con cada uno de los ya seleccionados es menor que otro umbral, UMB2. Cuando los grupos están compuestos por menos de cierto número de días, se considera que ya no tienen representatividad suficiente y se detiene el proceso.

La repetición de este proceso utilizando diferentes valores de UMB1, genera diferentes conjuntos de posibles semillas. Y entre esas semillas habrá que seleccionar las definitivas.

Las pruebas iniciales mostraron que existen algunos clusters naturales bien definidos en la población (por ejemplo los representados por los tipos 1, 2, 4, 13..., ver figuras 7 a 24). Son repartos que presentan fronteras muy claramente definidas, con muchos días muy similares entre sí. Sin embargo, otros tipos de reparto presentan mucha variabilidad, de forma que los días que a ellos se asignan presentan similitudes más bajas. En unos casos (5, 8, 17, 18...), se debe a que no presentan fronteras espaciales definidas (cuencas, cordilleras...), posiblemente derivado de su carácter convectivo. En otros casos (14, 15...), aunque las fronteras aparecen más claras, la extensión de la precipitación les confiere una gran complejidad morfológica, que disminuye las posibilidades de encontrar días con bajas medidas de disimilitud con ellos. Por eso, algunas semillas aparecerán con valores altos de UMB1, pero otras sólo lo harán reduciendo dicho umbral.

El método original de Petisco fija los valores de UMB1 y UMB2, tras probar diferentes combinaciones, generando un conjunto de semillas prácticamente definitivo, y evitando así la posterior selección subjetiva. La diferencia fundamental entre este método y el aquí empleado, estriba en que aquí se usa un (único) valor muy alto de UMB2, para que aparezca un gran número de puntos de alta "densidad", de modo que sólo se dejan de seleccionar los grupos casi iguales a alguno ya seleccionado. Además, se utilizan simultáneamente varios valores de UMB1, por el explicado motivo de existir diferencias de similitud interna entre los diferentes "clusters" naturales.

Como inconveniente teórico y práctico, este método modificado presenta la necesidad de hacer una selección de semillas, lo que incrementa la subjetividad del método (aunque sólo en esta fase de selección de tipos, pues la asignación, es decir, la aplicación de la clasificación a períodos cualesquiera, es totalmente automática). Pero en cambio, presenta un abanico enormemente más amplio de posibles semillas, incluyendo algunas que, por su naturaleza, sólo aparecen con valores muy bajos de UMB1, y otras que sólo aparecen con valores muy altos (pues al bajarlos, se funden con otros grupos, formando uno solo a partir de varios). El motivo de la modificación introducida es el de evitar el compromiso de fijar unos valores de UMB1 y UMB2 que pudieran favorecer la aparición de uno u otro de los tipos de semillas recién referidos.

Además de estas, hay otras diferencias de menor importancia entre ambos métodos, que no requieren comentario.

Se ha optado por utilizar valores de UMB1 de 0.95, (es decir, disimilitudes de  $D=1-0.95=0.05$ ), 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, y 0.55, y un único  $UMB2=0.95$ , de manera que sólo se rechaza un grupo si es prácticamente igual a alguno ya seleccionado. Y el número de días mínimo se fijó en 20 (sobre los 10957 del período 1961-1990).

Con estos valores, se obtuvieron 9 grupos de semillas, cada uno con el siguiente número de ellas, por orden decreciente del UMB1: 18, 40, 76, 92, 68, 38, 24, 16 y 9 semillas. La selección de las semillas definitivas entre todas estas es subjetiva, tras su representación gráfica y análisis visual. Los criterios de selección son los siguientes:

- Se deben eliminar aquellas semillas que resulten redundantes, que son la mayoría, pues el  $UMB2=0.95$  es muy poco restrictivo. Otras veces, las semillas son rechazadas por ser algo intermedio o suma de otras ya seleccionadas. Las semillas seleccionadas deben ser lo suficientemente diferentes entre sí como para que la asignación posterior sea fiable. Si dos semillas son excesivamente similares, pequeñas variaciones entre dos días muy parecidos podrían clasificarlos en tipos diferentes.
- También se eliminarán aquellas semillas cuya extraña configuración y escaso número de elementos permitan suponer que se trata de una repetición de repartos de precipitación debida al azar, más que a existir un origen sinóptico común.
- Las semillas seleccionadas, deben, si es posible, cubrir conjuntamente la práctica totalidad de la geografía estudiada, para que cualquier día con precipitación sinóptica concentrada en una zona amplia pueda ser clasificado fiablemente.
- Si un tipo de reparto aparece para varios valores de UMB1, de seleccionarse alguno como semilla, sería el grupo correspondiente al mayor valor del mismo, para conseguir una mayor similitud interna (evidentemente, al disminuir el valor de UMB1, habrá un aumento del número de días asignados y una disminución de la similitud entre ellos y el centroide).
- Deben satisfacer las exigencias derivadas de la experiencia, de modo que queden representados los repartos de precipitación que se conocen subjetivamente, especialmente

los de origen sinóptico.

#### 4.5.2) Asignación de días a las semillas

Una vez seleccionadas las semillas, se procede a clasificar los días del período de estudio, simplemente asignando cada día a la semilla con la que tenga mayor similitud, de acuerdo a la medida de disimilitud seleccionada. Si el día no supera un cierto umbral UMB3 de similitud con ninguna semilla, se considera inclasificado.

### 4.6. SELECCIÓN DE LAS VARIABLES Y LA MEDIDA DE DISIMILITUD

Para optimizar la utilización de este método de "clustering", que desde las primeras pruebas ofreció unos excelentes resultados, deben seleccionarse adecuadamente la medida de disimilitud y las variables que representen los campos. En un primer paso, se preseleccionan unas semillas, con el método descrito, aplicado sobre un campo de datos y una medida de disimilitud que se consideren *a priori* aceptables. A partir de esas semillas se generan diferentes clasificaciones con el fin de seleccionar la mejor combinación de medida de disimilitud-variables que representan el campo de precipitación. Cada clasificación tendrá asociada una de estas combinaciones. Los centroides con los que medir disimilitudes se calculan en cada caso como media de los días que componen las semillas, representados por los campos que correspondan)

La selección de la mejor combinación, generado un número suficiente de clasificaciones, se reduce a determinar qué clasificación es la más acertada. Y para ello se ha utilizado de nuevo el criterio subjetivo.

Para comparar dos o más clasificaciones, se analizan visualmente los días que no clasifiquen todas en el mismo tipo. El estudio de más de 2000 discrepancias ha permitido sacar las siguientes conclusiones:

- Las distancias y pseudodistancias, ponderadas o no, clasifican considerablemente peor que los índices de correlación. La razón de esto es que se consideran repartos de precipitación similares aquellos que presenten formas de reparto parecidas, aunque sus niveles de precipitación sean diferentes. Y el coeficiente de correlación evalúa el parecido de las formas, independientemente de los niveles, pero éstos tienen una importancia capital en la distancia o pseudodistancia.
- Cuando los campos se representan por coeficientes de componentes principales, los índices de correlación ponderados por la varianza explicada por cada componente funcionan apreciablemente mejor que si no se ponderan. En algunos campos, la primera componente puede explicar el 25% de la varianza y la última el 0.25%, y es lógico que las diferencias en los coeficientes de aquella pesen más que las de los de esta.
- Las componentes principales sin rotar (CPSs) funcionan mucho peor que las rotadas ortogonalmente (CPRs), y estas peor que las rotadas oblicuamente (CPOs). Este hecho, constatado en la comparación de clasificaciones, concuerda con lo que cabía esperar, en virtud de las explicaciones teóricas del apartado 4.1.
- En contra de lo presumible, en general funcionan mejor los campos sin transformar que los de raíz cúbica de precipitación, aunque, como se ha comentado, el suavizado que

supone esta transformación hace que la correlación entre series correspondientes a pares de observatorios sea mayor para valores transformados.

- **La combinación seleccionada fue la del índice de correlación sobre 22 componentes principales de los valores de precipitación sin transformar, rotadas oblicuamente, ponderando por la varianza que explica cada componente.**
- La clasificación correspondiente al índice de correlación sobre valores sin transformar de precipitación en puntos de rejilla, ofrece unos excelentes resultados, salvo para el tipo 1, sin precipitaciones apreciables, al que asigna muchos menos días que otras clasificaciones. Se debe a que este índice, como se ha dicho, atiende a las formas, y cuando un día presenta, precipitación sólo en un punto y muy escasa, por ejemplo en San Sebastián, el índice lo asigna al tipo de precipitación concentrada en la zona, en este caso al tipo 4 (ver figuras 7 y 10), pues, ignorando el nivel de precipitación, la forma se parece más a la del tipo 4 que a la del 1, que es prácticamente plana. En cambio, la CPO correspondiente a precipitaciones en esa zona, en este caso la 6 (ver figura 6 a)), abarca varios puntos, es decir, es

$$\text{Cos}_6 = c_{6,1}P_1 + c_{6,2}P_2 + \dots + c_{6,j}P_j + \dots + c_{6,203}P_{203}$$

con valores altos de  $c_{6,j}$  para varios puntos de la zona cantábrica. Y por ello su coeficiente  $\text{Cos}_6$  no tomará un valor apreciable salvo si hay precipitación en varios de esos puntos. Por eso con CPOs se asigna mejor al tipo 1 que con valores en punto de rejilla. Sin embargo, exceptuando los días de estas características, esta clasificación asigna al resto de tipos con el mismo acierto que la combinación seleccionada, sin presentar apenas discrepancias con ella. Si se supera este problema, esta clasificación puede ser la más adecuada, por su sencillez de aplicación, tanto práctica como teórica, al no exigir la selección de componentes principales, etc.

#### 4.7. PROCESO FINAL DE CLASIFICACIÓN

Una vez seleccionada la medida de disimilitud y las variables a utilizar, se vuelve a ejecutar con ellas el método de Petisco modificado, y se hace la selección de las semillas definitivas, siguiendo las pautas indicadas, y la clasificación final, asignando cada día del período considerado a la semilla más parecida. Se ha utilizado por tanto un valor  $\text{UMB3}=0$ , de manera que la clasificación sea completa, sin días inclasificados. Así se forman grupos menos homogéneos, pero no se desaprovecha la información que supone el saber que cierto día, aunque es poco similar a todas las semillas, es más similar a la del tipo al que se ha asignado que a las del resto.

Con esto queda finalizada la clasificación de superficie, que ha generado unos tipos de reparto de precipitación, representados en las figuras 7 a) a 24 a), y la serie de los días del período 1941-1990 clasificados en esos tipos. Las figuras 25 a 30 representan las frecuencias anuales de aparición de los diferentes tipos a lo largo del período.

## 5. RESULTADOS

### 5.1. TIPOS DE REPARTO DE PRECIPITACIÓN

Son los que aparecen en las figuras 7 a) a 24 a). Representan el reparto de precipitación de un día imaginario, de "máxima densidad" relativa, de manera que hay un máximo de días reales similares a él en más de cierto umbral. La precipitación en cada punto de rejilla de estas semillas o tipos, se calcula como media de las precipitaciones en ese punto de todos los días que las componen. Representan, por tanto, la precipitación acumulada de un día tipo de ese reparto, en litros por metro cuadrado.

Para resaltar su significación sinóptica, se acompaña cada tipo con los campos medios de presión a nivel del mar y geopotencial de 500mb. Hecha la clasificación del período 1961-1990, se consideran para calcular estos campos medios el 50% de los días de cada tipo que más se parecen a sus semillas respectivas. Así se evita incluir en las medias días que no pertenecen claramente a ningún tipo, pero que han sido clasificados al más parecido. Representan entonces los campos de un día tipo (ligeramente diferente al del párrafo anterior, pues en su cálculo entran algunos días que no constituyen las semillas), en mb y m. Si se han usado más días para calcular las medias que en el caso anterior, es para mostrar que, aún reduciendo la similitud interna de los grupos (dentro de un orden, cabe suponer que la media de todos los días del grupo sea menos satisfactoria), se obtienen buenos resultados.

Recordando lo expuesto en el apartado 3, actualmente se dispone de 9 campos de variables atmosféricas. La presión a nivel del mar describe la configuración en la baja troposfera. El geopotencial de 500mb informa sobre el forzamiento dinámico en la troposfera media. La elección de ambas variables es sólo provisional, aunque ambas tienen la ventaja de su uso tradicional en climatología sinóptica y de mejores y más largos bancos de datos.

Uno de los alicientes de este tipo de aproximación indirecta a la generación de clasificaciones sinópticas puede ser esta: la posibilidad de proyectar la clasificación de efectos terrestres sobre cualquier campo de variables atmosféricas.

Desde el punto de vista de determinación de la precipitación, SLP y Z500 aportan información sobre la direccionalidad de los flujos y forzamientos dinámicos. Puede echarse de menos la selección de un campo referente a la inestabilidad de carácter termodinámico. Se construyó un campo sustrayendo para cada punto el campo de T700 del de T850. Pretendía captar el gradiente térmico en la troposfera baja, pero no funcionó coherentemente. En un futuro se probará con T850-T500, ya que esta variable se incluye en la última versión del CD-ROM citado en el apartado 3, de inminente publicación.

A continuación, se comenta sucintamente la significación de los 18 tipos seleccionados.

\* **TIPO 1: Sin Precipitaciones Apreciables.** En superficie, el anticiclón centrado en el golfo de Vizcaya afecta a toda la Península. En 500mb, una potente dorsal indica fenómenos de subsidencia.

\* **TIPO 2: Precipitaciones en Galicia.** En superficie, el anticiclón de las Azores afecta al centro y sur peninsulares. En Galicia, el flujo es ciclónico del oeste-suroeste. Esta ausencia de componente norte ocasiona que la vertiente cantábrica quede a la sombra de las precipitaciones,

concentrándose éstas en la Galicia atlántica. En altura, el flujo es zonal del oeste-suroeste, debilitándose mucho hacia el sur.

\* **TIPO 3: Precipitaciones en Galicia y Vertiente Cantábrica.** El anticiclón se centra en las Azores, condicionando junto con las bajas presiones británicas un flujo del noroeste sobre la Península, que en la zona norte del país tiene clara curvatura ciclónica (asociada a circulación de frentes por el Cantábrico). En 500mb, el flujo es oeste-noroeste.

\* **TIPO 4: Precipitaciones en el Cantábrico Oriental.** El anticiclón centrado al este de las Azores afecta a las vertientes atlántica y cantábrica. Las precipitaciones son generadas mayoritariamente por el forzamiento topográfico que ejerce la Cordillera Cantábrica sobre las masas húmedas atlánticas que circulan sobre ella. Las lluvias descargan en la zona oriental debido a la menor influencia anticiclónica en la región, con respecto al Cantábrico occidental. En altura, el flujo es del noroeste y la Península se ve afectada por la advección de vorticidad negativa de la dorsal atlántica (forzamiento dinámico negativo).

\* **TIPO 5: Precipitaciones en el Interior Norte Peninsular.** En superficie, la baja térmica africana se extiende al interior peninsular (formando la baja térmica ibérica). En altura existe forzamiento dinámico por la pequeña vaguada con eje en las inmediaciones de las costas portuguesas. Estas condiciones suelen coincidir con inestabilidad termodinámica en las zonas mesetarias más altas del interior peninsular (inmediaciones del Sistema Ibérico y La Mancha), originándose fuertes lluvias convectivas. Las diferencias entre este tipo y el 17, Precipitaciones en el Interior Sur Peninsular, no son claras en estos campos medios. El campo de 500mb del tipo 17 apunta a representar situaciones de bloqueo sobre Gran Bretaña, que vienen casi siempre acompañadas de una anomalía negativa de geopotencial sobre Iberia. En estas condiciones, el forzamiento dinámico más intenso corresponde a la parte sur de España.

\* **TIPO 6: Precipitaciones en Pirineos.** En superficie, la situación está poco definida, apuntándose la posible existencia de baja térmica. En altura, aparece una ligera vaguada con el eje en el centro oeste peninsular. El débil forzamiento dinámico positivo (zona de advección de vorticidad de la vaguada) afecta al nordeste peninsular. La topografía pirenaica facilita el disparo de procesos convectivos.

\* **TIPO 7: Precipitaciones en Cataluña.** En superficie, el flujo no está muy definido. En 500mb una vaguada débil tiene su eje sobre los Pirineos. El forzamiento dinámico positivo actúa débilmente sobre Cataluña. Si existe inestabilidad termodinámica, se producen precipitaciones convectivas.

\* **TIPO 8: Precipitaciones en el Bajo Valle del Ebro.** En superficie, existe un débil flujo mediterráneo sobre la costa catalana. En altura, una vaguada muy acentuada que es casi una depresión cerrada, se sitúa en el eje peninsular. En estas condiciones, hay un potente forzamiento dinámico en el nordeste peninsular. El resultado son fuertes precipitaciones convectivas en la cuenca baja del Ebro.

\* **TIPO 9: Precipitaciones en la Comunidad Valenciana.** En superficie, el flujo es de levante, perpendicular a la costa valenciana, debido a las situaciones del anticiclón europeo y la baja térmica africana. Aire húmedo del Mediterráneo es adveccionado sobre los contrafuertes montañosos de la región. En altura, una depresión centrada sobre la Península genera una fuerte advección de vorticidad positiva en niveles altos. El resultado suele ser la formación de núcleos convectivos, a veces mesoescalares, que generan fuertes precipitaciones.

\* **TIPO 10: Precipitaciones en Baleares.** Una baja mediterránea en superficie se insinúa sobre Cerdeña, y su posición relativa con respecto al anticiclón europeo, genera flujo del nordeste.

Si esto ocurre en otoño o principios de invierno, cuando el Mediterráneo está relativamente más caliente que el aire europeo que se le advecta, se desatan fenómenos convectivos, acentuados por el forzamiento topográfico del relieve insular. La precipitación sólo cae en Baleares, debido a que en el resto de la zona nordeste española, el forzamiento negativo de la dorsal en 500mb que se aproxima es demasiado fuerte.

\* **TIPO 11: Precipitaciones en el Sudeste Peninsular.** En superficie, el flujo es de levante con una situación muy similar a la del tipo 9. En altura, una depresión centrada en el Estrecho de Gibraltar, genera un potente forzamiento dinámico en el sudeste peninsular, focalizando aquí el núcleo de las precipitaciones convectivas.

\* **TIPO 12: Precipitaciones en el Estrecho y Andalucía Occidental.** En superficie, una baja débil en el Golfo de Cádiz, advecta aire húmedo atlántico con componente sur en el área del Estrecho. En altura, una pequeña depresión en el suroeste peninsular genera un forzamiento dinámico positivo en la zona, provocando precipitación, que se ve mayorada en las serranías gaditanas.

\* **TIPO 13: Precipitaciones en el Sudoeste Peninsular (Ábrego).** En superficie, una borrasca se centra frente a las costas portuguesas advectando flujo del suroeste en el sudoeste peninsular. Dicho flujo puede traer frentes o fronteras cálidos asociados, que son los responsables de la precipitación. En altura, la vaguada de 500mb yace un poco al oeste de la baja superficial.

\* **TIPO 14: Precipitaciones en Toda la Vertiente Atlántica y Pirineos.** En superficie, una potente borrasca al oeste de la zona marítima de Gran Sol advecta sus frentes sobre la fachada atlántica peninsular, dejando precipitaciones en todas las cuencas atlánticas. La España cantábrica y mediterránea no recibe casi precipitación por encontrarse a sotavento del flujo. En esta zona, solamente los Pirineos occidentales reciben precipitaciones de importancia, ya que, aunque el aire llega aquí bastante seco, sufre un importante forzamiento topográfico. Este mismo forzamiento explica los máximos relativos de precipitación en los principales macizos montañosos de las cuencas atlánticas. En altura, existe un intenso flujo zonal.

\* **TIPO 15: Precipitaciones en el Oeste peninsular.** En superficie, la borrasca se sitúa al oeste de Gran Sol, como en el caso anterior, pero con menor gradiente barométrico y flujo claramente del suroeste, que no alcanza la zona del Estrecho. Destaca la precipitación en la vertiente sur de la Sierras de Gata y Gredos, debido al fuerte forzamiento orográfico que presentan a este flujo. En altura aparece una profunda vaguada, generando un flujo generalizado del suroeste sobre la península.

\* **TIPO 16: Precipitaciones en el Noroeste peninsular.** En superficie, la borrasca se centra al oeste de Irlanda y advecta sus frentes sobre la esquina noroeste peninsular, con una clara componente oeste-suroeste. Esta dirección del flujo deja a sotavento la vertiente cantábrica, disminuyendo aquí notablemente la precipitación. Una cuña anticiclónica afecta a la mitad sur peninsular, impidiendo que los frentes la alcancen. En altura, el flujo es zonal con componente suroeste, y adquiere curvatura anticiclónica en el sur y este peninsulares.

\* **TIPO 17: Precipitaciones en el Interior Sur peninsular.** Las características son muy similares a las del tipo 5, y las escasas diferencias advertidas ya se han comentado en aquel apartado.

\* **TIPO 18: Precipitaciones en el Norte peninsular.** En altura, la configuración es muy parecida a la del tipo 6. Una pequeña vaguada con eje sobre el oeste de la Península genera un débil forzamiento dinámico en el centro y este de el País. Sin embargo, en superficie, el flujo, aunque suave, tiene una clara componente norte sobre la vertiente cantábrica. Como resultado

aprecen chubascos convectivos en la Cordillera Cantábrica y alto Ebro.

## 5.2. APLICACIÓN DE LA CLASIFICACIÓN AL PERÍODO 1941-1990

Se presentan las frecuencias ( $^{\circ}/_{00}$ ) anuales de aparición de cada reparto de precipitación en el período 1941-1990 (figs 25 a 30). Las tendencias reflejadas en estas series (por ejemplo, la disminución de frecuencia del tipo 14, precipitaciones en toda la vertiente atlántica y Pirineos, a partir de mediados de los 70) no pueden considerarse más que primeras conclusiones. Este Servicio trabaja actualmente en analizar los intervalos de confianza de las frecuencias anuales, para determinar la significación estadística de dichas tendencias. Además, se pretende completar el período de datos hasta 1995 y alargar así las series, para su mayor utilidad en la detección de un posible cambio de clima. Si se incluyen las series en este apartado de resultados, es con carácter meramente orientativo, y para dar una idea de las posibilidades que ofrecen este tipo de clasificaciones.

La fig 31 representa la frecuencia total de cada tipo en el período ( $^{\circ}/_{00}$ ) y su distribución en los períodos estacionales (%). Las estaciones aquí utilizadas son de tres meses, comenzando el invierno el 1 de diciembre. Las limitaciones anteriores apenas pesan aquí, por haberse estimado las frecuencias sobre poblaciones mucho mayores.

En lo referente a la frecuencia total, el tipo 1 (sin precipitaciones apreciables) es con mucho el más frecuente, con un 28%. Los tipos 2, 3 y 4, genéricamente galaico-cantábricos, representan individualmente en torno al 8%, y conjuntamente el 23%, bastante próximo al tipo 1. El tipo 7, de precipitaciones en Cataluña, también es bastante frecuente.

Respecto a las frecuencias estacionales:

- El tipo 1 presenta un claro máximo estival, como era presumible.
- Los tipos 3 y 4 (precipitaciones cantábricas) y los 2, 12, 13, 14, 15 y 16 (precipitaciones atlánticas), de indudable origen frontal, presentan mínimos muy claros en verano y en general máximos invernales. Esos mínimos estivales son mucho más evidentes en los tipos 12, 13 y 14, que son generados por borrascas desplazadas hacia el sur (ver figs 7 b) a 24 b)). El tipo 15 es ligeramente discordante del resto, al no presentar un máximo invernal.
- Comportamiento inverso presentan los tipos 5, 6, 7, 8, 17 y 18, con claros mínimos invernales y máximos estivales (salvo el 17, con máximo primaveral). Este hecho concuerda con la naturaleza convectiva atribuida en los análisis atmosféricos del apartado 5.1.
- Por fin, los tipos 9, 10 y 11, de carácter mediterráneo, presentan gran actividad otoñal, asociada con la inestabilidad que se produce en esa época, con las aguas del Mediterráneo calientes, y frecuentes forzamientos dinámicos y bajas temperaturas, en la troposfera media.

## 6. AGRADECIMIENTOS

Se agradece la colaboración al Centro de Proceso de Datos de la Universidad Complutense de Madrid, por las facilidades prestadas en la utilización de sus medios informáticos, y a todo su personal, por su gran disposición y por la asesoría teórica y práctica prestada en todo momento y con gran eficacia, sin cuya colaboración ejemplar este proyecto se hubiera visto muy dificultado. Se agradece en especial la asesoría en programación de D. Pedro Cuesta.

Del mismo modo se agradece la colaboración de D. Luis Balairón Ruiz, Jefe del Servicio de Análisis e Investigación del Clima del INM, y del personal de este Servicio: D. Eduardo Petisco, D. José María Martín, D. Javier López y D. Javier Rodríguez Izcara. En especial, la cesión del programa desarrollado por D. Eduardo Petisco, que ha sido de enorme utilidad. Así mismo, se agradece la asesoría de D. Ricardo Riosalido, Jefe del Servicio de Técnicas y Análisis de Predicción.

Se agradece a D. Ángel Fernández Cancio y a D. José Luis Allué, del Instituto Nacional de Investigaciones Agrarias, la asesoría técnica y la confianza personal depositada.

## 9. BIBLIOGRAFÍA.

- Ministerio de Obras Públicas y Transporte e Instituto Geográfico Nacional, 1992: Atlas Nacional de España. Sección II. Climatología.
- Borén, R., J. Ribalaygua y L. Balairón, 1995: Método analógico de simulación de escenarios climáticos a escala comarcal. Informe N° 4. Servicio de Análisis e Investigación del Clima, Instituto Nacional de Meteorología. Madrid. "In press".
- Buell C.E., 1979: On the Physical Interpretation of Empirical Orthogonal Functions. Preprints of the Sixth Conference on Probability and Statistics in Atmospheric Science. Banff, Alberta American Meteorological Society, 112-117.
- Calvo Sánchez, J., 1993: Clasificación de flujos en 500 hPa sobre la Península Ibérica. Nota técnica N° 32 del Servicio de Predicción Numérica del Instituto Nacional de Meteorología.
- Douglas, A.V. and P.J. Englehart, 1981: On a statistical relationship between rainfall in the central equatorial Pacific and the subsequent winter precipitation in Florida. Monthly Weather Review. 109, 2377-2382.
- Font Tullot, I., 1983: Climatología de España y Portugal. Instituto Nacional de Meteorología. Ministerio de Transportes, Turismo y Comunicaciones. Madrid.
- Júdez, L., 1989: Técnicas de Análisis de Datos Multidimensionales. M° Agricultura Pesca y Alimentación, Madrid.
- Lamb, H.H., 1972: British Isles weather types and a register of daily sequences of circulation patterns, 1861-1971. Geophysical Memoir 116, HMSO, London.
- Lanzante, J.R. and R.P. Harnack, 1982: Specification of United States Summer Season Precipitation. Monthly Weather Review, Vol.110, 1843-1850.
- Linés Escardó, A., 1981: Perturbaciones típicas que afectan a la Península y precipitaciones asociadas. Publicaciones del Instituto Nacional de Meteorología, serie A, N° 80. Madrid.
- Mass, C.F., H.J. Edmon, H.J. Friedman, N.R. Cheney, and E.E. Recker, 1987: The use of compact discs for the storage of large meteorological and oceanographic data sets. Bulletin of the American Meteorological Society. Vol 68(12), 1556-1558.
- Petisco, E. y Jose M<sup>a</sup> Martín, 1995: Caracterización de la Circulación General Atmosférica en la Península Ibérica y Baleares. Informe N°5, del Servicio de Análisis e Investigación del Clima. Instituto Nacional de Meteorología. Madrid. "In press".
- Preisendorfer, R.W., 1988: Principal Component Analysis in Meteorology and Oceanography. Elsevier. New York.
- Richman, M.B. and Lamb, P.J., 1985: Climatic Pattern Analysis of Three- and Seven-Day Summer Rainfall in the Central United States: Some methodological Considerations and

- a Regionalization. *Journal of Climate and Applied Meteorology*. Vol 24, 1325-1343.
- Sánchez Rodríguez, J., 1993: Situaciones Atmosféricas en España. Serie de monografías. MOPT. Instituto Nacional de Meteorología. Madrid.
  - SAS Institute Inc., 1990: SAS/STAT User's Guide, Version 6, Fourth Edition. North Caroline.
  - Von Storch, H., E.Zorita and U.Cubasch, 1993: Downscaling of climate change estimates to regional scales: an application to Iberian rainfall in winter time. *J. Climate* 6, 1161-1171.
  - White D., M. Richman and B. Yarnal, 1991: Climate Regionalization and Rotation of Principal Components. *International Journal of Climatology*, Vol 11, 1-25.
  - Wigley, T.M.L., P.D. Jones, K.R. Briffa and G. Smith, 1990: Obtaining Sub-grid Scale Information from Coarse-Resolution General Circulation Model Output. *Journal of Geophysical Research*, Vol 95, 1943-1953.
  - Willmott, C.J., 1977: A component analytic approach to precipitation regionalization in California. *Arch. Met. Geoph. Biokl.* Vol B24, 269-281.
  - Zorita, E., J. Hughes, D.Lettenmaier, and H.von Storch,1993: Stochastic downscaling of regionalcirculation patterns for climate model diagnosis and estimation of local precipitation. Max-Planck-Institute für Meteorologie, Report 109.

**ANEXO**



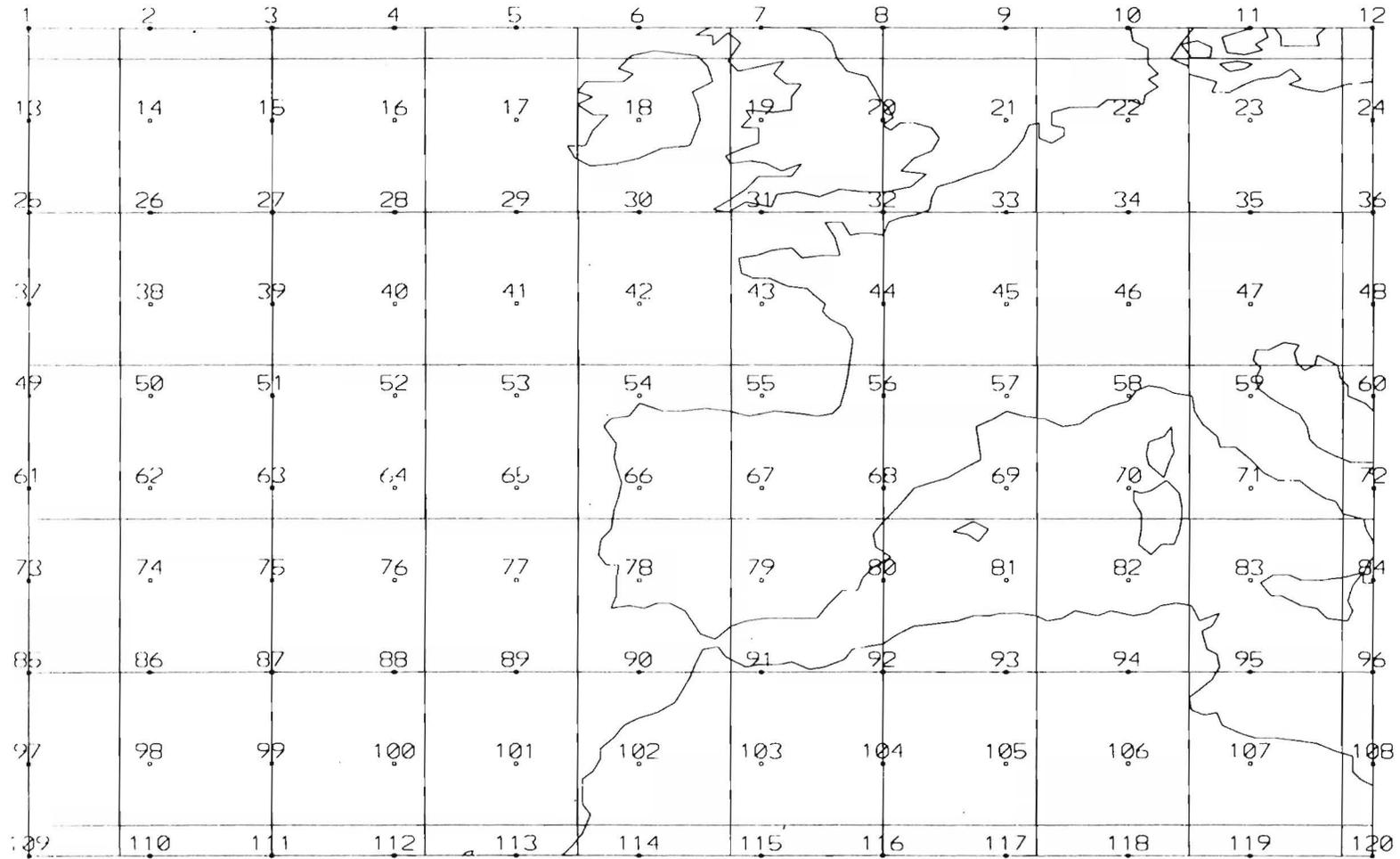


Figura 1: 120 puntos de grid para campos de altura.



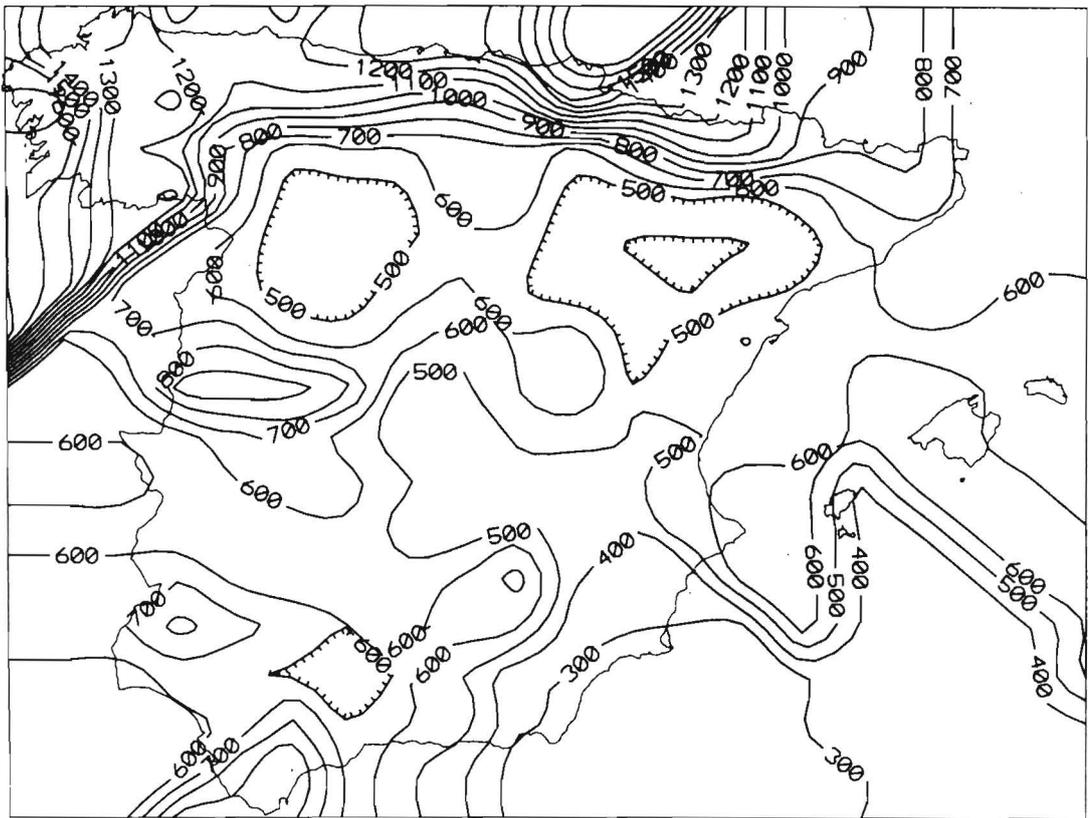


Figura 4: Precipitación anual media del período 1961-1990, estimada como media de los campos de precipitación en 203 puntos de rejilla, a partir de los 10957 días del período.

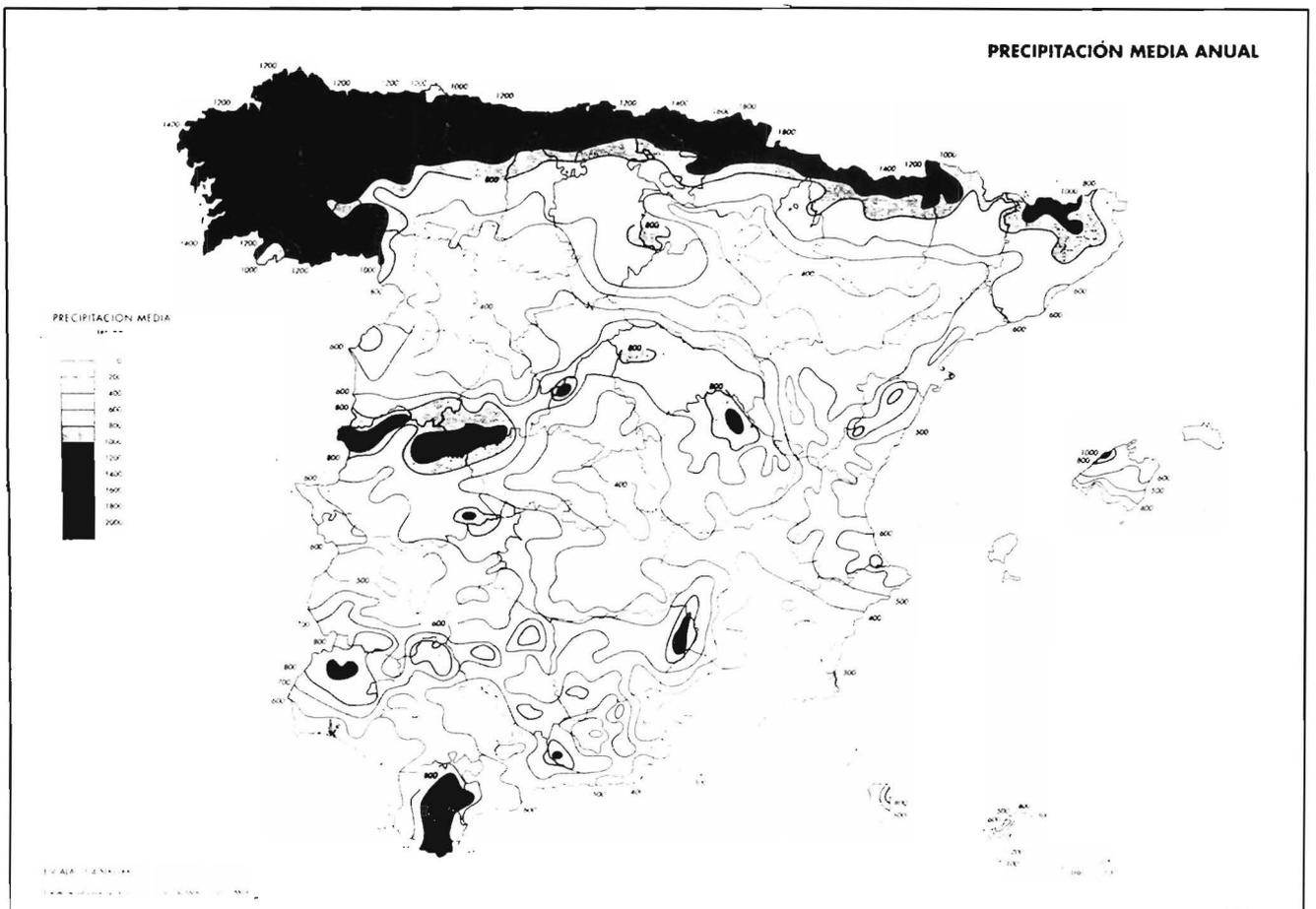


Figura 5: Precipitación media anual del período 1961-1990, estimada en el Atlas Nacional del Clima.

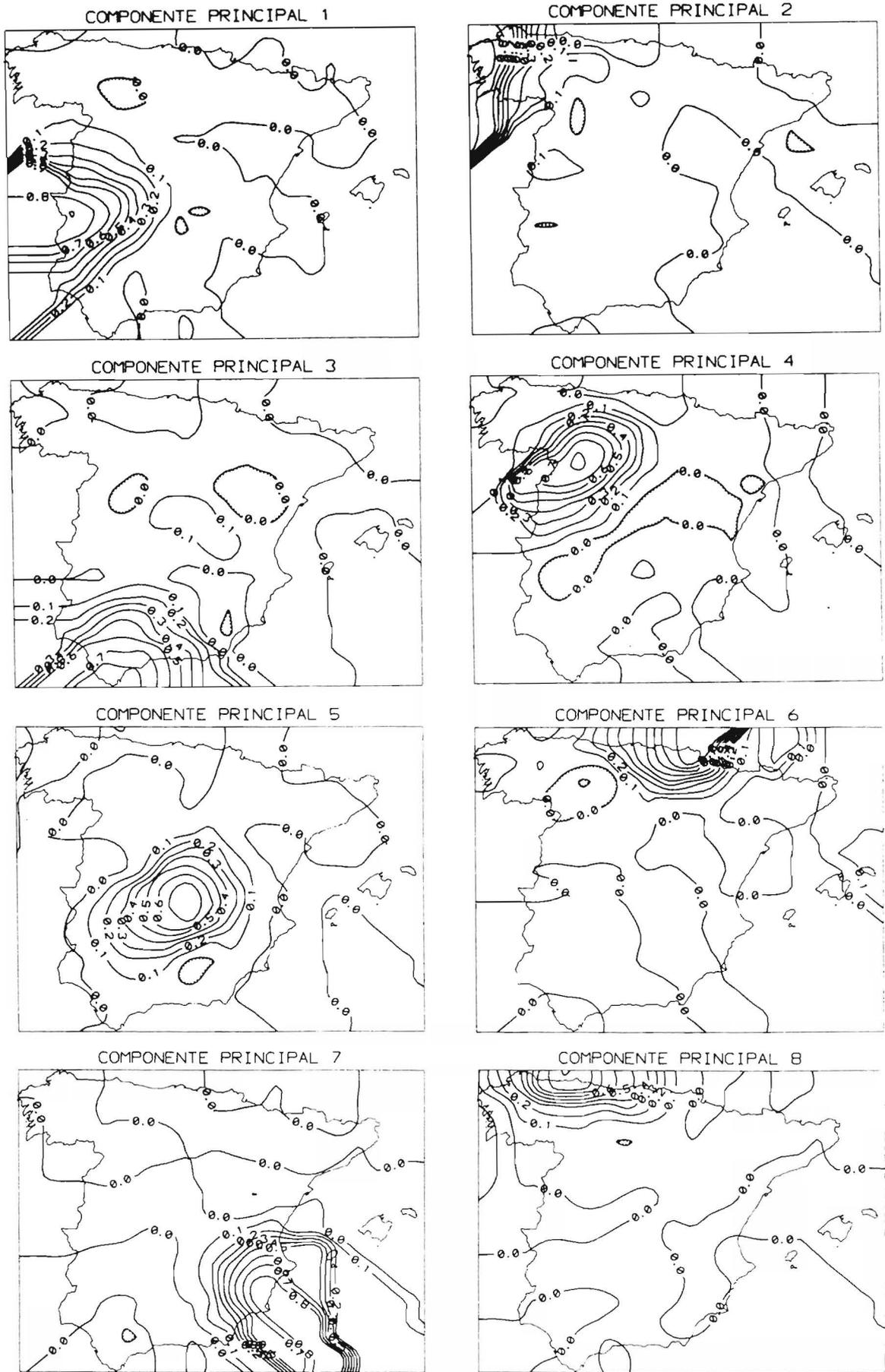


Fig. 6 a): "Factor loadings patterns" de las componentes principales rotadas oblicuamente 1 a 8.

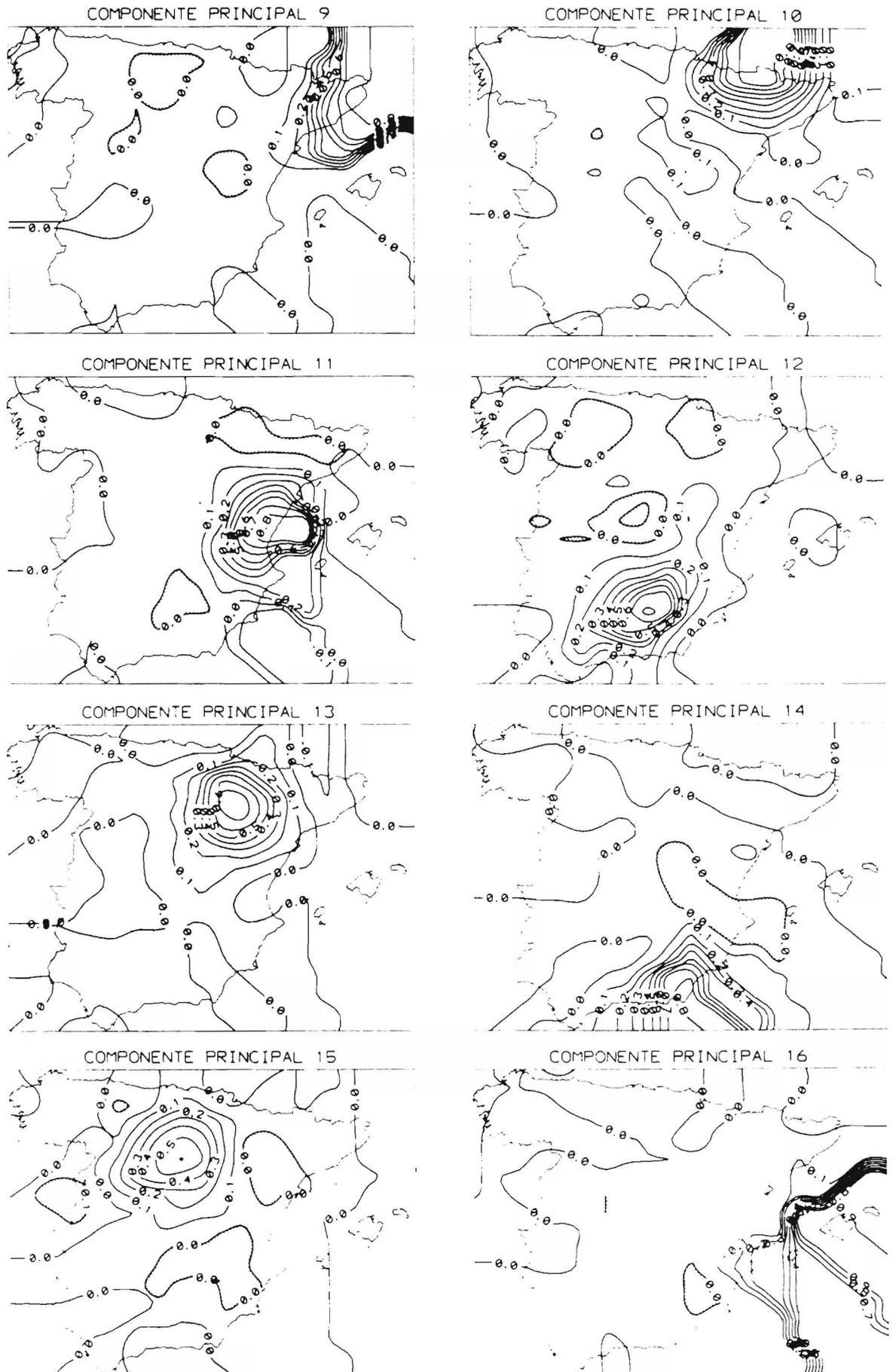


Fig. 6 b): "Factor loadings patterns" de las componentes principales rotadas oblicuamente 9 a 16.

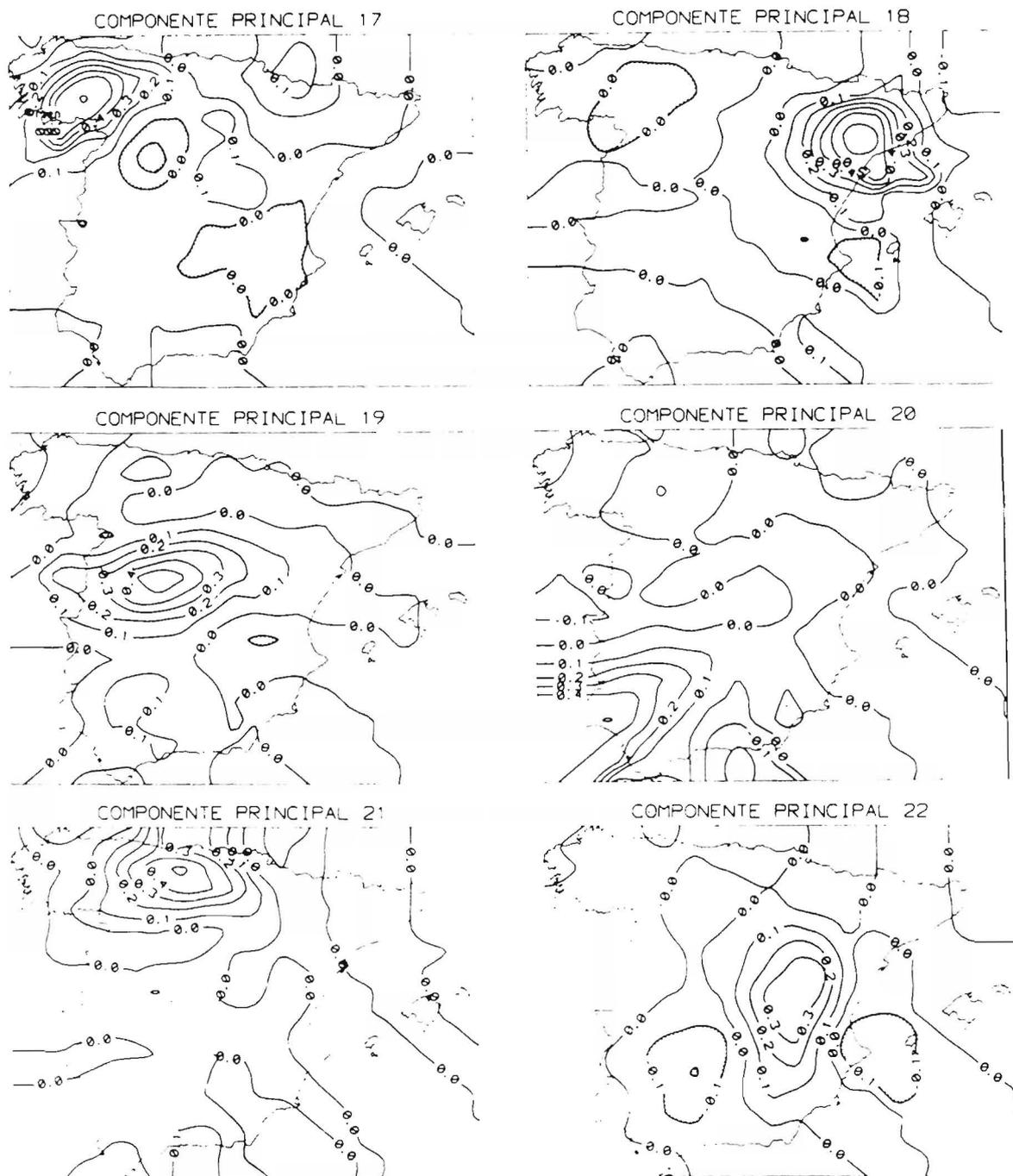


Fig. 6 c): "Factor loadings patterns" de las componentes principales rotadas oblicuamente 17 a 22.

PORCENTAJE DE VARIANZA TOTAL EXPLICADA

CP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
%	5.5	4.8	4.5	4.0	3.6	3.2	3.1	3.0	2.9	2.9	2.9	2.7	2.7	2.3	1.9	1.7	1.7	1.7	1.7	1.2	1.2	1.1

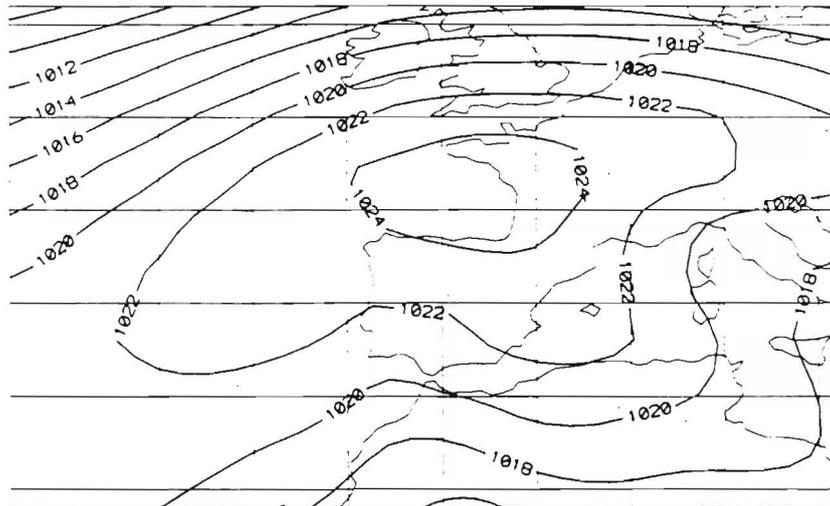
TIPO 1

a)



PRESION A NIVEL DEL MAR

b)



GEOPOTENCIAL 500mb

c)

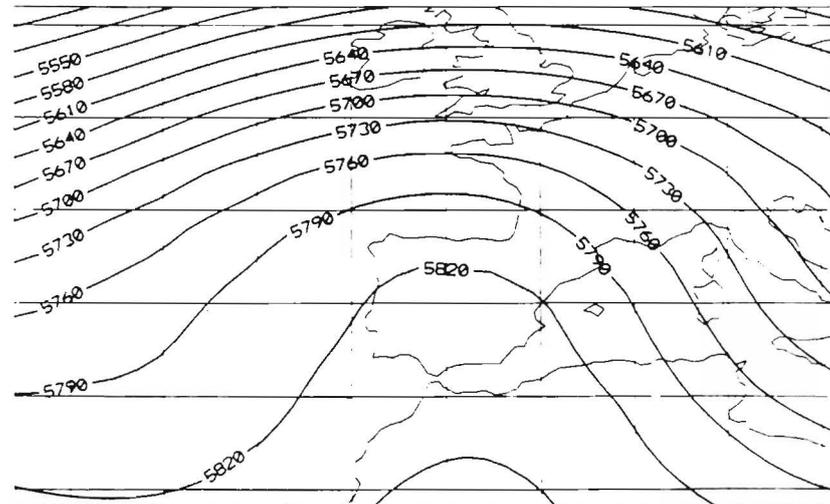


Figura 7: Tipo 1: Sin Precipitaciones Apreciables. a) Reparto de precipitación. b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 2

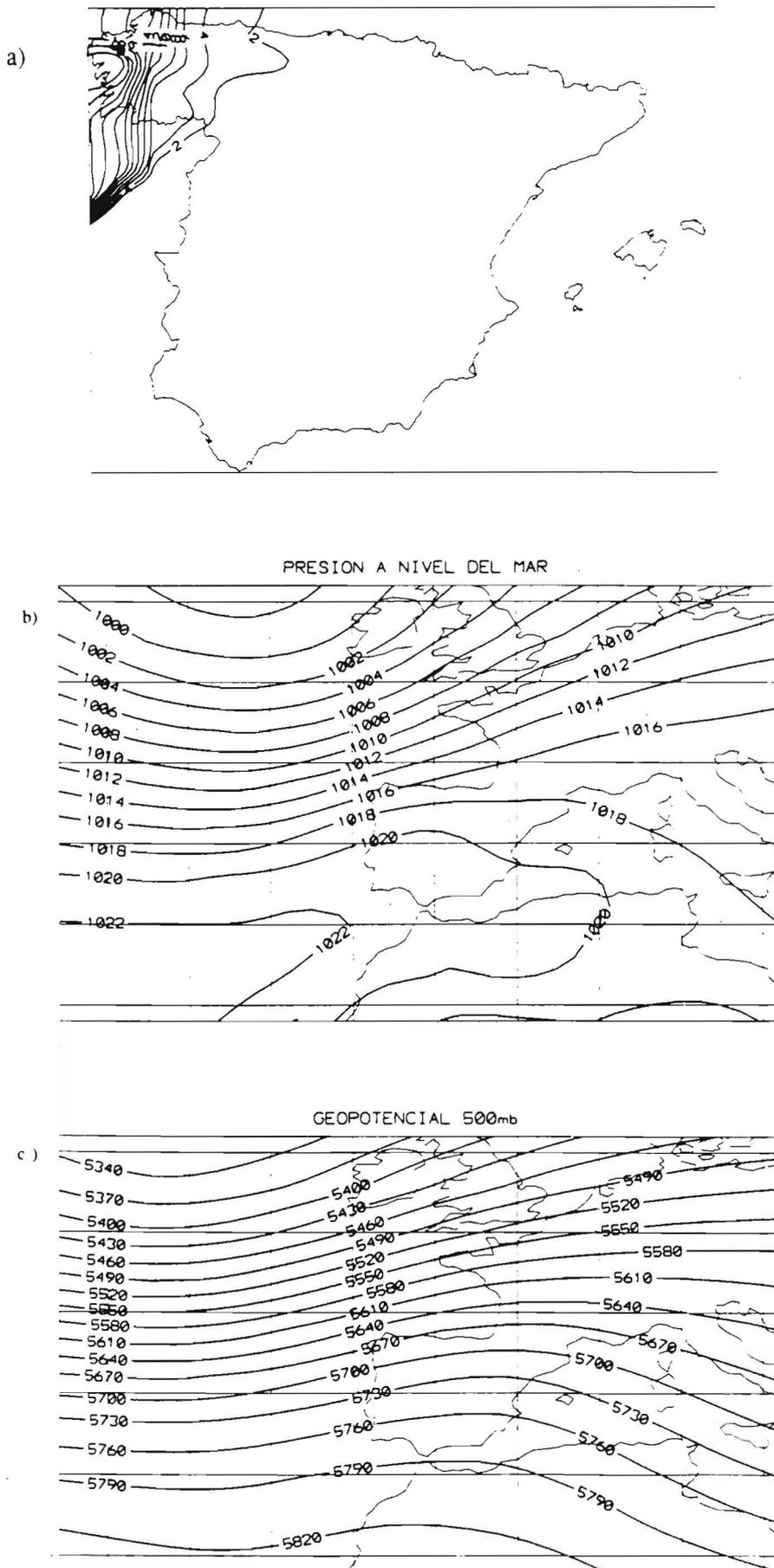


Figura 8: Tipo 2: Precipitaciones en Galicia. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 3

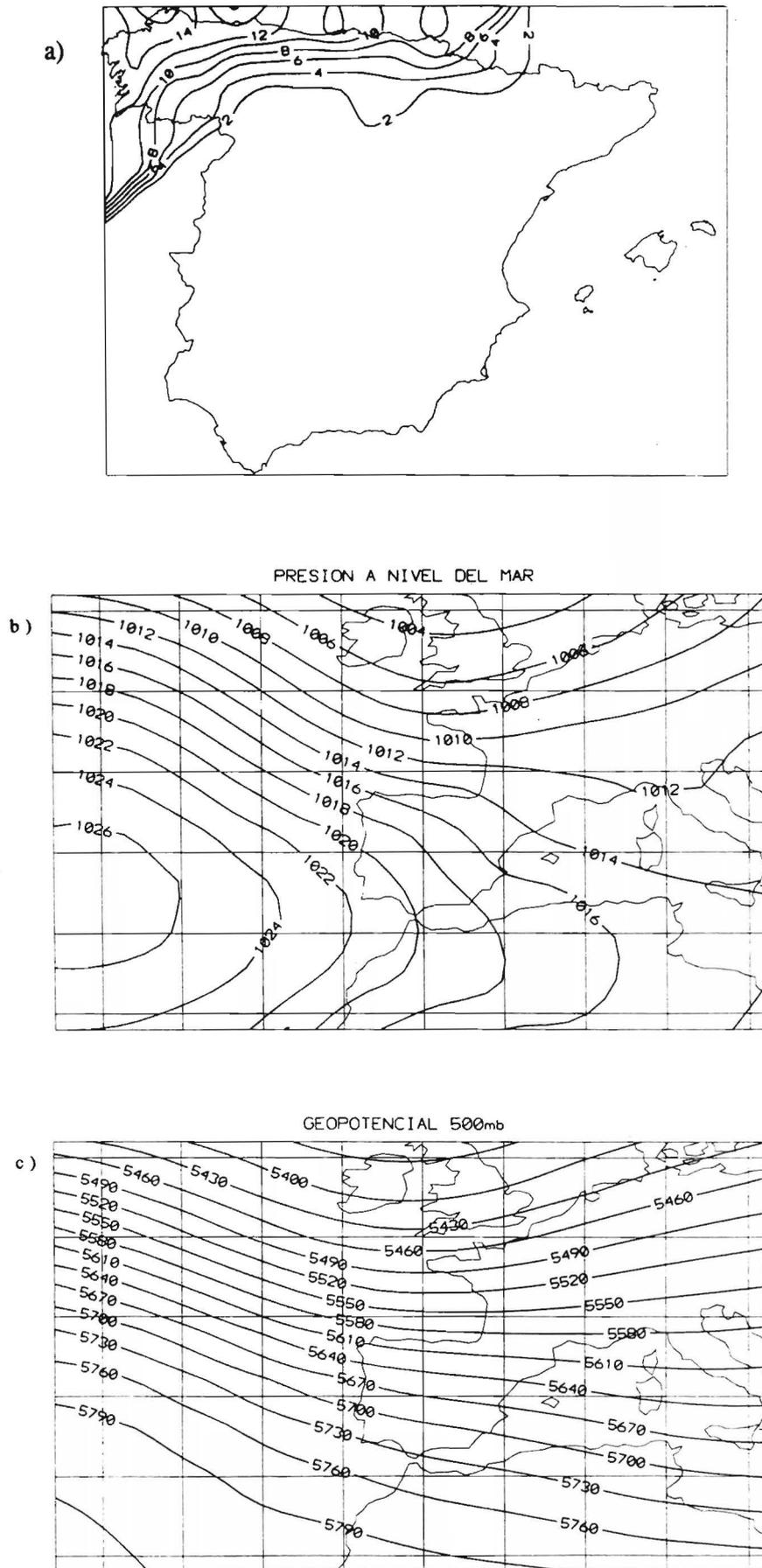


Figura 9: Tipo 3: Precipitaciones en Galicia y Vertiente Cantábrica. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 4

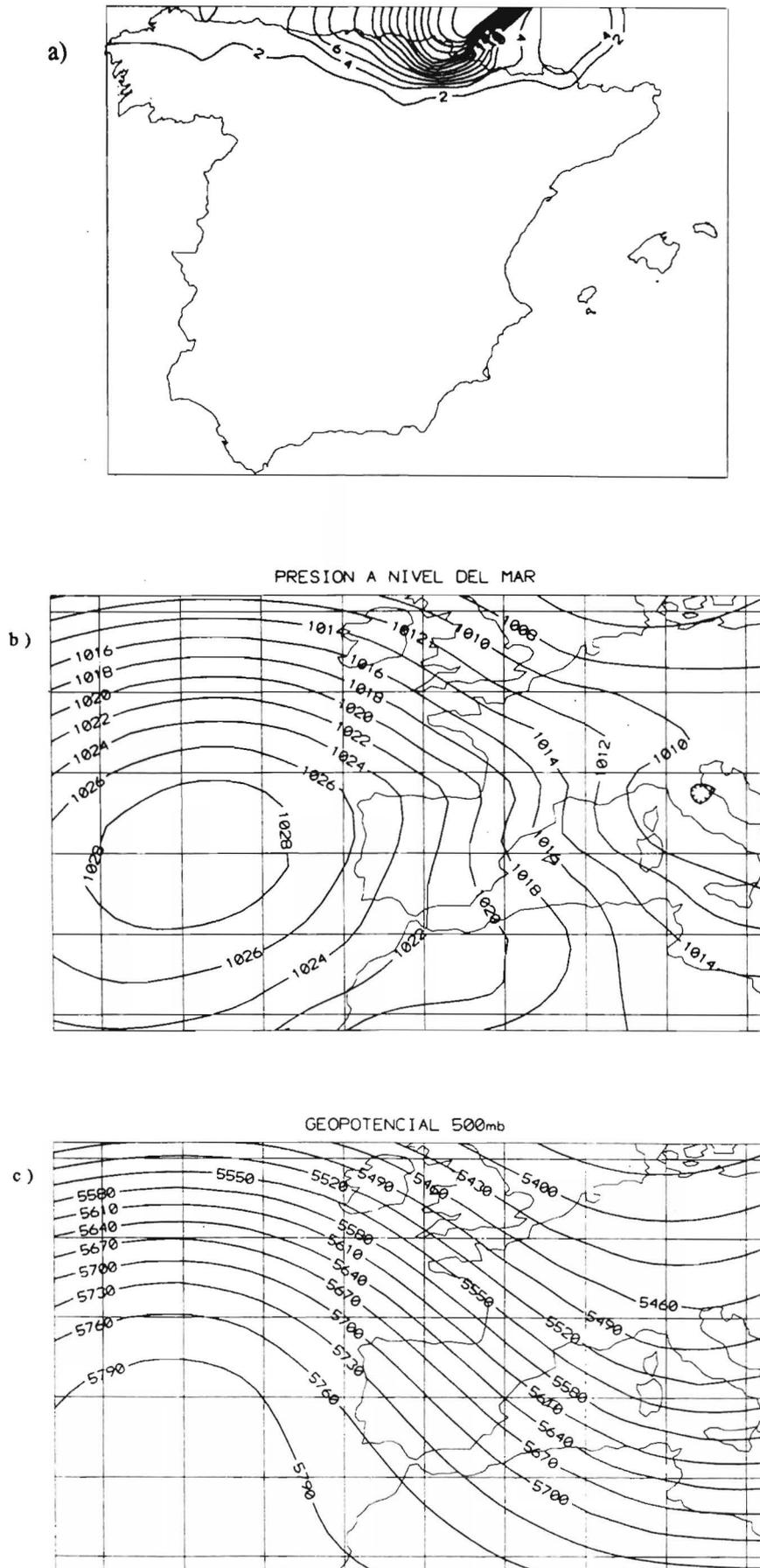


Figura 10: Tipo 4: Precipitaciones en el Cantábrico Oriental. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

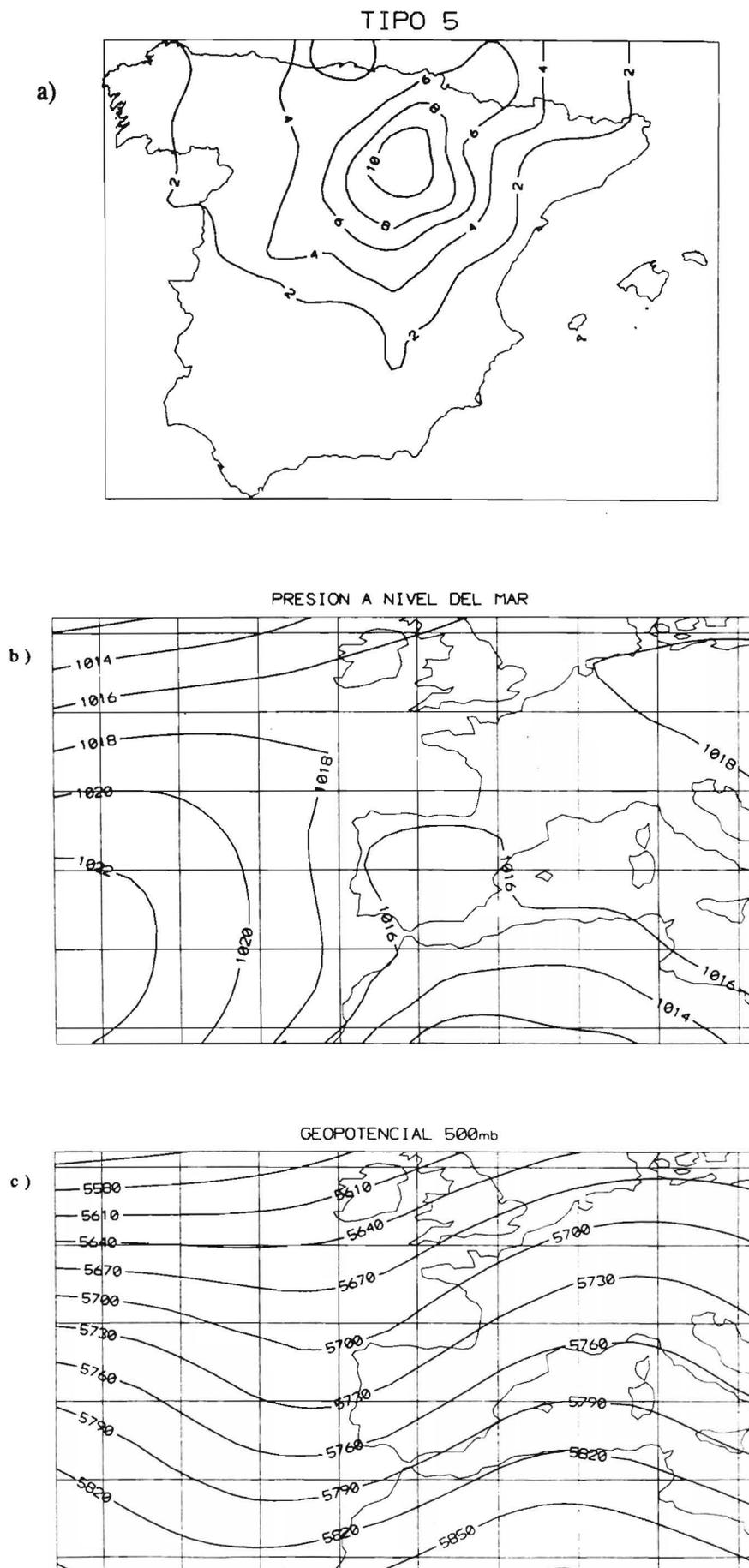
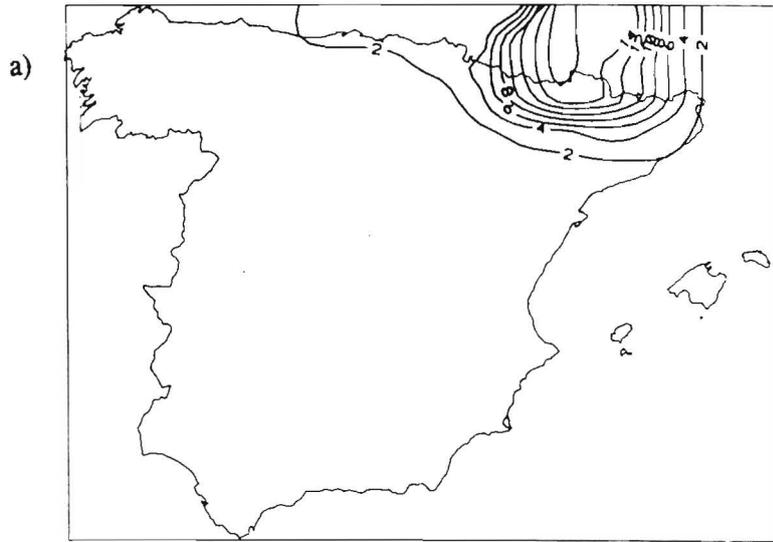
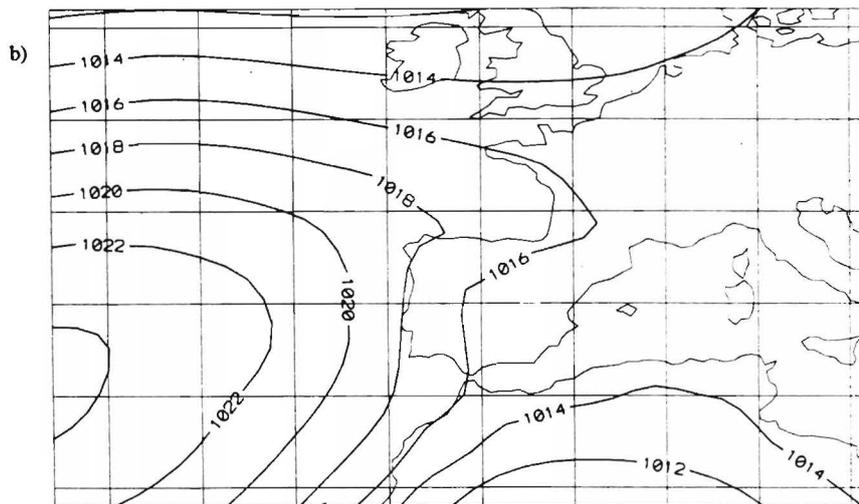


Figura 11: Tipo 5: Precipitaciones en el Interior Norte Peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 6



PRESION A NIVEL DEL MAR



GEOPOTENCIAL 500mb

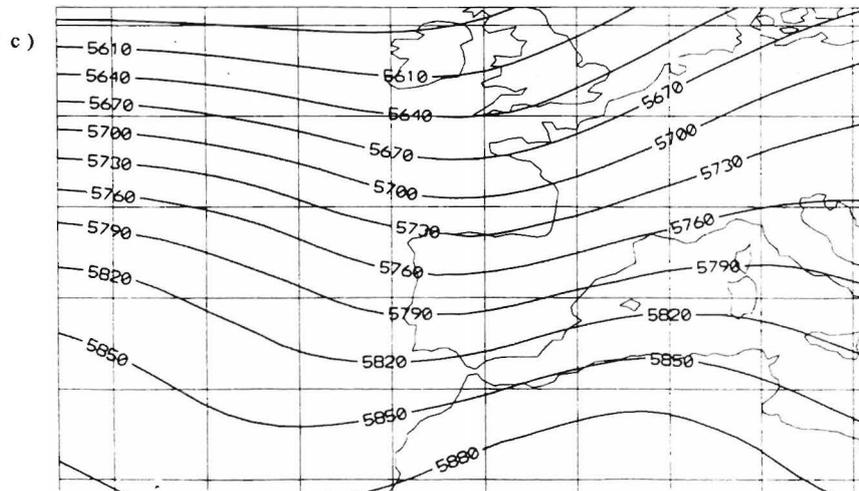
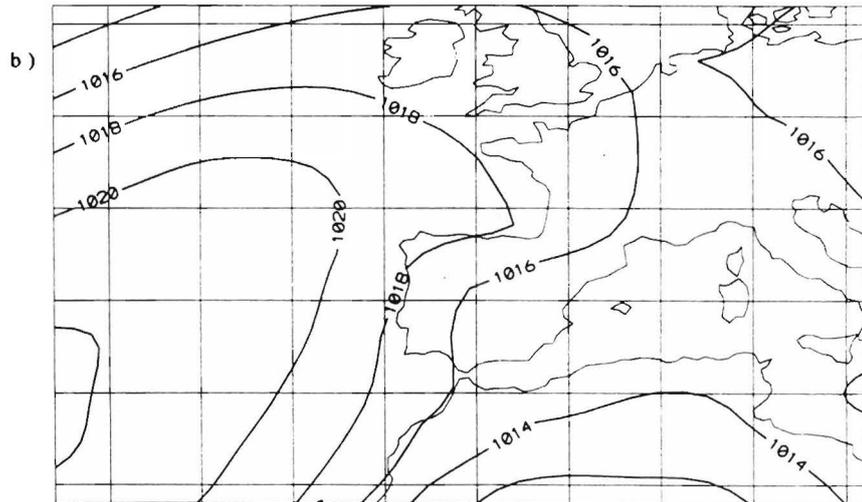


Figura 12: Tipo 6: Precipitaciones en Pirineos. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 7



PRESION A NIVEL DEL MAR



GEPOTENCIAL 500mb

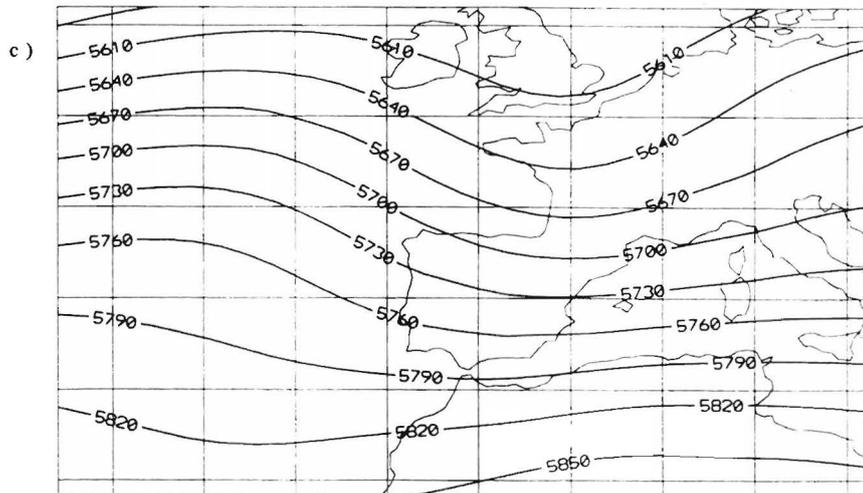
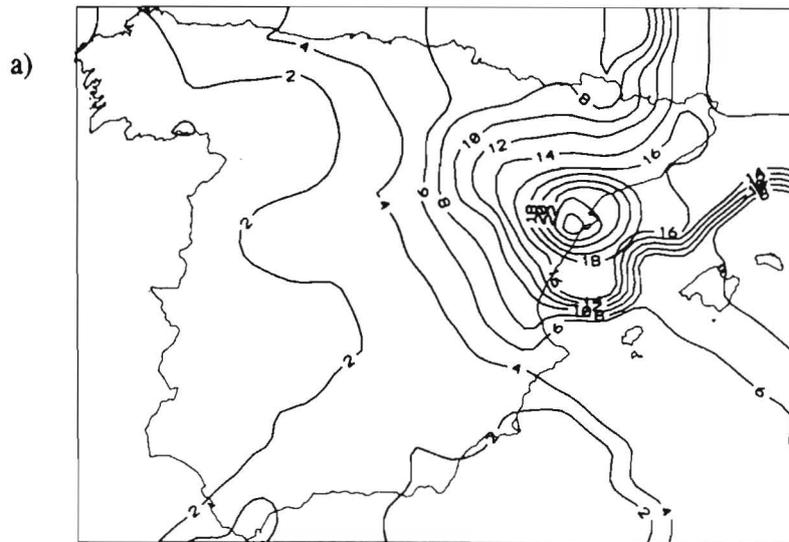
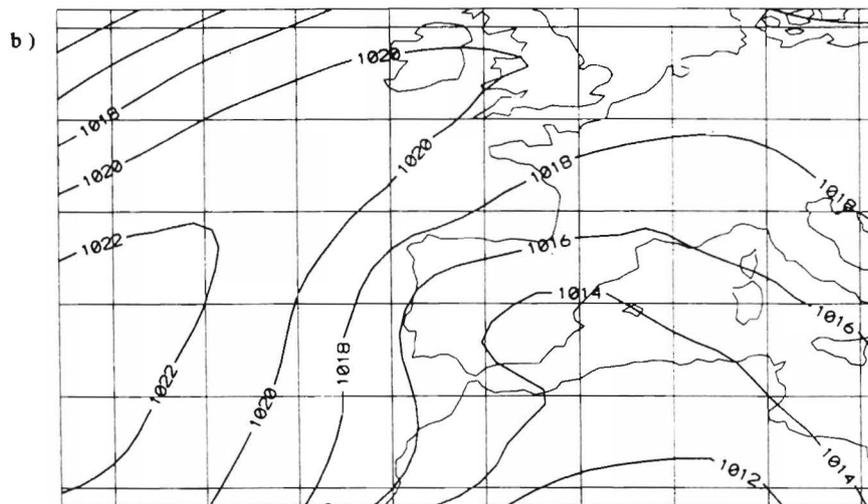


Figura 13: Tipo 7: Precipitaciones en Cataluña. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 8



PRESION A NIVEL DEL MAR



GEOPOTENCIAL 500mb

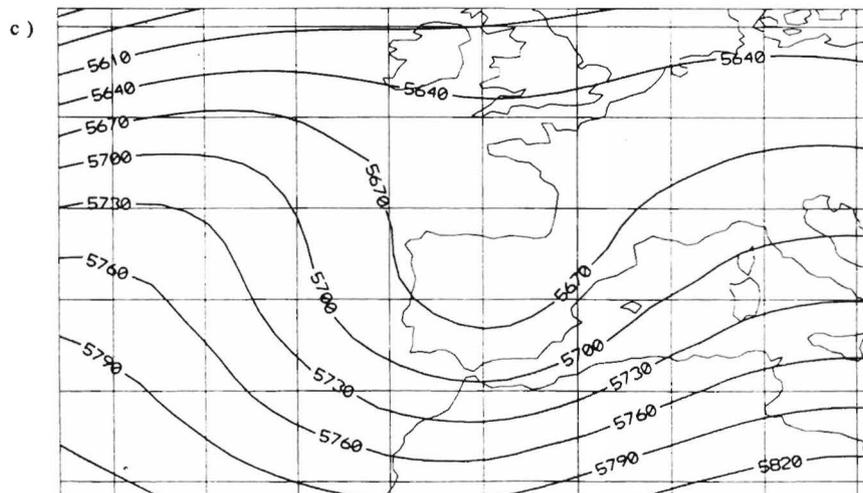


Figura 14: Tipo 8: Precipitaciones en el Bajo Valle del Ebro. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 9

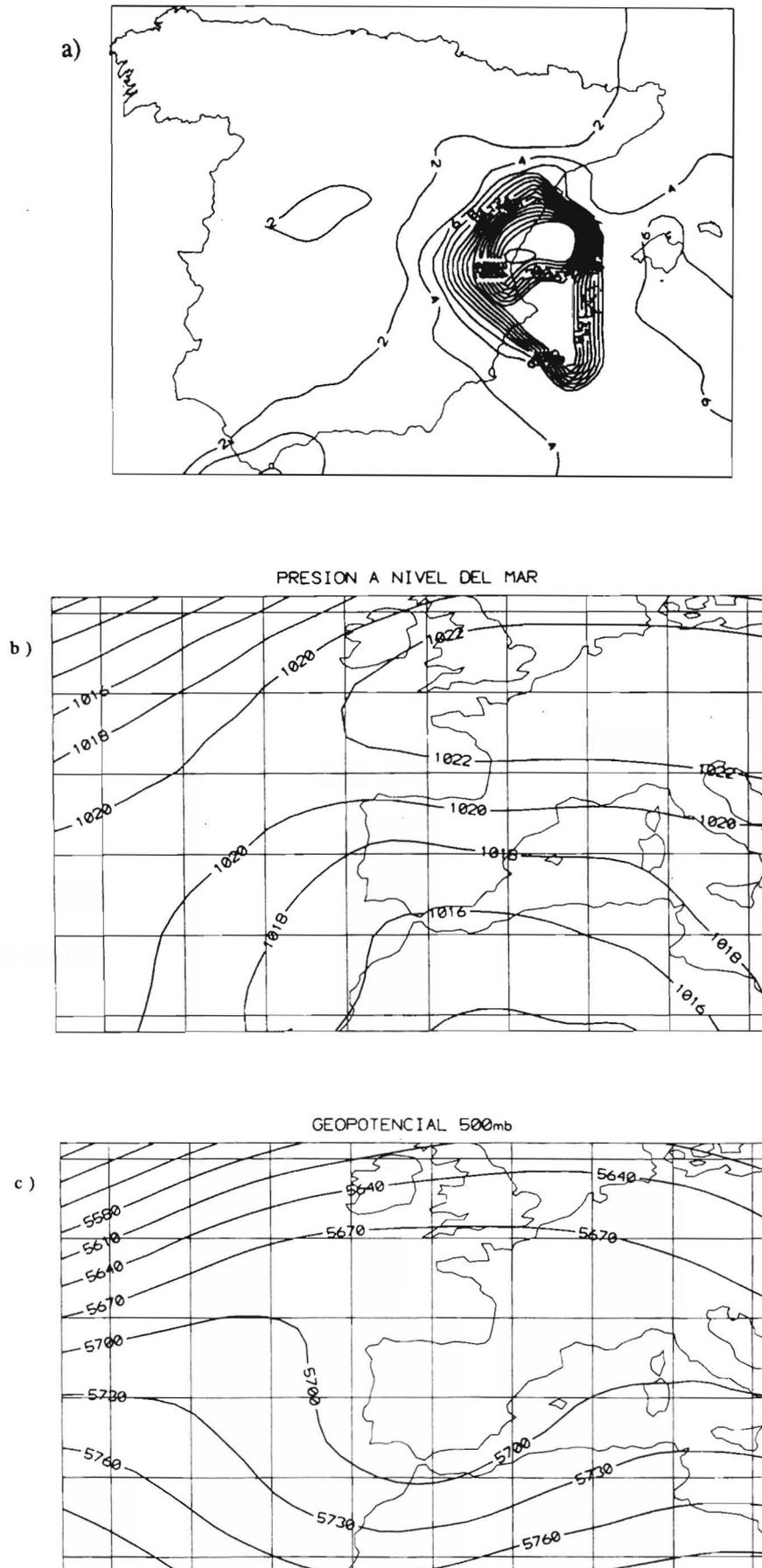


Figura 15: Tipo 9: Precipitaciones en la Comunidad Valenciana. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 10

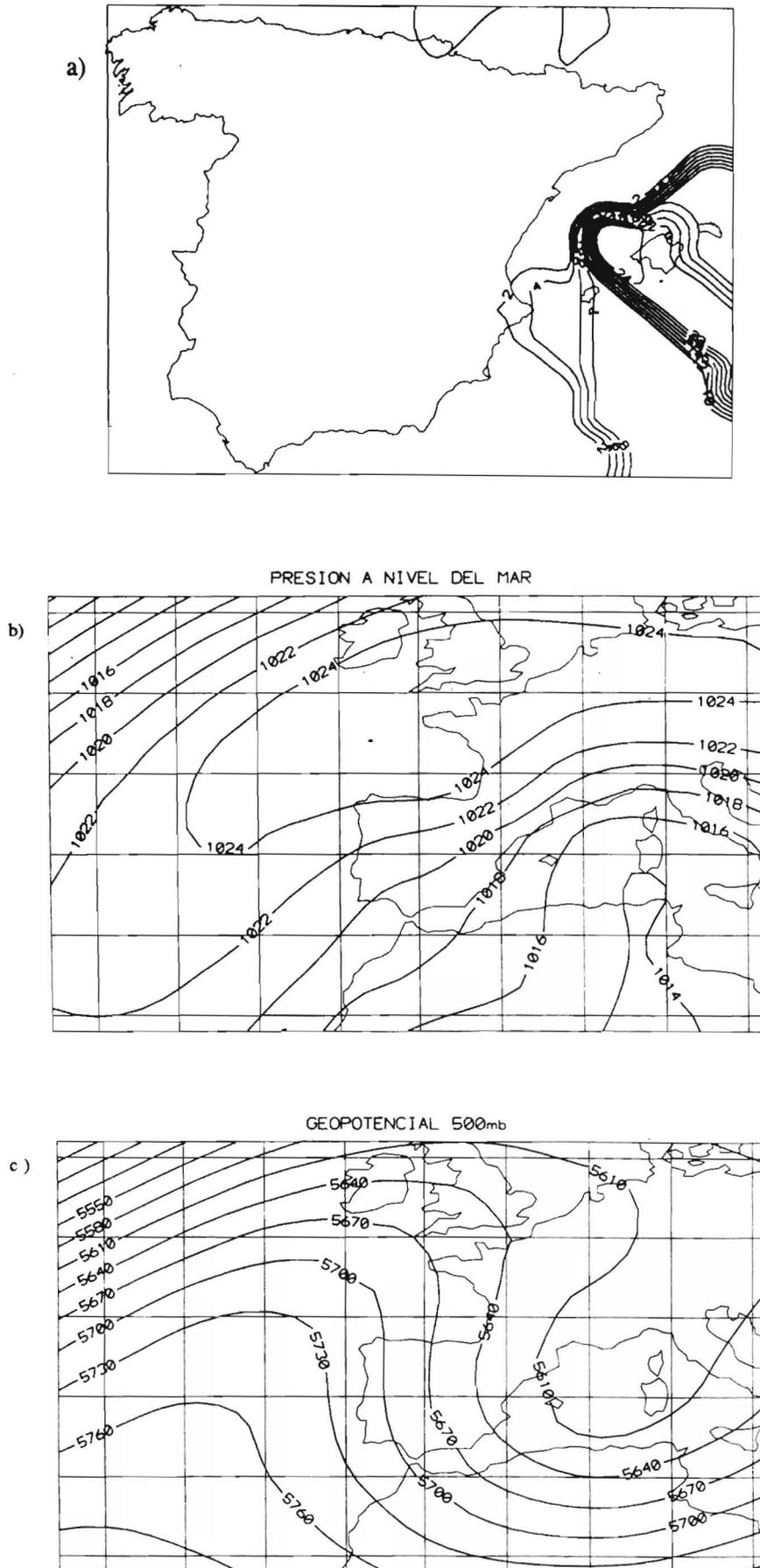
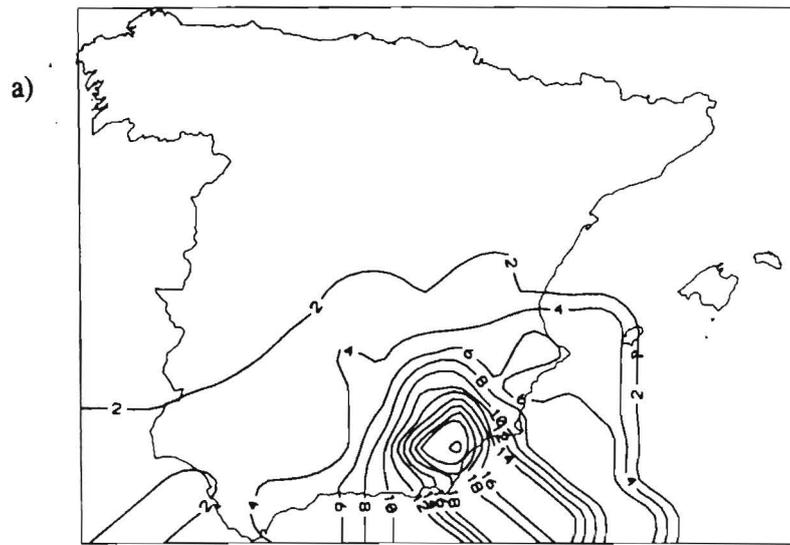
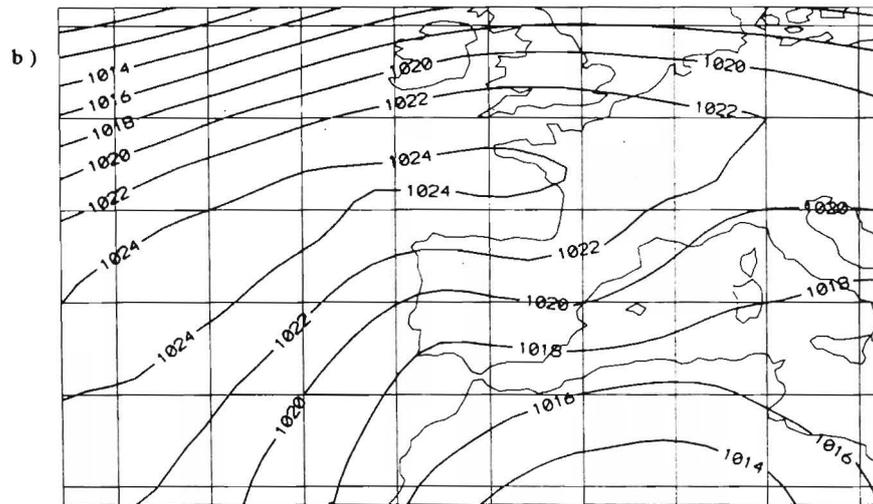


Figura 16: Tipo 10: Precipitaciones en Baleares. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 11



PRESION A NIVEL DEL MAR



GEPOTENCIAL 500mb

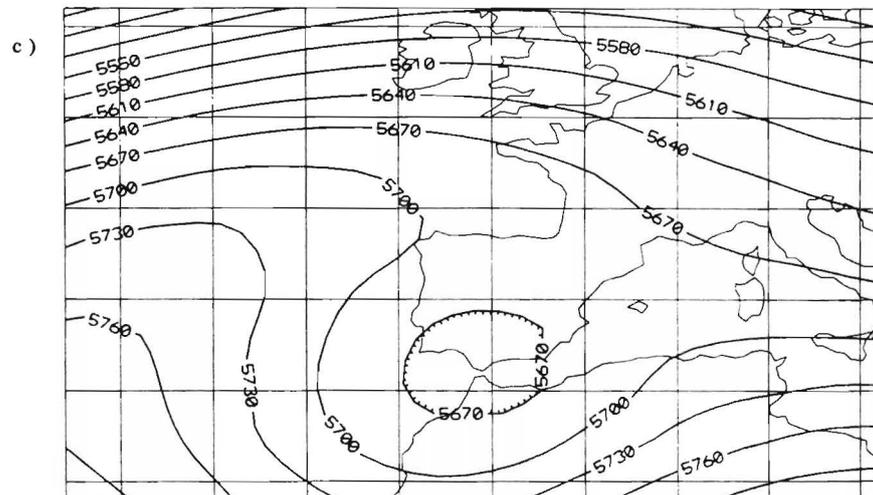


Figura 17: Tipo 11: Precipitaciones en el Sudeste Peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 12

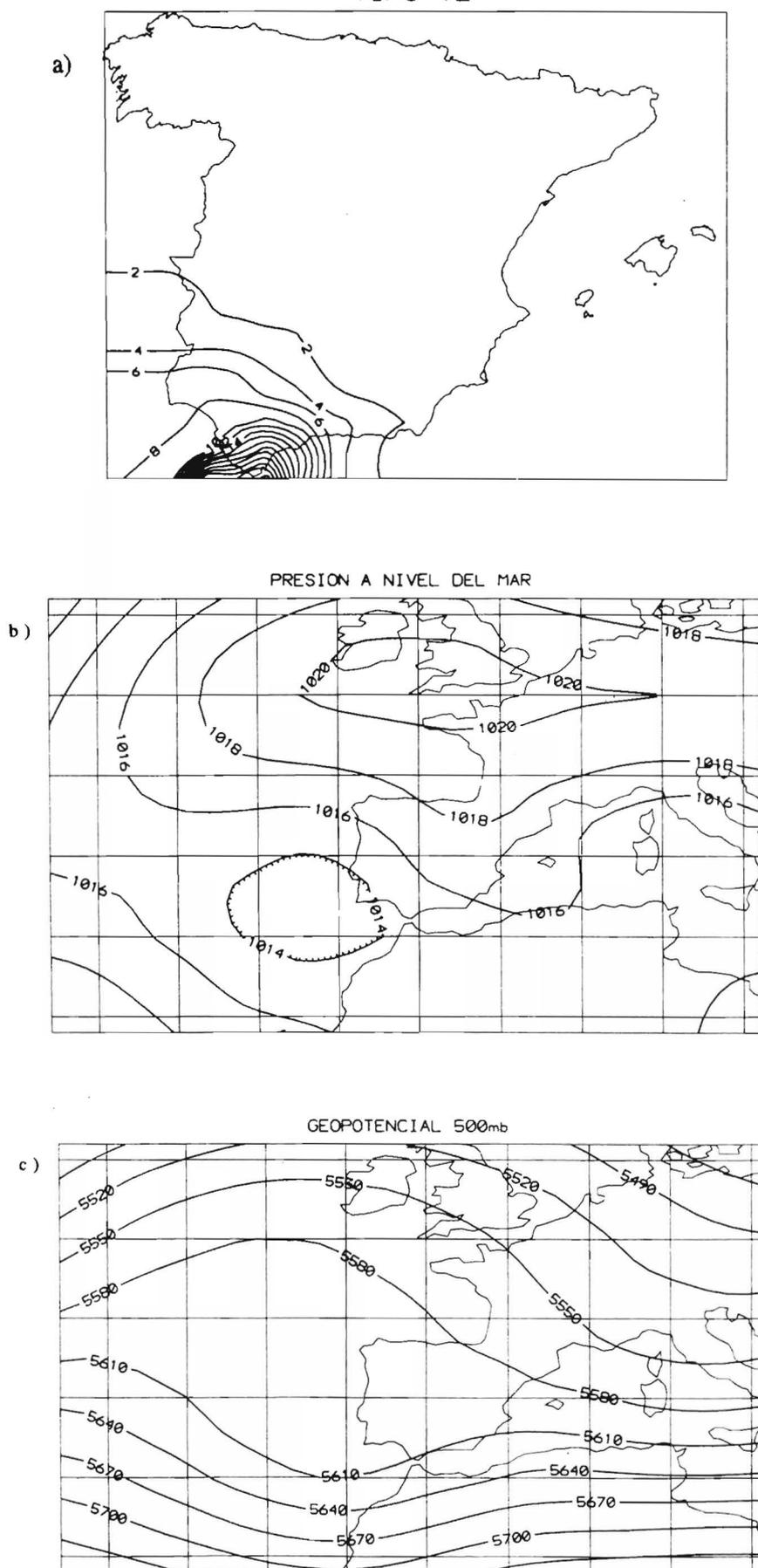
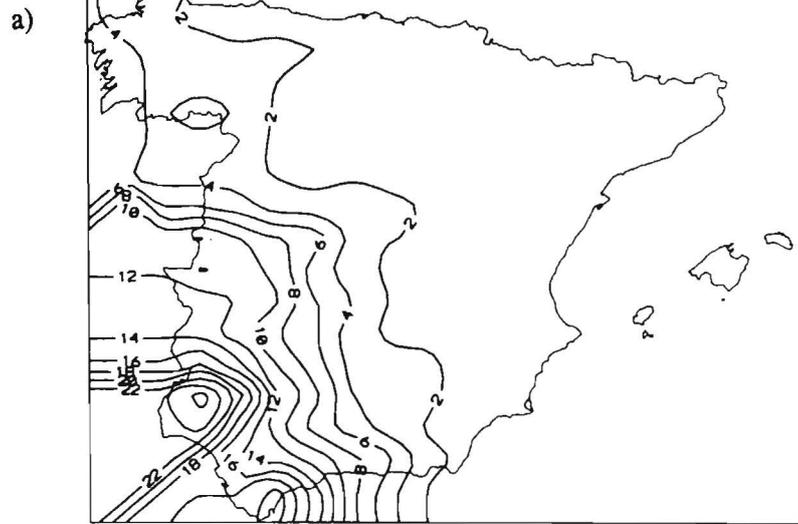
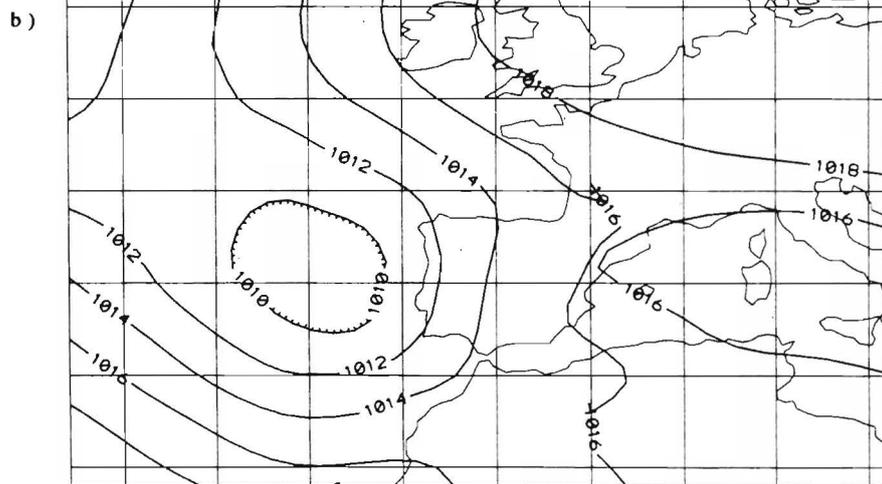


Figura 18: Tipo 12: Precipitaciones en el Estrecho y Andalucía Occidental. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 13



PRESION A NIVEL DEL MAR



GEOPOTENCIAL 500mb

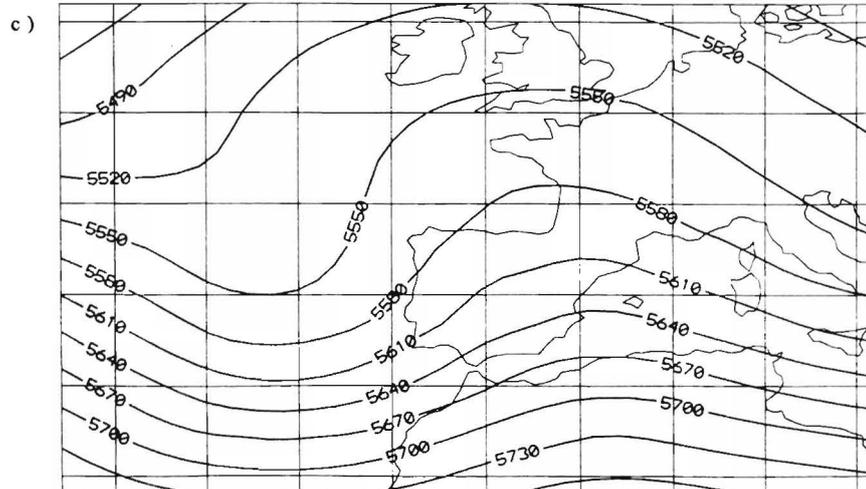
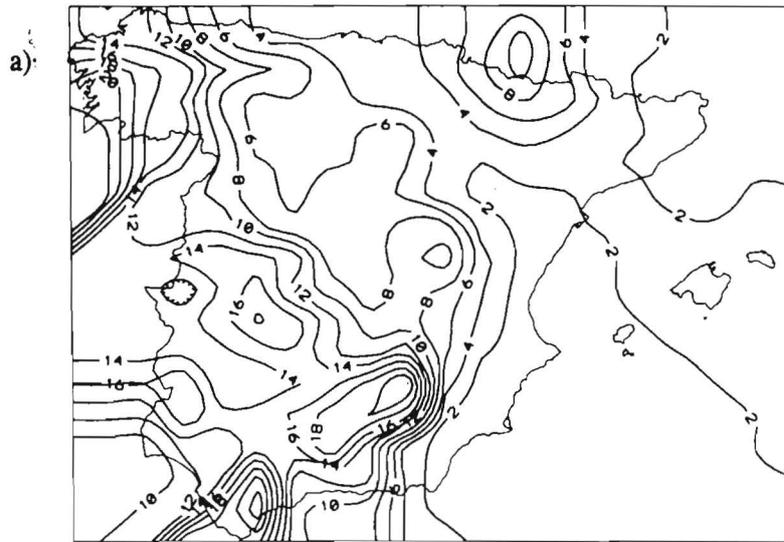
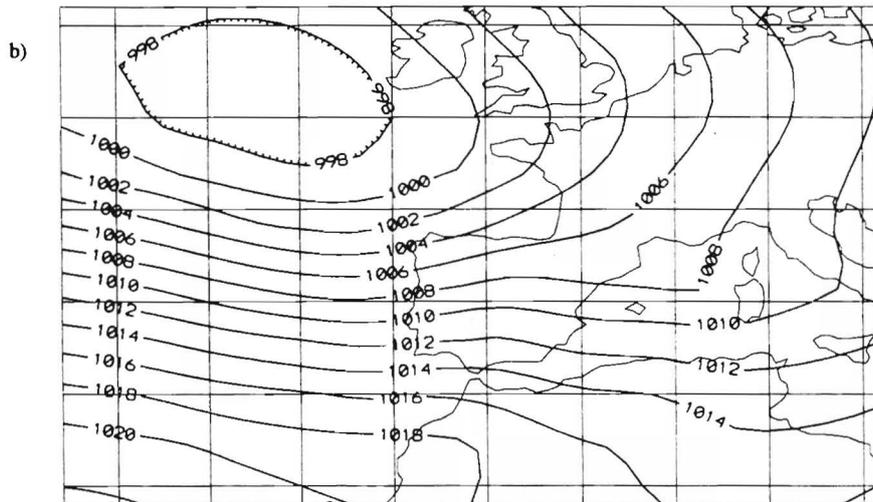


Figura 19: Tipo 13: Precipitaciones en el Sudoeste Peninsular (Ábrigo). a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 14



PRESION A NIVEL DEL MAR



GEOPOTENCIAL 500mb

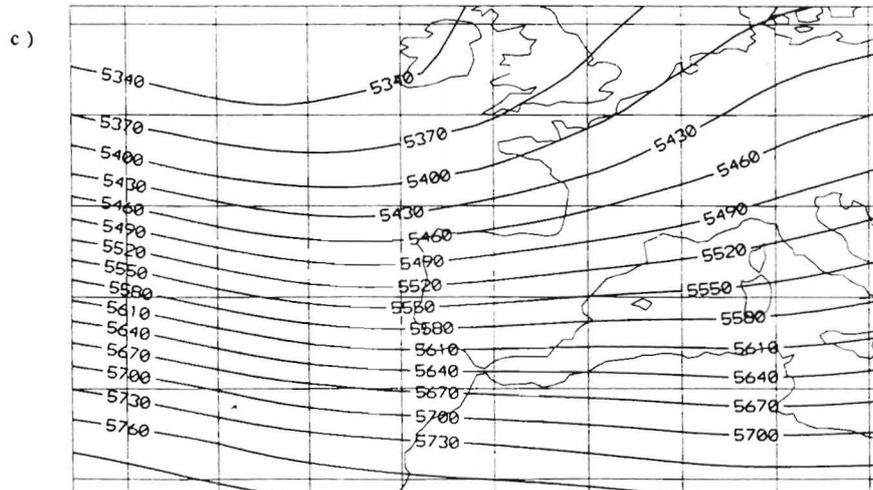


Figura 20: Tipo 14: Precipitaciones en Toda la Vertiente Atlántica y Pirineos. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 15

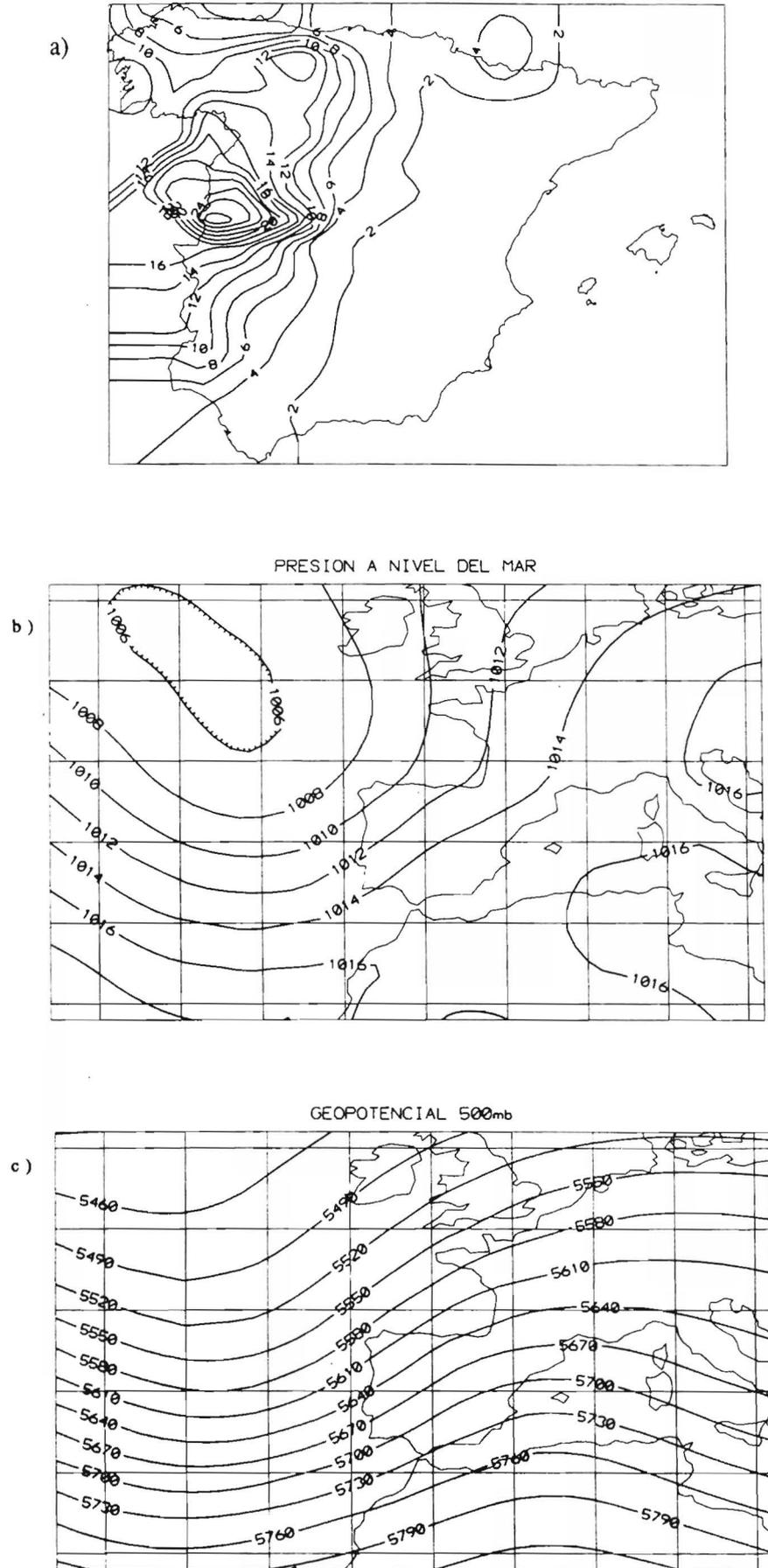
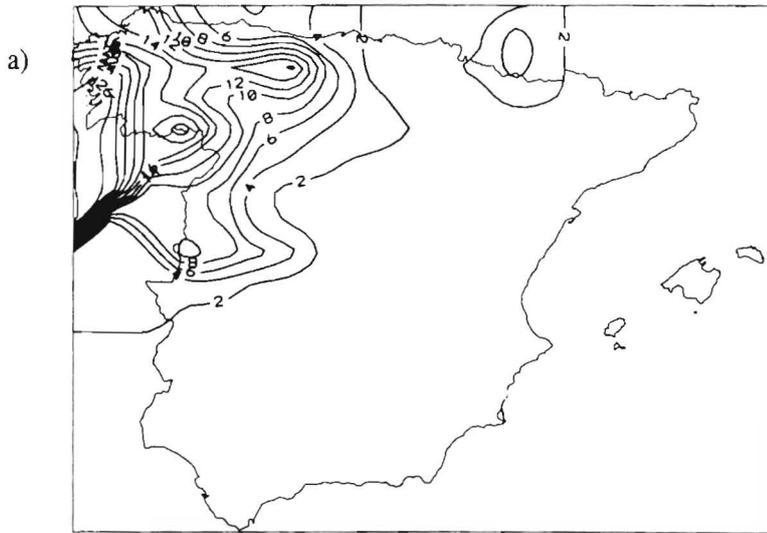
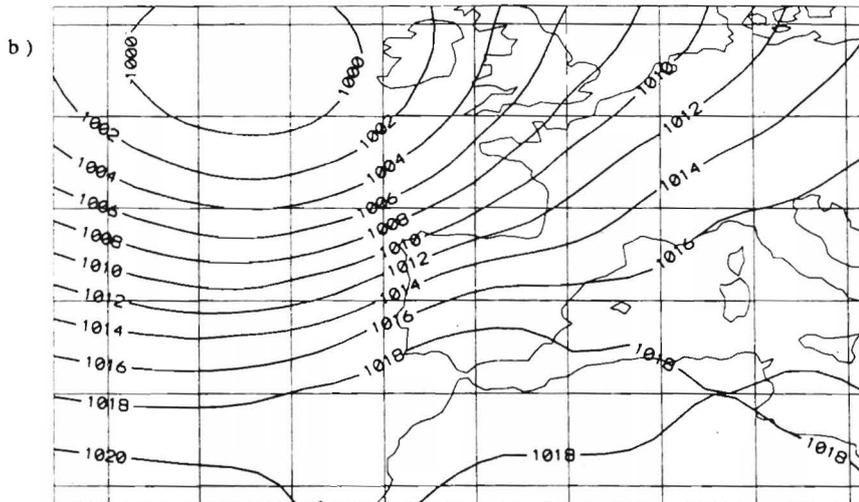


Figura 21: Tipo 15: Precipitaciones en el Oeste peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 16



PRESION A NIVEL DEL MAR



GEOPOTENCIAL 500mb

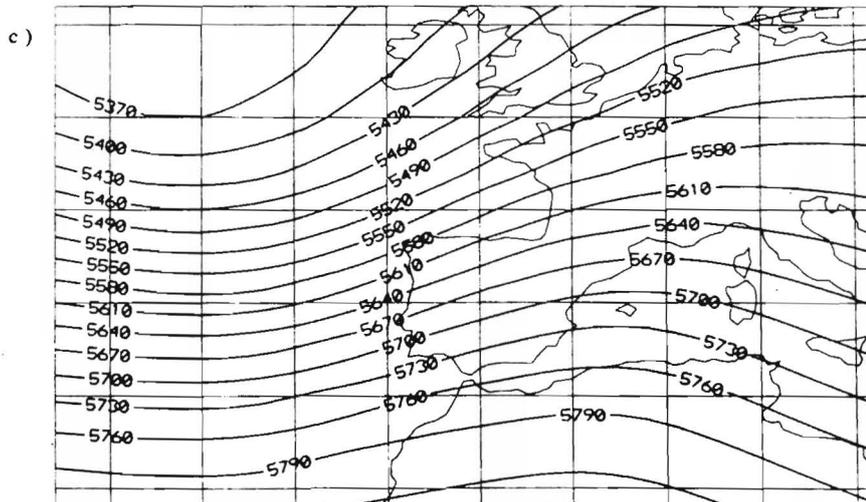


Figura 22: Tipo 16: Precipitaciones en el Noroeste peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 17

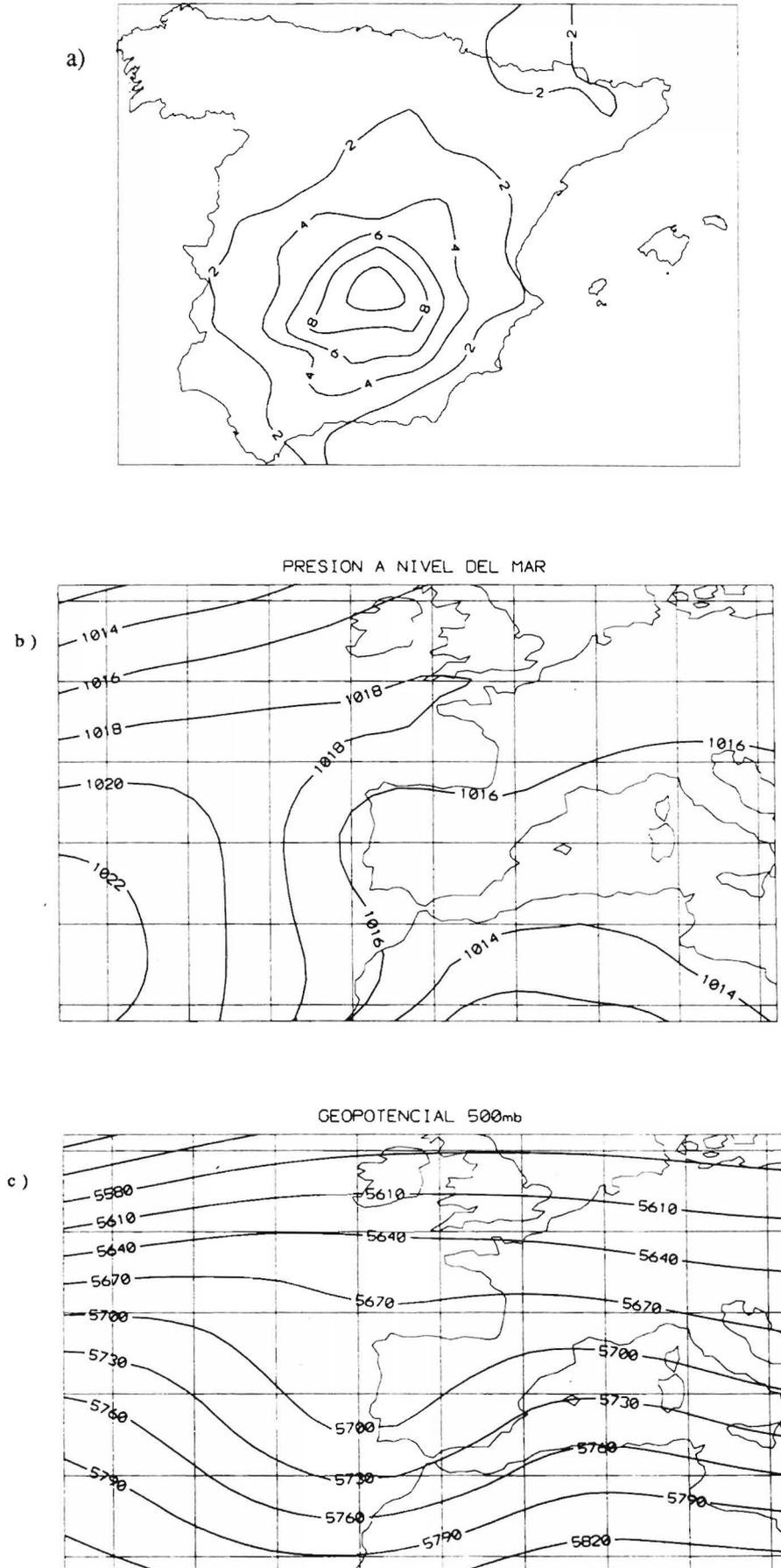


Figura 23: Tipo 17: Precipitaciones en el Interior Sur peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

TIPO 18

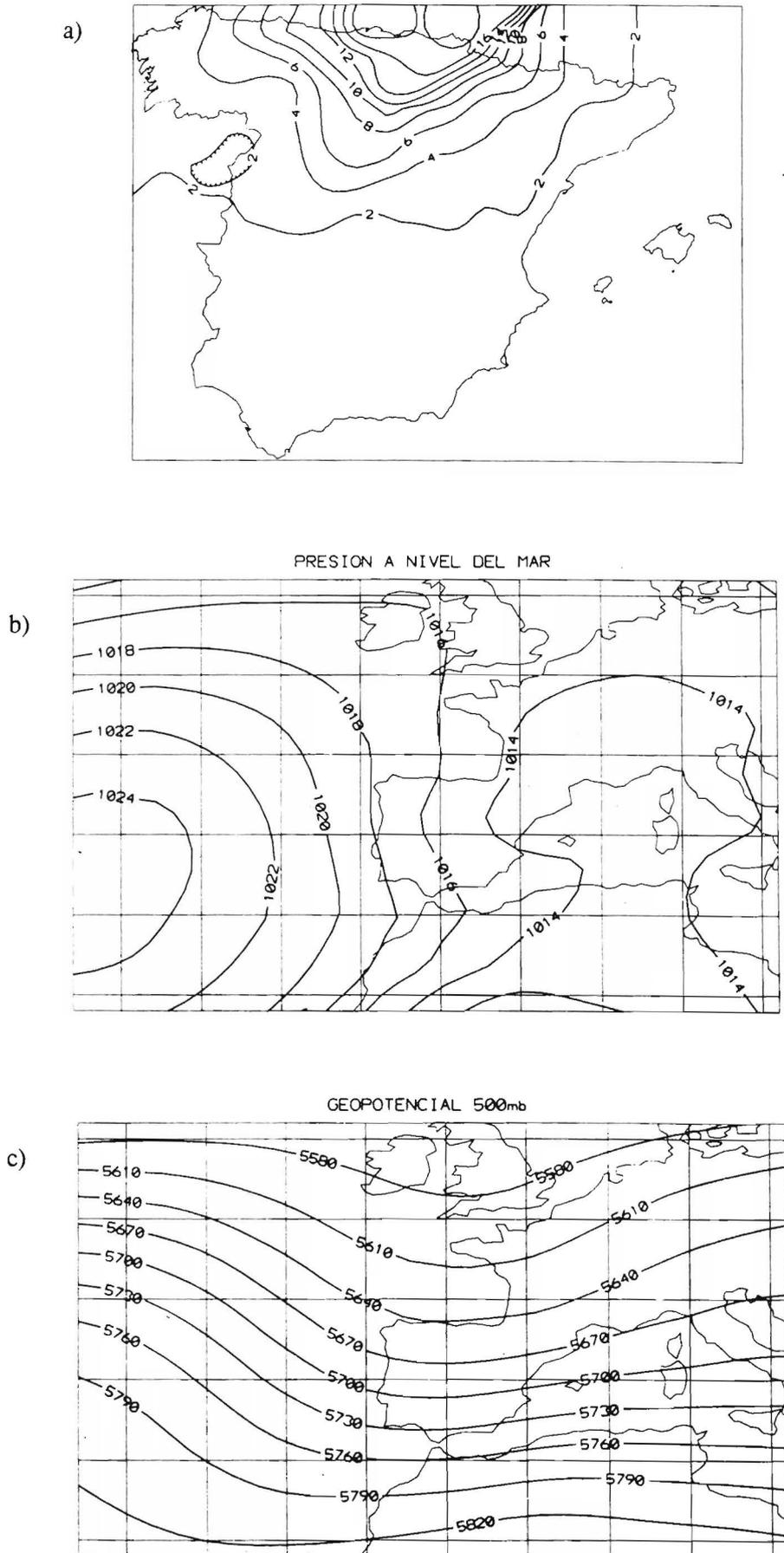


Figura 24: Tipo 18: Precipitaciones en el Norte peninsular. a) Reparto de precipitación, b) Presión a nivel del mar, c) Geopotencial de 500mb.

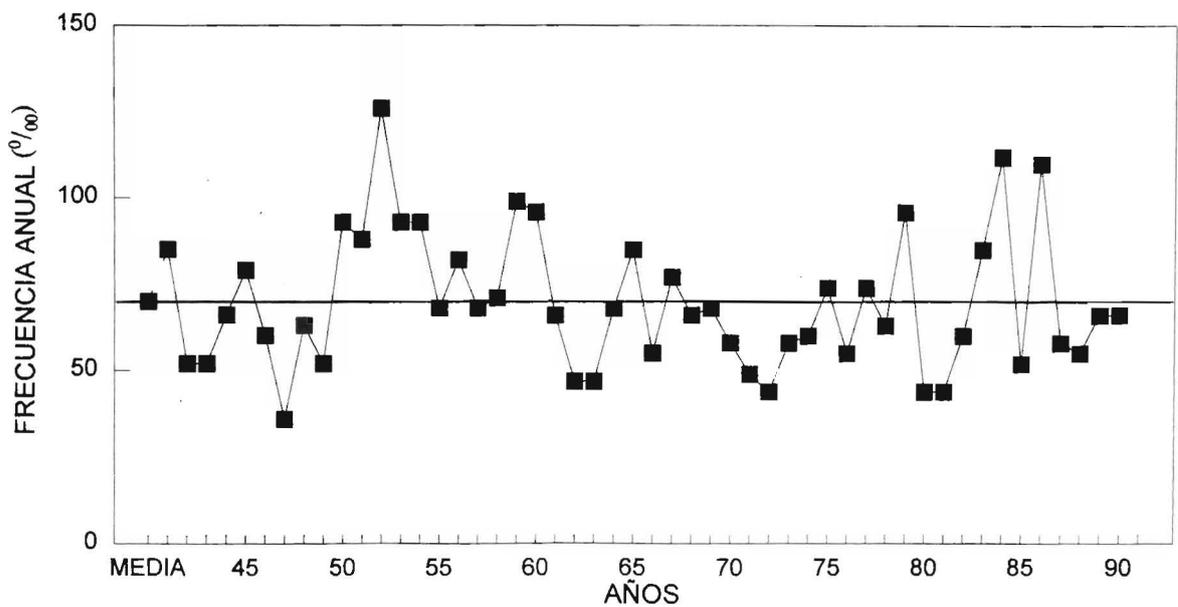
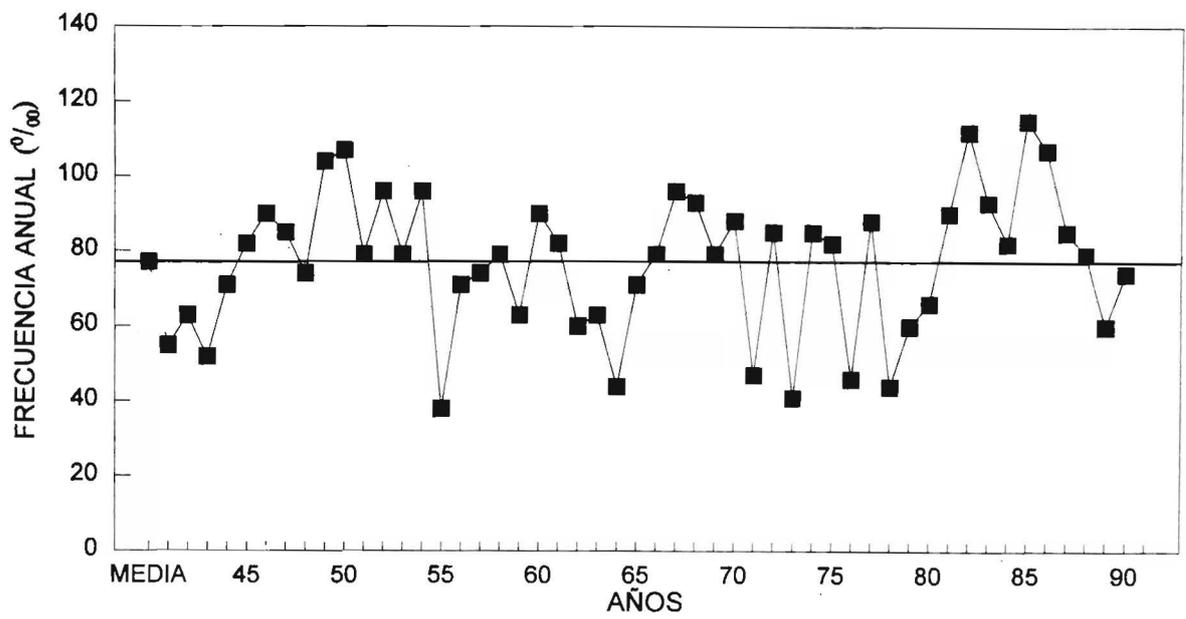
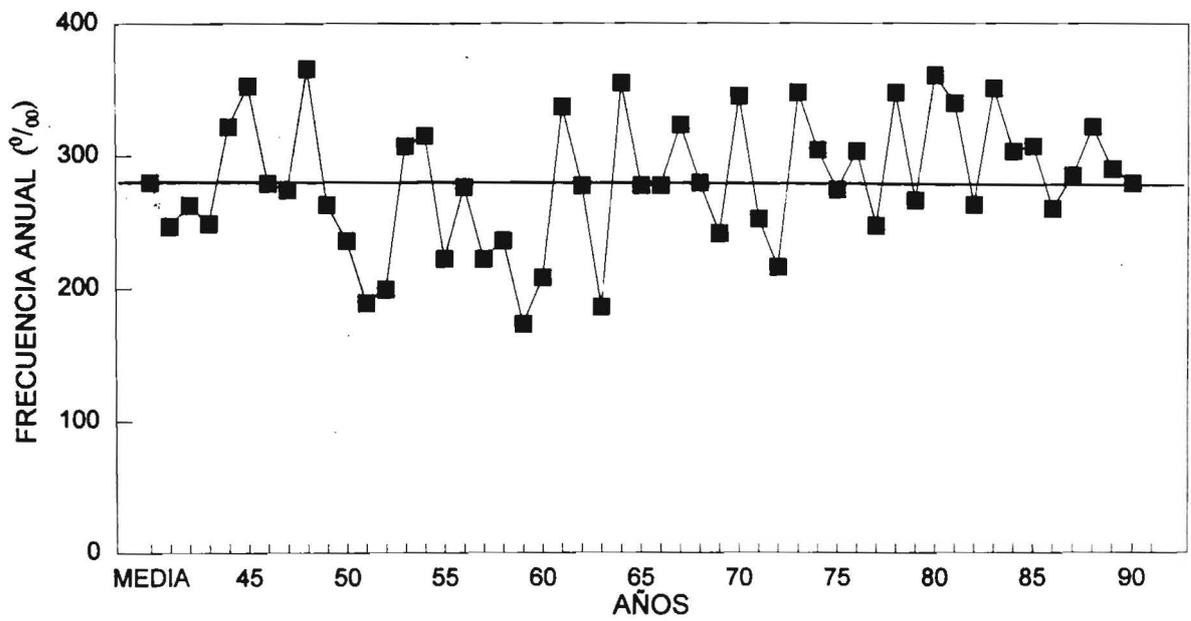


Fig. 25: Frecuencia (‰) anual y media del período 1941-1990: a) Tipo 1, b) Tipo 2, c) Tipo 3.

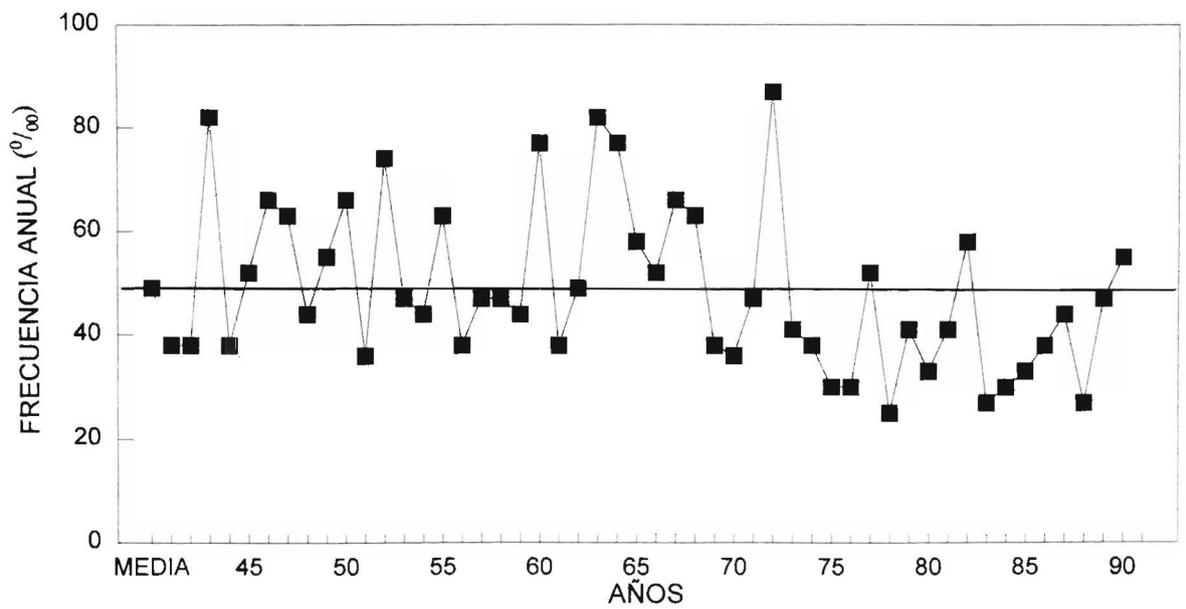
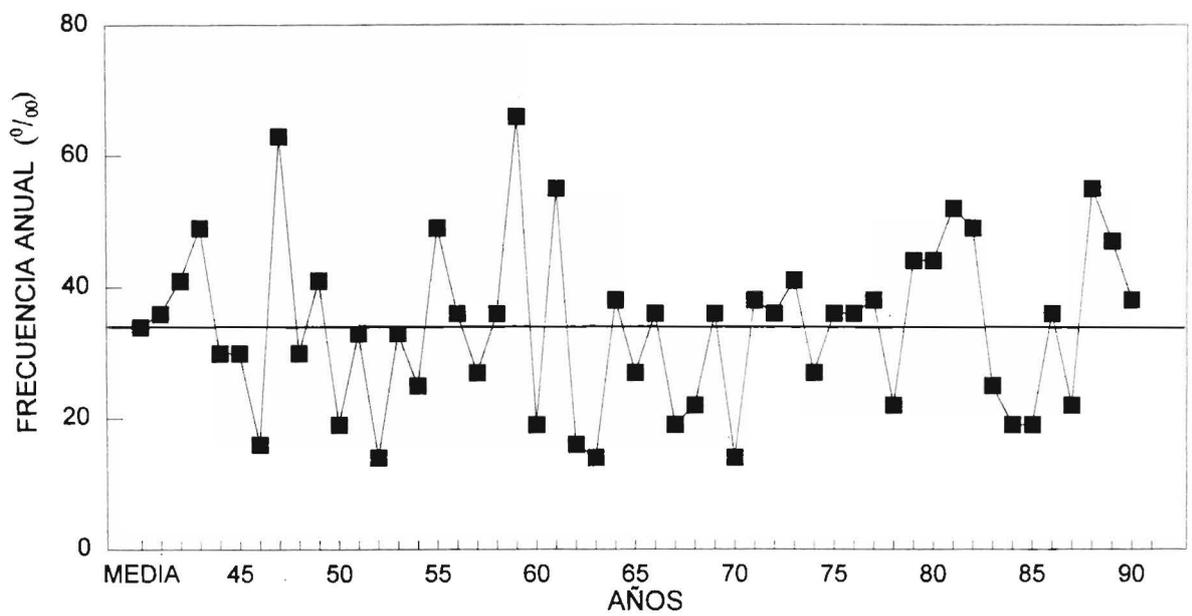
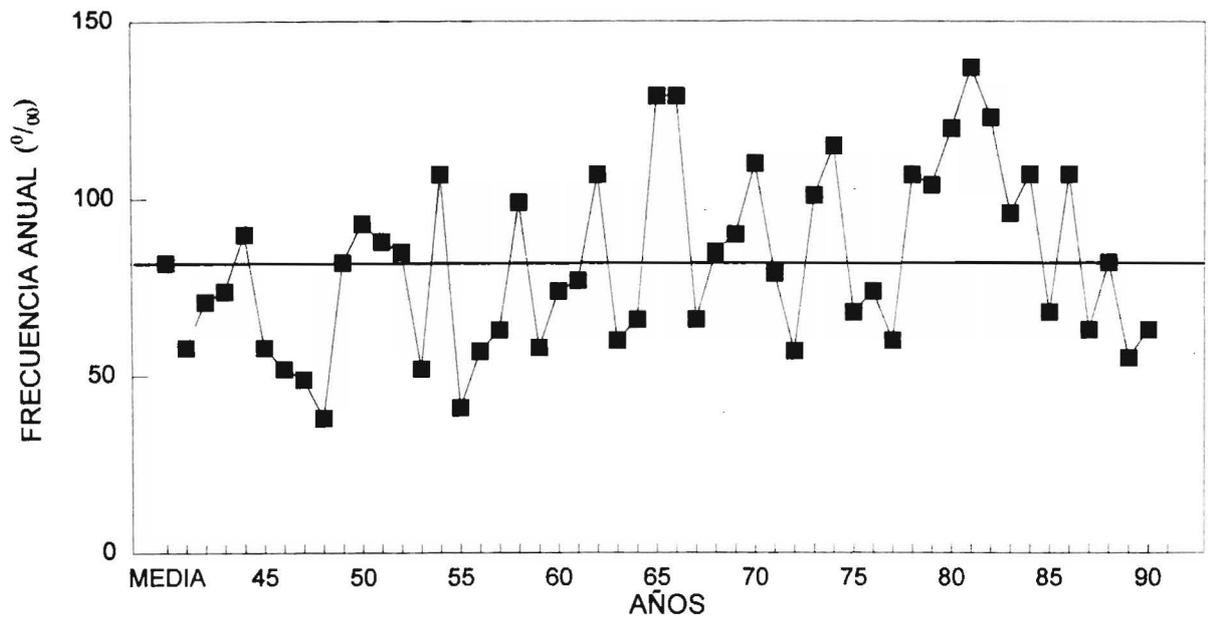


Fig. 26: Frecuencia (‰) anual y media del período 1941-1990: a) Tipo 4, b) Tipo 5, c) Tipo 6.

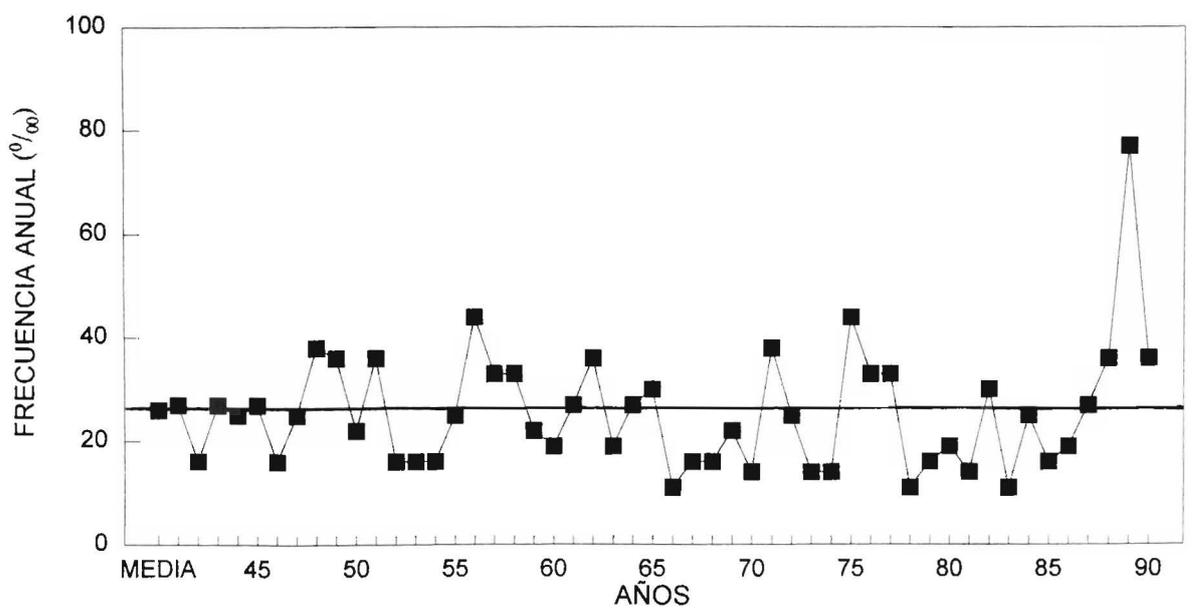
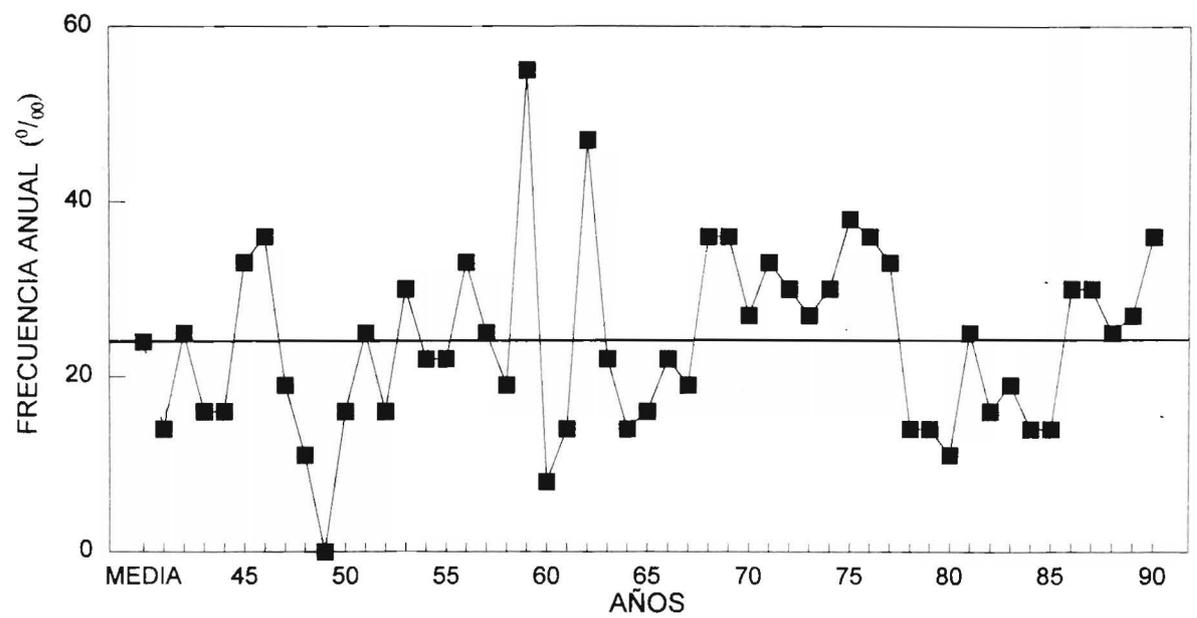
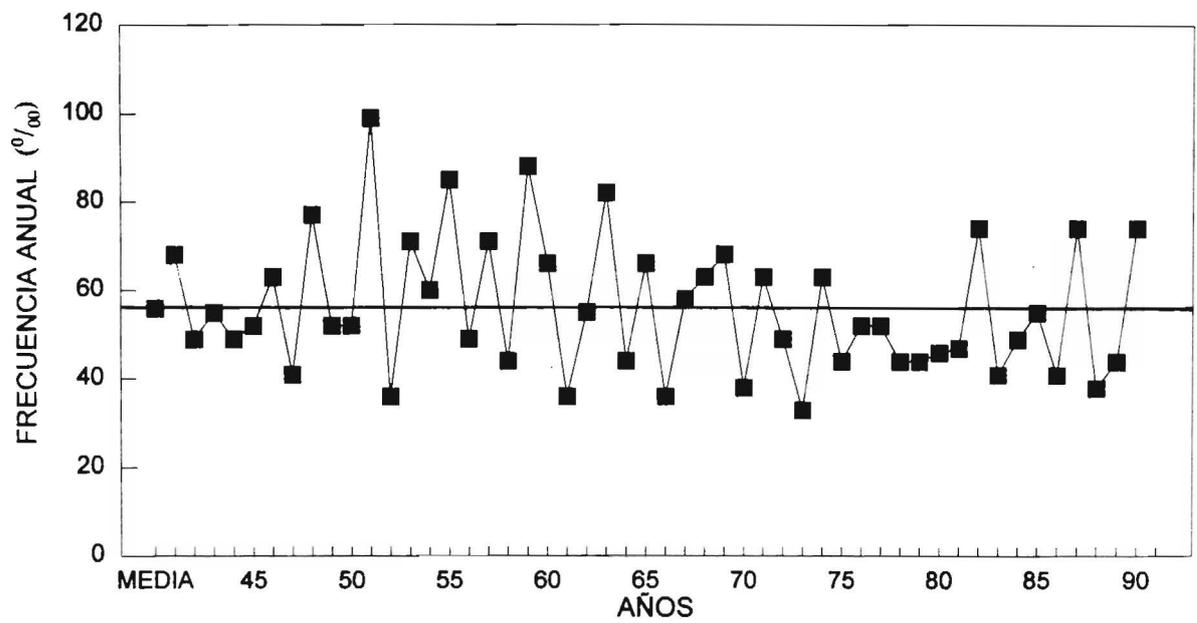


Fig. 27: Frecuencia (‰) anual y media del período 1941-1990: a) Tipo 7, b) Tipo 8, c) Tipo 9.

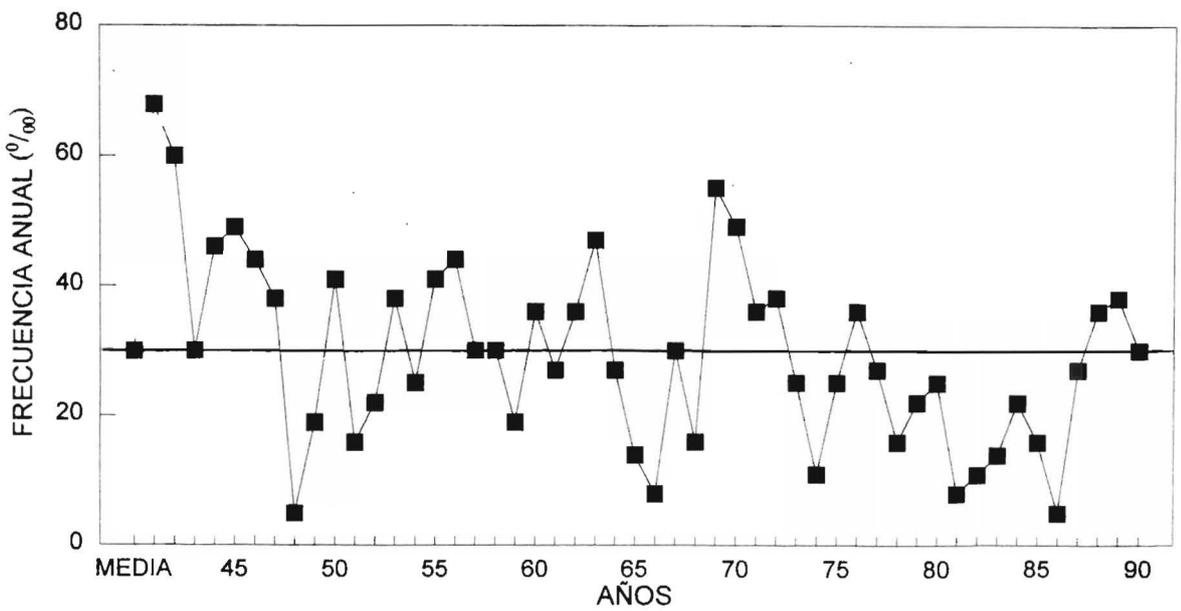
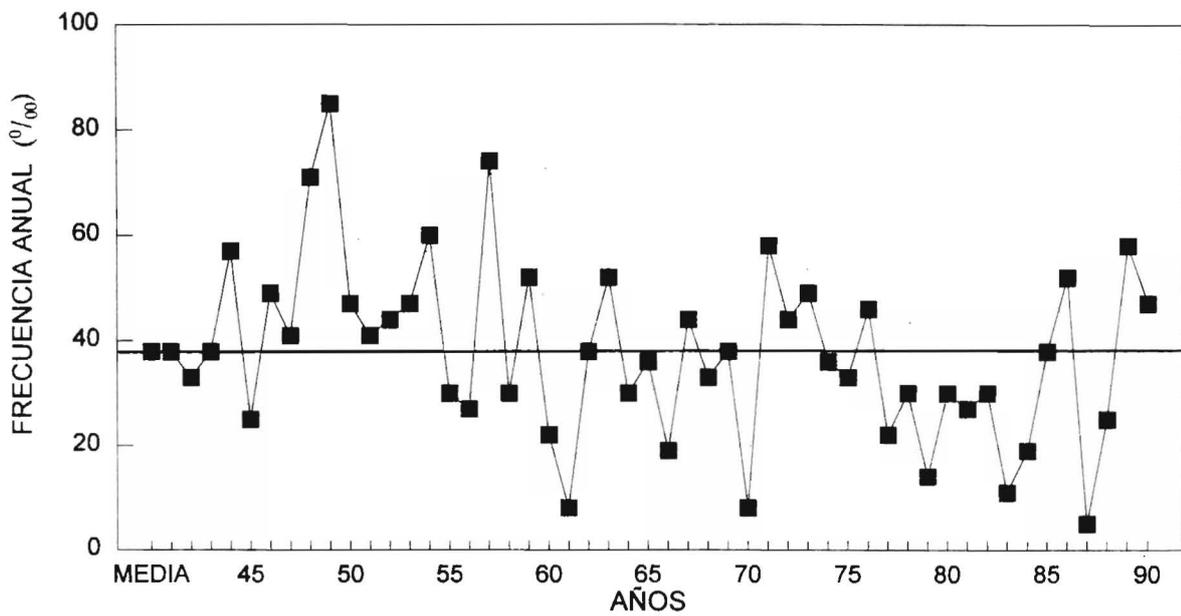
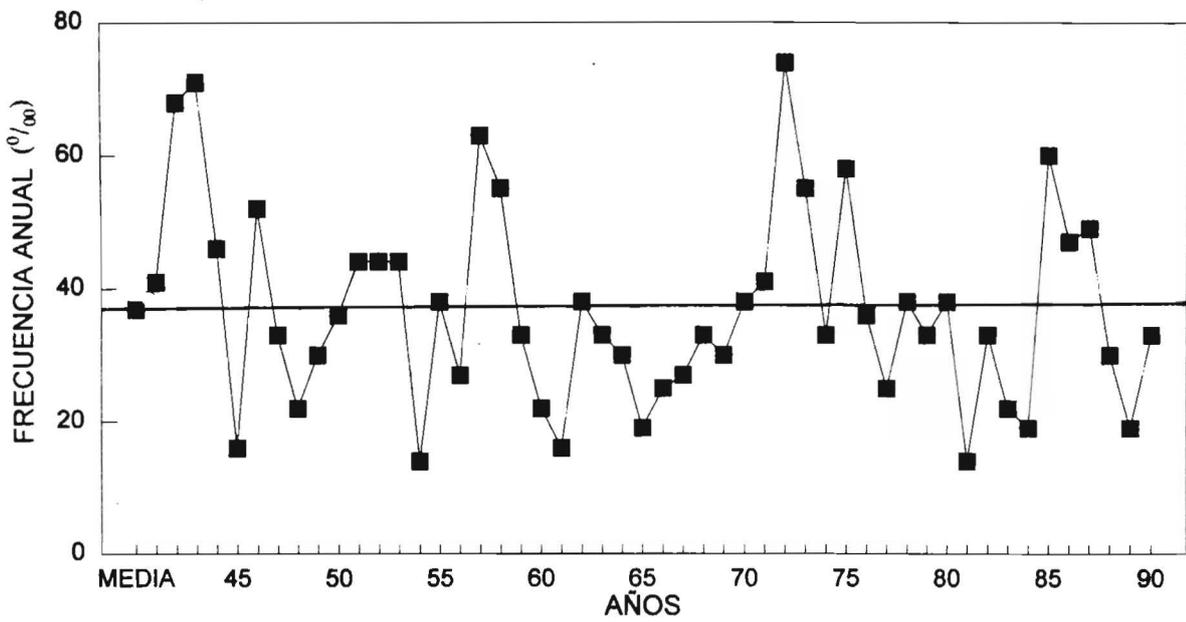


Fig. 28: Frecuencia (‰) anual y media del período 1941-90: a) Tipo 10, b) Tipo 11, c) Tipo 12.

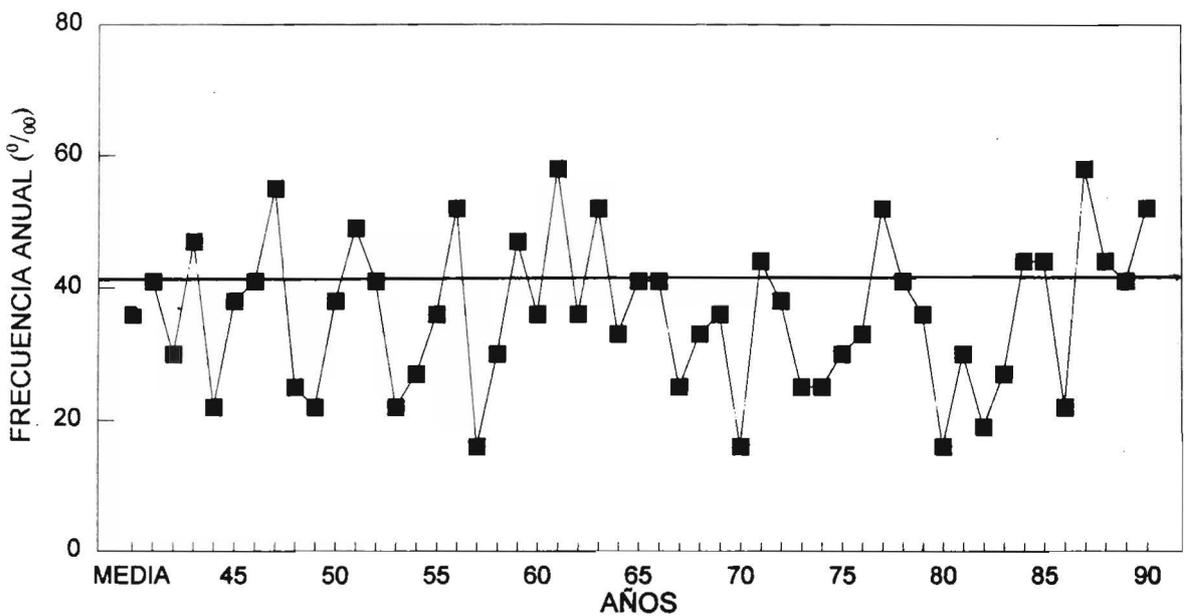
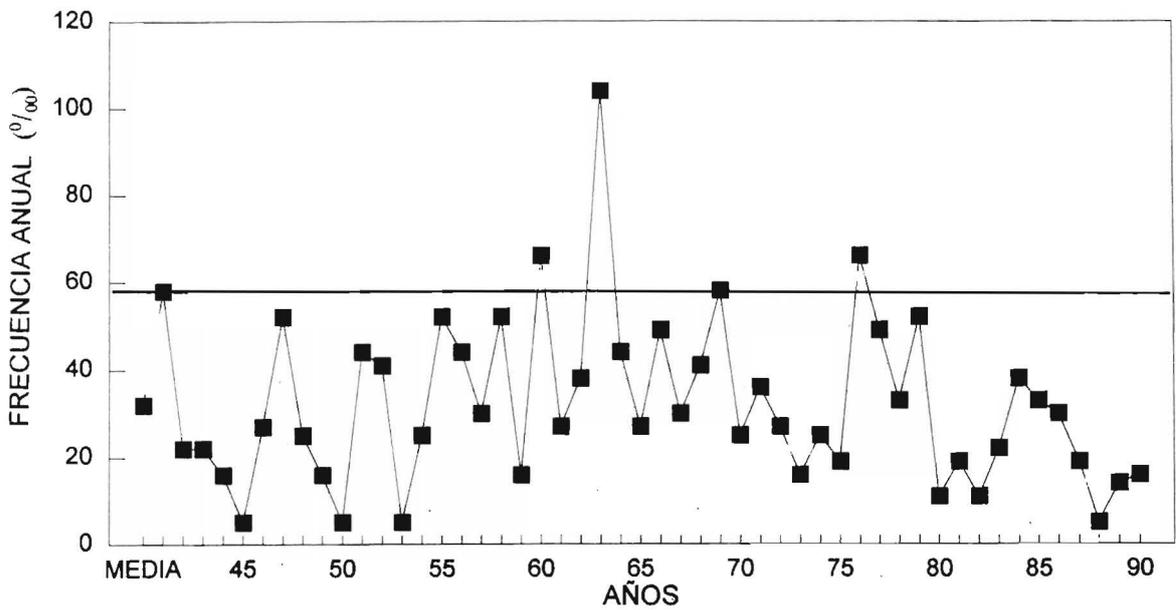
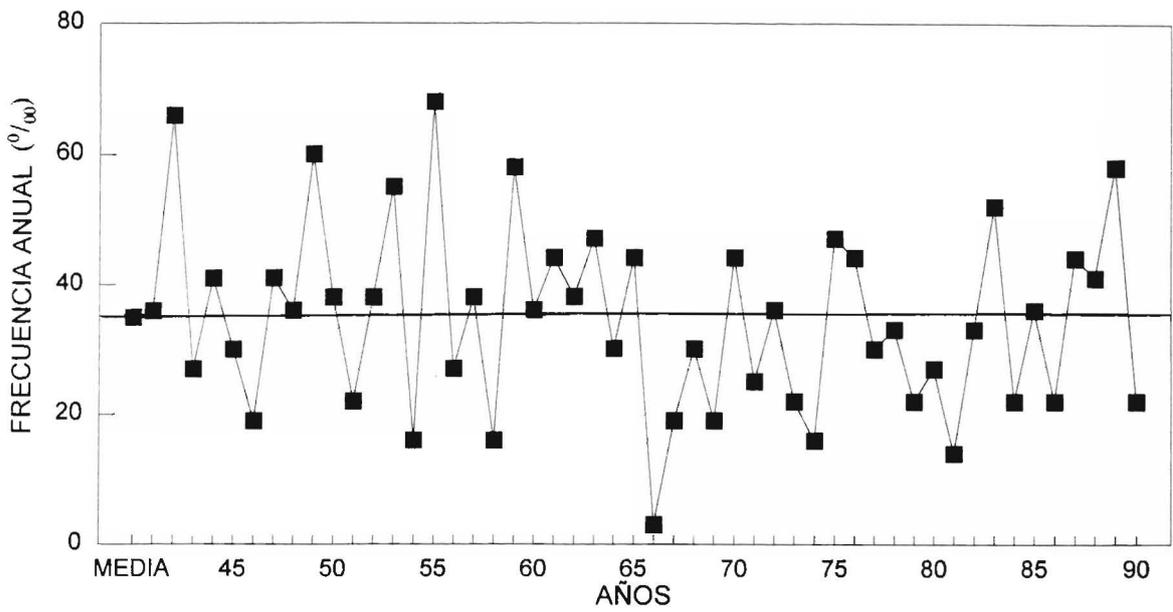


Fig. 29: Frecuencia (‰) anual y media del período 1941-90: a) Tipo 13, b) Tipo 14, c) Tipo 15.

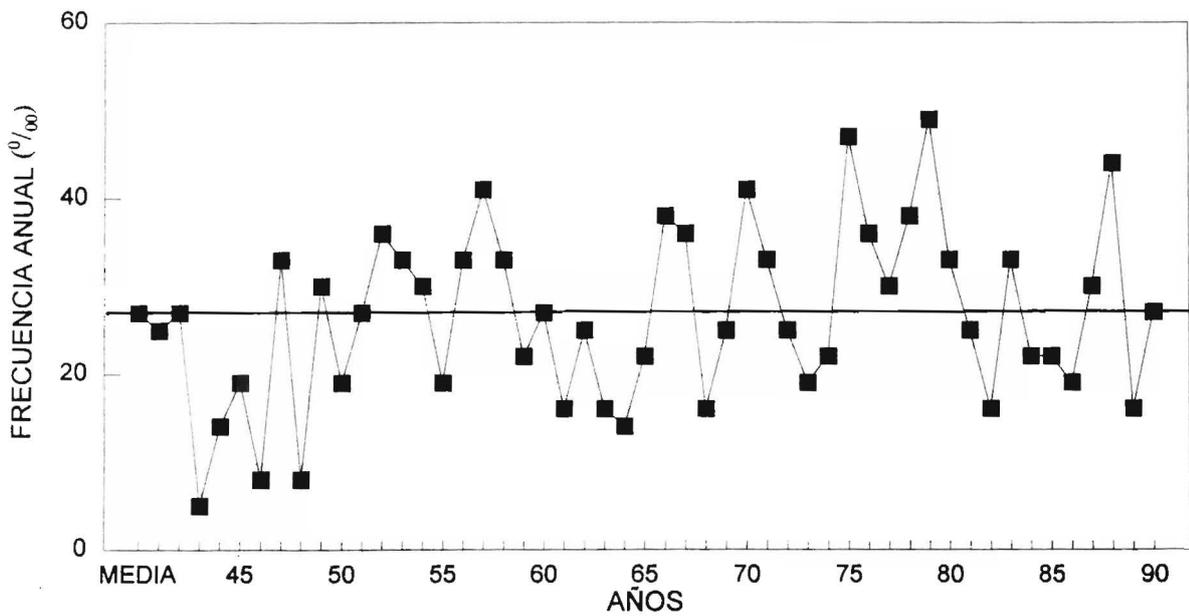
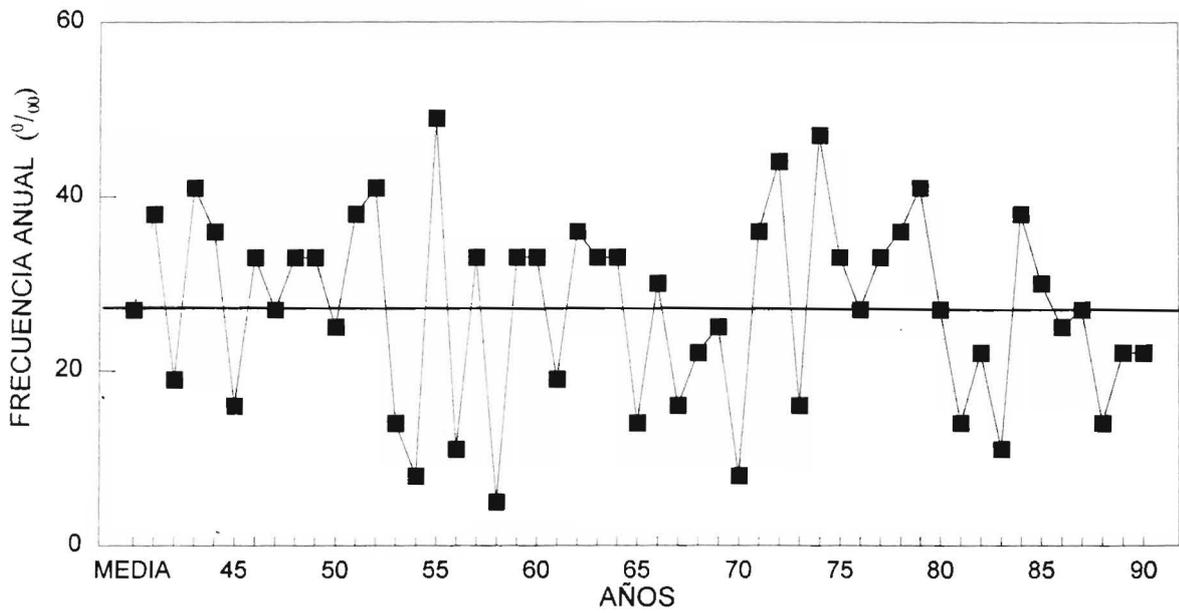
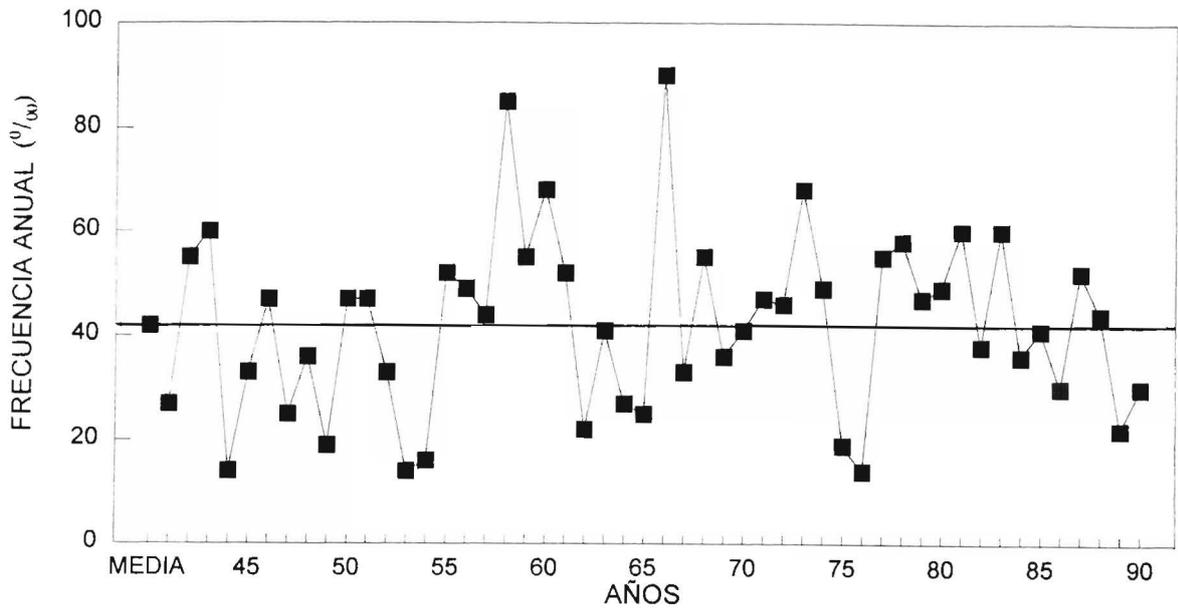


Fig. 30: Frecuencia (‰) anual y media del período 1941-90: a) Tipo 16, b) Tipo 17, c) Tipo 18.

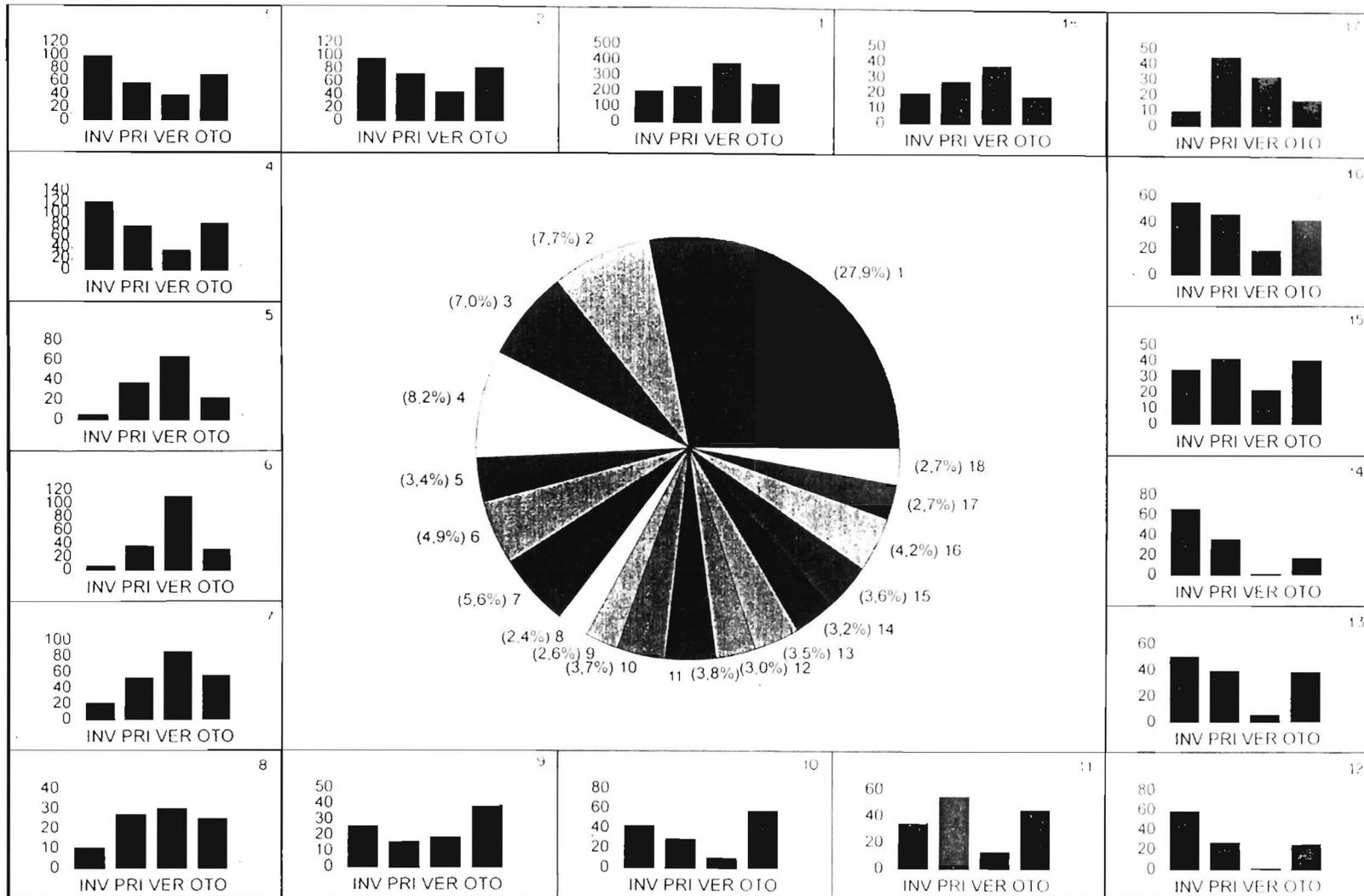


Figura 31: Diagrama de sectores: Frecuencia (%) media de cada Tipo para el período 1941-1990.  
 Diagramas de barras: Frecuencia (%) estacional de cada Tipo, del mismo período.

