



Classificació de tipus de circulació atmosfèrica: proposta metodològica i aplicacions

Pere Esteban Veà

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Departament de Geografia Física i Anàlisi Geogràfica Regional

Facultat de Geografia i Història

UNIVERSITAT DE BARCELONA

TESI DOCTORAL

**CLASSIFICACIÓ DE TIPUS DE CIRCULACIÓ ATMOSFÈRICA:
PROPOSTA METODOLÒGICA I APLICACIONS**

Memòria presentada per

PERE ESTEBAN VEA

Per optar al grau de Doctor en Geografia

Programa de Doctorat en Geografia Física

Bienni 1999-2001

Director de tesi

DR. JAVIER MARTÍN VIDE

Catedràtic de Geografia Física

UNIVERSITAT DE BARCELONA

Barcelona, març de 2012

Dedicada a tots aquells que m'estimo,
especialment els meus pares.

AGRAÏMENTS

Primer de tot voldria agrair als meus pares, Maricel i Pere, el seu infinit i valuós suport incondicional. Us mereixeu el millor, espero que aquest granet de sorra ajudi a engrandir la vostra felicitat. També a la meva germana Estela, referent de capacitat de superació davant les dificultats, així com a la Gemma per animar-me sempre a tirar endavant aquest projecte i compartir amb mi, sempre amb l'interruptor positiu i total generositat, els encerts i errors del dia a dia. Tenir-vos a prop és el millor que em podia passar a la vida.

També el meu agraïment als amics, que sempre heu sabut donar consells oportuns quan els necessitava, tant si les coses van bé com malament. Moltes gràcies a l'Edu, l'Aniol, l'Olga, l'Eli, el Marc i a tota la colla de la UEC i l'AEC i la colla de la Casa Roja.

Rosanna Fatsini, Núria Espelta i Aniol Rafel, moltíssimes gràcies per la excel·lent revisió del document.

Des de la perspectiva científica, el meu agraïment a aquells que heu col·laborat en els treballs científics aquí presentats o que m'heu acollit durant les estades a l'estranger: Phillip D. Jones (CRU), Claire Goodess (CRU), Cary Mock (USC), Karl Birkeland (MSU), Katty Hansen (MSU), Montserrat Mases, Miquel Ninyerola (UAB), Marc Prohom (SMC), Nicolau Pineda (SMC), Laura Trapero (CENMA), Xavier Soler (SMC) Christoph Beck (UniA), Andreas Philipp (UniA) i Stefanie Seubert (UniA).

Vull també fer palès el luxe que ha estat compartir cinc anys del projecte COST 733 amb tots els seus participants, i especialment amb els abans esmentats Christoph Beck i Andreas Philipp (UniA), Mathias Demuzere (KULeuven), Radan Huth (IAP), Massimiliano Pasqui (IBIMET), Christel Prudhome (CEH), Ole Einar Tveito (NMI), Domingo Rasilla (UNICAN) i Maria Asunción Pastor i Maria Jesús Casado (AEMET). Gràcies pels vostres valuosos consells i l'excel·lent tracte. L'aprenentatge en aquests anys ha estat molt gran, especialment quant a la professionalitat científica, l'entusiasme per a la recerca, la capacitat (sovint ingent) de treball i el positivisme.

També vull agrair a la Generalitat de Catalunya i a la Universitat de Barcelona haver-me concedit la beca predoctoral 256FI durant el 2000-2003 i que m'ha permès formar-me com a científic i de gaudir d'estades molt valuoses a centres estrangers.

Tot aquest treball no hauria estat possible sense la confiança i la llibertat que el professor Javier Martín-Vide ha dipositat en mi. A aquest li agraeixo la predisposició i l'ajuda d'aquests anys. Sense la seva col·laboració, consells, paciència i recolzament no hauria adquirit una visió tan ampla de la Climatologia actual, ni hauria gaudit del projecte COST733, ni hauria completat aquest treball. Gràcies Javier per haver-me donat la oportunitat de pertanyer al Grup de Climatologia de la Universitat de Barcelona. Al seus membres també el meu agraïment, i especialment al Marc Prohom i a en David Montserrat, amb qui vaig compartir despatx durant els anys del Parc Científic i on es van donar moments impossibles d'esborrar de la memòria. També, al departament de Geografia Física i Anàlisi Geogràfica Regional, gràcies especialment al Lothar Schulte i al Xavi Úbeda pel seu suport, ajuda i cordialitat.

Finalment, dono gràcies també als meus companys de feina de l'Institut d'Estudis Andorrans perquè m'han donat ànims i m'han facilitat el camí per culminar el projecte. Potser deixaré de parlar de classificacions, però no us escapareu del *jet-stream*, les inversions de subsidència, les sèries homogeneïtzades i del canvi climàtic. Endavant amb la bona feina que s'està fent des del CENMA!

TAULA DE CONTINGUTS

AGRAÏMENTS	5
TAULA DE CONTINGUTS	7
TAULA DE FIGURES	9
LLISTAT D'ACRÒNIMS	11
PREFACI.....	13
1. INTRODUCCIÓ: CONCEPTES EN CLIMATOLOGIA SINÒPTICA.....	15
2. MÈTODES DE CLASSIFICACIÓ DELS TIPUS DE CIRCULACIÓ ATMOSFÈRICA (CTC).....	27
2.1. MÈTODES MANUALS O SUBJECTIUS.....	28
2.1.1. CLASSIFICACIÓ DE LAMB [1972]	29
2.1.2. CLASSIFICACIÓ DE HESS-BREZOWSKY GROSSWETTERLAGEN.....	30
2.1.3. CLASSIFICACIÓ DE PÉDELABORDE.....	31
2.1.4. CLASSIFICACIONS A LA PENÍNSULA IBÈRICA	32
2.1.5. CLASSIFICACIONS CENTRADES ALS PIRINEUS I ANDORRA.....	33
2.2. MÈTODES MIXTOS	36
2.2.1. MÈTODE DE JENKINSON I COLLISON	36
2.2.2. OGWT – OBJECTIVACIÓ DEL HESS-BREZOWSKY GROSSWETTERLAGEN ALEMANY...	40
2.2.3. LA CLASSIFICACIÓ PER PROTOTIPUS.....	41
2.3. MÈTODES BASATS EN CORRELACIONS	44
2.3.1. MÈTODE DE LUND [1963]	44
2.3.2. MÈTODE DE KIRCHHOFER [1973].....	46
2.4. MÈTODES BASATS EN L'ANÀLISI DE COMPONENTS PRINCIPALS	50
2.4.1. ESTABLIMENT DE PATRONS DE BAIXA FREQUÈNCIA AMB L'ACP – LA ROTACIÓ DELS FACTORS.....	61
2.4.2: L'ESTRUCTURA DE LA MÀTRIU TIPUS S (S-MODE) PER A LA CTC.....	67
2.4.3. L'ESTRUCTURA DE LA MÀTRIU TIPUS T (T-MODE) PER A LA CTC.....	70
2.4.4. ALTRES VARIANTS DE L'ACP PER A LA CTC.....	72
2.4.5. ALTRES APLICACIONS INTERESSANTS DE L'ACP	74
2.5. MÈTODES D'AGRUPACIÓ OPTIMITZADA O <i>CLUSTERING</i>	76
2.5.1. EL QUICK CLUSTER [HARTIGAN 1975].....	80

2.5.2. SANDRA: SIMULATED ANNEALING AND DIVERSIFIED RANDOMIZATION CLUSTERING	82
2.5.3. SOM: SELF-ORGANIZING MAPS	83
2.5.4. COSA: CLUSTERING OBJECTS ON SUBSETS OF ATTRIBUTTES	85
2.6. ALTRES MÈTODES DE DE CLASSIFICACIÓ DE TIPUS DE CIRCULACIÓ	86
2.7. QUIN MÈTODE DE CLASSIFICACIÓ DE TIPUS DE CIRCULACIÓ S'HA D'UTILITZAR?	88
3. HIPÒTESI DE TREBALL I OBJECTIUS DE LA TESI	89
4. MÈTODE DE CLASSIFICACIÓ DELS TIPUS DE CIRCULACIÓ: PROPOSTA METODOLÒGICA	91
4.1. TRIA DEL MÈTODE DE CLASSIFICACIÓ	91
4.2. PROPOSTA METODOLÒGICA: MÈTODE DE LES PUNTUACIONS EXTREMES	95
4.2.1. ESTANDARDITZACIÓ ESPACIAL DE LES DADES	96
4.2.2. ESTABLIMENT DEL NOMBRE DE GRUPS I DELS SEUS CENTROIDES	99
4.2.3. CLASSIFICACIÓ DELS CASOS	101
5. APLICACIÓ DE LA METODOLOGIA A CASOS DIVERSOS PER A LA SEVA VALIDACIÓ .	103
6. COMPENDI D'ARTICLES.....	105
6.1. PATRONS DE CIRCULACIÓ ATMOSFÈRICA RELACIONATS AMB DIES DE NEVADA INTENSA A ANDORRA (PIRINEUS). <i>INTERNATIONAL JOURNAL OF CLIMATOLOGY</i> 25, 319-329. ANY 2005.....	105
6.2. CATÀLEG DIARI DE LA CIRCULACIÓ ATMOSFÈRICA PER SECTOR DE L'EUROPA OCCIDENTAL UTILITZANT TÈCNIQUES ESTADÍSTIQUES MULTIVARIANTS. <i>INTERNATIONAL JOURNAL OF CLIMATOLOGY</i> 26, 1501-1515. ANY 2006.....	119
6.3. MODELITZACIÓ ESPACIAL DE LA TEMPERATURA DE L'AIRE I LA PRECIPITACIÓ A ANDORRA (PIRINEUS) A PARTIR DE TIPUS DE CIRCULACIÓ DIARIS. <i>THEORETICAL AND APPLIED CLIMATOLOGY</i> , 2009.....	137
6.4. TIPUS DE CIRCULACIÓ RELACIONATS AMB L'ACTIVITAT DE LLAMPS A CATALUNYA I EL PRINCIPAT D'ANDORRA. <i>PHYSICS AND CHEMISTRY OF THE EARTH</i> , 2010.....	153
7. CONCLUSIONS I DISCUSSIÓ	163
7.1. PEL QUE FA AL MÈTODE DE CLASSIFICACIÓ	164
7.2. PEL QUE FA A LES NEVADES INTENSES A ANDORRA	167
7.3. PEL QUE FA A LA CLASSIFICACIÓ A L'EUROPA OCCIDENTAL	170
7.4. PEL QUE FA A LA CARTOGRAFIA CLIMÀTICA AMB SIG	171
7.5. PEL QUE FA ALS PERÍODES D'ACTIVITAT DE LLAMPS SIGNIFICATIUS.....	173
8. BIBLIOGRAFIA	175

TAULA DE FIGURES

FIGURA 1.1: Exemple d'una classificació de tipus de circulació (CTC).	16
FIGURA 1.2: El <i>Downscaling</i> busca establir relacions objectives entre dues escales	17
FIGURA 1.3: Cartografia dels valors mitjans de canvi en percentatge de la precipitació per al període 2090-2099 amb relació al 1980-1999	18
FIGURA 1.4: Resultat de regionalitzar les Illes Balears amb la variable temperatura i utilitzant l'anàlisi de components principals en S-mode.....	21
FIGURA 1.5: Adaptació al català dels tipus de clima de Gerstengarbe i Werner [2007] per a tot el món	23
FIGURA 1.6: El patró de la Oscil·lació de l'Atlàntic Nord respon a un dipol de pressió a l'Atlàntic Nord.....	24
FIGURA 1.7: Com classifiquem en grups aquests <i>smilies</i> ?	25
FIGURA 2.1: Imatge de H. H. Lamb, autor d'una de les classificacions manuals que ha tingut més repercussió en la Climatologia mundial	29
FIGURA 2.2: Principals trajectòries ciclòniques establertes per Pédelaborde a l'Europa occidental, i de les quals destaca algunes com a més rellevants pels Pirineus	34
FIGURA 2.3: Cartografia sinòptica associada a l'advecció septentrional segons el catàleg de Raso [1999] centrat a Andorra	35
FIGURA 2.4: Distribució recomanada dels punts de malla en relació amb el sector objecte d'estudi en el mètode de Jenkinson i Collison.	37
FIGURA 2.5: Formulació del mètode de Jenkinson i Collison.....	37
FIGURA 2.6: Esquema descriptiu sobre el procediment a seguir per a aplicar el mètode dels prototipus.....	43
FIGURA 2.7: Imatge de l'ordinador IBM 704. Lund, en el seu article de 1963, cita explícitament la utilització d'aquesta màquina.....	45
FIGURA 2.8: Reproducció dels dos tipus principals obtinguts per Lund en el seu treball de 1963.	46
FIGURA 2.9. L'ACP consisteix a canviar els eixos de coordenades que fem per descriure les nostres dades i així optimitzar i simplificar la seva explicació.	52
FIGURA 2.10.: Un simple gràfic de l'evolució dels valors propis d'una mateixa mostra, però amb diferents mides	55

FIGURA 2.11.: Exemple de representació gràfica de l'Scree Test	59
FIGURA 2.12.: La rotació de factors reajusta els nous eixos de coordenades creats amb l'ACP.....	63
FIGURA 2.13: Cartografia dels coeficients (<i>loading patterns</i>) obtinguts per	65
FIGURA 2.14. Taula amb els modes matricials que s'utilitzen a l'ACP i que, quan treballem amb informació meteorològica, combina tres tipus d'entitats diferents: espai (malla de dades o xarxa d'estacions), temps (series diàries, mensuals...) i camp (variable meteorològica com temperatura, pressió en superfície,...).	66
FIGURA 2.15: Característiques del mode S, extret de la FIGURA 2.14.....	68
FIGURA 2.16: El procés habitual seguit quan fem una classificació amb l'ACP amb matriu tipus S.....	69
FIGURA 2.17: Característiques del mode T, extret de la FIGURA 2.14.....	70
FIGURA 2.18. Exemple de dendrograma, representació gràfica del procés que segueixen els mètodes jeràrquics aglomeratius per classificar una mostra de casos... ..	78
FIGURA 2.19: Reproducció del funcionament de classificació d'un mètode d'optimització global.	79
FIGURA 2.20: El mètode de K-mitjanes resulta inestable a l'hora de convergir els seus resultats un cop fetes les iteracions a partir de diferents llavors de partida i amb la mateixa mostra.....	81
FIGURA 2.21: Esquema clàssic del funcionament d'una xarxa neuronal unidireccional.	84
FIGURA 2.22: Classificació de Rasilla obtinguda el 2003.....	87
FIGURA 4.1: Esquema gràfic que representa la proposta metodològica que es fa en el present treball.	95
FIGURA 4.2: Variacions estacionals de la circulació atmosfèrica entre gener (esquerra) i juliol (dreta).	97
FIGURA 4.3: Esquema on es mostra el mètode de les puntuacions extremes per establir el nombre de grups d'una classificació i calcular els seus centroides.....	101

LLISTAT D'ACRÒNIMS

Com que el present treball està format per textos tant en català com en anglès, hi ha acrònims associats a noms en català i altres a mots en anglès; aquest darrer apareix en parèntesi quan això ocorre per a les dues llengües.

ACP (PCA): anàlisi de components principals / *Principal Component Analysis*
AS: *Asian Summer Pattern*
CAPE: *Convective Available Potential Energy*
CG: núvol-terra / *Cloud to Ground*
CL: mètode d'agrupació optimitzada / *Clustering*
COSA: *Clustering Objects on Subsets of Attributes*
COST: *European Cooperation on Science and Technology*
CP (PC): component principal / *Principal Component*
CRECIT: Centre de Recerca en Ciències de la Terra
CRU: *Climatic Research Unit*
CTC: classificació de tipus de circulació / *Circulation Types Classification*
MDE (DEM): model digital d'elevacions / *Digital Elevation Model*
DPP: probabilitat diària de precipitació / *Daily Precipitation Probability*
E: est
ECMWF: *European Centre for Medium-Range Weather Forecasts*
EMULATE: *European and north atlantic daily to MULTidecadal climATE variability*
FEDA: Forces Elèctriques d'Andorra
EOF: funcions ortogonals empíriques / *Empirical Orthogonal Functions*
GPH: altura geopotencial / *Geopotential height*
GWL: *Grosswetterlagen* (mot alemany)
hPa: hectopascals
IC: llamp dins de núvol / *intra-cloud flashes*
ICC: Institut Cartogràfic de Catalunya
INM: Instituto Nacional de Meteorología
IPCC: Panell Intergovernamental per al Canvi Climàtic / *Intergovernmental Panel for Climate Change*
IQR: Rang interquartílic / *Interquartilic Range*
JC: *Jenkinson & Collison*
LF: baixa freqüència / *Low frequency*
m/s: metres per segon
MCBF: modes de circulació de baixa freqüència / *modes of low-frequency circulation variability*
MCG(GCM): model de circulació general / *General Circulation Model*
MCR: model de circulació regional
MDMn: temperatura mínima mitjana diària / *Mean Daily Minimum Temperature*
MDMx: temperatura màxima mitjana diària / *Mean Daily Maximum Temperature*
MDP: precipitació mitjana diària / *Mean Daily Precipitation*
MSU: *Montana State University*
N: nord
NAO: oscil·lació de l'Atlàntic Nord / *North Atlantic Oscillation*
NCAR: *National Center of Atmospheric Research*
NCEP: *National Centers for Environmental Prediction*

NCEP NFL: *Final analyses of the Global Tropospheric Analyses*
NE: nord-est
NNW: xarxes neuronals / *Neural Networks*
NPO: oscil·lació del Pacífic Nord
NW: nord-oest
O₃: ozó
°C: Graus centígrads
OGWT: objectivització de *Grosswetterlagen* alemany
PM₁₀: partícules en suspensió de diàmetre 10 micròmetres
PNA: *Pacific /North American Pattern*
PSO: processos sinòptics objectius
PSPA: *Principal Sequence Pattern Analysis*
QC: control de qualitat / *Quality Check*
REES: reducció estadística d'escala / *Statistical Downscaling*
RMSE: error mitjà quadràtic / *Root Mean Square Error*
RPC: components principals rotats / *Rotated Principal Components*
S: sud
SAFIR: *Satellite for Information Relay*
SE: sud-est
SEOFS: *Simplified Empirical Ortogonal Functions*
SIG (GIS): Sistema d'Informació Geogràfica / *Geographical Information System*
SLP: pressió atmosfèrica en superfície / *Sea Level Pressure*
SOM: *Self-Organizing Maps*
SPSS: *Statistical Package for Social Sciences*
SW: sud-oest
SZ: *Subtropical Zonal Pattern*
TC (CT): tipus de circulació atmosfèrica / *Atmospheric Circulation Types*
TROTREP: *Tropospheric Ozone and its Precursors*
USA o US: *United States of America*
USC: *University of South Carolina*
USGS: *United States Geophysical Survey*
UTC (GMT): temps universal coordinat / *Greenwich Mean Time*
UTM: Universal Transversal de Mercator
W: oest
WeMO: oscil·lació de la Mediterrània occidental

PREFACI

Amb el present document es mostra la recopilació d'un llarg període de treball en el camp de la Climatologia sinòptica que començà l'any 2000 i que junt amb les inquietuds ja existents sobre la Meteorologia i la Climatologia de muntanya, conclou ara amb la presentació d'aquesta memòria. Ha estat un llarg temps de formació científica molt valuós i amb un recorregut variat. L'inici, amb la incorporació al Grup de Climatologia de la Universitat de Barcelona i la concessió de la beca FI de la Generalitat de Catalunya, va representar un primer període de treball intens per abordar un camp de recerca força desconegut en aquell moment, la classificació de la circulació atmosfèrica amb mètodes automàtics o semiautomàtics, i que va demanar de visitar centres a l'estranger com la *Climatic Research Unit d'Anglaterra*, a Norwich (CRU). Posteriorment, en voler centrar part del treball en zones d'orografia complexa (i concretament a Andorra, gràcies al suport de l'Institut d'Estudis Andorrans), es va obrir la possibilitat de visitar la Universitat de Carolina del Sud, a Columbia (USC), i la de l'estat de Montana, a Bozeman (MSU), ja que molts avenços en el camp de la neu, les allaus i la Meteorologia de muntanya s'estaven fent a l'altra banda de l'Atlàntic. D'aquí sorgiren les primeres publicacions que s'adjunten en aquest treball, però paral·lelament s'inicià la participació en el projecte COST733, que responia en el seu *Memorandum of Understanding* exactament al tipus de recerca que estàvem duent a terme nosaltres: optimitzar la classificació de la circulació atmosfèrica a resolució diària. Amb el temps veiem clarament que ha valgut la pena esperar a la finalització de COST733, ja que l'aprenentatge i la feina feta en l'àmbit d'aquesta acció europea han ampliat moltíssim el nostre coneixement en els mètodes de classificació de la circulació atmosfèrica, ha permès consolidar la feina feta en els anys previs, i ha possibilitat treballar amb altres especialistes europeus. Això ens ha mostrat de primera mà l'estat en què es troba la Climatologia sinòptica. Fruit de tot aquest recorregut, creiem que aquest treball posa sobre la taula tres aportacions principals:

1. Una extensa sinopsi sobre conceptes de Climatologia sinòptica i mètodes de classificació de la circulació atmosfèrica.

2. Una nova proposta metodològica basada en una de les aproximacions més reconegudes: l'anàlisi de components principals.
3. L'aplicació d'aquesta metodologia a diferents camps actuals de les ciències atmosfèriques, amb aportacions també concretes en cadascun: Meteorologia de muntanya, variabilitat i canvi climàtic, cartografia climàtica i riscos meteorològics.

En definitiva, ens trobem davant un treball de tesi doctoral en el nucli del qual rau l'objectiu d'avançar des d'una perspectiva metodològica, però que no pot evitar aplicar la metodologia que s'hi proposa a realitats meteorològiques i climàtiques centrades principalment en sectors de muntanya i relacionades amb l'escala de detall, sense deixar de banda un dels temes centrals de la recerca climàtica actual: el canvi climàtic.

Esperem, doncs, que tots aquells que consulteu aquest document hi trobeu eines o resultats d'interès, i us animem a treballar en el camp de la Climatologia sinòptica i, concretament, en el de les classificacions de tipus de circulació amb tot el seu ampli ventall de mètodes i aplicacions.

1. INTRODUCCIÓ: CONCEPTES EN CLIMATOLOGIA SINÒPTICA

Classificar ha estat i és una eina molt emprada dins de les ciències per estudiar fenòmens i processos tant naturals com socials, i així avançar en el seu coneixement. Les ciències atmosfèriques no n'han estat cap excepció, i la Climatologia sinòptica n'és una branca que ha emprat sovint aquesta metodologia per establir catàlegs que resumeixen el comportament de variables atmosfèriques.

Una d'aquestes aproximacions és la classificació de tipus de circulació (CTC – *circulation types classification*), i que l'entendem com l'establiment d'un catàleg format per un seguit de tipus de circulació atmosfèrica (TC – *atmospheric circulation types*). Aquest catàleg, que s'estableix a partir d'una sèrie temporal de mapes del temps (normalment camps de pressió en superfície (SLP) o de geopotencial (GPH), resumeix de la forma més òptima possible la circulació atmosfèrica característica d'un sector en estudi. Dit d'una altra manera, la CTC és un mètode, principalment, per categoritzar la continuïtat de la circulació atmosfèrica en un nombre raonable i manejable de tipus de circulació discrets [Beck i Philipp; 2010]. A través d'aquesta discretització podem analitzar aquests tipus de circulació establerts, així com altres variables (normalment meteorològiques en superfície), a escala local, regional, etc. (FIGURA 1.1), acomplint el principal objectiu associat a la Climatologia sinòptica (*synoptic climatology*) tal com s'enfoca actualment: integració del comportament simultani de la dinàmica atmosfèrica i de la resposta acoblada en el medi ambient superficial [Yarnal *et al.*, 2001].

Ja avancem, tot i que hi entrarem molt més a fons al llarg del present treball, que considerem com a treballs clau per tenir una bona visió dels mètodes de classificació en Climatologia sinòptica, i especialment els de CTC, els de Yarnal [1993], Barry i Carleton [2001], Yarnal *et al.*, [2001] i el més recent de Huth *et al.*, [2008].

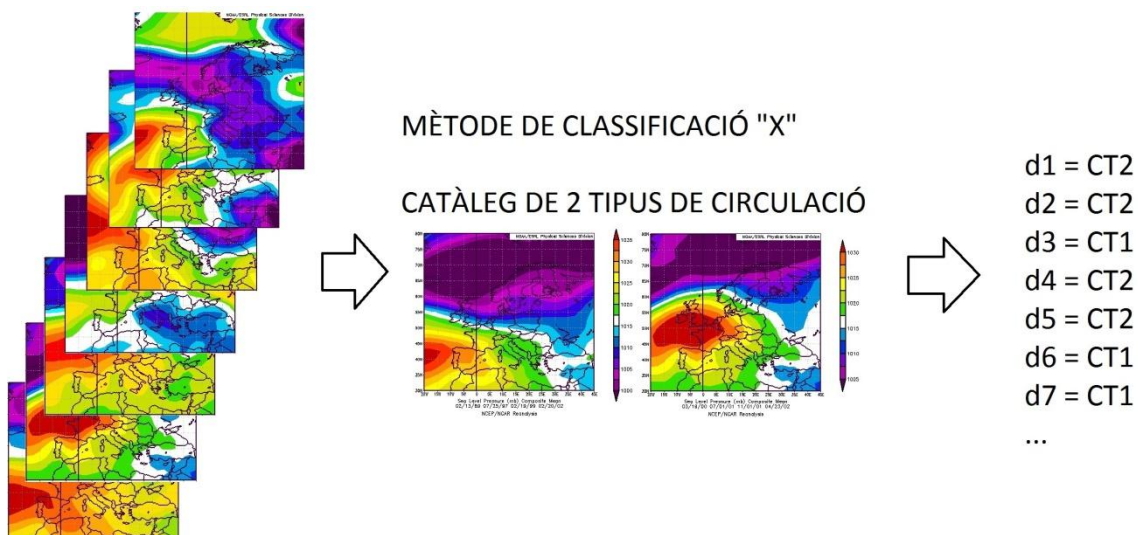


FIGURA 1.1: Exemple d'una classificació de tipus de circulació (CTC). A partir d'un mètode X i d'una sèrie de dies amb ocurrència d'incendis forestals al principat d'Andorra, podem separar dos grups de dies per similitud en les característiques de la circulació atmosfèrica (mitjana dels casos per grup): dorsal de l'anticicló de les Açores sobre els Pirineus (vent de nord-oest a nord sobre els Pirineus), i anticicló a l'oest de França (vent del nord a nord-est sobre els Pirineus). Tenim, doncs, una CTC amb dos TC o CT (*circulation types*) associats als incendis forestals a Andorra. Font: NOAA [2010] i basat en Esteban i Gallego [2011].

Al llarg de les darreres dècades les aplicacions de les CTC han estat diverses. Però des del nostre punt de vista, i a dia d'avui, aquesta tècnica la considerem de molt interès dins d'un procés més ampli, l'anomenada **reducció estadística d'escala** (REES – *statistical downscaling*). Benestad *et al.*, [2008] ho defineixen com el procés d'establir un lligam entre l'estat d'una variable representativa a escala petita (normalment hemisfèrica, sinòptica, mesoescalar) i l'estat d'una variable representativa d'una escala molt més gran (normalment local, puntual). És aquí on volem puntualitzar, sent geogràficament correctes (i al revés del que s'acostuma a utilitzar), que escala petita correspon a escales de poca resolució, de manera que parlàriem, per al cas dels CTC, de, per exemple, la distribució del camp de pressió sobre l'Europa occidental, mentre que escala gran s'ajusta al concepte d'escala local o alta resolució, i es referiria a les mesures d'una estació meteorològica com la de Ransol al principat d'Andorra o a la d'un sector de territori amb alta densitat i variabilitat d'informació (FIGURA 1.2).

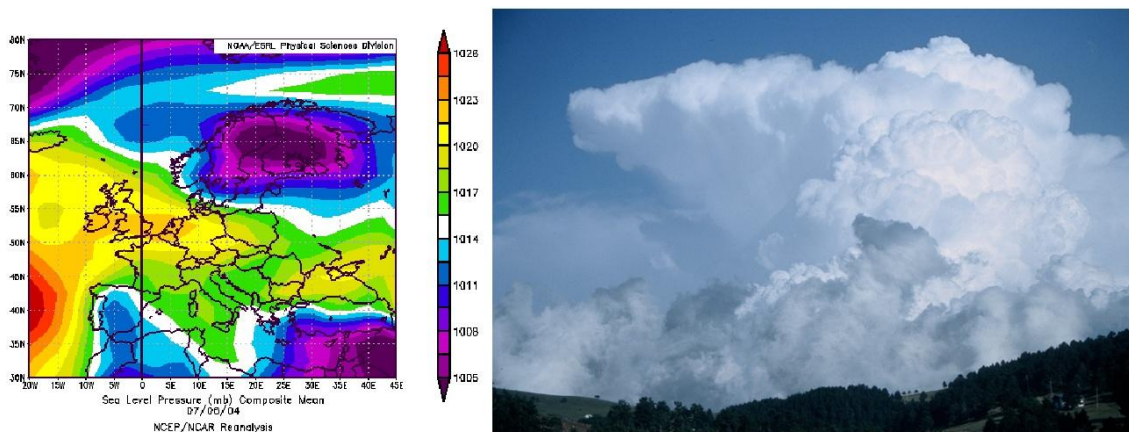


FIGURA 1.2: El *Downscaling* busca establir relacions objectives entre dues escales meteorològiques (l'escala petita, el mapa sinòptic, i l'escala gran, una tempesta a una vall pirinenca). Un TC de baixa tèrmica sobre la península Ibèrica, representat aquí pel dia 6 de juliol de 2004, va sovint associat a la formació de tempestes als Pirineus Orientals. Quina probabilitat tenim de tenir tempesta als Pirineus Orientals en situació de baixa tèrmica? Mapa: NOAA [2010]. Fotografia: Pere Esteban.

Segons Benestad *et al.*, [2008] aquesta relació entre dues escales diferents es pot establir de forma objectiva a través de diferents metodologies tant de caràcter lineal com la regressió múltiple o les correlacions canòniques, així com amb enfocaments no lineals com els mètodes d'analogies o els mètodes de classificació. És precisament en aquestes tècniques no lineals on entra en joc l'ús de tècniques de CTC. Més detalls sobre les tècniques estadístiques emprades per al REES es poden trobar també a Von Storch i Zwiers [1999] i a Wilks [2006].

L'establiment d'aquestes relacions a diferents escales amb el REES ens obre dues grans línies de treball:

CLIMATOLÒGICA: Fa referència a la creació d'escenaris futurs de canvi climàtic amb l'ús dels **models de circulació general** (MCG – *Global Circulation Models*) o **models de circulació regional** (MCR – *Regional Circulation Models*). Podent desenvolupar-se escenaris regionals i locals que permetin l'estudi dels impactes del canvi climàtic [Yarnal *et al.*, 2001]. Si establim un model estadístic que relacioni la circulació atmosfèrica actual a l'escala dels MCG amb el clima i els fenòmens naturals a escala local, aquesta mateixa relació es pot aplicar a les prediccions fetes per aquests models

de canvi climàtic, que aportaria informació del comportament del clima a escala local en el futur i per a diferents escenaris [Wilby *et al.*, 2004]. Cal no oblidar, això sí, que les relacions estadístiques que es puguin establir en l'actualitat no necessàriament s'han de mantenir en el futur. Com a exemple d'aquest treball volem destacar Demuzere *et al.*, [2009] on emprà les classificacions per abordar el camp de la contaminació atmosfèrica.

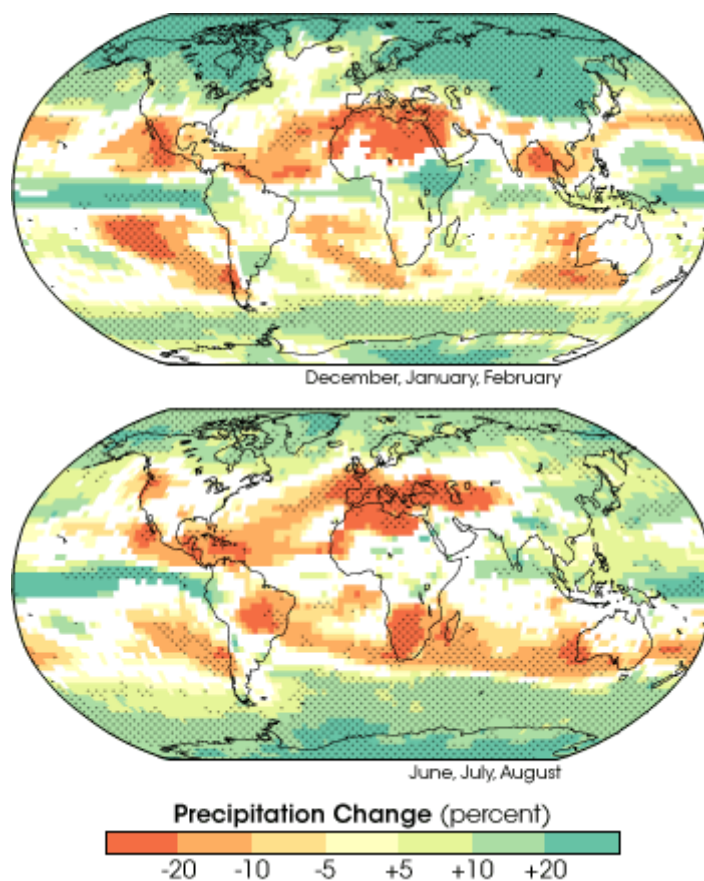


FIGURA 1.3: Cartografia dels valors mitjans de canvi en percentatge de la precipitació per al període 2090-2099 amb relació al 1980-1999 (hivern i estiu) a partir dels valors obtinguts per diferents models de circulació general. Aquests s'empren per projectar variacions futures de variables climàtiques, però els importants requeriments de càlcul van en detriment de la resolució espacial dels resultats. Amb les tècniques de *downscaling* podem derivar aquesta informació a l'escala local i així interpretar de forma més acurada aquesta informació per a la nostra zona d'estudi. Font: IPCC [2007].

METEOROLÒGICA: Una segona aplicació és la realització de **previsions meteorològiques híbrides**, que combinen les sortides dels models numèrics amb la informació estadística (climatologia) per optimitzar la predicció de fenòmens que depenen de processos físics complexos i que ocorren a escales menors que la resolució dels models dinàmics (precipitació, boires, tempestes...) [Gutiérrez *et al.*, 2004]. Molt interessant és l'exemple recent publicat per Peña *et al.* [2011] amb relació als episodis de vent fort a Catalunya.

Sembla, d'altra banda, que l'aplicació de les CTC per a fer **prediccions estacionals** també està avançant amb força [Chabot *et al.*, 2008].

Tots aquests conceptes i les seves interrelacions ens permeten veure que l'establiment de les relacions estadístiques òptimes entre la circulació atmosfèrica i el clima local és una de les peces angulars de la recerca climàtica actual. Veiem, doncs, que com en un trencaclosques, un correcte establiment dels principals tipus de circulació atmosfèrica que regeixen un sector és clau per a obtenir uns bons resultats, si abordem recerques o aplicacions com les que s'han exposat anteriorment. El present treball aborda, en gran part, aquest darrer punt.

Per classificar la circulació atmosfèrica es treballa a partir de camps de variables sobre un sector. Tradicionalment parlem de mapes del temps a escala sinòptica, però a dia d'avui normalment es treballa amb **malles de dades** (*grids*) que ens permeten tenir llargues sèries temporals de cartografia atmosfèrica en format numèric: en són coneguts exemples les reanàlisis del NCEP/NCAR americà (*National Centers for Environmental Prediction/National Center for Atmospheric Research*) (<http://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml> - agost 2011) o del ECMWF europeu (*European Centre for Medium-Range Weather Forecasts*) (http://www.ecmwf.int/research/era/do/get/Reanalysis_ECMWF - agost 2011), conegut com a ERA40.

Existeixen altres enfocaments en Climatologia sinòptica que també es basen en les classificacions per caracteritzar el clima. En aquest sentit, la classificació simultània de diferents variables observades en estacions meteorològiques (escala local-puntual) ens permet diferenciar el concepte de **tipus de temps** (*weather types*), referit a les combinacions més habituals de diverses variables atmosfèriques (temperatura, precipitació, humitat, radiació solar...) per a una o diverses estacions meteorològiques. El treball de referència d'aquesta aproximació és encara ara el de **Kalkstein et al. [1987]**. Aquests *weather types* són combinacions de variables meteorològiques que es poden relacionar posteriorment amb una determinada distribució del camp de pressió o geopotencial, de manera que finalment acabem també obtenint classificacions de la circulació atmosfèrica. **Huth et al. [2008]** destaquen que cal diferenciar-ho de la classificació de **masses d'aire** (*air mass classifications*), en aquest cas la classificació també és de diverses variables, però principalment a nivells baixos i mitjans de l'atmosfera.

Hem de dir que les classificacions de tipus de temps i de masses d'aire no són gaire freqüents a la bibliografia. Pel que fa les primeres, considerem interessant de citar el treball de **Huth et al., [1993]**, on estableix aquestes tipologies per a l'estació meteorològica Prague-Clementinum emprant dades diàries hivernals de temperatura mitjana de l'aire, amplitud tèrmica, humitat relativa de l'aire, força del vent, les components zonal i meridiana del vent, nuvolositat i tendència de la temperatura (diferència de la temperatura del dia respecte al dia anterior). Pel que fa les classificacions de masses d'aire, destaca el treball de **Bower et al. [2007]** que categoritzen l'Europa occidental en sis masses d'aire (polar seca, seca moderada, tropical seca, polar humida, humida moderada i tropical humida) a partir de 48 estacions per al període 1974-2000.

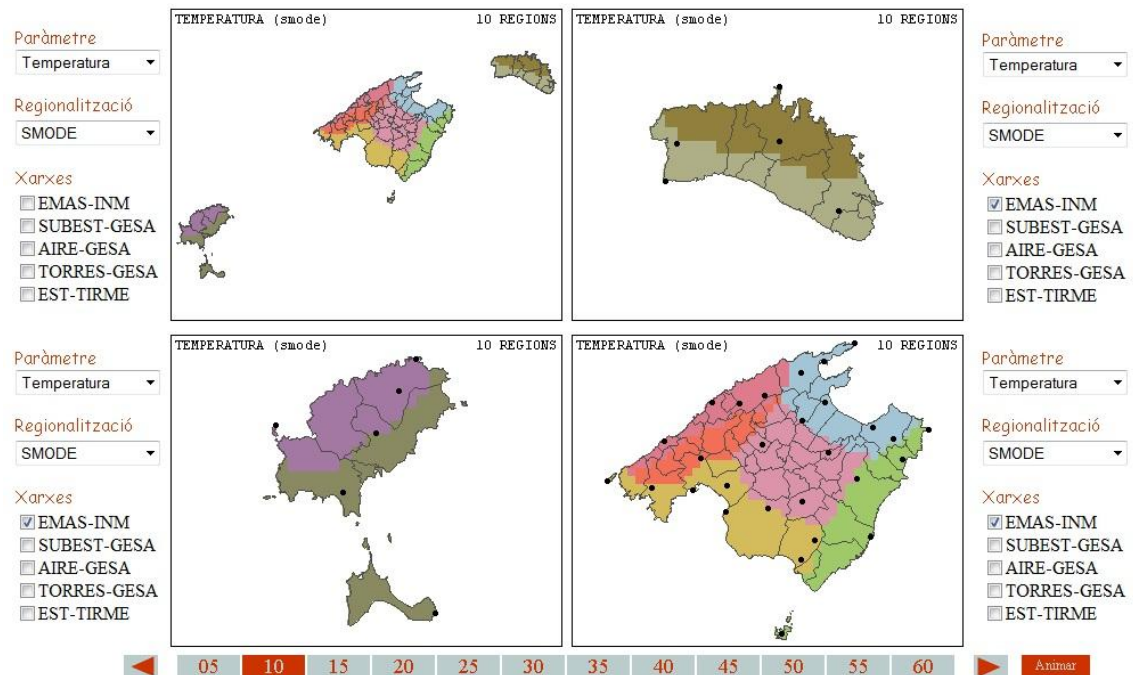
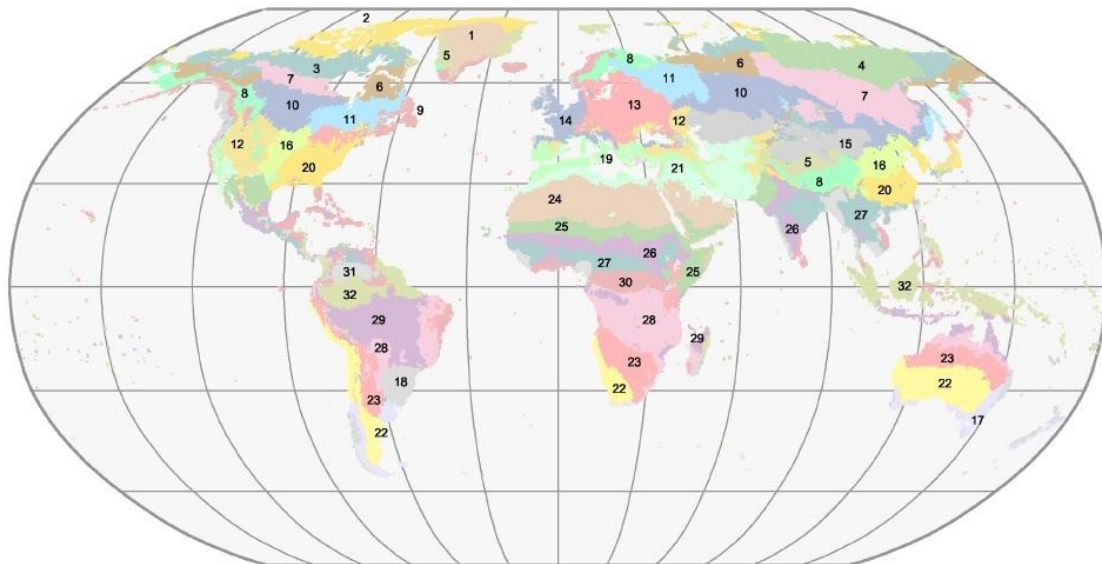


FIGURA 1.4: Resultat de regionalitzar les Illes Balears amb la variable temperatura i utilitzant l'anàlisi de components principals en S-mode (vegeu apartat 2.4 del present treball per a més detalls). A la web <http://mm5forecasts.uib.es/OCLIBreg/> es pot regionalitzar climàticament l'àmbit geogràfic de les Illes Balears a partir de les dades d'estacions meteorològiques i l'ús de mètodes automàtics de classificació. En aquest cas, és a partir de la nostra experiència que establím un total de deu regions climàtiques amb comportament diferent de la temperatura en aquesta zona.

Una tercera aproximació en Climatologia que es basa en la classificació és la **regionalització climàtica** (*climate regionalization*), que estableix àrees geogràfiques amb propietats climàtiques similars (FIGURA 1.4). Són les divisions climàtiques tradicionals, que avui dia també han trobat una eina molt consistent en les possibilitats d'aplicació de mètodes estadístics amb computadores potents.



Tipus de clima



TIPUS DE CLIMA	Temperatura			Amplitud tèrmica	Precipitació		
	Mitjana anual	Mitjana estiu	Mitjana hivern		Total anual	Total estiu	Total hivern
1	molt fred	molt fred	molt fred	gran	mod. sec	sec	mod. sec
2	molt fred	fred	molt fred	molt gran	sec	molt sec	sec
3	molt fred	fred	molt fred	molt gran	sec	sec	sec
4	molt fred	fresc	molt fred	extrema	sec	sec	sec
5	fred	fred	fred	gran	sec	sec	mod. sec
6	fred	fresc	molt fred	molt gran	mod. sec	sec	mod. sec
7	fred	temperat	molt fred	extrema	mod. sec	mod. sec	sec
8	fresc	fresc	fred	gran	mod. sec	mod. sec	mod. sec
9	fresc	fresc	fred	mitjana	humit	mod. sec	humit
10	fresc	temperat	fred	molt gran	mod. sec	sec	mod. sec
11	fresc	temperat	fred	molt gran	mod.humit	mod. sec	mod.humit
12	temperat	temperat	fresc	gran	mod. sec	sec	mod. sec
13	temperat	temperat	fresc	gran	mod.humit	mod. sec	mod.humit
14	temperat	temperat	temperat	mitjana	mod.humit	sec	humit
15	temperat	càlid	fred	molt gran	sec	molt sec	sec
16	temperat	càlid	fresc	gran	mod.humit	mod. sec	sec
17	temperat	temperat	temperat	mitjana	mod.humit	sec	humit
18	càlid	temperat	càlid	mitjana	humit	mod.humit	humit
19	càlid	càlid	temperat	mitjana	mod. sec	molt sec	humit
20	càlid	càlid	temperat	gran	humit	mod.humit	humit
21	càlid	molt càlid	temperat	gran	sec	molt sec	mod. sec
22	càlid	càlid	càlid	mitjana	sec	sec	mod. sec
23	càlid	càlid	càlid	mitjana	mod. sec	mod. sec	sec
24	molt càlid	molt càlid	càlid	mitjana	molt sec	molt sec	molt sec
25	molt càlid	sempre càlid		mitjana	sec	sec	sec
26	molt càlid	sempre càlid		petita	mod.humit	mod.humit	sec
27	molt càlid	sempre càlid		petita	humit	humit	sec
28	molt càlid	sempre càlid		petita	humit	mod.humit	sec
29	molt càlid	sempre càlid		molt petita	humit	humit	mod.humit
30	molt càlid	sempre càlid		molt petita	humit	mod.humit	humit
31	molt càlid	sempre càlid		petita	molt humit	molt humit	mod.humit
32	molt càlid	sempre càlid		molt petita	molt humit	humit	molt humit

PARÀMETRE						
Temperatura (°C)	molt fred	fred	fresc	temperat	càlid	molt càlid
Mitjana anual	< -10	< -5	< 5	< 12	< 22	> 22
Mitjana estiu	< 0	< 8	< 12	< 20	< 26	> 26
Mitjana hivern	< -20	< -5	< 0	< 10	< 20	> 20
Amplitud tèrmica	molt petita	petita	mitjana	gran	molt gran	extrema
	< 5	< 10	< 20	< 30	< 40	> 40
Precipitació (mm)	molt sec	sec	mod. sec	mod. humit	humit	molt humit
Total anual	< 50	< 300	< 600	< 1000	< 2000	> 2000
Total estiu	< 60	< 200	< 300	< 500	< 1000	> 1000
Total hivern	< 10	< 50	< 100	< 150	< 350	> 350

FIGURA 1.5: Adaptació al català dels tipus de clima de Gerstengarbe i Werner [2007] per a tot el món (inici de la figura a la pàgina anterior). Al planisferi podem veure la distribució espacial dels trenta-dos tipus de clima establerts amb la classificació. A les taules, per una banda queden definits els criteris que caracteritzen les definicions climàtiques a partir de llindars relacionats amb dos paràmetres, la temperatura i la precipitació; criteris que finalment s'associen a cadascun dels trenta-dos climes per caracteritzar-los.

Com a treballs de regionalització climàtica en el nostre sector geogràfic podem citar, per exemple, l'establiment de patrons espacials de precipitació diària a Catalunya [Serra *et al.*, 1996], el treball de Romero *et al.* [1999a], on es regionalitza la precipitació a la façana mediterrània de la península Ibèrica amb l'ús de l'anàlisi de components principals (vegeu secció 2.4), o el de Peñarrocha *et al.* [2002], que estableix regions pluviomètriques amb dades de precipitacions intenses al País Valencià. I per a un sector més ampli i sense centrar-se només en la precipitació, destaca el treball de Gerstengarbe i Werner [2007], on fan una classificació climàtica a escala planetària (FIGURA 1.5) emprant mètodes d'agrupació no jeràrquics (vegeu secció 2.5) a partir de sèries meteorològiques de temperatura i precipitació.

Un altre aspecte a tenir en compte a l'hora de treballar amb tipus de circulació és diferenciar-los bé respecte als anomenats **modos de circulació de baixa freqüència** (MCBF-*modes of low-frequency circulation variability*). Aquest concepte respon a aquelles estructures atmosfèriques (normalment a escala hemisfèrica o sinòptica) que amb la seva major o menor presència i persistència faciliten o dificulten l'ocurrència dels diferents tipus de circulació atmosfèrica. Dit d'una altra manera: els patrons de

baixa freqüència són les traces principals que presenta la circulació atmosfèrica en un determinat moment; són els maons dels quals es compon la circulació atmosfèrica [Huth *et al.*, 2008]. En aquest sentit, d'un mateix mode de circulació de baixa freqüència se'n poden derivar diversos tipus de circulació.

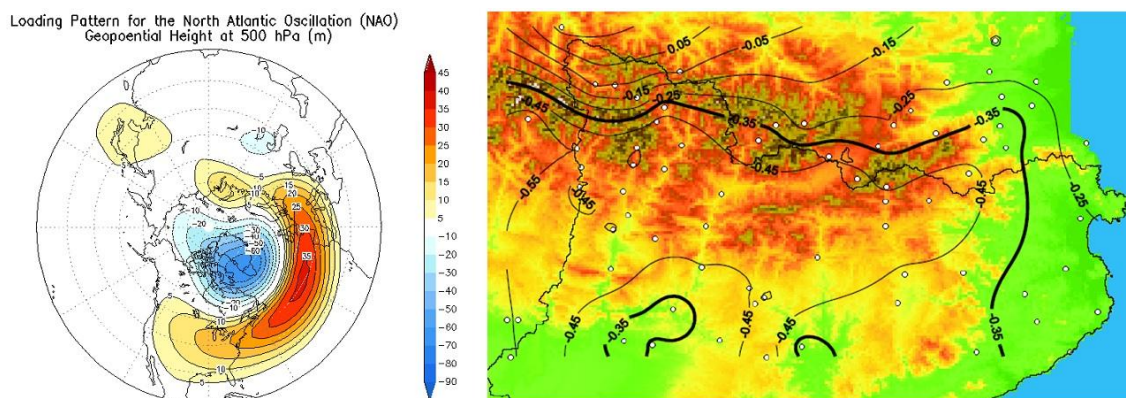


FIGURA 1.6: El patró de la Oscil·lació de l'Atlàntic Nord respon a un dipol de pressió a l'Atlàntic Nord. D'aquí s'estableix un índex que presenta valors positius quan respon al patró mostrat, i negatius quan s'inverteix. Valors de NAO negativa estan significativament correlacionats (inferiors a -0,35) amb les precipitacions al vessant sud Pirinenc, precipitacions que van associades a determinats patrons de circulació sovint de component meridiana. Figura esquerra: NCEP [2010]. Figura dreta: extret de Esteban *et al.*, [2002].

Precisament, com que el concepte de MCBF respon a la idea de trobar peces bàsiques que modulin el comportament de la circulació, l'exemple més conegut a la nostra zona correspondria a l'oscil·lació de l'Atlàntic Nord (NAO – *North Atlantic Oscillation*). En aquest sentit, segons quin sigui el seu comportament, es facilita la major presència de tipus de circulació associats als vents de l'oest, mentre que a més pot facilitar una major presència de situacions meteorològiques associades a la circulació meridiana o de bloqueig (FIGURA 1.6). S'han realitzat diferents treballs sobre quina és l'estructura i el comportament de la NAO, però volem destacar els resultats de Huth *et al.* [2006] sobre la influència que hi té l'activitat solar. També volem posar èmfasi en el treball de Garcia-Sellés *et al.* [2010], on s'estableix la relació entre la NAO, i també de la WeMO (oscil·lació de la Mediterrània occidental- *Western Mediterranean Oscillation*), amb l'ocurrència d'allaus de neu de grans dimensions als Pirineus Orientals. Aquest tipus de

treballs són de gran interès de cara al desenvolupament de les prediccions estacionals que citàvem anteriorment.

Més detalls sobre l'establiment dels patrons de baixa freqüència els trobarem al punt 2.4.1. amb la descripció de l'anàlisi de components principals, que és el mètode amb què sovint s'aborda aquesta identificació d'estructures de circulació.

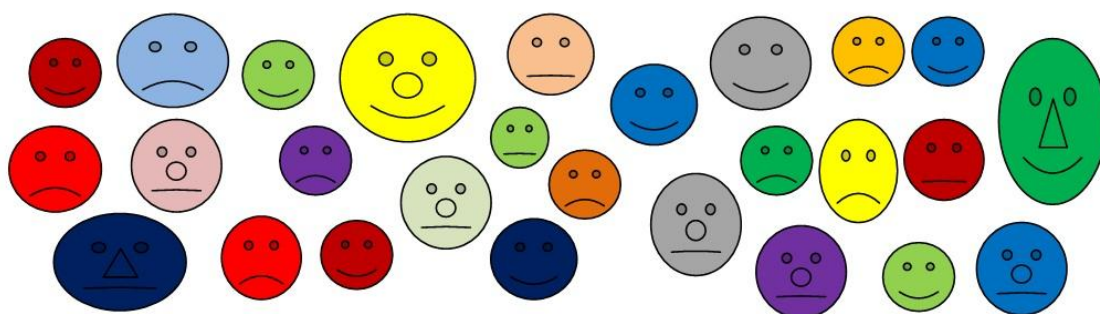


FIGURA 1.7: Com classifiquem en grups aquests *smilies*? Per color? Per somriure? Per nas? Per grandària? Cada cop, en augmentar la variabilitat i la continuïtat, la discriminació de grups es fa més complicada. Els diferents mètodes de classificació existents busquen optimitzar aquesta capacitat de separació, tenint en compte que potser hi haurà casos que difícilment es poden arribar a associar clarament a algun grup.

Finalment, i tornant de nou a l'establiment de TC, creiem important recordar que l'atmosfera no és precisament un element estàtic, i que intentar reflectir el seu dinamisme i continuïtat a través de l'establiment de patrons discrets és un objectiu certament complicat (FIGURA 1.7). Així, algunes classificacions s'han basat en la detecció dels principals centres d'acció i l'establiment de la seva morfologia i distribució més habitual, mentre que altres treballs han optat per l'estudi del moviment i les trajectòries d'aquests centres d'acció. Reflectir la dinàmica atmosfèrica a través de patrons discrets és, doncs, una problemàtica encara no superada satisfactòriament, però que actualment impulsa els climatòlegs a desenvolupar noves fórmules de classificació cada cop més òptimes. És en aquest àmbit on s'emmarca el projecte **COST 733 - Harmonisation and Applications of Weather Type Classifications for European Regions** [cost733.met.no/index.htm], i que ha aportat resultats sobre els

avantatges i inconvenients de les diferents aproximacions existents, així com de les millors opcions, si n'hi ha, per resoldre les problemàtiques abans exposades.

És que l'acció europea **COST 733** per a l'harmonització i l'aplicació de les classificacions de tipus de temps a Europa representa l'esforç de síntesi més important i actualitzat que s'ha fet sobre aquesta temàtica, i concretament sobre la classificació de tipus de circulació (tot i dir-se *weather type classifications* en el nom del projecte).

2. MÈTODES DE CLASSIFICACIÓ DELS TIPUS DE CIRCULACIÓ ATMOSFÈRICA (CTC)

L'estudi de la circulació atmosfèrica a través de l'establiment de tipus de circulació s'ha abordat sobretot des de la primera meitat del segle XX. L'evolució de les metodologies en les últimes dècades ha estat estretament lligada a l'evolució dels ordinadors, i especialment a la seva capacitat de càlcul. En aquest sentit, els enfocaments que s'han tractat per optimitzar les classificacions han estat diversos. A continuació anirem descrivint aquestes diferents modalitats per classificar la variabilitat atmosfèrica.

Ens ha semblat adient adoptar unes tipologies generals molt similars a les presentades a [Philipp *et al.* \[2009\]](#) i que deriven dels treballs de l'acció COST 733 citada al final del capítol anterior. Pel que fa a les metodologies que s'expliquen en cada tipologia, hem abordat aquelles que ens han semblat de més interès d'acord amb l'experiència de l'autor del present treball.

2.1. MÈTODES MANUALS O SUBJECTIUS

Són els mètodes amb què es van iniciar les CTC tan bon punt van aparèixer els mapes del temps a principi del segle XIX [Barry i Carleton, 2001], i on l'investigador classifica la informació sinòptica (mapes del temps) en grups i segons unes categories predeterminades [Yarnal, 1993] per a una regió concreta. Així doncs, la classificació resultant depèn del **criteri i coneixement de les dades i de l'àrea d'estudi de l'investigador**, que confereix a aquest mètode una marcada subjectivitat, fet que va acompanyat pel caràcter feixuc del procés i la important despesa de temps normalment requerida.

Un dels punts més delicats és la determinació de les categories en les quals classificar els mapes. Establir massa categories, que permetin una alta cohesió dels grups, pot tenir el problema que alguna agrupació presenti una mostra molt petita de casos, mentre que, si establim un grup massa reduït de categories, alguns grups poden incloure alta varietat (variabilitat) de les situacions [Barry i Carleton, 2001], amb tot el que pot implicar en Climatologia una opció o una altra. D'altra banda, la majoria de classificacions manuals no només cataloguen la mostra dels mapes, sinó que normalment hi associen el tipus de temps que es dona en el punt o regió d'interès per a l'investigador.

El mètode manual de classificació té com a principal inconvenient la **subjectivitat** aportada als resultats pel classificador, amb la possibilitat de donar relleu a situacions transitòries o poc freqüents. Alhora comporta treballar amb sèries llargues de mapes, amb una important despesa de temps i els canvis de criteri que es poden donar durant el llarg procés de classificació. En aquest sentit, si una classificació la fan diferents experts, correm el risc d'introduir inhomogeneïtats en els resultats [Cahynová i Huth, 2009]. Amb tot, encara que el mètode manual no sigui transferible, l'expert aporta una experiència i coneixement interessant i útil per a la classificació automàtica.

2.1.1. CLASSIFICACIÓ DE LAMB [1972]

És el treball de classificació de caràcter manual probablement més extensament tractat i citat. És un catàleg sinòptic per al sector de les illes Britàniques (50°-60°N / 10°W-2°E), originalment a partir de mapes diaris de pressió atmosfèrica en superfície, i posteriorment també amb mapes a 500 hPa [Kelly *et al.*, 1997], i tenint en compte l'evolució i el moviment dia a dia dels centres d'acció. Originalment cobria el període 1861-1971, en el qual Lamb (FIGURA 2.1) va identificar set categories o tipus **bàsics** (anticiclònic, ciclònic, oest, nord-oest, nord, est, sud), a part dels casos **híbrids** (per exemple, un dia amb ciclònic-oest o anticicló-nord-est), o els **no classificables** (situacions caòtiques, les incompatibles al mateix temps sobre la regió, o les que presenten canvis ràpids durant un dia). Veiem, doncs, que el mètode estableix els tipus, principalment, a partir de la component del vent i la presència d'altres o baixes pressions sobre un punt o regió d'estudi, que en el cas de Lamb, com hem dit, eren les illes Britàniques.

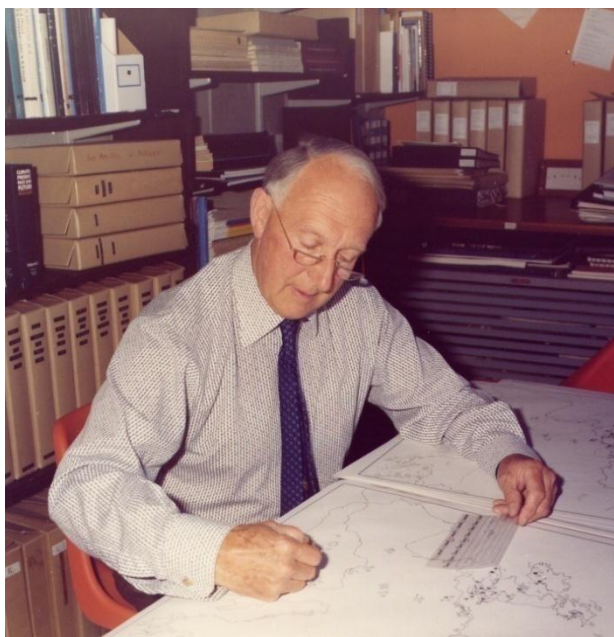


FIGURA 2.1: Imatge de H. H. Lamb, autor d'una de les classificacions manuals que ha tingut més repercussió en la Climatologia mundial. Aquestes imatges de climatòlegs amb llapis i mapes cada cop són menys freqüents. Font: Climatic Research Unit (CRU) <http://www.cru.uea.ac.uk/cru/people/lamb/>

Com que Lamb, a l'hora d'establir la classificació, també té en compte els vents dominants (no inclou els sistemes de vents locals, com ara les brises) i el temps resultant en superfície, la classificació no només es pot considerar de patrons de circulació (*circulation types*), sinó que com el títol del seu treball indica, també seria una classificació de tipus de temps (*weather types*). Així doncs, s'ajusta molt bé a la realitat climàtica de les illes Britàniques, però es difícil d'exportar a altres latituds. Precisament, de cara a resoldre aquest fet se n'ha fet l'automatització (vegeu apartat 2.2.1).

2.1.2. CLASSIFICACIÓ DE HESS-BREZOWSKY GROSSWETTERLAGEN

Una altra aproximació interessant és la que parteix del concepte alemany *grosswetter* (temps a gran escala). A partir de l'aplicació d'aquest concepte va aparèixer la classificació anomenada *Grosswetterlagen* (GWT) [Baur *et al.*, 1944; Hess i Brezowsky, 1952; Gerstengarbe *et al.*, 1999], centrada sobre Europa i el sector nord-est de l'Atlàntic Nord (24°-70°N/30°W-45°E), i caracteritzada per la distribució de la pressió en superfície i a nivells mitjans de la troposfera [Barry i Perry, 1973]. La tendència general del flux atmosfèric o moviment i trajectòria dels centres d'acció (el que s'ha anomenat en anglès *steering*) a tot el sector en estudi és el criteri que cal seguir per a establir el tipus de circulació [Yarnal, 1993], en comptes de les direccions del vent en superfície sobre una àrea concreta, com planteja Lamb. D'aquesta manera, cada tipus o categoria vindrà definit per la distribució mitjana a la zona d'estudi de la pressió en superfície al llarg d'un període de temps (tres dies) durant el qual la posició dels centres d'acció i dels règims de circulació variïn poc [Barry i Carleton, 2001], i les condicions del temps no presentin cap canvi o pocs al sector d'estudi [Barry i Perry, 1973]. En aquest sentit, els tipus establerts en aquesta classificació representen pròpiament règims de circulació més que no pas un conjunt de patrons de circulació mitjans diaris [James, 2006]. A partir d'aquest plantejament, es van establir vint-i-nou categories, que responen a tres grans grups: zonal, meridional o mixt, a part dels casos no classificables (categoria 30).

Aquesta és una classificació molt emprada a tots els països de l'Europa central i molt sovint es coneix estrictament com a classificació de Hess-Brezowsky, i a dia d'avui està implementada de forma operativa al Servei Meteorològic Alemany i amb inici de la sèrie el 1881.

Les principals deficiències que se li consideren són la manca de bons resultats per als sectors llunyans al centre de l'àrea d'estudi; és a dir, la classificació no s'ajusta prou bé a tot el continent europeu (especialment als sectors sud-est i nord-est d'Europa), algunes limitacions a l'hora d'establir prou diferències en tipus de l'oest, i alguna excessiva representació d'alguns tipus de bloqueig que podrien ser reduïts a un de sol [James, 2006].

2.1.3. CLASSIFICACIÓ DE PÉDELABORDE

Un tercer exemple que volem citar i que en el seu moment va marcar precedents sobre el tarannà de la climatologia a Catalunya i a Espanya és el de **Pédelaborde [1957]**, que descriu els patrons sinòptics de circulació atmosfèrica amb relació a la *Bassin Parisien*, que respondria al sector nord i occidental de França. En aquest treball, que pretén fer una Climatologia de l'àrea d'estudi, l'autor considera quatre passos per dur-la a terme:

1. La recerca de les característiques generals de la circulació atmosfèrica sobre la regió
2. La descripció dels tipus de temps
3. La definició del clima, a partir de la totalitat dels tipus de temps
4. La recerca de característiques regionals i locals d'aquest clima, en funció de les característiques geogràfiques de la zona d'estudi

El primer d'aquests passos, que respondria a la classificació de la circulació atmosfèrica, no només té en compte la distribució de la pressió en superfície o el component del vent, sinó que inclou en la caracterització variables com la circulació en

altura, el tipus de massa d'aire, la trajectòria de les pertorbacions i, fins i tot, la importància de la distribució i morfologia del territori en el comportament de la circulació. Pédelaborde, més que establir un catàleg de patrons, presenta diferent cartografia lligada a la circulació, amb la intenció de descriure'n els principals aspectes i el seu interès per a la zona en anàlisi. És en el posterior establiment dels tipus de temps quan realment es realitza un catàleg de tipus, on es considera la component del vent junt amb factors estacionals i les característiques de la massa d'aire associada (per exemple, tipus ciclònic de sud-oest a l'estiu: les pertorbacions profundes). És, doncs, una aproximació marcadament descriptiva, no independent dels tipus de temps, i seguint els patrons característics de l'escola de geografia francesa.

2.1.4. CLASSIFICACIONS A LA PENÍNSULA IBÈRICA

Si ens centrem al sector sud-oest d'Europa, i especialment a l'àrea de la península Ibèrica, hi ha uns quants catàlegs de patrons de circulació que s'han dut a terme a diferents escales espacials i per a diferents zones. Les classificacions en aquest àmbit territorial presenten una major complexitat respecte als centrats al continent Europeu per diferents aspectes, especialment per la localització de la península Ibèrica entre la zona temperada i la subtropical, i també per la presència de la mar Mediterrània. Així, la manca de definició del gradient de pressió, com també la possible no concordança entre les configuracions superficials amb les de nivells alts, prenen marcada importància a la nostra zona [Fernández *et al.*, 2003].

Les classificacions al sector peninsular més citades són la de Clavero i Raso [1979] amb divuit patrons de circulació de pressió en superfície i geopotencial a 500 hPa i alhora agrupats en quatre classes (situacions convectives o situacions advectives combinades amb situacions més o menys definides de subsidència anticiclònica o ascendència ciclònica). També la de Font-Tullot [1983], amb vint-i-tres tipus per a Espanya i Portugal, i a partir del plantejament segons el qual les característiques atmosfèriques de la península Ibèrica depenen de l'estat de circulació general (Font-Tullot estableix un primer nivell d'agrupació basant-se en l'expansió i la contracció del vèrtex

circumpolar, per passar a una discriminació final determinada pel grau d'ondulació del vèrtex). També trobem molt estesa la de [Martín-Vide \[1990\]](#), que identifica setze situacions sinòptiques que amb major o menor freqüència afecten la península Ibèrica i les Balears, i que es poden agrupar en tres categories principals a partir de la comparació dels mapes de pressió en superfície i de geopotencial a 500 hPa (advectives, no advectives, i entre les dues anteriors).

2.1.5. CLASSIFICACIONS CENTRADES ALS PIRINEUS I ANDORRA

Com que els resultats presentats en aquest treball (vegeu secció 5) se centren principalment en la serralada pirinenca i sobretot, al principat d'Andorra, hem volgut citar dos exemples de treballs centrats en aquesta zona.

D'una banda, la classificació de [Pédelaborde \[1959\]](#), on es fa una descripció de la circulació sinòptica amb relació als Pirineus francesos i relacionada amb dades observades a estacions meteorològiques (utilitza tres anys de dades, 1951-1953, amb la qual cosa no podem parlar de catàleg de circulació). El mateix autor, tenint en compte la longitud del període d'anàlisi, ja parla d'estudi preliminar en espera que més endavant se'n pugui fer un treball més exhaustiu. Pédelaborde, basant-se en la circulació a 500 hPa i en superfície, estableix deu tipus de circulació per al sector sinòptic pirinenc basats en un seguit de trajectòries de caire ciclònic més freqüents sobre l'Europa occidental i el tipus de massa d'aire associada (FIGURA 2.2). El posterior anàlisi de dies concrets associats a les situacions, acompanyats dels valors termopluiomètrics observats en algunes estacions meteorològiques i del càlcul de les freqüències estacionals completa l'aportació de Pédelaborde. Volem destacar que en aquest treball es fa notar la importància del relleu (que exerceix un *rôle tyrannique* segons l'autor) a l'hora de governar tots els processos aerològics associats als diferents tipus de circulació.

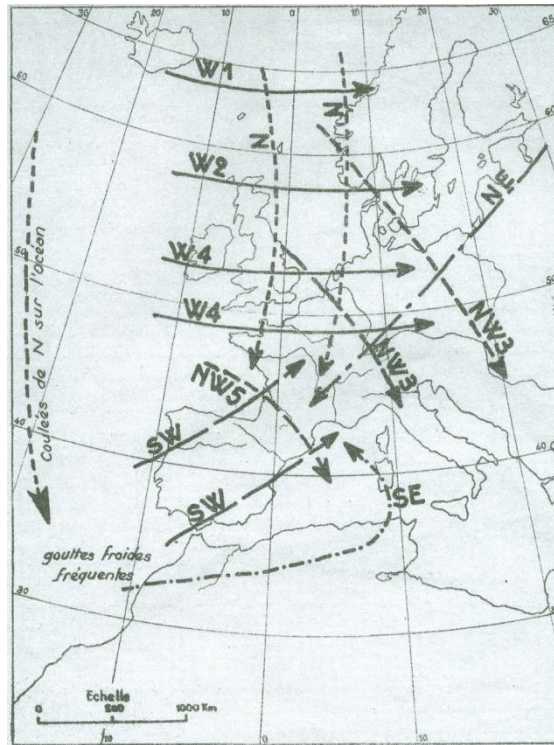


FIGURA 2.2: Principals trajectòries ciclòniques establertes per Pédelaborde a l'Europa occidental, i de les quals destaca algunes com a més rellevants pels Pirineus: nord quan és prou vigorós (N), nord-oest sobre la trajectòria 3 quan és prou intens i meridional, (NW3), nord-oest sobre la trajectòria 5 (NW5), tipus ciclònic sobre la trajectòria 4 característic de l'hivern i sovint només cua de front (W4), tipus ciclònic del sud-oest, que segons quin sigui afecta sectors més orientals o occidentals de la serralada (SW), i tipus ciclònic del sud-est, rellevant al sector de la serralada més proper a la Mediterrània (SE).

El segon exemple és la classificació de [Raso \[1999\]](#), que descriu deu situacions principals de circulació en superfície i a 500 hPa pel cas del principat d'Andorra (quatre considerades convectives i sis, advectives), tractant el tipus de massa d'aire associada, la freqüència anual del patró establert i les temperatures associades als principals observatoris andorrans. Concretament, les situacions convectives corresponen a l'anticicló centrat, depressió centrada, anticicló en superfície i depressió en altura, pantà baromètric, mentre les advectives són la zonal o temperada, l'advecció subzonal freda, la septentrional (FIGURA 2.3), la nord-oriental o continental europea, l'advecció continental africana, i la subzonal càlida.

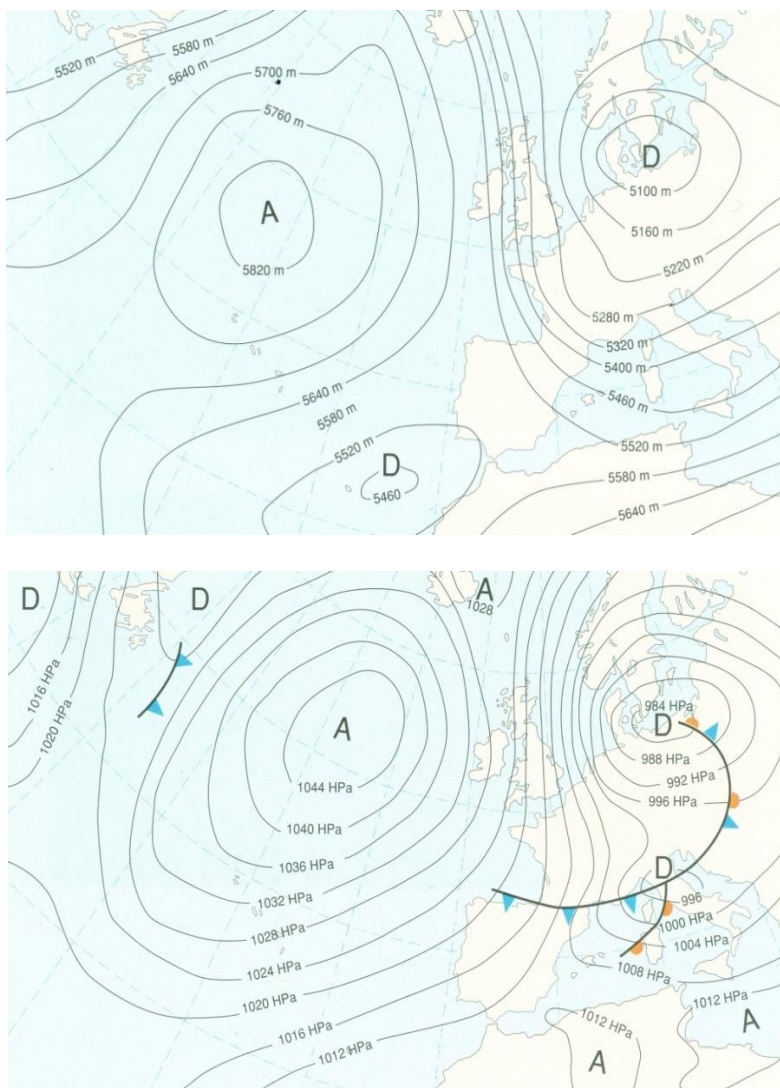


FIGURA 2.3: Cartografia sinòptica associada a l'advecció septentrional segons el catàleg de Raso [1999] centrat a Andorra. A dalt, geopotencial a 500 hPa, a sota el mapa d pressió en superfície. Segons l'autor, aquesta situació s'estima que ocorre per sobre del 7% a l'hivern, mentre que la seva presència estival pot ser inferior al 2% .

2.2. MÈTODES MIXTOS

Amb l'increment vertiginós de la capacitat de càlcul dels ordinadors al llarg de la segona meitat del segle XX i les possibilitats que això aporta en el terreny de les CTC, alguns investigadors han desenvolupat mètodes automàtics que es basen en mètodes manuals ja existents; és a dir, parlaríem de l'objectivació de mètodes manuals com alguns dels anteriorment exposats. Això fa que inicialment s'estableixin un criteris previs de referència per a la classificació; criteris que es poden basar en l'ús de criteris matemàtics/geomètrics, en la mateixa experiència del climatòleg, o en la generació de prototipus de circulació.

2.2.1. MÈTODE DE JENKINSON I COLLISON

Una de les aproximacions, i pel que fa a la bibliografia, menys discutides, és la classificació objectiva de patrons de circulació atmosfèrica de [Jenkinson i Collison \[1977\]](#), que citarem com a JC. Inspirada en la de caire subjectiu de [Lamb \[1972\]](#) sobre les illes Britàniques vista a l'apartat 2.1.1, es basa en l'ús d'índexs de circulació a partir del càlcul dels valors de flux i vorticitat geostrofica als punts d'una malla regular de pressió en superfície centrada sobre un sector en estudi (vegeu figures 2.4 i 2.5 amb el *grid* de treball i la formulació). Amb el càlcul del flux caracteritzem sobre la nostra zona d'estudi la direcció general del vent per veure si tenim una situació advectiona i de quina component, o en canvi, si tenim una situació centrada. Amb la vorticitat, podem discriminar entre situacions ciclòniques o anticiclòniques. La metodologia també preveu la possibilitat de casos no classificables, un desavantatge al qual s'afegiria la impossibilitat d'emprar aquest mètode per a altres variables i nivells atmosfèrics. La mitjana dels mapes associats a un tipus ens permetria cartografiar el patró.

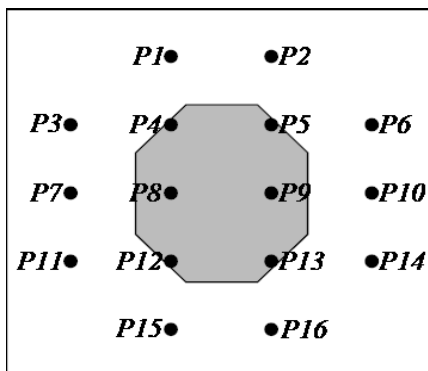


FIGURA 2.4: Distribució recomanada dels punts de malla en relació amb el sector objecte d'estudi en el mètode de Jenkinson i Collison. Els punts estan numerats amb relació a la formulació. Com podem veure, la zona d'estudi (àrea en gris) ha d'estar centrada respecte a la malla.

FIGURA 2.5: Formulació del mètode de Jenkinson i Collison.

(P1-P16 estan referits als punt de *grid* emprats que apareixen a la FIGURA 2.4)

FLUX DE L'OEST

$$W=0,5(P12+P13)-0,5(P4+P5)$$

FLUX DEL SUD

$$S=A[0,25(P5+2*P9+P13)-0,25(P4+2*P8+P12)]$$

on $A=1/\text{COS}(\text{Latitud})$

FLUX RESULTANT

$$F=(S^2+W^2)^{0,5}$$

DIRECCIÓ DEL FLUX

$$\text{DIR}=\text{TAN}^{-1}(W/S) \text{ (afegir } 180^\circ \text{ si } W>0 \text{ o } S>0)$$

VORTICITAT DE COMPONENT OEST

$$ZW=B[0,5(P15+P16)-0,5(P8+P9)]-C[0,5(P8+P9)-0,5(P1+P2)]$$

on $B=\text{SIN}(\text{Latitud})/\text{SIN}(\text{Latitud}-\text{ResolucioGrid})$ on $C=\text{SIN}(\text{Latitud})/\text{SIN}(\text{Latitud}+\text{ResolucioGrid})$

VORTICITAT DE COMPONENT SUD

$$ZS=D[0,25(P6+2*P10+P14)-0,25(P5+2*P9+P13)-0,25(P4+2*P8+P12)+0,25(P3+2*P7+P11)]$$

on $D=0,5(\text{COSLatitud})^2$

VORTICITAT RESULTANT

$$Z=ZW+ZS$$

SI $|Z|<F$, EL FLUX ÉS DIRECCIONAL PUR (N, NE, E, SE, S, SW, W, NW)

SI $|Z|>2F$, EL CAS SERÀ CICLÒNIC C SI $Z>0$ O ANTICICLÒNIC A SI $Z<0$

SI $F<|Z|<2F$, EL CAS SERÀ HÍBRID (PER EXEMPLE AE)

SI $F<6$ I $|Z|<6$, CAS NO CLASSIFICAT

Dessouky i Jenkinson [1975] ja havien emprat aquesta aproximació sobre Egipte amb dades diàries i mensuals, relacionant els resultats amb determinades situacions meteorològiques importants per a aquell sector. Jones *et al.* [1993], amb la intenció d'actualitzar la classificació de Lamb, i també per donar-hi un format numèric que en faciliti l'ús en estudis relacionats amb els models de circulació general (MCG), empra la formulació de JC, sent en realitat l'article que recupera, dona a conèixer, i avalua aquesta metodologia. Diversos treballs posteriors al de Jones *et al.* [1993], sobretot plantejats des de la visió de la REES o *downscaling* estadístic, han utilitzat aquesta aproximació per a catalogar la circulació. Així, Conway *et al.* [1996] estableixen relacions amb precipitació diària a les illes Britàniques, mentre que Kilsby *et al.* [1998], i centrant-se en Anglaterra i Gal·les, generen models de regressió amb intenció predictiva de la precipitació, i són part de les variables independents els índexs de circulació obtinguts amb JC. Per la seva banda, Linderson [2001] ho aplica al sud d'Escandinàvia, volen profunditzar en el comportament de la precipitació i la temperatura a escala local. Per a Portugal, Trigo i DaCamara [2000] estableixen el comportament de la circulació atmosfèrica i la precipitació entre 1946 i 1990, realitzant una climatologia d'aquest meteor respecte a cada patró, i també analitzant-ne la variabilitat al llarg del període d'estudi. D'altra banda, a través del projecte europeu TROTREP (*Tropospheric Ozone and its Precursors*) s'establien els índexs de circulació diaris (a les 00, 06, 12 i 18 h UTC) i els corresponents catàlegs de mapes per a Europa central amb el mètode de JC [Goodes, 2001] amb l'objectiu de relacionar la circulació atmosfèrica i el comportament de l'ozó en superfície en aquell sector. Força més recent és el treball de Demuzere *et al.* [2009], que empra aquesta CTC per abordar la problemàtica de les concentracions de O₃ i PM₁₀ a Holanda.

Per al sector de la península Ibèrica, Goodes i Palutikov [1998] relacionen la precipitació diària per a la conca del riu Guadalentín (Múrcia) amb els diferents patrons de circulació atmosfèrica obtinguts amb JC, per finalment, a través d'un model de circulació general (MCG), obtenir escenaris futurs pel que fa al comportament de la precipitació a la zona d'estudi. Posteriorment, Spellman [2000] realitza la classificació per a tota la península Ibèrica, caracteritza el comportament anual i estacional dels patrons establerts, i els relaciona amb la precipitació hivernal de diferents

observatoris. A les seves conclusions, Spellman fa notar alguns dels problemes de la metodologia.

Així, fruit de les dimensions geogràfiques, les seves característiques físiques i, en definitiva, la variabilitat climàtica peninsular, les conseqüències meteorològiques sobre el territori seran força diferents en un mateix patró, també pel que fa al flux, amb la qual cosa unes determinades situacions no es compliran arreu alhora. D'altra banda, Spellman denota les dificultats en caracteritzar la circulació en superfície quan, en determinats moments de l'any, certes conseqüències meteorològiques com les pluges intenses responen més a estructures presents en altura que no a configuracions superficials. En aquest sentit, [Martín-Vide \[2001\]](#), i reforçant les conclusions d'Spellman, aprofundeix sobre les limitacions del mètode JC per al cas de la península Ibèrica:

- La possible ocurrència de dos tipus de patrons alhora fruit de les característiques físiques peninsulars, especialment la seva extensió geogràfica.
- L'important nombre de dies no classificats, sobretot a causa de la impossibilitat d'establir valors de flux o vorticitat en situacions de gradient molt dèbil. Serien casos especialment estivals, com els dies amb pantà baromètric.
- Una excessiva freqüència estival de les situacions depressionàries. La causa seria la presència de la baixa tèrmica, que requeriria el reforç dels mapes en altitud per a una millor caracterització dels casos.
- La poca significació dels valors de pressió atmosfèrica en superfície amb relació a la presència de la precipitació. A la franja mediterrània, moltes situacions de pluges intenses es donen amb valors marcadament anticiclònics.

Una classificació derivada d'aquesta metodologia de JC és el treball de caracterització de la circulació atmosfèrica peninsular dut a terme a l'*Instituto Nacional de Meteorología* espanyol [[Martín Herreros, 2001](#)], que combina els índexs calculats amb el mètode de JC amb els desenvolupats per Murray i Lewis als anys 60-70. Segons

Martín-Herreros, Murray i Lewis desenvolupen quatre índexs (P, S, M i C) que ens donen indicacions sobre el grau de bloqueig de la situació, el grau de circulació de component sud, la circulació meridiana (nord o sud), i el grau de ciclogènesi.

2.2.2. OGWT – OBJECTIVACIÓ DEL HESS-BREZOWSKY GROSSWETTERLAGEN ALEMANY

Aquest mètode proposat per James [2006, 2007] treballa amb el camp de pressió en superfície i el geopotencial a 500 hPa, exigint a més una persistència mínima de tres dies, tal com passa al mètode del *Grosswetterlagen* (GWL - vegeu secció 2.1.2). El procés d'objectivació del mètode de Hess-Brezowsky seguit ha estat el següent:

1. Utilitzant dades diàries de reanàlisi en format de malla (ECMWF ERA40 1957-2002), s'han calculat per a cadascun dels tipus del catàleg GWL els mapes mitjans a SLP i 500 hPa, separat per hivern i estiu, i segons el catàleg manual ja existent.
2. S'estableix per a cada GWL quin sector de l'àrea d'estudi és més rellevant. Això es fa a partir de la distribució espacial de la desviació estàndard pels camps d'anomalies normalitzades de cada GWL, de manera que permet establir una àrea de doble pes a l'hora d'establir correlacions en passos posteriors.
3. Es calcula el catàleg objectiu a partir de correlacions entre els GWL calculats anteriorment i els mapes diaris suavitzats (se'ls aplica una mitjana mòbil de cinc dies als camps a classificar), i s'assigna a aquell GWL amb major coeficient de correlació, independentment que la correlació sigui baixa (no hi ha casos no classificables).
4. Finalment, com que cada tipus de GWL ha de tenir una durada no inferior a tres dies, seqüències de menor longitud són filtrades i se substitueixen per l'alternativa més adient. Aquesta alternativa, relativament complexa, es detalla a James [2007].

Veiem, doncs, que, un cop establerts els GWT, el mètode és completament objectiu, alhora que planteja una possible forma d'objectivar altres classificacions manuals existents a altres regions. James [2007] la presenta com una tècnica, tot i que un xic complexa, suficientment robusta com per ser aplicada a les sortides dels models de circulació general (MCG). En aquest sentit, utilitzant el reanàlisi de l'ERA40 i les sortides en *ensemble* del model global del *Hadley Centre* anglès, l'autor analitza la freqüència i la variabilitat dels patrons de circulació a escala sinòptica per a la regió d'Europa-Atlàntic Nord. James conclou que, tot i el que apunten els resultats, l'alta variabilitat interanual força a utilitzar períodes més llargs de dades abans no es puguin extreure conclusions sòlides sobre els canvis de la circulació atmosfèrica sobre Europa en un clima canviant.

De moment, no tenim referències d'altres treballs que emprin aquesta metodologia o l'adoptin a altres sectors geogràfics.

2.2.3. LA CLASSIFICACIÓ PER PROTOTIPUS

Aquesta aproximació ha estat desenvolupada per Beck [2000] i es basa en l'ús de deu patrons de circulació predefinitos o *prototypes* a partir de la generalització dels vint-i-nou establerts al *Grosswetterlagen*. Recordem que els patrons del GWL es defineixen per la posició geogràfica dels centres d'acció més importants i per la localització i extensió de les línies frontals, de manera que Beck assumeix que els pot caracteritzar segons el grau de zonalitat, meridianitat i vorticitat dels camps de pressió en superfície sobre la zona en estudi. El passos a seguir són [Beck et al., 2007] (FIGURA 2.6):

Es defineixen tres prototipus de pressió en superfície per a la zona d'estudi. Han d'incloure vint-i-cinc punts de malla, i han de ser representacions idealitzades d'un flux oest-est, est-oest i d'una baixa centrada.

1. Es calculen les correlacions espacials entre els tres prototipus i els casos a classificar. Aquestes correlacions espacials s'adopten com a coeficients de zonalitat (*Z*), meridianitat (*M*) i vorticitat (*V*) respectivament.

2. Els deu tipus de circulació basats en el GWT s'obtenen a partir de les diferents combinacions d'aquests coeficients de correlació. Així, l'anticicló (A) i la baixa (B) centrada s'estableixen amb els valors màxims del coeficient V (negatiu i positiu, respectivament). La resta de tipus es defineixen a partir de Z i M , establint el tipus d'advecció al centre de la zona d'estudi:

$$\begin{aligned}
 Z=1/M=0 &\rightarrow \text{oest (W) a est (E)} \\
 Z=0,7/M=0,7 &\rightarrow \text{sud-oest (SW) a nord-est (NE)} \\
 Z=0/M=1 &\rightarrow \text{sud (S) a nord (N)} \\
 Z=-0,7/M=0,7 &\rightarrow \text{sud-est (SE) a est (E)/nord-oest (NW)} \\
 Z=-1/M=0 &\rightarrow \text{est (E) a oest (W)} \\
 Z=-0,7/M=-0,7 &\rightarrow \text{nord-est (NE) a sud-oest (SW)} \\
 Z=0/M=-1 &\rightarrow \text{nord (N) a sud (S)} \\
 Z=0,7/M=-0,7 &\rightarrow \text{nord-oest (NW) a sud-est (SE)}
 \end{aligned}$$

3. Tots els casos que no s'han classificat com a A o B s'assignen, doncs, a algun dels vuit tipus advectionals; assignació que es realitza a partir de la distància euclidiana mínima resultant dels coeficients Z i M amb relació als tipus preestablerts.
4. Una nova divisió dels vuit tipus direccionals es pot fer en funció que les condicions siguin anticiclòniques o ciclòniques i en funció de quin sigui el signe del coeficient V .

Aquesta metodologia té un recorregut molt més breu que el de JC i ha estat encara poc aplicada. De moment, destaca el seu ús en l'avaluació de la importància dels canvis en la freqüència i en la dinàmica dels tipus de circulació per al sector d'Europa i l'Atlàntic Nord en el comportament del clima centreeuropeu per al període 1780-1995 [Beck *et al.*, 2007].

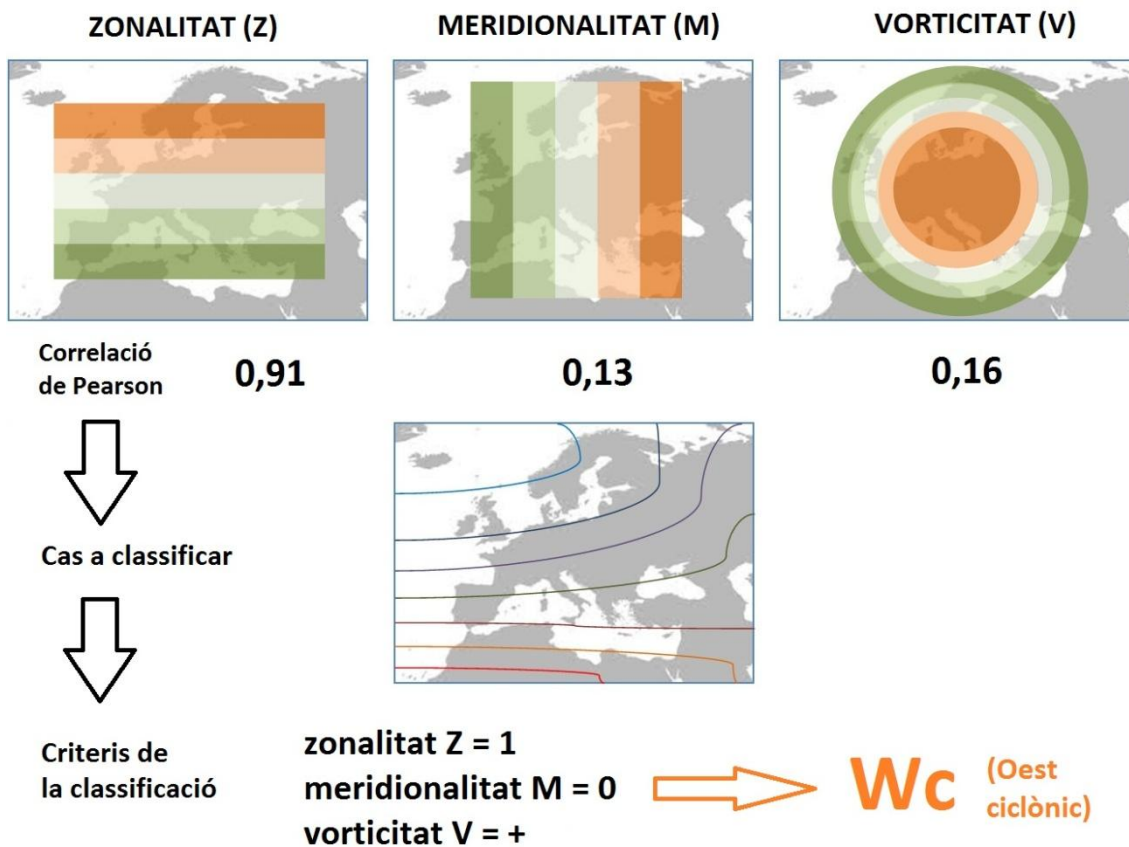


FIGURA 2.6: Esquema descriptiu sobre el procediment a seguir per a aplicar el mètode dels prototipus. Les correlacions d'un cas respecte als tres models (zonal, meridiana i centrada) permeten classificar la circulació atmosfèrica en deu tipus. Figura adaptada amb permís de C. Beck i A. Philipp.

2.3. MÈTODES BASATS EN CORRELACIONS

Són una de les tècniques més simples i importants desenvolupades pels científics de l'atmosfera, i es basen en les similituds establertes per aquesta tècnica estadística per caracteritzar els diferents grups sinòptics [Barry i Carleton, 2001]. Segons Yarnal [1993], la seva simplicitat, i el fet que la màquina reproduceix el mètode manual i genera patrons de circulació similars en categories discretes, va fer aquestes tècniques força populars.

2.3.1. MÈTODE DE LUND [1963]

Basat en l'aplicació de la correlació de Pearson entre els *grids* de tots els mapes a tractar, es busca aquell cas que presenti un major nombre de mapes, del total de la mostra a classificar, amb un grau de similitud elevat. Aquest mapa, un cop establert, passa a ser el mapa clau (*Key-day*) per al primer grup o patró, que integrarà tots aquells casos que correlacionin amb aquest mapa clau per sobre d'un llindar concret (l'autor proposa una correlació de 0,7). Un cop s'ha creat aquest primer grup, és repeteix el procediment per buscar un nou mapa clau amb els casos (mapes) restants; seqüència que es va repetint fins que no queden casos a classificar o només queden dies no classificables (no correlacionen elevadament amb cap cas).

La fórmula per establir les correlacions és la següent:

$$r_{xy} = \frac{\sum_{i=1}^N [(x_i - \bar{X})(y_i - \bar{Y})]}{\sum_{i=1}^N (x_i - \bar{X})^2 \sum_{i=1}^N (y_i - \bar{Y})^2}$$

On x_i representa la variable (estandarditzada amb la mitjana i desviació del mapa al qual correspon) als N punts del primer mapa a comparar, mentre que y_i és la mateixa variable (estandarditzada) als mateixos punts del segon mapa. Per la seva banda, \bar{X} i \bar{Y} són les mitjanes dels N punts del *grid*.



FIGURA 2.7: Imatge de l'ordinador IBM 704. Lund, en el seu article de 1963, cita explícitament la utilització d'aquesta màquina, cosa que denota l'esforç que requeria una classificació automàtica. El mateix càlcul, a dia d'avui, el podem fer amb la potència d'un telèfon mòbil. Imatge extreta de <http://www.retroist.com/2010/03/08/retroist-access-monday-ram-%E2%80%93-the-ibm-704/>

En el seu treball, Lund [1963] classifica (amb un ordinador IBM 704, FIGURA 2.7) els mapes diaris de pressió en superfície de cinc hiverns a partir d'un *grid* de vint-i-dos punts de malla sobre el nord-est dels Estats Units (FIGURA 2.8). L'autor, a partir del llindar de correlació de 0,7, estableix deu mapes clau, i a continuació, relaciona els resultats amb dades de precipitació, radiació solar i nevades a Boston. Com explica Yarnal [1993], Lund va ensenyar els deu patrons obtinguts a experimentats predictors del temps, i les consideracions fetes per aquests especialistes coincidien amb els patrons establerts pel mètode per a les diferents conseqüències meteorològiques. Com a exemple de treball amb aquest mètode tenim el de Bischoff i Vargas [2003], que empren l'aproximació de Lund per establir lligams entre la circulació atmosfèrica al GPH de 1.000 i 500 hPa amb períodes extrems pel que fa a temperatura i precipitació al sector meridional de l'Amèrica del Sud.

Amb tot, cal dir que diversos autors han destacat certes mancances en el mètode de Lund, i per exemple Barry i Carleton [2001] denoten que aquest mètode només classifica normalment entre un 60 % i un 80 % dels mapes analitzats.

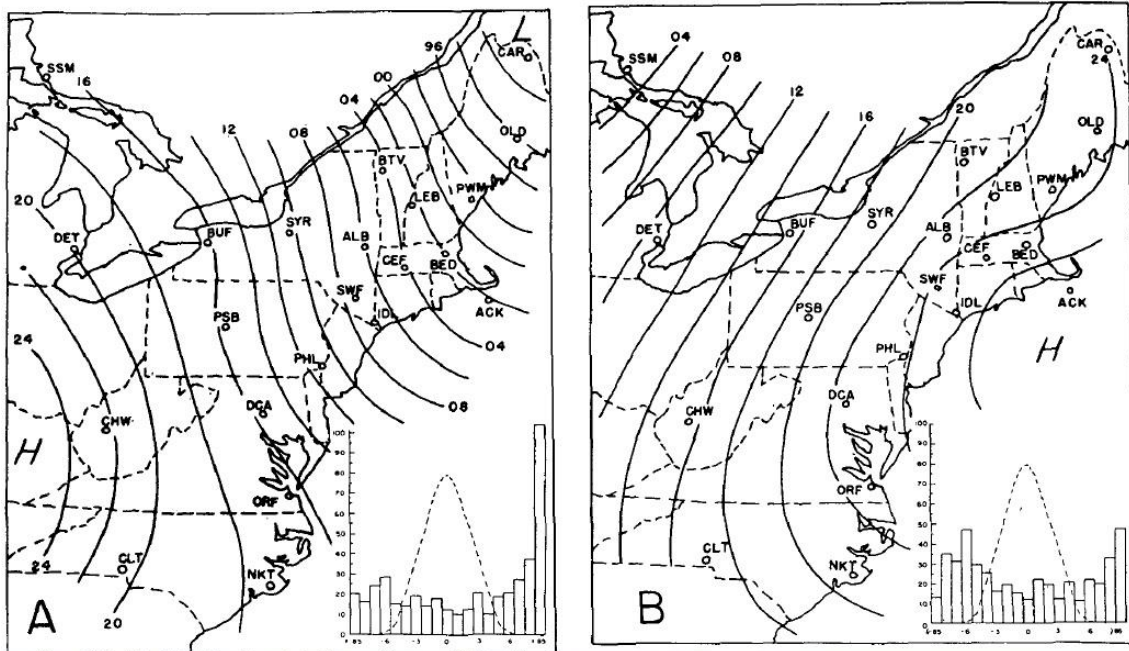


FIGURA 2.8: Reproducció dels dos tipus principals obtinguts per Lund en el seu treball de 1963. S'hi pot observar la distribució irregular de la malla de treball, així com el gràfic de freqüències dels coeficients de correlació obtinguts.

2.3.2. MÈTODE DE KIRCHHOFER [1973]

Intentant optimitzar aquest tipus de tècniques, Kirchhofer va establir una segona aproximació per classificar els patrons a partir de correlacions entre els mapes, augmentant la capacitat de classificació a un 90 % del total de la mostra [Yarnal, 1984]. En comptes d'utilitzar la correlació directa entre els *grids*, l'autor estableix les relacions entre els mapes a partir de la suma dels quadrats de les diferències segons la següent formulació que hem extret de Yarnal [1993] i Barry i Carleton [2001]:

$$S = \sum_{i=1}^N (Z_{xi} - Z_{yi})^2$$

On S seria el valor de puntuació de Kirchofer, Z_{xi} és el valor estandarditzat del punt i el dia x , Z_{yi} el valor estandarditzat del punt i el dia y , i N el nombre de punts de malla.

Els valors líndars del valor S per a classificar els mapes en un grup haurien de ser com a màxim $1.0N$. Aquest valor de $1.0N$ correspondria a un líndar de correlació de 0,5 pel mètode de Lund, segons va establir Willmott *et al.*, 1987 [extret de Barry i Carleton 2001]. Cal dir que també és possible calcular subvalors de similitud per a determinats sectors del *grid* tractat, creant valors de puntuació per a les columnes (S_c) i les files (S_f) amb la mateixa fórmula abans presentada, cosa que permet identificar diferències significatives en casos que podrien agrupar-se si tractem tot el mapa alhora [Yarnal i White, 1987]. Valors de S_c i S_f d'entre $1,0n$ i $1,4n$ (sent n el nombre de punts de *grid* a la columna o fila respectivament) serien, com també citem més endavant, els recomanables. Brinkmann [2000] també considera aquesta possibilitat, que dos mapes poden ser molt similars, però difereix de forma marcada en una petita part de la finestra tractada (presència d'elements importants de circulació a petita escala), de manera que afecta clarament l'homogeneïtat interna dels grups finalment obtinguts. Aquest autor aplica la metodologia de Kirchoffer a sis subfinestres del *grid* utilitzat, havent millorat els resultats obtinguts anteriorment en la classificació de tot el *grid* [Brinkmann, 1999].

Pel que fa a l'estandardització de Z_{xi} i Z_{yi} , aquesta seria com la de Lund, sent la mitjana i la desviació del dia x o y al qual pertany el punt i respectivament. Si treballem amb subfinestres, no hem de caure en l'error d'estandarditzar respecte a tota la malla, ja que, com diu Blair [1998], hi introduïm biaixos que altres autors no han previst, i hem d'estandarditzar, doncs, respecte a la mitjana i la desviació de la subfinestra corresponent.

També és interessant tenir en compte l'aportació de Kaufmann *et al.* [1999], que per avaluar la significació dels patrons identificats amb la tècnica de Kirchofer varien sensiblement aquesta estandardització, ja que treballen amb temperatures màximes; variable que consideren que no es pot tractar com espacialment estacionaria com la pressió en superfície (és a dir, que la variabilitat en el territori segueix uns patrons molt poc regulars i amb canvis marcats, com també el seu ritme anual). Així, aquests autors

estandarditzen utilitzant la mitjana i desviació del punt i el dia x o y , respecte als valors del dit punt i al llarg de tota la sèrie que empren per a l'estudi (parlaríem, doncs, d'estandardització temporal). D'altra banda, en aquest mateix estudi, [Kauffmann et al. \[1999\]](#) denoten la influència en el mètode de classificació de Kirchhofer del llinar utilitzat per a la classificació, així com de la mida de la mostra. En aquest sentit, [Barry i Carleton \[2001\]](#), i a partir de les consideracions dels treballs de [Key i Crane \[1986\]](#) i [Yarnal i White \[1987\]](#), plantegen les següents recomanacions per als mètodes de correlacions de Lund i Kirchhofer:

- Els patrons de circulació s'han d'identificar amb mostres de com a mínim mil casos (*grids*) per poder obtenir resultats estables.
- El llinar de similitud triat entre els casos ha de ser prou rigorós per evitar massa diversitat interna en els grups establerts. Un llinar de correlació de 0,7 sembla suficient quan treballem amb dades de pressió en superfície, mentre que el llinar de 0,9 sembla més adient per mapes de 700 hPa. En el cas de la S de Kirchhofer, llinars entre 1 i 1,4 vegades el nombre de punts per fila (o columna) de la malla emprada (recordem els valors de puntuació S_f i S_c), semblen produir resultats satisfactoris des del punt de vista del total de casos classificats, encara que pot generar gran quantitat de patrons (al voltant de cinquanta).
- El nombre òptim de patrons a establir per descriure les característiques sinòptiques d'una àrea depèn de diverses consideracions. Una és l'efectivitat que tinguin les categories establertes a l'hora de descriure el clima de la zona d'estudi. Amb tot, l'experiència suggereix que entre vint i trenta patrons és un nombre difícil de tractar de cara a la seva interpretació i estudi. Més que intentar reduir el nombre de grups disminuint els llinars de similitud, sembla preferible agrupar les categories inicials mitjançant mètodes objectius com ara els mètodes d'agrupació (vegeu apartat 2.5) o amb un criteri subjectiu que s'ajusti a la problemàtica en estudi.

En definitiva, tres decisions subjectives de l'investigador afecten principalment els resultats. Així, la mida (espacial i temporal) i resolució de malla emprada, el llinar de

similitud per a les agrupacions i el nombre màxim de categories acceptat (com el mínim de dies inclosos als grups per considerar-los correctes) [Yarnal, 1993] afecten la qualitat dels resultats, i sobretot en aquell punt que tota classificació de mapes ha d'optimitzar: una alta homogeneïtat interna dins els grups i alhora una màxima heterogeneïtat externa entre els grups. La discussió sobre aquests aspectes no està resolta. Amb tot, Huth *et al.* [2008] fan notar la bona capacitat d'aquests mètodes a l'hora de separar els grups.

Finalment, i pel que fa a altres treballs que hem consultat i que han utilitzat l'aproximació del mètode de correlacions en alguna de les variants vistes i per a la classificació de patrons de circulació, tenim l'exemple de Yarnal [1984], que caracteritza la circulació atmosfèrica diària a 500 hPa amb relació als canvis de massa de la glacera Sentinel (Colúmbia Britànica, el Canadà), o el d'El-Kadi i Smithson [1996, 2000], que amb aquestes tècniques aborden el tractament de la persistència i les transicions dels patrons de circulació atmosfèrica en superfície amb relació al sector sinòptic de les illes Britàniques.

Els mètodes de correlacions es presenten, doncs, com un bona aproximació per a l'establiment de catàlegs de patrons de circulació atmosfèrica. Amb tot, hem de valorar el fet que no sempre estiguem treballant amb tota la mostra alhora (a mesura que avancem en el procediment d'agrupació, ja no utilitzem els casos classificats), com també les problemàtiques tractades per alguns autors pel que fa a la sensibilitat del *grid* i del llindar de similitud emprats. Huth *et al.* [2008] també denoten la tendència a crear grans grups junt amb altres de molt petits, alhora que molts casos no classificats. Nosaltres considerem un dèficit la manca d'algun mecanisme intern en el mètode que orienti sobre el nombre de grups adient a considerar.

2.4. MÈTODES BASATS EN L'ANÀLISI DE COMPONENTS PRINCIPALS

L'anàlisi de components principals (a partir d'ara ACP – en anglès *Principal Component Analysis*) és una tècnica estadística multivariant (treball en N dimensions) basada en els principis matemàtics dels vectors propis, i que trobem implementada en molts programaris de caire estadístic (SPSS, SAS, R...). Vegem-ne algunes definicions per després entendre millor com podem utilitzar aquesta tècnica per a la CTC:

[Jolliffe, I.T. 1986] La idea central de l'anàlisi de components principals és la de reduir la dimensionalitat d'un conjunt de dades en les quals hi ha un important nombre de variables interrelacionades, alhora que retenim el màxim possible de la variància present en aquest conjunt de dades. Aquesta reducció s'assoleix amb la transformació a un nou conjunt de variables, els components principals, que no estan correlacionades, i que s'ordenen de manera que unes poques primeres retenen molta de la variabilitat present en totes les variables originals.

[Wilks, D.S., 2006] Normalment, les dades de camps atmosfèrics o geofísics presenten destacades correlacions entre les variables, i l'ACP ho resol amb una representació molt més compacta de la seva variació.

[Von Storch, H., Zwiers, W.; 2001] Les funcions ortogonals empíriques (*Empirical Orthogonal Functions*, EOF) és una tècnica multivariant que serveix per derivar els patrons dominants de variabilitat d'un camp estadístic.

(El concepte EOF s'utilitza també per identificar l'ACP. Amb tot, hi ha autors que consideren rellevant diferenciar entre EOF i ACP, ja que corresponen a dos mètodes semblants de càlcul dels vectors propis però basats en normalitzacions diferents [Jolliffe 1987, Mestas-Nuñez 2000]. En aquest sentit, fins i tot alguns autors han identificat avantatges d'una opció o l'altra quan treballem amb modes de circulació de baixa freqüència (vegeu apartat 2.4.1)).

En definitiva, l'ACP identifica i simplifica la informació comuna bàsica d'un conjunt de dades: volem detectar aquelles estructures espacials més importants (principals) i que

són necessàries perquè es doni la totalitat de configuracions espacials presents als mapes del temps que volem classificar. Aquestes morfologies més comunes que n'obtenim no són tipus de circulació, sinó cartografia no dimensional que reproduïx les formes bàsiques (estructures principals de circulació) que es donen a l'àrea d'estudi; estructures que amb una bona i coherent interpretació es poden relacionar amb els centres d'acció, la seva distribució, i la circulació atmosfèrica derivada. Són els **modos principals de variació**.

Anem a desgranar una mica més el procés introduint-hi conceptes bàsics necessaris per entendre l'aplicabilitat d'aquesta tècnica en Climatologia, i sobretot per anar aclarint com podem abordar la CTC.

- a. El principal objectiu de l'ACP com a eina estadística és el de **reduir la informació** per facilitar el treball amb grans volums de dades. Aquesta informació de partida presenta una estructura matricial on tenim unes variables (columnes) i unes observacions (files).
- b. La transformació de la matriu original a un nou format de **noves variables** anomenades **components o factors** es realitza a través d'una simple transformació lineal. Així doncs, hi haurà valors que permetin relacionar la nova informació de partida amb la informació original: són els coeficients o saturacions (*loadings*) i les puntuacions (*scores*).
- c. Aquesta transformació lineal es basa en l'establiment de **correlacions** (o covariances) entre les variables originals (columnes de la matriu original).
- d. Els components (noves variables) creades són un conjunt de vectors ortogonals entre ells (els anomenats vectors propis) que optimitzen la representació de la informació original **eliminant la redundància** provocada per la interrelació entre moltes de les variables originals (FIGURA 2.9).
- e. L'**ortogonalitat** que presenten aquestes noves variables o components implica una correlació zero entre elles. Així, cada component portarà associat un

percentatge de variància explicada, que denotarà la seva **importància** en la més òptima i simple interpretació del total d'informació original.

- f. Per poder reduir la informació original se seleccionen unes poques de les noves variables o components. El grau d'informació associat a aquests **components principals retinguts** ve representat pel percentatge del total de variància explicada que sumen, i que ha de ser força elevat.

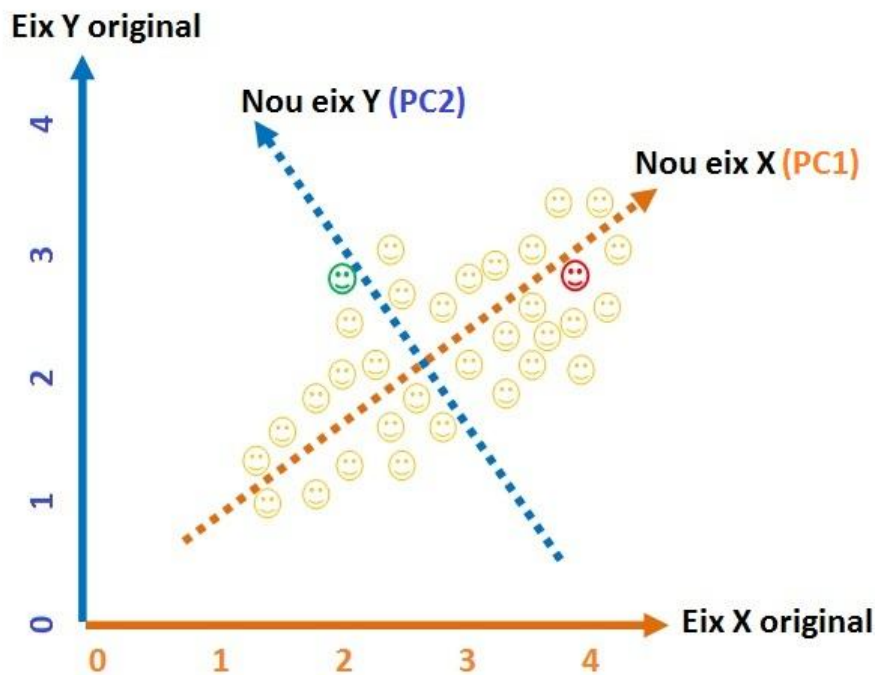


FIGURA 2.9. L'ACP consisteix a canviar els eixos de coordenades que fem per descriure les nostres dades i així optimitzar i simplificar la seva explicació. Observem com els nostres *smilies* vermell (3,9:2,9) i verd (1,9:2,9) presenten valors elevats, si tenim en compte les coordenades que tenim amb els eixos originals (format original de les dades), valors que no fan fàcil diferenciar-los. En canvi, quan apliquem l'ACP i forçem els nous eixos, les coordenades que ens descriuen els nostres objectes s'optimitzen. Ara l'*smilie* vermell (2:-0,1) està molt ben representat pel nou eix X (el que seria el component principal 1, PC1), mentre que l'*smilie* verd (-0,1:1) ara està molt ben representat pel nou eix Y (el que seria el component principal 2, PC2). Amb les noves coordenades, no és ara més fàcil diferenciar els dos *smilies*? Hem simplificat l'explicació dels nostres objectes ajustant-nos millor a la distribució de la seva variància.

- g. Si hem fet una correcta selecció dels components significatius, només treballarem amb el **senyal de la mostra**. La resta explica poca variància i representa el **soroll de la mostra**.
- h. Els components principals retinguts seran representatius dels **principals modes de variació de les dades**. Si treballem amb informació espacial, seran representatius dels principals patrons de distribució de la variable tractada a la zona d'estudi.
- i. Els **coeficients o saturacions (loadings)** i les **puntuacions (scores)** ens permeten veure les característiques (temporals i espacials en el nostre cas) dels components principals, i alhora veure en quin grau queda representat cada cas original (els casos reals) per aquests components principals; és a dir, pels modes de variació espacial.

Tot i que molts autors consideren aquesta aproximació per a l'establiment de patrons de circulació una tècnica objectiva, durant el procés s'han de prendre un seguit de decisions que poden afectar els resultats finals, alhora que també hi alguns aspectes metodològics a tenir presents:

A. LA MIDA DE LA MOSTRA A TREBALLAR:

Un aspecte a tenir present a l'hora de treballar amb l'ACP està relacionat amb la mida de la mostra (dimensions de la matriu). La regla bàsica genèrica que es considera en aquest punt és que hi ha d'haver més observacions que variables per obtenir resultats estables quan apliquem l'ACP. *Hair et al. [2000]* parlen d'un nombre d'observacions cinc vegades el nombre de variables.

Pel que fa només al nombre mínim d'observacions (files de la matriu, i dies en el nostre cas), *Kline [1994]* considera que cent és un valor mínim suficient (a partir de tres-cents serien mostres grans), encara que reconeix la marcada subjectivitat del

llindar establert. En canvi, **Hair et al. [2000]** recomanen que no es treballi amb menys de cinquanta, encara que també consideren preferible que siguin un centenar o més.

En el mateix sentit, és de gran interès en el camp de la Climatologia el treball **de Lana i Mills [1994]**, que, a partir d'una mostra de dimensions destacades (vint-i-sis variables i dues mil vuitanta-dues observacions), avaluen el nombre mínim d'observacions necessàries perquè els resultats de la classificació de tipus de temps que estan realitzant siguin estables. Els autors estableixen mostres aleatòries de diferents mides (cent, dos-cents, tres-cents... mil) amb ACP i mètodes de *clustering* (CL, vegeu apartat 2.5), i estudien l'evolució en l'estabilitat dels resultats a mesura que varia la mida de la mostra. Lana i Mills descarten nombres inferiors a cent observacions a causa de l'excessiva degeneració que presenten, i conclouen que entre tres-cents i quatre-cents observacions s'establiria el mínim necessari per obtenir uns resultats estables, cosa que ens pot estalviar importants esforços de càlcul segons l'ordinador amb què treballem i la mostra original que fem (FIGURA 2.10).

Observant, doncs, les consideracions que aporten les fonts consultades, volem plantejar aquests intervals (aproximat i genèric) de suficiència de la mida de la mostra:

- Mostra inferior a cinquanta observacions: insuficient
- Entre cinquanta i tres-cents observacions: mostra molt petita, suficient però amb cautela
- Entre tres-cents i mil observacions: mostra suficient
- Mostra superior a mil: mostra bona

Per la seva banda, **Von Storch i Hannoschöck [1985]** tracten les propietats estadístiques de l'anàlisi de components principals amb mostres petites. Aquests autors conclouen que, si la ràtio de nivells de llibertat de la mostra és 0,5 o menys, els valors propis obtinguts es presentaran esbiaixats.

Podem observar, doncs, que la mida de la mostra amb què treballem a l'ACP és un punt de marcada importància alhora d'obtenir resultats estables. Sobre la base

d'aquestes diverses consideracions, molts autors recomanen que, quan treballem amb mostres petites, tinguem en compte la desmesurada influència que puguin tenir algunes observacions individuals [Jolliffe, 1993]; és a dir, que tinguem en ment aquest fet a l'hora d'extreure'n conclusions.

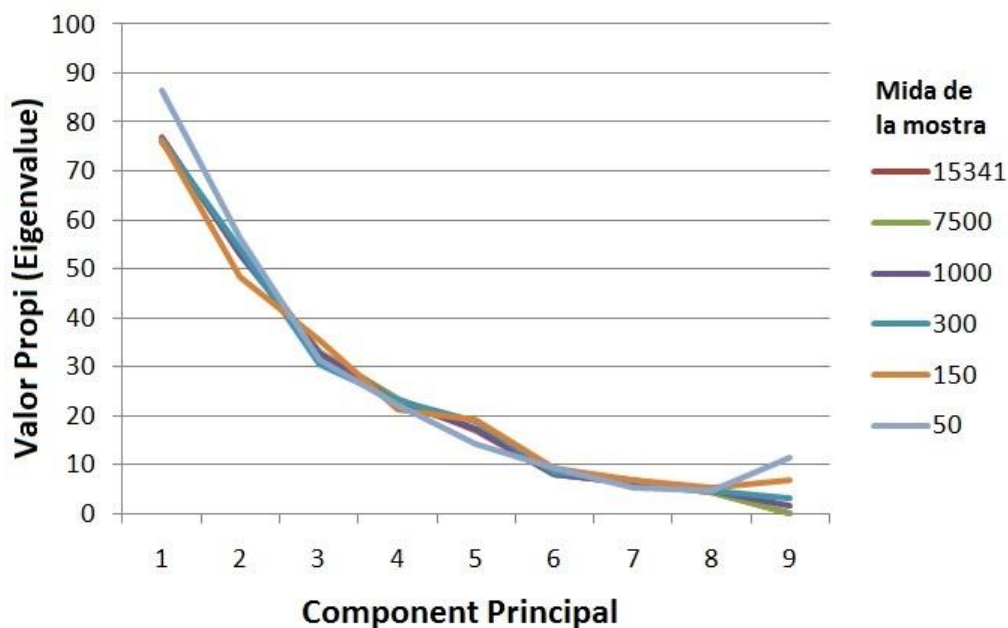


FIGURA 2.10.: Un simple gràfic de l'evolució dels valors propis d'una mateixa mostra, però amb diferents mides (nombre de casos), ens ensenya com va variant l'estabilitat dels resultats. Les mostres més grans (7500, 15341) ja no presenten diferències en els seus valors propis.

B. COM S'ESTABLEIX LA INTERRELACIÓ ENTRE LES VARIABLES

Si recordem, l'ACP busca els principals modes de variació de les nostres dades, de manera que obtinguem unes noves variables que ens permetin reduir el volum d'informació i ens facilitin la seva interpretació. La manera de fer-ho és establint relacions entre aquestes variables que tenim de partida (les columnes de la matriu), de manera que, a partir de mesurar el seu grau de similitud obtindrem els components o noves variables, una forma més òptima d'explicar la base de dades original. Aquesta mesura de semblança en la variació s'estableix a partir del càlcul de les covariàncies o de les correlacions entre les variables originals, tot i que hi ha altres possibilitats (els

productes creuats). A la bibliografia científica trobem que s'utilitzen les dues opcions, però anem a veure quins elements són els més importants per a tenir en compte en un cas o en l'altre:

COVARIÀNCIES:

Optar per la matriu de covariàncies ens permet reflectir millor les desviacions espacials de les dades. El desavantatge és que els patrons resultants tendiran a concentrar les estructures espacials rellevants a les àrees amb màxima variància en detriment de la seva situació real per qüestions climàtiques. Aquest fet pot ser força problemàtic en aquelles mostres que presentin grans contrastos de variància entre les seves variables, provocant a més que aquelles zones que tenen els valors més alts acabin influenciant excessivament en els primers components als resultats de l'ACP [Richman, 1981] [Richman, 1986] [Jackson, 1991] [Yarnal, B., 1993] [Jolliffe, I. T., 1990]. Pensem que, si hem optat per l'estructura tipus S (vegeu apartat 2.4.2), una variable (columna) és la successió temporal de valors de, per exemple, pressió en superfície sobre un mateix punt. Només cal pensar en un punt de latituds mitjanes (per exemple, les illes Britàniques), on al llarg de l'any es poden donar valors de pressió força baixos pel pas de depressions profundes, com també s'hi poden centrar, especialment a la meitat freda de l'any, anticiclons de marcada profunditat.

CORRELACIONS:

Amb l'ús d'aquesta opció, extraïem la mitjana de la variable per després dividir-la per a la seva desviació; és a dir, estandarditzem de forma automàtica les variables (mitjana 0 i desviació 1). Això permet, doncs, solucionar el problema de diferències marcades pel que fa les variàncies de les variables. En Climatologia aquest fet ens el podem trobar tant amb diferents variables que tinguin la mateixa unitat de mesura, com també amb variables que presentin diferents unitats ($^{\circ}\text{C}$, m/s, hPa...). La matriu de correlacions elimina aquesta variabilitat que hi pugui haver [Jackson, 1991], de manera que, a diferència de la matriu de covariàncies, no s'introdueixen desviacions en la localització dels centres d'acció [Richman, 1981]. D'altra banda, si treballem amb la matriu de tipus S (vegeu apartat 2.4.2) i optem per les correlacions, l'estandardització de les variables fa que les comparacions entre els mapes originals i els tipus obtinguts

refleixin la seva forma, però no la seva intensitat [Richman, 1981] [Yarnal, 1993], amb els possibles desavantatges que això pot comportar segons quins siguin els nostres objectius des d'un punt de vista climàtic.

C. MÈTODE PER TRIAR EL NOMBRE DE COMPONENTS PRINCIPALS A RETENIR

Amb l'ús de l'ACP obtenim unes noves variables, tantes com teníem inicialment (nombre de columnes) i que optimitzen la representació numèrica de la informació original. N'obtenim, doncs, un seguit de components que s'ordenen des d'una primera que és la que millor resumeix tota la matriu de dades original (té la major variància explicada, recordeu la FIGURA 2.9), fins a unes últimes que aporten molt poca o gairebé nul·la informació (valors molt baixos de variància explicada). A partir del principi que un pocs components, les primeres, ens permeten treballar amb un alt grau de la informació original (totalitzen una gran part de la variància total), l'ACP planteja la possibilitat de reduir la mostra pel que fa a la mida, facilitant així el treball i la interpretació. Dit d'una altra manera, que seleccionarem o retindrem aquelles noves variables o components que ens aportin informació suficient per interpretar la mostra original (representarien el senyal de la mostra), mentre que deixariem de banda aquelles components restants que portin associada molt poca explicació (que se les considera representatives del soroll). Fer una bona selecció d'aquestes noves variables o components realment representatives i significatives, i poder així deixar de banda el soroll que no ens interessa per establir els principals patrons de variació de les dades és un punt importantíssim, ja que pot afectar de manera destacada els resultats finals [Jolliffe 1993].

A la bibliografia científica trobem gran varietat de mètodes per realitzar aquesta selecció de les noves variables més importants o components principals. Aquests mètodes presenten variada complexitat, i Preisendorfer *et al.* [1981] i Jackson [1991] fan una extensa exposició de moltes de les opcions existents.

No s'han trobat avantatges clars dels mètodes més complexos sobre els més simples o descriptius [Jolliffe, 1993], i és per això que la gran majoria d'autors opten pels

mètodes senzills d'implementar. A continuació en descriure'm els més importants, sobretot pel que fa al seu major ús en publicacions de Climatologia:

CRITERI DEL PERCENTATGE DE VARIÀNCIA EXPLICADA:

És l'aproximació més simple de les emprades per seleccionar els components a retenir. És basa a establir un llindar mínim de variància total acumulada que han d'explicar els components retinguts. No hi ha cap valor establert, però alguns autors recomanen, per a les ciències naturals, un valor del 95% o que cap de les components seleccionades aportí menys del 5% de la variància total original [Hair *et al.*, 2000]. Llindars del 80% o del 90% també han estat proposats, però moltes aplicacions climàtiques compten amb centenars de variables, fet que sovint ens obliga a acostar-nos amb percentatges inferiors [Jolliffe 1993].

CRITERI DE KAISER O DE L'ARREL LATENT:

No deixa de ser la mateixa aproximació que el criteri de variància explicada, però en comptes d'utilitzar els valors acumulats de percentatge, es basa en els valors propis. La metodologia planteja que una component principal és significativa quan explica, almenys, la mateixa variància que una variable original [Serra, 1994]. Això voldria dir la selecció d'aquells components que presentin valors propis més grans que 1 (per al cas de la matriu de correlacions, és la variància que aporta originalment cada cas, i això significa que aquell component retingut representa com a mínim una de les variables originals) [Yarnal, 1993], tot i que també es planteja el llindar de 0,7, segons quins siguin els nostres objectius. Aquest mètode acostuma a retenir massa components quan treballem amb més de cinquanta variables originals, mentre que tendeix a quedar-se curt quan aquest nombre de partida és inferior a vint [Hair *et al.*, 2000]. Amb tot, té l'avantatge de garantir la pèrdua d'una petita part de la variació original [Jolliffe, 1993].

SCREE TEST O CRITERI DEL CONTRAST DE CAIGUDA:

És probablement el mètode més utilitzat a la bibliografia, junt amb el criteri de North que presentem després. Introduït per Catell [1966], es basa en la representació gràfica del valor propi corresponent a cada component. El resultat acostuma a ser una corba

de perfil logarítmic, amb una marcada caiguda vertical inicial amb relació als components més explicatius, un esglaió de canvi de pendent, i una part final de la corba gairebé plana i relacionada amb els components molt poc explicatius (FIGURA 2.11). El punt (o component) on es dona el canvi de pendent i s'inicia la part horitzontal i plana de la línia es considera que marca el màxim de components a retenir. En definitiva, el mètode reté tots aquells components fins al punt que, afegir-ne un de nou, aporta molt poca informació (variància) nova. No sempre el punt de tall apareix clar als nostres ulls, i és per això que, com diu Serra [1994], en cas de dubte millor agafar un component trivial que perdre'n de significatius.

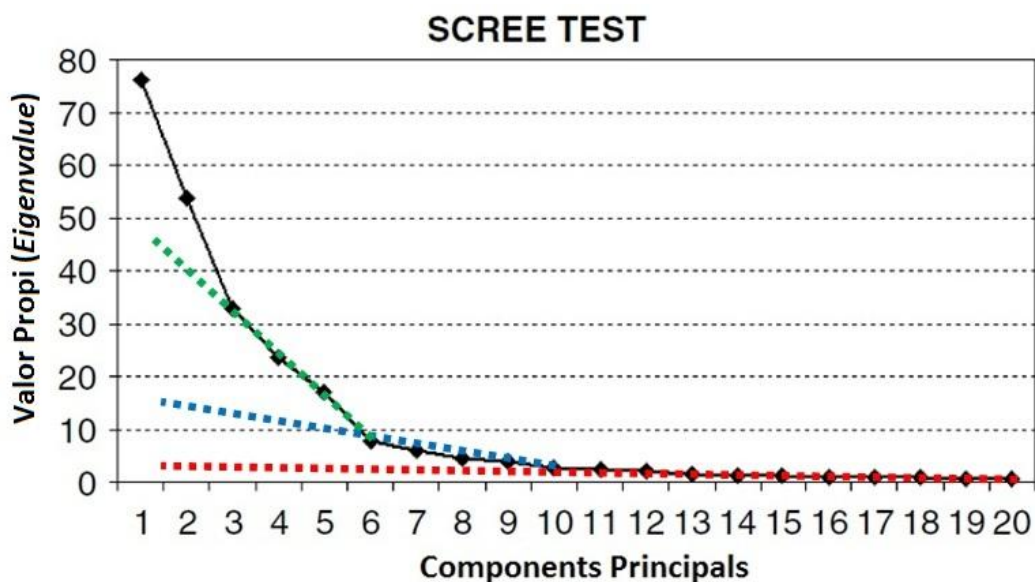


FIGURA 2.11.: Exemple de representació gràfica de l'Escree Test. Els canvis de pendent ens marquen els possibles punt de tall per decidir el nombre de components a triar. Podem observar que en aquest exemple apareixen tres possibles punts (3, 6 i 10), a partir dels quals es consideraria que els components principals representen el soroll de la mostra. Quin punt triar? Reforçar la decisió amb altres criteris com el de la variància explicada tenint clar quin és l'objectiu del nostre treball ens ajudarà molt per a fer la tria.

DIAGRAMA LEV:

Molt similar a l'anterior, però en comptes de representar gràficament cada component versus el seu valor propi, el que es fa és representar cada component respecte al

logaritme neperià del valor propi (LEV: *log eigenvalue vs eigenvector value*). El mètode fou plantejat per Craddock i Flood [1969], i genera una línia amb tres sectors: un de primer amb tendència lineal, un d'intermedi amb progressió geomètrica, i una tercera part de nou amb tendència lineal. Els autors parteixen de la màxima que en Meteorologia els components lligats al soroll presenten una progressió geomètrica, de manera que haurem de determinar el nombre de components a retenir segons el nostre criteri (quantitat de components que considerem òptim, grau de soroll que estem disposats a assumir...) en algun dels punts del sector en progressió geomètrica.

CRITERI DE NORTH:

Plantejada per North *et al.* [1982], es basa en l'ús de l'error que pot anar associat a la determinació dels valors propis quan treballem, sobretot, amb mostres petites. Podem, doncs, determinar un interval d'error per a cada valor propi. Reproduïm la formulació tal com la presenta Serra [1994] per la seva simplicitat:

El càlcul de l'error admissible a un valor propi seria:

$$\varepsilon\lambda = \lambda\sqrt{(2/N)}$$

on ε seria l'error associat al valor propi λ i N el nombre de variables considerades.

L'interval d'error associat a cada valor propi λ seria:

$$[\lambda - \varepsilon\lambda, \lambda + \varepsilon\lambda]$$

Aleshores, si en la seqüència d'interval dels valors propis trobem que hi ha intervals que es trepitgen (la diferència entre ells és 0), voldrà dir que aquests valors propis s'han degenerat, de manera que no s'han de seleccionar.

CARTOGRAFIA DE LES COMPONENTS PRINCIPALS:

Quan representem les components principals, és a dir, cartografiem els patrons de variació de les nostres dades, ens apareixen estructures que localitzen en l'espai la distribució de la variabilitat de la informació. Quan representem components principals que corresponen al soroll de la mostra, difícilment podrem detectar-hi una estructura espacial mínimament interpretable, ja que el gradient espacial dels

coeficients serà molt baix o nul. En aquest sentit, detectar a partir de quina component principal això ocorre també ens orientarà sobre on pot estar el líndar per on separem el senyal respecte del soroll de la mostra.

Veiem, doncs, que existeixen diverses formes d'establir els components principals que volem retenir. En aquest sentit, *White et al. [1991]* recomanen, vist que no hi ha cap mètode que és present com a òptim per diferenciar el senyal del soroll de la mostra, basar la decisió final en la "convergència d'evidències"; és a dir, a extreure una decisió final a partir del resultat de diferents tècniques.

Altres aspectes tractats per alguns autors a l'hora de valorar l'impacte que poden tenir en els resultats de l'ACP és la irregularitat i forma de la malla de dades [*Karl et al., 1982, O'Lenic i Lizevey, 1988*] i especialment l'ús de la rotació de factors per a la interpretació dels resultats; aspecte que abordem en apartats posteriors.

Independentment d'això, i per resoldre altres dubtes i principalment des de la perspectiva metodològica, recomanem entre les diverses publicacions existents el manual de *Hair et al., [1999]*, el manual universitari de *López i Lizares [2000]*, i, ja enfocat a les ciències atmosfèriques, la segona i més recent edició del manual sobre mètodes estadístics de *Wilks [2006]*.

2.4.1. ESTABLIMENT DE PATRONS DE BAIXA FREQUÈNCIA AMB L'ACP – LA ROTACIÓ DELS FACTORS

Si fem un cop d'ull als treballs que han emprat l'ACP per a l'estudi de la circulació atmosfèrica, veurem que ha estat i continua sent una eina important en Climatologia, especialment en la detecció de **patrons de baixa freqüència** (*low-frequency modes*). I és que la propietat de l'ACP d'acumulació de la variància a uns pocs components

principals ortogonals ha permès als climatòlegs detectar el que podríem considerar els maons (modes principals) de la circulació atmosfèrica [Huth *et al.*, 2008].

Tot i que coneixem indirectament l'existència de treballs anteriors com el de Fukuoka el 1951, el de Lorenz el 1956 [Richman, 1981] o el de Hare el 1958 [Barry i Carleton, 2001], als anys 60 ja trobem dues aportacions importants en l'estudi de la circulació atmosfèrica amb l'ús de l'ACP. Kutzbach [1967] caracteritza i interpreta els principals modes de variació de la mitjana mensual de pressió en superfície, temperatura, i precipitació sobre els Estats Units i el Canadà per a vint-i-cinc geners entre 1941 i 1965, que és d'interès per l'aplicació de la metodologia de l'ACP a les tres variables tant individualment com en conjunt, amb la consegüent discussió sobre els canvis d'estructura espacial que això comporta, i que conclou que els patrons establerts són consistents amb el coneixement que es té del comportament d'aquestes variables. D'altra banda, Craddock i Flood [1969] estableixen i interpreten els principals patrons de circulació atmosfèrica a 500 hPa per a gran part de l'hemisferi nord (a partir de 30°N) amb dades diàries, apostant per a la utilitat d'aquestes metodologies en la diagnosi i prognosi de la variabilitat en la circulació atmosfèrica. Ja a la dècada dels 80, Rogers [1981] estudia la variabilitat estacional a partir de dades diàries de pressió en superfície i de geopotencial a 500 hPa sobre l'hemisferi nord entre 1946 i 1977, intentant optimitzar la detecció de les anomalies de circulació. A part de la interpretació espacial dels resultats, on l'autor troba similituds amb treballs anteriors en la identificació de patrons com l'oscil·lació de l'Atlàntic Nord (NAO) o la del Pacífic Nord (NPO), Rogers analitza les sèries temporals obtingudes per detectar, entre d'altres, anys amb determinades conseqüències climàtiques fruit de la persistència d'un determinat patró al qual estan associats. La mateixa metodologia és aplicada posteriorment a l'hemisferi sud per a l'estiu i l'hivern [Rogers i Van Loon, 1982], però amb una sèrie temporal més curta (1972-1977).

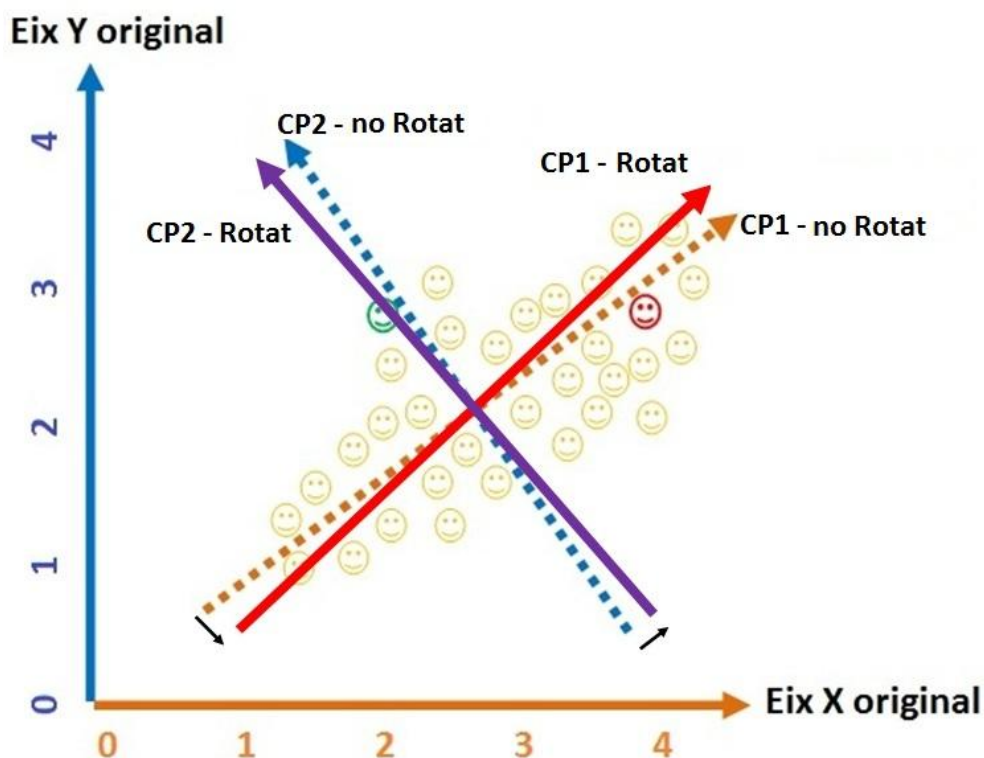


FIGURA 2.12.: La rotació de factors reajusta els nous eixos de coordenades creats amb l'ACP (els components principals). Amb aquest canvi, els primers components passen a explicar menys variància, informació que es redistribueix cap a altres components principals. En el nostre cas, com que els nous eixos mantenen un angle de 90º entre ells, parlarem de rotació orthogonal, la més emprada (l'anomenat mètode VARIMAX)

Tot i que hi havia hagut debats sobre la metodologia, els anys 80 es caracteritzen per l'aparició de treballs a favor i en contra del que s'anomena **rotació de factors** (FIGURA 2.12), procediment de reorientació dels eixos de les components (relaxació del forçament que busca l'ACP per representar millor la variància de la mostra) que es considera que permetria fer realment interpretables climàticament els resultats de l'ACP. Treballs com els de [Richman \[1981, 1986\]](#) o [Cohen \[1983\]](#) denoten les possibles mancances de la interpretació dels resultats no rotats, i s'introdueix la discussió plantejada anys abans per [Buell \[1975, 1979\]](#) sobre la predictabilitat dels resultats no rotats i la possibilitat que només siguin simples artefactes estadístics. En aquest sentit, [Horel \[1981\]](#) realitza un estudi comparatiu de gran interès. Així, a partir d'un treball de [Wallace i Gutzler \[1981\]](#) que estableixen els principals modes de variació de la circulació atmosfèrica a 500 hPa a l'hemisferi nord, Horel va més enllà i utilitza la rotació per

veure els canvis que puguin presentar els resultats. Com l'autor considera a les seves conclusions, la no rotació no necessàriament ha de significar un resultat erroni, però l'ús d'aquestes tècniques de transformació lineal dels resultats de l'ACP facilita l'establiment de relacions entre els patrons obtinguts i el coneixement observacional que es té de la circulació atmosfèrica.

Aquesta discussió sobre les limitacions, utilitat i interpretació correcta de resultats rotats o no rotats és recurrent i extensa. En aquest sentit, el treball de referència és el de [Richman \[1986\]](#) on tot i fer notar que l'opció no rotada és òptima quan busquem estrictament la reducció de dades, exposa clarament les deficiències dels resultats no rotats (ja hem citat anteriorment l'exemple de la predictibilitat dels resultats fruit d'un major pes de les dimensions del domini més que no pas per la variació de les dades (*domain shape dependence*)), discussió posteriorment ampliada per [Jolliffe \[1987\]](#), [Richman \[1987\]](#), [Legates \[1991\]](#), [Richman \[1993\]](#), i [Legates \[1993\]](#).

Posteriorment, [Mestas-Núñez \[2000\]](#) aborda un aspecte molt interessant relatiu a l'efecte que té en l'ortogonalitat espacial i temporal dels resultats rotats el tipus de normalització aplicada prèviament a les dades, i que serà segons apliquem les funcions ortogonals empíriques–EOF o l'anàlisi de components principals–ACP/PCA, aproximacions que sovint es consideren una mateixa. I és que la primera aproximació, les EOF, serien més òptimes en el seu ús quan volem utilitzar els resultats rotats per definir índexs temporals a la regió d'estudi ja que no estarien forçats a preservar les correlacions en el temps. En canvi, l'ACP, que si preserva la correlació temporal i relaxa l'ortogonalitat espacial, semblaria la millor aproximació per resoldre els problemes plantejats per [Richman \[1986\]](#) referents a la predictibilitat dels resultats, i en aquest sentit, aquesta aproximació seria preferible quan abordem la identificació de modes de circulació.

Un treball molt citat i de marcada entitat que empra la rotació dels factors és el de [Barston i Lizevey \[1986\]](#), que realitzaren un exhaustiu i extens estudi on s'identifiquen

els principals patrons de baixa freqüència de l'hemisferi nord a 700 hPa a escala mensual. Destaca, en els seus resultats, la identificació de la NAO com a únic patró hemisfèric que es dona al llarg de tot l'any (FIGURA 2.13), mentre que altres estructures de circulació com la PNA (*Pacific/North American Pattern*), la SZ (*Subtropical Zonal Pattern*) o l'AS (*Asian Summer Pattern*), entre moltes altres, no només queden clarament identificades en la seva localització i distribució espacial, sinó que els autors n'estableixen l'estacionalitat i persistència al llarg de l'any. La comparació amb altres treballs anteriors, i especialment amb els de Wallace i Gutzler [1981] i Horel [1981], demostra la bondat dels resultats climàtics obtinguts, com la potencialitat de l'ús de l'ACP per localitzar i caracteritzar aquest tipus d'estructures de circulació atmosfèrica, especialment quan es considera la rotació de les components principals.

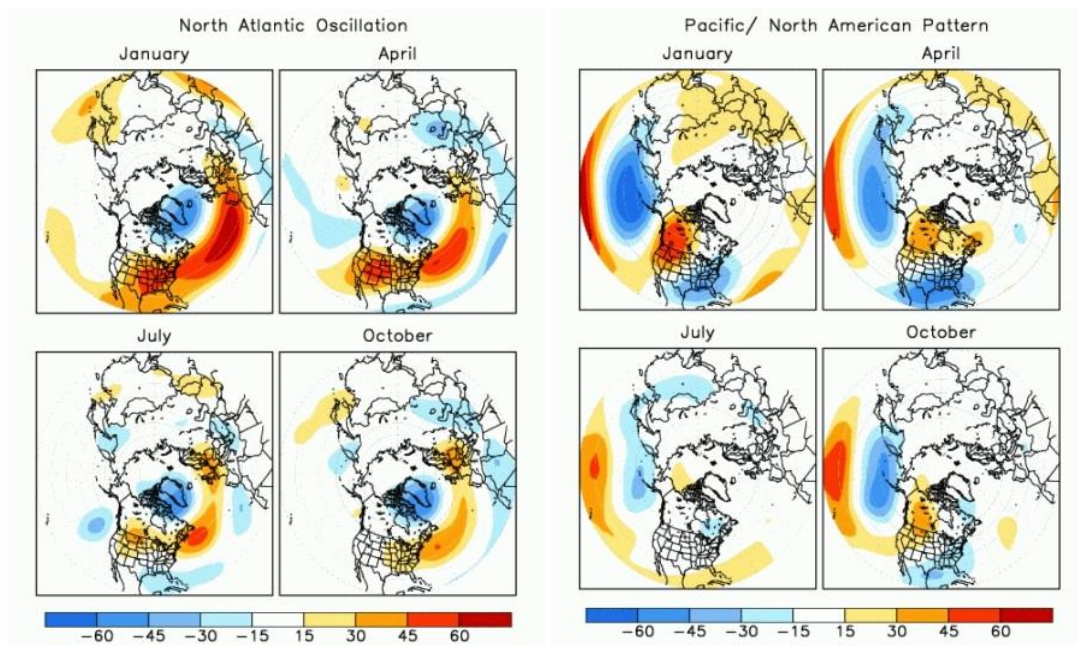


FIGURA 2.13: Cartografia dels coeficients (*loading patterns*) obtinguts per Barnston i Lizevey [1986] amb l'anàlisi de components principals al mes de gener, abril, juliol i octubre. Els mapes, doncs, representen la correlació temporal entre les anomalies estandarditzades de cada punt de malla (geopotencial a 700 hPa) i la sèrie temporal de patró de teleconnexió d'aquell mes. Els patrons representats són, a l'esquerra, l'oscil·lació de l'Atlàntic nord (NAO), i a la dreta, l'oscil·lació del Pacífic nord (PNA). Font: <http://www.cpc.ncep.noaa.gov/>

MODE	MATRIU DE DADES ORIGINALS	MATRIU DE CORRELACIONS O COVARIANCES	MATRIU DE COEFICIENTS DELS COMPONENTS PRINCIPALS	MATRIU DE PUNTUACIONS DELS COMPONENTS PRINCIPALS
O	TEMPS VARIABLES Estacions	TEMPS Correlacions	CP Coeficients o Saturacions "Loadings"	CP VARIABLES Puntuacions factorial "Scores"
P	VARIABLES Estacions	VARIABLES Correlacions	CP Coeficients o Saturacions "Loadings"	CP TEMPS Puntuacions factorial "Scores"
Q	ESTACIONS VARIABLES Temps	ESTACIONS Correlacions	CP Coeficients o Saturacions "Loadings"	CP VARIABLES Puntuacions factorial "Scores"
R	VARIABLES Temps	VARIABLES Correlacions	CP Coeficients o Saturacions "Loadings"	CP ESTACIONS Puntuacions factorial "Scores"
S	ESTACIONS TEMPS Variable	ESTACIONS Correlacions	CP Coeficients o Saturacions "Loadings"	CP TEMPS Puntuacions factorial "Scores"
T	TEMPS ESTACIONS Variable	TEMPS Correlacions	CP Coeficients o Saturacions "Loadings"	CP ESTACIONS Puntuacions factorial "Scores"

FIGURA 2.14. Taula amb els modes matricials que s'utilitzen a l'ACP i que, quan treballem amb informació meteorològica, combina tres tipus d'entitats diferents: espai (malla de dades o xarxa d'estacions), temps (series diàries, mensuals...) i camp (variable meteorològica com temperatura, pressió en superfície,...). A partir de la seva combinació es poden construir diferents tipus de matrius de dades originals que observem a la primera columna. Per a la classificació de patrons de circulació es treballa amb el mode S (que considera les correlacions espacials d'una variable al llarg del temps) i amb la seva transposada, mode T (considera les correlacions entre períodes de temps per a una variable i al llarg de l'espai). L'estructura tipus S també s'empra per a la regionalització climàtica, mentre que la tipus R s'usa per obtenir classificacions de tipus de temps. Adaptat de Richman [1986]

Fins ara hem vist alguns exemples que s'han anat succeint durant les passades dècades, aproximacions orientades a detectar els patrons de baixa freqüència. Aquests treballs coincideixen metodològicament en un aspecte molt important: treballen amb

una estructura de la matriu original del tipus S (*S-mode*) (vegeu FIGURA 2.14). Això significa que a la matriu de dades sobre la qual s'aplica l'ACP, les variables (columnes de la matriu) són els punts de malla repartits uniformement o relacionats amb estacions de mesura reals, mentre que les observacions (files de la matriu) són els casos en què s'ha donat el fenomen en estudi (dies, mesos, anys...). Si tenim en compte que la metodologia de l'ACP estableix les relacions entre les variables (les columnes de la matriu) per obtenir els resultats, podem veure que el que s'està interrelacionant són les sèries temporals dels punts de la malla (vegeu FIGURA 2.14, columna 2, fila 5). Això fa que els coeficients obtinguts siguin valors de correlació entre els punts de malla i els components principals. Aquests coeficients (FIGURA 2.14, columna 3, fila 5), doncs, es poden cartografiar (valors per a cada punt de malla), de manera que obtenim mapes d'aquests patrons de baixa freqüència; és a dir, les principals estructures espacials a les quals responen les nostres dades originals.

2.4.2: L'ESTRUCTURA DE LA MATRIU TIPUS S (S-MODE) PER A LA CTC

Hem vist a l'apartat anterior que podíem obtenir amb l'ACP i l'estructura de la matriu tipus S cartografia climàtica que representa les principals característiques que defineixen la circulació atmosfèrica a la nostra zona d'estudi amb l'obtenció dels modes de variació de baixa freqüència. Això vol dir que no hem obtingut una CTC, ja que no tenim una assignació dels nostres casos reals a un tipus concret que els agrupi.

Hem vist també que els coeficients resultants de l'ACP són valors de correlació de Pearson entre els punts de malla i els components principals (FIGURA 2.15, columna 3). En aquest sentit, els altres valors obtinguts, les puntuacions factorials, quantifiquen com de representat es veu cada cas real (dies, per exemple) per cada un dels components principals (FIGURA 2.15, columna 4). En definitiva, les puntuacions factorials són valors de proximitat de cada cas a les diferents estructures principals detectades per l'ACP.



FIGURA 2.15: Característiques del mode S, extret de la FIGURA 2.14. Per classificar els nostres casos i fer una CTC utilitzarem les puntuacions factorialis que s'obtenen amb relació a cadascuna de les components principals (CP). Amb tot, es requerirà algun mètode d'agrupacions per dur a terme aquest procés de generació de grups.

És important en aquest punt entendre que un cas de la mostra es pot veure molt ben representat per un component principal (puntuacions elevades per a aquell component i baixes per als altres), de manera que la seva estructura espacial s'ajustarà força al que representa aquest mode de variació. Però, amb tot, també és possible que les característiques espacials d'un cas concret siguin més aviat una combinació de diferents components (valors baixos o alts per a tots els components); és a dir, necessitem diverses d'aquestes peces bàsiques de la nostra variabilitat per construir el nostre cas real. Això ens dóna, en definitiva, uns valors d'associació (representativitat) de cada cas a cada component principal, els quals podem utilitzar per classificar les files de la matriu de puntuacions (FIGURA 2.15, columna 4); és a dir, els nostres casos (per exemple, dies). En realitat, cada seqüència de valors no deixa de ser un llistat de coordenades que ens permeten situar cadascun dels nostres casos en l'espai multivariant (recordeu la FIGURA 2.12). A partir d'aquí, podem utilitzar tècniques d'agrupació o *clustering* (apartat 2.5) per distribuir-los en grups similars d'estructura espacial i obtenir el CTC amb la mitjana dels casos de cada grup. Tot aquest procés queda resumit a la FIGURA 2.16.

Cal dir, amb tot, que aquest pas, des que s'obtenen les puntuacions factorialis i es distribueixen els casos en grups, no és tan directe: cal triar el nombre de grups, així com el mètode d'agrupacions a emprar. Aquest és un dels punts claus a l'hora d'emprar aquests mètodes, i un dels aspectes centrals que s'ha volgut resoldre en aquest treball, com veurem a la proposta metodològica (secció 4).



FIGURA 2.16: El procés habitual seguit quan fem una classificació amb l'ACP amb matriu tipus S. En color marró hi ha els passos associats a l'aplicació de l'ACP, mentre que en verd hi ha opcions més habituals que s'empren per a classificar els casos a partir de la matriu de puntuacions factorials. Més detalls sobre alguns d'aquests passos es donen a l'apartat 2.5 (mètodes d'agrupacions) i a l'apartat 4.2 (selecció de les dades, ús de la matriu de correlacions, rotació dels factors, selecció dels components a retenir, i cartografia dels components principals)

Diversos autors han utilitzat aquesta aproximació basada inicialment en l'ACP en mode S i finalitzada amb alguna tècnica d'agrupacions (vegeu secció 2.5). Així [Kidson \[1994\]](#) classifica nou anys d'anàlisis diàries de geopotencial a 1.000 hPa al sector de Nova Zelanda amb el mètode d'agrupacions de Ward i el de K-mitjanes (vegeu l'apartat 2.5) per classificar els casos a partir de les puntuacions factorials obtingudes amb l'ACP. [Galambosi et al. \[1996\]](#) aborden qüestions hidrològiques al sud-oest dels Estats Units fent la classificació dels casos amb K-mitjanes i prèvia determinació subjectiva del nombre de grups, mentre que [Corte-Real et al. \[1998\]](#) estableixen els tipus de circulació al sector sinòptic de Portugal per analitzar la variabilitat de la precipitació. Finalment, [Ekström et al. \[2002\]](#) i [Ekström et al. \[2004\]](#) treballen sobre l'erosió del vent a Suècia i Austràlia amb un dels enfocaments més habituals: ACP en mode S,

matriu de correlacions, rotació ortogonal *varimax* i agrupació dels casos amb el mètode de K-mitjanes.

N'hi ha més exemples, i en aquest sentit queda clar que l'eina d'ACP en mode S s'ha convertit molt sovint com un pas previ a l'ús de tècniques de *clustering* i que permet simplificar les variables de partida, reduint la dificultat de discriminació/càlcul als algoritmes d'agrupació.

2.4.3. L'ESTRUCTURA DE LA MATRIU TIPUS T (T-MODE) PER A LA CTC

Existeix una altra possibilitat, que correspon a la matriu transposada de l'anterior, la de Tipus T (*T-mode*) on les relacions s'estableixen entre la sèrie de dades corresponents a cada dia, any, etc.; és a dir, entre els mapes del temps (les columnes de la matriu, vegeu FIGURA 2.17, columna 2). Veient això, és fàcil pensar que aquesta és l'opció més òptima per establir patrons de circulació atmosfèrica; és precisament la que parteix de la comparació dels mapes per trobar les configuracions més explicatives. I és que, si observem la corresponent matriu de coeficients (FIGURA 2.17, columna 3), i com són valors de correlació de Pearson (hem de treballar amb matriu de correlacions), es fa molt fàcil associar els casos a cada component i construir el nostre CTC sense recórrer a mètodes d'agrupacions com quan treballem amb mode S.

MODE	MATRIU DE DADES ORIGINALS		MATRIU DE CORRELACIONS O COVARIANCES		MATRIU DE COEFICIENTS DELS COMPONENTS PRINCIPALS		MATRIU DE PUNTUACIONS DELS COMPONENTS PRINCIPALS	
		TEMPS		TEMPS		CP		CP
T	ESTACIONS	Variable	TEMPS	Correlacions	TEMPS	Coefficients o Saturacions "Loadings"	ESTACIONS	Puntuacions factorials "Scores"

FIGURA 2.17: Característiques del mode T, extret de la FIGURA 2.14. Per a classificar els nostres casos i fer una CTC utilitzarem els valors de correlació respecte als components principals que ens aporta la matriu de coeficients o saturacions

Hi ha treballs que comparen les diferents opcions que planteja l'ACP i altres tècniques per a la CTC. Així, [Huth \[1996\]](#) analitza cinc aproximacions diferents, entre les quals hi ha l'ACP amb la matriu **tipus S**, i seguida de dos possibles aproximacions de CL (l'*average linkage* com a mètode jeràrquic i el K-means com a no jeràrquic, vegeu apartat 2.5), i l'ACP amb la matriu **tipus T** amb classificació final dels dies basant-se en els valors dels coeficients (com dèiem, valors de correlació de Pearson de la sèrie temporal de casos respecte a les estructures espacial representades per les components principals). Huth conclou que, encara que tots els mètodes tractats presenten avantatges i inconvenients, l'ús del **mode T** és el que millor reproduceix l'estructura física de les dades i també és el que menys depèn de les decisions que es prenen durant la classificació, considerant-lo així com el mètode més prometedor per a les CTC.

Altres autors també consideren l'ús del mode T com la millor opció, com [Richman \[1981\]](#), que, tot i realitzar la classificació amb el tipus S, fa una aposta de futur en l'ús del tipus T per millorar l'establiment dels patrons.

En aquest sentit, volem destacar l'aportació de [Compagnucci i Richman \[2008\]](#) on en un article amb el títol "Pot l'ACP obtenir patrons de circulació atmosfèrica o de teleconnexió?" aprofundeixen en la interpretació que podem fer dels resultats obtinguts amb matriu de tipus S i de tipus T, i valoren si realment podem obtenir-ne resultats equivalents. D'entre els molts aspectes que ressalten aquests autors, demostren de nou que quan treballem amb matriu tipus S podem interpretar els resultats com a patrons de teleconnexió (recordem l'apartat 2.4.1), a diferència de la matriu tipus T, que permet interpretar els seus resultats com a patrons de flux, per exemple, com a tipus de circulació. En aquest sentit, és molt rellevant saber que quan obtenim la cartografia dels components principals a partir del mode S només els sectors amb saturacions elevades són interpretables espacialment i temporal, de manera que fem una bona identificació climàtica dels centres d'acció més rellevants de la zona d'estudi (recordem que ens cal un pas posterior per acabar la CTC). En canvi, amb matriu tipus T podem interpretar de forma conjunta i completa la cartografia de les components (el mapa de les puntuacions factorials); és a dir, obtenim directament

un catàleg de les principals estructures de la circulació junt amb la seva distribució al llarg del temps (sèrie de saturacions).

Així doncs, sembla clar que optar pel mode T seria la millor opció per a fer un CTC basant-nos en l'ACP, però resulta curiós veure que la major part dels articles que hem pogut consultar abans del 2004 treballin amb el tipus S seguit d'un mètode d'agrupacions. [Huth \[2000\]](#) en dóna probablement la raó més important: les aplicacions del **mode T** amb grans bases de dades requereixen un alt potencial de càlcul dels ordinadors, fet no sempre a l'abast de la mà de molts investigadors. Afortunadament, això ha anat variant, i ja s'ha convertit durant els darrers anys en una aproximació força més estesa.

Treballs que volem destacar que prenen l'opció en mode T per a la CTC són [Romero et al., \[1999a\]](#) per a la classificació del patrons diaris lligats a episodis de precipitacions intenses a la façana mediterrània de la península Ibèrica; [Jacobeit et al. \[2003\]](#) per estudiar la variabilitat del sector europeu de l'Atlàntic nord des de mitjan segle XVII, o el de [Maheras et al. \[1999\]](#) per caracteritzar els patrons de circulació sobre el Mediterrani relacionades amb les anomalies mensuals càlides i fredes més importants. Més recent és el de [García-Sellés et al. \[2009\]](#) que caracteritza els episodis de grans allaus de neu als Pirineus Orientals.

2.4.4. ALTRES VARIANTS DE L'ACP PER A LA CTC

Com que les problemàtiques a abordar a partir de la classificació de la circulació atmosfèrica són diverses i de variada complexitat, l'aparició de noves variants metodològiques que es basen en l'ACP és relativament freqüent. Ens ha semblat interessant de citar-ne dues: els SEOFS (*Simplified Empirical Ortogonal Functions*) [\[Hannachi et al., 2006\]](#) i el PSPA (*Principal Sequence Pattern Analysis*) [\[Compagnucci et al., 2001\]](#).

Segons Jolliffe *et al.* [2002], la transformació de les components principals convertint els coeficients a 0 i 1 a través de la rotació de factors no és l'única manera de simplificar les nostres dades. És per això, i per evitar aspectes considerats sovint controvertits com la decisió del nombre de factors a rotar, o quin tipus de rotació utilitzar (també existeixen rotacions factorials de tipus oblic que trenquen l'ortogonalitat), entre d'altres, que plantegen els SEOFS [Hannachi *et al.*, 2006]. Aquesta aproximació busca aprofitar els avantatges de l'ACP no rotat i rotat alhora a partir de la successiva maximització de la variància i forçant la seva ortogonalitat i simplicitat. En definitiva, els SEOFS simplifiquen les dades originals obtenint coeficients petits o grans per a cada patró; és a dir, coeficients molt propers a 0 o a 1 i sense valors intermedis, fet que no es produeix quan treballem amb ACP o ACP rotats. Aquest forçament s'obté a partir de l'ús d'un paràmetre llindar que s'afegeix als procediments habituals de l'ACP, optimització que requereix major capacitat de càlcul i que dificulta la seva implementació. Tot i això, i segons mostren els autors, els resultats són satisfactoris (en l'exercici presentat a Hannachi *et al.* [2006] s'identifiquen la NAO i altres patrons coneguts) i per això consideren l'aproximació d'utilitat si volem identificar pocs patrons per interpretar.

Per la seva banda, el PSPA busca classificar l'evolució dels sistemes atmosfèrics. Així, i actuant com una extensió de la classificació en mode T, s'aplica a camps atmosfèrics consecutius en el temps i ens permet determinar l'evolució dels centres d'acció. Així doncs, és una eina útil per obtenir CTC que a més incloguin l'evolució temporal habitual dels tipus establerts. La matriu d'entrada que s'empra, com dèiem, segueix exactament el model del tipus T mostrat a l'apartat anterior, però estenen les columnes de la matriu amb valors de tants dies com llarga es vulgui la seqüència de la classificació. En el seu treball, Compagnucci *et al.* [2001] realitzen la classificació dels camps de geopotencial a 1.000 hPa de dos-cents cinquanta-sis dies consecutius considerant seqüències de cinc dies, i conclouen que els resultats són molt satisfactoris quan s'analitzen els tipus de circulació obtinguts. Amb tot, sí que denoten una

dificultat: la complicació que representa la tria de components a retenir si volem implementar la rotació de factors. En aquest sentit, i sense saber-ho amb cap certesa, potser aquesta és la raó per la qual no s'han trobat altres referències rellevants posteriors on s'apliqui aquest mètode, fet estrany, tenint en compte l'atractiu evident que té aquesta aproximació.

Finalment, volem citar el treball de [Hannachi et al. \[2007\]](#), on es fa una revisió de moltes de les tècniques existents (inclosa alguna de les anteriors o d'altres de similars) per a la classificació de camps atmosfèrics a partir de les EOF. Recordem que aquest nom és utilitzat sovint per fer referència de forma genèrica a les tècniques basades en els vectors propis, es a dir, com a sinònim de l'ACP (vegeu apartat 2.4.1).

2.4.5. ALTRES APLICACIONS INTERESSANTS DE L'ACP

Finalment, volem puntualitzar que hi ha altres aplicacions de l'ACP lligades a l'estudi de la circulació atmosfèrica, com per exemple la reconstrucció històrica, a través de l'ús de les sèries temporals, de la distribució espacial de la temperatura [[Mann et al., 1998](#)], el camp de pressió en superfície [[Luterbacher et al., 2002a](#)], o d'índexs de circulació com la NAO [[Luterbacher et al., 2002b](#)]. Una altra possibilitat que ofereix l'ACP és la detecció de canvis de la morfologia i persistència de la circulació atmosfèrica després de l'ocurrència de fenòmens puntuals que afecten el comportament de l'atmosfera, com per exemple les erupcions volcàniques [[Prohom et al., 2003](#)] o l'activitat solar [[Lopez-Bustins et al., 2007](#)], entre altres.

En definitiva, l'ACP és una tècnica de caràcter semiautomàtic o assistida per ordinador (estem obligats a prendre certes decisions al llarg del procés) que inicialment va ser més orientada a la detecció de teleconnexions i d'anomalies de circulació a partir de l'establiment dels principals modes de variació de la circulació atmosfèrica. Molt útil també per a la reducció i la simplificació de les dades per facilitar la interpretació de la

informació, la manca d'uns resultats en format de camps atmosfèrics reals quan treballem amb mode S no ens permet, a priori, l'establiment de catàlegs de circulació tal com s'han plantejat habitualment en Climatologia, si és que no ho acompanyem amb un mètode de *clustering* o agrupacions que assigni els casos als tipus establerts. Amb tot, la possibilitat d'un canvi d'estructura de la matriu de partida (de tipus S a tipus T) i la millora de capacitat de càlcul dels ordinadors ha permès que l'ACP per si sola pugui considerar-se també una eina de CTC.

D'altra banda, és important en aquesta tècnica conèixer el seu funcionament, ja que, depenent de les opcions preses, els resultats se'n poden veure afectats i fer errar les nostres interpretacions.

2.5. MÈTODES D'AGRUPACIÓ OPTIMITZADA O CLUSTERING

Aquests mètodes multivariants consisteixen en la creació de grups de casos o observacions (*Clustering- CL*) de la manera més òptima possible a partir del càlcul de les distàncies entre ells en l'espai multidimensional. Vegem definicions de diferents autors:

[Everitt, BS; 1980] L'anàlisi clúster s'utilitza per assignar un conjunt d'observacions dins de grups que tenen característiques similars mesurades a partir d'un conjunt de variables mesurables.

[Hair, J.F. et al., 2000] L'anàlisi clúster és la denominació d'un grup de tècniques multivariades el propòsit de les quals és agrupar objectes basant-se en les característiques que posseeixen. [...] Els conglomerats d'objectes resultants haurien de mostrar un alt grau d'homogeneïtat interna (dins dels conglomerats) i un alt grau d'heterogeneïtat externa (entre conglomerats).

[Wilks, D.S., 2006] L'anàlisi clúster tracta la separació de dades entre grups les identitats de les quals no es coneixen a priori. En general, el nombre correcte de grups en els quals les dades s'han de classificar no es coneix amb anterioritat. Millor dit, és el grau de similitud i diferència entre les observacions individuals que s'utilitzen per definir els grups, i per assignar els integrants d'aquests grups.

[Barry, R.G. and Carleton, A.M., 2001] La classificació serveix al propòsit d'anomenar grups de coses, com també el d'agrupar coses per semblança, per relació, o ambdues. L'objectiu fonamental de tot mètode de classificació és el d'obtenir la mínima variabilitat dins dels grups i la màxima diferència entre els grups.

Amb aquestes definicions presentades, extraurem algunes conclusions puntuals a fi de tractar amb més detall la metodologia:

- a. L'objectiu principal del mètode d'agrupacions és el d'obtenir **grups d'objectes reals similars** per facilitar la seva anàlisi.

- b. Aquests objectes hauran de presentar un conjunt de valors associats que els defineixin perquè així, quan siguin comparats amb els mateixos valors de la resta d'objectes, puguem **establir similituds i diferències**.
- c. Aquests valors associats es poden considerar les coordenades de l'objecte a l'espai multivariant. Mirant les **distàncies** (que poden ser lineals, és a dir, distàncies euclidianes) entre els objectes en aquest espai, podem establir major o menor proximitat, interpretable com a major o menor similitud.
- d. Aquestes consideracions han d'aconseguir obtenir grups el màxim de compactes interiorment (el que s'anomena **alta homogeneïtat interna**), el que vindria a significar distàncies petites en l'espai multivariant entre els objectes d'aquell grup.
- e. D'altra banda, els límits de distància i/o la quantitat de grups han de permetre, alhora, que els diferents grups creats siguin el màxim de diferents entre ells (**alta heterogeneïtat externa**), el que voldria dir que les distàncies a l'espai multivariant entre aquests grups o *clusters* sigui el més gran possible.

Fins els darrers anys, en el camp de les ciències atmosfèriques, s'ha parlat sempre de dos principals mètodes d'agrupacions o *Clustering*: els mètodes jeràrquics i els mètodes iteratius [Yarnal, 1993; Barry i Carleton 2001; Hair *et al.*, 2000]. Però recentment, Jolliffe i Philipp [2010] fan una distinció més acurada en quatre grups:

MÈTODES JERÀRQUICS AGLOMERATIUS: aquí s'inclouen els mètodes que, després de considerar tots els objectes com un possible grup, els van agrupant en diferents etapes segons la seva proximitat en l'espai multivariant. Aquest procediment es representa normalment amb dendrogrames (FIGURA 2.18) que finalitzen quan tots els casos queden units en un mateix grup. Els mètodes més coneguts i emprats en Climatologia són el mètode de Ward i el del veí més proper (*single-linkage*). Com a exemple tenim el treball de Bednorz [2008] que utilitza el mètode de Ward per classificar les nevades intenses a cotes baixes de Polònia i Alemanya, obtenint tres tipus de circulació que

confirmen una forta correlació negativa entre la ocurrència d'aquest fenomen i els valors positius de l'índex NAO.

Cal dir que, tot i que per si sols es poden emprar per obtenir una CTC (només cal triar en quin punt del dendrograma creiem que es dóna el nombre òptim de grups), és més habitual que s'empri com a pas previ d'un mètode iteratiu (mètodes d'optimització global com K-means que veurem després). Aquest ús previ a un mètode global serviria per triar el nombre de grups o *clusters* que ha de tenir la nostra CTC. Aquesta opció és la que empra [Kidson \[1994\]](#) per classificar la circulació atmosfèrica al sector de Nova Zelanda.

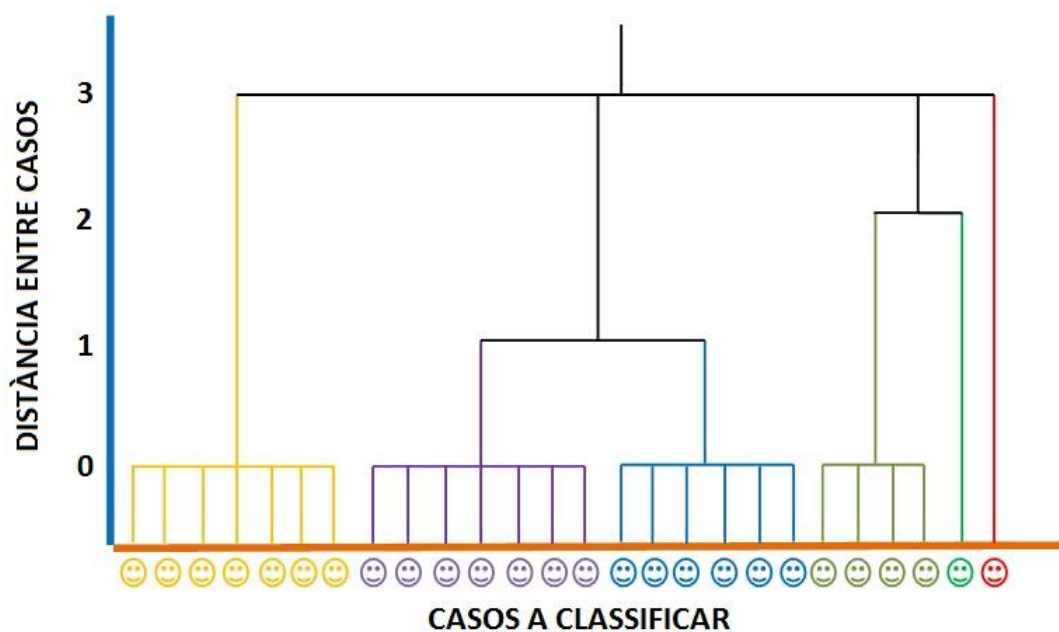


FIGURA 2.18. Exemple de dendrograma, representació gràfica del procés que segueixen els mètodes jeràrquics aglomeratius per classificar una mostra de casos. Els nostres *smilies* es podem diferenciar clarament en sis grups. Però les diferències entre ells no sempre són les mateixes, ja que alguns grups de casos són del mateix color i només els diferencia la tonalitat (és el cas del lila i el blau o del verd oliva i el verd clar). També és interessant veure un cas ben diferenciat de la resta, el vermell.

MÈTODES JERÀRQUICS DIVISIUS: Seria el procés invers a l'anterior, començant per un sol *cluster*. En general presenten els mateixos problemes que els aglomeratius però, malgrat tot, com se cita en [Jolliffe i Philipp \[2010\]](#), poden ser avantatjosos com a de

recurs informàtic quan el nombre de grups a obtenir és molt menor que la mostra de casos, alhora que hi hauria menys probabilitat de fer passos erronis que influïssin en el resultat final. Hem de dir que en la nostra cerca bibliogràfica no n'hem trobar cap exemple.

MÈTODES D'OPTIMITZACIÓ GLOBAL: S'associen normalment als mètodes anomenats de K-mitjanes (K-means), una de les aproximacions que més apareix a la bibliografia de CTC (FIGURA 2.19). S'acostuma a partir de la definició d'un nombre de grups òptim i que, localitzats a l'espai multidimensional (centroïdes), incorporaran tots aquells casos que tinguin més propers tenint en compte el criteri de distància triat (normalment l'eulidiana). Per aconseguir això, el criteri de qualitat de la classificació sovint es basa en aconseguir la màxima variabilitat entre els grups (*between-group variability*) i la mínima dins els grups (*within-group variability*), fet que pot requerir d'iterar el procediment d'assignació fins que, teòricament, s'aconsegueixi l'anomenat òptim global.

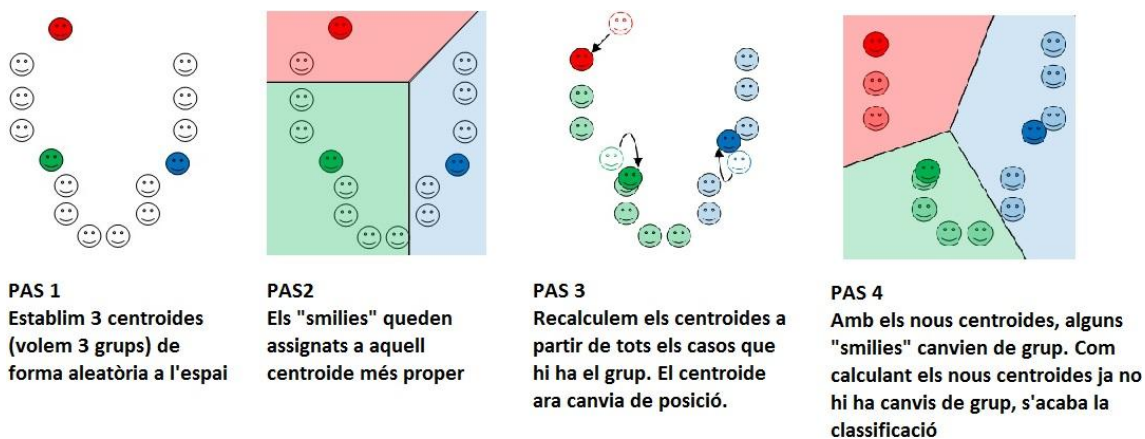


FIGURA 2.19: Reproducció del funcionament de classificació d'un mètode d'optimització global. El mètode de les K-mitjanes (*K-means*) és el nom que rep sovint aquesta aproximació. Adaptat a partir de http://en.wikipedia.org/wiki/K_means.

A diferència dels procediments anteriors, on la interpretació del dendrograma divisiu o aglomerador ens pot orientar sobre el nombre de grups, els mètodes d'optimització no ens aporten informació prèvia sobre el nombre de grups (es a dir, de tipus de circulació) que ha de tenir la nostra classificació. Es per això, com s'ha comentat anteriorment, que sovint els mètodes jeràrquics s'utilitzen com a tècnica exploratòria que permet definir el nombre de centroides a utilitzar com a referència en el mètode iteratiu, com era el cas de [Kidson \[1994\]](#) a Nova Zelanda. Altres autors han definit aquest nombre de grups i centroides de partida de forma subjectiva [[Bogardi et al., 1993](#); [Galambosi et al., 1996](#)].

MÈTODES BASATS EN MODELS: En aquestes tècniques és determinant el coneixement de la forma del núvol de punts de les nostres variables distribuïdes en l'espai multivariant, és a dir, la forma dels nostres *clusters*. Això permet estimar paràmetres per a aquestes distribucions i així generar les assignacions amb mètodes estadístics. Sobre aquesta metodologia tampoc hem trobat documentació al llarg de la nostra recerca bibliogràfica i no hem pogut entrar-hi amb més detalls, excepte conèixer-ne recentment la seva existència gràcies a [Jolliffe i Philipp \[2010\]](#).

2.5.1. EL QUICK CLUSTER [HARTIGAN 1975]

Probablement és l'algoritme d'optimització més emprat en ciències atmosfèriques, i segurament també en altres disciplines. I és que, a part de la seva simplicitat, està implementat a programaris d'ús molt estès com l'SPSS.

Anteriorment hem vist la definició genèrica dels mètodes K-mitjanes, que a partir d'uns centroides o llavors a l'espai multivariant assignen els casos per mesures de proximitat, procés que s'itera amb un nou càlcul dels centroides i noves assignacions fins que s'obté un resultat estable. El que caracteritza l'algoritme de [Hartigan \[1975\]](#) és la seva simplicitat i rapidesa, fet que ha facilitat que estigui implementat al software comercial SPSS o sigui l'opció per defecte al software lliure R.

La bibliografia és extensa pel que fa a treballs que classifiquen amb K-mitjanes (suposem que amb aquest algoritme del *quick cluster*, ja que no s'acostuma a explicitar en els articles). Un exemple clàssic és el de [Corte-Real et al., \[1998\]](#) que obté patrons de circulació per a Portugal relacionats amb la precipitació. Un altre treball de força interès ha estat el d'[Ekström et al., \[2002\]](#) que classifica a resolució diària i per mitjà de la definició d'un índex de severitat de l'erosió, l'efecte del vent sobre zones agrícoles de sòl sorrenc al sud de Suècia.

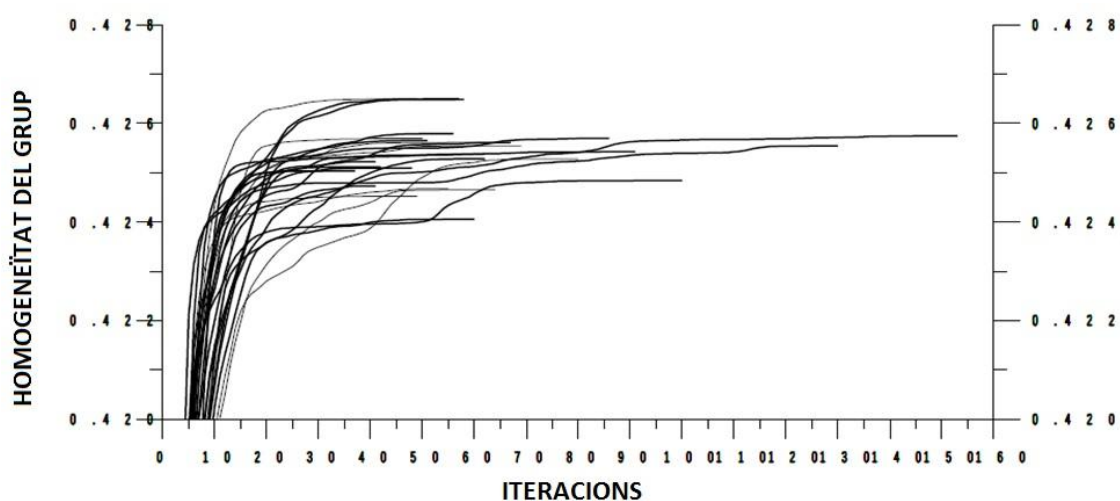


FIGURA 2.20: El mètode de K-mitjanes resulta inestable a l'hora de convergir els seus resultats un cop fetes les iteracions a partir de diferents llavors de partida i amb la mateixa mostra segons es deriva dels assajos fets al departament de Geografia de la Universitat d'Augsburg (Alemanya). Adaptat amb permís d'Andreas Philipp.

Però hi ha un aspecte important que alguns autors han fet notar en relació al mètode iteratiu de K-mitjanes, i és que realment mai no es troba un òptim global, ja que el procediment de classificació es veuria afectat per quins són els centroides de partida i també per l'ordre en què s'utilitzen [\[Philipp 2007\]](#) (FIGURA 2.20). En aquest sentit, l'ús de centroides de partida aleatoris no seria la millor opció de cara a obtenir una classificació estable. A més, es recomanaria l'establiment de forma prèvia d'aquest nombre de grups i els seus centroides amb la màxima solidesa climàtica possible. [Huth \[1996\]](#) fa notar també una altra limitació d'aquest mètode: la tendència a fer grups

d'igual dimensió (similar nombre de casos), tendència clarament artificial quan estem treballant amb la classificació de la circulació atmosfèrica.

2.5.2. SANDRA: SIMULATED ANNEALING AND DIVERSIFIED RANDOMIZATION CLUSTERING

Aquesta tècnica no jeràrquica busca, com en el cas de l'aproximació K-mitjanes, la minimització del sumatori de les distàncies euclidianes entre les classes, però amb la diferència que s'utilitzen un seguit d'estratègies per aproximar-nos millor a un òptim global [Philipp 2009], que hem vist que no era possible amb els mètode no jeràrquics convencionals.

La novetat se centra en els anomenats reassignaments erronis (*wrong re-assignments*), és a dir, objectes que s'han de treure del grup més proper depenent d'una probabilitat P que és elevada a l'inici del procés però que va decreixent a poc a poc. En aquest sentit, si el procés convergeix en un resultat de poca qualitat, alguns objectes podran ser assignats de nou emprant aquest valor P i el resultat de la classificació milloraria. Perquè aquest valor P es redueixi de forma lenta i progressiva, s'empra un paràmetre de control T (per a la temperatura) que davalla en funció d'un factor de refredament C (*Cooling factor*) a cada iteració [Philipp et al., 2010]. És la tècnica del temperament simulat (*simulated annealing*).

Aquest procediment de millora dels resultats es veu recolzat per una repetició del procés de mil vegades amb selecció aleatòria de les llavors de partida i del seu ordre a l'hora d'iniciar els càlculs (*diversified randomization clustering*).

Aquest mètode de classificació ha despertat molt interès entre la comunitat científica, especialment des de la seva utilització en el projecte EMULATE (*European and north atlantic daily to MULTidecadal climATE variability* [

<http://www.cru.uea.ac.uk/projects/emulate/>). En aquest sentit, són força interessants els resultats sobre els tipus de circulació obtinguts i el comportament de la temperatura per al període 1850-2003 al sector de l'Atlàntic Nord i Europa a partir de malles de pressió atmosfèrica diària reconstruïdes en aquest projecte europeu [Philipp *et al.*, 2007].

D'altra banda, la seva principal limitació són els requeriments de capacitat de càlcul demanats pel procediment seguit, que en alguns casos poden limitar la seva aplicació.

Finalment cal fer notar que aquesta metodologia (en general, tots els mètodes d'optimització) també permet la classificació de seqüències de casos tal i com ja hem plantejat amb l'Anàlisi de Components Principals (vegeu 2.4.4). En aquest sentit la comparativa realitzada per Philipp [2009] entre la classificació de seqüències amb l'ACP i mètodes d'optimització conclou bons resultats en els dos casos, i destaca els beneficis que pot tenir l'ús de seqüències en comptes de casos individualitzats, per exemple, en treballs de reducció estadística d'escala (*statistical downscaling*). Alguns treballs en aquesta direcció i orientats als riscos naturals (incendis forestals, allaus de neu...) s'han presentat en congressos científics recents [Esteban *et al.*, 2010].

2.5.3. SOM: SELF-ORGANIZING MAPS

El SOM (*Self Organizing maps* / mapes organitzats automàticament), és un mètode basat en els principis de les xarxes neuronals (*Neural networks* o NNW) [Cavazos, 2000, Hewitson i Crane 2002].

En analogia a l'estructura del cervell dels éssers vius, les xarxes neuronals estan compostes per un gran nombre d'unitats simples de procés (neurones) connectades entre si en base a una topologia (arquitectura) definida (FIGURA 2.21). En aquest sentit, les xarxes neuronals són sistemes intel·ligents la base de coneixement de les quals és la mateixa estructura de la xarxa i els pesos apresos. Així, un cop escollida la topologia de connexió, la funcionalitat de la xarxa ve donada per la funció concreta de

cada neurona i els pesos de les connexions entre elles, amb la qual cosa la xarxa aprèn i generalitza el coneixement adquirit fins a donar un resultat [Gutiérrez *et al.*, 2004].

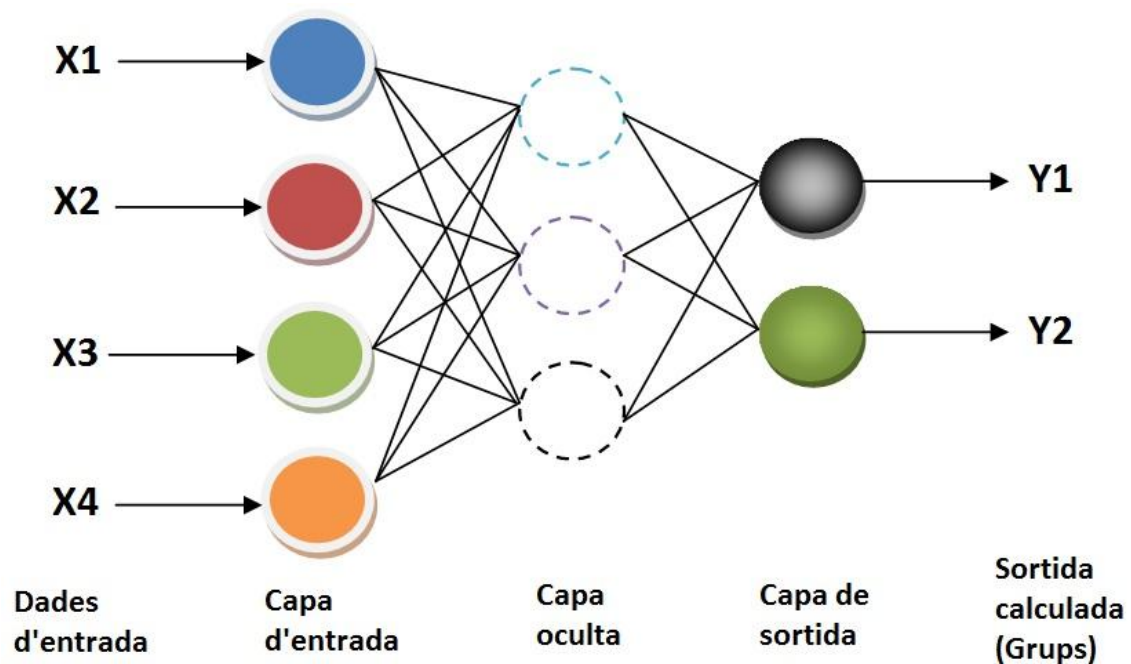


FIGURA 2.21: Esquema clàssic del funcionament d'una xarxa neuronal unidireccional. Les relacions d'aprenentatge poden ser molt més complexes, amb iteracions o addició de condicionants externs, per exemple. Pel que fa al cas del self organizing maps, on es segueix la metodologia de Kohonen, el veïnatge espacial també es té present en l'aprenentatge del sistema.

Concretament en CTC habitualment es treballa amb l'aproximació de Kohonen [Philipp *et al.*, 2010, Reusch 2010]. La diferència amb d'altres algorismes és que el procés d'entrenament inclou un mecanisme auto-organitzatiu de veïnatge de manera que els veïns en una retícula de dos dimensions (que s'utilitza per organitzar i visualitzar les dades de forma intuïtiva i eficient) són també similars a l'espai real [Gutiérrez *et al.*, 2004].

2.5.4. COSA: CLUSTERING OBJECTS ON SUBSETS OF ATTRIBUTES

És un mètode similar a K-means, però que permet que diferents variables (per exemple zones geogràfiques de la malla) tinguin diferents pesos que poden variar de forma contínua de cara al càlcul de les distàncies entre els *clusters*. Això fa el mètode molt flexible, ja que hi ha molts paràmetres que es poden afinar [Jolliffe i Philipp 2010]. Tots els detalls sobre la metodologia, de la qual no coneixem cap aplicació en Climatologia, es pot trobar a Friedman i Meulman [2004].

2.6. ALTRES MÈTODES DE DE CLASSIFICACIÓ DE TIPUS DE CIRCULACIÓ

Existeixen altres aproximacions per a l'establiment de catàlegs de patrons de circulació atmosfèrica. Algunes combinen alguns dels mètodes que hem vist fins ara, mentre que altres introdueixen noves formes d'establiment dels catàlegs.

Així, [Frakes i Yarnal \[1997\]](#) plantegen un mètode híbrid de classificació: la combinació del mètode manual i el de correlacions. La problemàtica dels mètodes de Lund i Kirchhofer, especialment des del vessant de l'establiment dels mapes clau, es replantejaria en aquesta nova opció mitjançant l'establiment manual dels grups que seran les llavors per al mètode de correlacions.

El mateix plantejament ha estat abordat per [Rasilla \[2003\]](#) en el cas de la península Ibèrica, però en aquesta ocasió combinant el mètode manual (basat en l'establiment de les categories segons criteris de moviment i trajectòria dels centres d'acció, curvatura del camp de pressió, i procedència de l'advecció), per després emprar l'ACP i el CL (algoritme no jeràrquic de K-mitjanes) per obtenir un catàleg final de deu tipus (FIGURA 2.22) i classificar tots els casos.

Una altra opció híbrida és la plantejada per [Maheras et al., \[2000\]](#), que empra mètodes de topologia i geometria per a la classificació dels mapes. Maheras intenta combinar els avantatges dels mètodes objectius i els manuals en set passos que se centren en la identificació de les altes i baixes pressions a la zona en estudi. En realitat, aquesta aproximació es podria considerar basada, en certa mesura, en tècniques d'anàlisi discriminant.

Finalment, de Cuba van sorgir els PSO (Processos Sinòptics Objectius) [\[Fernández et al., 2003\]](#), desenvolupats per l'escola cubana de Meteorologia, basades en bona mesura en les aproximacions de tipus discriminant ja emprades per altres autors. La novetat de les PSO és la introducció de seqüències de mapes en l'establiment dels tipus de circulació, de manera que cada patró o PSO obtingut està compost per una seqüència de tres dies (D-1, D, D+1) de pressió en superfície i geopotencial a 500 hPa,

posteriorment relacionable amb variables en superfície (en el cas del treball consultat, es tracta la precipitació de la península Ibèrica a la tardor).

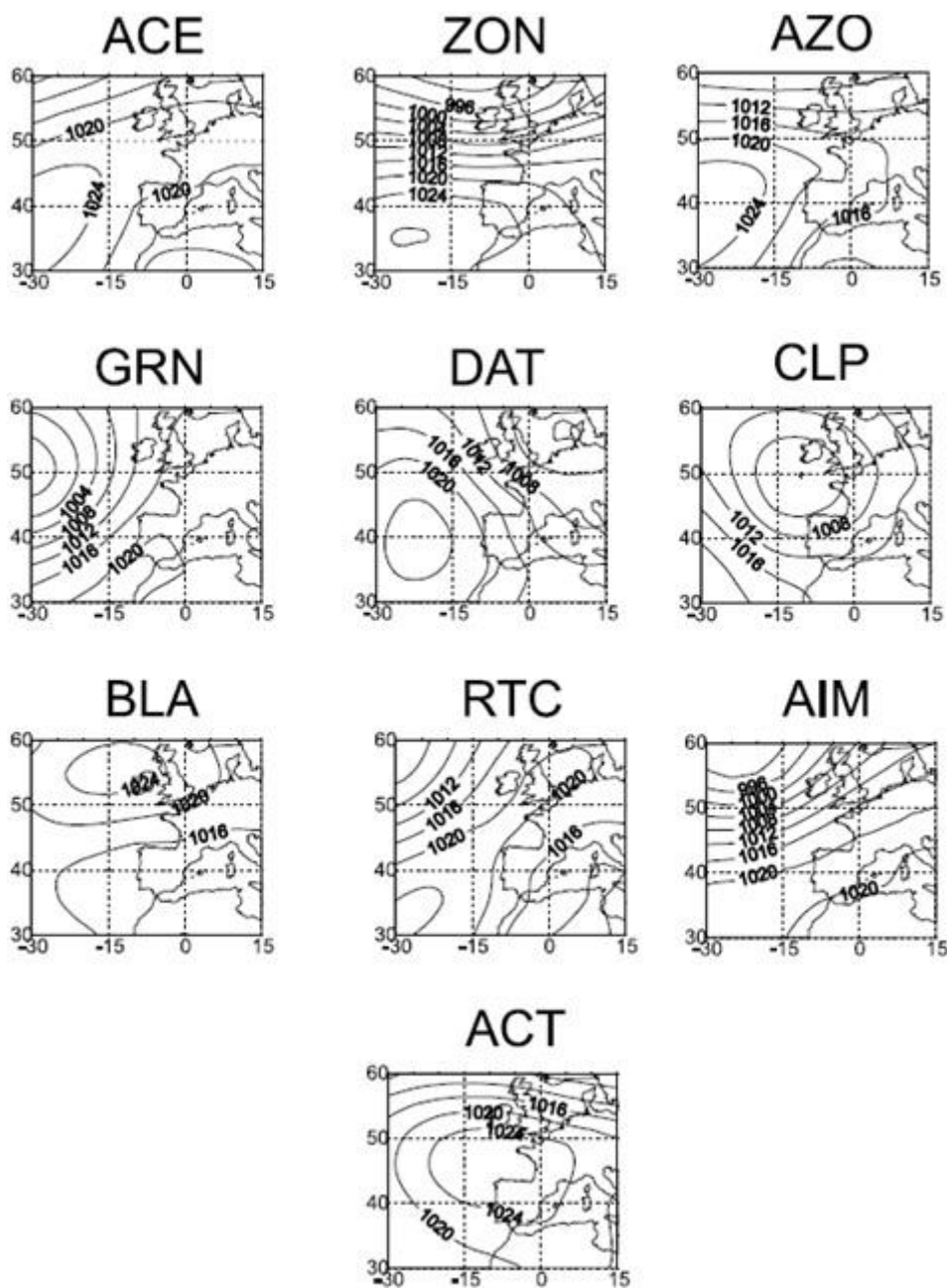


FIGURA 2.22: Classificació de Rasilla obtinguda el 2003. Aquest deu tipus de circulació obtinguts pel sector de l'Europa Occidental están pensats sobretot en l'àmbit de la Península Ibèrica. L'aproximació metodològica combina l'experiència subjectiva del climatòleg junt amb les capacitats del mètode de K-mitjanes, (Extret de Rasilla [2003]).

2.7. QUIN MÈTODE DE CLASSIFICACIÓ DE TIPUS DE CIRCULACIÓ S'HA D'UTILITZAR?

Amb totes les possibilitats exposades, quedar clar que hi ha moltes opcions per abordar una Classificació de Tipus de Circulació (CTC). En l'ampli ventall de descripcions prèvies s'han anat fent rellevants alguns avantatges i inconvenients de les diferents metodologies, problemàtiques que normalment quan es donen en un mètode sovint es resolen en el següent i viceversa. És a partir d'aquí, i per ajudar a aquesta tria, que plantejem un llistat de criteris que considerem que pot ser útil a l'hora d'escollir aquella classificació que ens sembli, de forma teòrica, més adient per a les característiques de la problemàtica que volem resoldre:

- a. Que sigui el màxim d'objectiva possible. – és necessari que el classificador tingui coneixements sobre la climatologia de la regió estudiada?
- b. Que estigui ja força consolidada entre la comunitat científica i que hi hagi bibliografia suficient. – està contrastada científicament per diversos autors?
- c. Que sigui aplicable a diferents nivells i variables atmosfèriques. – podem treballar sobre qualsevol variable atmosfèrica?
- d. Que classifiqui tots els casos. – com es tracten els casos fora de rang? Hi ha molts casos no classificables?
- e. Que sigui aplicable a qualsevol sector geogràfic, per a diferents resolucions i a diferents escales espacials i temporal. – és molt particular en un determinat espai geogràfic?
- f. Que no sigui excessivament costosa computacionalment. – disposem de capacitat de càlcul per poder fer la classificació?

3. HIPÒTESI DE TREBALL I OBJECTIUS DE LA TESI

Com hem vist, existeixen moltes metodologies per a generar catàlegs de la circulació atmosfèrica. El ventall de problemàtiques de caràcter meteorològic, climàtic i geogràfic que es poden abordar és ampli, fet que queda evidenciat amb molta de la bibliografia existent. Amb tot, els sectors d'orografia complexa on hi ha una elevada variabilitat meteorològica i climatològica a dia d'avui encara representen un problema a l'hora d'abordar-los, ja sigui des de la perspectiva de la predicció meteorològica com també de la seva caracterització climàtica i la generació d'escenaris de canvi climàtic. Així, ni els models meteorològics de caràcter dinàmic ni, especialment, els models de circulació general o regional emprats per projectar el canvi climàtic són prou acurats en sectors com, per exemple, els Pirineus o Andorra. En aquest sentit, en aquest treball partim de la següent hipòtesi: **L'ús de les classificacions de tipus de circulació es presenta com una eina molt eficient per abordar estudis de Climatologia sinòptica a sectors de relleu complex o quan volem aprofundir en les relacions entre la circulació atmosfèrica i l'escala local. Utilitzar-les ens permetrà deixar enrere les tècniques manuals, sobretot si volem treballar amb grans volums de dades i obtenir catàlegs molt pràctics de forma eficient sense patir per la capacitat de càlcul dels nostres ordinadors.**

En aquest sentit, s'han plantejat els següents objectius en el present treball:

A. OBJECTIUS PRINCIPALS

- I. Conèixer les metodologies existents per a la classificació de la circulació atmosfèrica i triar un mètode preferent de classificació entre el ventall existent, mostrant també les diferents possibilitats d'aplicació de les Classificacions de Tipus de Circulació.
- II. Fer una proposta metodològica per resoldre les mancances parcialment o completa del mètode triat i adaptar-lo preferentment, si és necessari, a l'escala diària per poder treballar en àmbits com els riscos meteorològics o la predicció meteorològica a zones d'orografia complexa.

- III. Validar el funcionament de la proposta metodològica per mitjà de diferents exemples meteorològics i climàtics, centrant-nos prioritàriament en el sector geogràfic del principat d'Andorra.

B. OBJECTIUS SECUNDARIS

- i. Caracteritzar sinòpticament les nevades intenses al principat d'Andorra per aprofundir en els seus trets meteorològics i climàtics. Obtenir, doncs, un catàleg que sigui d'utilitat de cara a la predicció meteorològica a la muntanya.
- ii. Obtenir un catàleg de la circulació atmosfèrica per a l'Europa Occidental i caracteritzar el seu comportament estacional. Analitzar també la seva variabilitat recent i veure si els patrons establerts presenten tendències de canvi significatives.
- iii. Veure si la discriminació a escala sinòptica de tipus de circulació és traslladable a l'escala local a partir de la generació de cartografia climàtica de temperatura i precipitació. La seva aplicació al principat d'Andorra també ens ha d'aportar un millor coneixement de les característiques climàtiques d'aquestes dues variables en aquest sector del Pirineu Oriental.
- iv. Caracteritzar les situacions sinòptiques amb activitat elèctrica en el sector d'Andorra i Catalunya i derivar productes que puguin ser d'utilitat de cara a la predicció meteorològica.

4. MÈTODE DE CLASSIFICACIÓ DELS TIPUS DE CIRCULACIÓ: PROPOSTA METODOLÒGICA

4.1. TRIA DEL MÈTODE DE CLASSIFICACIÓ

Seguint els criteris plantejats a l'apartat 2.7 d'aquest treball, hem fet la tria del mètode que ens ha semblat més adient per a la classificació de tipus de circulació atmosfèrica. Anem seguint els criteris i anem donant-los resposta:

- a. Que sigui el màxim d'objectiva possible –requereix de coneixements sobre la climatologia de la regió que és objecte d'estudi per part del classificador?*

En aquest punt, els mètodes de Jenkinson i Collison (apartat 2.2.1.) i el dels Prototipus (2.2.3) podrien ser les dues opcions més adequades ja que es poden considerar absolutament objectives. Les Xarxes Neuronals també es podrien tenir en compte des d'aquesta perspectiva, i l'ACP a partir de la matriu T també s'aproximaria força a la total objectivitat.

- b. Que estigui ja força consolidada entre la comunitat científica i que hi hagi bibliografia suficient – està contrastada científicament per diversos autors?*

La metodologia que es presenta com la més consolidada des de la perspectiva de la seva utilització en treballs científics correspon al mètode de l'anàlisi de components principals amb mode S i es combinada amb mètodes d'agrupacions tant jeràrquics com iteratius (2.5). Altres aproximacions també molt emprades són el mètode de Jenkinson i Collison (2.2.1) i l'ACP amb mode T (apartat 2.4.3).

En el sentit contrari, s'han presentat propostes metodològiques molt interessants com el mètode del Prototipus, el mètode SANDRA, els *Extended*

EOFs o el mètode d'agrupacions COSA. Però en aquests casos com en d'altres, difícilment s'han emprat de forma mínimament extensiva.

- c. *Que sigui aplicable a diferents nivells i variables atmosfèriques – podem treballar sobre qualsevol variable atmosfèrica?*

Aquí la majoria de mètodes no presenten dificultats però, en canvi, sí que el mètode de Jenkinson i Collison presenta una important mancança ja que tal com està definit ens limita a treballar només amb pressió en superfície o amb algun nivell de geopotencial (aquí ja necessitaríem adaptar la formulació).

- d. *Que classifiqui tots els casos – com es tracten els casos fora de rang? Hi ha molts casos no classificables?*

El tractament dels *outliers* o de valors molt allunyats de la centralitat de la mostra només ens dona dues alternatives, especialment quan treballem amb mostres de grans dimensions: o acceptem l'existència de casos inclassificables o assumim que poden haver-hi casos associats a grups tot i que no estiguin òptimament representats. Mètodes com els de correlacions sovint es critiquen per l'elevat nombre de casos no classificats que s'obtenen, aspecte que també pot ocórrer amb Jenkinson i Collison. D'altra banda, trobem les metodologies basades en l'ACP, els mètodes d'agrupacions o les Xarxes Neuronals, que *a priori* classifiquen tots els casos. I diem *a priori* perquè no es pot descartar (tot i que és molt difícil) l'obtenció de grups amb un sol cas.

- e. *Que sigui aplicable a qualsevol sector geogràfic, per a diferents resolucions i a diferents escales espacials i temporals – podem classificar a diferents escales meteorològiques?.*

En aquest punt són especialment sensibles els mètodes manuals que, com hem vist, depenen absolutament del coneixement del clima de la regió estudiada pel classificador. Això dificulta molt que els criteris utilitzats siguin transportables a altres zones. Derivat d'això, altres mètodes que també presenten mancances en aquest punt són algunes de les aproximacions basades en mètodes manuals. És el cas de l'Objectivació del *Groswaterlagen* alemany. En canvi, mètodes com els de correlacions, els basats en l'ACP o els mètodes d'agrupacions, poden aplicar-se arreu i per resolucions espacialment i temporalment variables.

f. Que no sigui excessivament costosa computacionalment – disposem de capacitat de càlcul per a poder fer la classificació?

Aquí, les principals limitacions sorgeixen per a mètodes com les xarxes Neuronals, SANDRA o els SEOF. I si a més ens plantejem el treball amb ordinadors casolans i amb matrius de grans dimensions, també el mètode ja consolidat d'ACP basat en el mode T ens pot generar problemes.

Sospesat tot això, el mètode de l'ACP a partir de la matriu tipus T va ser el triat per treballar. Els motius de la tria van ser la seva elevada objectivitat, el fet d'estar basat en una metodologia robusta i ben establerta (ACP), la seva aplicació a diversos casos, l'existència de treballs comparatius (destaca [Huth \[1996\]](#)) i el fet que ens permetés treballar amb qualsevol variable atmosfèrica si fos el cas que no s'ajustés al nostre problema el tradicional ús de la pressió en superfície o d'algun nivell de geopotencial.

Però, i el problema computacional? Efectivament ens vam trobar amb aquesta limitació quan vam voler portar endavant els nostres càlculs (aplicació a exemples concrets que es mostren en els diferents articles que componen el nucli d'aquest treball). El problema computacional ens va limitar especialment a l'hora d'abordar el càlculs a realitzar en el treball presentat a l'apartat 6.2, en el qual es treballa amb una matriu de grans dimensions. A més, amb l'estructura tipus T i la màquina de què disposàvem ens va resultar impossible calcular-ho. Aquí és on ens vam adonar de la

importància que té la manca d'accés a maquinària de càlcul potent, i que un gran avantatge a tenir en compte quan valorem un mètode de CTC és precisament que requereixi de pocs recursos computacionals.

I com resolíem aquest problema si no podíem canviar de màquina? És aquí on ens vam plantejar la possibilitat de buscar una alternativa a l'ACP amb matriu T, però una alternativa que s'allunyés el mínim possible d'aquesta metodologia. I és per això que vam optar per treballar amb l'ACP amb matriu S, però aportant noves propostes metodològiques dins d'aquesta aproximació que emulessin l'enfocament basat en el mode T.

4.2. PROPOSTA METODOLÒGICA: MÈTODE DE LES PUNTUACIONS EXTREMES

Com hem dit anteriorment, les limitacions de càlcul ens han portat finalment a treballar amb l'ACP amb matriu tipus S, mètode al que hem introduït algunes modificacions.

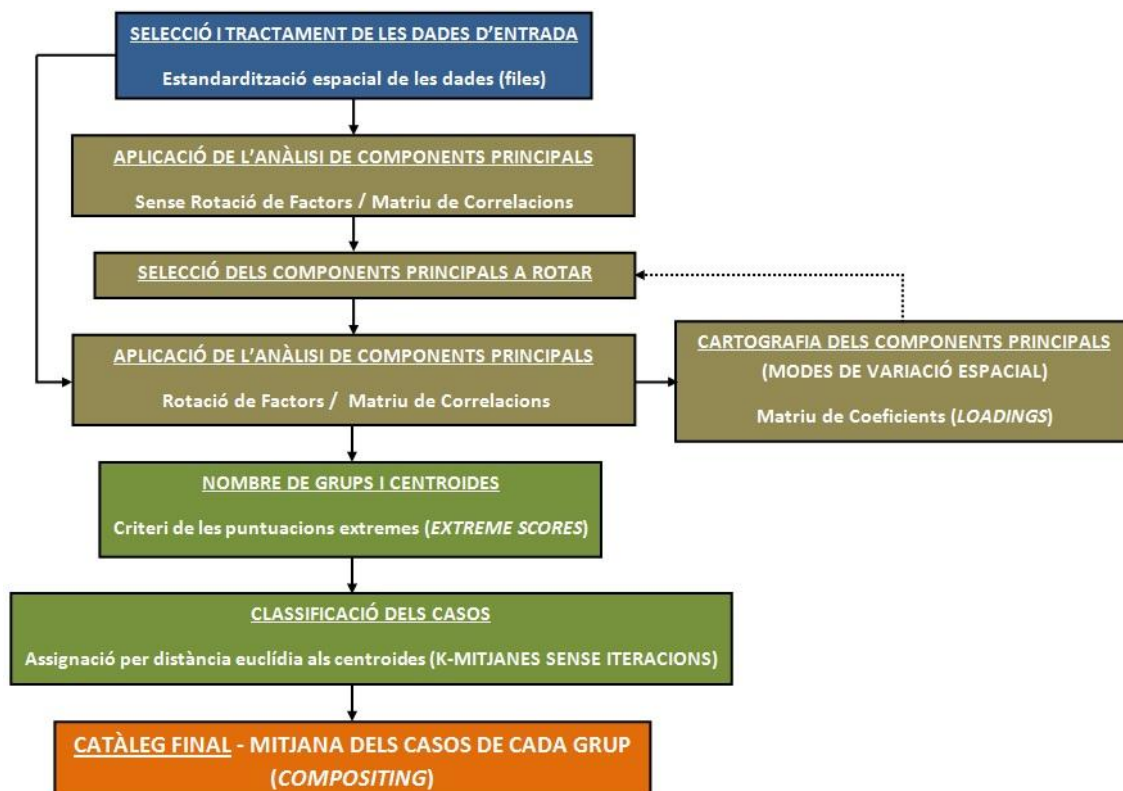


FIGURA 4.1: Esquema gràfic que representa la proposta metodològica que es fa en el present treball. Adaptat a partir del procediment habitual de classificació dels tipus de circulació amb matriu tipus S (FIGURA 2.16), mostra el procés seguit en aquest treball tenint en compte els aspectes proposats en relació a l'estandardització de les dades, l'establiment del nombre de grups i els seus centroides, i l'assignació dels casos als grups finals.

A la FIGURA 2.16 de l'apartat 2.4.2 havíem vist el procediment seguit habitualment quan s'empra la matriu tipus S per classificar, sempre completat d'un mètode d'agrupacions (jeràrquic i/o iteratiu) per triar el nombre de grups, els centroides, i fer l'assignació final dels casos. A la FIGURA 4.1 podem observar l'esquema metodològic

que descriu el mètode proposat en aquest treball. Aquest segueix en gran mesura l'ACP amb matriu tipus S però inclou algunes variacions. Analitzem aquells passos en els quals hem introduït noves propostes:

4.2.1. ESTANDARDITZACIÓ ESPACIAL DE LES DADES

Quan treballem amb mapes de circulació atmosfèrica hem de tenir en compte alguns aspectes de caràcter estrictament climàtic que poden afectar els resultats i que hem d'evitar introduir en l'anàlisi. El més important de tots són els canvis estacionals que presenta aquesta circulació, especialment pel que fa a les variacions d'intensitat dels centres d'acció. Molts autors citen, en parlar de la dinàmica anual de les configuracions de pressió o de geopotencial, intensitats més marcades a l'hivern que a l'estiu, quan les diferències entre els pols i l'equador són més profundes [Barry i Carleton, 2001], [Yarnal, 1993]. D'aquí sorgeix el concepte d'**extracció de l'efecte estacional**, que es referiria al filtratge d'aquests canvis al llarg de l'any per fer comparables tots els mapes de la mostra. Com diu Yarnal [1993], amb l'extracció de l'efecte estacional sobre els valors absoluts i el gradient de pressió, només introduïm en l'anàlisi mapes comparables, encara que siguin de mesos i estacions diferents.

S'han realitzat diferents aproximacions per aconseguir aquest aspecte, generalment basades en:

$$\text{Estandarditzacions: } Z_i = (x_i - \bar{X}) / S$$

$$\text{Anomalies: } A_i = x_i - \bar{X}$$

On Z_i és el valor estandarditzat (sense unitats) del punt de malla o *gridpoint* i , A_i l'anomalia (mateixa unitat que la dada original) del punt de malla o *grid-point* i , x_i és el valor original observat en aquell mateix punt i . D'altra banda banda, \bar{X} (mitjana) i S (dispersió tipus) s'estableixen de diferent manera segons trobem a la bibliografia.

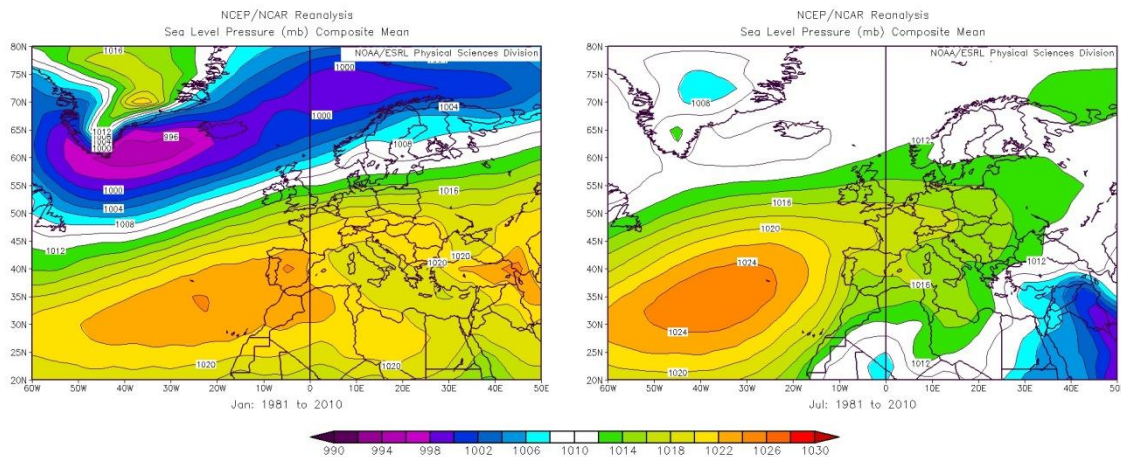


FIGURA 4.2: Variacions estacionals de la circulació atmosfèrica entre gener (esquerra) i juliol (dreta). La mitjana climàtica (1981-2010) de la pressió en superfície del gener ens mostra la baixada d'Islàndia ben formada i profunda, mentre que l'anticicló de les Açores es presenta més desdibuixat i s'estén cap al continent en forma de pont anticiclònic. Al juliol, l'anticicló de les Açores s'estructura molt, alhora que estén la seva influència més al nord. En canvi, la baixa d'Islàndia queda molt desdibuixada. Tot això fa que els gradients bàrics siguin més marcats sobre l'Atlàntic Nord durant el període hivernal, mentre que al juliol es relaxin en aquesta zona i es reforcin en d'altres sectors com la costa occidental africana. Font: NCEP/NCAR

Així, **Cohen [1983]** utilitza l'estandardització, sent \bar{X} i S la mitjana mensual i la desviació tipus mensual dels *grids* respectivament.

Una altra aproximació és la d' **Ekström et al., [2002]**, que estimen el primer harmònic per a la sèrie temporal de *grids* i estableixen així la durada mitjana dels episodis sinòptics. El valor d'aquest harmònic s'utilitza per fer una mitjana mòbil sobre aquesta sèrie de malles de punts suavitzant, doncs, la sèrie. Així, **Ekström et al., [2002]** creuen que un període de setze dies per a el primer harmònic és significant en un 61% de les sèries temporals, de manera que finalment apliquen un filtre de mitjana mòbil de quinze dies. En definitiva, \bar{X} serà la mitjana mòbil de quinze dies, sent el dia central aquell al qual correspon el punt de malla que estem transformant, per obtenir finalment anomalies.

Rasilla [2003] fa una aproximació similar, però amb l'establiment de forma subjectiva de tretze dies com el període per a la suavització, per acabar finalment transformant els punts de malla en anomalies.

Per la seva banda, **Rogers [1981]** utilitza un període de trenta-un o trenta-dos anys per calcular la mitjana estacional de la sèrie de malles de punts, a partir de la qual calcula les anomalies dels punts de *grid* segons la seva distribució al llarg de l'any.

Força diferent és el cas de **Maheras et al., [1999]**, en l'estudi del qual el valor \bar{X} que s'utilitza és el de la mitjana mensual que presenta aquell punt que estem transformant, sense tenir en compte la resta de la malla, actuant de la mateixa manera amb les desviacions. Establiríem doncs, la Z_i del punt de malla X_i a partir del seu valor habitual al llarg de tota la sèrie o part d'ella.

En definitiva, podem considerar dues grans aproximacions per extreure l'efecte estacional a partir del càlcul dels valors estàndard o anomalies:

- En relació al conjunt del *grid*: s'estableixen mitjanes i/o desviacions del *grid* al qual pertany el punt a transformar. Aquestes mitjanes i desviacions es poden suavitzar amb mitjanes mòbils de diferent escala temporal (sobretot a partir de l'establiment de la durada normal d'una situació sinòptica) o, d'altra banda, es poden establir els seus valors mensuals, estacionals, anuals...
- En relació a la sèrie temporal del punt del *grid*: s'estableixen les mitjanes i les desviacions a partir de la sèrie temporal del punt de *grid* que estem transformant. Així establím les anomalies i els valors estàndards a partir dels valors habituals d'aquell punt en un període de temps concret, sigui diari, mensual, anual...

Quina és la millor opció? De nou hem de tenir present els objectius que persegueix el nostre estudi. Així, determinades estandarditzacions o càlcul d'anomalies, especialment relacionades amb l'ús de les mitjanes mòbils, tendiran a suavitzar (a fer més climàtics) els mapes a emprar. Aquest fet pot resultar interessant per alguns estudis, però pot ser contraproductiu en altres treballs amb intencions més meteorològiques que intentin tractar determinats riscos naturals. Pensem que certes situacions sinòptiques, relacionades amb episodis extrems (precipitacions intenses, per

exemple), poden tenir una durada temporal molt reduïda, i l'aplicació de determinades suavitzacions pot amagar la morfologia diària del patró.

D'altra banda, cal plantejar-se fins a quin punt ens pot interessar eliminar el cicle anual de les dades. En determinats estudis ens pot interessar diferenciar entre patrons similars de forma, però de marcada diferència pel que fa a la seva intensitat; fet que vindrà donat, en gran mesura, per la seva distribució anual. D'aquesta manera, no només classificaríem a partir de la morfologia similar, sinó que també introduiríem el gradient a l'anàlisi.

En els casos que ens ocupen, orientats a resoldre fenòmens meteorològics i climàtics relacionats amb l'escala diària, s'ha optat per l'estandardització espacial de les dades (files de la matriu). Això ha de permetre mantenir l'estructura espacial dels camps atmosfèrics però sense perdre la seva intensitat. Així doncs, tenint present al mateix temps que l'ús de la matriu de correlacions (veure punt b) estandarditza les columnes, és a dir, estandarditza temporalment, les diferències estacionals també queden corregides.

4.2.2. ESTABLIMENT DEL NOMBRE DE GRUPS I DELS SEUS CENTROIDES

Quan treballem amb l'ACP amb matriu tipus S , una decisió que no és fàcil de justificar tant si és emprant mètodes jeràrquics de classificació (interpretació del dendrograma), com mètodes globals (llavors aleatòries per fer córrer l'algoritme K-mitjanes) és decidir el nombre de grups i/o els centroides que es necessiten per a classificar els casos. És precisament en aquest àmbit que s'engloba la proposta que presentem a continuació.

La idea es basa en l'establiment dels centroides inicials requerits per a la classificació utilitzant la sèrie temporal de puntuacions factorials obtinguda un cop feta la rotació ortogonal (mètode VARIMAX). Recordem que quan treballem amb la matriu S , les puntuacions factorials o *scores* ens informen del grau de representativitat que tenen cadascun dels modes de variació associats a les components principals (cartografia dels coeficients o *loadings* rotats, que inclouen un significat físic) en relació als nostres

casos originals. En aquest sentit, els podem utilitzar per establir el nombre de tipus de circulació, així com els seus centroides.

Per a definir els tipus, es planteja el **criteri de les puntuacions extremes (*extreme scores*)**. Així, per cada fase (positiva i negativa) de cada component principal (CP) es crea un tipus de circulació associant-hi només aquells casos del total de la mostra que presenten valors alts absoluts de les puntuacions factorials per aquella CP i fase (normalment valors superiors a +2 i inferiors a -2, però que es poden variar a +1/-1 quan la mostra no és suficientment gran/té relativament poca variabilitat), i al mateix temps s'assignen valors baixos (entre +1 i -1) per a la resta de components principals (FIGURA 4.3). Seguint a grans trets aquest plantejament d'emprar llindars de les puntuacions factorials, també han treballat autors com [Tait i Fitzharris \[1998\]](#), [Birkeland *et al.*, \[2001\]](#) i [Kostopoulou \[2003\]](#).

Amb el plantejament fet, podem assumir que cadascun dels casos de la submostra que hem creat (normalment entre un 2,5 i un 5% del total) està molt ben representat per la CP amb la qual se'l relaciona, i que exclusivament aquest component principal és representatiu d'aquest cas que és fruit de l'ortogonalitat dels components principals rotats amb els quals treballem. En altres paraules, el mode de variació espacial associat a la fase d'un component principal ha de representar com a mínim, seguint el criteri abans exposat, un cas de la mostra excepte en el cas que aquesta fase no sigui un artefacte de l'ACP. Això, a més implica que potencialment podem tenir un catàleg de dues vegades el nombre de components principals retinguts (recordem que tenim en compte les fases positives i negatives). Aquest total de grups del nostre catàleg pot reduir-se si, com hem dit anteriorment, hi ha una fase que no tingui assignat un cas, sent un artefacte o un forçament necessari de l'ACP més que no pas un mode de variació climàticament rellevant.

Quan el nombre total de tipus queda establert, els corresponents centroides es calculen fent la mitjana dels casos assignats (mitjana de les puntuacions factorials). Aleshores la referència a l'espai multivariant per classificar és, ara sí, tota la mostra i completament així el nostre catàleg.

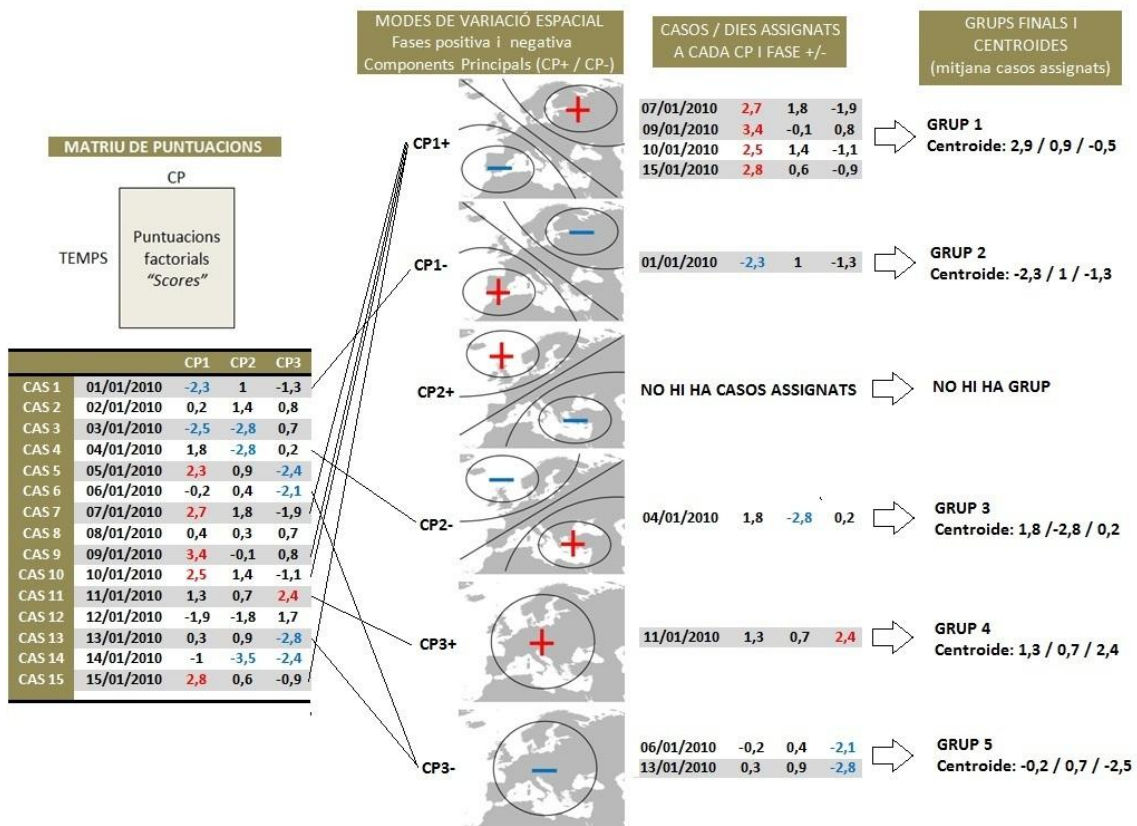


FIGURA 4.3: Esquema on es mostra el mètode de les puntuacions extremes per establir el nombre de grups d'una classificació i calcular els seus centroides. Només alguns casos de la mostra original compleixen el criteri de les puntuacions extremes, i aquest casos s'assignen a la corresponent fase (positiva o negativa) dels components retinguts i rotats (en aquest cas, tres). La mitjana dels casos assignats ens serveix per establir les coordenades dels centroides que serviran com a referència en l'espai multivariant per classificar la resta de la mostra. Com que a una de les fases (en aquest cas la CP2+) no se li ha assignat cap cas, del total de sis grups que potencialment podia tenir la nostra classificació, finalment en tindrà cinc.

4.2.3. CLASSIFICACIÓ DELS CASOS

En el punt anterior hem vist com obteníem el nombre de grups i els seus centroides. Per classificar la resta de casos només ens cal, per exemple, emprar un mètode d'optimització global com K-mitjanes on el nostre nombre de grups i els seus centroides facin la funció de punts de partida de la classificació.

Tot i això, i tenint en compte les consideracions fetes per autors com [Huth \[1996\]](#) i [Philipp \[2007\]](#) que parlen de tendència a crear grups d'igual dimensió o de la dificultat d'obtenir un òptim global per part del mètode (veure apartat 2.5.1), ens va semblar oportú no emprar iteracions en el procés final de classificació. Això vol dir, doncs, que l'assignació final de tots els casos als diferents grups es fa per càlcul de la distància euclídia, partint del supòsit que els centroides establerts són ja representatius de tipus de circulació ben establerta gràcies al procediment de les puntuacions extremes.

5. APLICACIÓ DE LA METODOLOGIA A CASOS DIVERSOS PER A LA SEVA VALIDACIÓ

Una proposta metodològica com la que es presenta en aquest treball no té millor prova que la seva aplicació en casos concrets en que es puguin fer rellevants tant els seus encerts com els seus defectes. En aquest sentit, s'ha optat per aplicar la proposta metodològica de les Puntuacions Extremes (*extreme scores*) en tres camps rellevants de la recerca climatològica actual:

- A. METEOROLOGIA DE MUNTANYA: EL CAS DE LES NEVADES INTENSES A ANDORRA.
- B. VARIABILITAT CLIMÀTICA I CANVI CLIMÀTIC: CATÀLEG DE CIRCULACIÓ DE L'EUROPA OCCIDENTAL, DISTRIBUCIÓ ANUAL I TENDÈNCIES RECENTS
- C. CARTOGRAFIA CLIMÀTICA: INTERPOLACIÓ ESPACIAL DE LA TEMPERATURA I LA PRECIPITACIÓ EN ZONES DE RELLEU COMPLEX A PARTIR DELS TIPUS DE CIRCULACIÓ.
- D. RISCOS METEOROLÒGICS: ACTIVITAT DE LLAMPS A ANDORRA I CATALUNYA

6. COMPENDI D'ARTICLES

6.1. PATRONS DE CIRCULACIO ATMOSFÈRICA RELACIONATS AMB DIES DE NEVADA INTENSA A ANDORRA (PIRINEUS). *INTERNATIONAL JOURNAL OF CLIMATOLOGY* 25, 319-329. ANY 2005.

Les nevades intenses a les regions de muntanya són sovint una causa directa de les allaus de neu. Situacions sinòptiques específiques són responsables d'aquests tipus de nevades extremes, fet que també es dona a Andorra, un país petit en extensió localitzat als Pirineus, entre França i Espanya. Basat en dies amb una intensitat de la nevada d'almenys 30cm de neu nova en 24 hores, el present estudi utilitza l'Anàlisi de Components Principals i tècniques de *clustering* per caracteritzar els tipus de circulació sinòptica per a aquests dies durant el període hivernal. L'àrea d'anàlisi correspon a la regió 30°N-60°N i 30°O-15°E i el període cobreix les temporades hivernals 1986-1987 fins a 2000-2001. La metodologia proposada inclou un processament previ de les dades consistent en l'estandardització espacial de les dades utilitzades posteriorment a l'ACP, una aproximació alternativa per triar els centroides i el nombre de grups pel procés d'agrupacions K-mitjanes, i el no ús de les iteracions d'aquest algoritme, de manera que l'assignació es fa directament amb distàncies euclidianes. Aquesta aproximació permet la classificació sinòptica de qualsevol dia amb nevades intenses, i en aquest sentit s'han construït els corresponents mapes mitjans per a cada tipus obtingut de les variables pressió en superfície, geopotencial a 500 hPa., i gruix geopotencial entre 1.000 i 500 hPa (les línies 5.270m, 5.400m i 5.520m). Els resultats mostren set patrons de circulació, la majoria d'ells amb component atlàntic del vent, així com d'altres de caire mediterrani que es pot combinar amb aire fred continental. Els resultats, en format de mapes del temps, poden ser útils per assistir en l'ús dels models meteorològics en casos de nevades intenses, i la classificació diària obtinguda obre noves possibilitats futures d'anàlisi detallat tant meteorològic com climàtic dels tipus establerts.

ATMOSPHERIC CIRCULATION PATTERNS RELATED TO HEAVY SNOWFALL DAYS IN ANDORRA, PYRENEES

PERE ESTEBAN,^{a,*} PHILIP D. JONES,^b JAVIER MARTÍN-VIDE^a and MONTSE MASES^c

^a *Group of Climatology, University of Barcelona, Catalonia, Spain*

^b *Climatic Research Unit, University of East Anglia, Norwich, UK*

^c *Centre de Recerca en Ciències de la Terra, IEA, Encamp, Andorra*

Received 26 January 2004

Revised 9 August 2004

Accepted 9 August 2004

ABSTRACT

Heavy snowfalls over mountain regions are often a direct cause of avalanches. Specific synoptic-scale atmospheric situations are responsible for these kinds of extreme snowfall event, and this is indeed the case for Andorra, a small country located in the Pyrenees, between France and Spain. Based on days with an intensity of at least 30 cm of snow in a 24 h period, the present study uses principal component analysis (PCA) and clustering techniques to characterize the synoptic circulation patterns for these days during the winter season. The area of analysis encompasses the region 30–60°N by 30°W–15°E and the period covers the winter seasons from 1986–87 to 2000–01. The methodology proposed involves a preprocessing approach consisting of a spatial standardization of the data used for the PCA, an alternative approximation to decide the centroids and the number of groups for the *K*-means clustering, and the rejection of the iterations for this algorithm. This approach enables the synoptic classification of every heavy snowfall day, and composite maps were constructed for sea-level pressure, 500 hPa geopotential height, and 1000–500 m thickness (the 5270 m, 5400 m and 5520 m contour lines). The results show seven circulation patterns, most of them with an Atlantic component of the wind, and others with a clear Mediterranean advection that could be combined with cold continental air. The results, as weather charts, could be a useful tool to assist meteorological models in heavy snowfall forecasting, and the day's classification obtained opens up future possibilities for detailed meteorological and climatological analysis of the established types. Copyright © 2005 Royal Meteorological Society.

KEY WORDS: heavy snowfall days; principal component analysis; clustering; atmospheric circulation patterns; Pyrenees; Andorra; avalanche forecasting

1. INTRODUCTION

In high-elevation mountain environments, snow avalanches present a severe risk to society. Snow avalanches impact the ski resort industry, outdoor activities such as snowmobiling and mountain climbing, homes, transportation networks, and vegetation. Andorra, a good example of an area affected by snow avalanche hazards, is located in the Pyrenees, between France and Spain (Figure 1). It is a small and hilly country with the highest altitudes around 2900 m (Figure 2). Andorra has a sub-Mediterranean climate, characteristic of the northern Mediterranean coast. Some continental influences that introduce changes in the typical seasonal precipitation regime, altitude, and south-facing location are clearly important with regard to the Andorran climate. Nevertheless, the low and well-oriented topography of the French side of the Pyrenees enables the easy entrance of Atlantic air masses from the north and northwest, especially affecting areas of high elevations and sometimes creating extreme snow avalanche events. A well-documented avalanche episode that took place near the ski resort of Arinsal in 1996 did not cause casualties, but it did cause heavy destruction of buildings.

* Correspondence to: Pere Esteban, Baldiri i Reixac s/n, Torre A 5^a planta, 08028, Barcelona, Catalonia, Spain; e-mail: pere.esteban@ub.edu

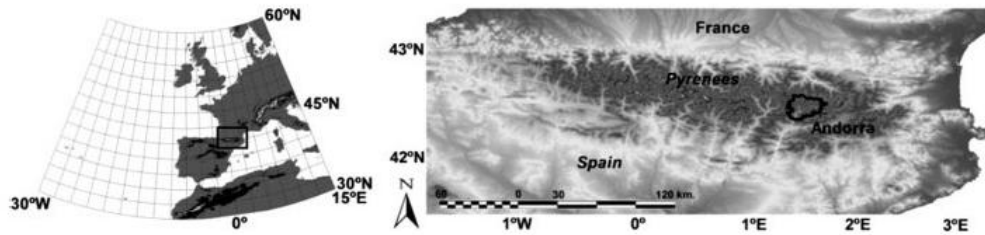


Figure 1. Left: the synoptic area considered for this study. A grid of 2.5° resolution, extracted from the National Centers for Environmental Prediction (NCEP)–National Center for Atmosphere Research (NCAR) reanalysis project, ensures a good resolution of atmospheric data, with 257 points. Right: the location over Andorra on the eastern side of the Pyrenees



Figure 2. The digital elevation model of Andorra shows the location of the weather stations used in this study (names quoted on the right-hand side of the figure). All of these stations are ski resorts, and the altitude and the corresponding coordinates are also shown

This extreme occurrence is an ideal example of an event that it would be useful to predict in order to mitigate avalanche hazards in Andorra (Esteban *et al.*, 2002).

There is no a universal precipitation threshold from which the beginning of snow avalanches is expected. A values of 30 cm of fresh snow is widely used, identified in Alta, Utah, when the precipitation is accompanied by wind (McClung and Schaerer, 1993). On the basis of the avalanche experience in the French Alps, Ancy and Charlier (1998) consider that snowfalls of more than 30 cm involve a moderate avalanche risk. For average precipitation intensity, some forecasters have quoted values of snow water equivalent of between 0.5 and 2.5 mm/h (McClung and Schaerer, 1993). Nevertheless, the avalanche community considers every threshold value only relevant to a certain mountain range and its climate regime, and it is difficult to establish a common threshold. In the eastern Pyrenees (where Andorra is located), the threshold of 30 cm of fresh snow in 24 h is an empirical rule of thumb for avalanche forecasters and researchers in Andorra (CRECIT, 2003) and Catalonia (ICC, 2003), involving the possibility of natural fresh snow avalanches on the steepest slopes. This is why this threshold of 30 cm of fresh snow in 24 h is the one used in this paper.

Several previous studies have examined the relationships between circulation patterns and snow avalanches. Younkin (1968) examined heavy precipitation events using a threshold of 4 in (10.16 cm) or more over a 12 h period, also considering synoptic patterns of 1000–500 hPa thickness, geopotential height at 1000 hPa and/or at 500 hPa, and temperature at 500 hPa for the western USA during the snow seasons from 1963 to 1967. Using a compilation of 130 years of avalanche deaths in Norway, Fitzharris and Bakkehøi (1986) examined the synoptic climatology with the Grosswetterlagen index of atmospheric circulation, together with the monthly sea-level pressure (SLP) anomaly maps of the 12 major avalanche winters. Plaut *et al.* (2001), in a similar approach but using multivariate statistics, characterized the 700 hPa large-scale circulation patterns related to heavy precipitation events (floods, intense snowfall and avalanches) over the Alpine subregion. Birkeland

and Mock (1996) examined atmospheric circulation patterns at 500 hPa levels based on heavy snowfall events greater than 32.8 cm at Bridger Bowl (Montana, USA); and Mock and Birkeland (2000) analysed the snow climatology of the mountain ranges of the western USA, regionalizing coastal, intermountain, and continental climates and using a binary avalanche and climate classification. These authors illustrated that spatial variations of avalanche climates are closely related to 500 hPa anomalies. Birkeland *et al.* (2001) characterized composite-anomaly maps at 500 hPa geopotential height for extreme avalanche events for four sites located in the western USA, based on a daily avalanche index that reflected avalanche frequency and size. These avalanche indices define abnormal avalanche events; principal component analysis (PCA) and clustering of meteorological variables enabled the classification of unique avalanche extremes. Birkeland and Mock (2001) examined a major snow avalanche cycle that affected the western USA in February 1986, analysing the avalanche, snowpack, and weather conditions for 20 sites. They also analysed 500 hPa anomaly maps and showed that the event resulted from an uncommonly persistent blocking pattern.

This paper identifies and describes the synoptic patterns of SLP and other atmospheric levels that cause heavy snowfall events in Andorra, using the threshold of ≥ 30 cm of fresh snow in 24 h. PCA and clustering techniques are used in the synoptic climatological analyses.

2. DATA AND METHODOLOGY

Daily weather data (at 08:00 h local time) from six Andorran ski resorts (Figure 2) were analysed in order to select days of heavy snowfalls. As mentioned above, a heavy snowfall day was defined as one having at least 30 cm or more of fresh snow in a 24 h period for at least one of the six stations used.

Data for this study cover 15 winter seasons (December–April) from 1986–87 to 2000–01, with reasonably good spatial coverage of weather stations in Andorra (Figure 2). SLP data at 18:00 GMT for the previous day were selected from the NCEP–NCAR reanalysis data (Kalnay *et al.*, 1996), encompassing the region 30–60°N latitude by 30°W–15°E longitude, with a 2.5° spatial resolution (Figure 1).

A PCA was conducted on the SLP data to derive the dominant patterns of variability from the raw fields (Richman, 1986; Yarnal, 1993). Later, clustering was used to classify the observations and to composite the circulation patterns maps (Yarnal, 1993).

PCA and clustering techniques have been widely used to obtain daily circulation pattern classifications related to several environmental problems. Analysing the obliquely rotated principal components as a map typing technique, Richman (1981) classified three-hourly SLP data obtained from the National Severe Storm Forecast Center, relating the results to precipitation events. Kidson (1994) clustered 9 years of daily analyses (at 00:00h UTC) of 1000 hPa geopotential height over the New Zealand region, identifying 13 synoptic groups, assuming a less than perfect representation of the low-pressure centres. Over the southwestern USA, Galambosi *et al.* (1996) applied these multivariate methodologies to daily 500 hPa pressure fields at 12:00 UTC, coupling PCA and clustering in order to describe and generate local hydrological variables. Applied to wind erosion events in sandy soil agricultural areas of southern Sweden, Ekstrom *et al.* (2002) classified daily SLP circulation data and identified (via the definition of an index of erosion severity) the patterns associated with large erosion damage. For the South Pacific–South America–South Atlantic region, Solman and Menéndez (2003) classified 34 years of winter daily 500 hPa geopotential height data, identifying five weather regimes as the most frequent large-scale circulation patterns. For the Iberian Peninsula area, Romero *et al.* (1999) identified 19 atmospheric circulation patterns related to significant rainfall days for Mediterranean Spain, combining the results with the precipitation spatial patterns obtained by the same authors via PCA and clustering techniques. Recently, Rasilla (2003) obtained 10 circulation patterns for the Iberian Peninsula, using PCA and the combination of manual and automated clustering. Other studies and applications can be found: Yarnal *et al.* (2001), who discuss the past, present, and future of the synoptic climatology, include several references on the application of this eigenvector-based analysis to this research field.

2.1. PCA

PCA is based on an S-mode data matrix (Richman, 1986), the variables being the grid points and the days being the observations. A preprocessing of the data is applied, scaling the rows, i.e. standardizing the observations from a spatial point of view. This procedure deseasonalizes the observations, keeping the fields' intensities in the analysis, and maintains the daily temporal scale of the original data. A correlation matrix provides the most efficient representation of variance in the data set (Barry and Carleton, 2001), and patterns of spatial correlation can be detected without possible domination by the grid points with the largest variances (Jolliffe, 1986). The number of components explaining a significant portion of the total variance (Kalkstein *et al.*, 1987; Barry and Carleton, 2001) was determined by means of the Scree test (Cattell, 1966), the explained variance criteria, the Kaiser method, and the North rule of thumb (North *et al.*, 1982). Next, the retained components were rotated using a varimax procedure to facilitate the spatial interpretation of the principal components (Yarnal, 1993) and avoid problems like domain shape dependence, but retaining the orthogonality constraint of the eigenvector model (Richman, 1986).

2.2. Clustering

PCA yielded component scores, which are standardized values that represent the relationship between the observations and the retained orthogonal components. The non-hierarchical *K*-means method is used to cluster the observations (Hair *et al.*, 1998). To decide on the number of groups and the centroids needed for the classification, we considered the spatial variation patterns established by the PCA, i.e. the principal components in the positive and negative phases as potential groups for circulation patterns eventually used for the *K*-means. To create the centroids of this groups, observations with high score values for a certain component were selected (values higher than +1 for the positive phase, or lower than -1 for the negative phase, but being between +1 and -1 for the remainder of the PCs retained) (Birkeland *et al.*, 2001; Tait and Fitzharris, 1998). Using this technique, some components in their positive or negative phase could not present any assigned case (day), and so this potential group (circulation pattern) was eliminated. The centroids were obtained from the average of the cases included into every group finally considered. Eventually, the *K*-means produces the final classification of all the observations (days) with similar distribution of the SLP fields. Non-iteration of the *K*-means clustering was used, as the centroids were well established by the first PCA grouping (reflecting a circulation pattern). Additionally, some experiments were performed, showing a clear trend of the iterations to create equal-sized groups.

Synoptic maps of the atmospheric circulation groups were constructed for SLP, 500 hPa geopotential heights, and 1000–500 hPa thickness (the '5270 m', '5400 m' and '5520 m' contour lines). The 1000–500 hPa thickness maps are used in meteorology to identify warm or cold air masses and their boundaries (fronts). When the thickness value becomes low enough, snow can reach the surface and, commonly, the 5400 m thickness is used as the non-snow/snow-line rule of thumb. Thicknesses of 5400 m or lower indicate that snow is most likely (50% of the time a 5400 m thickness will produce snow at elevations below 300 m) and thickness values greater than 5400 m usually indicate non-snow precipitation at low altitudes (Haby, 2003). In high-elevation regions, snow is common even when the 1000 to 500 hPa. thickness is 5460 m or 5520 m. This could be explained due to the high-elevation areas located at lower pressure levels, closer to mid-level colder air (Haby, 2003). Otherwise, for the area of northwestern Europe, Lamb (1955) obtained a fairly regular changeover of precipitation from unfrozen to frozen forms at the 5270 m line.

3. RESULTS

Six principal components were retained, based on the results of the different selection rule procedures (Figure 3), which together explain 89.7% of the total variance. Finally, seven groups representing the most relevant circulation patterns related to heavy snowfalls over Andorra (Figures 6 and 7) were identified from the PCA and clustering methods, permitting the classification of all the days into these groups. The seasonal distribution and evolution of the clusters during the period studied are shown in Figure 4. Table I shows the

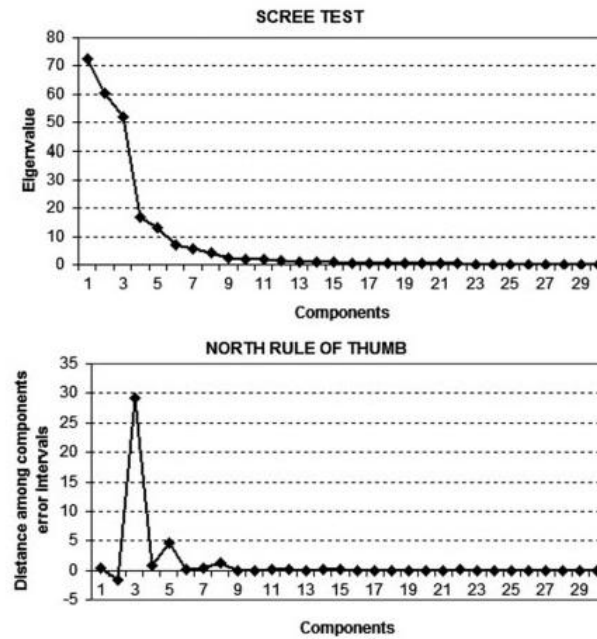


Figure 3. Scree test and North rule of thumb, two empirical techniques, used in the selection of principal components to be retained for rotation

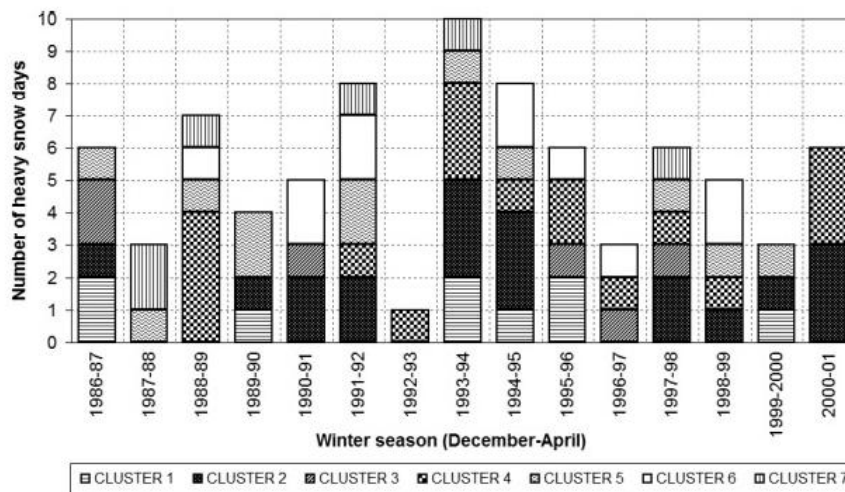


Figure 4. Winter-season frequencies of the number of heavy snowfall days with at least 30 cm of snow in a 24 h period from 1986–87 to 2000–01. The bars are shaded to reflect the number of days for every cluster

number of days classified in each cluster and their monthly distributions, and are represented as frequencies in Figure 5. Average and maximum snow precipitation in 24 h for every cluster are shown for three weather stations (Table II), reflecting the relationship between the circulation patterns obtained and the local conditions.

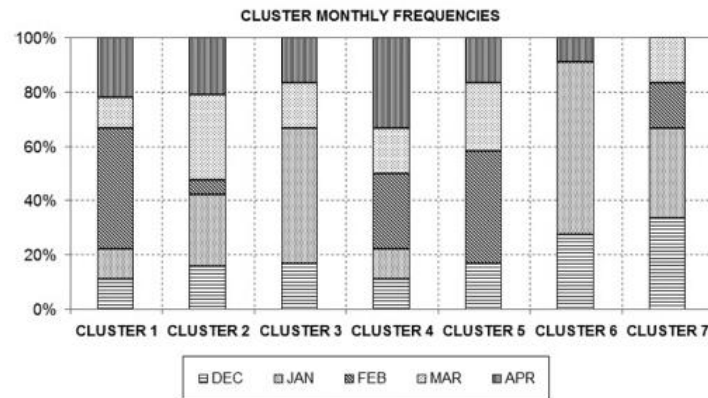


Figure 5. Monthly frequencies of the clusters obtained

The analysis of the weather charts shows that advective situations, with surface winds blowing from the northwest to the north, are the most frequent atmospheric patterns related to heavy snowfall days discussed in this research. These situations are followed by east to northeast patterns, and one situation with southerly flow. In all cases, three elements are relevant: moisture coming from the Atlantic or the Mediterranean, cold air advected from northern latitudes (especially when it comes from eastern continental Europe), and vertical motion (thanks to the combination of strong winds, the uplift produced by the Pyrenees orographic barrier, and the 500 hPa configurations, which are troughs or cut-off lows).

The first pair of charts (Figure 6, Cluster 1) shows a strong Azores high clearly extended over the Iberian Peninsula and a low pressure over the Mediterranean basin. This is a quite frequent situation, which occurs during the whole year. The orographic effect of the Pyrenees and the Alps generates and strengthens northerly winds over the area (genesis of the Mistral and Tramuntana local winds), which, in combination with the high levels of moisture of the Atlantic air mass, creates instability and precipitation. Additionally, the influence of the Siberian high over Europe, which usually appears as an anticyclonic bridge with the Azores high, means cold air advection over the study area, ensuring snow precipitation in the Pyrenees mountains and Andorra (see the 5270 m contour line affecting the Alps, and the 5400 m contour line clearly reaching the Pyrenees on the 1000–500 hPa thickness map). A similar explanation could be applied to the second cluster (Figure 6, Cluster 2). The Azores high is extended northwards, pumping cold and relatively humid air from the North Sea. On the other hand, a low pressure located over the Tyrrhenian Sea advects, as in the first cluster, very cold air from northeastern Europe. The latter is related to the winter semi-permanent Siberian high (note the southerly location of the 5400 m line on the 1000–500 hPa thickness map and the marked north–south configuration of the isohypses at the back of the deep trough). Cluster 6 (Figure 6) shows the last configuration involving cold continental air, but in this case mixing with the humid and relatively warm air pumped up by a low pressure located over the western Mediterranean basin. Instability is guaranteed with the 500 hPa cut-off low identifiable over the eastern Pyrenees, and the temperature is low enough in winter to generate heavy snow precipitation in the mountains, in spite of the warming influence of the Mediterranean. This weather pattern, which clearly affects the eastern side of the Pyrenees mountains and Andorra, thanks to the local configuration of the relief, would potentially receive large amounts of snow precipitation (Lefevre, 1998).

The third pair of maps (Figure 6, Cluster 3) shows a clearly different situation. A deep low pressure is located west of Andorra, over the northwest coast of Spain, with southerly surface winds carrying warm and humid Mediterranean air over the Pyrenees. High instability is ensured thanks to the location of a deep cold-core low at 500 hPa, reflecting the surface low pressure already mentioned. In such situations, very common in the winter half of the year, heavy precipitation (torrential in some cases) affects the southern side of the Pyrenees, the snow being quite abundant and humid, but affecting the highest elevations of the Pyrenees and

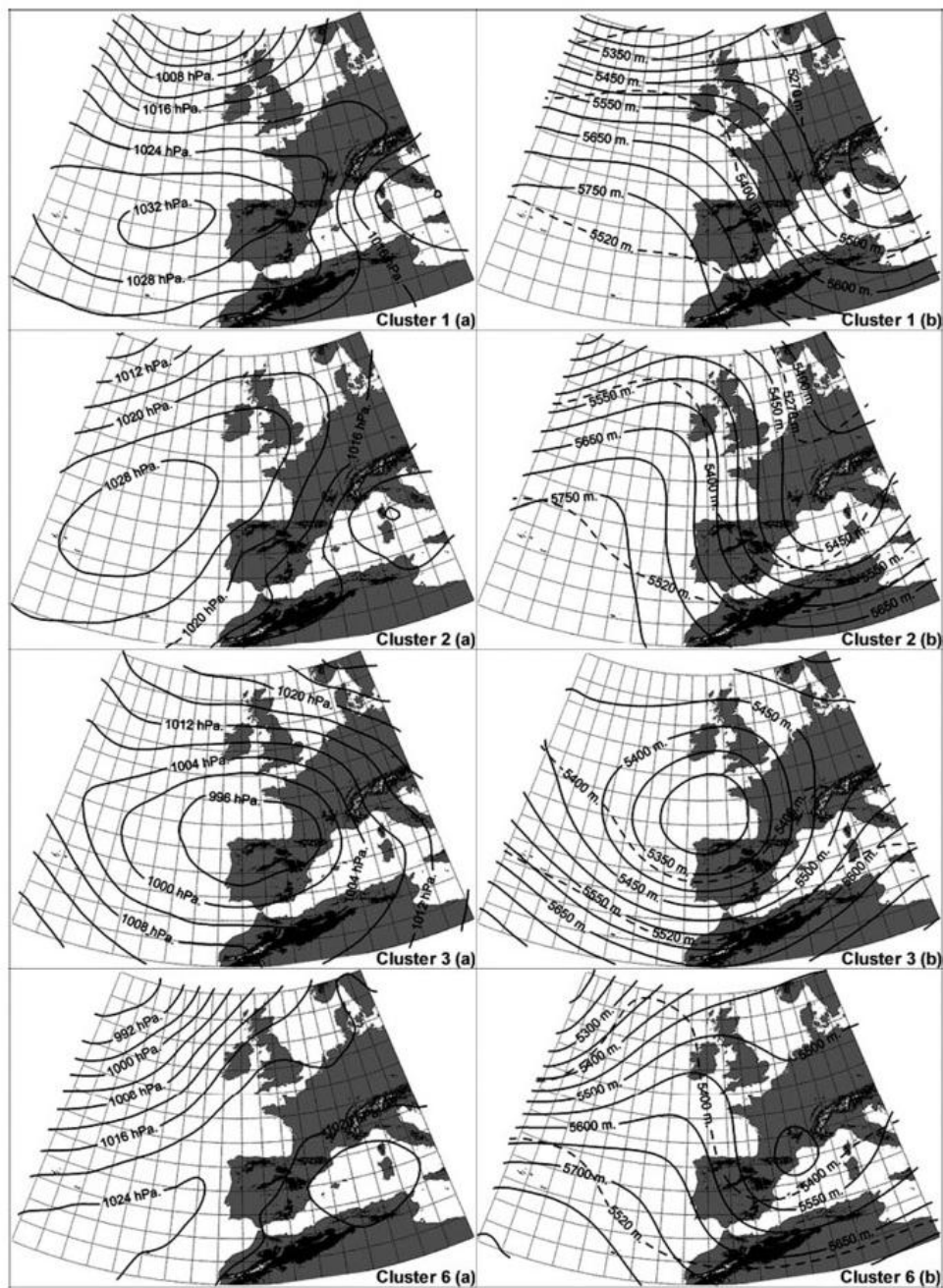


Figure 6. Synoptic circulation patterns based on multivariate techniques (Clusters 1 to 3 and 6) for heavy snowfall events over Andorra (1986–87 to 2000–01). The pairs of maps show: (a) SLP patterns; (b) 500 hPa geopotential heights (solid lines) and the 5270 m, 5400 m, and 5520 m lines of the 1000–500 thickness maps (dashed line)

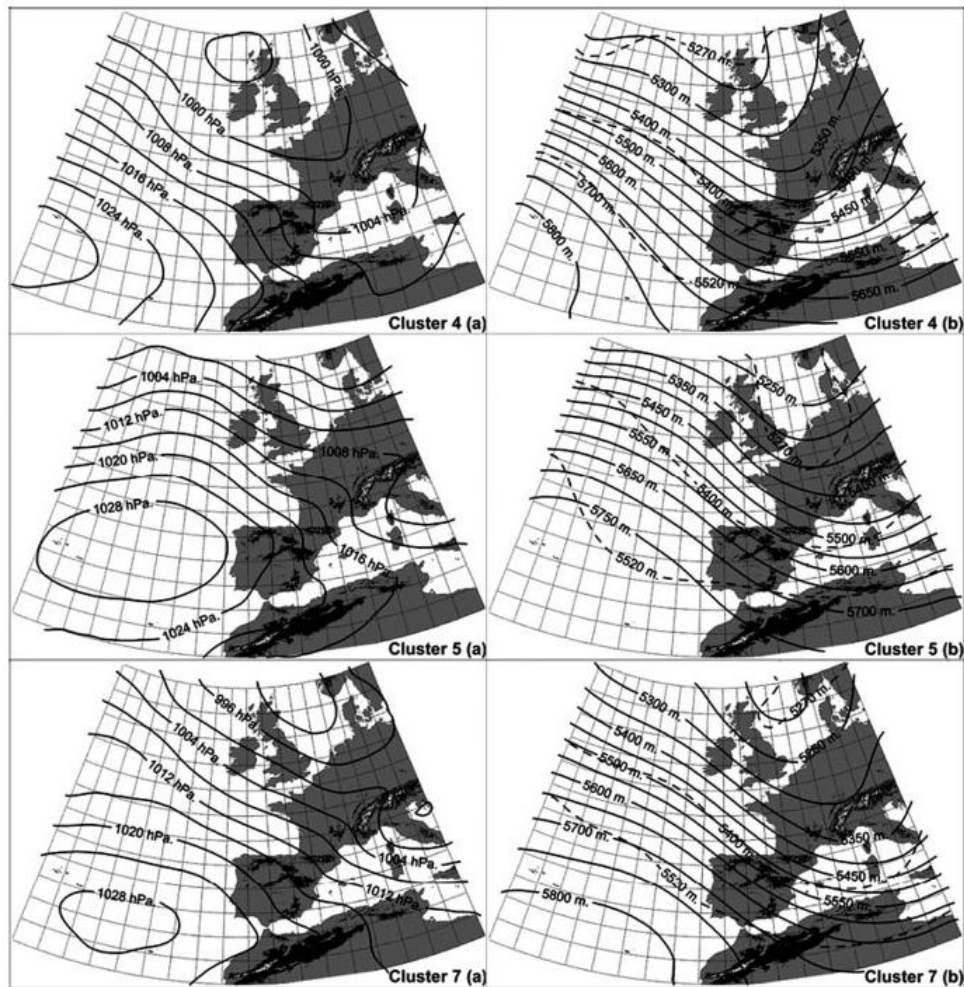


Figure 7. Synoptic circulation patterns based on multivariate techniques (Clusters 4, 5 and 7) for heavy snowfall events over Andorra (1986–87 to 2000–01). The pairs of maps show: (a) SLP patterns; (b) 500 hPa geopotential heights (solid lines) and the 5270 m, 5400 m, and 5520 m lines of the 1000–500 thickness maps (dashed line)

Andorra. Observe, in this regard, the location of the 5400 m line of 1000–500 hPa thickness, reflecting the warm air mass near the surface in contrast with low temperatures at high levels.

The three remaining situations (Clusters 4, 5 and 7) look very similar, and it is possible to think about grouping them into one cluster. In any event, it is interesting to analyse these three situations in detail, and some interesting differences are detected when this is done. Cluster 4 (Figure 7) shows the passage of a cold front over Andorra, related to a low pressure crossing the north of Europe, and blowing northwesterly winds over the Pyrenees. The typical precipitation related to an air mass border and the uplift effect ensure snow in wintertime. It is also interesting to observe the inflection of the 5400 m line in the 1000–500 hPa thickness map, denoting the arrival of cold air. On the other hand, in Cluster 7 (Figure 7) a northwest component of the wind is also identified, with in this case low pressures crossing central Europe, the Pyrenees being affected by these lows. It is also interesting to highlight the large gradient in pressure over the study area, reflecting

Table I. Number of heavy snow days per cluster and month

	Days per month					Total
	Dec	Jan	Feb	Mar	Apr	
Cluster 1	1	1	4	1	2	9
Cluster 2	3	5	1	6	4	19
Cluster 3	1	3	0	1	1	6
Cluster 4	2	2	5	3	6	18
Cluster 5	2	0	5	3	2	12
Cluster 6	3	7	0	0	1	11
Cluster 7	2	2	1	1	0	6
Total	14	20	16	15	16	81

Table II. Average and maximum snow precipitation for every cluster, registered in three of the weather stations used in this work (Ordino-Arcalis, Pas de la Casa, Arinsal)

	Snow precipitation in 24 h (cm)					
	Avg. Ordino	Max. Ordino	Avg. Pas Casa	Max. Pas Casa	Avg. Arinsal	Max. Arinsal
Cluster 1	36	57	27	65	23	32
Cluster 2	31	54	34	60	21	51
Cluster 3	40	65	44	75	45	70
Cluster 4	26	53	30	64	26	85
Cluster 5	28	43	23	42	17	50
Cluster 6	25	48	30	50	23	56
Cluster 7	37	63	34	46	25	45

strong winds with the likely implication of air mass uplift and snow blowing at mountain crests. Finally, Cluster 5 (Figure 7) again shows strong northwesterly winds over Andorra at all tropospheric levels, but with the study area being affected by high pressure values at the surface. The 500 hPa geopotential height and 1000–500 hPa thickness map are quite similar over Andorra in Clusters 5 and 7, with the study area always under the 5400 m line influence and strong northwesterly winds at the middle tropospheric levels.

4. CONCLUSIONS

Multivariate statistics (PCA and clustering) as a computer-assisted method has been used to describe seven circulation patterns that explain heavy snow precipitation days in Andorra based on the winter seasons between 1986–87 and 2000–01. Three of these patterns show advection of Atlantic or Mediterranean airmasses combined with the outbreak of cold air from continental Europe, as is well represented by the 5400 m contour line of the 1000–500 hPa thickness map. There is only one situation with clear southerly winds over Andorra, showing a deep low pressure at all tropospheric levels located over northwestern Spain, and affecting, with snow, only the highest elevations of Andorra. Finally, three maps represent advectations from the northwest, very important in Andorra (in spite of the south-facing location of the country in the Pyrenees mountains), but reflecting interesting differences in the surface maps. In all cases, a high gradient of the isobars is observed, indicating strong winds blowing over the Andorra area and suggesting a possible increase in avalanche hazard for some slopes with easy deposition of blown snow. These winds are also important for creating the uplift needed to support the vertical motions that help generate heavy snow events in the wintertime. Such an amount of fresh snow in a few hours (≥ 30 cm in 24 h) could seriously increase the

instability of the snowpack, leading to potential avalanche events. As a consequence, these maps, and those obtained from other variables, could be very useful for avalanche forecasters. With the days grouped, a deeper analysis of the meteorological variables that explain these events could be made in the future, thus establishing typical pattern sequences that are of interest in avalanche event analysis. This is, for example, the case of the Arinsal event of 8 February 1996 mentioned in Section 1 (Esteban *et al.*, 2002), showing first a situation represented by Cluster 3 (southerly winds), followed by a period of melt–freeze crust formation, and concluding with heavy snowfalls and strong northwesterly winds preceding the big avalanche, as represented in Clusters 4 and 7.

In addition, these circulation patterns could be used to evaluate possible future changes in the snowfall activity over the study area. Snow, an important resource for Andorra, could be seriously affected by climatic change and, by extension, the economy of Andorra would be impacted. A deeper analysis of the patterns in the context of climate variability, and considering the likely atmospheric circulation changes as suggested by general circulation models, might help in the appropriate management of this natural resource.

The method proposed offers new approaches to the eigenvector and clustering classification of circulation patterns. First, a spatial preprocessing of the original data is shown, allowing the introduction of the intensity of the fields to the analysis; second, a new way is shown, via PCA, to establish the centroids and number of groups for the *K*-means clustering algorithm. Finally, thanks to the spatial variation patterns established by the PCA, the *K*-means without iteration avoids equidistribution into the groups.

ACKNOWLEDGEMENTS

We would like to express our thanks to FEDA (Forçes Elèctriques d'Andorra), MeteoFrance, the Andorra ski resorts, Francesc Areny (Industry Department, Andorran Government), and Glòria Furdada (Geophysics and Geodynamics Department, University of Barcelona) for providing the data.

We also wish to extend thanks to the Catalan Government for supporting the Group of Climatology and the FI grant, the Barcelona University Science Park, the Spanish research project CICYT REN2001-2865-C02-01.

Finally, the excellent technical support by Ian Jolliffe, Cary Mock, Karl Birkeland, Marc Prohom and the two anonymous reviewers is gratefully acknowledged.

REFERENCES

- Ancey C, Charlier C. 1998. Les avalanches. In *Guide Neige et Avalanches: Connaissances, Pratiques, Sécurité*, Ancey C (ed.). Édisud: Aix-en-Provence: 87–126.
- Barry RG, Carleton AM. 2001. *Synoptic and Dynamic Climatology*. Routledge: London.
- Birkeland KW, Mock CJ. 1996. Atmospheric circulation patterns associated with heavy snowfall events, Bridger Bowl, Montana, U.S.A. *Mountain Research and Development* **16**: 281–286.
- Birkeland KW, Mock CJ. 2001. The major snow avalanche cycle of February 1986 in the western United States. *Natural Hazards* **24**: 75–95.
- Birkeland KW, Mock CJ, Shinker JJ. 2001. Avalanche extremes and atmospheric circulation patterns. *Annals of Glaciology* **32**: 135–140.
- Cattell RB. 1966. The scree test for the number of factors. *Multivariate Behavioral Research* **1**: 245–276.
- CRECIT. 2003. Center for Earth Sciences Research. www.iea.ad/crecit [July].
- Ekstrom M, Jonsson P, Barring L. 2002. Synoptic pressure patterns associated with major wind erosion events in southern Sweden. *Climate Research* **23**: 51–66.
- Esteban P, Mases M, Martín-Vide J. 2002. Climatology for avalanche forecasting: the Les Fonts avalanche of Arinsal, Andorra, February 8th of 1996. *Horitzó: Revista del Centre de Recerca en Ciències de la Terra* **2**: 10–19 (in Catalan).
- Fitzharris BB, Bakkehoi S. 1986. A synoptic climatology of major avalanche winters in Norway. *Journal of Climatology* **6**: 431–446.
- Galambosei A, Duckstein L, Bogardi I. 1996. Evaluation and analysis of daily atmospheric circulation patterns of the 500 hPa pressure field over the southwestern USA. *Atmospheric Research* **40**: 49–76.
- Haby J. 2003. The 540 line and precipitation type. www.theweatherprediction.com/habyhints/97/ [September].
- Hair JF, Anderson RE, Tatham RL, Black WC. 1998. *Multivariate Data Analysis*, fifth edition. Prentice-Hall International.
- ICC. 2003. Catalan Cartographic Agency. www.icc.es/allaus [July].
- Jolliffe IT. 1986. *Principal Component Analysis*. Springer-Verlag: New York.
- Kalkstein SL, Tan G, Skindlov JA. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Climate and Applied Meteorology* **25**: 717–730.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetma A, Reynolds R, Jeene R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Kidson JW. 1994. An automated procedure for the identification of synoptic types applied to the New Zealand region. *International Journal of Climatology* **14**: 711–721.

- Lamb HH. 1955. Two-way relationship between the snow or ice limit and 1000–500 mb thicknesses in the overlying atmosphere. *Quarterly Journal of the Royal Meteorological Society* **34**: 172–189.
- Lefevre C. 1998. Some characteristic situation. The Pyrenees. In *La Meteorología de Montaña*, Thillet JJ (ed.). Ediciones Martínez-Roca: Barcelona: 189–196 (in Spanish).
- McClung D, Schaerer P. 1993. *The Avalanche Handbook*. The Mountaineers: Seattle.
- Mock CJ, Birkeland KW. 2000. Snow climatology of the western United States mountain ranges. *Bulletin of the American Meteorological Society* **81**: 2367–2392.
- North GR, Bell TL, Calahan RF. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review* **110**: 699–706.
- Plaut G, Schuepbach E, Doctor M. 2001. Heavy precipitation events over a few Alpine sub-regions and the links with large-scale circulation, 1971–1995. *Climate Research* **17**: 285–302.
- Rasilla DF. 2003. Aplicación de un método de clasificación sinóptica a la Península Ibérica. *Investigaciones Geográficas* **30**: 27–45.
- Richman MB. 1981. Obliquely rotated principal components: an improved meteorological map typing technique? *Journal of Applied Meteorology* **20**: 1145–1159.
- Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.
- Romero R, Sumner G, Ramis C, Genovés A. 1999. A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *International Journal of Climatology* **19**: 765–785.
- Solman SA, Menéndez CG. 2003. Weather regimes in the South American sector and neighbouring oceans during winter. *Climate Dynamics* **21**: 91–104.
- Tait AB, Fitzharris BB. 1998. Relationship between New Zealand rainfall and south-west Pacific pressure patterns. *International Journal of Climatology* **18**: 407–424.
- Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press: London.
- Yarnal B, Comrie AC, Frakes B, Brown DP. 2001. Developments and prospects in synoptic climatology. *International Journal of Climatology* **21**: 1923–1950.
- Younkin RJ. 1968. Circulation patterns associated with heavy snowfall over the western United States. *Monthly Weather Review* **96**: 851–853.

6.2. CATÀLEG DIARI DE LA CIRCULACIÓ ATMOSFÈRICA PER SECTOR DE L'EUROPA OCCIDENTAL UTILITZANT TÈCNIQUES ESTADÍSTIQUES MULTIVARIANTS. *INTERNATIONAL JOURNAL OF CLIMATOLOGY* 26, 1501-1515. ANY 2006.

Aquest estudi fa servir l'anàlisi de components principals i tècniques d'agrupació per caracteritzar la circulació sinòptica a escala diària a la regió 30°N-60°N i 30°O-15°E per al període 1960-2001, i utilitza el reanàlisi del NCEP/NCAR. La metodologia usada implica un processament previ de les dades consistent en una estandardització espacial de les dades utilitzades en l'ACP en matriu tipus S, matriu de correlacions i rotació ortogonal Varimax de les components retingudes. Per al procés d'agrupacions, un mètode alternatiu basat en els resultats de l'ACP (emprant les fases positiva i negativa de les components principals) per a decidir els centroides i el nombre de grups per a l'ús de l'algoritme K-mitjanes, però sense iteracions (assignació directa per distància euclidiana i sense reclassificacions). S'han obtingut 20 tipus de tipus de circulació a partir de les dades de pressió en superfície i amb tots els dies classificats. Per a cada tipus s'ha representat el mapa mitjà de pressió en superfície i de geopotencial a 500 hPa, la distribució mensual dels casos i la variabilitat al llarg del període. Els resultats són consistents amb el coneixement subjectiu de la circulació atmosfèrica de la zona en estudi i sembla una aproximació fàcil i accessible per a l'anàlisi sinòptica de fenòmens de curta durada (entre setmanal i diari). A més, la metodologia apareix com a aplicable a totes les àrees climàtiques del món, ja que caracteritza prou correctament la complexa circulació de les latituds mitjanes, incloses les situacions de gradient de pressió baix en superfície que es donen al sector de la mediterrània en el període estival.

DAILY ATMOSPHERIC CIRCULATION CATALOGUE FOR WESTERN EUROPE USING MULTIVARIATE TECHNIQUES

PERE ESTEBAN,^{a,b,*} JAVIER MARTIN-VIDE^{a,*} and MONTSERRAT MASES^c

^a *Group of Climatology, University of Barcelona, Catalonia, Spain*

^b *Servei Meteorològic de Catalunya, Barcelona, Catalonia, Spain*

^c *Centre de Recerca en Ciències de la Terra, IEA, Encamp, Andorra*

Received 1 June 2005

Revised 7 October 2005

Accepted 11 April 2006

ABSTRACT

This study uses principal component analysis (PCA) and clustering (CL) techniques to characterize the daily surface synoptic circulation patterns over the region 30°N–60°N by 30°W–15°E, for the period 1960–2001 using the NCEP/NCAR Reanalysis Project data. The methodology used, previously proposed by Esteban *et al.* (2005) to identify heavy snowfall days over Andorra (Pyrenees), involves a pre-processing approach consisting of a spatial standardization of the data used for the rotated PCA on S-mode analysis and correlation matrix, and Varimax rotation of the retained components. For the clustering process, an alternative approximation based on the PCA results (positive and negative phases) to decide the centroids and the number of groups for the K-means clustering is shown, followed by the rejection of the iterations for this non-hierarchical algorithm. Twenty sea level pressure (SLP) circulation patterns were obtained, with all days classified. The composite maps for SLP and 500 hPa geopotential height (GPH), the monthly distribution, and the long-term variability for each of the circulation patterns is obtained. The results are consistent with the subjective knowledge of the daily atmospheric circulation over the area treated, and seem to be an easy and accessible way for the synoptic studies of short-lived (from weekly to daily) phenomena. In addition, this methodology appears to be applicable to all climatic areas of the world, as it characterizes well enough the complex circulation variability at mid-latitudes, including the surface low-pressure gradient patterns normally located over the Mediterranean in summer. Copyright © 2006 Royal Meteorological Society.

KEY WORDS: daily atmospheric circulation patterns; western Europe; principal component analysis; K-means clustering; spatial standardization; extreme scores

1. INTRODUCTION

The classification of atmospheric circulation is a topic that has been treated by climatologists for decades. As Yarnal *et al.* (2001) define synoptic climatology nowadays, the establishment of empirical relationships between atmospheric circulation and local weather is a main aspect. In this way, atmospheric circulation catalogues try to explain environmental or human phenomena (i.e. heavy snowfalls, droughts, landslides, soil erosion, pollution, mortality, among many others) using weather map statistics. The interest in improving these methods in weather and, especially, in climate change forecasting on a local scale, has generated projects like Atmospheric Circulation Classification and Regional Downscaling (ACCORD, 2005) (<http://www.cru.uea.ac.uk/cru/projects/accord/>), Statistical and Regional dynamical Downscaling of Extremes for European regions (STARDEX, 2005) (<http://www.cru.uea.ac.uk/cru/projects/stardex/>), or the Harmonization and Applications of Weather Types Classifications for European Regions (COST Action 733, 2005) (http://cost.cordis.lu/src/action_detail.cfm?action=733), partly or completely focused on the analysis and development of atmospheric circulation classifications.

* Correspondence to: Pere Esteban and Javier Martín-Vide, Montalegre 6-8, 08001, Barcelona, Catalonia, Spain;
e-mail: pere.esteban@meteocat.com

Traditional manual subjective methods such as those proposed by Lamb (1972) for the British Isles, or Hess and Brezowsky (1952) for Europe, are today replaced by or combined with objective or semi-automated techniques permitting the analysis of large amounts of data with less time and effort. The correlation-based method of Lund (1963) and the sum-of-squares method of Kirchhofer (1973), good reproducers of manual classifications (Yarnal, 1993; Barry and Carleton, 2001) and discussed by some authors (Yarnal and White, 1987; Yarnal *et al.*, 1988), have been applied to the analysis of precipitation and temperature (Bischoff and Vargas, 2003), mass balance of glaciers (Yarnal, 1984), or persistence and transition between circulation patterns on a synoptic scale (El-Kadi and Smithson, 2000). Another approximation is that proposed by Jenkinson and Collison (1977), which can be considered as the automated version of Lamb's weather type classification scheme (Jones *et al.*, 1993) and is based on a set of indices related to the direction and vorticity of the geostrophic flow calculated on an SLP gridded data. Several authors have used this classification method over recent years, particularly to describe or downscale precipitation as the works proposed by Conway *et al.* (1996) over the British Isles, Goodess and Palutikov (1998) for south-eastern Spain, Kilsby *et al.* (1998) for England and Wales, Trigo and DaCamara (2000) for Portugal, Spellman (2000) for the Iberian Peninsula, and Linderson (2001) for southern Sweden. Other examples are shown by Hellström (2005) to distinguish extreme and non-extreme precipitation patterns in Sweden, or the TROTREP project (Goodess, 2001) for analysing tropospheric ozone. Some methodological limitations were introduced by Martín-Vide (2001), pointing out some deficiencies in this method, especially on identifying typical summer low-pressure patterns over the western Mediterranean.

There are other classification approximations, such as that developed by Maheras *et al.* (2000), or the recent fast increase on the use of neural networks (Cavazos, 2000; Hewitson and Crane, 2002) for downscaling applications.

But the methodology most applied and discussed for circulation pattern classification is based on the use of multivariate statistics, principal component analysis (PCA) and clustering techniques (CL). These methodologies are being used by different scientific subjects for data reduction, interpretation, and classification purposes. In climatology they have been applied to the classification of weather types (Huth *et al.*, 1993; Davis *et al.*, 1997; Bejarán and Camilloni, 2003), climate regionalization (White *et al.*, 1991; Serra *et al.*, 1996), atmospheric circulation reconstructions (Luterbacher *et al.*, 2002), circulation anomalies associated with natural climatic variability as those produced by large volcanic eruptions (Prohom *et al.*, 2003), climate series reconstruction (Mann *et al.*, 1998), and many others. Scientific discussion on these statistical methods and their climatic applications is extensive (Yarnal, 1993; Wilks, 1995; Von Storch and Zwiers, 1999; Barry and Carleton, 2001; Yarnal *et al.*, 2001; Zwiers and Von Storch, 2004), and relates to data normalization, spatial distribution and minimum sample size (Karl *et al.*, 1982; Von Storch and Hannoschöck, 1985; Lana and Mills, 1994; Mestas-Núñez, 2000), the significance of principal components (Overland and Preisendorfer, 1982; North *et al.*, 1982), the rotation of principal components and their climatic interpretation (Buell, 1975, 1979; Horel, 1981; Richman, 1986; Jolliffe, 1987; Richman, 1987; O'Lenic and Lizevey, 1988; Legates, 1991; Richman, 1993; Cheng *et al.*, 1995; Dommenget and Latif, 2001; Jolliffe *et al.*, 2002), or the comparison and evaluation of different hierarchical or non-hierarchical CL techniques (Kalkstein *et al.*, 1987; Gong and Richman, 1995; Mimmack *et al.*, 2001). This multivariate approximation to create circulation classifications has been evaluated and compared with other methods by different authors (Key and Crane, 1986; Bretherton *et al.*, 1992; Wallace *et al.*, 1992), emphasizing the in-depth analysis made by Huth (1996), comparing the Lund and Kirchhofer methods previously cited, and concluding that better results are reached with multivariate methods.

2. PCA AND CL FOR CLASSIFYING DAILY ATMOSPHERIC CIRCULATIONS

PCA and CL techniques have been widely used to study atmospheric circulation in different temporal and spatial scales, and are specially known in their application on the detection of principal modes of circulation, defining monthly, seasonal, or annual patterns such as NAO (North Atlantic Oscillation) or PNA

(Pacific/North American Pattern). Also, relationships between atmospheric circulation and surface variables such as precipitation (Tait and Fitzharris, 1998; Quadrelli *et al.*, 2001) and temperature (Maheras *et al.*, 1999) have been established.

Since the 1980s and 1990s, interest in the analysis of different environmental problems from the synoptic climatology's point of view, particularly related to extreme events, has widely increased the interest of researchers in looking for the establishment of relationships between atmospheric circulation and daily atmospheric data. Many methods have been used, and different variables investigated such as heavy snowfalls (Birkeland and Mock, 1996). One of the first papers using PCA and CL in this way corresponds to Richman (1981), relating circulation patterns to precipitation events. Richman showed satisfactory results from this technique for daily and hourly data classification, and began an in-depth methodological discussion on aspects such as the mode of decomposition, the type of dispersion matrix, or the type of rotation to use in these kinds of classifications. Bogardi *et al.* (1993) clustered, via a non-hierarchical K-means method, the standardized 500 hPa geopotential height data over USA for describing precipitation over eastern Nebraska. The same CL method was used by Kidson (1994) for classifying 9 years of daily analyses (at 00:00 UTC) of 1000 hPa geopotential height over the New Zealand region. Kidson, on the basis of S-mode and covariance dispersion matrix in PCA, used Ward's hierarchical CL to determine the number of groups and their centroids for the K-means algorithm. Over south-western USA, Galambosi *et al.* (1996) applied these multivariate methodologies to daily 500 hPa pressure fields at 12:00 UTC (diamond structured grids), coupling PCA (S-mode of decomposition and correlation matrix) and CL (K-means) in order to describe and generate local hydrological variables. Galambosi *et al.* (1996), as did Bogardi *et al.* (1993), selected the number of groups for the K-means method in a subjective way, denoting this as a typical subjective step in the automated classification procedure. In this way, and to analyse precipitation variability in Portugal in relation to large-scale circulation, Corte-Real *et al.* (1998) applied an information measure based on the precipitation series for determining the optimal number of groups for K-means, eliminating subjectivity in this step of the classification. With T-mode and correlation matrix in PCA, and K-means CL previously oriented by Ward's hierarchical method, Romero *et al.* (1999a) identified 19 atmospheric circulation patterns at two levels (925 and 500 hPa) related to significant rainfall days for Mediterranean Spain. The patterns obtained have been related to the precipitation spatial patterns obtained by the same authors via PCA and CL techniques for the area treated (Romero *et al.*, 1999b), and recently changes have been estimated in these precipitation patterns for the late twenty-first century using the GCM ECHAM-OPYC3 (Sumner *et al.*, 2003). Applying it to wind erosion events in sandy soil agricultural areas of southern Sweden, Ekström *et al.* (2002) classified daily SLP circulation data and identified (via the definition of an Index of Erosion Severity) the patterns associated with large erosion damage. Ekström *et al.* used probably the most extended steps for the PCA and CL classifications: S-mode, correlation matrix, Varimax orthogonal rotation, and K-means CL, recently reproduced by Ekström *et al.* (2004) for Australian dust storm analysis. Recently, in order to identify heavy snowfall day SLP patterns in the central Pyrenees, Esteban *et al.* (2005) considered a spatial standardization of the previous data for S-mode PCA, and suggested an approximation for selecting the number of clusters and their centroids for use in the K-means method without iterations.

Many other works have been proposed, showing different approximations on daily circulation pattern classification via PCA and CL techniques (Plaut and Simonnet, 2001; Plaut *et al.*, 2001; Skinner *et al.*, 2002; Solman and Menéndez, 2003; Rasilla, 2003; Müller *et al.*, 2003). This extensive bibliography shows the robustness of the method for daily circulation pattern classifications, but does not yet clarify what the best options to be chosen are, leaving this decision to the researcher. However, many publications highlight the association with the results obtained and their dependence on the selected steps. This is the reason why more research and classifications are needed, not only for analysing environmental problems but also for improving classifications on the short temporal scale. In this way, this paper uses the methodology proposed by Esteban *et al.* (2005), with good results for extreme events analysis (heavy snowfalls), and applies it to the classification of a long temporal period on a synoptic scale and for the full year simultaneously.

3. DATA AND METHODOLOGY FOR THE DAILY CATALOGUE

Daily SLP data at 18:00 GMT were selected from the NCEP/NCAR Reanalysis data (Kalnay *et al.*, 1996), encompassing the region from 30°N–60°N latitude by 30°W–15°E longitude, with a 2.5° spatial resolution (Figure 1) and for the 42-year period 1960–2001.

As Esteban *et al.* (2005) proposed for classifying heavy snowfall days, a PCA was conducted on the SLP data to derive the dominant patterns of variability from the raw fields (Richman, 1986; Yarnal, 1993). Later, K-means CL was used to classify the observations and to composite the circulation pattern maps (Yarnal, 1993). The details of the methodology are explained below:

3.1. PCA

As the number of variables (days) would largely exceed the number of observations (grid points) and also owing to the high computation requirements of T-mode (Huth, 2000), an S-mode data matrix is used (Richman, 1986), the variables being the grid points and the days being the observations. Data is deseasonalized via spatial standardization (scaling the rows), which keeps the fields' gradient in the analysis and maintains the daily temporal scale of the original data. A correlation matrix provides the most efficient representation of variance in the data set (Barry and Carleton, 2001), and patterns of spatial correlation can be detected without possible domination by the grid points with the largest variances (Jolliffe, 1986). Scree test (Cattell, 1966) (Figure 2), the explained variance criteria, the Kaiser method, and the North rule of thumb (North *et al.*, 1982) (Figure 2) were used to retain the components explaining a significant portion of the total variance (Kalkstein *et al.*, 1987; Barry and Carleton, 2001), being later rotated using a Varimax procedure. The rotation of PCs facilitates the spatial interpretation of the principal components (Yarnal, 1993) and avoids problems like the domain shape dependence, but retains the orthogonality constraint of the eigenvector model (Richman, 1986). Clear differences between the non-rotated and rotated spatial structures and the easy identification of the artificial configurations predicted by Buell (1975) in the non-rotated results (Figure 3) show that rotation of the PCs is acceptable in our analysis.

3.2. Clustering

The non-hierarchical K-means method is used to cluster the observations (Hair *et al.*, 1998) via the component scores obtained from PCA. To decide on the number of groups and the centroids needed for the classification, usually a decision not easy to justify, we considered the spatial variation patterns (modes) established by the PCA, i.e. the principal components retained and rotated in their positive and negative phases, as potential groups for circulation patterns classification. The centroids are calculated by averaging

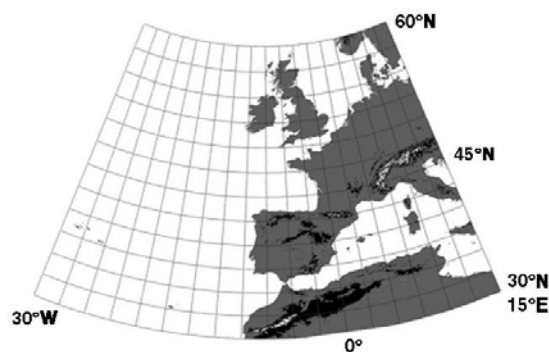


Figure 1. The synoptic area considered for this study. A grid of 2.5° resolution, extracted from the NCEP/NCAR Reanalysis project, ensures a good resolution of atmospheric data with 257 points

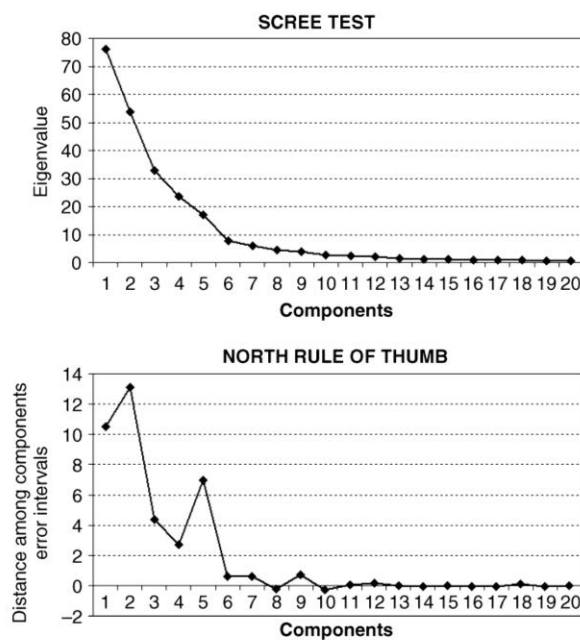


Figure 2. Scree test and North rule of thumb, two empirical techniques, used in the selection of principal components to be retained for rotation

the days that fulfil the principle of the 'extreme scores' for a certain pattern and phase: observations with high score values for a certain component (values higher than +2 for the positive phase, or lower than -2 for the negative phase) but with low score values for the remainder (between +2 and -2) were selected (Tait and Fitzharris, 1998; Birkeland *et al.*, 2001; Kostopoulou, 2003). Using this technique, if there is no real case that could be assigned to the PC that we are checking (i.e. there is no day with a close spatial structure), we will consider this component and phase as an artificial result on the PCA analysis, thus eliminating this potential group (circulation pattern). Briefly, the extreme scores procedure establishes the number of groups and their centroids for the K-means method, but also tries to be a filter to avoid artefacts in the final classification if the Varimax rotation previously used in the PCA process has not done it as well as expected. Eventually, the K-means produces the final classification of all the observations (days) with a distribution similar to the SLP fields. In this non-hierarchical classification process iterations are avoided, as the centroids were well established by the first PCA grouping (reflecting a circulation pattern) and because of the trend to equal the sizing of the groups of the iterations on the K-means CL procedure (Huth, 1996).

4. RESULTS

On the basis of the results of the different selection rule procedures (Figure 2), ten principal components were retained, which together accounted for 92.5% of the total variance. Finally, all days have been classified into 20 groups (all the potentials obtained in the PCA procedure), which represent the most relevant circulation patterns over the study area. Sea level pressure and 500 hPa maps, their annual distribution, and their evolution through the period studied (1960–2001) are shown in Figure 4(a), (b) and (c) for the 20 clusters identified. Table I shows some descriptions of every pattern. Next, a brief commentary on each pattern (all distributed into five general groups) in terms of morphology and seasonality, and later their recent trends, are presented. The clusters are numbered following the order obtained on the K-means output (Figure 4(a), (b), (c)).



Figure 3. Spatial structure of the non-rotated (PC) and rotated (RPC) principal components retained via different selection rules

4.1. Morphology and annual distribution

4.1.1. Westerly patterns (36.71%). These are listed below:

- Winter Westerly (CL1): Strong west flow near surface as well as in 500 hPa. Low pressure to the north-west of the British Isles. Winter pattern, practically non-existent in summer.
- Summer Westerly in Atlantic façade (CL3): Moderate west flow in the Atlantic façade with low-pressure gradient over the Mediterranean basin near surface. Summer pattern practically non-existent from November to February.
- Mediterranean high (CL5): Anticyclone in the Mediterranean Sea and Iceland low producing south-west flow over most of Western Europe. Very low frequency in summer.
- West–north-west advection (CL7): Azores anticyclone and a low over Scandinavia producing a strong west–north-west flow near surface as well as in 500 hPa over most Europe. Quite well distributed throughout the year.
- North-west advection in Atlantic façade (CL10): Quite similar to pattern 7. Azores anticyclone and a low to the north of the British Isles producing a strong north-west flow over Atlantic Europe façade near surface. Low-pressure gradient in central and south Europe. Minimum frequency in winter.

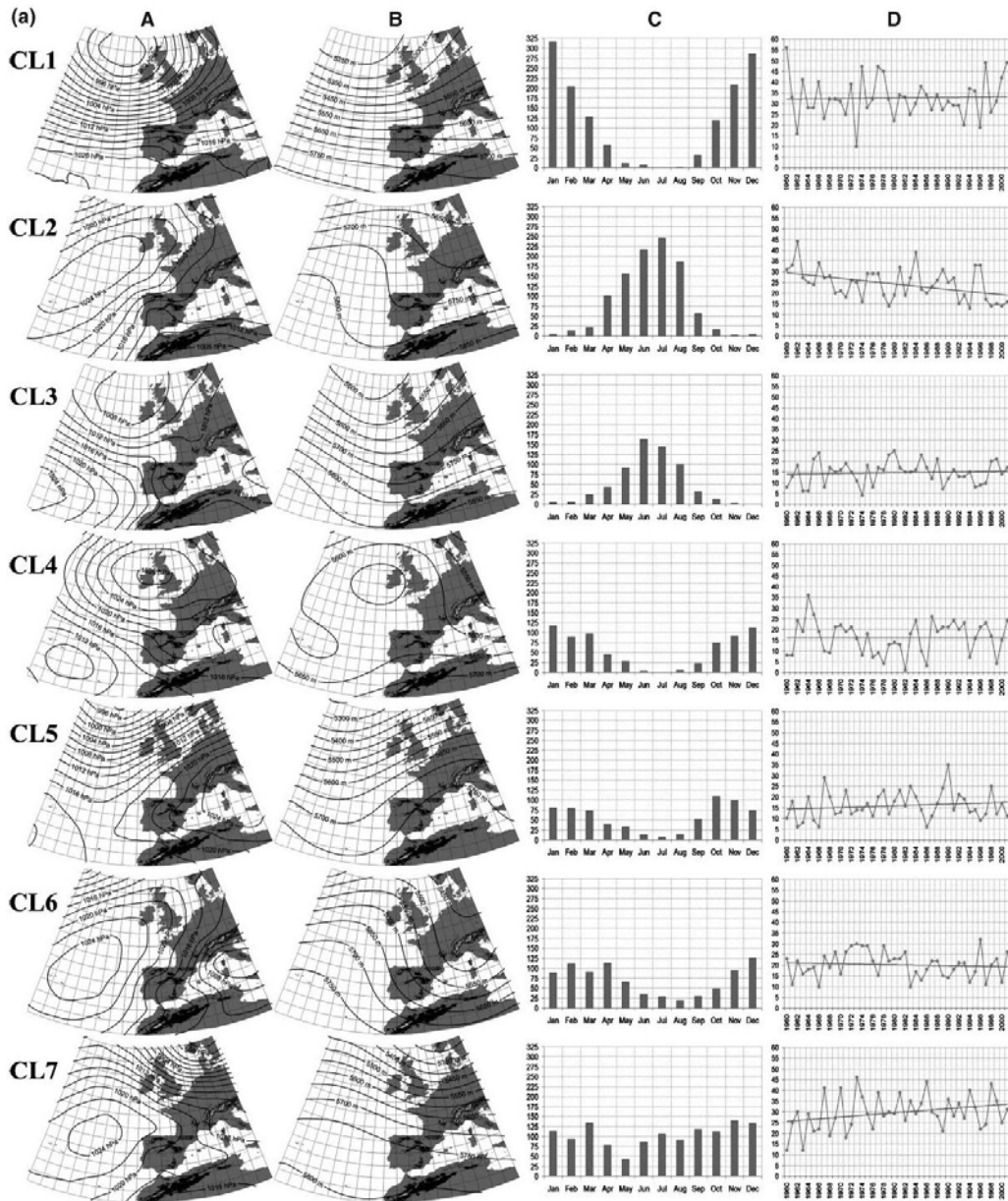


Figure 4. (a) Synoptic circulation patterns based on multivariate techniques (1960–2001) for western Europe: Clusters 1 to 7. (A) Sea level pressure. (B) 500 hPa geopotential. (C) Annual distribution of number of days belonging to a selected cluster. (D) 42-year variability of the number of days belonging to a selected group/cluster. (b) Synoptic circulation patterns based on multivariate techniques (1960–2001) for western Europe. Clusters 8 to 14. (A) Sea level pressure. (B) 500 hPa geopotential. (C) Annual distribution of number of days belonging to a selected cluster. (D) 42-year variability of the number of days belonging to a selected group/cluster. (c) Synoptic circulation patterns based on multivariate techniques (1960–2001) for western Europe. Clusters 15 to 20. (A) Sea level pressure. (B) 500 hPa geopotential. (C) Annual distribution of number of days belonging to a selected cluster. (D) 42-year variability of the number of days belonging to a selected group/cluster

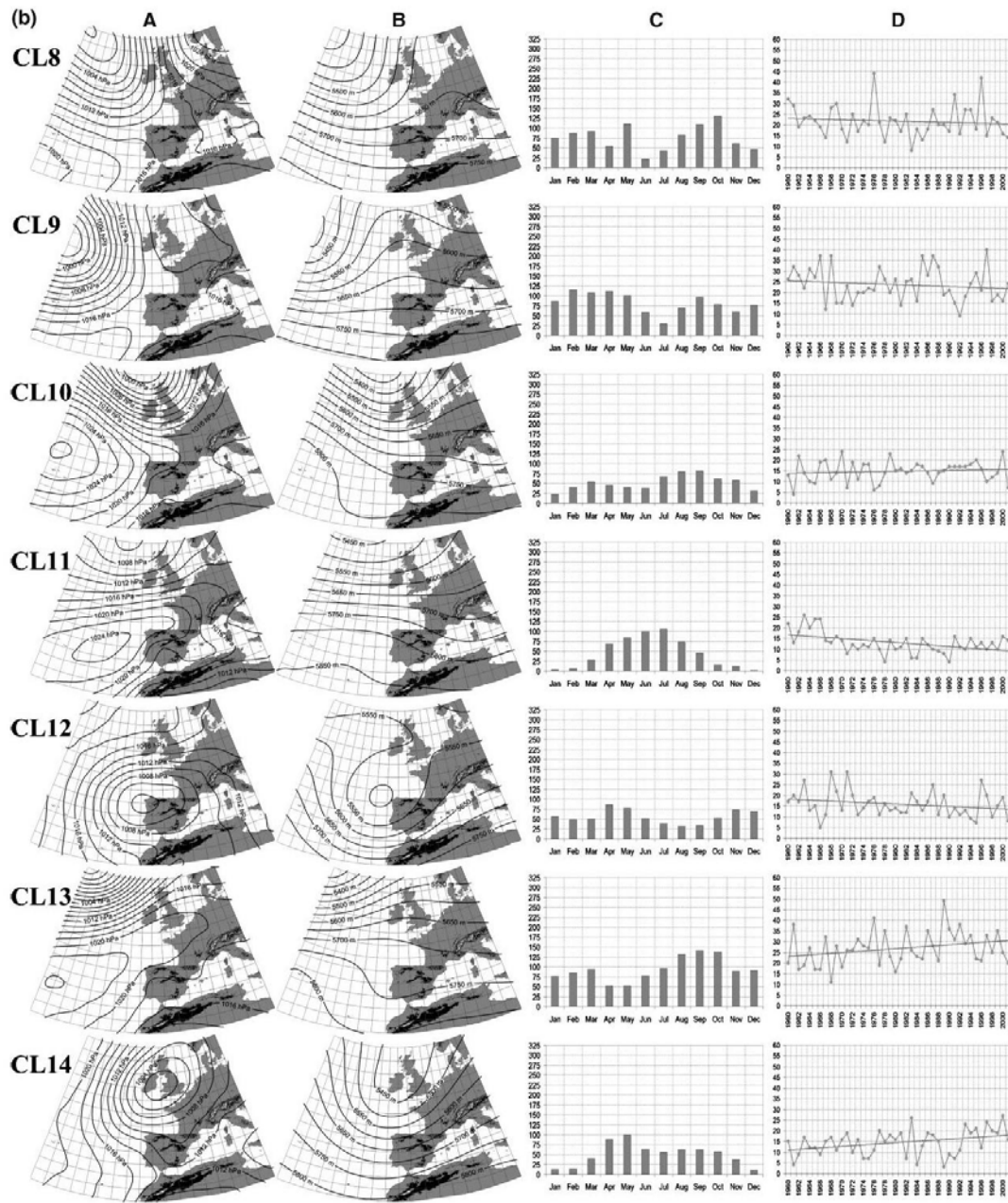


Figure 4. (Continued)

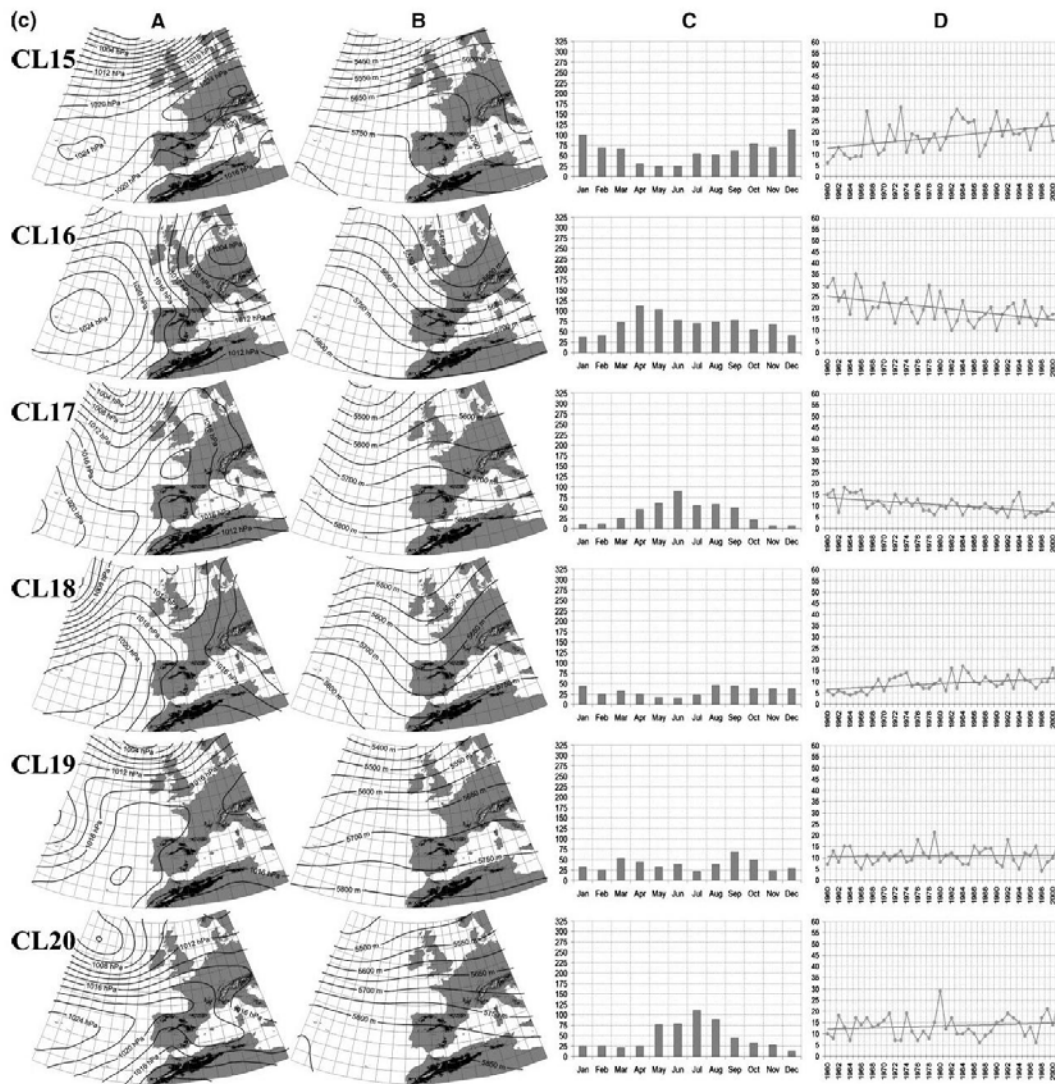


Figure 4. (Continued)

- Azores anticyclone I (CL11): Azores anticyclone extended by a ridge to Central Europe. West flow in north Europe and east winds in south Iberia near surface. Summer pattern practically non-existent in winter.
- Azores anticyclone II (CL20): Very similar to pattern 11. Azores anticyclone extended by a ridge to central Europe. Strong west flow in mid-latitudes over the Atlantic Ocean in 500 hPa. East winds in south Iberia near surface. Summer pattern.

4.1.2. Meridional patterns (20.10%). These are listed below:

- North Atlantic anticyclone (CL2): North Atlantic anticyclone affecting the British Isles and northeast flow over most of Western Europe. Summer pattern practically non-existent from November to January.

Table I. Description of the circulation patterns obtained: cluster number, total number of days, frequency in percentage, month with maximum number of days, month with minimum, year with maximum, year with minimum

Pattern	Number of days	Frequency (%)	Month max	Month min	Year max	Year min
Cluster 1	1367	8.91	January	July	1960	1973
Cluster 2	1018	6.64	July	November	1962	1994
Cluster 3	619	4.03	June	December	1980	1974
Cluster 4	682	4.45	January	July	1964	1982
Cluster 5	673	4.39	October	July	1990	1962/66/86
Cluster 6	851	5.55	December	August	1996	1966/83
Cluster 7	1244	8.11	November	May	1973	1960/63
Cluster 8	912	5.94	October	June	1976	1983
Cluster 9	992	6.47	February	July	1997	1992
Cluster 10	620	4.04	September	January	2000	1961
Cluster 11	545	3.55	July	December	1963	1978/90
Cluster 12	668	4.35	April	August	1971	1966
Cluster 13	1127	7.35	September	April/May	1989	1968
Cluster 14	608	3.96	May	December	2000	1989
Cluster 15	746	4.86	December	May	1973	1960
Cluster 16	828	5.4	April	January	1965	1982/90
Cluster 17	436	2.84	June	November/December	1963	1992/95
Cluster 18	387	2.52	August	June	1984	1961/64/67
Cluster 19	454	2.96	September	July	1979	1998
Cluster 20	564	3.68	July	December	1980	1987/97

- Mediterranean low/northerly advection (CL6): Azores anticyclone extending to the British Isles and a low over Italy producing north flow over western Europe and northeast flow in central Europe. Minimum frequency in summer.
- Central Europe low/northerly advection (CL16): Azores anticyclone and central Europe low near surface producing a north–northwest flow over western Europe. Weak circulation index. Maximum frequency in spring.
- Trough (CL18): Trough in 500 hPa with the axis extended from Scotland to Canary Islands. Azores anticyclone and low-pressure gradient in east Spain and west France. Minimum frequency in late spring and early summer.

4.1.3. Surface low gradient patterns (19.61%) These are listed below:

- North Atlantic low (CL9): Deep low over the north Atlantic Ocean and low-pressure gradient in Europe. Relatively equinoctial.
- Azores ridge (CL13): Azores anticyclone projecting a long, wide ridge to central Europe. Deep depression in high Atlantic latitudes. Low gradient of the isobars in the Mediterranean basin. Preferentially late summer and autumn pattern.
- Iberian thermal low (CL17): Iceland low with a trough over the Atlantic Ocean. Well-configured thermal depression in Iberia and relatively low-pressure gradient over the continent near surface. Moderate-weak west flow in 500 hPa. Summer pattern with very low frequency from November to February.
- Anticyclone bridge (CL19): Lax anticyclonic surface field joining Central Europe and an area to the south-west of Iberia. Strong west flow over the north of the British Isles. West weak flow in 500 hPa. Relatively equinoctial pattern.

4.1.4. Blocking patterns (15.25%). They are listed below:

- British high (CL4): Blocking pattern with the anticyclone over the British Isles. Ω -shaped configuration in 500 hPa. Winter pattern, practically non-existent in summer.

- Southerly flow (CL8): Atlantic west flow diffluent over Atlantic façade. Low to Iceland and anticyclone over Scandinavia. South component winds in most Europe near surface. Preferentially equinoctial.
- Central Europe high (CL15): Central European anticyclone connected with Azores high. West component flow in north Europe and east component flow near surface in south Europe. Maximum frequency in winter.

4.1.5. Cyclonic patterns (8.32%) They are listed below:

- North-west Iberian low (CL12): All-level tropospheric depression in northwest of the Iberian Peninsula. Southerly winds over the Alps, and east component flow in central and north Europe. Quite well distributed throughout the year.
- British low (CL14): All-level tropospheric depression centred over the British Isles. South component flow in Central Europe and Scandinavia. Very low frequency in winter.

4.2. Long-term variability

A trend analysis of the frequency of each pattern for the period 1960–2001 has been made by means of the *t*-test (using the AnClim software Štěpánek, 2003). Seven of the twenty patterns show a significant trend at the 95% confidence level (Table II). Most of those are patterns of little frequency. There is, however, a common pattern, the CL2, which is the second most frequent and that has been named ‘North Atlantic anticyclone’, with a negative significant trend (−2.585/decade). This is the summer pattern that produces northerly flow over most of western Europe (east in south Iberia). Its negative trend means that surface meridional flows have become less important in terms of frequency during the summer. The same result can be established from the negative significant trend of pattern 16, named ‘Central Europe Low-Northerly advection’, that also produces northerly flow over western Europe, with maximum frequency in spring and summer. Finally, five of the seven patterns with significant trends, negative or positive, show a seasonal maximum in summer or in spring; meanwhile the most frequent winter patterns, CL1, ‘Winter Westerly’, and CL4, ‘British High’, do not show any trend.

Table II. Results of the application of the *t*-test to the resulting clusters for identifying significant trends over the period 1960–2001. Seven patterns show significant trends at 95% confidence level. Their corresponding regression line and the trend expressed over 10 years is also presented

Pattern	Significance (95%)	Regression line	Trend 10 years
Cluster 1	No	–	–
Cluster 2	Yes	$y = 29.796 - 0.258x$	−2.585
Cluster 3	No	–	–
Cluster 4	No	–	–
Cluster 5	No	–	–
Cluster 6	No	–	–
Cluster 7	No	–	–
Cluster 8	No	–	–
Cluster 9	No	–	–
Cluster 10	No	–	–
Cluster 11	Yes	$y = 17.016 - 0.188x$	−1.879
Cluster 12	No	–	–
Cluster 13	No	–	–
Cluster 14	Yes	$y = 10.804 + 0.171x$	1.708
Cluster 15	Yes	$y = 12.400 + 0.249x$	2.494
Cluster 16	Yes	$y = 25.376 - 0.263x$	−2.633
Cluster 17	Yes	$y = 13.876 - 0.162x$	−1.625
Cluster 18	Yes	$y = 6.537 + 0.124x$	1.245
Cluster 19	No	–	–
Cluster 20	No	–	–

5. CONCLUSIONS

Multivariate statistics (PCA and CL) as a computer-assisted method have been used to obtain 20 circulation patterns that explain SLP and 500 hPa principal configurations over western Europe for the whole year and in a daily temporal resolution. The seasonality and trend of every cluster during the period studied is also briefly shown.

The method used offers a new approach to the eigenvector- and clustering-based classification of circulation patterns oriented to short-term phenomena analysis as heavy precipitation. In this way, trying to clarify some decisions not yet well established in this multivariate methodology has been one of the main objectives of this paper. First, a spatial pre-processing of the original data is done, allowing the introduction of the intensity of the fields to the analysis and, second, a new approach is shown, using the 'extreme scores' via PCA, to create the centroids and the number of groups for the K-means CL algorithm. Finally, thanks to the spatial variation patterns obtained by the PCA, the K-means without iteration avoids equidistribution into the groups that could distort the climatic meaning of the clusters obtained.

Otherwise, other steps in the PCA and clustering classification method are not well clarified. We would like to highlight the selection of the components to be retained in the PCA, a step that clearly influences the results of the 'extreme scores' proposal.

The resulting 20 atmospheric circulation patterns, also distributed into 5 general categories – westerly (36.71% of days), meridional (20.10%), surface low gradient (19.61%), blocking (15.25%), and cyclonic (8.32%) – show the most frequent distribution of the principal SLP configurations, reflecting the daily reality of the original data and establishing a detailed catalogue of surface and 500 hPa configurations for western Europe. Centres of action like the Azores anticyclone, depending on their extension and the location of their centre, produce different circulation patterns over the area studied, which the authors consider interesting to distinguish in a daily classification. On the other hand, patterns of low surface gradient, highly relevant in a climatic and meteorological sense for the southern sector of the area studied (especially over the Mediterranean basin), are well represented in the catalogue, without reducing the correct detection of the typical configurations at the northern side of western Europe. In this way, a visual analysis of the patterns over the British Isles shows a good coincidence with Lamb's patterns for this area. In this aspect, this methodology looks suitable for all climatic areas of the world. Nevertheless, a higher resolution of the grid would have allowed the detection of configurations of reduced spatial extension, such as those that affect the Mediterranean basin, but would have also forced us to consider a larger number of patterns for the same area. Considering the highlighted limitations of the method shown in this paper, such as the importance of the grid used or the number of principal components retained, we emphasize a more methodological discussion for the improvement of circulation pattern classification techniques for daily phenomena analysis.

Future work linking every pattern with the surface meteorological data observed and obtaining continuous climatic cartography (via geographic information systems) for a sector or the full area studied could help to clarify aspects like spatial behaviour or climate change impact in relation to daily phenomena, especially with the more frequent patterns that show a significant trend during recent decades, such as cluster CL2. Finally, and related to these trends obtained, a deeper analysis of the temporal and spatial changes of the patterns during the period analysed is also an interesting future work.

ACKNOWLEDGEMENTS

Our thanks are due to Marc Prohom (Catalan Meteorological Service), Phil Jones (Climatic Research Unit), Ian Jolliffe (University of Reading), and the CRECIT staff (Marta and Natalia) for their suggestions and support for the development of this work. We also thank the FI Grant of the Catalan Government and the Andorran authorities (Francesc Areny, Industry department) for financing our research project. We also wish to extend our thanks to the Spanish research project REN2001-2865-C02-01 and to the Catalan Government for supporting the Group of Climatology. Finally, we would like to thank the useful comments of Peter Bissolli (German Meteorological Service) and the two anonymous reviewers.

REFERENCES

- ACCORD. (Atmospheric Circulation Classification and Regional Downscaling). 2005. <http://www.cru.uea.ac.uk/cru/projects/accord/>.
 Action Cost 733 (Harmonisation and Applications of Weather Types Classifications for European Regions). 2005. http://cost.cordis.lu/src/action_detail.cfm?action=733.
- Barry RG, Carleton AM. 2001. *Synoptic and Dynamic Climatology*. Routledge: London.
- Barston AG, Lizevey RE. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly weather review* **115**: 1083–1126.
- Bejaran RA, Camilloni IA. 2003. Objective method for classifying air masses: an application to the analysis of Buenos Aires (Argentina) urban heat island intensity. *Theoretical and Applied Climatology* **74**: 93–103.
- Birkeland KW, Mock CJ. 1996. Atmospheric circulation patterns associated with heavy snowfall events, Bridger Bowl, Montana, U.S.A. *Mountain Research and Development* **16**: 281–286.
- Birkeland KW, Mock CJ, Shinker JJ. 2001. Avalanche extremes and atmospheric circulation patterns. *Annals of Glaciology* **32**: 135–140.
- Bischoff SA, Vargas WM. 2003. The 500 and 1000 hPa weather types circulations and their relationship with some extreme climatic conditions over southern South America. *International Journal of Climatology* **23**: 541–556.
- Bogardi I, Matyasovszky I, Bardossy A, Duckstein L. 1993. Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns. *Journal of Geophysical Research* **98**: 16653–16667.
- Bretherton CS, Smith C, Wallace JM. 1992. An intercomparison of methods for finding coupled patterns in climate data. *Journal of Climate* **5**: 541–560.
- Buell CE. 1975. The topography of the empirical orthogonal functions. *Preprints Fourth Conference on Probability and Statistics in Atmospheric Sciences*. American Meteorological Society: Tallahassee, FL; 188–1193.
- Buell CE. 1979. On the physical interpretation of empirical orthogonal functions. *Preprints Sixth Conference on Probability and Statistics in Atmospheric Sciences*. American Meteorological Society: Banff; 112–117.
- Cattell RB. 1966. The scree test for the number of factors. *Multivariate Behavioral Research* **1**: 245–276.
- Cavazos T. 2000. Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans. *Journal of Climate* **13**: 1718–1732.
- Cheng X, Nitsche G, Wallace JM. 1995. Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analyses. *Journal of Climate* **8**: 1709–1713.
- Conway D, Wilby RL, Jones PD. 1996. Precipitation and air flow indices over the British Isles. *Climate Research* **7**: 169–183.
- Corte-Real J, Qian B, Xu H. 1998. Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation. *International Journal of Climatology* **18**: 619–635.
- Davis RE, Stanmeyer TM, Jones GV. 1997. A synoptic climatology of tornadoes in Virginia. *Physical Geography* **18**: 383–407.
- Dommenget D, Latif M. 2001. A cautionary note on the interpretation of EOFs. *Journal of Climate* **15**: 216–225.
- Ekström M, Jönsson P, Barring L. 2002. Synoptic pressure patterns associated with major wind erosion events in southern Sweden (1973–1991). *Climate Research* **23**: 51–66.
- Ekström M, McTainsh GH, Chappell A. 2004. Australian dust storms: temporal trends and relationships with synoptic pressure distributions (1960–90). *International Journal of Climatology* **24**: 1581–1599.
- El-Kadi AKA, Smithson PA. 2000. Probability assessment of pressure patterns over the British Isles. *International Journal of Climatology* **20**: 1351–1369.
- Esteban P, Jones PD, Martín-Vide J, Mases M. 2005. Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *International Journal of Climatology* **25**: 319–329.
- Galambosi A, Duckstein L, Bogardi I. 1996. Evaluation and analysis of daily atmospheric circulation patterns of the 500 hPa pressure field over the southwestern USA. *Atmospheric Research* **40**: 49–76.
- Gong X, Richman MB. 1995. On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *Journal of Climate* **8**: 897–931.
- Goodess CM, Palutikof JP. 1998. Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology* **10**: 1051–1083.
- Goodess CM. 2001. European circulation indices for use in tropospheric ozone studies. *Tropospheric Ozone and its Precursors (TROTREP) Project*. Final report to University of Berne: Norwich, UK.
- Hair JF, Anderson RE, Tatham RL, Black WC. 1998. *Multivariate Data Analysis*, 5th edn. Prentice-Hall International: Englewood Cliffs, New Jersey, USA.
- Hellström C. 2005. Atmospheric conditions during extreme and non-extreme precipitation events in Sweden. *International Journal of Climatology* **25**: 631–648.
- Hess P, Brezowsky H. 1952. Katalog der grosswetterlagen Europas. *Berichte Deutsche Wetterdienstes* **15**(113): 39.
- Hewitson BC, Crane RG. 2002. Self-organizing maps: applications to synoptic climatology. *Climate Research* **22**: 13–26.
- Horel JD. 1981. A rotated principal component analysis of the interannual variability of the northern hemisphere 500 mb height field. *Monthly Weather Review* **109**: 2080–2092.
- Huth R, Nemesová I, Klimperová N. 1993. Weather categorization based on the average linkage clustering technique: an application to European mid-latitudes. *International Journal of Climatology* **13**: 817–835.
- Huth R. 1996. An intercomparison of computer-assisted circulation classification methods. *International Journal of Climatology* **16**: 893–922.
- Huth R. 2000. A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology* **67**: 1–18.
- Jenkinson AF, Collison FP. 1977. An initial climatology of gales over North Sea. *Synoptic Climatology Branch Memorandum n°62*. Meteorological Office: Bracknell.
- Jolliffe IT, Uddin M, Vines SK. 2002. Simplified EOFs three alternatives to rotation. *Climate Research* **20**: 271–279.
- Jolliffe IT. 1986. *Principal Component Analysis*. Springer-Verlag: New York.
- Jolliffe IT. 1987. Rotation of principal components: some comments. *International Journal of Climatology* **7**: 507–510.
- Jones PD, Hulme M, Briffa KR. 1993. A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology* **13**: 655–663.

- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetna A, Reynolds R, Jeene R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Kalkstein SL, Tan G, Skindlov JA. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Climate and Applied Meteorology* **25**: 717–730.
- Karl TR, Koscielny AJ, Diaz HF. 1982. Potential errors in the application of principal components (eigenvector) analysis to geophysical data. *Journal of Applied Meteorology* **21**: 1183–1186.
- Key J, Crane RG. 1986. A comparison of synoptic classification schemes based on 'objective' procedures. *Journal of Climatology* **6**: 375–388.
- Kidson JW. 1994. An automated procedure for the identification of synoptic types applied to the New Zealand region. *International Journal of Climatology* **14**: 711–721.
- Kilsby CG, Cowpertwait PSP, O'Connell E, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology* **18**: 523–539.
- Kirchhofer W. 1973. Classification of European 500 mb patterns. *Arbeitsbericht der Schweizerischen Meteorologischen Zentralanstalt* **43**: 205–207.
- Kostopoulou E. 2003. The relationships between atmospheric circulation patterns and surface climatic elements in the eastern Mediterranean. PhD thesis, University of East Anglia, Norwich.
- Lamb HH. 1972. *British Isles Weather Types and a Register of Daily Sequence of Circulation Patterns, 1861–1971*, *Geophysical Memoir 116*. HMSO: London.
- Lana X, Mills GF. 1994. Minimum sample size for synoptic weather type classification. Application to winter period data recorded on the catalan coast (north-east Spain). *International Journal of Climatology* **14**: 1051–1060.
- Legates DR. 1991. The effect of domain shape on principal component analyses. *International Journal of Climatology* **11**: 135–146.
- Linderson M. 2001. Objective classification of atmospheric circulation over southern Scandinavia. *International Journal of Climatology* **21**: 155–169.
- Lund IA. 1963. Map-pattern classification by statistical methods. *Journal of Applied Meteorology* **2**: 56–65.
- Luterbacher J, Xoplaki E, Dietrich D, Rickli R, Jacobeit J, Beck C, Gyalistras D, Schmutzet C, Wanner H. 2002. Reconstruction of sea level pressure fields over the eastern North Atlantic and Europe back to 1500. *Climate Dynamics* **18**: 545–561.
- Maheras P, Patrikas I, Karacostas Th, Anagnostopoulou Chr. 2000. Automatic classification of circulation types in Greece: methodology, description, frequency, variability and trend analysis. *Theoretical and Applied Climatology* **67**: 205–223.
- Maheras P, Xoplaki E, Davies T, Martin-Vide J, Barriandos M, Alcoforado MJ. 1999. Warm and cold monthly anomalies across the Mediterranean basin and their relationship with circulation; 1860–1990. *International Journal of Climatology* **19**: 1697–1715.
- Mann EM, Bradley RS, Hughes MK. 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**: 779–787.
- Martin-Vide J. 2001. Limitations of an objective weather-typing system for the Iberian peninsula. *Weather* **56**: 248–250.
- Mestas-Núñez AM. 2000. Orthogonality properties of rotated empirical modes. *International Journal of Climatology* **20**: 1509–1516.
- Mimmack GM, Mason SJ, Galpin JS. 2001. Choice of distance matrices in cluster analysis: defining regions. *Journal of Climate* **14**: 2790–2797.
- Müller GV, Compagnucci R, Núñez MN, Salles A. 2003. Surface circulation associated with frost in the wet Pampas. *International Journal of Climatology* **23**: 943–961.
- North GR, Bell TL, Cahalan RF, Moeng FJ. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review* **110**: 699–706.
- O'Lenic EA, Lizevey RE. 1988. Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields. *Monthly Weather Review* **116**: 1682–1689.
- Overland JE, Preisendorfer RW. 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**: 1–4.
- Plaut G, Simonnet E. 2001. Large-scale circulation classification, weather regimes, and local climate over France, the Alps and western Europe. *Climate Research* **17**: 303–324.
- Plaut G, Schuepbach E, Doctor M. 2001. Heavy precipitation events over a few Alpine sub-regions and the links with large-scale circulation, 1971–1995. *Climate Research* **17**: 285–302.
- Prohom MJ, Esteban P, Jones PD, Martín-Vide J. 2003. Surface atmospheric circulation over Europe following major tropical volcanic eruptions, 1780–1995. In *Volcanism and the Earth's Atmosphere, Geophysical Monograph Series 139*, Robock A, Oppenheimer C (eds); 273–281.
- Quadrelli R, Lazzzeri M, Cacciamani C, Tibaldi S. 2001. Observed winter Alpine precipitation variability and links with large-scale atmospheric patterns. *Climate Research* **17**: 275–284.
- Rasilla DF. 2003. Aplicación de un método de clasificación sinóptica a la Península Ibérica. *Investigaciones Geográficas* **30**: 27–45.
- Richman MB. 1981. Obliquely rotated principal components: an improved meteorological map typing technique? *Journal of Applied Meteorology* **20**: 1145–1159.
- Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.
- Richman MB. 1987. Rotation of principal components: a reply. *Journal of Climatology* **7**: 511–520.
- Richman MB. 1993. Comments on: 'The effect of domain shape on principal components analyses'. *International Journal of Climatology* **13**: 203–218.
- Romero R, Sumner G, Ramis C, Genovés A. 1999a. A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *International Journal of Climatology* **19**: 765–785.
- Romero R, Ramis C, Guijarro JA. 1999b. Daily rainfall patterns in the Spanish Mediterranean area: an objective classification. *International Journal of Climatology* **19**: 95–112.
- Serra C, Fernandez Mills G, Peniágo MC, Lana M. 1996. Winter and autumn daily precipitation patterns in Catalonia, Spain. *Theoretical and Applied Climatology* **54**: 175–186.

- Skinner WR, Flannigan MD, Stocks BJ, Martell DL, Wotton BM, Todd JB, Mason JA, Logan KA, Bosch EM. 2002. A 500 hPa synoptic wildland fire climatology for large Canadian forest fires, 1959–1996. *Theoretical and Applied Climatology* **71**: 157–169.
- Solman SA, Menéndez CG. 2003. Weather regimes in the South American sector and neighbouring oceans during winter. *Climate Dynamics* **21**: 91–104.
- Spellman G. 2000. The application of an objective weather-typing system to the Iberian peninsula. *Weather* **55**: 375–385.
- STARDEX. (Statistical and Regional dynamical Downscaling of Extremes for European regions). 2005. <http://www.cru.uea.ac.uk/cru/projects/stardex/>.
- Štěpánek P. 2003. *AnClim-Software for Time Series Analysis*. Department Of Geography, Faculty of Natural Sciences, Masaryk University: Brno.
- Sumner GN, Romero R, Homar V, Ramis C, Alonso S, Zorita E. 2003. An estimate of the effects of climate change on the rainfall of Mediterranean Spain by the late twenty first century. *Climate Dynamics* **20**: 789–805.
- Tait AB, Fitzharris BB. 1998. Relationship between New Zealand rainfall and south-west Pacific pressure patterns. *International Journal of Climatology* **18**: 407–424.
- Trigo RM, DaCamara CC. 2000. Circulation weather types and their influence on the precipitation regime in Portugal. *International Journal of Climatology* **20**: 1559–1581.
- Von Storch H, Hannoschöck G. 1985. Statistical aspects of estimated principal vectors (EOFs) based on small sample sizes. *Journal of Climate and Applied Meteorology* **24**: 716–724.
- Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University press: Cambridge, UK.
- Wallace JM, Smith C, Bretherton CS. 1992. Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies. *Journal of Climate* **5**: 561–576.
- White D, Richman M, Yarnal B. 1991. Climate regionalization and rotation of principal components. *International Journal of Climatology* **11**: 1–25.
- Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences: An Introduction, International Geophysics Series 59*. Academic Press: St Louis, Missouri, USA.
- Yarnal B, Comrie AC, Frakes B, Brown DP. 2001. Developments and prospects in synoptic climatology. *International Journal of Climatology* **21**: 1923–1950.
- Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press: London.
- Yarnal B. 1984. Synoptic scale atmospheric circulation over British Columbia in relation to the mass balance of Sentinel Glacier. *Annals of the Association of American Geographers* **74**: 375–392.
- Yarnal B, White DA. 1987. Subjectivity in a computer-assisted climatology I: classification results. *Journal of Climatology* **7**: 119–128.
- Yarnal B, White DA, Leathers DJ. 1988. Subjectivity in a computer-assisted climatology II: relationships to surface climate. *Journal of Climatology* **8**: 227–239.
- Zwiers FW, Von Storch H. 2004. On the role of statistics in climate research. *International Journal of Climatology* **24**: 665–680.

6.3. MODELITZACIÓ ESPACIAL DE LA TEMPERATURA DE L'AIRE I LA PRECIPITACIÓ A ANDORRA (PIRINEUS) A PARTIR DE TIPUS DE CIRCULACIÓ DIARIS. *THEORETICAL AND APPLIED CLIMATOLOGY*, 2009.

Basat en el catàleg de circulació atmosfèrica presentat a l'article anterior (apartat 6.2) per a Europa occidental, s'ha obtingut per a Andorra la cartografia d'alta resolució de temperatura mitjana diària de les màximes, temperatura diària mitjana de les mínimes, precipitació mitjana diària i probabilitat diària de precipitació. Aquests mapes associats a cada tipus de circulació s'han construït utilitzant l'altitud, la latitud, la continentalitat i la radiació solar com a predictors de la regressió múltiple. Les sèries de temperatura diària i precipitació utilitzades són provinents d'estacions meteorològiques andorranes, franceses i catalanes/espanyoles, i se'ls ha passat un procés de control de qualitat. Els resultats confirmen la complexitat de la distribució espacial dels fenòmens meteorològics als sectors de muntanya com Andorra, i mostren la importància de la influència mediterrània i atlàntica en el clima d'aquest país pirinenc. D'altra banda, diferents tests han mostrat que l'ús de les freqüències dels tipus de circulació millora els resultats de la interpolació.

Spatial modelling of air temperature and precipitation for Andorra (Pyrenees) from daily circulation patterns

P. Esteban · M. Ninyerola · M. Prohom

Received: 30 April 2007 / Accepted: 3 April 2008 / Published online: 31 May 2008
© Springer-Verlag 2008

Abstract Based on the daily sea level pressure (SLP) circulation catalogue obtained by Esteban, Martin-Vide and Mases, *Int J Climatol* 26:1501–1515, (2006) for Western Europe, high-resolution maps of daily maximum and minimum temperature, mean daily precipitation and daily precipitation probability have been obtained for Andorra (Pyrenees). The 20 daily-circulation patterns cover the period 1960–2001 and were generated using new approaches based on the rotated principal component analysis (PCA) and clustering technique. The final maps of Andorra associated with each circulation pattern have been constructed using altitude, latitude, continentality and solar radiation as multiple regression predictors (Ninyerola, Pons and Roure,

Int J Climatol 20:1823–1841, 2000). The daily temperature and rainfall series used from Andorran, French and Catalan/Spanish weather stations have been checked for data quality. The results confirm the complexity of the spatial distribution of meteorological phenomena over mountainous areas such as in Andorra, and show the importance of the Mediterranean and Atlantic influence upon the climate of this country of the Pyrenees. On the other hand, different tests have been made showing that the classification results could improve the resulting interpolated climate maps by the use of the circulation-pattern frequencies.

1 Introduction

Based on Barry and Perry (1973), Yarnal (1993) and Yarnal et al. (2001), Ustrnul (2006) has recently defined synoptic climatology as the subject that combines atmospheric circulation with climate and other elements of the environment, enabling the explanation and forecasting of many phenomena and processes that take place on the atmosphere, as well in the whole environment. In this respect, classifying the atmospheric circulation is becoming one of the most used techniques for these purposes, and that is why Esteban et al. (2006) highlight the interest of the atmospheric circulation catalogues as a good tool for explaining environmental and human phenomena using weather-map statistics. Here we would like to emphasise, as have many other researchers, the efforts made by Yarnal (1993) and Yarnal et al. (2001) for summarising and discussing the techniques, the developments and the prospects in synoptic climatology, including the role of geographic information systems (GIS).

One of the main objectives of using GIS in climatology is to construct maps for highlighting spatial properties of

P. Esteban (✉)
Snow and Mountain Research Centre,
Andorran Research Institute,
Av Rocafort 21–23, Edifici Molí, 3r pis,
AD600 St. Julià de Lòria, Principality of Andorra
e-mail: pesteban.cenma@iea.ad

P. Esteban · M. Prohom
Group of Climatology, University of Barcelona,
Montalegre 6,
Barcelona 08001, Catalonia, Spain

M. Prohom
e-mail: mprohom@meteocat.com

M. Ninyerola
Department of Animal Biology, Plant Biology and Ecology,
Autonomous University of Barcelona,
Cerdanyola del Vallès 08193, Catalonia, Spain
e-mail: miguel.ninyerola@uab.cat

M. Prohom
Area of Climatology, Meteorological Service of Catalonia,
C Berlin 28-46, 4rt,
Barcelona 08029, Catalonia, Spain

climate data that can be valuable for the analysis of physical and human phenomena. Examples in climate research are the spatial distribution of climate elements in relation to terrain features and different air masses or the usefulness of distinguishing between various weather types when constructing climate maps for avoiding possible erroneous interpretations related to the use of annual/seasonal/monthly averages (Bissolli and Müller-Westermeier 2005). On the other hand, these relationships are clearly very interesting for climate-change studies from the downscaling-techniques point of view. Tveito and Ustmul (2003) treated all of these aspects, pointing out that this approach should have great potential in describing spatial patterns of weather and climate elements, but this capability has not yet been explored properly. Advances in this area are the principles that this paper follows.

1.1 Atmospheric circulation classification techniques

Circulation-pattern classifications have been applied to different spatial and temporal scales and atmospheric variables. Recently, thanks to the improvement of the weather-station networks and the availability of gridded circulation data (NCEP/NCAR Reanalysis, ERA40 database, among others), many authors are studying the daily scale trying to make progress in the knowledge of extreme events such as heavy precipitations, local phenomena like air pollution, or other weather and climate impacts on society. As is well known, establishing statistical relationships between the synoptic and the local scale, clearly improve the interpretation and the conclusions that we can extract from global climate models (GCM) for the evaluation of climate change when trying to solve spatial resolution limitations of the dynamical models.

Different classification techniques are being used for obtaining atmospheric circulation catalogues. Some of these methodologies are called subjective methods, others are considered as being semi-automated or objective. Detailed information about these techniques and their applications can be found in Yarnal (1993), Barry and Carleton (2001) or in the results of the ACCORD (2000) and STARDEX (2005) research projects. In the last few years, many examples have appeared using different synoptic classification techniques. Philipp et al. (2005) classify reconstructed sea level pressure (SLP) dataset extending back to 1850 using a clustering based on the simulated annealing and multistart techniques. Esteban et al. (2006) identify heavy snowfall circulation patterns over Andorra with principal component analysis (PCA) and clustering techniques, proposing a new method for selecting the optimal groups. Hope et al. (2006) use self-organising maps to classify the winter circulation over southwest Western Australia. James (2007) proposes the automated and objec-

tive version of the Hess and Brezowsky Grosswetterlagen over Europe. Seibert et al. (2007) study heavy precipitation events in Austria with a trajectory clustering method. In this context, the main objectives of the Action COST 733 (2007) are to obtain extensive information about synoptic classification methods, methodological recommendations for this purpose, and evaluation of its results (Huth et al. 2005).

1.2 Spatial interpolation techniques in climatology

Spatial interpolation techniques have been successfully applied to regional mapping on monthly mean temperature and precipitation fields (Daly et al. 1994; Hutchinson 1995; Lennon and Turner 1995; Goodale et al. 1998; Ninyerola et al. 2000; Price et al. 2000; Brunsdon et al. 2001; Perry and Hollis 2005). However, some models and monitoring processes often require finer spatial resolution maps as input data. In this direction, some efforts have been made in order to capture the spatial pattern of daily meteorological data. In the case of minimum and maximum temperature mapping, several approaches exist (Chessa and Delitala 1997; Jarvis and Stuart 2001), even in complex mountainous areas (Dodson and Marks 1997; Bolstad et al. 1998; Hasenauer et al. 2003; Hunter and Meentemeyer 2005; Stahl et al. 2006) as is our case. On the other hand, daily precipitation has barely been mapped (Hasenauer et al. 2003; Hunter and Meentemeyer 2005). Monthly maps present lower average evaluation errors (from about 0.5 to 1°C) than daily maps (from about 1 to 2.5°C). Nevertheless, by using synoptic circulation patterns, we are able to improve these results and obtain daily maps with lower errors at the same time that they capture some spatial patterns which classic interpolation methods do not detect.

1.3 Spatialisation of climate elements by means of circulation pattern classifications

Few papers have appeared joining the principles and results previously mentioned. The first general conclusions in that direction could be found in the advances of Action COST 719 (2007) dealing with the use of GIS in climatology and meteorology (Tveito and Ustmul 2003). Tveito (2002) obtained preliminary results on the spatial distribution of winter temperatures in Norway for the Lamb circulation types, using residual kriging as the interpolation method, whose results were considered satisfactory by the author. Bissolli and Müller-Westermeier (2005) established monthly, spatial precipitation patterns for Germany using the objective weather-types classification method of the German Meteorological Service. This classification consists of 40 types based on wind direction in 700 hpa, cyclonality in 950 and 500 hpa, and the tropospheric humidity. The climatic precipitation surface series are interpolated based on the

inverse distance approximation, and as the authors conclude “the ensuing maps are not in contradiction to what is expected from meteorology”. Similar to the previous works is the approximation of Ustrnul (2006), interpolating daily temperatures in Poland also using the *Grosswetterlagen* of the German Meteorological Service. This author, like Tveito (2002), uses residual kriging, obtaining acceptable results except for the non-advective anticyclonic types during wintertime: the thermal inversions introduce relatively great variability on the results. Finally, Baker Perry and Konrad (2006) analyse the snowfalls on windward slopes related to northwestern flows over the southern Appalachians in the USA. Various topographic and geographic variables are considered and statistical relationships are established. The authors conclude that multivariate regression modelling implemented in GIS can be an effective way to map snowfall patterns associated with specific wind directions, particularly when strong topographic control is evident.

The study presented in this paper is developed over Andorra and the surrounding area, located in the central part of the Pyrenees massif, between France and Spain (Fig. 1). Andorra is a small (468 km²) and hilly country with a mean altitude of 2,044 m, ranging between 837 and 2942 m. (Fig. 2). It is included in the Mediterranean climate type, but continentality, altitude and the south-facing location in the Pyrenees range introduce important changes with regard to the temperature and precipitation regimes. In this way, the low and well-oriented topography of the French side of the Pyrenees also enables easy entrance of Atlantic air masses from the north and northwest, especially affecting areas of the high-elevation sites. This geographical complexity makes Andorra an interesting area for climate studies, and looks like a good area in relation to the objective of this paper, obtaining high-resolution daily climatic maps of precipitation and temperature on the basis of a circulation-pattern catalogue of Western Europe. The results also try to show the usefulness of this methodology for climate research and the improvement of the interpolated values in climatic maps if the circulation-patterns

classification results are also considered in the regression model used.

2 Data and methods

A three-step chain makes up the process followed for obtaining the results (see also the flow diagram of Fig. 3).

2.1 Atmospheric circulation catalogue

The classification used in this study is the one presented by Esteban et al. (2006), where 20 SLP types are obtained based on PCA and non-hierarchical clustering. The catalogue is obtained for the 30°N/30°W–15°E region and for the period 1960–2001, using the 2.5° gridded data of the NCEP/NCAR reanalysis database (Kalnay et al. 1996; Fig. 1). This method is also described in Esteban et al. (2005), but applied to extreme events (heavy snowfalls).

For the PCA procedure, the S-mode data matrix is used (Richman 1986), the variables being the grid points and the days being the observations. Previously, data is deseasonalised via spatial standardisation (scaling the rows) that keeps the fields’ gradient in the analysis and maintains the daily temporal scale of the original data. A correlation matrix provides the most efficient representation of variance in the dataset (Barry and Carleton 2001), and patterns of spatial correlation can be detected without possible domination by the gridpoints with the greatest variances (Jolliffe 1986). The Scree test (Cattell 1966), the explained variance criteria, the Kaiser method and the North rule-of-thumb (North et al. 1982) were used to retain the components explaining a significant portion of the total variance (Kalkstein et al. 1987; Barry and Carleton 2001), later being rotated using a Varimax procedure. The orthogonal rotation of principal components (PCs) facilitates the spatial interpretation of the PCs (Yarnal 1993), showing clear differences between the non-rotated and rotated spatial structures. The easy identification of the artificial config-

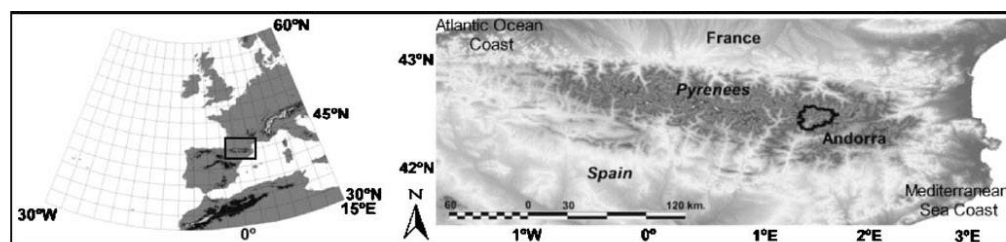


Fig. 1 *Left:* Synoptic area used in the circulation patterns classification. *Right:* Geographical location of Andorra in the Pyrenees range

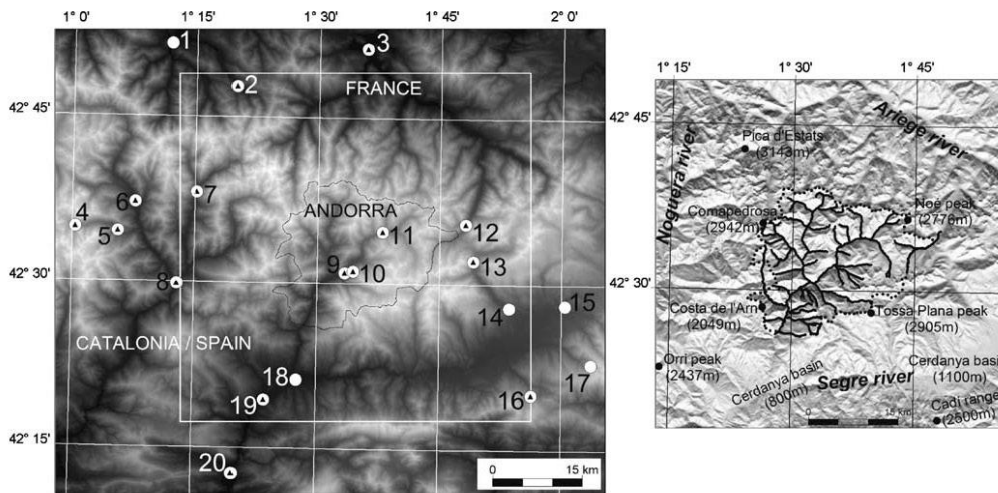


Fig. 2 *Left:* Digital elevation model of Andorra and its surrounding area. *White dots* stand for pluviometric stations and *black triangles* for thermo-pluviometric stations. The model has been run using the entire DEM but we only show the *inner squared area* in the final maps.

Right: Annual solar radiation map with shadowed aspects in *dark tones* and sunny aspects in *light tones*. *Black lines* represent the Andorran boundary and the main river streams. The main geographical elements are also drawn

urations predicted by Buell (1975) in the non-rotated results, showed that rotation of the PCs is recommendable in our analysis.

Initially, the iterative K-means method is the one considered to cluster the observations (Hair et al. 1998) via the component scores obtained from PCA. To decide on the number of groups and the centroids needed for the classification, usually a decision which is not easy to justify, the spatial variation patterns (modes) established by the PCA are considered. This means to understand the principal components retained and rotated in their positive and negative phases like potential groups for circulation patterns classification. The centroids are calculated by averaging the days that fulfil the principle of the “extreme scores” for a certain pattern and phase: observations with high-score values for a certain component (values higher than +2 for the positive phase, or lower than -2 for the negative phase), but with low-score values for the remainder (between +1 and -1) were selected (Tait and Fitzharris 1998; Birkeland et al. 2001; Kostopoulou 2003). Using this technique, if there is no real case that could be assigned to the PC that we are checking (i.e. there is no day with a close spatial structure), we will consider this component and phase as an artificial result in the PC analysis, thus eliminating this potential group (circulation pattern). Briefly, the extreme-scores procedure establishes the number of groups and their centroids for the K-means method, but also tries to be a filter to avoid artefacts in the

final classification if the Varimax rotation previously used in the PCA process has not done it as well as expected. Eventually, the K-means produces the final classification of

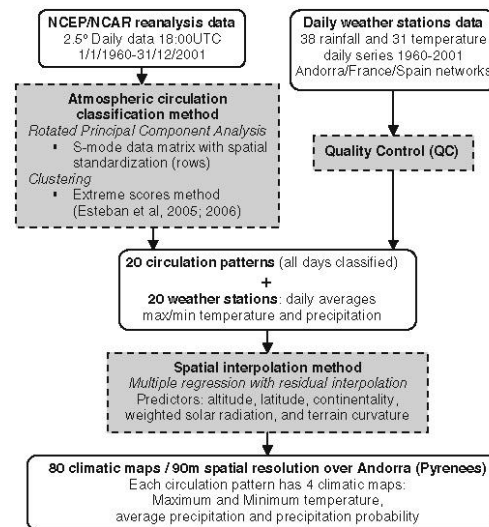


Fig. 3 Flow diagram showing the methodological sequence used to obtain the 80 climatic maps related to 20 circulation patterns

all of the observations (days) with a distribution similar to the SLP fields. After some experiments, the non-iterated result was chosen, considering that the centroids were well established by the first PCA grouping (reflecting real circulation patterns). On the other hand, some authors have pointed out that the iterations commonly tend to equal the size of the groups (Huth 1996), probably related with a strong underlying structure, as a continuum, of the data.

2.2 Check for data quality

Daily temperature and rainfall series around Andorra were collected from three different sources: the National Meteorological Institute (INM) for the Catalan series, Météo-France for the French data and FEDA (Forces Elèctriques d'Andorra) for the Andorran series. Initially, 38 daily total-rainfall series and 31 daily maximum and minimum surface air temperature series were selected for their geographical distribution and data availability, encompassing the period 1960–2001 (see Table 1). As a previous step to the quality control, some missing daily-precipitation values from the Catalan series were digitised as a result of the recovery of the original sources (pluviometric cards). As a result of this process, 843 daily rainfall data were recovered and were included in the initial series to be analysed.

In this section, we describe the quality control (QC) and assurance that is designed and applied to identify erroneous data resulting from sensors and observation sources. Some QC steps are common to both daily rainfall and temperature series, while some others are specific to only one variable.

Gross errors This method compares daily rainfall and temperature values from individual stations with established extreme values. Here, the extremes given by Gleason (2002) to compare the stations' daily values are used. Thus, for precipitation, the extremes are 1,828 and 0 mm, while for temperature they are 50 and -40°C. Data values greater than (less than) the highest (lowest) values defined are flagged as erroneous and are not used in the following QC steps.

Outliers detection In this step, rainfall and temperature amounts are treated separately. First, daily rainfall values are transformed into monthly totals. Then, an outlier is defined as those monthly rainfall totals exceeding the third quartile by more than three times the interquartile range (IQR), and at the same time exceeding the 5σ threshold of the mean. The same method is used for daily maximum and minimum temperature values, but considering monthly means in spite of monthly totals. Once detected, outliers

Table 1 Weather stations used in this study, temporal coverage and main results for the QC procedure

Weather station	Altitude (m)	Data coverage (extreme years)	Total P data	Ratio of erroneous P data	Total T data	Ratio of T erroneous data
1-SEIX	520	1960–2001	15,341	0.000	—	—
2-AULUS LES BAINS	733	1960–2001	15,341	0.002	30,682	0.000
3-TARASCON SUR ARIEGE	475	1960–2001	15,090	0.002	29,858	0.000
4-ESTANY DE SANT MAURICI	1,920	1965–2000	7,302	0.004	16,424	0.000
5-ESPOT	1,310	1965–1991	7,184	0.000	13,454	0.002
6-ESTERRI D'ÀNEU	940	1960–2001	12,570	0.005	22,904	0.003
7-TAVASCAN	1,100	1960–1994	12,290	0.007	19,272	0.008
8-LLAVORSÍ	850	1960–1999	13,826	0.000	25,432	0.005
9-ANDORRA - CENTRAL	1,150	1960–2001	15,341	0.000	32,868	0.008
10-ENGOLASTERS	1,640	1960–2001	15,341	0.000	32,868	0.004
11-RANSOL	1,640	1960–2001	15,341	0.000	32,868	0.005
12-L'HOSPITALES-PRÈS-LANDORRE	1,425	1961–2001	14,671	0.000	30,682	0.000
13-PORTE-PUYMORENS	1,620	1961–2001	6,662	0.005	12,418	0.000
14-TOR DE QUEROL	1,250	1967–1987	7,487	0.000	—	—
15-LLÍVIA	1,260	1960–1977	6,575	0.000	—	—
16-LA MOLINA	1,704	1960–1998	13,317	0.007	26,948	0.000
17-VALLCEBOLLERA	2,100	1967–1987	7,487	0.004	—	—
18-LA SEU D'URGELL	692	1965–1993	10,389	0.003	—	—
19-ADRALL	642	1960–1996	12,876	0.002	26,000	0.002
20-ORGANYÀ	540	1960–1999	14,573	0.000	19,511	0.005

Ratio of erroneous P or T data is expressed as: total number of erroneous data/total available data

are manually verified (or refused) by comparing the value with the five nearest series. Those values classified as erroneous are substituted for missing values (i.e. for temperature the daily value is replaced, while for precipitation the whole month is taken, as sometimes it is difficult to identify individual rainfall spells over this mountainous area).

'False' 0 The aim here is to distinguish those months registering null rainfall (common in the Mediterranean climate) from those without any observation. The method suggested by Abaurrea et al. (2004) is then followed. A "0" value is considered 'suspicious' if (1) at least 75% of the nearest series (six as optimum) have registered precipitation >15 mm and/or (2) the two nearest stations have exceeded 10 mm. When at least one of the rules is achieved, the daily value for the whole month affected is replaced by a missing value. For daily temperature data, when a whole month shows a 'flat line' of 0°C, the values are replaced by missing data.

Internal consistency check (daily temperature values) At this step, two rules are checked, quite similar to those proposed by Reek et al. (1992) and Feng et al. (2004): (1) internal inconsistency, which identifies errors such as the maximum temperature being cooler than the minimum temperature; (2) a 'flat-line' check, which identifies data of the same value for at least six consecutive days. We have chosen the six-day threshold because in most of the series the accuracy of the records is 1°C.

Temporal consistency check (daily temperature values) Here, day-to-day variations in maximum and minimum temperatures are analysed. Hence, a maximum step-change greater than or equal to 25°C is considered as being erroneous, while a step of 20°C is considered as being suspicious. As in previous steps, daily values outside this range are manually checked and edited on a case-by-case basis.

Date displacement (daily precipitation values) As a result of the different data sources used (from three different countries), observation methods or accuracy may have varied in time. For the aim of our research, the exact date where the rainfall data is recorded by the observer is crucial. To detect changes in the criterion of noting down the date where the observation was made, a relative index between the candidate and an auxiliary series is obtained (Abaurrea et al. 2004). First, the yearly sum of the daily (t) rainfall differences between the candidate series (y) and the auxiliary series (x) is obtained, and expressed as absolute values. Then, the same yearly sum of daily differences is calculated but displacing the candidate series one day

forward. Finally, the ratio between the two sums is obtained, as follows:

$$\frac{\sum |y_t - x_t|}{\sum |y_{t+1} - x_t|}$$

The closer the index is to 0, the greater the similarity is between both series. The same procedure should be applied, but displacing the candidate series one day backward:

$$\frac{\sum |y_t - x_t|}{\sum |y_{t-1} - x_t|}$$

To confirm the findings, the method is applied with at least another nearby series.

For the entire QC checks considered, the detected erroneous data were replaced by a missing value. As a result of the QC procedure and the data temporal coverage, 20 weather stations were finally selected, most of them containing daily temperature and rainfall data. Figure 2 shows the geographical distribution of the weather stations, while Table 1 indicates data coverage and total.

2.3 Spatial interpolation

Checked weather station data has been grouped together based on the 20 circulation patterns obtained by Esteban et al. (2006). Thus, a multiple regression with the residual interpolation method has been applied to obtain surfaces of mean daily precipitation (MDP), daily precipitation probability (DPP), mean daily minimum temperature (MDMn) and mean daily maximum temperature (MDMx).

Next the characteristics and methods applied to develop the predictor fields are explained: altitude, latitude, continentality (friction cost), weighted solar radiation, and terrain curvature. In order to include the stations surrounding Andorra, we have combined the 5-m DEM from this country with the 90-m DEM from the Shuttle Radar Topography Mission (USGS 2007). Before combining, we have resampled the 5–90-m DEM and also reprojected the DEMs in the UTM-31N horizontal reference system. The resulting mosaic map extent has been: MinX=353,970, MaxX=412,650, MinY=4,683,300 and MaxY=4,741,710. DEM has also been the basis to obtain other predictor variables, especially the continentality and the solar radiation surfaces.

Latitude and terrain curvature were computed using simple GIS tools: latitude is the linear distance to the raster lower row and terrain curvature is the second derivative from the altitude. Continentality has been computed as a function of cost distances from sea to each target raster cell using a cost-push algorithm (Eastman 1987–1997). This means building a friction surface (obtained from the DEM) and a source feature indicating from where we want to model the

air-mass motion. In this case, we have used a single source feature indicating the Mediterranean as well as the Atlantic coasts. The resulting friction map is obtained once the algorithm has assigned the lower cost value to each cell.

Finally, we have computed a DEM-based potential solar radiation model (Pons and Ninzerola 2008). This surface takes into account astronomical data (Earth-Sun position throughout the year), geographical data (topographical exposure and shadowing) and considers a uniform optical density with clear conditions (0.288). From monthly solar radiation maps, weighted solar radiation maps for each one of the 20 synoptic circulation patterns have been computed. These weights are based on the frequency in which a pattern appears in different months (see histogram in Figs. 4 and 6, column c). This weighted radiation surface is a refinement of a mean annual map.

Cross-validation For each synoptic circulation pattern we have performed n regression models each time using $n-1$

stations following a leave-one-out cross-validation technique to assess the quality of the ensuing maps through RMS error. After a test bed, the best model is the one that uses inverse distance weighting to interpolate the residuals using altitude, latitude, friction and weighted solar radiation as predictor variables. Also, in order to compare the performance of the multiple regression method, we have applied the same methodology, but interpolating the weather stations using a technique that does not take into account geographical information such as splines.

3 Results

Applying the methods previously explained, 80 climatic maps at 90-m spatial resolution were obtained, i.e. each of the 20 circulation patterns of the catalogue have detailed maps of MDP, DPP, MDMn and MDMx over the Andorran area. The ensuing high-resolution patterns can be physically interpreted in a reasonable way.

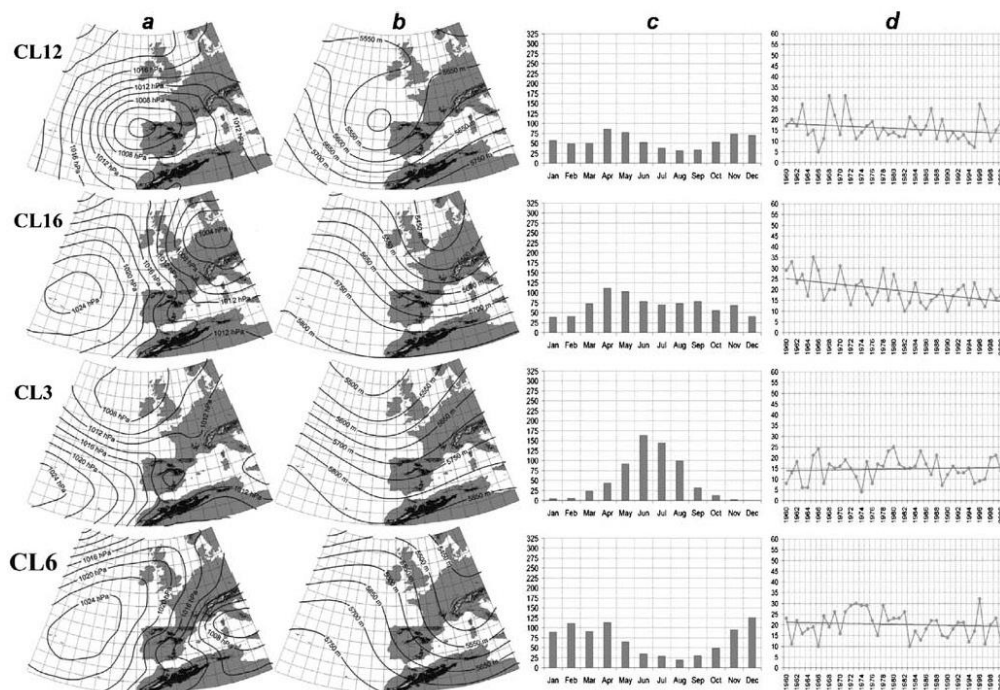


Fig. 4 Circulation patterns results for clusters (CL) 3, 6, 12 and 16. **a** Sea level pressure map, **b** 500 hPa geopotential height, **c** annual distribution of the days and **d** interannual variability

3.1 Subjective evaluation of the results

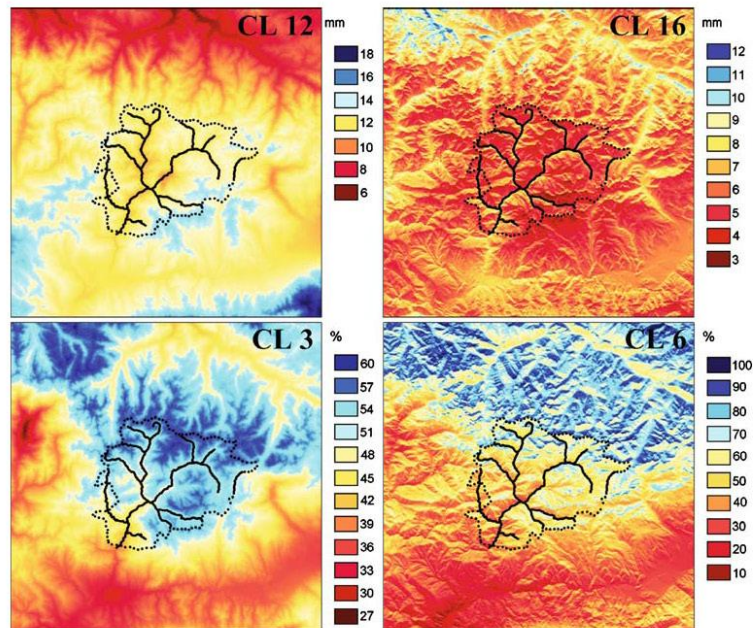
3.1.1 Mean daily precipitation (MDP) and daily precipitation probability (DPP)

The maps obtained for these variables show interesting spatial patterns related to the distribution of precipitation and temperature in relation to orography and the Mediterranean and Atlantic influences, or the variation of the amount and probability of precipitation with altitude. In this way, northerly/southerly advections clearly show strong differences between the northern/southern sides of the Pyrenees range, emphasising the role of these mountains as an effective orographic barrier to wet air masses. This can be observed in Fig. 4 (circulation patterns) and Fig. 5 (climatic maps) where the southerly flow (cluster 12: called north-west Iberian low by Esteban et al. 2006) and the northerly advection (CL16: central Europe low) over the area are shown. Thus, should be pointed out the latitudinal changes of the precipitation amount at the southerly pattern (CL12), which is very effective over the area analysed as is shown by the rainfall averages, and closely related to high levels of absolute humidity transported by the air mass. In addition, we can also observe the typical increase of the average values with altitude (see the maximum values over the Cadi range and the southern part of Andorra). On the contrary, extremely low precipitation averages (pluviomet-

ric shadow) can be observed in the low valleys to leeward of the range, in the French part. The same conclusion could be made for CL16, but the climatic maps are less visually evident. In this way, maximum values are now located on the French sector at the north face of the Pyrenees. The effect of the orography is less evident in this case, and precipitation shows a more homogeneous pattern despite the presence of a Föhn effect at low altitudes at the southern side (called Fogony in the area; Gómez 2001). These effects can also be seen on the basis of four weather stations; for example, changes in the distribution of the number of days of precipitation depending on the type of advection (Fig. 8c) southerly, Fig. 8d) northerly can be detected in addition to the effect of altitude. See also the relative important frequency of days with more than 50 mm for CL12 (Fig. 8c).

These strong gradients also appear in the DPP maps, where, in some advective patterns, the range of probabilities exceeds 80%. Such is the example of the maps shown in Figs. 4 and 5 and corresponding to CL 6 (Mediterranean low/northerly advection), where 80-100% values appear over the French area and over the mountain crests, while only 20% values can be found at low levels at the southern side (see the Cerdanya basin). The other pattern shown (CL3: Summer-westerly wind in the Atlantic façade) is also considered an advection from a general point of view, but in summertime, when the highest frequencies are shown,

Fig. 5 Precipitation maps obtained for clusters (CL) 3, 6, 12 and 16



the occurrence of a thermal low is common over the Iberian Peninsula. Because of that, convective activity is dominant over the Pyrenees, reflected by a clear extension of high precipitation probabilities over the southern and highest areas of Andorra, contrasting with a diminishing of the probabilities over the French valleys in comparison with CL6.

3.1.2 Mean daily minimum and maximum temperature (MDMn, MDMx)

The results in temperature also look promising. Hot and cold pools over the valleys, altitudinal gradients and differences between sunny and shaded aspects are found. The patterns shown for minimum temperature (Figs. 6 and 7) reflect a warm-season northerly advection (CL2: North Atlantic anticyclone) and an anticyclonic pattern (CL5: Mediterranean high). For the first case, the coldest average values of all of the patterns are detected over the Andorran area, the lowest ones being at the highest altitudes, as is expected. On the other hand, a Föhn effect could probably be

considered again: a 9.5°C value is obtained at 540m over the French valleys and at 740 m in the Cerdanya (Catalan side).

In relation to CL 5, the effect of thermal inversion is found and positive gradients with altitude can be observed. Despite that, the authors have detected some data gaps such as the presence of this effect in probably wrong places (in higher altitudes than expected, see near Ransol weather station). Similar problems also appear in the results of Ustmul (2006), and can be associated with not enough density of the weather stations or problems for the independent variables to model this behaviour (cold pools). Otherwise, it should be considered that we are working with an annual average; monthly or seasonal analysis will improve the representation of this process. This is supported by looking at the monthly anomaly graphs (Fig. 8f), showing the presence of these inversions during winter for the lowest weather stations, while high positive anomalies are obtained for the remaining ones (see Hospitalet weather station, at 1,425 m on the French side).

Maximum temperatures usually present fewer problems to be spatially modelled. The most interesting elements to

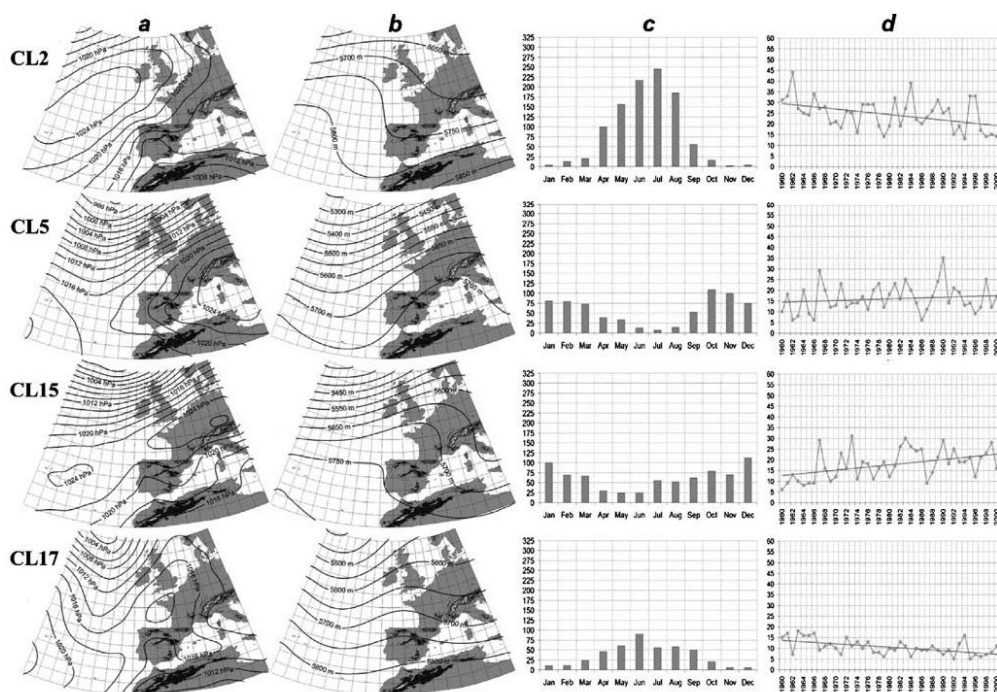
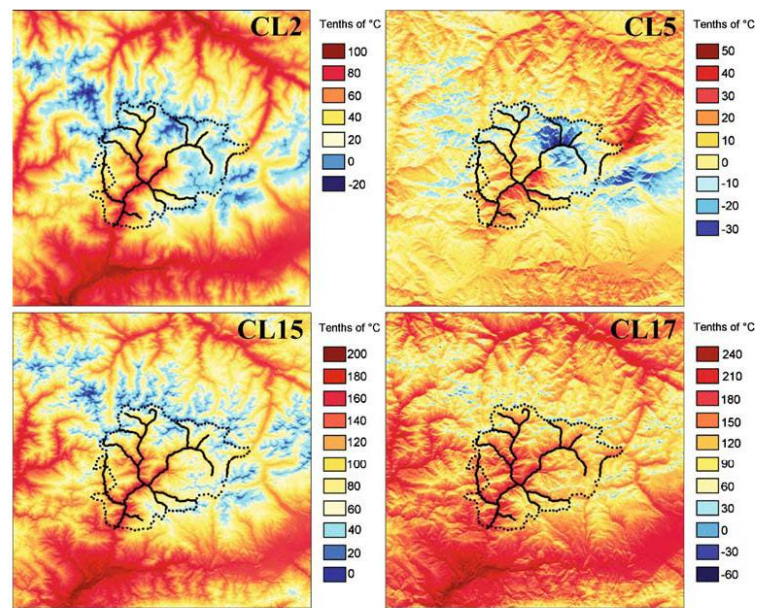


Fig. 6 Circulation patterns results for clusters (CL) 2, 5, 15 and 17. a Sea level pressure map, b 500 hPa geopotential height, c annual distribution of the days and d interannual variability

Fig. 7 Temperature maps obtained for clusters (CL) 2, 5, 15 and 17



be commented on are the presence of heat pools. Looking at CL 17 (Iberian thermal low, Fig. 6), this effect can be visually detected over the valleys in the southern area, especially in La Cerdanya basin (Fig. 7). It is less important in pattern 15 (Central Europe high), a weather type more frequent in winter (Fig. 6, column c). We would like to add that thermal inversions are not detected for this anticyclonic example in the corresponding minimum temperature map. Furthermore, positive anomalies are obtained only for Adrall weather station (642 m, south face) for the entire year (Fig. 4g), again showing a possible Föhn effect over the southern-side valleys. Finally, it is also interesting to mention that the differences for maximum temperatures between sunless slopes in the same valley are well represented, as can be observed by taking the Andorran rivers as reference.

3.2 Objective evaluation of the results

For evaluation purposes, we have used the averaged RMSE (root mean squared error) for the entire 20 synoptic circulation patterns. We have obtained a averaged cross-validation RMSE, in the case of temperature, between 0.6 and 1.7°C, and between 0.7 and 1.9 mm in the case of precipitation (see Table 2). Considering that daily fields are more difficult to model than monthly climatic averages, we could compare these results with the ones obtained in the case of monthly mean temperature mapping in Catalonia

(NE of Spain), which range between 0.5 1°C (Ninyerola et al. 2000). These authors used a fit set of 60% of the weather stations and a test set of the remaining 40%.

These kinds of comparisons could be difficult with other interpolation studies because of the different evaluation error measurements and interpolation approaches. However, it could be useful to review these data to provide a framework. In the case of temperature mapping in complex mountainous areas, Dodson and Marks (1997) obtained a MAE (mean absolute error) ranging between 1.2 1.3°C for the Columbia River basin. Bolstad et al. (1998) obtained a MAE between 1.4 2.3°C for the Appalachian Mountains. Hasenauer et al. (2003) obtained a MAE between 0.8 1°C in Austria. Hunter and Meentemeyer (2005) obtained a MAE between 1.7 2°C in California. Stahl et al. (2006) obtained a MAE 1.2 and 1.9°C in British Columbia. Hasenauer et al. (2003) and Hunter and Meentemeyer (2005) obtained average MAE values of 2.5 mm. in the case of precipitation.

Fig. 8 a-d Frequency distribution of daily precipitation amounts for clusters 3, 6, 12 and 16. e-f Monthly minimum temperature anomalies for clusters 2 and 5. g-h) Monthly minimum temperature anomalies for clusters 15 and 17. Calculated values at four weather sites: Adrall (642 m), Ransol (1,640 m), L'Hospitalet-près-l'Andorre (1,425 m) and Tarascon sur Ariege (475 m). Temperature anomalies calculated taking the period 1961–1990 as reference. Months with less than 20 days within a month have been not considered

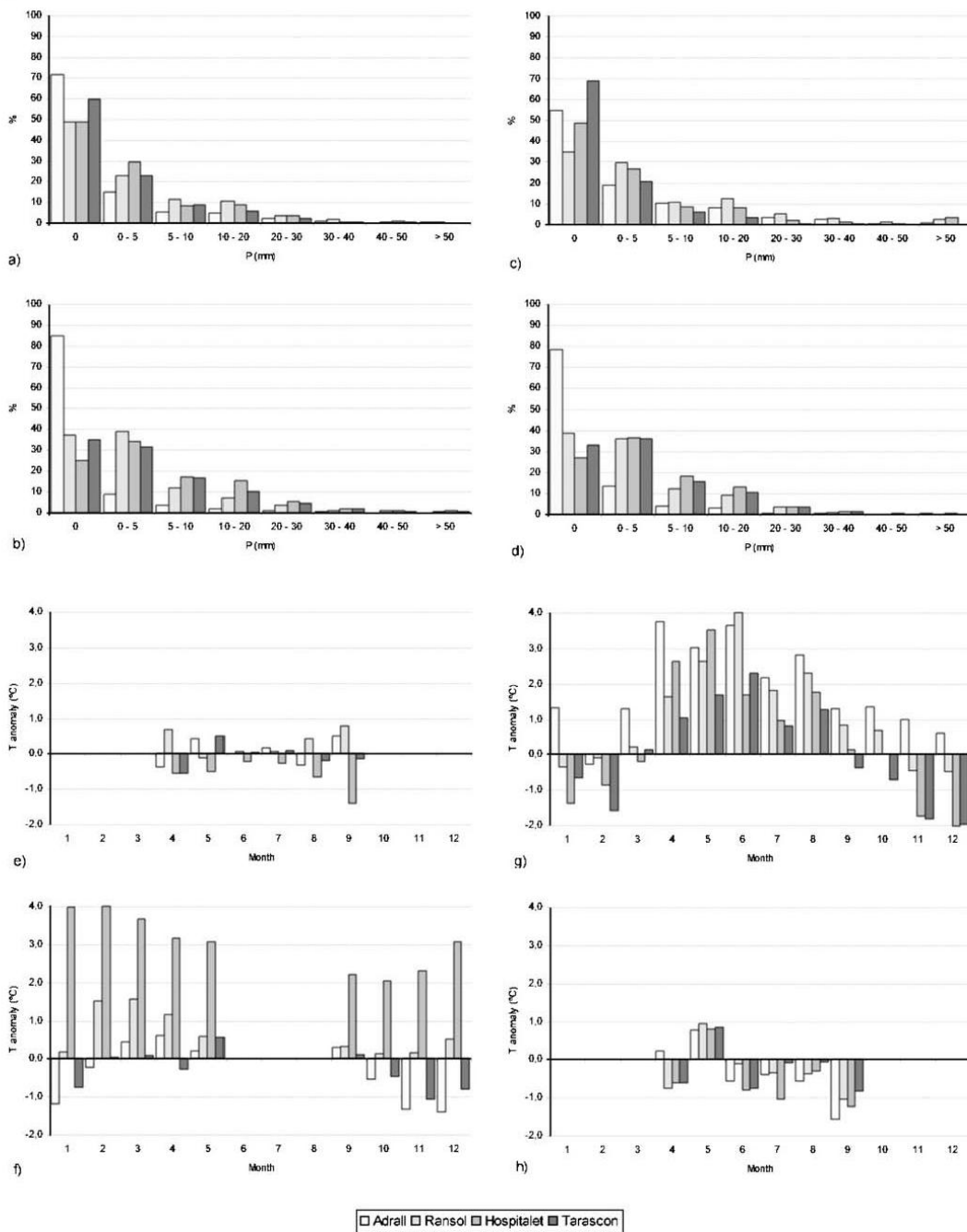


Table 2 RMSE values obtained from leave-one-out cross-validation for the proposed interpolation method (Reg) and for splines (Sp)

Cluster	MDMn (tenths of °C)		MDMx (tenths of °C)		MDP (mm)		DPP (%)	
	Reg	Sp	Reg	Sp	Reg	Sp	Reg	Sp
CL1	1.2	1.7	0.8	2.5	1.2	1.3	5.9	6.8
CL2	1.1	2.0	1.3	2.9	0.9	1.1	4.7	5.1
CL3	1.2	2.0	0.7	2.6	0.8	1.1	5.5	7
CL4	1.3	1.8	1.3	2.5	0.7	1.1	4.3	4.8
CL5	1.6	2.7	1.5	2.6	1.2	1.6	1.6	1.7
CL6	0.6	1.7	1.5	2.5	1.7	1.5	9.0	9.3
CL7	1.2	1.7	1.7	2.5	0.9	1.1	6.3	6.9
CL8	1.3	1.9	1.3	2.7	0.8	1.1	4.1	4.9
CL9	1.1	1.8	1.0	2.6	0.9	1.1	4.7	5.3
CL10	1.0	1.7	1.2	2.6	1.4	2.0	5.6	5.7
CL11	1.1	1.9	1.4	2.9	0.9	1.3	6.5	6.9
CL12	1.2	2.0	1.0	3.0	1.5	2.5	6.0	6.9
CL13	1.5	1.9	1.1	2.5	0.9	1.1	4.0	4.7
CL14	1.0	1.9	1.4	3.0	0.9	1.3	5.8	5.3
CL15	1.0	1.5	1.7	2.4	1.3	1.3	3.6	4
CL16	0.7	1.7	1.4	2.7	1.3	1.2	9.3	9.6
CL17	0.9	1.8	1.3	2.9	1.1	1.4	5.1	5.1
CL18	1.6	2.0	1.1	2.3	1.9	2.3	6.7	6.9
CL19	1.4	2.0	1.4	2.7	0.8	1.3	5.8	6.7
CL20	1.1	1.7	1.4	2.5	0.8	2.3	4.7	5.6
Mean average	1.1	1.9	1.3	2.6	1.1	1.5	5.5	6.0
Mean deviation	0.3	0.2	0.3	0.2	0.3	0.6	1.7	1.8

MDMn daily minimum temperature, *MDMx* daily maximum temperature, *MDP* daily precipitation, *DPP* daily precipitation probability and *CL* cluster/circulation pattern

Shaded values correspond to the spatial patterns shown in Figs. 5 and 7. Mean average and mean deviation correspond to the RMSE of all circulation patterns

In relation to other studies that distinguish among circulation patterns, Ustrnul (2006) obtained better results for monthly averages of the daily mean temperature in Poland. The range of RMSE goes from 0.37 to 0.73°C in January, and from 0.32 to 0.67°C in July.

In order to provide an improvement measure, we have compared our results with a splines interpolation which does not include geographical information (see Table 2). In the case of mean temperature, the average RMSE is 0.7°C (minimum temperatures) and 1.1°C (maximum temperatures) lower when using geographical information. In the case of rainfall, the difference, in the same direction, is of 0.5 mm. Moreover, in this case, the standard deviation of the RMSE is higher in the splines approach, showing that some circulation patterns (i.e. CL12 and CL18) are not well modelled by a splines performance.

Finally, we would like to highlight that worse RMSE results for certain patterns could be related to the internal variability of the cluster obtained in the circulation pattern classification. This is the case of pattern 12 in MDP, influenced by the difficulty of classifying low pressures over the Mediterranean area (Esteban et al. 2006), which are mainly included in this group.

4 Conclusions

Eighty detailed climatic maps of meteorological variables over Andorra (Pyrenees) in relation to a circulation-pattern catalogue over Western Europe has been obtained, with satisfactory results, both from a subjective (description of spatial features of climate from various sources) and objective (use of the RMSE) point of view. In this way, the daily variability of the atmospheric variables mapped is faithfully represented in relation to the main atmospheric circulation types.

We would like to emphasise the capability of this methodological proposal to model local effects, such as the increasing values of mean daily precipitation with altitude, and depending on the face orientation on the mountain range, heat pools and pluviometric shadows, and differences between sunny/shady slopes. Worse results are obtained for anticyclonic circulation patterns and the wintertime cool pools and thermal inversions expected. Otherwise, the authors consider that monthly or seasonal averages will surely improve the climatic maps. In addition, discussion on new independent variables to be included in the regression model and, especially, a better spatial

coverage of the weather stations in middle and high elevations, should be considered in future research. In this way, automatic weather stations are being installed in the highest areas of Andorra and will probably improve the results in the near future.

From an objective evaluation point of view, we have compared the RMSE of the maps obtained through the proposed methodology to the RMSE of the maps built using a splines interpolation (that does not use ancillary geographical data). Regression methodology provides lower RMSE values in all the cases. In addition, in the case of precipitation, the RMSE standard deviation is also lower when we introduce geographical information. This means that there are several synoptic patterns (especially CL 10) that are not well modelled when splines are applied.

In this way, from the circulation classification method's point of view, it is very interesting to highlight that the introduction of the frequencies of the patterns into the radiation values (monthly weighted radiation) improve the RMSE, showing that the inclusion of new advances on weather type classification will help for obtaining better results. More consistency in the clusters' internal homogeneity/external heterogeneity would surely improve the climatic maps' RMSE. More research is needed in this application of synoptic climatology, and the results here obtained suggest that this tool will improve our knowledge on climate-change impact, climate variability and anomalous-events analysis.

Acknowledgements The authors wish to thank Météofrance (French Meteorological office), Météocat (Meteorological Service of Catalonia) and INM (Spanish Meteorological Institute) for meteorological data availability; Javier Martín-Vide (Group of Climatology, University of Barcelona) and Montse Mases (CENMA-Snow and Mountain Research Centre, Andorra) for their scientific and technical support. To the SH164–630-VE01-CA02 project of the Spanish "Ministerio de Educación y Ciencia" and especially to the Action COST 733 on "Harmonisation and Applications of Weather Types Classifications for European Regions". Finally we would like to thank to Radan Huth (Institute of Atmospheric Physics, Czech Republic) and to the two anonymous reviewers for their useful comments.

References

- Abaurrea J, Asín J, Cebrián C, Centelles A (2004) Metodología para el control de calidad y homogeneidad de una base de datos de precipitación diaria (in Spanish). *Publ Spanish Climatol Soc* 4:431–440
- ACCORD (Atmospheric Circulation Classification and Regional Downscaling) (2000) Final report. <http://www.cru.uea.ac.uk/cru/projects/accord/>. Cited April 2007
- Action Cost 719 (2006) The use of geographical information systems in climatology and meteorology. <http://www.knmi.nl/samenw/cost719/>. Cited April 2007
- Action Cost 733 Harmonisation and Applications of Weather Types Classifications for European Regions. http://cost.cordis.lu/src/action_detail.cfm?action=733. Cited April 2007
- Baker Perry L, Konrad CE (2006) Relationships between NW flow snowfall and topography in the southern Appalachians, USA. *Climate Res* 32:35–47
- Barry RG, Carleton AM (2001) *Synoptic and dynamic climatology*. Routledge, London
- Barry RG, Perry AH (1973) *Synoptic climatology, methods and applications*. Methuen, London, 555 pp
- Birkeland KW, Mock CJ, Shinker JJ (2001) Avalanche extremes and atmospheric circulation patterns. *Ann Glaciol* 32:135–140
- Bissolli P, Müller-Westermeier G (2005) The spatial distribution of precipitation in Germany for different weather types. Proceedings from the 5th Annual Meeting of the European Meteorological Society, Session AW8, Weather Types Classifications. Utrecht, The Netherlands, 12–16 September 2005
- Bolstad PV, Swift LL, Collins F, Régnière J (1998) Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains. *Agr Forest Meteorol* 91:161–176
- Brunsdon C, McClatchey J, Unwin DJ (2001) Spatial variations in the average rainfall-altitude relationship in Great Britain: an approach using geographically weighted regression. *Int J Climatol* 21:455–466
- Buell CE (1975) The topography of the empirical orthogonal functions. Reprints Fourth Conference on Probability and Statistics in Atmospheric Sciences, Am. Meteorol. Soc., Tallahassee, FL, pp 188–193
- Cattell RB (1966) The scree test for the number of factors. *Multiv Behav Res* 1:245–276
- Chessa PA, Delitala AM (1997) Objective analysis of daily extreme temperatures of Sardinia (Italy) using distance from sea as independent variable. *Int J Climatol* 17:1467–1485
- Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J Appl Climatol* 33:140–158
- Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Res* 8:1–20
- Eastman JR (1987–1997) *Idrisi. Users' guide*, Clark University, Worcester, MA
- Esteban P, Jones PD, Martín-Vide J, Mases M (2005) Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *Int J Climatol* 25:319–329
- Esteban P, Martín-Vide J, Mases M (2006) Daily atmospheric circulation catalogue for Western Europe using multivariate techniques. *Int J Climatol* 26:1501–1515
- Feng S, Hu Q, Qian W (2004) Quality control of daily meteorological data in China, 1951–2000: a new dataset. *Int J Climatol* 24:853–870
- Gleason E (2002) *Global daily climatology network, V1.0*. National Climate Data Center, Ashville, NC
- Gómez B (2001) El fogony (efecto föhn) en el valle pirenaico de Sort (in Spanish). *Treballs de la Societat Catalana de Geografia*, no. 52, Institut d'Estudis Catalans, Barcelona, pp 311–320
- Goodale CL, Aber JD, Ollinger SV (1998) Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Res* 10:35–49
- Hair JF, Anderson RE, Tatham RL, Black WC (1998) *Multivariate data analysis*, 5th edn. Prentice-Hall, Englewood Cliffs, NJ
- Hasenauer H, Merganicova K, Petritsch R, Pietsch S, Thornton PE (2003) Validating daily climate interpolations of complex terrain in Austria. *Agr Forest Meteorol* 119:87–107
- Hope PK, Drosowsky W, Nicholls N (2006) Shifts in the synoptic systems influencing southwest Western Australia. *Clim Dynam* 26:751–764
- Hunter R, Meentemeyer RK (2005) Climatologically-aided mapping of daily precipitation and temperature. *J Appl Meteorol* 44:1501–1510

- Hutchinson MF (1995) Interpolating mean rainfall using thin plate smoothing splines. *Int J Geogr Inf Syst* 9:385–403
- Huth R (1996) An intercomparison of computer-assisted circulation classification methods. *Int J Climatol* 16:893–922
- Huth R, Ustrnul Z, Dittmann E, Bissolli P, Pasqui M, James P (2005) Inventory of circulation classification methods and their applications in Europe within the COST 733 Action. Proceedings from the 5th Annual Meeting of the European Meteorological Society, Session AWS, Weather Types Classifications. Utrecht, The Netherlands, 12–16 September 2005
- James PM (2007) An assessment of European synoptic variability in Hadley Centre Global Environmental models based on an objective classification of weather regimes. *Clim Dynam* 27:215–231
- Jarvis CH, Stuart N (2001) A comparison among strategies for interpolating maximum and minimum daily air temperatures, part I: the selection of 'guiding' topographic and land cover variables. *J Appl Meteorol* 40:1060–1074
- Jolliffe IT (1986) *Principal component analysis*. Springer, New York
- Kalkstein SL, Tan G, Skindlov JA (1987) An evaluation of three clustering procedures for use in synoptic climatological classification. *J Clim Appl Meteorol* 25:717–30
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetma A, Reynolds R, Jeene R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *B Am Meteorol Soc* 77:437–471
- Kostopoulou E (2003) The relationships between atmospheric circulation patterns and surface climatic elements in the eastern Mediterranean. PhD Thesis, University of East Anglia, Norwich, UK
- Lennon JJ, Turner JR (1995) Predicting the spatial distribution of climate temperature in Great Britain. *J Anim Ecol* 64:370–392
- Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modeling of temperature and precipitation through GIS techniques. *Int J Climatol* 20:1823–1841
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. *Mon Weather Rev* 110:699–706
- Perry M, Hollis D (2005) The generation of monthly gridded datasets for a range of climatic variables over the UK. *Int J Climatol* 25:1041–1054
- Philipp A, Jacobeit J, Della-Marta PM (2005) Classifying reconstructed daily pressure patterns for the period 1850 to 2003 in the North-Atlantic-European region by simulated annealing clustering. Proceedings from the 5th Annual Meeting of the European Meteorological Society. Session AWS, Weather Types Classifications, Utrecht, The Netherlands, 12–16 September 2005
- Pons X, Ninyerola M (2008) Mapping a topographic global solar radiation model implemented in a GIS and calibrated with ground data. *Int J Climatol* (in press)
- Price DT, McKenney DW, Nalder IA, Hutchinson MF, Kesteven JL (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agr Forest Meteorol* 101:81–94
- Reek T, Doty SR, Owen TW (1992) A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *B Am Meteorol Soc* 73:753–762
- Richman MB (1986) Rotation of principal components. *J Climatol* 6:293–335
- Seibert P, Frank A, Formayer H (2007) Synoptic and regional patterns of heavy precipitation in Austria. *Theor Appl Climatol* 87:139–153
- Stahl K, Moore RD, Floyer J, Asplin M, McKendry I (2006) Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agr Forest Meteorol* 139:224–236
- STARDEX (Statistical and Regional dynamical Downscaling of Extremes for European regions) final report (2005) <http://www.cru.uea.ac.uk/cru/projects/stardex/>. Cited April 2007
- Tait AB, Fitzharris BB (1998) Relationship between New Zealand rainfall and south-west Pacific pressure patterns. *Int J Climatol* 18:407–424
- Tveito OE (2002) Spatial distribution of winter temperatures in Norway related to topography and large-scale atmospheric circulation. In: IAHS PUB Workshop, Brasilia, Brazil, 20–22 November 2002
- Tveito OE, Ustrnul Z (2003) A review of the use of large-scale atmospheric circulation classification in spatial climatology, met. no. Report 10/03 KLIMA
- USGS (US Geological Survey) (2007) Mapping the world in 3D. <http://srtm.usgs.gov/>. Cited April 2007
- Ustrnul Z (2006) Spatial differentiation of air temperature in Poland using circulation types and GIS. *Int J Climatol* 26:1529–1546
- Yarnal B (1993) *Synoptic climatology in environmental analysis*. Belhaven, London, 195 pp
- Yarnal B, Comrie AC, Frakes B, Brown DP (2001) Developments and prospects in synoptic climatology. *Int J Climatol* 21:1923–1950

6.4. TIPUS DE CIRCULACIÓ RELACIONATS AMB L'ACTIVITAT DE LLAMPS A CATALUNYA I EL PRINCIPAT D'ANDORRA. *PHYSICS AND CHEMISTRY OF THE EARTH*, 2010.

En aquest treball es caracteritzen els patrons de circulació a resolució temporal 6 horària relacionats amb activitat substancial de llamps a Catalunya i Andorra entre el gener de 2003 i el desembre de 2007. Les dades de grid emprades han estat els *NCEP Final Analyses of the Global Tropospheric Analyses* a una resolució d'1° i pel sector 35°N-48°N a 5°O-8°E. La informació d'activitat de llamps s'ha observat pel sistema SAFIR de detecció de llamps operat pel Servei Meteorològic de Catalunya (SMC), sistema que cobreix la zona d'estudi. S'han obtingut nou tipus de circulació basats en l'anàlisi de components principals. L'aproximació de les puntuacions extremes (*extreme scores*) s'ha emprat prèviament a l'assignació de tots els casos per obtenir el nombre final de tipus de circulació i els seus centroides. Les diferències trobades en els patrons mitjans de pressió en superfície (SLP) ens han permès agrupar els tipus en tres patrons generals, tenint en compte la seva escala/el seu origen dinàmic. El primer grup de tipus mostra la diferent distribució dels centres d'acció a escala sinòptica relacionats amb l'ocurrència de llamps. El segon està lligat a la dinàmica mesoescalar, sobretot induïda per la presència de la serralada dels Pirineus. El tercer grup mostra els tipus amb baix gradient de pressió en superfície i en els quals l'activitat de llamps és una conseqüència de les dinàmiques tèrmiques (brises de costa i muntanya).

A part de reforçar la consistència dels grups obtinguts, l'anàlisi de la classificació resultant millora el nostre coneixement sobre la distribució geogràfica i els factors de gènesi de l'activitat tempestuosa a la zona d'estudi, i aporta informació complementària per donar suport a la predicció meteorològica.



Circulation types related to lightning activity over Catalonia and the Principality of Andorra

N. Pineda^a, P. Esteban^{b,c,*}, L. Trapero^b, X. Soler^{a,c}, C. Beck^d

^aMeteorological Service of Catalonia (SMC), Berlín 38, Barcelona E-08029, Catalonia, Spain

^bSnow and Mountain Research Centre of Andorra - Institut d'Estudis Andorrans (CENMA/IEA), Avda. Rocafort 21-23 Edifici Moll, Sant Julià de Lòria AD600, Andorra

^cGroup of Climatology - University of Barcelona, Laboratori de Climatologia, Montalegre 6, Barcelona E-08001, Catalonia, Spain

^dInstitute of Geography, University of Augsburg, Universitaetsstrasse 10 Augsburg D-86135, Germany

ARTICLE INFO

Article history:

Received 3 March 2009

Received in revised form 26 November 2009

Accepted 29 December 2009

Available online 14 January 2010

Keywords:

Atmospheric circulation types

Principal Component Analysis

Lightning activity

Type sequences

Frequency of occurrence

ABSTRACT

In the present study, we use a Principal Component Analysis (PCA) to characterize the surface 6-h circulation types related to substantial lightning activity over the Catalonia area (north-eastern Iberia) and the Principality of Andorra (eastern Pyrenees) from January 2003 to December 2007. The gridded data used for classification of the circulation types is the NCEP Final Analyses of the Global Tropospheric Analyses at 1° resolution over the region 35°N–48°N by 5°W–8°E. Lightning information was collected by the SAFIR lightning detection system operated by the Meteorological Service of Catalonia (SMC), which covers the region studied. We determined nine circulation types on the basis of the S-mode orthogonal rotated Principal Component Analysis. The “extreme scores” principle was used previous to the assignation of all cases, to obtain the number of final types and their centroids. The distinct differences identified in the resulting mean Sea Level Pressure (SLP) fields enabled us to group the types into three main patterns, taking into account their scale/dynamical origin. The first group of types shows the different distribution of the centres of action at synoptic scale associated with the occurrence of lightning. The second group is connected to mesoscale dynamics, mainly induced by the relief of the Pyrenees. The third group shows types with low gradient SLP patterns in which the lightning activity is a consequence of thermal dynamics (coastal and mountain breezes).

Apart from reinforcing the consistency of the groups obtained, analysis of the resulting classification improves our understanding of the geographical distribution and genesis factors of thunderstorm activity in the study area, and provides complementary information for supporting weather forecasting. Thus, the catalogue obtained will provide advances in different climatological and meteorological applications, such as nowcasting products or detection of climate change trends.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction and objectives

Circulation-type classifications have been applied to different spatial and temporal scales and atmospheric variables. Recently, due to the improvement of weather-station networks and the availability of gridded circulation data (NCEP/NCAR Reanalysis, ERA40 database, etc.), many authors are dealing with the daily scale in an attempt to further our knowledge of extreme events such as heavy precipitation, local phenomena like air pollution, or other weather- and climate-related impacts on society. As is well-known, establishing statistical relationships between the synoptic and the local scale, clearly improves the interpretation and conclusions that can be obtained from General Circulation Models (GCM) in the evaluation of climate change, due to the fact that they attempt to solve

limitations referring to the spatial resolution of the dynamical models. Furthermore, weather forecasting applications (among others) can also be derived from these statistical techniques (Huth et al., 2008). Thus, the main objective of the Action COST-733 (2009) on Harmonisation and applications of weather type Classifications for European Regions (<http://www.cost733.org/>) is to “achieve (a) general numerical method(s) for assessing, comparing and classifying weather situations in Europe, scalable to any European (sub) region with time scales between 12 h and 3 days and spatial scales of ca. 200–2000 km, applicable for a number of applications”. Methods such as ones based on correlations, Principal Components Analysis, Clustering Techniques or Neural Networks are being tested in this European project. See Huth et al. (2008) and Philipp et al. (2010) for an extensive review of these methods and applications.

Reap (1994) pointed out that circulation types can be used to analyse typical weather patterns related with thunderstorms, as well as the spatial distribution of lightning activity. Moreover, climatological studies on lightning can be useful in operational

* Corresponding author. Address: Snow and Mountain Research Centre of Andorra - Institut d'Estudis Andorrans (CENMA/IEA), Avda. Rocafort 21-23 Edifici Moll, Sant Julià de Lòria AD600, Andorra. Tel.: +376 742630.
E-mail address: pesteban.cenma@iea.ad (P. Esteban).

settings, as they can be used to develop probability equations for predicting thunderstorm distributions associated with the major flow patterns (Lericos et al., 2002). Several researchers have analyzed lightning distributions as a function of the prevailing large-scale flow over different regions in the world. For example, over central Florida (US), López and Holle (1987) showed its dependence on the prevailing low-level flow, and Lericos et al. (2002) have found that some flow patterns were more conducive to cloud-to-ground (CG) flashes than others. In Europe, Hagen and Finke (1999) have related the probability of thunderstorms in Germany to weather patterns based on a low-level flow classification. In Spain, Tomás et al. (2004) have found that convective processes can be associated with a small number of general synoptic situations obtained with the Jenkinson and Collison (1977) classification method. Following these analysis of lightning distributions as a function of the atmospheric circulation type, the aim of this paper is to improve our knowledge of thunderstorm activity, from the dynamical and statistical point of view, over Catalonia and the Principality of Andorra. This objective will be achieved through a classification of atmospheric circulation types associated with substantial lightning activity in 6-h periods, in order to obtain the main Sea Level Pressure (SLP) types and associated information such as: monthly and 6-h frequencies, sequence types, and lightning activity spatial distribution. Furthermore, for weather forecasting purposes, we analyzed some meteorological variables and indexes, and calculated the frequencies of occurrence of substantial lightning activity for SLP configurations showing a high degree of similarity to the centroid patterns of the circulation types.

A feature that must be considered for the analysis is the particular orography of the study area: the coastline oriented from northeast to southwest with two mountain ranges of around 1000 m altitude running roughly parallel to it, delimited by the Pyrenees (with maximums heights of 3000 m and where Andorra is located) to the North and the Ebro river valley to the West (Fig. 1). Air flows are highly influenced by these orographic features and their direction and speed can be modified. Some important effects are: blockage of the air flow by mountain barriers or generation of channelling winds through river valleys and consequently the presence of convergence areas inland.

2. Data and methods

2.1. Lightning data

Lightning information was collected by a SAFIR lightning detection system (Richard and Lojou, 1996) operated by the Meteorolog-

ical Service of Catalonia (SMC). Three sensors, covering the region of Catalonia and its contiguous areas, including the Principality of Andorra, make up the network (Fig. 1). The SAFIR system combines interferometry (in the very high frequency, VHF) for detecting intra-cloud flashes (IC) with direction finding (Low Frequency, LF) for discriminating cloud-to-ground flashes (CG). The SMC-SAFIR system presents a detection efficiency of around 90% and a lightning localization accuracy of approximately 2–3 km (Pineda and Montanyà, 2009). We grouped the lightning activity of the 5 years analyzed (2003–2007), into 6-h periods, according to four daily NCEP analyses hours (00 h–06 h–12 h–18 h UTC). Thus, the 6-h periods are from 21 h to 03 h, from 03 h to 09 h, etc. We only analyzed the 6-h periods with over 200 CG, which provided a sample of more than 460 6-h periods covering 82% of the CG flash counts recorded in Catalonia and Andorra from January 2003 to December 2007.

2.2. Numerical model data

The gridded data used for classification of the circulation types is the NCEP Final Analyses of the Global Tropospheric Analyses (NCEP FNL data; FNLDOC/NOAA/NWS/NCEP, 2000) at 1° resolution over the region 35°N–48°N by 5°W–8°E. The meteorological field classified is Sea Level Pressure (SLP).

2.3. Circulation type classification

The classification method used in this study is the one presented by Esteban et al. (2005, 2006, 2009) based on Principal Component Analysis (PCA) with the “extreme scores” variant for clustering all the cases. The method follows the well-known S-mode based procedure (Richman, 1986; Yarnal, 1993) with previous deseasonalisation of the data via spatial standardisation (scaling the rows). Additionally, the correlation matrix is used to provide the most efficient representation of variance in the dataset (Barry and Carleton, 2001), and the orthogonal Varimax rotation of Principal Components (PCs) is applied to facilitate the spatial interpretation of the PCs (Yarnal, 1993). The Scree test criterion and the North rule-of-thumb are the methods used for retaining the relevant PCs.

Based on the standardized PC scores, the decision regarding the number of groups (CTs) and the definition of their initial centroids is based upon the principle of “extreme scores”. The basic idea of this approach entails using observations exhibiting “extreme scores” on one of the retained and rotated PCs (usually values higher than +2 for the positive phase, lower than –2 for the negative phase) while having low score values on the remaining PCs (usu-

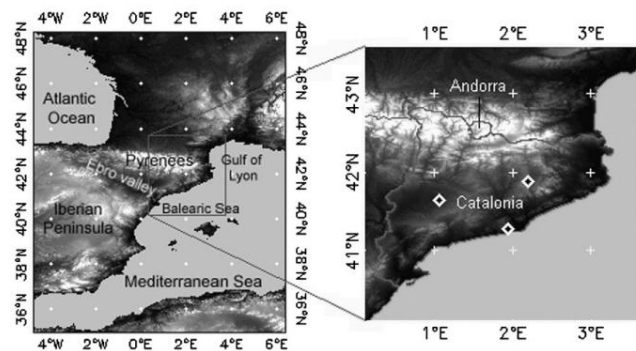


Fig. 1. (a) Area of interest, NE of the Iberian Peninsula. (b) Catalonia and Andorra with the lightning detection sensors (diamonds) of the Meteorological Service of Catalonia (SMC).

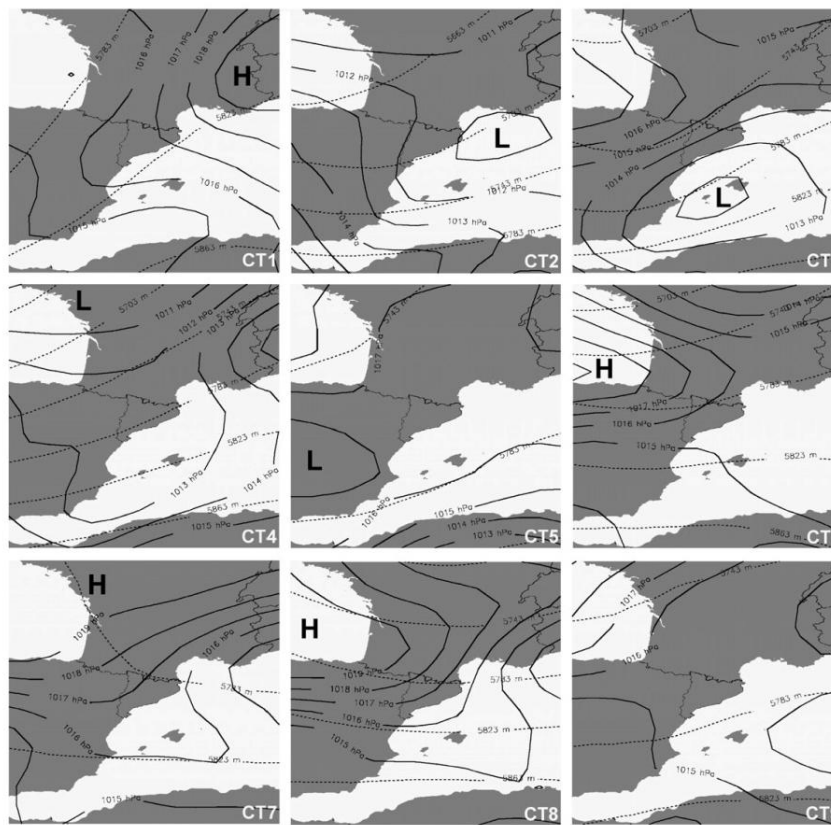


Fig. 2. SLP synoptic circulations types (solid lines) and 500 hPa geopotential heights (dashed lines) based on multivariate techniques (centroids of clusters 1–9 for 6-h periods with a threshold of over 200 cloud-to-ground flashes).

ally values between +1 and –1) for defining the initial centroids of the CTs of the final classification catalogue. Subsequently, these centroids are calculated by averaging all observations that can be assigned to one PC according to “extreme scores” criteria (Tait and Fitzharris, 1998; Birkeland et al., 2001; Kostopoulou, 2003). Normally 2.5–5% of all cases are used to obtain the initial centroids (4.10% for the classification presented in this paper). PCs to which no observation can be assigned (i.e. there is no day with a close spatial structure), are considered as an artificial result from the PCA, and consequently this potential group (circulation type) is eliminated. Thus, this procedure avoids artefacts in the final classification that may result when relying solely on the methods for retaining the main components and the Varimax rotation applied during the preceding PCA. Finally, all observations (6-h lightning periods) are assigned to the nearest initial centroid in terms of the Euclidean distance (calculated with the standardized scores values) for obtaining the definitive CTs. It should be noted that the commonly used *K*-means iterative method was not applied as the last step for classifying all the observations, assuming that the centroids were well-established using the “extreme scores” criterion (the centroids related to the PCs have a spatial structure close to the observed 6-h circulation cases), and no optimisation

procedure was required. Details of this method can also be found in Philipp et al. (2010).

3. Results

Nine SLP circulation types (CT) were identified with the above mentioned PCA and clustering method, which are shown in Fig. 2. Additionally, to facilitate the meteorological interpretation, the mean geopotential height at 500 hPa for each CT has been added to the maps. Fig. 3 shows the CT monthly frequencies, the frequencies for 6-h periods and the lightning activity of each pattern. It also shows transition frequencies among CTs between consecutive 6-h periods. Fig. 4 shows the spatial distribution of flashes determined over all the CG activity of the 6-h periods related to each CT.

3.1. CT1: Easterly flow

High pressures over continental Europe, generating Mediterranean winds (east to southeast) over the Catalonian coast and a south-westerly flow at 500 hPa causing significant shear (Fig. 2). The frequencies obtained show evident concentration of this

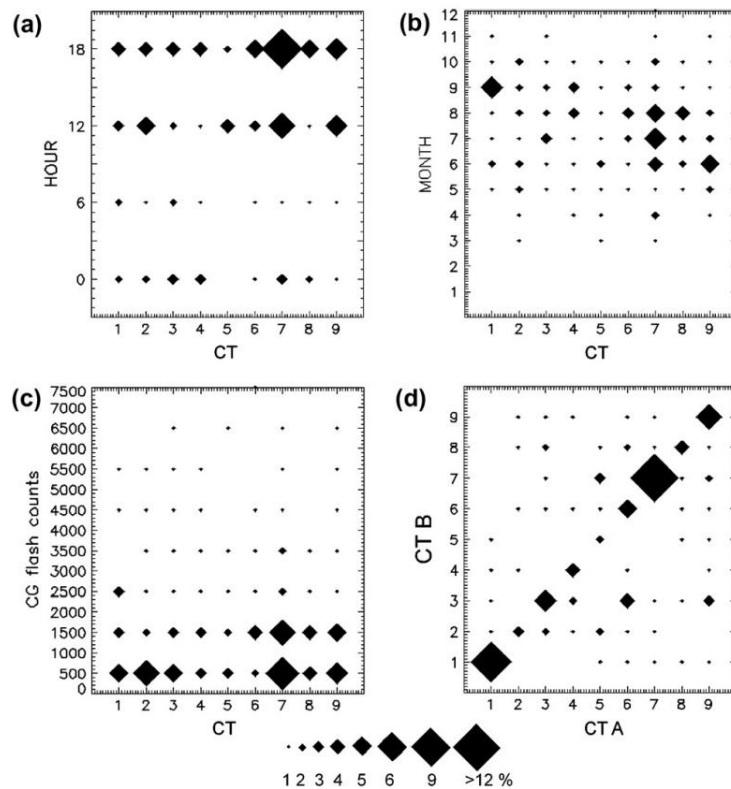


Fig. 3. (a) Six-hourly and (b) monthly frequencies of the clusters obtained, (c) distribution of circulation types (CTs) by cloud-to-ground flashes (CG), and (d) transition frequencies among CTs between consecutive 6-h periods (CT A-CT B).

pattern in September and a homogeneous hourly distribution (Fig. 3). The spatial distribution of flashes related to this circulation type exhibits a major concentration of lightning along the Mediterranean coastline (Fig. 4). This distribution, together with the homogeneous hourly distribution across the day, points to a maritime origin of the thunderstorms. The Mediterranean Sea supplies warm and humid air through onshore flow in the lower troposphere, especially during autumn (the so called "Mediterranean Air Mass" by Jansà (1959)), combined with the advance of a wave trough at 500 hPa. Such conditions can lead to heavy rainfall events, which are frequent on the Catalanian coastline at the beginning of autumn, and which sometimes affect the Pyrenees, depending on the orographic forcing required for instability (Llasat and Puigcerver, 1994).

3.2. CT2: Westerly flow with Mediterranean low

The monthly distribution of this pattern has distinct maximums in the spring and autumn months. This type represents a frontal system coming from the west and crossing the study area. In addition, the arrival of the associated air mass over the Mediterranean could generate a dynamical low, on average centered over the Gulf of Lyon. Spatial lightning distribution for this CT shows a mixture of the typical spring and autumn patterns in the region. As in spring, thunderstorms tend to occur in the Pyrenees and in nearby

continental Catalonia; in autumn the activity is located in the Mediterranean Sea and on the Catalanian coastline.

3.3. CT3: Mediterranean low

Dynamical low over the Mediterranean Sea, located on average in the Balearic Sea and pumping humid air over the Catalanian coast. This pattern presents intense lightning activity, mainly during the 6-h periods at 18 h and 00 h UTC. Occurrence thereof, despite the summer maximums, is well distributed over all months with substantial lightning activity in our dataset. The spatial lightning distribution for this CT shows that this Balearic Sea dynamical low generates lightning activity mainly along the Catalanian coast, especially the southern part. The other maximum of activity, located in continental Catalonia, is related to the channelling of the Mediterranean humid air in river valleys and subsequent orographic uplift further inland (mainly on the southern slopes of the Pyrenees).

3.4. CT4: Westerly flow

As in CT2, a frontal system is crossing the area, but in this case with a more pronounced south-westerly component of winds. Frequency maxima appear in August and September and during night time. Thunderstorms usually develop in the west

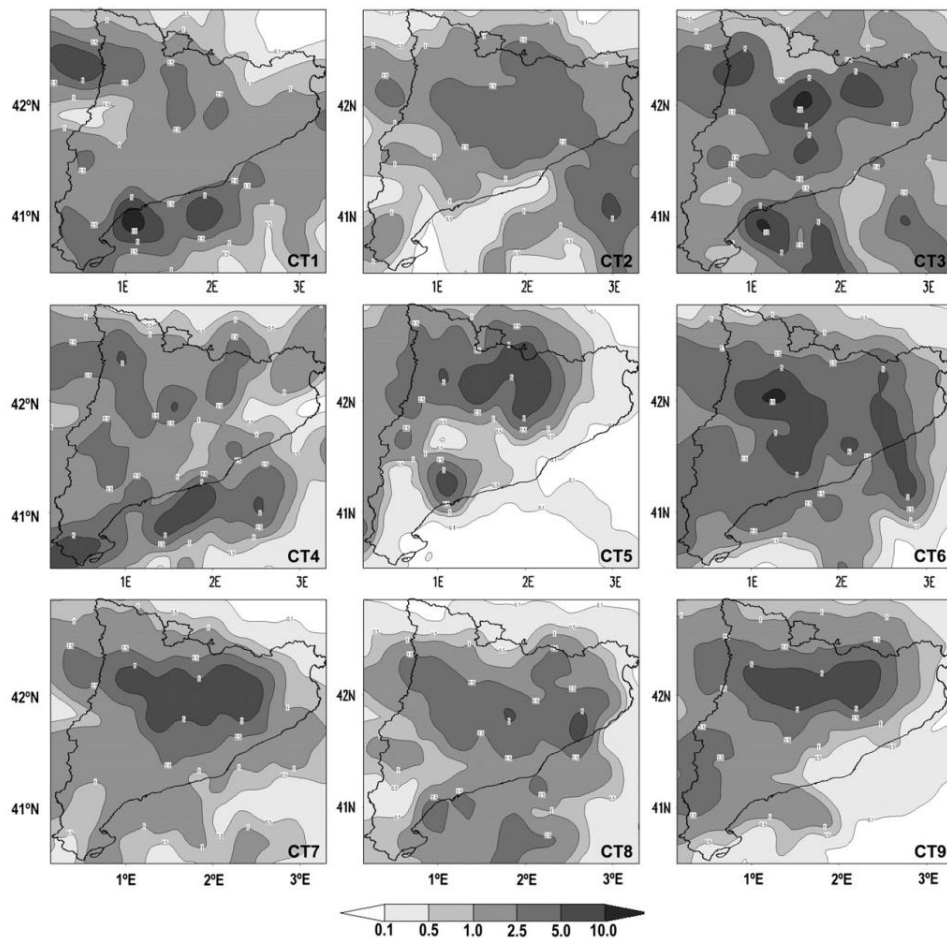


Fig. 4. Spatial lightning distributions over Catalonia and Andorra for each CT. Isolines are calculated after rating the total CG counts of each CT for the number of its corresponding 6 h-periods in a mesh of 0.1°.

of the study area (mainly in Aragon) and reach Catalonia in a mature stage, affecting the inland areas. Subsequently, contact between the frontal system and the Mediterranean Sea, which is warm and humid at the end of summer and autumn, reinforces the convection once again.

3.5. CT5: Iberian thermal low

Low-pressure gradient over the area and presence of a thermal low centered over the Iberia Peninsula. The annual maxima are in late spring and at the beginning of summer. It is interesting to highlight the lack of this type at night and in the morning; this diurnal evolution pattern is clearly related to the solar radiation cycle. The spatial lightning distribution for this CT has a minimum activity over the sea and shows high activity in mountain regions, related to local convection in spring and early summer, when solar heating over the relief generates higher vertical temperature gradi-

ents than the free atmosphere. Hence, this CT shows maxim values over Andorra and the topmost areas in the Pyrenees, mainly in its eastern half. This is a well-known circulation pattern over Iberia (Martín-Vide, 2005; Esteban et al., 2006), which is not always sufficiently clearly identified by classification methods.

3.6. CT6: Cold front

Low-pressure systems crossing Europe; the study area is affected by the southern part of the frontal system. Subsequently, the Azores high directs northerly winds over the Pyrenees. Monthly maximum registered in August and daily maxima at the afternoon–evening hours. The spatial lightning distribution for this CT shows thunderstorms affecting almost the whole region from west to east, probably due to the westerly flow at 500 hPa. It is important to highlight that the local maximum in the east is related to one single extreme event.

Table 1

Mean and standard deviation values for temperature difference between 850 and 500 hPa, (TD850-500), CAPE and Lifted Index (LI) for each circulation pattern (CT).

	TD850-500 (°C)	Cape (J/kg)	LI
CT1 – Easterly flow	27.7 (2.1)	365.7 (379.9)	-0.7 (2.2)
CT2 – Westerly flow & Mediterranean low	28.1 (2.2)	236.8 (335.9)	0.4 (2.5)
CT3 – Mediterranean low	27.6 (2.4)	323.7 (397.2)	-0.2 (2.5)
CT4 – Westerly flow	28.9 (2.5)	476.9 (458.2)	-1.1 (2.5)
CT5 – Iberian thermal low	29.0 (2.1)	231.2 (276.3)	0.1 (2.2)
CT6 – Cold front	28.4 (2.0)	378.6 (406.7)	-0.7 (2.7)
CT7 – North-eastern flow	28.4 (2.1)	275.0 (343.7)	0.2 (2.6)
CT8 – Azores anticyclone	26.8 (2.0)	270.1 (363.7)	0.6 (2.6)
CT9 – Non-gradient pattern	29.1 (1.8)	297.1 (408.4)	-0.2 (2.4)

3.7. CT7: North-eastern flow

Azores high extended over the Atlantic and Western Europe, causing north to east winds over the area. Lightning activity is concentrated over the southernmost chains of the Pyrenees, showing high amounts of lightning flashes in a large area. In this situation, the Pyrenees have an important orographic effect, triggering a low-pressure centre on the leeward side over Catalonia and Andorra (orographic dipole effect), where important convergence zones can be generated and further initiate convective processes (Bessemoulin et al., 1993). In addition to this orographic effect, the diurnal solar heating effect reinforces thunderstorm activity at the 12 h and 18 h UTC periods over the southern area of the Pyrenees range. Concerning Andorra, the area is affected by low lightning activity. This pattern represents 22% of the 6-h periods, and is also the most frequent one in summer, including most of the lightning activity.

3.8. CT8: Azores anticyclone

This centre of action spreads over the study area, and northerly winds blow over the Pyrenees. The orographic dipole effect is also present in this type, but a westerly to north-westerly flow at 500 hPa reduces thunderstorm activity over the Pyrenees area (CT with the lowest values over Andorra). It is well-known that the convergence is concentrated over the north-eastern area of Catalonia, represented by a small maximum on the lightning distribution maps. The occurrence frequencies maximum appears in summer and during the evening-night time period.

3.9. CT9: Non-gradient pattern

This type is very typical over the Mediterranean (in Catalan *Pantà Baromètric*, i.e. barometric swamp), showing very weak SLP gradients over the western Mediterranean. Its distinct daily cycle is associated with diurnal variations in solar radiation with maxima between 12 and 18 h. The seasonal maximum is concentrated in late spring and the beginning of summer. The spatial lightning distribution for this type is quite similar to Type 7, where convection and thunderstorms are located in mountainous regions (southern area of Pyrenees range). Furthermore, a secondary maximum is located in the inland areas of the southern half of Catalonia. Moderate values are observed over Andorra, mainly in its southern half.

3.10. Circulation pattern persistence

We analyzed the transitions between circulation types in two consecutive 6-h periods (2CT) in order to detect relationships between CTs and to analyse their persistence. A total of 43% of the 6-h samples used in the classification of circulation types are followed by another important lightning event, while 57% are “solitaire” events. As can be seen in Fig. 3d, of all the possible 2CT

combinations (CT A–CT B), the highest frequencies occur in sequences of the same CT. The most frequent sequence is CT7–CT7, presenting 15% of the cases, followed by CT1–CT1 (11%). The reason for these high percentages is persistent anticyclonic configurations, which are characteristic of both CTs. For instance, a 5-day episode for CT1 appeared during September 2006. Moreover, the low occurrence of the CT5–CT5 sequence (Iberian thermal low) can be attributed to the patent midday nature of this thermal pattern; which relatively increases the frequency of the CT5–CT7 sequence, because the thermal low over Iberia disappears in the afternoon-evening. The secondary maximum for CT6–CT3 and CT9–CT3 sequences is noteworthy. Both cases show similar evolutions, present a weak SLP gradient over the study area and are followed by the formation of a Mediterranean low. The first sequence is known in terms of weather forecasting and synoptic climatology over the study area. In summer the contact of cold air, associated with a frontal system (CT6), with the Mediterranean air mass favours the generation of cyclogenesis (CT3) and triggers atmospheric instability. This is not the case for the CT9–CT3 sequence, in which the results obtained can likely be conditioned by the size of the dataset for the period analyzed. It is also interesting to analyse when these 2CT sequences occur. A total of 54% of the 2CT combinations start at 12 h, indicating that the majority of thunderstorms starting in the early afternoon extend their lifetime until night. From the 2CT starting at noon, 50% do so with CT2, CT5 and CT7 configurations and thunderstorm formation is therefore mainly related to the diurnal radiation cycle.

3.11. Atmospheric indices characterizing circulation types

We calculated some instability indices, such as temperature difference between 850 and 500 hPa (TD850-500), Convective Available Potential Energy (CAPE) and Lifted Index (LI), for all the 6-h periods and the 7 cells corresponding to Catalonia and Andorra in the NCEP FNL data (Table 1).

The highest mean values of TD850-500 obtained (CT9 29.1 °C and CT5 29.0 °C) correspond to circulations types in which convection does not rely on dynamic factors, but rather is due to thermal contrasts between low and middle levels. These two types also exhibit lower CAPE values (231.2 J/kg and 297.1 J/kg respectively), indicating the lack of dynamic factors in convection triggering under these CTs. Furthermore, atmospheric instability (higher CAPE, lower LI) is highest for Westerly flow (CT4: CAPE 476.9 J/kg and LI -1.1), followed by cold front (CT6: CAPE 378.6 J/kg and LI -0.7) and Easterly flow (CT1: CAPE 365.7 J/kg and LI -0.7), which indicates that convection under these regimes is dominated mainly by dynamic factors; thus, values of TD850-500 in CT1 are quite low (27.7 °C). Overall, we have observed that in CTs for which thermal dynamics constitutes the most significant factor (i.e. CT5, CT9) the values of TD850-500 from June to August fluctuate at around 29 °C (with ±2 °C deviation), whereas the remaining CTs present lower values, close to 27 °C (±2, 2 °C deviation).

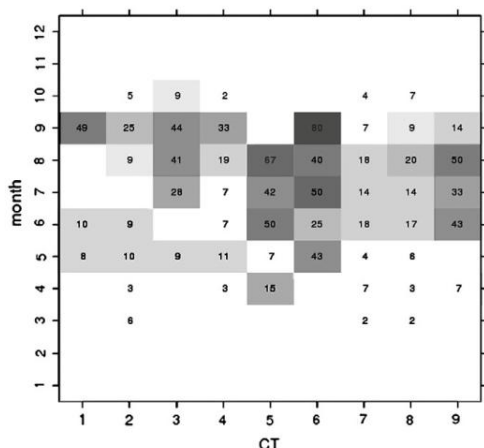


Fig. 5. Monthly frequencies of occurrence (%) of over 200 CG flashes for 6-h SLP fields from the 2003–2007 period similar (exhibiting a pattern correlation value higher than 0.8) to one of the nine centroids from Fig. 2.

Finally, it is worth pointing out that we observed similarities in the spatial distribution of lightning and the Atmospheric Instability Indices. For instance, the Easterly flow (CT1) shows higher lightning activity in the south (see Fig. 4), where the values of CAPE are higher in this CT (660.2 J/kg). Also, in the North-eastern flow (CT7), the CAPE and LI mean values (345.4 J/kg and -0.5 respectively) indicate that atmospheric instability is greater in the northern half of the study area, where lightning activity has proven to be the most intense. In CTs like the Easterly flow (CT1) and the Mediterranean low (CT3), convection is concentrated along the south coast of Catalonia, as a result of dynamic factors such as eastern advection or the presence of lows.

3.12. Frequencies of occurrence of enhanced lightning activity for SLP fields similar to CT centroids

From results presented to date, it is difficult to establish how typical the SLP-composites (centroids) representing the CTs are for periods of enhanced lightning activity, i.e. the frequency with which synoptic situations similar to one of the centroids are related to more than 200 flashes during the corresponding 6-h period. Thus, we performed an additional analysis in order to determine how often SLP fields resembling one of the lightning cluster centroids are related to enhanced lightning activity (over 200 flashes during the respective 6-h period). Using the same dataset as for the circulation type classification, for each of the nine centroids derived from the classification of 6-h SLP fields related to lightning activity (see Fig. 2), we calculated the pattern correlations (Pearson correlation coefficients) between the centroid and any 6-h SLP field from the whole sample (7304 cases). All cases (with and without enhanced lightning activity) were then assigned to the most similar of the nine CT centroids in terms of pattern correlation exceeding a value of 0.8. Following this approach, 50% of all lightning events and 39% of all non-lightning events can be assigned to the CT centroids derived from the classification of 6-h periods presenting lightning activity. This appears to indicate that not all synoptic situations related to lightning activity are clearly represented by the centroids shown in Fig. 2 and that these centroids (unsurprisingly) do not reflect the whole range of synoptic situations realized during the whole period from 2003 to 2007.

Subsequently, for each centroid the frequency of occurrence of enhanced lightning activity – for SLP configurations similar to the centroid – was calculated as the fraction of lightning cases assigned to the respective centroid. Fig. 5, which shows the resulting monthly frequencies of occurrence of enhanced lightning activity for the nine centroids from Fig. 2, reveals distinctly high frequencies of occurrence (>40%), especially for SLP configurations similar to the centroids of the low-gradient types CT5 (June–August) and CT9 (June and August), the cold-front type CT6 (May and July–September), the Mediterranean low represented by CT3 (August, September) and the Easterly flow situation from CT1 (September). However, it should be mentioned that some of the distinctly high relative frequencies of occurrence values in Fig. 5 (CT6 in September, CT5 in August and CT9 in August) rely on very low numbers of cases assigned to the respective centroids.

4. Conclusions

Based on 6-h periods of substantial lightning activity (over 200 flashes) over Catalonia and the Principality of Andorra from 2003 to 2007, we obtained a nine-type circulation catalogue. The classification method used in this study is based on an S-mode Principal Component Analysis with the “extreme scores” variant for clustering all the cases.

Clear differences can be found on observing the average SLP fields obtained for each type, and these differences become more distinct when other parameters such as frequency (monthly and 6-h), spatial distribution of flashes and pattern persistence are analyzed. Consequently, it can be concluded that three main groups of patterns are obtained from the scale/dynamical origin point of view: those related to synoptic dynamics (CT1, CT3, CT4, CT6), i.e. resulting from the location of the centres of action and the pass of frontal systems; those related to mesoscale dynamics (CT2, CT7, CT8), i.e. the formation of mesoscale systems as a consequence of orographic effect (mainly the Pyrenees); those related to thermal dynamics associated with solar heating (CT5, CT9), i.e. as a consequence of the effect of coastal and mountain breezes triggering instability.

It is also interesting to highlight the analysis of the persistence of the types obtained, which reinforces the consistency of the results. It shows how types characterized by anticyclonic configurations are more likely to persist, especially in mid summer, while thermal patterns usually show short sequences and are concentrated between midday and evening.

Furthermore, the use of atmospheric instability indices related to each pattern did not provide the results expected with regard to improving the characterization of the types obtained. However, the results of TD850–500 (combined with other factors like orography or meso scale flows) could help to improve weather forecasting during summer months, especially when CT associated with solar heating are present (CT5 and CT9).

Nonetheless, there is some similarity between the spatial distribution of the indices calculated and the location of lightning activity, particularly in the types related to dynamical factors. We will perform a painstaking analysis in this sense in the near future using radiosonde data over the study area, which will enable us to improve the results in terms of mean and deviation of CAPE and LI values obtained with the FNL data.

Finally, the results obtained from the monthly frequencies of occurrence of substantial lightning activity according to the SLP fields can be considered to be relevant with respect to lightning forecast issues. It should not be forgotten that some of the distinctly high occurrence frequencies are estimated on the basis of a very low number of cases. It can, however, be concluded that representatives of the group of CTs reflecting synoptic dynamics (CT1,

CT3, CT6), as well CTs characterized by weak gradients (CT5 and CT9) exhibit a close relationship with enhanced lightning activity, at least during some months between June and September. To the contrary, all CTs reflecting mesoscale dynamics (CT2, CT7 and CT8) generally show lower values of frequencies of occurrence.

To conclude, our classification method is useful for the identification of relevant circulation types related to non-continuous events such as lightning activity, and for deriving valuable climatological and meteorological information. The procedure presented in our paper could also be applied to other phenomena, including extreme events, thus helping to increase our knowledge of their genesis mechanisms, and to create “rules-of-thumb” for weather forecasters.

Acknowledgments

The authors wish to thank Thomas Krennert of the Central Institute for Meteorology and Geodynamics (ZAMG), Austria, Oriol Argemí and Joan Bech from the Meteorological Service of Catalonia (SMC) and Javier Martín-Vide from the University of Barcelona (2009 SGR 443 GRUP DE CLIMATOLOGÍA) for their technical support. We would also like to thank the two reviewers for their useful comments, which were of considerable help in improving the text.

References

- Action Cost 733, February 2009. Harmonisation and Applications of Weather Types Classifications for European Regions. <<http://www.cost733.org>>.
- Barry, R.G., Carleton, A.M., 2001. Synoptic and Dynamic Climatology. Routledge, London.
- Bessemoulin, P., Bougeault, P., Genovés, A., Jansà, A., Puech, D., 1993. Mountain pressure drag during PVREX. *Beiträge zur Physik der Atmosphäre* 66, 305–325.
- Birkeland, K.W., Mock, C.J., Shinker, J.J., 2001. Avalanche extremes and atmospheric circulation patterns. *Annals of Glaciology* 32, 135–140.
- DOC/NOAA/NWS/NCEP, 2000. NCEP FNL Global Tropospheric Analyses, 1x1, daily 1999jul30–present, updated daily. Published by the CISL Data Support Section at the National Centre for Atmospheric Research, Boulder, CO (ds083.2).
- Esteban, P., Jones, P.D., Martín-Vide, J., Mases, M., 2005. Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *International Journal of Climatology* 25, 319–329.
- Esteban, P., Martín-Vide, J., Mases, M., 2006. Daily atmospheric circulation catalogue for Western Europe using multivariate techniques. *International Journal of Climatology* 26, 1501–1515.
- Esteban, P., Ninyerola, M., Prohom, M., 2009. Spatial modelling of air temperature and precipitation for Andorra (Pyrenees) from daily circulation patterns. *Theoretical and Applied Climatology* 96, 43–56.
- Hagen, M., Finke, U., 1999. Motion characteristics of thunderstorms in southern Germany. *Meteorological Applications* 6, 227–239.
- Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J., Tveito, O.E., 2008. Classification of atmospheric circulation patterns – recent advances and applications. *Annals of the New York Academy of Sciences* 1146, 105–152. doi:10.1196/annals.1446.019.
- Jansà, J.M., 1959. La masa de aire mediterránea. *Revista de Geofísica* 18, 35–50.
- Jenkinson, A.F., Collison, F.P., 1977. An initial climatology of gales over the North Sea. Synoptic Climatology Branch Memorandum No.62. Meteorological Office, Bracknell.
- Kostopoulou, E., 2003. The relationships between atmospheric circulation patterns and surface climatic elements in the eastern Mediterranean. PhD. Thesis, University of East Anglia, Norwich.
- Lericos, T.P., Fuelberg, H.E., Watson, A.L., Holle, R.L., 2002. Warm season lightning distributions over the Florida Peninsula as related to synoptic patterns. *Weather and Forecasting* 17, 83–98.
- Llasat, M.C., Puigcerver, M., 1994. Meteorological factors associated with floods in the north-eastern part of the Iberian Peninsula. *Natural Hazards* 9, 81–93.
- López, R.E., Holle, R.L., 1987. The distribution of lightning as a function of low-level wind flow in central Florida. NOAA Tech. Memo. ERL ESG-28, National Severe Storms Laboratory, Norman, OK, 43 pp.
- Martín-Vide, J., 2005. Los mapas del tiempo. Mataró: Davinci, 219pp.
- Philipp, A., Bartholy, J., Beck, C., Erpicum, M., Esteban, P., Fettweis, X., Huth, R., James, P., Jourdain, S., Kreienkamp, F., Krennert, T., Lykoudis, S., Michalides, S., Pianko, K., Post, P., Rassilla Álvarez, D., Spekat, A., Tymvios, F.S., 2010. COST 733CAT – a database of weather and circulation type classifications. *Physics and Chemistry of the Earth* 35, 360–373.
- Pineda, N., Montanyà, J., 2009. Lightning detection in Spain: the particular case of Catalonia. In: Betz, H.D., Schumann, U., Laroche, P. (Eds.), *Lightning: Principles, Instruments and Applications: Review of Modern Lightning Research*. Springer, p. 641.
- Reap, R.M., 1994. Analysis and prediction of lightning strike distributions associated with synoptic map types over Florida. *Monthly Weather Review* 122, 1698–1715.
- Richard, P., Lojou, J.Y., 1996. Assessment of application of storm cell electrical activity monitoring to intense precipitation forecast. In: 10th Int. Conf. on Atmospheric Electricity, Osaka, Japan, pp. 284–287.
- Richman, M.B., 1986. Rotation of principal components. *Journal of Climatology* 6, 293–335.
- Tait, A.B., Fitzharris, B.B., 1998. Relationship between New Zealand rainfall and south-west Pacific pressure patterns. *International Journal of Climatology* 18, 407–424.
- Tomás, C., de Pablo, F., Rivas Soriano, L., 2004. Circulation weather types and cloud-to-ground flash density over the Iberian Peninsula. *International Journal of Climatology* 24 (1), 109–123.
- Yarnal, B., 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press, London.

7. CONCLUSIONS I DISCUSSIÓ

Amb els treballs presentats anteriorment, podem fer dos grans blocs de conclusions que cal considerar. Així, d'una banda, hi ha les conclusions associades pròpiament al mètode de les puntuacions extremes per classificar la circulació atmosfèrica i, de l'altra, les conclusions associades a les problemàtiques climàtiques o meteorològiques plantejades a cadascun dels articles.

7.1. PEL QUE FA AL MÈTODE DE CLASSIFICACIÓ

Tal com es deia als apartats 3 i 5 del present treball, l'aplicació del **mètode de les puntuacions extremes** a casos com els presentats anteriorment seria la forma de validar el funcionament d'aquesta proposta metodològica per a la classificació de tipus de circulació.

Veient els treballs realitzats, considerem que els resultats obtinguts són força satisfactoris tant des d'una perspectiva meteorològica com climàtica. Així, s'ha resolt satisfactòriament l'obtenció de tipus de circulació representatius de la resolució diària, alhora que són capaços d'incorporar les variacions estacionals pròpies de cada tipus pel que fa als canvis de gradient bàric o el seu comportament interanual. És per això que creiem que s'han assolit els objectius derivats de la hipòtesi de treball principal plantejada a l'apartat 3: L'ús de les classificacions de tipus de circulació es presenta com una eina molt eficient per abordar estudis de Climatologia sinòptica a sectors de relleu complex o quan volem aprofundir en les relacions entre la circulació atmosfèrica i l'escala local. D'aquesta manera, es poden deixar enrere les tècniques manuals, sobretot si volem treballar amb grans volums de dades i obtenir catàlegs molt pràctics de forma eficient sense patir per la capacitat de càlcul dels nostres ordinadors.

En aquest sentit, la seva aplicabilitat tant en el camp del canvi climàtic com en el de la cartografia climàtica ha quedat demostrada amb els resultats obtinguts als articles 2 i 3. Alhora, la capacitat del mètode per treballar amb diferents resolucions espacials i temporals s'ha comprovat amb la seva aplicació al cas del llamps.

A més a més, el fet de no haver-ne obtingut resultats absurds o clarament erronis mostra també la capacitat de la classificació de tipus de circulació de reproduir de forma simplificada i correcta el comportament normal de la dinàmica atmosfèrica en un sector d'estudi. Així, tot i partir de l'ACP amb una matriu tipus S, i tenint en compte

les consideracions fetes per [Compagnucci i Richman \[2008\]](#) (vegeu apartat 2.4.3 del present treball), el mètode de les puntuacions extremes presentat en aquest treball basaria la seva classificació dels casos en una identificació prèvia dels centres d'acció més rellevants a la zona d'estudi i, sense fer-ne cap interpretació física (sinòptica), passaria a classificar la mostra a partir de detectar els casos més representatius de l'activitat d'aquells centres d'acció al llarg del període d'estudi, i després assignar la resta dels casos mitjançant la distància euclídia.

Evidentment, molts aspectes ens poden servir per trobar punts febles a la metodologia: alta dependència dels mètodes de selecció de components, subjectivitat en el llinar de selecció de les puntuacions extremes (depenent de la mida i variabilitat de la mostra), o possible desvirtuació del catàleg obtingut amb el criteri de les puntuacions extremes, fruit de forçar la classificació final de tots els casos. Tot i això, i tenint presents els resultats i consideracions de l'Acció COST733, la qualitat dels resultats que es poden obtenir amb aquesta metodologia no s'allunya de la dels altres mètodes existents, alhora que tampoc hi ha cap aproximació metodològica que destaquí clarament. En aquest sentit, pensem que l'aproximació presentada és prou robusta per a ser considerada com a mètode per a la classificació de tipus de circulació (CTC), tant amb aplicacions de caràcter meteorològic com climàtic. Si a més tenim en compte aquests resultats de l'Acció COST733 que no determinen cap mètode com el millor, es fa recomanable que qualsevol treball de Climatologia sinòptica basat en les CTC consideri la conjunció (*ensemble*) de resultats obtinguts amb diferents metodologies, entre les quals es pot emprar el mètode de les puntuacions extremes.

Per això, apuntem com a proper pas d'interès a fer després del present treball el d'abordar l'aplicació de diferents metodologies de CTC per resoldre problemàtiques tant meteorològiques com climàtiques com les presentades en aquesta tesi doctoral i comparar-les. D'altra banda, i vista la complexitat d'alguns fenòmens naturals, com ara les allaus de neu, entre d'altres, la classificació de seqüències de circulació tal com s'ha mostrat als apartats 2.4.4 o 2.5.2 pot ser un altre camp de les CTC per a desenvolupar

properament. En aquest sentit, i com a reconeixement a Andreas Philipp i Christoph Beck com a principals impulsors i desenvolupadors, així com a tots els col·laboradors, serà de molta utilitat el programari lliure COST733class desenvolupat recentment (geo23.geo.uni-augsburg.de/cost733class-1.1), i en el qual, junt amb moltes altres aproximacions, hi ha implementada la **metodologia de les puntuacions extremes** presentada en aquest treball.

7.2. PEL QUE FA A LES NEVADES INTENSES A ANDORRA

L'anàlisi dels set tipus de circulació obtinguts mostra que les situacions advectives amb vents en superfície bufant de nord-oest a nord són les estructures de circulació més freqüents associades a les nevades intenses en el sector d'estudi (Andorra). Aquestes situacions estan seguides per tipus de l'est al nord-est i per un tipus de component sud.

En tots aquests patrons hi ha tres elements rellevants: humitat provinent de la Mediterrània o l'Atlàntic, advecció d'aire fred provinent de latituds més elevades (especialment quan aquest ve de l'Europa continental), i moviments verticals fruit de la combinació de vents forts, l'ascensió provocada per l'efecte del relleu pirinenc, i la configuració a 500 hPa amb solcs o baixes en alçada.

Un aspecte que creiem interessant de comentar és que la classificació obtinguda ha permès diferenciar tipus de circulació semblants pel que fa a la seva estructura de circulació (mateixa component), però diferenciats pels gradients de pressió o per la distribució espacial dels valors de pressió, fet que creiem d'interès quan es vol donar un enfocament meteorològic a aquest tipus de catàlegs.

Com a tipus de circulació que volem destacar tindriem el 3, associat a l'advecció de component meridiana sobre la serralada pirinenca. Aquest tipus destaca per aportar les màximes mitjanes de precipitació de neu a les tres estacions tractades (Ordino, 40 cm; Pas de la Casa, 44 cm; Arinsal, 45 cm), alhora que dos dels tres màxims absoluts (65 cm a Ordino i 75 cm a Pas de la Casa). Pel cas d'Arinsal el tipus 3 va associat al segon màxim absolut (70 cm), que denota la rellevància d'aquesta situació a tot el Principat.

Amb relació a aquest tipus de circulació i segons la premsa andorrana, del 28 al 30 de gener de l'any 2006 es va donar un episodi que a Andorra va requerir voladures a

diferents zones de sortides d'allaus, es va produir un ferit per allau al sector de Pas de la Casa, i la carretera entre Hospitalet-près-l'Andorre i Pas de la Casa va patir també els efectes d'un despreniment de neu. Al Pirineu català va ser molt rellevant l'activitat d'allaus, i es van produir despreniments que van superar els límits establerts per la cartografia existent; va haver-hi afectació d'edificis, de pistes d'esquí i es van destruir parts de boscos centenaris. L'existència d'una capa feble prèvia formada per gebre en superfície, junt amb la important acumulació de neu humida associada a aquest episodi, va afavorir les allaus generalitzades i de grans dimensions [Esteban *et al.*, 2006].

Unes altres situacions que cal destacar són la 4 i la 7, ambdues de direcció nord-oest, i que, com ja se cita a l'article, correspondrien a les condicions atmosfèriques ocorregudes a la gran allau d'Arinsal del 8 de febrer de 1996 [Esteban *et al.* 2002]. Precisament, per a l'estació d'Arinsal, el màxim històric registrat (85 cm) correspon a aquest tipus de circulació.

Aquestes dues situacions, junt amb el tipus 5, representen les configuracions associades a les adveccions del nord-oest, d'important impacte també sobre el Principat. En aquest sentit, s'han donat situacions els anys recents amb precipitacions abundants de neu que podrien assignar-se a aquests tipus de circulació. Així, destaquen l'episodi del 10 i 11 de març de 2006 (CL7, es requerí fer diferents voladures en zones d'allaus d'Andorra, alhora que es van tancar alguns accessos) i l'episodi del 10 i 11 de febrer de 2009 (CL5), que donà nevades molt abundants fins a totes les cotes dins del principat d'Andorra [Trapero i Esteban, 2010].

Interessant també és el tipus 2, amb una component aparentment més continental, però amb marcada entrada d'aire fred i la presència d'una baixa a la Mediterrània occidental. Una situació meteorològica rellevant recent que es podria assignar a aquest tipus de circulació va ser l'episodi de nevades intenses del 26 al 31 de gener de 2003, amb acumulacions que van superar de llarg els 30 cm / 24 h a moltes zones, i

que va anar associat a activitat d'allaus en sectors com la Vall d'Aran [Esteban *et al.*, 2007] i Andorra. Concretament, durant l'episodi i els dies posteriors (fins al dia 6 de febrer) es van donar dues grans allaus en zones boscoses de les valls de Canillo i el Serrat, quatre allaus més que van deixar aïllades tres-centes persones al Serrat durant sis hores. Es va fer el desallotjament de les Salines, les Bordes d'Arans i Aina, i voladures a les zones de sortida de les Fonts, Arinsal i Soldeu, segons la premsa andorrana.

7.3. PEL QUE FA A LA CLASSIFICACIÓ A L'EUROPA OCCIDENTAL

Amb aquest treball s'han obtingut vint tipus de circulació que representen de forma força acurada les principals configuracions de circulació atmosfèrica a l'Europa occidental entre 1960 i 2001.

Una de les característiques que convé destacar del catàleg obtingut és que respecta la resolució de les dades de partida; és a dir, que obtenim patrons realistes amb la resolució diària. Veiem com centres d'acció com l'anticicló de les Açores, depenent de la seva localització i extensió, poden produir diferents tipus de circulació, fet d'interès quan volem una classificació que s'ajusti a la resolució diària i no tant a la mensual o estacional (aquí ja parlaríem potser de patrons de circulació). A més, les situacions de baix gradient de pressió, molt importants des de la perspectiva climàtica i meteorològica de l'àrea mediterrània, es veuen ben representades en el catàleg.

Un altre dels aspectes que cal remarcar dels resultats són les tendències calculades pels diferents tipus per al període d'estudi (1960-2001); tendències que en alguns casos han resultat significatives. En aquest sentit, semblaria que les situacions de bloqueig sobre el sector continental europeu haurien guanyat pes al llarg de les darreres dècades. D'altra banda, i ja pensant en les implicacions pluviomètriques que pot haver tingut això al nostre sector geogràfic més proper, creiem interessant fer notar l'evolució decreixent patida pel CL2, el CL11 i el CL16 (implicacions sobretot als Pirineus), així com l'increment del CL15 (implicacions sobretot a la façana mediterrània de la península Ibèrica), coherent amb el comportament seguit per l'Oscil·lació de la Mediterrània Occidental (WeMO) [Martin-Vide i López-Bustins 2006].

7.4. PEL QUE FA A LA CARTOGRAFIA CLIMÀTICA AMB SIG

Els resultats obtinguts amb la cartografia climàtica associada als vint tipus de circulació mostren que la discriminació feta en la circulació atmosfèrica a partir de l'establiment d'un catàleg de tipus de circulació permet també caracteritzar el comportament de variables climàtiques com ara la precipitació o la temperatura a escala local i en sectors d'orografia complexa. En aquest sentit, es fan evidents les capacitats d'aquesta aproximació metodològica a l'hora de fer *downscaling* o reduccions d'escala, que denota així el seu potencial en l'ús tant meteorològic com climàtic. Això obre les portes a abordar la generació d'escenaris de canvi climàtic en sectors on les sortides directes dels models de circulació general (MCG) o els regionals (MCR) no donen encara respostes satisfactòries.

També volem fer notar no només la qualitat dels resultats de les interpolacions a través dels RMSE obtinguts, sinó també els resultats mostrats a la FIGURA 8 de l'article. Així, si considerem la distribució espacial de les estacions, queden ben reflectits els gradients nord-sud existents a la zona, fruit del relleu i la diferent exposició (vessant nord/sud del Pirineus) dels emplaçaments. A més, creiem convenient destacar els canvis de comportament estacional que mostren alguns tipus i en pocs quilòmetres de distància; tipus que resulten càlids una part de l'any i freds la resta, quan parlem de temperatures mínimes (molt evident en el tipus 15), variació que també depèn d'on estigui situada l'estació (cotes més elevades o fons de vall), com es veu al gràfic de la temperatura mínima del tipus 5.

Aquests detalls també són identificables quan mirem la variable precipitació, i en aquest sentit és de gran interès fixar-se en la cua del gràfic c de la FIGURA 8, on s'apunta la presència de precipitacions abundants en el tipus 12.

Finalment, les limitacions previsibles d'un model lineal com els models de regressió múltiple prenen relleu en situacions complexes, com ara les inversions tèrmiques (tipus 5), mentre que els seus avantatges ens permeten obtenir valors molt raonables

de variables meteorològiques en sectors amb manca d'observacions, i en aquest cas a l'alta muntanya, on no existeixen registres històrics de longitud suficient.

7.5. PEL QUE FA ALS PERÍODES D'ACTIVITAT DE LLAMPS SIGNIFICATIUS

Els nou tipus de circulació obtinguts en aquest catàleg mostren les situacions relacionades amb la presència significativa d'activitat de llamps al sector nord-est de la península Ibèrica i per a períodes de sis hores. En aquest sentit, cal fer notar l'àmbit relativament reduït de la zona sobre la qual s'ha realitzat la classificació (i a una resolució de malla d'1º), alhora que l'alta resolució temporal dels casos treballats.

Dels resultats obtinguts destaquen els diferents productes derivats de les agrupacions fetes i que aporten un avenç significatiu en el coneixement tant climàtic com meteorològic dels llamps a Catalunya i Andorra, i tant des d'una perspectiva espacial com temporal. Per exemple, als gràfics de la FIGURA 3 del treball, on es fa palesa la major concentració de l'activitat elèctrica a partir de migdia i durant la tarda, alhora que els mesos d'estiu i d'inici de tardor concentren també els màxims d'ocurrència. En aquest punt és força interessant veure com el tipus 1, el llevant, presenta el màxim d'activitat elèctrica al setembre.

La distribució espacial que mostren els mapes de la FIGURA 4 també aporten una mica més de llum sobre el comportament de les diferents situacions, que demostra com és d'important el relleu per a l'ocurrència d'aquests fenòmens tant quan es dóna el dipol orogràfic, per la presència dels Pirineus (sobretot els tipus 7 i 8), com quan tenim situacions amb baix gradient bàric en superfície (tipus 5 i 9).

Finalment, i ja en un enfocament marcadament meteorològic, cal tenir molt presents els resultats mostrats a la TAULA 1, valors molt orientatius sobre la utilitat real que puguin tenir a la nostra zona índexs d'inestabilitat, com ara la diferència de temperatura entre 850 i 500 hPa (TD850-500), el CAPE o el *Lifted Index* (LI). Alhora, les probabilitats d'ocurrència calculades i mostrades a la TAULA 5 fan rellevant també el

potencial de les eines de classificació de la circulació atmosfèrica per a desenvolupar productes de predicció estadística.

8. BIBLIOGRAFIA

- Baur F., Hess P., Nagel H. 1944. "Kalender der Großwetterlagen Europas 1881-1939". *Forschungsinstitut für langfristige Wettervorhersage*, Bad Homburg.
- Barry R. G., Perry A. H. 1973. *Synoptic Climatology, Methods and Applications*. Methuen i Co. Ltd. 555 p
- Barry R. G., Carleton A. M. 2001. *Synoptic and Dynamic Climatology*. Routledge.
- Barston A. G., Lizevey R. E. 1987. "Classification, seasonality and persistence of low-frequency atmospheric circulation patterns". *Monthly weather review*. 115: 1083-1126.
- Beck C. 2000. "Zirkulationsdynamische Variabilität im Bereich Nordatlantik-Europa seit 1780". *Würzburger Geographische Arbeiten* 95.
- Beck C., Jacobeit J, Jones P. D. 2007. "Frequency and within-type variations of large scale circulation types and their effects on low-frequency climate variability in Central Europe since 1780". *International Journal of Climatology* 27: 473-471.
- Beck C. i Philipp A. 2010. "Evaluation and comparison of circulation type classifications for the European domain". *Physics and Chemistry of the Earth* 35: 374-387.
- Bednorz E. 2008. "Synoptic reasons for heavy snowfalls in the Polish-German lowlands". *Theoretical and Applied Climatology* 92, 3-4: 133-140.
- Bejaran R. A., Camilloni I. A. 2003. "Objective method for classifying air masses: an application to the analysis of Buenos Aires, Argentina. Urban heat island intensity". *Theoretical and Applied Climatology*. 74: 93-103.
- Benestad, R. E., Hanssen-Bauer, I. And Chen, D. 2008. Empirical-Statistical Downscaling. *World Scientific Publishers*, 215pp.
- Birkeland K. W., Mock C. J., Shinker J. J. 2001. "Avalanche extremes and atmospheric circulation patterns". *Annals of Glaciology*. 32: 135-140.
- Bischoff S. A., Vargas W. M. 2003. "The 500 and 1000 hPa weather types circulations and their relationship with some extreme climatic conditions over southern South America". *International Journal of Climatology* 23: 541-556.
- Blair D. 1998. "The Kirchhofer technique of synoptic typing revisited". *International Journal of Climatology* 18: 1625-1635.
- Bogardi I., Matyasovszky I., Bardossy A., Duckstein L. 1993. "Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns". *Journal of Geophysical Research*. 98: 16653-16667.
- Bower D., McGregor G. R., Hannah D. M., Sheridan SC. 2007. "Development of a spatial synoptic classification scheme for western Europe". *International Journal of Climatology* 27: 2017-2040.

- Brinkmann W. A. R. 1999. "Within-type variability of 700 hPa winter circulation patterns over the Lake Superior basin". *International Journal of Climatology* 19:41–58
- Brinkmann W. A. R. 2000. "Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation". *International Journal of Climatology* 20: 839-852.
- Buell C. E. 1975. "The topography of the empirical orthogonal functions. Preprints fourth Conference on Probability and Statistics in Atmospheric Sciences, Tallahassee", *American Meteorological Society*: 188-193
- Buell C. E. 1979. "On the physical interpretation of empirical orthogonal functions. Preprints sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff", *American Meteorological Society*: 112-117
- Cattell R. B. 1966. "The scree test for the number of factors". *Multivariate Behavioral Research*. 1: 245-276
- Cahynová M., Huth, R. 2009. "Enhanced lifetime of atmospheric circulation types over Europe: fact or fiction?" *Tellus* 61A: 407-416.
- Cattell R. B. 1966. "The Scree test for the number of factors". *Multivariate Behavioral Research* 1: 245-76
- Cavazos T. 2000. "Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans". *Journal of Climate*. 13: 1718-1732.
- Chabot E. 2008. "Prévisibilité des régimes de temps à échelle saisonnière". Technical Report TR/CMGC/08/65, SUC au CERFACS, URA CERFACS/CNRS No1875, Tolosa, França.
- Clavero P., Raso J. M. 1979. "Catálogo de tipos sinópticos para un estudio climático del este de la península Ibérica y Baleares". A *Aportacions en homenatge al geògraf Salvador Llobet*. Departament de Geografia, Universitat de Barcelona: 55-62.
- Cohen S. 1983. "Classification of 500 mb height anomalies using obliquely rotated principal components". *Journal of Climate and Applied Meteorology* 22, 1975.
- Compagnucci R. H., Araneo D., Canziani P. O. 2001. "Principal sequence pattern analysis: a new approach to classifying the evolution of atmospheric systems". *International Journal of Climatology* 21: 197-217.
- Compagnucci R. H., Richman M. B. 2008. "Can principal component analysis provide atmospheric circulation or teleconnection patterns?" *International Journal of Climatology* 28: 703-726.
- Conway D., Wilby R. L., Jones P. D. 1996. "Precipitation and air flow indices over the British Isles". *Climate Research* 7: 169-183.
- Corte-Real J., Qian B., Xu H. 1998. "Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation".

International Journal of Climatology. 18: 619-635.

- Craddock J. M., Flood C. R. 1969. "Eigenvectors for representing the 500-mb geopotential surface over the Northern Hemisphere". *Quarterly Journal of the Royal Meteorological Society* 95: 576-593.
- Demuzere M., Trigo R. M., Vila-Guerau de Arellano J. 2009. "The impact of weather and atmospheric circulation on O₃ and PM₁₀ levels at a rural mid-latitude site". *Atmospheric Chemistry and Physics* 9: 2695-2714.
- Dodson R., Marks D. 1997. "Daily air temperature interpolated at high spatial resolution over a large mountainous region". *Climate Research* 8: 1-20.
- Ekström M., Jönsson P., Barring L. 2002. "Synoptic pressure patterns associated with major wind erosion events in southern Sweden. 1973-1991". *Climate Research* 23: 51-66.
- Ekström M., McTainsh G. H., Chappell A. 2004. "Australian dust storms: temporal trends and relationships with synoptic pressure distributions 1960-90". *International Journal of Climatology* 24: 1581-1599.
- El-Kadi AKA, Smithson P. A. 2000. "Probability assessment of pressure patterns over the British Isles". *International Journal of Climatology*. 20: 1351-1369.
- Dessouky T. M., Jenkinson A. F. 1975. "An objective daily catalogue of surface pressure, flow, and vorticity indices for Egypt and its use in monthly rainfall forecasting". *Meteorological Research Bulletin, Egypt* 11: 1-25.
- Esteban P., Soler X., Prohom M., Planchon O. 2002. "La distribución de la precipitación a través del índice NAO. El efecto del relieve a escala local: El Pirineo Oriental". Publicaciones de la Asociación Española de Climatología - AEC, Serie A, nº 3, pp. 25-34
- Esteban P. 2007. "La complexitat de la Meteorologia de muntanya". *Revista Catalana de Geografia* IV època, volum XII, 30: <http://www.rcg.cat/articles.php?id=86>
- Esteban P., Martí G., García C., Aran M., Gavalda J., Moner I. 2007. "Heavy Snowfalls and avalanche activity over eastern Pyrenees: a study of two extreme cases. Proceedings Alpine Snow Workshop", *Forschungsbericht Munich*. Alemanya, 5-6/10/2006: 44-48.
- Esteban P., Beck C., Philipp A. 2010. "Classification of circulation type sequences applied to snow avalanches over the eastern Pyrenees (Andorra and Catalonia)". *Geophysical Research Abstracts* 12, EGU2010-11868.
- Esteban P., Gallego N. 2011. "Els incendis forestals a Andorra 1962-2010: Climatologia sinòptica i anàlisi estadística". *Revista del CENMA* 6: 14-21.
- Everitt B. S. 1980. "Cluster Analysis. Second Edition". Heineman Educational Books Ltd.
- Fernández A. J., Martín-Vide J., Díaz Y. A., Mestre A. 2003. "Aplicación de los procesos sinópticos objetivos a la Península Ibérica en otoño". *Investigaciones Geográficas* 31: 37-65.

- Frakes B., Yarnal B. 1997. "A procedure for blending manual and correlation-based synoptic classifications". *International Journal of Climatology* 17: 1381–1396.
- Friedman H. F., Meulman J. M. 2004. "Clustering Objects on Subsets of Attributes". *Journal of the Royal Statistical Society. Series B*, vol. 66, 4: 815-849.
- Font Tullot, I. 1983. "Climatología de España y Portugal". Instituto Nacional de Meteorología.
- Galambosi A, Duckstein L, Bogardi I. 1996. "Evaluation and análisis of daily atmospheric circulation patterns of the 500 hPa pressure field over the southwestern USA". *Atmospheric Research* 40: 49-76.
- García-Sellés C., Martí G., Oller P., Moner I., Gavaldà J., Martínez P., Peña J. C. 2009. "Major avalanche occurrence at regional scale and related atmospheric circulation patterns in the Eastern Pyrenees", *Cold Regions Science and Technology*, 59: 106-118.
- García-Sellés C., Peña J. C., Marti G., Oller P., Martínez P. 2010. "WeMOI and NAOi influence on major avalanche activity in the eastern Pyrenees". *Cold Regions Science and Technology*, 64: 137-145.
- Gerstengarbe F. W., Werner P. C., Rüge U. 1999. "Katalog der Grosswetterlagen Europas nach Paul Hess und Helmuth Brezowsky 1881-1998". *Deutscher Wetterdienst*, Offenbach am Main.
- Gerstengarbe F. W., Werner P. C. 2007. "Der rezente Klimawandel. In: Der Klimawandel-Einblicke, Rückblicke und Ausblicke. Hrsg.: Endlicher W., Gerstengarbe F. W., Potsdam-Institut für Klimafolgenforschung und Humboldt-Universität zu Berlin: 34-43.
- Goodess C. M., Palutikov J. P. 1998. "Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling". *International Journal of Climatology*. 10: 1051-1083.
- Goodess C. M. 2001. "European circulation indices for use in tropospheric ozone studies". Final report to University of Berne. Tropospheric Ozone and its Precursors - TROTREP project.
- Gutiérrez J. M., Cano R., Cofiño A. S., Sordo C. M. 2004. "Redes probabilísticas y Neuronales en las Ciencias Atmosféricas". Ministerio de Medio Ambiente. 279 pp.
- Hair J., Anderson R., Tatham R., Black W. 2000. *Análisis multivariante*. 5ª Edición. Prentice Hall. 799 pp.
- Hannachi A., Jolliffe I. T., Stephenson D. B., Trendafilov N. 2006. "In search of simple structures in climate: symplifying EOFs". *International Journal of Climatology* 26: 7-28.
- Hannachi A., Jolliffe I. T., Stephenson D. B. 2007. "Empirical orthogonal functions and related techniques in atmospheric science: A review". *International Journal of Climatology* 27: 1119-1152.

- Hartigan J. A. 1975. "*Clustering algorithms*". Wiley.
- Haby, J. 2003. "The 540 line and precipitation type - Habyhints 97". www.theweatherprediction.com
- Hess P., Brezowsky H. 1952. "Katalog der Grosswetterlagen Europas". *Berichte des Deutscher Wetterdienstes*: 15 (113).
- Hewitson B. C., Crane R. G. 2002. "Self-organizing maps: applications to synoptic climatology". *Climate Research* 22: 13-26.
- Horel J. D. 1981. "A rotated principal component analysis of the interannual variability of the northern hemisphere 500 mb height field". *Monthly Weather Review*. 109: 2080-2092.
- Huth R. 1996. "An intercomparison of computer-assisted circulation classification methods". *International Journal of Climatology*. 16: 893-922.
- Huth R. 2000. "A circulation classification scheme applicable in GCM studies". *Theoretical and Applied Climatology*. 67: 1-18.
- Huth R., Nemesová I., Klimperová N. 1993. "Weather categorization based on the average linkage *clustering* technique: an application to european mid-latitudes". *International Journal of Climatology* 13: 817-835.
- Huth R., Beck C., Philipp A., Demuzere M., Ustrnul Z., Cahynová M., Kyselý J., Tveito O. E. 2008. "Classifications of atmospheric circulation patterns: recent advances and applications". *Annals of the New York Academy of Science* 1146: 105-152.
- Huth R., Pokorná L., Bochníček J., Hejda P. 2006. "Solar cycle effects on modes of low-frequency circulation variability". *Journal of Geophysical Research* D22107, doi:10.1029/2005JD006813
- IPCC. 2007. "Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Core Writing Team". Pachauri, R. K, Reisinger A. eds. IPCC, Geneva, Switzerland, 104 pp.
- Jacobeit J., Wanner H., Luterbacher J., Beck C., Philipp A., Sturm K. 2003. "Atmospheric circulation variability in the North-Atlantic-European area since the mid-seventeenth century". *Climate Dynamics* 20: 341-352
- James P. M. 2006. "An assessment of European synoptic variability in Hadley Centre Global Environmental models based on an objective classification of weather regimes". *Climate Dynamics* 27: 215-231
- James P. M. 2007. "An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe". *Theoretical and Applied Climatology* 88: 17-42
- Jenkinson A. F., Collison FP. 1977. "An initial climatology of gales over North Sea". *Synoptic Climatology Branch Memorandum* n. 62, Meteorological Office, Bracknell.
- Jolliffe I. T. 1986. "Principal Component Analysis". Springer-Verlag.

- Jolliffe I, T. 1987. "Rotation of principal components: somme comments". *Journal of climatology* 7: 507-510.
- Jolliffe I. T. 1990. "Principal component analysis: a beginners's guide. I Introduction and applications". *Weather* 45:375-382.
- Jolliffe I. T. 1993. "Principal component analysis: a beginners's guide. II Pitfalls and Extensions". *Weather* 48: 246-253.
- Jolliffe I. T, Uddin M., Vines S. K. 2002. "Simplified EOFs-three alternatives to retain". *Climate Research* 20: 271-279.
- Jolliffe I. T., Philipp A. 2010. "Some recent ideas in cluster analysis". *Physics and Chemistry of the Earth* 35: 309-315.
- Jones P. D., Hulme M., Briffa K. R. 1993. "A comparison of Lamb circulation types with an objective classification scheme". *International Journal of Climatology*. 13: 655-663.
- Kalkstein S. L., Tan G., Skindlov J. A. 1987. "An evaluation of three clustering procedures for use in synoptic climatological classification". *Journal of Climate and Applied Meteorology*. 25: 717-30.
- Karl T. R., Koscielny A. J., Diaz H. F. 1982. "Potential errors in the application of principal components. Eigenvector. Analysis to geophysical data". *Journal of Applied Meteorology*. 21: 1183-1186.
- Kaufmann R. K., Snell S. E., Gopal S., Dezzani R. 1999. "The significance of synoptic patterns identified by the Kirchhofer Technique: a Monte Carlo approach". *International Journal of Climatology* 19: 619–626.
- Kelly P. M., Jones P. D., Briffa KR. 1997. "Classifying the winds and weather. In *Climates of the British Isles*", Hulme M., Barrow E. (eds). Routledge: London.
- Key J., Crane R. G. 1986. "A comparison of synoptic classification schemes based on objective procedures". *Journal of Climatology*. 6: 375-388.
- Kidson J. W. 1994. "An automated procedure for the identification of synoptic types applied to the New Zealand region". *International Journal of Climatology*. 14: 711-721.
- Kilsby C. G, Cowpertwait P. S. P., O'Connell E., Jones P. D. 1998. "Predicting rainfall statistics in England and Wales using atmospheric circulation variables". *International Journal of Climatology*. 18:523-539.
- Kirchhofer W. 1973. "Classification of European 500 mb patterns". *Arbeitsbericht der Schweizerischen Meteorologischen Zentralanstah*. 43: 205-207.
- Kline P. 1994. "An easy guide to factor analysis". Routledge.
- Kostopoulou E. 2003. "The relationships between atmospheric circulation patterns and surface climatic elements in the eastern Mediterranean". PhD. Thesis, University of East Anglia, Norwich.
- Kutzbach J. E. 1967. "Empirical eigenvectors of sea-level pressure, surface

temperature, and precipitation complexes over North America". *Journal of Applied Meteorology* 6: 791.

- Lamb H. H. 1972. "British Isles weather types and a register of daily sequence of circulation patterns, 1861-1971". *Geophysical Memoir* 116, HMSO, London.
- Lana X., Mills G. F. 1994. "Minimum sample size for synoptic weather type classification. Application to winter period data recorded on the catalan coast (north-east Spain)". *International Journal of Climatology*. 14: 1051-1060.
- Legates D. R. 1991. "The effect of domain shape on principal component analyses". *International Journal of Climatology* 11: 135-146
- Legates D. R. 1993. "The effect of domain shape on principal component analyses. A reply". *International Journal of Climatology* 13: 219-228.
- Linderson M. 2001. "Objective classification of atmospheric circulation over southern Scandinavia". *International Journal of Climatology*. 21: 155-169.
- López C, Lizares P. 2000. "Anàlisi multivariable de dades estadístiques". Col·lecció Materials 93. Universitat Autònoma de Barcelona.
- López-Bustins A., Esteban P., Labitzke K., Langematz U. 2007. "The Role of the Stratosphere in Iberian Peninsula Rainfall: a Preliminary Approach in February". *Journal of Atmospheric and Solar Terrestrial Physics* 69 (12): 1471-1484.
- Lund I. A. 1963. Map-pattern classification by statistical methods. *Journal of Applied Meteorology*. 2: 56-65.
- Luterbacher J., Xoplaki E, Dietrich D., Rickli R., Jacobeit J., Beck C., Gyalistras D., Schmutzet C., Wanner H. 2002. *Climate Dynamics*. 18: 545-561.
- Maheras P., Patrikas I., Karacostas Th., Anagnostopoulou Chr. 2000. "Automatic classification of circulation types in Greece: methodology, description, frequency, variability and trend analysis". *Theoretical and Applied Climatology*. 67: 205-223.
- Maheras P., Xoplaki E., Davies T., Martin-Vide J., Barriendos M., Alcoforado M. J. 1999. "Warm and cold monthly anomalies across the Mediterranean basin and their relationship with circulation 1860-1990". *International Journal of Climatology*. 19: 1697-1715.
- Mann E. M., Bradley R. S., Hughes M. K. 1998. "Global-scale temperature patterns and climate forcing over the past six centuries". *Nature*. 392: 779-787
- Martín-Herreros J. M. 2001. "Caracterización de la circulación atmosférica en un entorno de la Península Ibérica mediante cuatro índices reticulares". Nota técnica nº 8 del Servicio de Variabilidad y Predicción del Clima (INM).
- Martín-Vide J. 2001. "Limitations of an objective weather-typing system for the Iberian peninsula". *Weather* 56: 248-250.
- Martín-Vide J. 1990. "Mapas del Tiempo: Fundamentos, Interpretación e imágenes de satélite". Oikos. Tau, 170 p.
- Martín-Vide J., López-Bustins J. A. 2006. "The Western Mediterranean

- Oscillation and rainfall in the Iberian Peninsula". *International Journal of Climatology* 26: 1455-1475.
- Mestas-Núñez A. M. 2000. "Orthogonality properties of rotated empirical modes". *International Journal of Climatology*. 20: 1509-1516.
 - North G. R., Bell T. L., Cahalan R. F., Moeng F. J. 1982. "Sampling errors in the estimation of empirical orthogonal functions". *Monthly Weather Review*. 110: 699-706.
 - O'Lenic E. A, Lizevey R. E. 1988. "Practical considerations in the use of rotated principal component analysis. RPCA. In diagnostic studies of upper-air height fields". *Monthly Weather Review*. 116: 1682-1689.
 - Overland J. E., Preisendorfer. 1982. "A significance test for principal components applied to a cyclone climatology". *Monthly Weather Review*. 110: 1-4
 - Pédelaborde P. 1957. *Le Climat de la Bassin Parisien*. Ed. Génin.
 - Pédelaborde P. 1959. *Circulation atmosphérique sur les Pyrénées françaises*. Centre de Documentation Universitaire.
 - Peña J. C., Aran M., Cunillera J., Amaro J. 2011. "Atmospheric circulation patterns associated with strong wind events in Catalonia". *Natural Hazards and Earth System Sciences* 11: 145-155.
 - Peñarrocha D., Estrela M. J., Millán M. 2002. "Classification of daily rainfall patterns in a mediterranean area with extreme intensity levels: the Valencia region". *International Journal of Climatology* 22: 677-695.
 - Philipp A., Della-Marta P. M., Jacobeit J., Fereday D. R., Jones P. D., Moberg A., Wanner H. 2007. "Long term variability of daily North Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering". *Journal of Climate* 20: 4065-4095.
 - Philipp 2009. "Comparison of principal component and cluster analysis for classifying circulation patterns sequences for the European domain". *Theoretical and Applied Climatology* Volume 96, 1-2: 31-41.
 - Phillip A. and 18 authors. 2010. "Cost733cat—A database of weather and circulation type classifications". *Physics and Chemistry of the Earth* Volume 35: 9-12.
 - Preisendorfer R. W., Zwiers F. W., Barnett T. P. 1981. "Foundations of Principal Component Selection Rules". SIO Reerence Series 81-4, Scripps Institution of Oceanography, 192pp
 - Prohom M. J., Esteban P., Jones P. D., Martín-Vide J. 2004. "Surface atmospheric circulation over Europe following major tropical volcanic eruptions, 1780-1995". *Volcanism and the Earth's Atmosphere*. Robock, A. and Oppenheimer, C. Eds. Geophysical monograph series. 139: 273-281.
 - Rasilla D. F. 2003. "Aplicación de un método de clasificación sinóptica a la Península Ibérica". *Investigaciones Geográficas*. 30: 27-45.

- Raso Nadal J. M. 1999. "El clima d'Andorra". *Monogràfics de Geografia*. Govern d'Andorra, 139 pp.
- Reusch, D. B. 2010. "Nonlinear climatology and paleoclimatology: Capturing patterns of variability and change with Self-Organizing Maps", *Physics and Chemistry of the Earth*, 35: 329-340
- Richman M. B. 1981. "Obliquely Rotated Principal Components: An Improved Meteorological Map Typing Technique?" *Journal of Applied Meteorology*. 20: 1145-1159.
- Richman M. B. 1986. "Rotation of Principal Components". *Journal of Climatology*. 6: 293-335.
- Richman M. B. 1987. "Rotation of principal components: a reply". *Journal of Climatology*. 7: 511-520.
- Richman M. B. 1993. Comments on: "The effect of domain shape on principal components analyses". *International Journal of Climatology*. 13:203-218.
- Romero R., Ramis C., Guijarro J. A. 1999a. "Daily rainfall patterns in the Spanish Mediterranean area: an objective classification". *International Journal of Climatology*. 19: 95-112.
- Romero R., Sumner G., Ramis C., Genovés A. 1999b. "A classification of the atmospheric circulation patterns producing significant daily rainfall in the spanish mediterranean area". *International Journal of Climatology*. 19: 765-785.
- Rogers J. C. 1981. "Spatial variability of seasonal sea level pressure and 500mb height anomalies". *Monthly Weather Review* 109: 2093-2106.
- Rogers J. C., Van Loon, H. 1982. "Spatial variability of sea level pressures and 500 mb height anomalies over the Southern Hemisphere". *Monthly Weather Review* 110: 1375-1392.
- Serra, C. 1994. "Análisis en componentes principales y algoritmos de clasificación automática: aplicación al estudio del comportamiento de la precipitación en Catalunya y a la obtención local de tipos de tiempo". Tesis doctoral, Universitat de Barcelona.
- Serra C., Fernández Mills G., Periago M. C., Lana X. 1996. "Winter and Autumn Daily Precipitation Patterns in Catalonia, Spain". *Theoretical and Applied Climatology* 54: 175-186.
- Spellman G. 2000. "The application of an objective weather-typing system to the Iberian peninsula". *Weather*. 55: 375-385.
- Tait A. B, Fitzharris B. B. 1998. "Relationship between New Zealand rainfall and south-west Pacific pressure patterns". *International Journal of Climatology*. 18: 407-424.
- Trapero L., Esteban P. 2010. "Modelisation of northerly snow episodes over Andorra (Pyrenees) using WRF". *European Meteorological Society (EMS)-Annual Meeting*.
- Trigo R. M., DaCamara C. C. 2000. "Circulation weather types and their

influence on the precipitation regime in Portugal". *International Journal of Climatology*. 20: 1559-1581.

- Von Storch H., Hannoschöck G. 1985. "Statistical aspects of estimated principal vectors. EOFs. based on small sample sizes". *Journal of Climate and Applied Meteorology* 24: 716-724.
- Von Storch H., Zwiers F. W. 1999. *Statistical analysis in climate research*. Cambridge University press.
- Wallace J. M., Gutzler D. S. 1981. "Teleconnections in the geopotential height field during the Northern Hemisphere winter". *Monthly Weather Review* 109: 784-812.
- White D., Richman M., Yarnal B. 1991. "Climate regionalization and rotation of principal components". *International Journal of Climatology*. 11: 1-25.
- Wilby R. L., Charles S. P., Zorita E., Timbal E., Whetton P., Mearns LO. 2004. "Guidelines for use of climate scenarios developed from statistical downscaling methods". Supporting material of the Intergovernmental Panel on Climate Change (Task group on Data and Scenario Support for Impacts and Climate Analysis).
- Wilks, D. S. 2006. "Statistical Methods in the Atmospheric Sciences". Second Edition. *International Geophysics Series*. Elsevier-Academic Press.
- Yarnal B., Comrie A. C., Frakes B., Brown D. P. 2001. "Developments and prospects in synoptic climatology". *International Journal of Climatology* 21: 1923-1950.
- Yarnal B., White D. A. 1987. "Subjectivity in a computer-assisted climatology I: classification results". *Journal of Climatology* 7:119-128.
- Yarnal B. 1984. "Synoptic-scale atmospheric circulation over British Columbia in relation to the mass balance of Sentinel Glacier". *Annals of the Association of American Geographers* 74: 375-392.
- Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press.
- Zwiers F. W., Von Storch H. 2004. "On the role of statistics in climate research". *International Journal of Climatology*. 24: 665-680.

Llocs web:

- cost733.met.no/index.htm
- geo23.geo.uni-augsburg.de/cost733class-1.1
- www.cpc.ncep.noaa.gov/
- www.cru.uea.ac.uk/cru/people/lamb/
- www.cru.uea.ac.uk/projects/emulate/
- www.ecmwf.int/research/era/do/get/Reanalysis_ECMWF
- www.emc.ncep.noaa.gov/gmb/ssaha/indices/nao_load.gif

- www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml
- www.retroist.com/2010/03/08/retroist-access-monday-ram-%E2%80%93-the-ibm-704/