



Predicción probabilista

27

DOI: [10.31978/014-18-009-X.27](https://doi.org/10.31978/014-18-009-X.27)

CARLOS SANTOS BURGUETE

Centro Nacional de Predicción (CNP), Agencia Estatal de Meteorología (AEMET)

ANEXO: SERGI GONZÁLEZ HERRERO

Delegación territorial (DT) en Cataluña, AEMET

ANEXO: ÁLVARO SUBÍAS DÍAZ-BLANCO Y ALEJANDRO ROA ALONSO

Área de Técnicas y Aplicaciones de Predicción (ATAP), AEMET

El mundo es tan impredecible. Cosas suceden de repente, inesperadamente. Queremos sentirnos en control de nuestra propia existencia. En algunas cosas lo estamos, en otras no lo estamos. Nos regimos por las fuerzas del azar y la casualidad.

frase atribuida – PAUL AUSTER

La *probabilidad* es un modelo rico y razonable para describir y comprender muchos aspectos del mundo físico. Los Sistemas de Predicción por Conjuntos (SPC) proporcionan información útil y honesta cuando los miramos como sistemas de predicción intrínsecamente probabilistas. ¿Cómo se interpreta adecuadamente esta información para hacer predicciones, ya sean generales o específicas? Veremos en este capítulo que los SPC deben usarse cuidadosamente, porque sin previo entrenamiento pueden chocar con el «sentido común». Bien usados, nos brindan información cuantitativa, detallada y explícita sobre la incertidumbre de la atmósfera y, por ende, de su predecibilidad.

Palabras clave: sistemas de predicción por conjuntos aplicados a la predicción de corto y medio plazo, predicción probabilista, agrupamiento, dispersión, mapas de probabilidad, EPSgramas, predicción de riesgos, fenómenos adversos, fenómenos extremos, predecibilidad dependiente del flujo.

Imagen parte superior: *cumulonimbus mamma*. Ciudad Universitaria, Madrid, hacia el nordeste, 22 de mayo de 2007, a las 20:13. Fotografía de JOSÉ ANTONIO QUIRANTES CALVO.

Advertencia. En este capítulo se presentan numerosos productos para la predicción, aunque solamente unos pocos están disponibles para el público general en los sitios web correspondientes. En los sitios web de los servicios meteorológicos no se publica toda la información debido a políticas de datos. Además, la información probabilista es menos susceptible de ser publicada, pues todavía falta costumbre para ello y a menudo se piensa que el público no está preparado, lo cual es un eterno tema interno de discusión. En este libro se alternarán productos, la mayoría restringidos, del sitio web del European Centre for Medium-range Weather Forecasts -Centro Europeo de Predicción a Plazo Medio- (ECMWF) [17] y, de la Agencia Estatal de Meteorología (AEMET), del sitio web y de su Intranet. Esperamos fervientemente que las políticas de datos y productos tiendan hacia la apertura en un futuro cercano.

27.1 Incertidumbres propias en predicción del tiempo

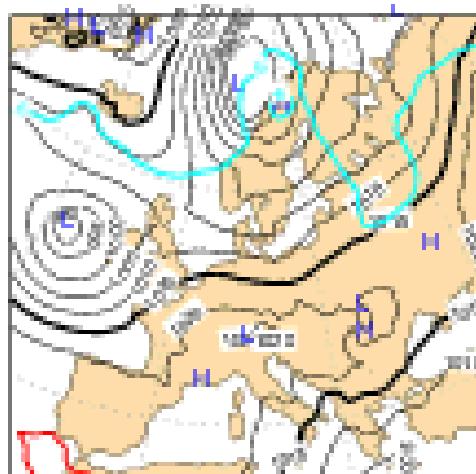
Este capítulo presenta los principios de la predicción probabilista orientada principalmente a la predicción del tiempo atmosférico en el corto y medio plazo, desde el D+1 hasta, como mucho, un D+10 o un D+15, rangos que suelen englobarse en lo que se denomina comúnmente *predicción del tiempo*. La predicción en escalas superiores como la mensual, la estacional, la anual o, más allá, las proyecciones del clima, son ámbitos desarrollados en otros capítulos de este libro (cap. 28 en la página 447 y cap. 29 en la página 477). Como se describía con cierto detalle en el capítulo sobre predecibilidad (cap. 12), cada rango de predicción tiene unas incertidumbres propias con diferentes pesos en diferentes aspectos físicos. Así, la predicción del tiempo es más un problema de predecibilidad de primera clase, es decir, con mucho peso en las condiciones iniciales mientras que, en el otro extremo de la escala, las proyecciones del clima son más un problema de predecibilidad de segunda clase, es decir, con mucho peso en las condiciones de contorno o forzamientos (ver sec. 12.2.3 en la página 158). Las escalas intermedias están a caballo entre ambos problemas: las predicciones mensuales, estacionales, anuales o decadales tienen sus pesos más repartidos entre ambos aspectos.

27.2 Productos crudos

Como hemos ido viendo en capítulos anteriores, los SPC se componen de miembros que individualmente son *predicciones deterministas* y, en principio, nada impide disponer de esa pléthora de información determinista [6]. Esta información tiene utilidades diversas: puede ayudar a los predictores «clásicos» a entender los SPC, puede proporcionar información complementaria a la propiamente probabilista, etc. Dado un punto de rejilla del SPC en el que tenemos N predicciones (correspondientes a un SPC de N miembros), llamamos *productos crudos* a aquellos en los que se muestran tan sólo esos N valores directamente. Tenemos tres ejemplos básicos de productos crudos: *sellos*, *penachos* y *espaguetis*.

27.2.1 Sellos

En los *mapas meteorológicos clásicos* se muestran una o más variables meteorológicas interesantes para el diagnóstico o la predicción del tiempo. Uno de los más usuales es el famoso «mapa de isobaras», que representa la presión reducida al nivel del mar (*mean sea level pressure* o MSLP). Esas isobaras muestran estructuras interesantes, como los anticlones, las borrascas, etc. A menudo se acompaña con la temperatura; pero no con la temperatura en superficie, que es una variable muy «orográfica» (sigue fielmente la orografía) y por tanto no aporta una información tan rica. Se suele acompañar con la temperatura en el nivel de 850 hPa, que es un campo algo más suave e informa mejor de la distribución térmica en niveles bajos.



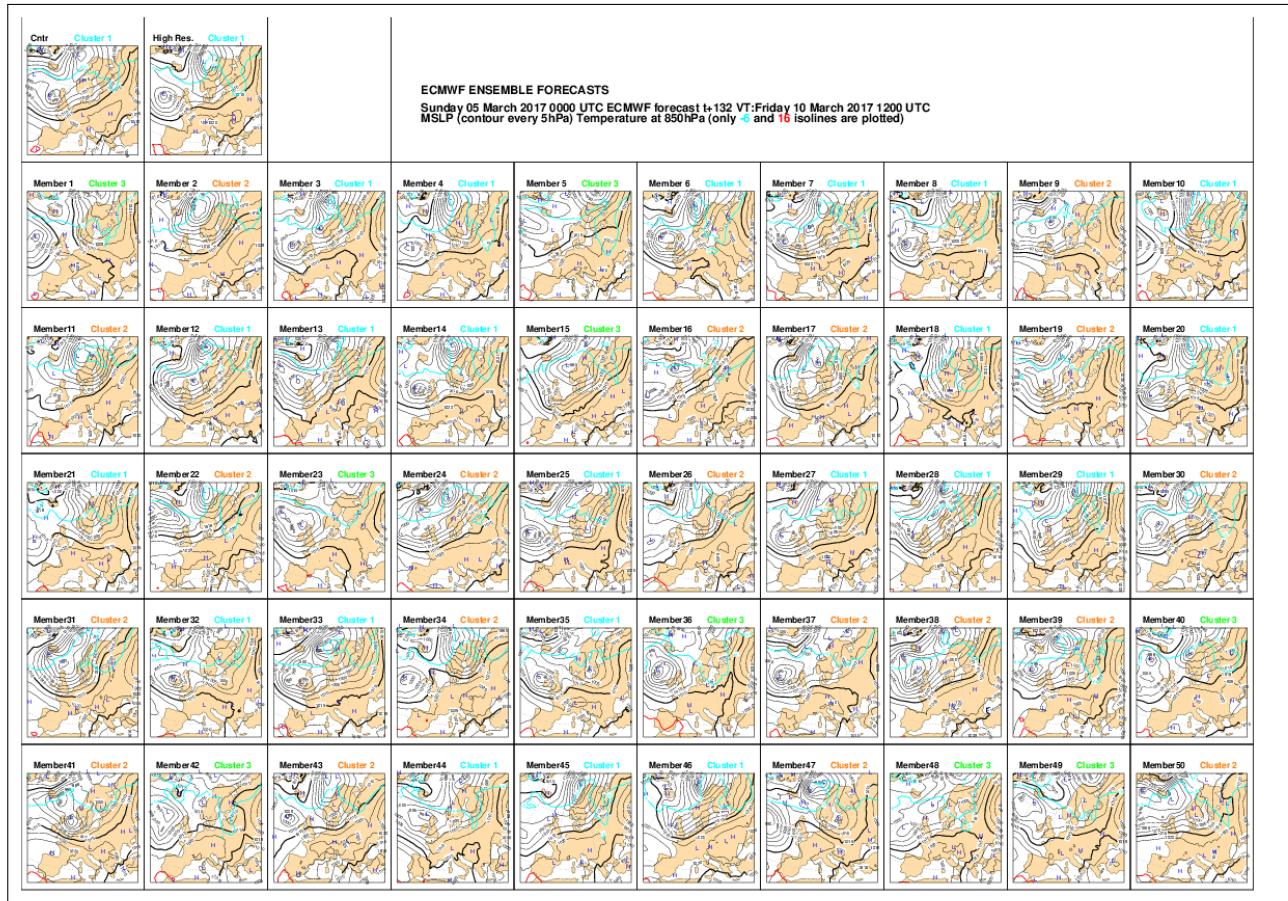


Figura 27.2: Sellos de «presión» y «temperatura» para un D+5 (ver texto). ECMWF 2017.

En la Figura 27.1 en la página anterior tenemos un ejemplo: mapa de presión y temperatura del modelo ECHRES (sec. 19.2 en la página 291) del ECMWF (cap. 19 en la página 289). Las isobaras son líneas negras, pintadas cada 4 hPa, y sólo se pintan dos isotermas para no «embarullar», en azul la de 16 °C y en rojo la de -6 °C. Este mapa combinado es una prime-rísima aproximación para diagnosticar o predecir la situación atmosférica basándonos en los niveles bajos. La presión y la temperatura son salidas, usualmente, de un modelo determinista. Si disponemos de un SPC que conste de N miembros, entonces podríamos disponer de N mapas distintos de presión y temperatura similares a éste.

Los *sellos* son mosaicos ordenados de mapas meteorológicos de todos los miembros de un SPC (ver Figura 27.2). Estos sellos incorporan el *miembro de control* (si lo hay, ver sec. 13.5 en la página 174), el *modelo determinista de mayor resolución* (si lo hay y si procede) y el resto de miembros (llamados miembros *perturbados* en algunos SPC). Estos sellos proporcionan un abanico de escenarios atmosféricos posibles, por lo que la dificultad principal es que el predictor/usuario tiene que tratar con una enorme cantidad de informa-

ción. El ECMWF muestra en su sitio web sellos de «presión» y «temperatura» de su ECENS junto con su ECHRES.

La Figura 27.2 muestra un ejemplo de estos sellos, con la presión reducida al nivel medio del mar (MSLP) y la temperatura en 850 hPa, para un D+5 (previsión para un «día más cinco»).

Esta imagen, en particular, está realizada con la pasada 2017030500 UTC, alcance T+132, es decir, válida para 2017031012 UTC (ver sec. 10.2 en la página 142). Arriba a la izquierda tenemos el miembro de control del ECENS (aquel que se ha corrido con condiciones iniciales sin perturbar, y por tanto «parecido» en principio al ECHRES). A su lado, el ECHRES, determinista de mayor resolución.

El resto despliega, en 5 filas y 10 columnas, los 50 miembros perturbados del ECENS (sec. 13.5 en la página 174). Para simplificar esta avalancha de información se pintan sólo dos isotermas en cada sello (como decíamos, la de -6 °C y la de 16 °C). Si se dispone de información complementaria de agrupamiento (*clustering*, ver 27.4.1 en la página 406), entonces

puede resultar muy interesante saber en qué grupo (cluster) ha caído cada miembro, y esa información consta, en principio, en cada sello.

La similitud de estos $1+1+50=52$ escenarios atmosféricos que brindan los sellos es una primera indicación sobre la dispersión del SPC. Cuanto más se parezcan, menos dispersión presentará el SPC y podremos atribuir a la atmósfera menos incertidumbre y, por ende, más predecibilidad. Si, por el contrario, los diferentes escenarios muestran diferencias notables y/o numerosas, entonces el SPC mostrará más dispersión y podremos atribuir a la atmósfera más incertidumbre y, por ende, menos predecibilidad.

27.2.2 Espaguetis

Resulta imposible superponer los mapas meteorológicos de todos los miembros de un SPC en una sola imagen: resultaría absolutamente ilegible. Para poder representar muchas predicciones (muchos miembros) en el mismo mapa, existe un truco: puede hacerse seleccionando solamente una misma isolínea de cada miembro, y pintando esa isolínea, de forma que el mapa tendrá N versiones de esa isolínea. Así se podrá observar, a grandes rasgos, el comportamiento de todos los miembros sobre una zona determinada. Se suelen dibujar mapas de espaguetis de campos dinámicos (e. g. altitud geopotencial) acompañados de otros campos.

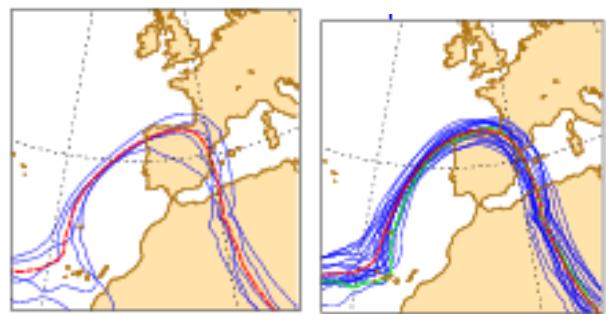


Figura 27.3: Izquierda: grupo con unos pocos espaguetis, en azul, y control, en rojo. Derecha: grupo con muchos espaguetis y control. AEMET 2017.

Puede dibujarse en un color diferente el miembro de control (si lo hay), y dibujarse, también en otro color, el modelo determinista de mayor resolución (si lo hay). En la Figura 27.3 podemos observar dos ejemplos de espaguetis seleccionando la isolínea de 5820 mgp de altitud geopotencial; a la izquierda de la figura, unos pocos miembros en azul y control en rojo; a la derecha de la figura, muchos miembros en azul, control en rojo y el modelo determinista de mayor resolución en verde. Igual que los penachos (sección 27.8.2 en la página 432), estos mapas proporcionan información sobre la dispersión del SPC, es decir, la incertidumbre y, por ende, la predecibilidad atmosférica. Si de entre todos los espaguetis seleccionamos unos pocos que se parecen y que por lo tanto representan configuraciones sinópticas parecidas, observamos que aunque son parecidos, todavía presentan ligeras diferencias. En la Figura 27.4 se muestra un ejemplo de mapas de espaguetis típicamente utilizados en AEMET por los predictores.

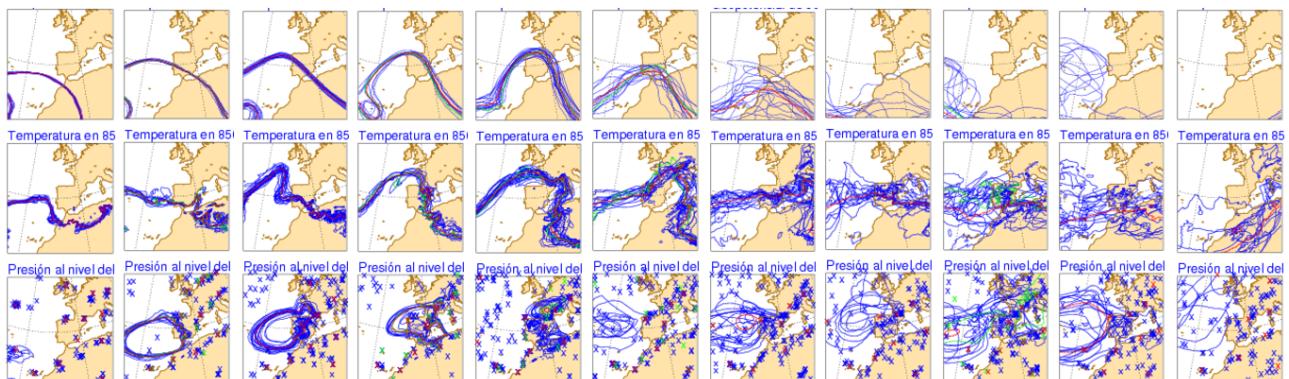


Figura 27.4: Montaje de 11 días de espaguetis. D+0 a la izquierda y sucesivamente. Al tratarse de un montaje específico para la figura, los títulos no pueden leerse completos. AEMET 2017.

En tres filas se muestran, respectivamente, la altitud geopotencial en 500 hPa (solamente la isohipsa de 5820 mgp para cada miembro, por tanto se muestran 52 isohipsas), temperatura en 850 hPa (solamente la isoterma de 10 °C), y presión reducida al nivel del mar (solamente la isobara 1028 hPa). En esta figura, por simplicidad, hemos condensado la información usual, mostrando 11 días en 11 columnas: desde D+1 uno por uno hasta D+10 y después D+15. Puede observarse muy bien que las diferentes isolíneas están, al principio (columnas de la izquierda), muy juntas y son muy similares. A medida que el plazo de predicción aumenta y nos alejamos en el tiempo (columnas más hacia la derecha), las isolíneas van presentando cada vez más diferencias y más notables. Llega un plazo en el que hay tal «barullo» que parece un plato de espaguetis. Puede observarse, así mismo, que la altitud geopotencial en 500 hPa, al ser un campo más «dinámico», va desarrollando diferencias menos ruidosas que la temperatura en 850 hPa y ésta, a su vez, presenta también menos ruido que la presión reducida al nivel del mar. Las diferencias aportan información, detallada *espacialmente*, sobre la *dispersión* y su naturaleza meteorológica. Si, por ejemplo, la zona de diferencias es una vaguada de geopotencial con el eje posicionado en diferentes lugares, sabemos que la incertidumbre va asociada, en principio, a la zona de más inestabilidad baroclinica, etc. Estos conceptos se desarrollan con más detalle en el apartado [27.12 en la página 413](#).

Todos los productos crudos comparten la misma desventaja: una dificultad inherente para manejar la enorme cantidad de información que brindan. Es preciso sintetizarla e interpretarla por otros caminos. Esos caminos son los productos probabilistas.

27.2.3 Penachos

Se desarrollan más adelante en este mismo capítulo pues, aunque encajarían aquí por categoría, conceptual y didácticamente resultan más útiles en otro apartado (sec. [27.8.2 en la página 432](#)).

27.3 El papel de la resolución

La resolución horizontal de los modelos ha sido siempre una cuestión crucial a la hora de interpretarlos apropiadamente, conocer sus fortalezas y debilidades y sacarles el mejor fruto que pueden dar. Cuanto

más fina la resolución horizontal, más fina es la dinámica del modelo, más realistas su parametrizaciones (incluso hay procesos que dejan de requerir parametrización pues se resuelven a nivel de rejilla del modelo), más realista su orografía, etc. Siempre que un modelo sufre una mejora en resolución horizontal (que suele ir acompañada de otras mejoras) esperamos que «resuelva» mejor los fenómenos meteorológicos, aunque también nos dice la experiencia que puede haber otros problemas. A menudo, un aumento de resolución requiere una temporada de prueba, pues puede ir acompañado de representación ruidosa de algunos fenómenos en particular, como por ejemplo rachas de viento a sotavento de barreras montañosas, cuya intensidad puede exagerar el modelo notablemente en comparación con la realidad. Ahí está la mano del predictor para evaluar el realismo de estas representaciones. En cualquier caso, un mismo modelo con dos resoluciones horizontales diferentes puede mostrar resultados con mayores o menores diferencias según la situación meteorológica. Cuando, independientemente de la resolución, el modelo muestra resultados similares, decimos que la resolución horizontal no juega un papel crítico. Por el contrario, cuando hay diferencias notables en el resultado dependiendo de la resolución horizontal, es más claro que ésta desempeña un papel crucial: decimos entonces que en esos casos, *cuando hay diferencias debido únicamente a este factor, la resolución horizontal juega un papel crítico*.

27.3.1 Comparación control-determinista

En aquellos SPC donde haya un miembro de control y, además, se disponga de un modelo determinista de mayor resolución horizontal que el SPC, puede realizarse un «chequeo» interesante. Un ejemplo cercano es el SPC del ECMWF (ECENS). Se compara el miembro de control del SPC con el citado modelo determinista de mayor resolución ECHRES (el modelo determinista era la única herramienta en predicción hasta hace relativamente poco, y en predicción clásica no había modelos atmosféricos). Dado que el miembro de control tiene menor resolución horizontal (18 km en 2017) que el determinista (9 km en 2017), pero como modelo es bastante equivalente (los dos se han corrido con el análisis sin perturbar), entonces la comparación entre ambos revela el *papel que juega la resolución horizontal*.

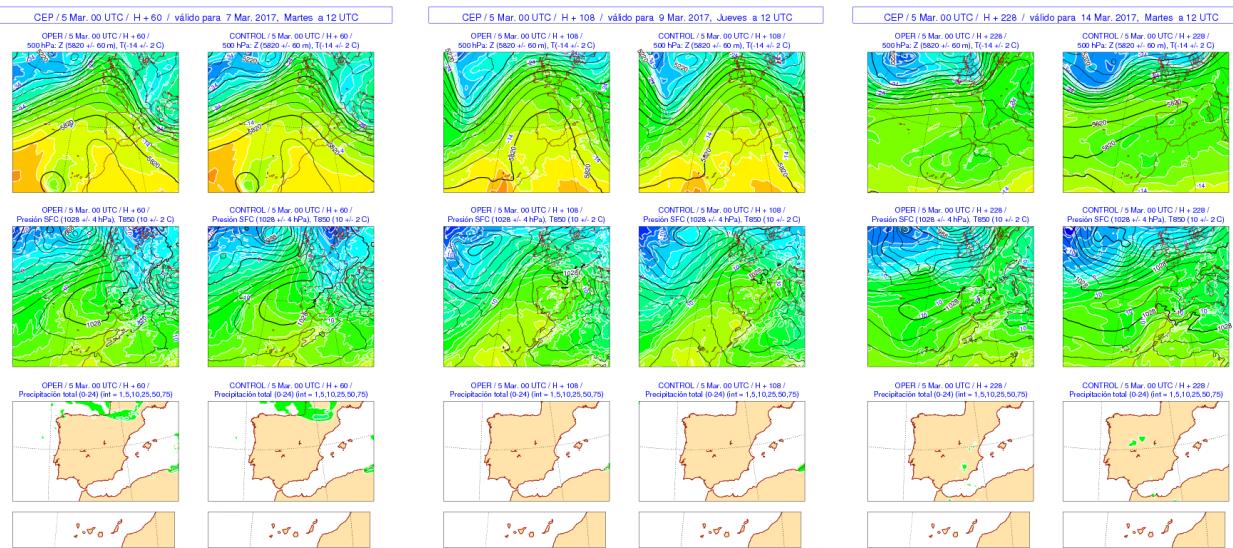


Figura 27.5: Comparación control-determinista para tres alcances de predicción: D+3, D+5, D+10. En cada alcance se muestran, en filas, un bosquejo de la atmósfera en niveles medios-altos (T y Z en 500 hPa), otro en niveles bajos (MSLP y T en 850 hPa) y la precipitación en 24 h. En columnas se muestran el determinista ECHRES a la izquierda y el control del ECENS a la derecha. En este caso control y determinista presentan pautas relativamente similares en todos los alcances (aunque las diferencias van aumentando con el alcance de predicción, en este caso no son diferencias notables) y se infiere que la resolución horizontal no juega un papel primordial en la situación (ver texto). AEMET 2017.

Si *control* y *determinista* presentan pautas similares, entonces la resolución horizontal no es algo crucial en la situación y podemos seguir el resto de la información del SPC sin más consideraciones (ejemplo en la Figura 27.5). Sin embargo, si presentasen pautas distintas, eso implicaría que la resolución horizontal jugaría un papel importante en la situación y, por tanto, podemos seguir la información del SPC pero con cierta precaución, dado que su resolución es menor y la resolución es importante en esa situación.

27.4 Estimación de la incertidumbre: agrupamiento

Mucho más allá de la información determinista de los miembros en los «productos crudos», los SPC proporcionan su aspecto más útil, dan «el do de pecho», cuando los miramos como sistemas de predicción intrínsecamente probabilistas. En este contexto, muchas de las salidas de los SPC reflejan esta naturaleza probabilista. La *probabilidad* es un modelo rico y razonable para describir y comprender muchos aspectos del mundo físico. Pero la interpretación de la información proporcionada (salidas) por los SPC debe aprenderse y usarse cuidadosamente más allá de la interpretación directa, porque (especialmente para los predictores y usuarios «deterministas») algunas interpretaciones

pueden chocar con su «sentido común» (anexo G en la página 1019).

La valoración de la incertidumbre atmosférica es un requisito previo importante en el proceso de predicción. Esta valoración puede realizarse con diferentes herramientas, complementarias y no excluyentes. Presentamos aquí las agrupaciones (grupos y supergrupos) y los mapas dispersión-promedio.

27.4.1 Grupos AEMET

Para la predicción operativa, el manejo de un producto como los sellos es inviable en la práctica: no se dispone del tiempo suficiente para analizar en detalle 51 miembros o escenarios, con todos los parámetros (presión, temperatura, etc.), alcances (plazos en el tiempo), etc. Para hacer tratable la información y resaltar partes predecibles, una primerísima aproximación natural a este problema es la *agrupación*: considerar similitudes entre los diferentes miembros, y agruparlos según un cierto criterio. El criterio de similitud puede ser una cierta *norma* o *métrica* (e. g. correlación entre campos, o diferencia cuadrática media) y, una vez agrupados los miembros, se promedian los miembros de cada grupo constituyendo nuevas «predicciones» o escenarios. El *agrupamiento* (*clustering*) es un mundo en sí

mismo dentro de la ciencia [7, 15, 16]; existen numerosos algoritmos de agrupamiento, según el propósito buscado, según la naturaleza de los elementos que queramos agrupar o el espacio en el que trabajemos, etc. No existe un método de agrupamiento que sea mejor que los demás, cada método tiene sus ventajas e inconvenientes y, según el contexto, podrá aplicarse un método mejor que otro.

En AEMET se realiza un *agrupamiento localizado* del ECENS (51 miembros) mediante ordenador. Es localizado (*target clustering*) porque la similitud entre miembros se calcula para un dominio regional que comprende la Península y Baleares (el ECENS es global, por lo que cualquier subconjunto es regional) o Canarias en otro agrupamiento distinto. El agrupamiento se computa mediante el llamado *algoritmo de Ward* [16], que pertenece a la familia de los llamados algoritmos jerárquicos [7]: la similaridad entre miembros se mide con la diferencia cuadrática media del campo de altitud geopotencial en 500 hPa (Z500) para una serie de alcances (T+120...T+168), de modo que el cálculo garantiza una «continuidad sinóptica». Por

debajo de un umbral predeterminado de diferencia cuadrática, se considera que dos miembros son suficientemente similares para pertenecer al mismo grupo. Aunque el número total de grupos (clusters) puede ser, en principio, variable, en AEMET se computan siempre seis grupos. El número de grupos, potencialmente variable aunque se fije en seis, dependería de tres factores principales: (i) la dispersión atmosférica propia del día, (ii) el umbral de diferencia cuadrática media escogido y (iii) el grado de modalidad (en principio, multimodal implica más grupos). Una vez agrupados los miembros, y asumiendo que la *varianza interna* de cada grupo es pequeña, se calcula el representante o *centroide* del grupo haciendo el valor medio de los miembros del grupo. Recuérdese que el promedio del SPC no se debe utilizar, en principio, o en todo caso utilizar con mucha precaución, pues puede no corresponder a ninguna situación meteorológica plausible (sec. 13.6.3 en la página 178). Sin embargo, el promedio de los miembros de un grupo sí que representará una situación meteorológica plausible, dado que la varianza interna del grupo es pequeña por construcción (Figura 27.6).

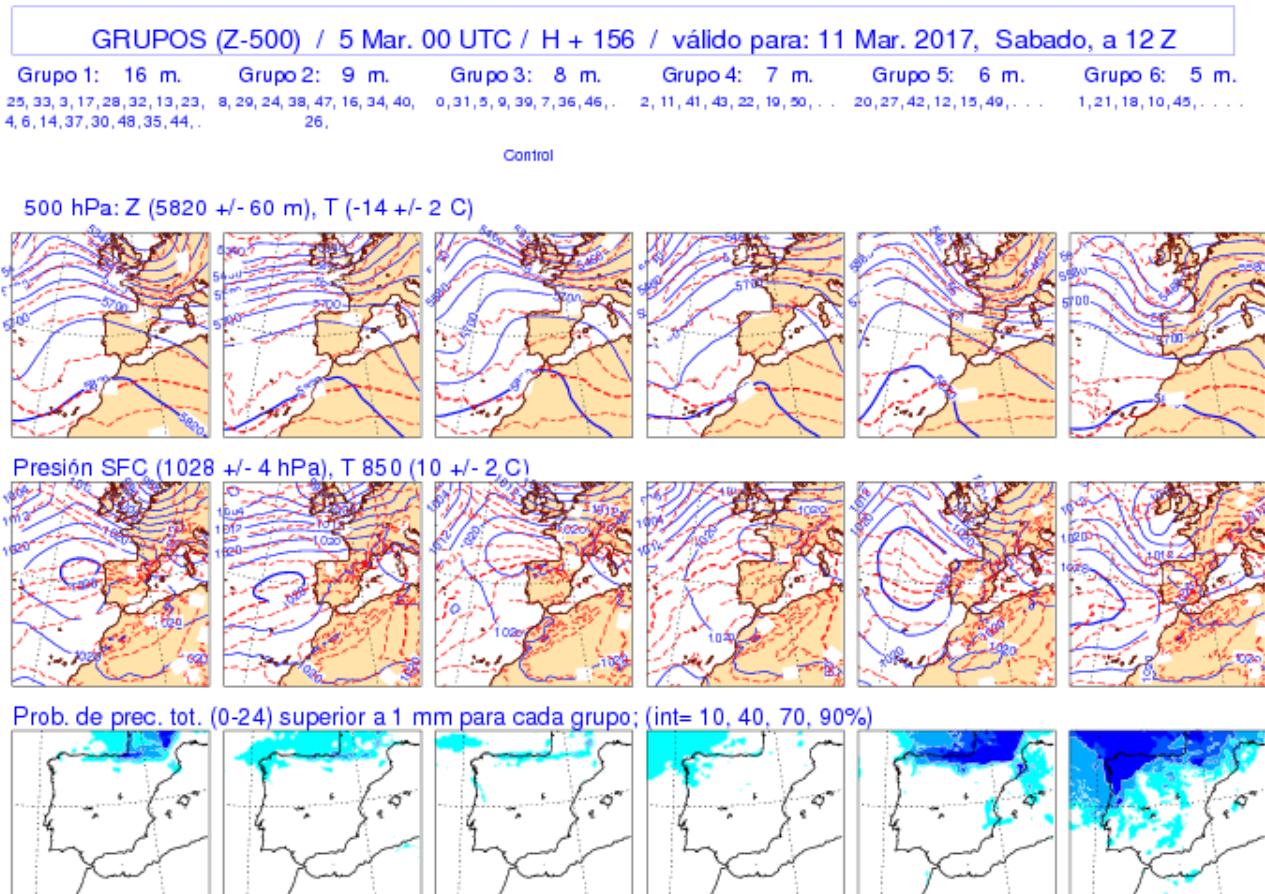


Figura 27.6: Grupos AEMET D+7 (ver texto). AEMET 2017.

Como ejemplo, la Figura 27.6 en la página anterior muestra los grupos elaborados en AEMET con el ECENS del día 20170305 pasada 00 UTC, alcance D+7, es decir, válido para el 20170311 a las 12 UTC. En la primera fila se muestra un resumen de la atmósfera en niveles medios, 500 hPa (altitud geopotencial y temperatura), en la segunda fila un compendio del comportamiento atmosférico en niveles bajos (presión reducida al nivel del mar y temperatura en 850 hPa; no se pinta la temperatura en superficie por ser un parámetro muy «orográfico», sigue la orografía demasiado fielmente, y por tanto no brinda información tan

útil) y en la última fila la probabilidad de precipitación acumulada en 24 horas superior a 1 mm (sec. 13.6.14 en la página 189). Las columnas muestran los diferentes grupos resultado del algoritmo de Ward. Para cada columna, en cabecera se muestra el número de miembros comprendidos en ese grupo, y cuáles son. Dado que asumimos que los miembros representan escenarios atmosféricos equiprobables (sec. 13.7.2 en la página 191), entonces ese número de miembros del grupo puede considerarse como el *peso* probabilista del grupo, de modo que la probabilidad de cada grupo sería proporcional a su peso.

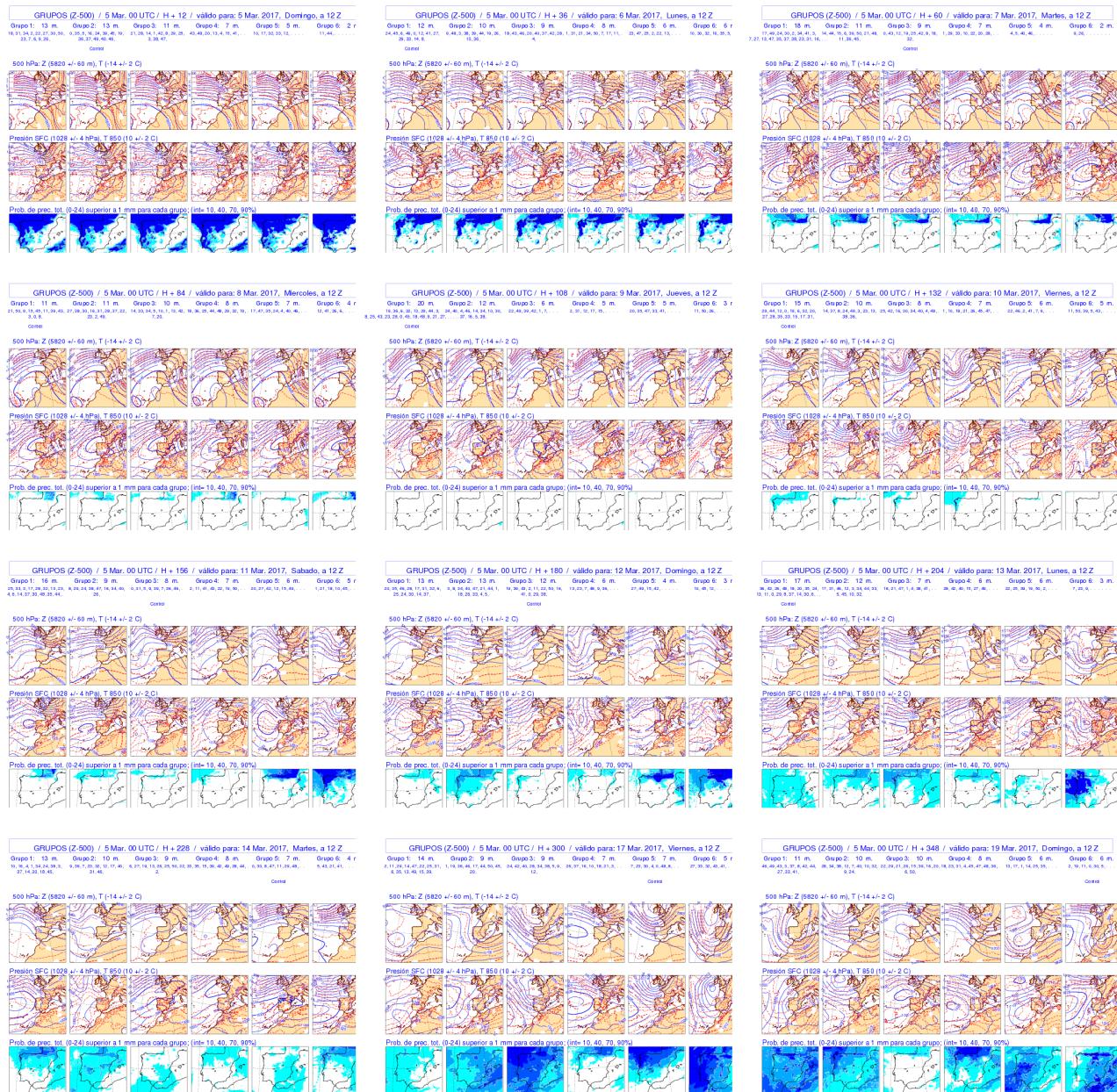


Figura 27.7: Grupos para la Península y Baleares desde el día D+1 hasta el D+10, y luego D+13 y D+15 (ver texto). AEMET 2017.

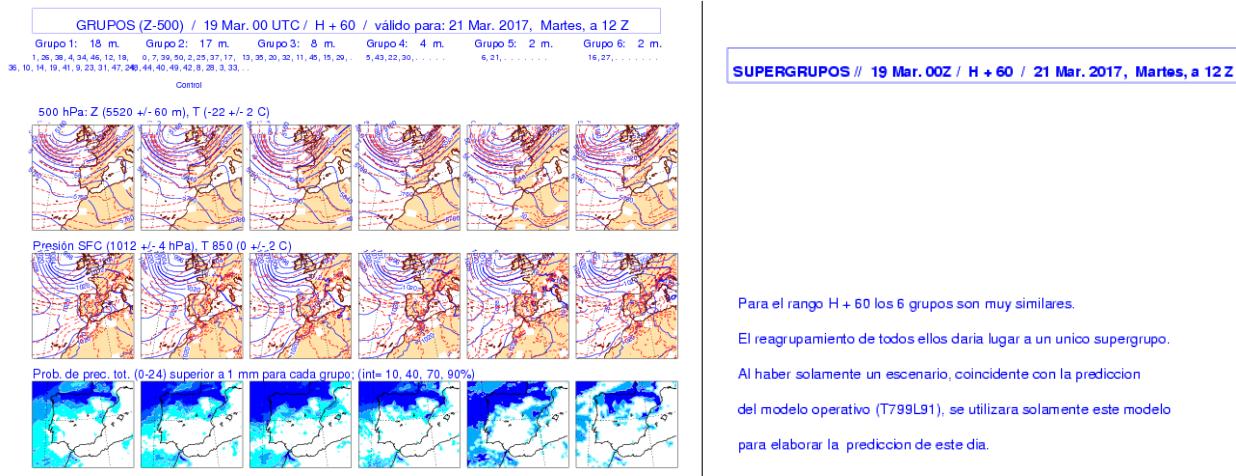


Figura 27.8: Grupos D+3 y supergrupos D+3, realizados con la pasada 2017-03-19 de las 00 UTC, válidos para 2017-03-21 a las 12 UTC. AEMET 2017.

Nótese que los grupos se muestran de más peso a la izquierda a menos peso a la derecha. Es importante conocer en qué grupo está comprendido el miembro de control, y se dispone precisamente de esa información: en este caso el control ha caído en el grupo 3. Obsérvese que con los grupos se dispone de información cuantitativa, detallada y explícita de la PDF atmosférica (sec. 13.6 en la página 176) en forma de mapas, en particular de la dispersión y de la incertidumbre. Es una información riquísima y con un gran valor sintético: se resume mucha información en 3x6 imágenes. El grupo 1 es el más probable, el siguiente el grupo 2, etc. El miembro de control está comprendido usualmente en el grupo 1 o en el 2, pero no siempre es así (como es el caso en este ejemplo). El escenario atmosférico que finalmente representó mejor lo que se observó (la «realidad») no tiene por qué estar comprendido en los grupos más probables y, aunque ciertamente suele estarlo, aquellos días del año en que esto no es así suelen ser más difícilmente predecibles (presentan menos predecibilidad). Una atmósfera muy estable (e. g. situación anticiclónica) suele corresponder a una PDF más bien *monomodal* del SPC, con un grupo 1 muy numeroso y los demás grupos con pocos miembros. Atmósferas más inestables presentan más dispersión y pueden ser más bien bimodales, trimodales, etc. Ese grado de *modalidad* (en el sentido de número de *modos*) tiene cierta correspondencia con la *incertidumbre* atmosférica, y por ende con la falta de *predecibilidad*. En ocasiones hay ingredientes muy finos para considerar: por ejemplo, no siempre la falta de dispersión implica poca incertidumbre; cuando el determinista de mayor resolución y el miembro de control no se parecen, entonces la resolución horizontal juega un

papel crítico, y la PDF que representa el SPC puede desviarse bastante del escenario del determinista. En esos casos hay poca predecibilidad aunque haya poca dispersión en el SPC. A lo largo de todo el alcance de predicción, los grupos suelen evolucionar aumentando la dispersión gradualmente y, aunque no siempre un día tiene menos dispersión que el siguiente, sí que suele ser la pauta general. La Figura 27.7 en la página anterior muestra los grupos elaborados en AEMET con el ECENS del día 20170305 pasada 00 UTC, alcances desde D+1 hasta D+15, es decir, válido desde el 20170306 a las 12 UTC hasta el 20170320 a las 12 UTC. Puede observarse como pauta general que los grupos del D+1 suelen ser todos similares. A medida que avanzamos en el plazo de predicción (D+2, D+3, ...) la dispersión atmosférica y la incertidumbre van creciendo a su vez, y los grupos van pareciéndose cada vez menos, de modo que la atmósfera es cada vez menos predecible. Los grupos para el D+7 ya presentan diferencias notables, y en el D+13 cualquier parecido entre ellos podría ser «pura coincidencia» si no supiéramos que se trata de predicciones para el mismo día. Es notable y señalamos aquí que en este contexto se tiene ya un soporte cuantitativo, detallado y explícito para afirmar que la atmósfera es menos predecible según avanza el alcance de predicción.

27.4.2 Supergrupos AEMET

En las Figuras 27.8 y 27.9 en la página siguiente se muestran ejemplos de grupos y, también, supergrupos correspondientes, generados en el CNP de AEMET. ¿En qué consisten estos supergrupos?

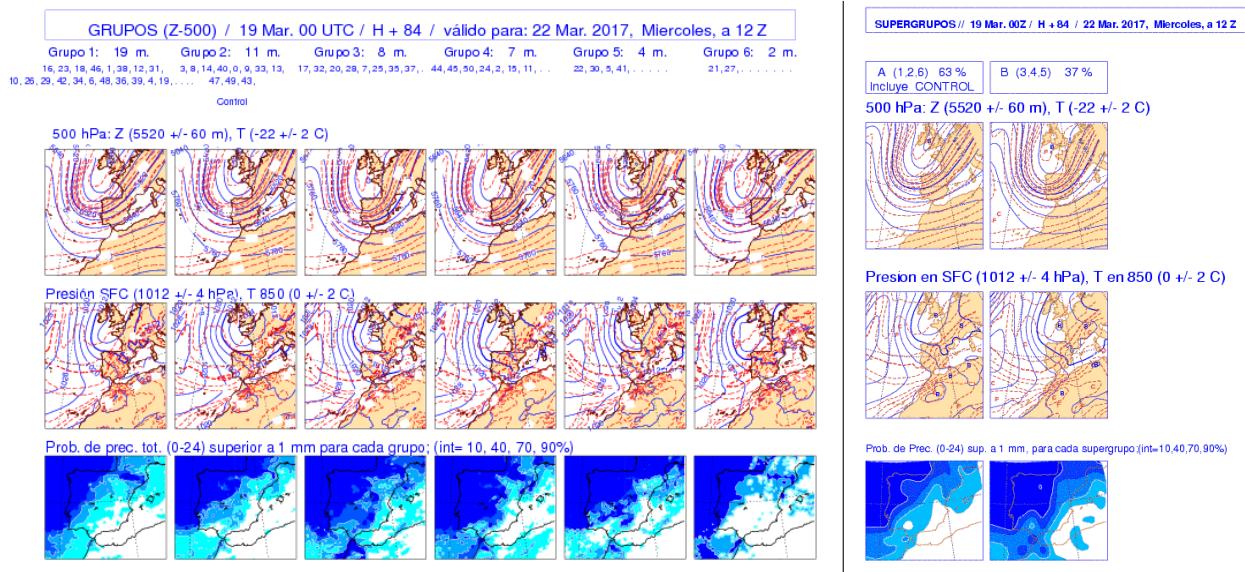


Figura 27.9: Grupos D+4 y supergrupos D+4, realizados con la pasada 2017-03-19 de las 00 UTC, válidos para 2017-03-22 a las 12 UTC. AEMET 2017.

En el procedimiento operativo de **AEMET**, una vez que se han computado automáticamente los grupos, el meteorólogo predictor examina estos grupos para los alcances D+3, D+4 y D+5, analizando para cada día diferencias a gran escala entre grupos en la situación atmosférica y su posible impacto en el tiempo sensible (viento, precipitación...). Una vez examinados, los «reagrupa» en «supergrupos». El número de supergrupos no es fijo, sino que dependerá de la dispersión y la incertidumbre del día en cuestión.

Cuando todos los grupos son muy similares, se genera un solo supergrupo, el día tiene muy poca incertidumbre, la atmósfera puede predecirse con el modelo determinista de mayor resolución ECHRES y no es necesario apoyarse en el ECENS; este sería el caso, típicamente, de las situaciones anticiclónicas. Muchos otros días hay dispersión suficiente para formar dos, o incluso tres, supergrupos; este sería el caso de días con paso de vaguada móvil, paso de frentes, etc. En casos excepcionales pueden llegar a formarse cuatro supergrupos; podría ser el caso de algunas depresiones aisladas en niveles altos (**depresión aislada de niveles altos (DANA)**) muy erráticas, por ejemplo. Igual que ocurre con los grupos, se cumple en general que la dispersión aumenta con el alcance de predicción, por lo que en el D+3 a menudo se forma un sólo supergrupo, en el D+4 uno o dos, y en el D+5 uno, dos o hasta tres en los casos usuales.

Pero este comportamiento, aunque sea una pauta general aproximada, no puede generalizarse y ahí está la ciencia y el arte del predictor. Podemos decir que

la dispersión aumenta y la predictibilidad disminuye con el alcance de predicción, pero no todos los días en el mismo grado ni al mismo ritmo. No se puede simplificar diciendo que «los SPC se usan en el medio plazo porque es ahí donde la atmósfera empieza a presentar más dispersión». Los SPC resultarán más útiles, a menudo, en el medio plazo que en el corto, pero habrá días, importantes por su dificultad, en que los SPC resulten utilísimos o incluso imprescindibles en el corto plazo. Como decíamos, estos supergrupos se deciden manualmente, y después mediante ordenador se calculan los *centroídes* correspondientes a esos supergrupos y se muestran en la intranet de **AEMET**. Quedan disponibles para el resto de predictores de **AEMET**, ahorrando una buena cantidad de tiempo y de trabajo, dado que la información sobre dispersión e incertidumbre queda aún más sintetizada que con los grupos.

En las Figuras 27.8 en la página anterior y 27.9 se muestran dos ejemplos de grupos y, también, supergrupos correspondientes, generados en el **CNP** de **AEMET** con el ECENS del día 20170319 pasada de 00 UTC, válidos para el D+3 y para el D+4, respectivamente. Para el D+3 hay poca dispersión y se genera un sólo supergrupo, por lo que se sugiere al personal de predicción operativa que utilice directamente el modelo determinista de mayor resolución. Para el D+4, sin embargo, hay dispersión suficiente como para generar dos supergrupos distintos, dos escenarios atmosféricos propiamente dichos en cuanto a modelo conceptual y tiempo presente.

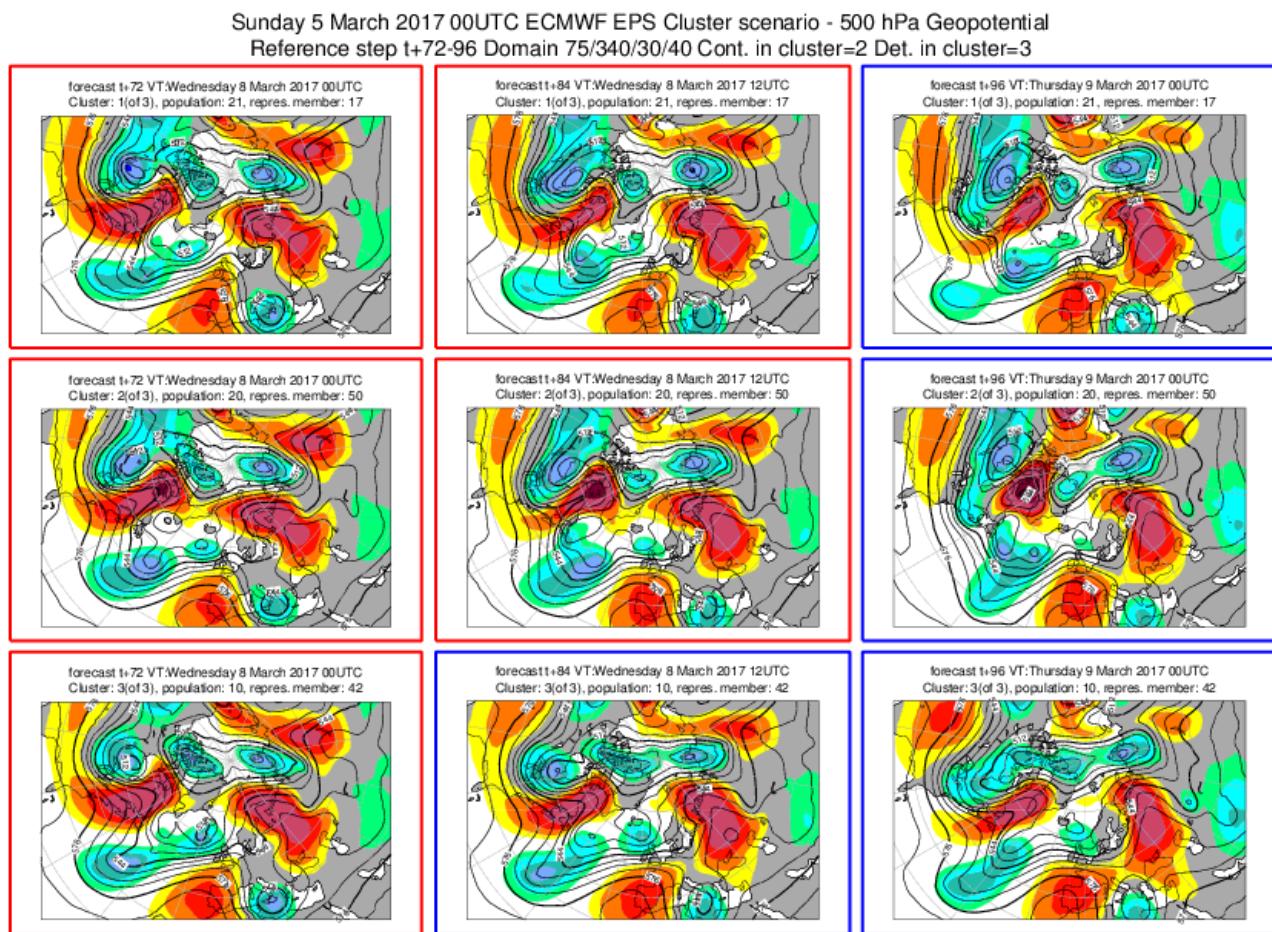


Figura 27.10: Grupos ECMWF. ECMWF 2017.

27.4.3 Grupos ECMWF

En el ECMWF se realiza otro tipo de agrupamiento, con características diferentes, que no se suele utilizar operativamente en [AEMET](#), pero se describe aquí por completar el enfoque. El dominio de agrupamiento es distinto, incorporando una buena parte del hemisferio norte (mucho más amplio que el dominio usado en [AEMET](#)). El algoritmo también es diferente: el número de grupos es variable, no fijo. Se buscan similitudes en el campo de altitud geopotencial en 500 hPa, en distintos intervalos de alcance de predicción: 72-96 h, 120-168 h, 192-240 h y 264-360 horas. Se ejecuta un procedimiento diario de agrupamiento para determinar escenarios de tiempo atmosférico, cada uno definido por el miembro del ECENS más cercano al centroide del grupo, lo cual aporta realismo. Se utiliza, además, un conjunto predefinido de cuatro regímenes climatológicos para cada estación del año, calculados con 29 años de *reanálisis* (ERA-Interim [2] y ERA-40 [19]), denominados *bloqueo*, *NAO positiva*, *NAO negativa* y *dorsal atlántica* (ver Figura 27.11 en la página siguiente y sec. 29.2.2 en la página 485). A cada escenario de tiempo se le atribuye un régimen

climatológico, aportando una medida de diferencias entre escenarios según el flujo de gran escala. De este modo, el usuario tiene dos informaciones útiles: el escenario y su atribución climatológica. En un futuro se darán medidas de calidad de las predicciones según el régimen, es decir, dependiendo del flujo. Ese contraste permitirá conocer la predecibilidad dependiente del flujo (sec. 12.7 en la página 164). En los paneles de grupos (Figura 27.10) se muestra la altitud geopotencial en 500 hPa y en 1000 hPa, en contornos sólidos, de los miembros del ECENS representativos de cada grupo (cada uno es el más cercano al centroide). En 500 hPa se muestran, además, las anomalías en colores. En aquellos casos en que la distribución del ECENS es suficientemente homogénea, el algoritmo de agrupamiento no puede agrupar de modo significativo: en esos casos se representa el miembro más cercano al promedio del ECENS. Cada escenario, como decíamos antes, está asociado a uno de los cuatro regímenes climatológicos de gran escala predefinidos, lo que se representa mediante un marco de color que lleva cada mapa: rojo para *bloqueo*, azul para *NAO positiva*, verde para *NAO negativa* y violeta para *dorsal atlántica*.

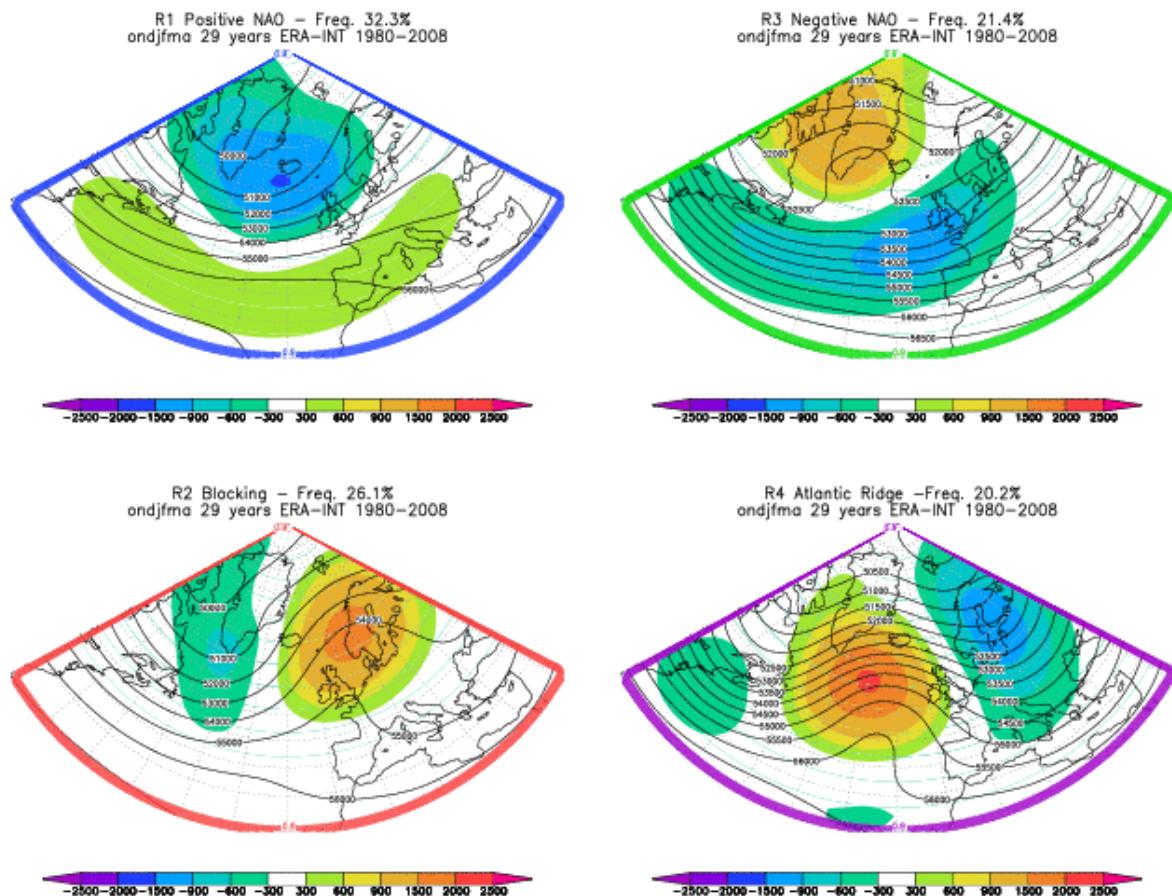


Figura 27.11: Regímenes climatológicos usados para atribuir a cada grupo. ECMWF 2017.

27.4.4 Nuevos grupos (2017) en AEMET

En noviembre de 2017 se implementó en AEMET una nueva técnica de agrupamiento, sustituyendo a la anterior junto con todos los paneles de presentación. Dado que la mayoría de casos meteorológicos presentados en este libro se han construido con las técnicas anteriores, se ha mantenido la descripción de la presentación antigua. En algún caso concreto se usa la nueva que, con mucho detalle, presentamos en el anexo de este capítulo (sec. 27.14 en la página 440).

27.5 Estimación de la incertidumbre: dispersión

27.5.1 Mapas de dispersión-promedio

Como se detalla en el capítulo de Sistemas de Predicción por Conjuntos (sec. 13.6.3 en la página 178)

el promedio aritmético del ensemble no es siempre una situación meteorológica plausible, pues es resultado de una operación estadística y no de un modelo atmosférico [1, 6]. Por tanto, se ruega encarecidamente no usar el promedio del ensemble en predicción operativa, salvo que se haga con mucha precaución. No obstante, una excepción notable son los mapas de dispersión-promedio que aquí presentamos, en los que el promedio se dibuja solo orientativamente. Junto a la *centralización*, la *simetría*, la *curtosis* y otras propiedades, la *dispersión* es una de las herramientas estadísticas conceptuales para sintetizar la información de un conjunto, a menudo amplio, de información. Asociamos en general la dispersión del SPC a la *incertidumbre atmosférica*, aunque en particular hay casos en los que esta relación requiere de matrices adicionales. En un punto de malla del SPC, se dispone de N predicciones para un parámetro y alcance de predicción. Se trata de una distribución de números, una cantidad grande de información que se requiere sintetizar.

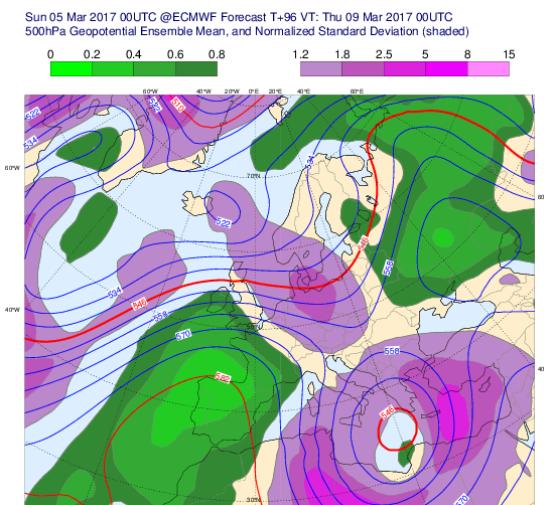


Figura 27.12: Ejemplo de mapa dispersión-promedio, en este caso dispersión normalizada y ECHRES sustituyendo al promedio. ECMWF 2017.

Dos cantidades que pueden sintetizar la información dada por esa distribución son la media, medida de centralización (sec. 13.6.3 en la página 178) y la desviación estándar, medida de dispersión (sec. 13.6.5 en la página 181). Extendiendo ese procedimiento a todos los puntos de la malla y dibujándolo, resultará un mapa de promedio y dispersión, abreviadamente dispersión-promedio, de los cuales hay diversas modalidades.

En la Figura 27.12 vemos un ejemplo de mapa de dispersión-promedio elaborado con el ECENS del ECMWF, de la altitud geopotencial en 500 hPa. Los mapas de esta variable ofrecen una orientación sobre el flujo atmosférico en altitudes en torno a los 5.5 km; las estructuras que aparecen en forma de ondulaciones, dorsales, vaguadas, etc., permiten al predictor, junto con mapas de otras variables en este y otros niveles, una primera composición 3D de la atmósfera para poder hacer diagnóstico y predicción. Las isolíneas muestran el promedio del SPC que, insistimos una vez más, debe mirarse con mucha precaución [6]. No está garantizado que el promedio sea una situación atmosférica plausible [1], ni siquiera realista de modo que, aunque a menudo el mapa pueda ser interpretado meteorológicamente a gran escala, puede haber estructuras de pequeña escala que tengan un sentido relativo y matizable. Se muestran algunas isolíneas en rojo como referencia, cada 6 decámetros (6 dam) y el resto en azul, cada decámetro (dam). Por otra parte se representa la desviación estándar como medida de dispersión. En este mapa no se representa la

desviación estándar tal cual, sino que se normaliza con respecto a la desviación de los últimos 30 días, para tener una idea ponderada de desviación relativa a la desviación promedio de esos últimos 30 días. Por tanto, la *desviación estándar normalizada* no tiene unidades, dado que está dividida entre una cantidad con sus mismas unidades. Esta desviación normalizada nos indicará que la desviación es similar, superior o inferior a la del último mes si su valor es similar, superior (tonos rosados) o inferior (tonos verdes) a 1, respectivamente. Esas zonas sombreadas verdes (rosas) en color más oscuro son aquellas en las que hay menos (más) desviación estándar relativa y por tanto donde los miembros del SPC se parecen más (menos), es decir, menos (más) dispersión relativa. De modo que en las zonas verdes asumimos menos incertidumbre atmosférica y, en principio, más predecibilidad. Por el contrario, en las zonas rosadas asumimos más incertidumbre atmosférica y, en principio, menos predecibilidad. La utilidad de estos mapas es, entonces, muy clara: nos muestran aquellas *áreas en las que tendremos una predecibilidad potencialmente mayor y menor comparativamente con el último mes* y, por tanto, *áreas en las que se deberá poner atención especial en el proceso de predicción*.

En la Figura 27.13 vemos un ejemplo diferente de mapa de dispersión-promedio elaborado con el ECENS del ECMWF, de la altitud geopotencial en 500 hPa. Diferente porque ahora se muestra el ECHRES en lugar del promedio del ECENS y porque ahora se muestra la desviación estándar tal cual. Las isolíneas muestran el modelo determinista ECHRES.

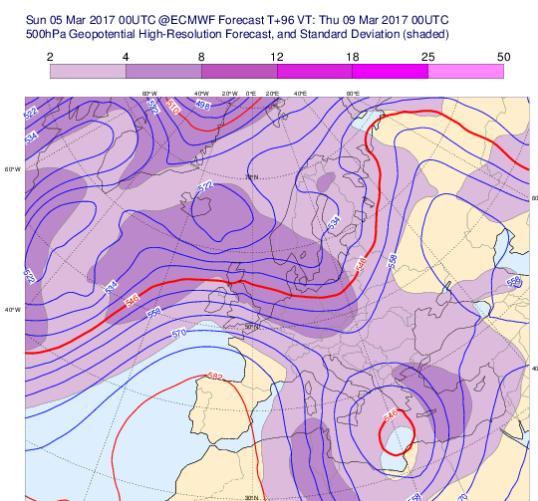


Figura 27.13: Ejemplo de mapa dispersión-promedio, en este caso dispersión sin normalizar y promedio directo del ECENS. ECMWF 2017.

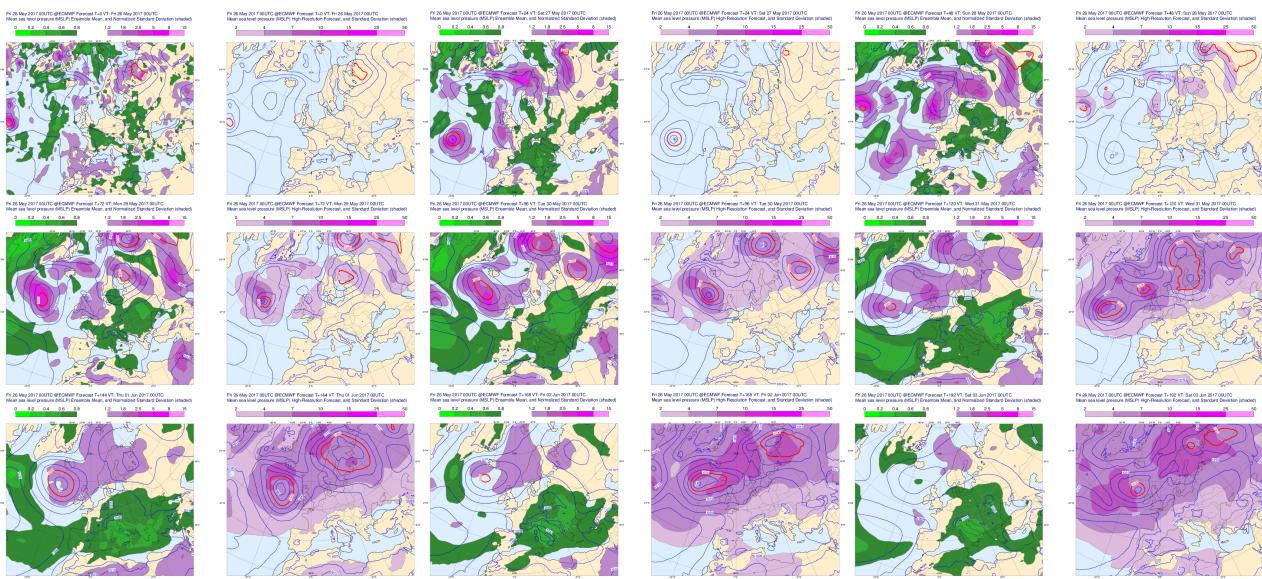


Figura 27.14: Mapas dispersión-promedio evolucionando con el alcance de predicción. ECMWF 2017.

Se muestran, como antes, algunas isolíneas en rojo como referencia, cada 6 décámetros (6 dam) y el resto en azul, cada dam. Por otra parte se representa la desviación estándar como medida de dispersión. En este mapa sí que se representa la desviación estándar tal cual. Sus unidades son las mismas que las de la variable en sí, en este caso dam. La zonas sombreadas en color más oscuro son aquellas en las que hay más desviación estándar y por tanto donde los miembros del SPC se parecen menos (los valores están más dispersos). En esas zonas asumimos más incertidumbre atmosférica y, en principio, menos predecibilidad. La utilidad de estos mapas es, entonces, muy clara y similar al ejemplo anterior, pero sin el ingrediente comparativo relativo a los últimos 30 días: nos muestran aquellas *áreas en las que tendremos una predecibilidad potencialmente menor* y, por tanto, áreas en las que se deberá poner atención especial en el proceso de predicción.

La evolución con el alcance de predicción suele ir unida a un aumento de dispersión: la incertidumbre atmosférica suele crecer con dicho horizonte y, por ende, disminuir la predecibilidad. Pero es importante saber que esto no es siempre así, y aquellos alcances en los que la dispersión, en vez de aumentar, disminuye, muestran señales muy interesantes para el proceso de predicción. Esta evolución con el alcance de la predicción la podemos ver en la Figura 27.14, donde se muestran en 3 filas y 3 columnas ordenadamente de derecha a izquierda y de arriba abajo los mapas de dispersión promedio, en sus dos variantes mostradas anteriormente, para los alcances de predicción desde D+0 («hoy») hasta D+8 («día más ocho»). Se observa

claramente cómo, en general y a gran escala, la dispersión, mapas sólo en rosa, va aumentando con el alcance: en el D+0 el color rosa no aparece, en el D+1 aparece alguna tímida sombra aquí y allá y después, progresivamente, van apareciendo cada vez más áreas con dispersión, mayor cada vez. Es importante resaltar que, para una región particular como podría ser una parte de la península ibérica, la dispersión no tiene por qué ir siempre a más: puede haber algún día en el que la dispersión baje o se mantenga relativamente constante. También destacamos que en ocasiones, la dispersión puede ser grande en un área en el panel derecho y, sin embargo, presentar valores iguales o inferiores a 1 en el panel izquierdo; ésto se debe a que, en predicciones hechas en días previos se dieron valores de dispersión en ese área iguales o mayores a los presentes.

Estos mapas de dispersión-promedio suelen utilizarse para valorar la incertidumbre atmosférica en diferentes zonas según los distintos alcances de predicción, incertidumbre en una escala sinóptica (procedimientos usuales en 2017), evaluando los patrones sinópticos de campos dinámicos en niveles medios-altos, e. g. altitud geopotencial en 500 hPa, o bien campos de niveles más bajos pero no demasiado ruidosos, e. g. presión reducida al nivel del mar, temperatura en 850 hPa, velocidad del viento en 850 hPa. Si se dibujaran, por ejemplo, para la temperatura a 2 m, que es un campo muy orográfico y ruidoso comparado con los anteriormente citados, la dispersión sería así mismo bastante ruidosa, lo que dificultaría mucho su interpretación y le restaría utilidad a la herramienta.

¿Cómo se interpretan estos mapas? En líneas generales, la interpretación se realiza identificando las áreas de mayor dispersión sobre los patrones sinópticos que presente la situación. Presentamos a continuación tres pautas sencillas y generales, ilustradas en las Figura 27.15, 27.17 y 27.17.

Eje de vanguardia. Si las áreas de mayor dispersión se presentan sobre un eje de vanguardia, esperaremos más incertidumbre en la intensidad de la vanguardia, dado que las zonas donde los miembros difieren más están sobre el eje. Puede visualizarse bien el razonamiento comparando espaguetis con dispersión, como en la figura.

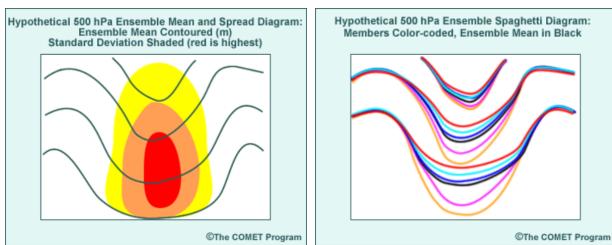


Figura 27.15: Interpretación conceptual de los mapas de dispersión. Eje de vanguardia. COMET.

Ambos flancos de una vanguardia. Si las áreas de mayor dispersión se presentan corriente arriba y abajo (a ambos lados de la vanguardia), esperamos más incertidumbre en la posición de dicha vanguardia, no en su intensidad.

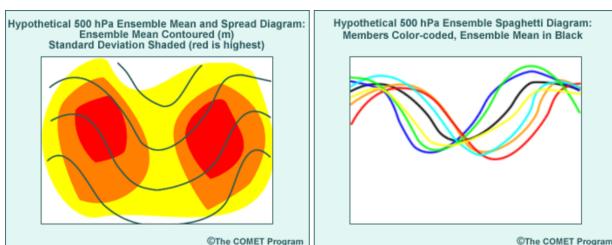


Figura 27.16: Interpretación conceptual de los mapas de dispersión. Ambos flancos de una vanguardia. COMET.

Un solo flanco de vanguardia. Si las áreas de mayor dispersión se presentan a un lado de la vanguardia solamente, entonces esperamos probablemente características menos predecibles, presentes a menudo en una escala menor, e. g. mesoscala de algún orden y ocultas por el procedimiento de suavizado al dibujar los campos.

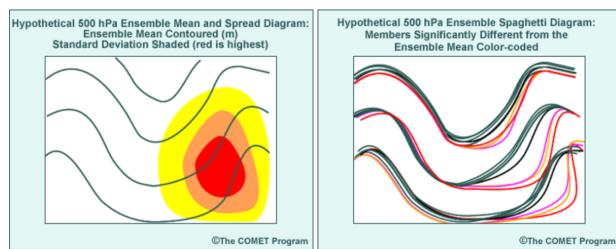


Figura 27.17: Interpretación conceptual de los mapas de dispersión. Un solo flanco de vanguardia. COMET.

Valores altos de desviación estándar suelen indicar una incertidumbre en la predicción también alta y, por ende, la fiabilidad de los mapas de la variable correspondiente, tanto la del modelo determinista de mayor resolución correspondiente (si lo hay) como la del promedio del SPC. Por otro lado, la desviación estándar normalizada pone en contexto el comportamiento general del SPC en el área mostrada en los últimos 30 días. Valores de desviación altos, aunque en general van asociados a mayor incertidumbre, si los días previos han sido altos también o incluso superiores, entonces la desviación normalizada mostrará valores relativamente bajos, aumentando ligeramente nuestra confianza en los mapas de la variable.

En el sitio web del ECMWF pueden consultarse mapas dispersión-promedio de cuatro variables meteorológicas diferentes del ECENS: presión reducida al nivel del mar, temperatura en 850 hPa, velocidad del viento en 850 hPa y altitud geopotencial en 500 hPa. Estos mapas se presentan (2017) en parejas: a la izquierda el promedio del ECENS con la desviación estándar normalizada y a la derecha, para comparar, el modelo determinista de mayor resolución ECHRES con la desviación estándar tal cual.

Como resumen diremos que estos mapas de dispersión-promedio son más específicos para la predicción operativa que para los usuarios, mostrando por un lado la dispersión (o dispersión normalizada) que tiene la predicción probabilista de un parámetro, y por otro, el valor del modelo determinista correspondiente (o del promedio del SPC). Aunque no está rigurosamente demostrada una relación directa entre mayor dispersión y menor predecibilidad, esta relación se asume ordinariamente por la experiencia en el uso de los SPC y tiene un uso orientativo. Así, en aquellas regiones donde la dispersión sea mayor, la incertidumbre será mayor, y el predictor puede poner especial atención en esas zonas.

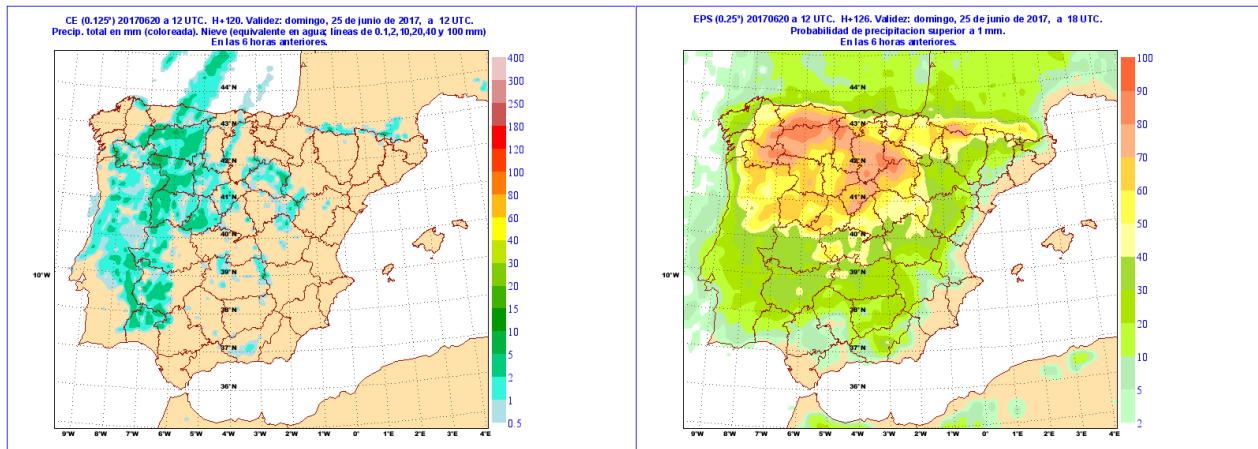


Figura 27.18: Comparación de un mapa de precipitación prevista determinista (izquierda) con el correspondiente mapa de probabilidad de precipitación prevista (derecha) (ver texto). AEMET 2017.

27.6 Predicción de eventos

27.6.1 Mapas de probabilidad

Los típicos productos meteorológicos de los modelos deterministas se muestran en un mapa, con una gama de colores, la gama de valores del parámetro meteorológico de interés. Por ejemplo, la precipitación prevista (entre 0 y 50 mm) para mañana entre las 06 y las 12 UTC (Figura 27.18 izquierda). En contraposición, los SPC muestran en un mapa, con una gama de colores, las diferentes probabilidades de que ocurra un evento meteorológico de interés. Por ejemplo, la probabilidad prevista de obtener precipitación superior a 1 mm para mañana entre las 12 y las 18 UTC (Figura 27.18 derecha).

Centremos nuestro interés en una variable meteorológica como la precipitación. La casuística de la precipitación puede desarrollarse según una serie de umbrales de precipitación (e. g. 1, 5, 10 y 20 mm) y sus correspondientes probabilidades de superación. Ahora fijemos el umbral de 1 mm y hagamos la pregunta ¿qué probabilidad tenemos de superar 1 mm de

precipitación? (algo así como preguntar ¿llueve o no llueve?). Pues con ayuda del SPC, en cada punto de rejilla tenemos un conjunto de valores previstos de precipitación (e. g. 51 valores). Si en un punto concreto de rejilla la mitad de esos valores superan 1 mm de precipitación, entonces la probabilidad, con un sencillo cálculo, es del 50%, para ese punto. Para otro punto será 70%, para otro 20%, para otros 0%, etc. Podemos dibujar esta idea en un mapa, coloreando con diferentes colores distintos valores de probabilidad. En la Figura 27.18 derecha se ilustra el resultado de este proceso. El color blanco es para probabilidades entre 0 y 2% («no llueve»), el verde playa entre 2 y 10% («apenas llueve»), gama de verdes entre 10 y 40% («no se descarta lluvia»), amarillo-naranja entre 40% y 70% («lluvia posible»), naranja-rojo entre 70% y 90% («lluvia probable») y, por último, rojo entre 90% y 100% («llueve seguro»). El mismo procedimiento puede realizarse con diferentes umbrales de precipitación: en la Figura 27.19 se muestran, completando la Figura anterior, los umbrales de 5, 10 y 20 mm, típicos en predicción operativa.

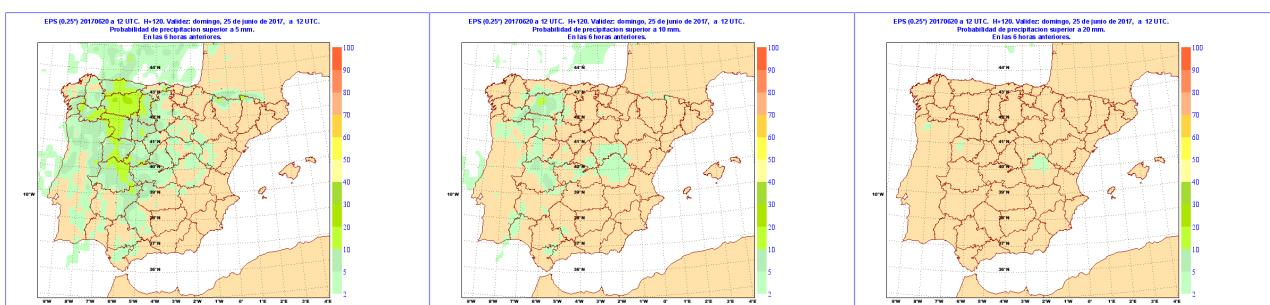


Figura 27.19: Mapas de probabilidad de precipitación prevista para superación de umbrales de 5, 10 y 20 mm (ver texto), completando la Figura 27.18. AEMET 2017.

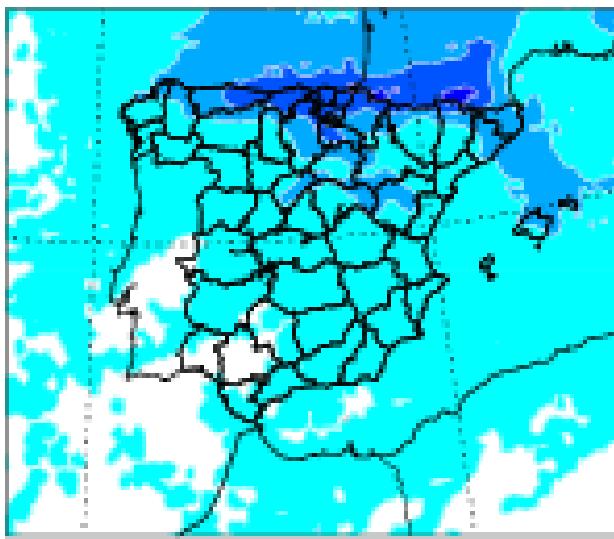


Figura 27.20: Mapa de probabilidad de precipitación prevista superior a 1 mm, según el estilo antiguo de colores: el color blanco es para probabilidades entre 0 y 10% («no llueve»), el celeste entre 10 y 40% («no se descarta lluvia»), azul entre 40% y 70% («lluvia posible»), azul oscuro entre 70% y 90% («lluvia probable») y, por último, azul muy oscuro entre 90% y 100% (llueve «seguro»). Ver texto. AEMET 2017.

Existían varios estilos de mapas de probabilidad usados en AEMET durante el proceso de escritura de este libro: por ejemplo, los mostrados anteriormente son más modernos. Otro estilo, más antiguo pero usado profusamente en los casos de este libro, utiliza gamas de azules para la probabilidad. Para precipitación, por ejemplo, el color blanco es para probabilidades entre 0 y 10% («no llueve»), el celeste entre 10 y 40% («no se descarta lluvia»), azul entre 40% y 70% («lluvia posible»), azul oscuro entre 70% y 90% («lluvia probable») y, por último, azul muy oscuro entre 90% y

100% (llueve «seguro»).

Para diferentes umbrales de precipitación y alcances de predicción tendremos diferentes mapas. La Figura 27.21 muestra abanicos de diversos mapas de probabilidad. Nubosidad y precipitación para diferentes umbrales (1, 5, 10, 20 mm) y alcances de predicción (D+3, D+4, D+5), tanto en la Península y Baleares como en Canarias.

Dado el fenómeno meteorológico de superación de un umbral (e. g. precipitación prevista > 5 mm / 24 h), el mapa de probabilidad muestra la distribución espacial de las probabilidades previstas proporcionadas por el SPC. Mapas similares pueden hacerse para los percentiles, aunque en AEMET se utilizan poco. Los mapas de probabilidad son una útil guía para el predictor, mostrando una primera aproximación de los lugares en los que esperamos con más o menos probabilidad que ocurrán eventos interesantes (e. g. precipitación mayor que 20 mm en 24 horas).

La «temperatura a 2 m» dada por los modelos suele ser una variable posprocesada, derivada de una interpolación entre el nivel más bajo del modelo, y la superficie misma. Es una variable tremadamente «orográfica» (influida por la «orografía»), por lo que en ese posproceso puede tenerse en cuenta la diferencia entre la orografía del modelo y la orografía verdadera. La Figura 27.22 en la página siguiente muestra abanicos de mapas de probabilidad relacionados con la tendencia de la temperatura, es decir, la diferencia de la temperatura con la temperatura del día anterior. A la izquierda, las probabilidades por umbrales de la tendencia de las máximas. A la derecha, lo mismo para la tendencia de las mínimas.

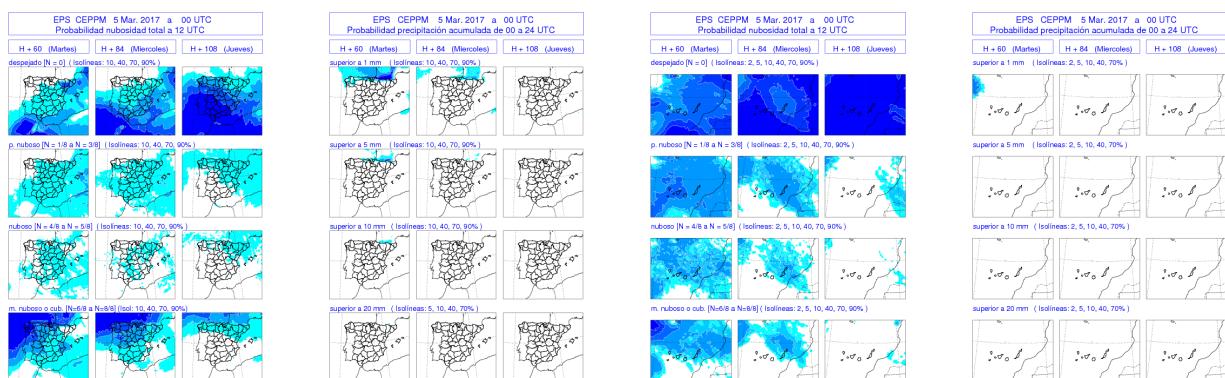


Figura 27.21: Probabilidades de nubosidad y precipitación. AEMET 2017.

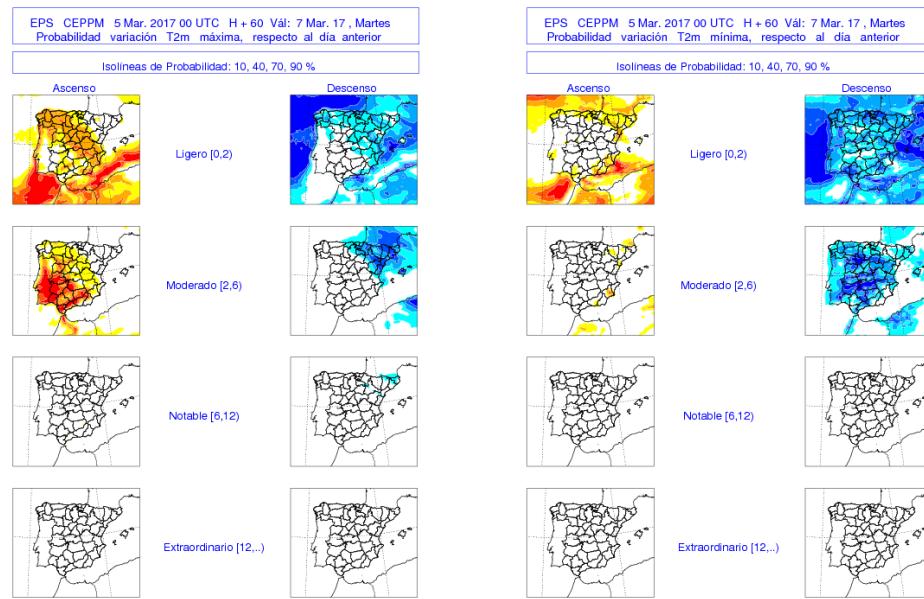


Figura 27.22: Probabilidades relacionadas con la tendencia de la temperatura (diferencia de la temperatura del día con la del día anterior). AEMET 2017.

En la Figura 27.23 se muestran otros mapas de probabilidad. De izquierda a derecha: «superación por encima» de distintos umbrales de temperatura, «rebasamiento por debajo» de umbrales de temperatura, viento y rachas de viento. Por sencillez, no se muestran otros mapas de probabilidad, e. g. convective available potential energy (CAPE), nieve acumulada en 24 horas, altura de olas de mar de viento, etc. Pueden dibujarse, y se dibujan, innumerables mapas de probabilidad, para muchas variables meteorológicas

diferentes y, para cada una de ellas, distintos umbrales, alcances de predicción y dominios geográficos. Una batería de información útil. La clave de la labor del predictor o del usuario experimentado es filtrar, entre la enorme cantidad de información, aquella que le puede ser más útil en cada situación. Se trata de algo similar a lo que se dice en la *frase del fontanero*, que «no cobra por apretar un grifo con una llave, sino por saber dónde hay que apretar».

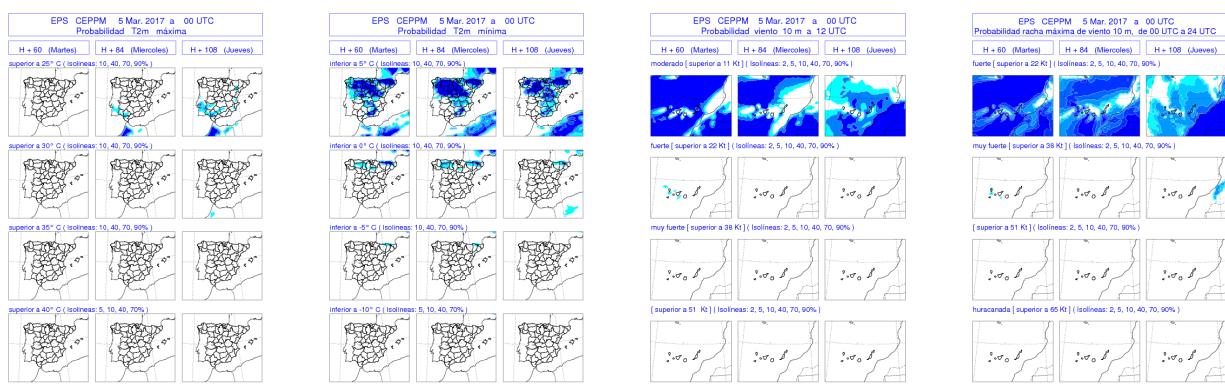


Figura 27.23: Otras probabilidades. AEMET 2017.

27.7 Predicción del riesgo

Los fenómenos *extremos* no siempre son *adversos*, aunque los fenómenos *adversos* suelen ser *extremos* [4]. Un fenómeno extremo es un fenómeno raro según un cierto criterio meteorológico de normalidad, mientras que un fenómeno adverso es un fenómeno con consecuencias potencialmente dañinas para la sociedad y, para valorar al fenómeno como adverso, el criterio es de carácter social, económico, etc. La idea de *extremo* o *anormal* como opuesto a *normal* es una idea muy intuitiva en meteorología de andar por casa, y se pueden oír frases como *este mes es más cálido de lo normal, estas lluvias han sido extremas*, etc. Pero entrando en detalles hace falta cierta precisión, pues una cosa es la idea intuitiva y otra es cómo sacarle punta a los SPC para tener una idea cuantitativa de *normalidad* o *anormalidad*. Las probabilidades en sí mismas no son indicativas de anormalidad; e. g. que la probabilidad de 20 mm de lluvia en 24 horas en un día de octubre sea del 40%, en Gandía no es un fenómeno anormal, pero sí lo sería en Almagro. Para poder hablar de *anormalidad* necesitamos dos ingredientes: primero, la predicción en sí misma, y segundo, un *historial de predicciones* que suele llamarse *climatología del modelo o del SPC*. De ese modo y, con técnicas matemáticas muy sofisticadas, se puede calcular si una predicción (e. g. para mañana) corresponde a un tiempo *normal* o si, por el contrario, puede catalogarse como de *anormal*, o incluso de *extrema*. En ese último caso, nos vendría de perlas saber *cuán extrema es* de algún modo. También se puede conocer esa información.

27.7.1 Climatología del ECENS (M-climate)

Un primer paso esencial es definir los márgenes de normalidad. Para ello no se toma la distribución climatológica de las observaciones, por presentar valores demasiado extremos que restan utilidad a las herramientas que se diseñan, que no suelen presentar señal [11]. En su lugar, se construye la climatología del modelo o del SPC, denominada *model climate* o *M-climate* en el ECMWF. Es función del alcance de predicción, para tener en cuenta derivas del modelo. Se utiliza un SPC especial denominado repredicción: se repredice el tiempo a 15 días vista, con fecha de partida la fecha de hoy (e. g. 15 de julio), del año pasado, del anterior y así los últimos 20 años; es decir, 20 repredicciones. La repredicción se ejecuta dos veces a la semana, lunes y jueves (2017), con 1 miembro de control y 10 miembros perturbados cada pasada. Usando 9 pasadas de esta repredicción, centradas en la fecha actual, cubrimos un abanico de unas 5 semanas. Juntando las 9 pasadas de los últimos 20 años, cada una con 11 miembros, tendremos un abanico de $9 \times 20 \times 11 = 1980$ valores previstos para definir los márgenes de tiempo representativo de «lo que es normal en esta época del año», quitando el posible ruido muestral que da la variabilidad diaria. Se usa la misma climatología, basada en la pasada de 00 UTC, para los productos de las pasadas de 00 UTC y 12 UTC para evitar inconsistencias (la pasada de 12 UTC es más antigua, etc. [23]). Esta metodología puede extenderse para los extremos hidrológicos [12, 13].

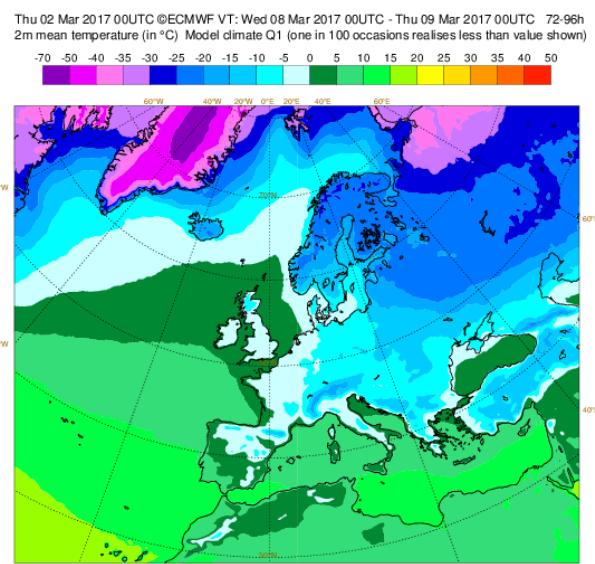
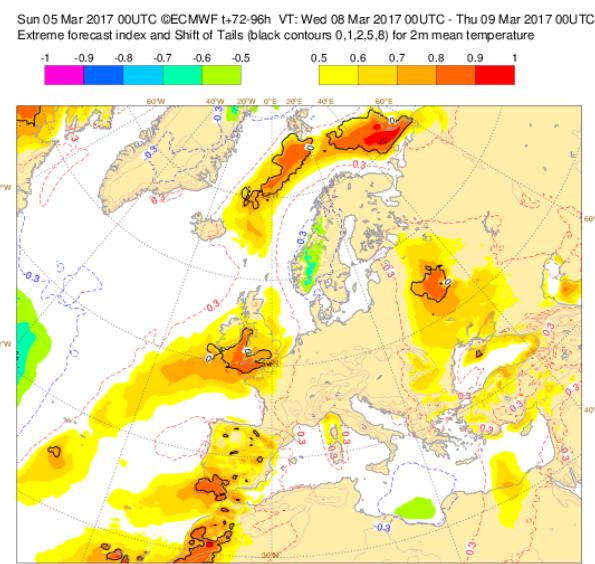


Figura 27.24: Índice de predicción extrema o Extreme Forecast Index (EFI). ECMWF 2017.

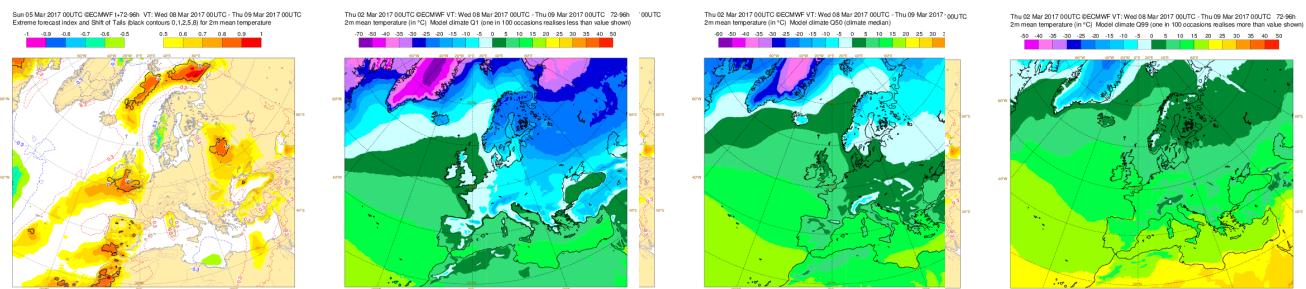


Figura 27.25: EFI para temperatura con el percentil climatológico Q_1 (izquierda) y recorte de otros dos percentiles climatológicos, Q_{50} (centro) y Q_{99} (derecha). ECMWF 2017.

27.7.2 Índice de predicción extrema (EFI)

Las siglas EFI significan Extreme Forecast Index (Índice de predicción extrema)[11]. El EFI es una cantidad que puede calcularse usando el SPC y su climatología (llamada la *climatología del modelo* o, para ser más precisos, del SPC) que no es otra cosa que el historial de predicciones. Para un lugar determinado, se compara la predicción probabilista (la famosa PDF, sec. 13.6 en la página 176) con ese historial, y se calcula un número, un índice, que dice *cuán extrema* es la predicción.

Extreme forecast index (EFI). Es una medida integral de la diferencia entre la distribución del SPC y la distribución de la climatología de dicho SPC (llamada a menudo climatología del modelo o *M-climate*).

$$EFI = \frac{2}{\pi} \int_0^1 \frac{p - F_f(p)}{\sqrt{p(1-p)}} dp \quad (27.1)$$

Así, se da una idea de *anormalidad* de la situación prevista sin necesidad de definir umbrales (ni en el espacio ni el tiempo). Si todos los miembros del SPC están por encima del máximo climatológico, entonces $EFI=+1$, si todos están por debajo del mínimo climatológico, entonces $EFI=-1$. Cuando más cerca de estos dos extremos, más *inusual* puede considerarse el tiempo previsto; cuando más cerca de 0, más *normal*.

El ECMWF muestra en su web (producto restringido) los parámetros EFI/SOT y la correspondiente climatología. En la Figura 27.24 en la página anterior podemos ver en el panel izquierdo el EFI coloreado en tonos amarillo/rojos para valores altos y en tonos fríos para valores bajos. Se pinta también el SOT (explica-

do más adelante en este mismo apartado) contorneado con isolíneas negras, así como la orografía [22]. En el panel derecho de la misma figura, el ECMWF presenta distintos percentiles de la climatología, para que el usuario pueda contrastar los valores climatológicos potencialmente extremos. En la Figura 27.25 vemos EFI para temperatura con distintos percentiles climatológicos. En la Figura 27.27 en la página siguiente se muestran EFI para viento y para precipitación acumulada.

EFI combinado. Para ofrecer ideas generales sobre potenciales fenómenos extremos de las variables de tiempo presente «todo en uno», se pueden dibujar distintos EFI diferenciando entre variables, usando distintos símbolos y colores. Símbolos rosas para el viento, símbolos verdes para la precipitación, sombra amarillo-naranja para temperaturas cálidas y celeste-azul para las frías. Para ayudar a la interpretación se superponen isohipsas de altura geopotencial en 1000 hPa (algo similar a la presión). Ver Figura 27.26.

Anomalous weather predicted by EPS:Sunday 05 March 2017 0000 UTC
1000 hPa Z ensemble mean (Wednesday 08 March 2017 1200 UTC)
and EFI values for Total precipitation,maximum 10m wind gust and mean 2m temperature (all 24h)
valid for 24hours from Wednesday 08 March 2017 0000 UTC to Thursday 09 March 2017 0000 UTC

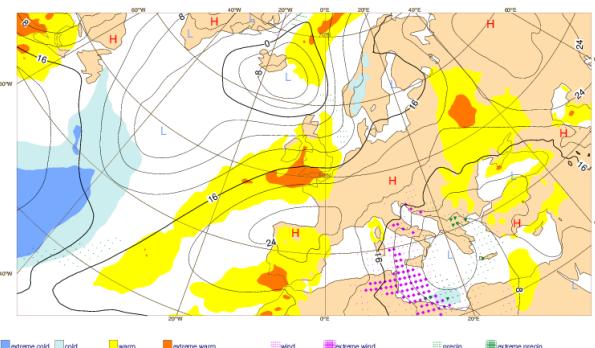


Figura 27.26: EFI combinado. ECMWF 2017.

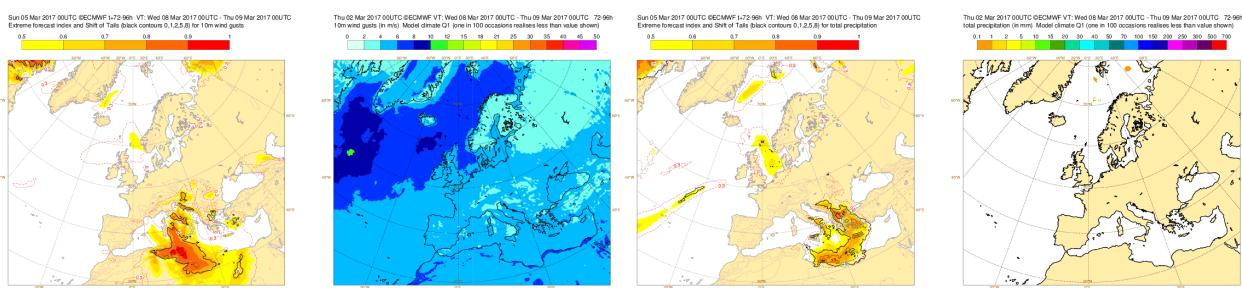


Figura 27.27: EFI para rachas de viento y para precipitación. ECMWF 2017.

27.7.3 Desplazamiento de cola SOT

Del inglés Shift of Tail (desplazamiento de cola) el *SOT* [10], es un índice complementario del *EFI*, proporcionando información sobre cuán extremo un evento extremo podría ser, dado un punto de referencia que es el percentil X-ésimo en la cola de la distribución.

Shift of tail (SOT). Fijando $X=90$, se compara la cola de la distribución del SPC con la cola de la distribución del M-clima. Para la parte alta de la cola X se computa siguiendo la relación:

$$SOT_{90} = \frac{Q_p(90) - Q_c(99)}{Q_c(99) - Q_c(90)} \quad (27.2)$$

donde $Q_f(90)$ es el percentil 90 de la distribución prevista del SPC, mientras que $Q_c(90)$ y $Q_c(99)$ son los percentiles 90 y 99 de la distribución de la climatología del ECENS.

Para las variables de temperatura se computa el SOT también para la cola inferior, tomando en ese caso $X=10$. Nótese la diferencia con el EFI, porque el SOT se computa para dar valores positivos también en la cola de abajo. La elección de 10 y 90 es relativamente arbitraria. El SOT 90 se contornea en trazo sólido y el SOT 10 en fino discontinuo. El SOT no está definido cuando $Q_c(90)=Q_c(99)$, lo cual conduce a mapas ruidosos e. g. de precipitación en regiones áridas y, para la nieve, cuando sucede ésto se fija SOT=-1 por consistencia. Valores positivos de SOT90 (SOT10) indican que las colas de la distribución prevista están por encima (debajo) del percentil 99 (percentil 1) de la climatología del ECENS. SOT=0 indica que el 10% de los miembros (5 de 51) cae por debajo de uno de estos percentiles. Valores EFI cerca de +1 y de -1

junto con valores positivos de SOT significan que se espera un evento muy inusual, más extremo cuanto más alto el SOT. Aunque teóricamente el SOT puede adquirir valores infinitos, se restringe gráficamente a un valor máximo de 8-10, siendo muy poco plausible que el percentil 90 de la distribución prevista por el ECENS exceda en más de 10 veces la distancia entre los percentiles 90 y 99 de la climatología.

Con respecto al EFI, se dibujan valores positivos y negativos para temperatura y *altura significativa* de ola y solamente positivos para el resto de variables, lo que concuerda conceptualmente con lo que en predicción se considera extremo o no. Como excepción tenemos la precipitación para los períodos 1-10, 1-15 y 10-15 días, donde se computan y muestran valores negativos, resultando enormemente útil para indicar períodos anormalmente secos. Nótese que, de todos modos, en zonas áridas precipitación nula no es un valor extremo. Para el período 10-15 días se dibujan también isolíneas de EFI de valores +0,1 y -0,1 pues pueden ser útiles aunque presenten señales pequeñas para alcances tan altos.

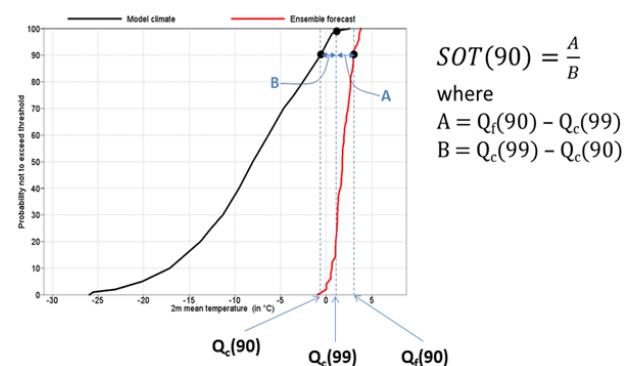


Figura 27.28: Desplazamiento de cola o Shift of Tail (SOT). ECMWF 2017.

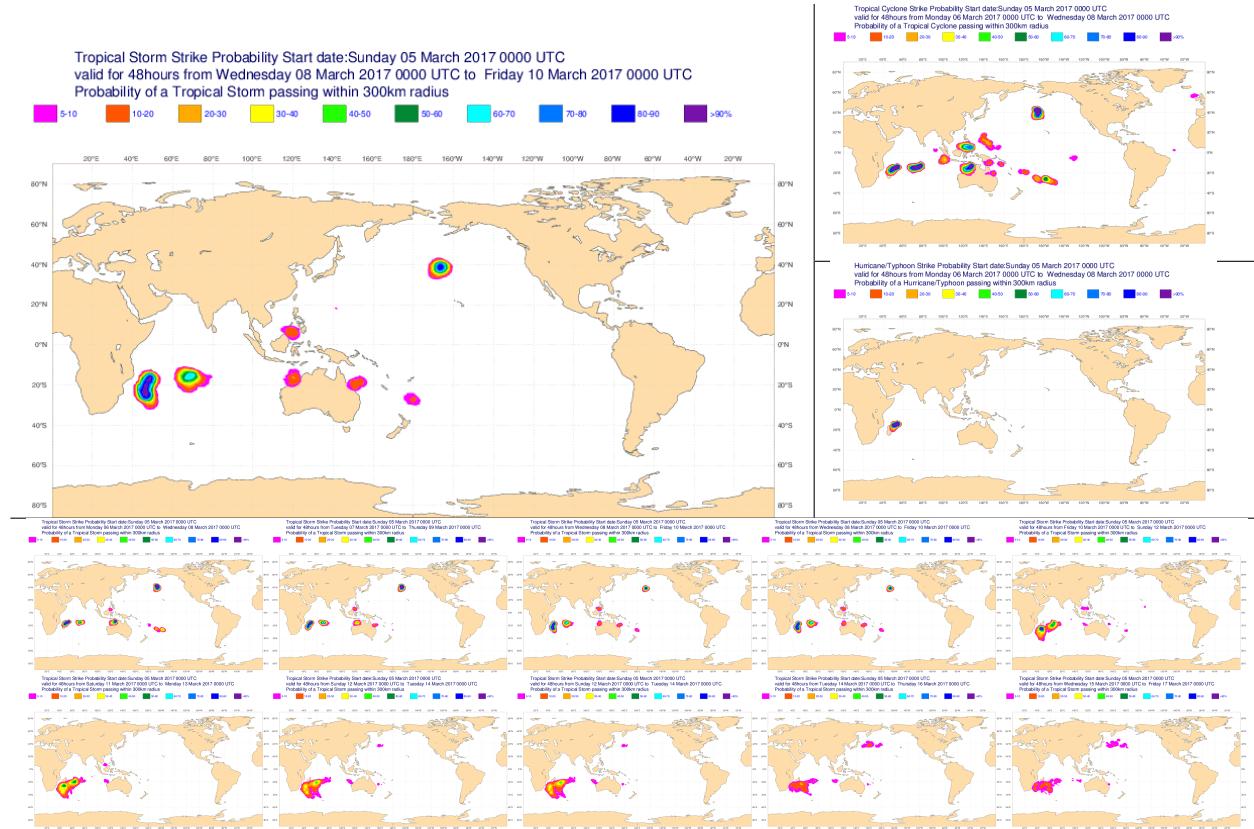


Figura 27.29: Mapas de estructuras tropicales. Arriba a la izquierda: probabilidad de encontrar tormentas tropicales. Arriba a la derecha, encima: probabilidad de encontrar ciclones tropicales. Arriba a la derecha, debajo: probabilidad de encontrar huracanes. Abajo: secuencia de probabilidad de encontrar tormentas tropicales desde el D+0 hasta el D+9. ECMWF 2017.

27.7.4 Estructuras tropicales

En el sitio web del ECMWF puede verse un apartado dedicado a la actividad de (a veces potenciales) ciclones tropicales. Se muestran, por un lado, mapas de probabilidad de impacto de ciclón o *strike probability*, basados en el número de miembros del ECENS que pasan por cada lugar en la previsión. En particular, se calcula la probabilidad de que pasen dentro de un radio de 300 km de la localización y en una ventana de tiempo de 48 h. Incorporando esta incertidumbre, con estos mapas puede evaluarse con eficacia el riesgo en diferentes áreas. Se generan estos mapas para tres categorías diferentes: ciclón tropical (vientos superiores a 8 m/s), tormentas tropicales (17 m/s) y huracanes-tifones (32 m/s). Se dispone de mapas globales, así como de 7 subregiones de interés. Se acompaña información de verificación. Ver Figura 27.29

Como botón de muestra, en las Figuras 27.30, 27.31 y 27.32 en la página siguiente mostramos algunos ma-

pas y gráficos previstos de la depresión extratropical Ophelia, ex huracán emigrado al NE desde los trópicos que pasó cerca de España en octubre de 2017, para comprobar la tremenda utilidad de estas herramientas de predicción probabilista.

Detección de ciclones. La identificación de ciclones en el proceso de predicción utiliza métodos objetivos computacionales [20]. El algoritmo busca un máximo de vorticidad en 850 hPa y localiza el mínimo local de presión en superficie más cercano. Se busca un núcleo cálido en la troposfera alta cerca del centro del ciclón para descartar otros fenómenos que no sean ciclones tropicales. Cuando se localiza un candidato en el proceso, se comprueba su continuidad en alcances sucesivos y se asigna una categoría de ciclón, tormenta o huracán, según los umbrales típicos de la escala Saffir-Simpson. Un algoritmo similar, adaptado para el Mediterráneo, se aplica en la detección temprana de *medicanes* (cap. 33 en la página 551).

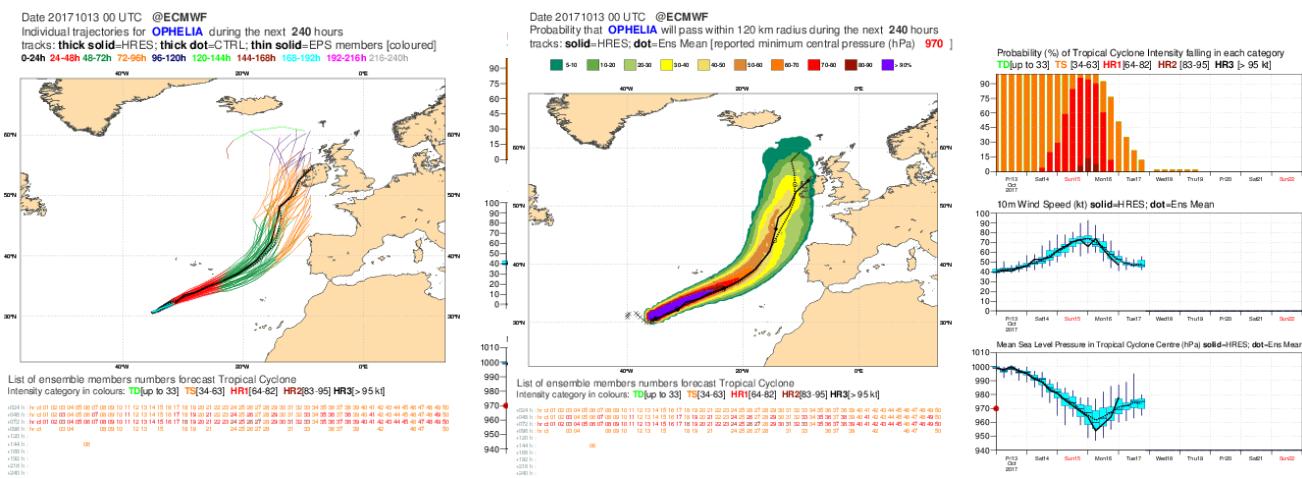


Figura 27.30: *Ophelia* a su paso cerca de la península ibérica. Izquierda: penachos. Derecha: probabilidades de impacto, junto con verificaciones. ECMWF 2017.

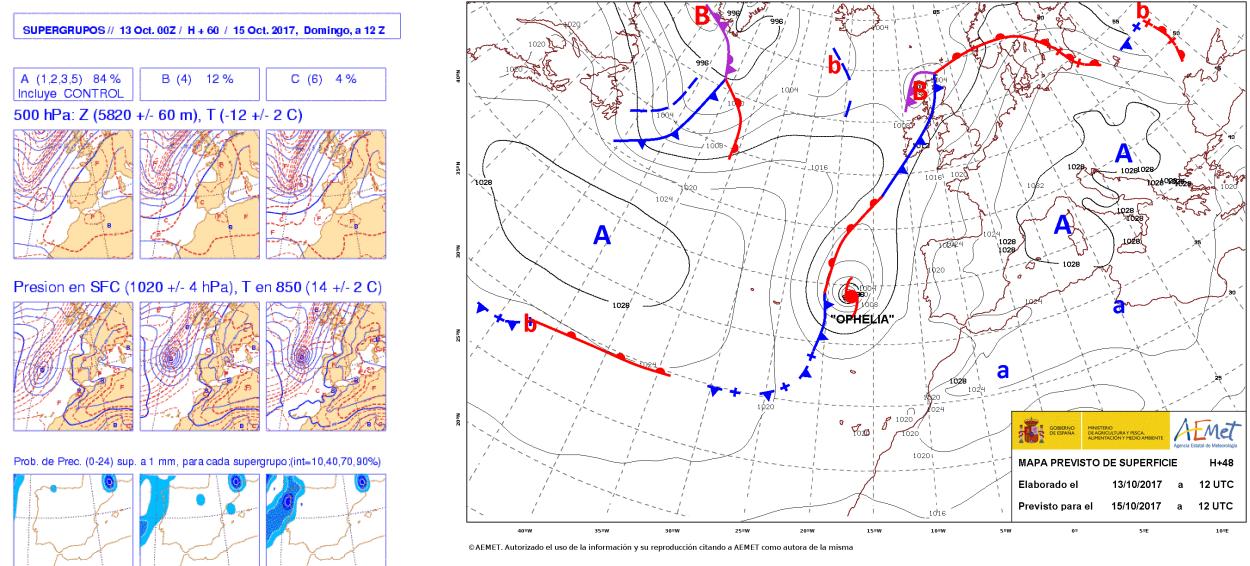


Figura 27.31: *Ophelia* a su paso cerca de la península ibérica. Izquierda: supergrupos para el 15-10-2017. Derecha: mapa previsto de superficie para esa misma fecha. AEMET 2017.

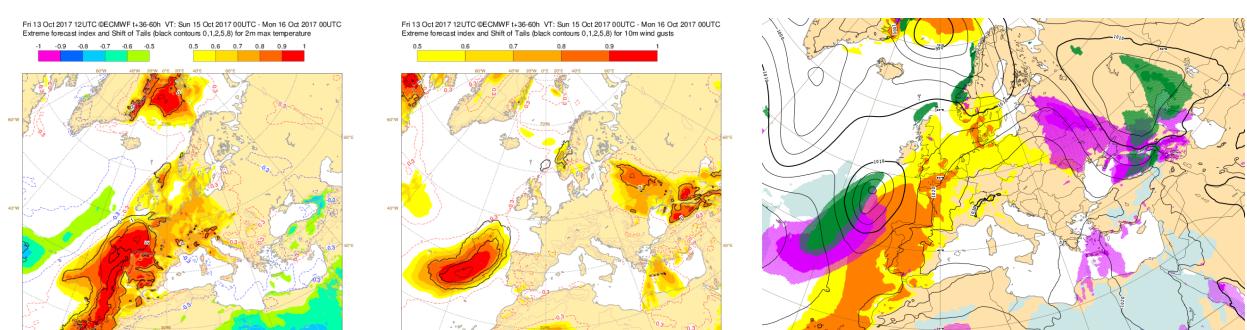
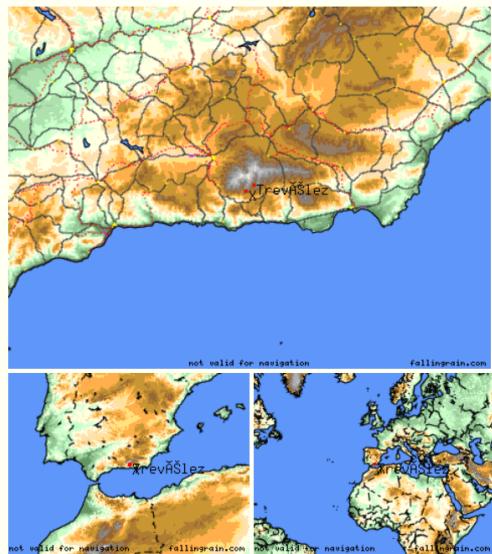


Figura 27.32: *Ophelia* a su paso cerca de la península ibérica, previsiones para el 15-10-2017. Izquierda: EFI de temperatura. Centro: EFI de viento. Derecha: EFI combinado. ECMWF 2017.

Trevélez

The grid point is located 8 km south-west of the required location.



Warning: a point being shown in the sea in the above map is not necessarily a sea point in the orography field used by the model. When a sea point is used to produce an EPS Metogram, the mention "(Sea Point)" is put in its title.

Note: the red marker with the place name (or a latitude value) denotes the chosen location; the other red marker, with the black cross, shows the location of the EPS grid point. If only one red marker is visible the chosen location is very close to the grid point.

Figura 27.33: Punto de malla más cercano a un punto geográfico de interés de localización arbitraria. ECMWF.

27.8 Productos en localizaciones

Prácticamente todos los productos vistos hasta ahora son mapas geográficos en los que se dibujan campos, normalmente probabilistas, derivados de algún ensemble o SPC. Para dibujarlos se utiliza algún software que toma los datos originales del SPC en sus puntos de malla y se realiza alguna computación en cada punto de malla. Esta computación toma los N valores correspondientes a N miembros en ese punto de malla. Ahora vamos algo más allá. Cuando nos interesa descender a la meteorología del punto geográfico, e. g. la famosa predicción de [AEMET por localidades](#), aparece una problemática previa, dado que el punto de interés no tiene por qué coincidir con algún punto de malla del SPC, sino que suele estar en el espacio intermedio entre cuatro puntos de malla (Figura 27.33).

¿Cómo obtenemos, en esos casos, los N valores previstos correspondientes al SPC en ese punto geográfico, habiendo una distancia hasta los puntos de malla? La técnica se llama *interpolación*. Esta técnica consiste en asumir que la variación espacial es, bien lineal, o bien una función conocida y calcular el valor o valores en el punto intermedio de interés. La interpolación tiene su propias dificultades:

La interpolación horizontal, en sí misma, puede dar valores meteorológicamente poco representativos si el punto de interés es de tierra y los puntos de malla más cercanos son de mar. Para ello deben seleccionarse puntos cercanos afines, si los hay.

La interpolación vertical es aún más delicada. El punto de interés tendrá a veces una altitud bien diferente de la de los puntos de malla circundantes. La variabilidad vertical de los parámetros meteorológicos es mucho mayor que la horizontal (sec. 6.3 en la página 77), de modo que una interpolación con resultados realistas debe hacer una corrección de los parámetros en la vertical, sobre todo de la temperatura, que es de los más sensibles. En el ejemplo dado, el punto de malla más cercano a Trevélez está 8 km al ESE. Al ser zona de montaña, con relieves muy acusados, la interpolación puede dar a menudo resultados poco realistas.

La orografía del modelo o del SPC puede diferir significativamente de la altitud física de la localización. El impacto de esta diferencia se corrige para la temperatura, ajustando 6.5 °C cada 1000 m, gradiente estándar de la [Organización de Aviación Civil Internacional \(OACI\)](#). En los productos se especifican las altitudes, tanto física como del modelo. En el caso de situaciones de inversión térmica, típicas en patrones anticiclónicos, las predicciones locales de temperatura deben tomarse con extrema precaución.

En esta línea, el aumento de resolución es beneficioso. El ECENS de 2017 ofrece sus salidas a una resolución horizontal de unos 18 km en nuestras latitudes, mientras que la versión anterior lo hacía a 32 km. Por su lado, ECHRES ha aumentado de 16 km a 9 km. En versiones aún anteriores, en la isla de Tenerife el Pico del Teide (3 718 m) quedaba circundado por cuatro puntos de malla marinos relativamente cercanos, por lo que era imposible obtener interpolaciones realistas. En los algoritmos actuales (2017), para la generación de cada producto local se consideran, en principio, los cuatro puntos de malla más cercanos a la localización específica. Para variables atmosféricas y, si estamos en tierra, de los cuatro puntos de malla circundantes se selecciona únicamente el punto de malla terrestre más cercano o, si sólo hay puntos marinos disponibles, entonces se toma el más cercano y se avisa de la contingencia en el título del producto. Por otro lado, para el oleaje tendríamos algo similar pero con puntos marinos aunque, si no hay puntos marinos circundantes, entonces el oleaje no se dibuja.

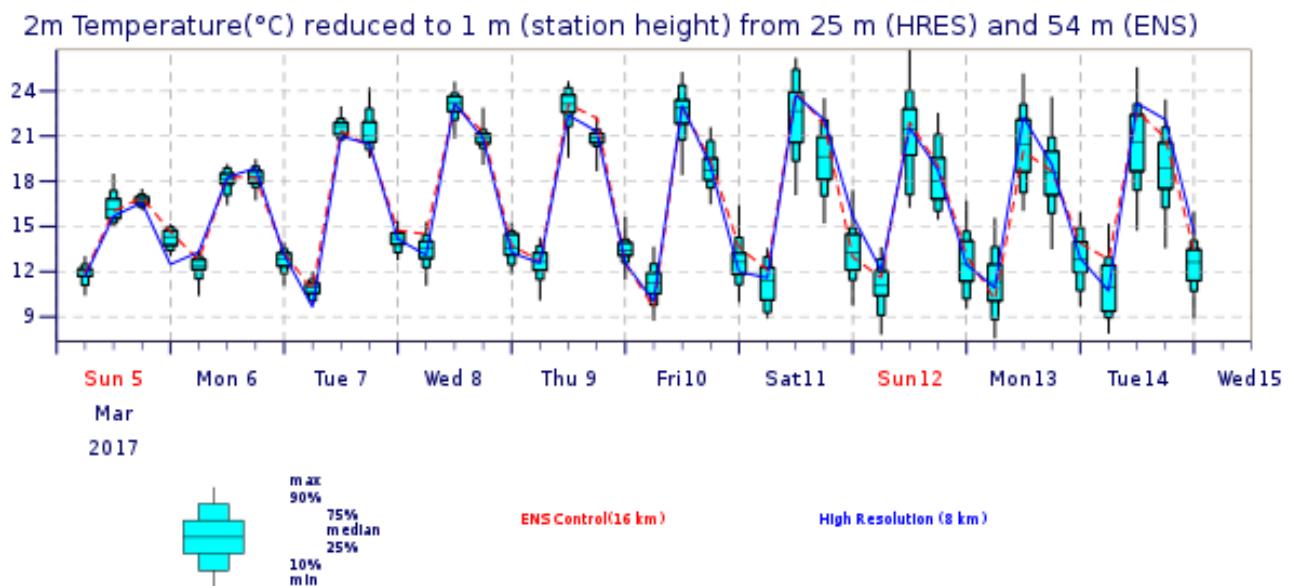


Figura 27.34: Box-plots de temperatura cada 6 horas para Huelva. ECMWF 2017.

27.8.1 EPSgramas

Los EPSgramas o meteogramas de SPC (mantendremos el término EPSgramas a lo largo del libro, por consistencia con la bibliografía más extendida) muestran información probabilista de la predicción para un cierto tiempo de una o varias variables meteorológicas en una localización específica. Son una extensión natural de los meteogramas tradicionales, pero construidos con información de los SPC. Los meteogramas tradicionales son gráficos en los que en el eje X se representan los días, con sus horas y en el eje Y se representa una variable meteorológica, que puede ser el viento, la temperatura, la humedad, etc. Los meteogramas tradicionales pueden mostrar la evolución en el tiempo de una variable meteorológica observada, o bien la previsión en el alcance de la predicción (tiempo futuro) de una variable meteorológica previsible. En el sitio web del ECMWF pueden encontrarse EPSgramas de varios tipos: de 10 días, de 10 días con oleaje, de 15 días, de 15 días con climatología, así como EPSgramas de EFI o de CDF (sec. 13.6.12 en la página 188 y sec. 27.7 en la página 419). Quizá los EPSgramas sean más populares para los usuarios finales que para la predicción operativa, aunque deben utilizarse con precaución.

Ingredientes generales. Los EPSgramas, al tener que incluir información de una SPC y no de un sólo modelo, son más complejos que los meteogramas tradicionales. Es decir, los EPSgramas representan información probabilista. La Figura 27.34 muestra un primer ejemplo sencillo de un EPSgrama de una sola variable,

la temperatura, con predicción probabilista para 10 días vista. Los ingredientes esenciales del EPSgrama son:

Título: variable meteorológica, en este caso temperatura a 2 m reducida a la altitud de la localización (extraída de una topografía global denominada GTOPO30 con 30 segundos de arco de resolución <https://lta.cr.usgs.gov/GTOPO30>) desde la altitud del modelo.

Eje X: 10 días divididos en intervalos de 6 h, por lo que hay cuatro muescas por día, correspondientes a las 00, 06, 12 y 18 UTC. Además de los días, se muestra debajo el mes y el año.

Eje Y: variable meteorológica, temperatura en este caso, por lo que en este caso en el eje está graduado en grados centígrados.

Curva azul: modelo determinista ECHRES (sec. 19.2 en la página 291). Al ser un solo modelo, la predicción es un solo valor y su evolución en el tiempo es una curva, que toma un valor distinto cada 6 h. La resolución horizontal se indica entre paréntesis, 9 km u 8 km, según los casos.

Curva roja: miembro de control del ECENS (sec. 19.3 en la página 293). Al ser un solo modelo, el de control, la predicción es un solo valor y su evolución en el tiempo es una curva, que toma un valor distinto cada 6 h. La resolución horizontal se indica entre paréntesis, 18 km o 16 km, según los casos.

Box-plots: miembros perturbados (50) del ECENS (sec. 19.3 en la página 293). Al ser 50 modelos,

la predicción es probabilista y su evolución en el tiempo se representa mediante box-plots, uno cada 6 h. Cada box-plot muestra de manera sintética la distribución de los 50 valores previstos: mediana, rango intercuartílico o IQR, 10-90 y min-max [21].

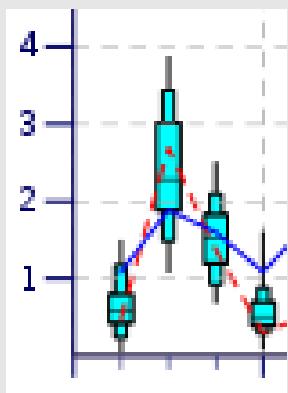


Figura 27.35: Cuatro box-plots consecutivos en el tiempo, uno cada seis horas. Cada box-plot muestra el IQR (caja central), la mediana (línea horizontal en su interior), el rango p10-p90 (cajas menores), el mínimo (línea vertical inferior) y el máximo (línea vertical superior). ECMWF 2017.

Box-plots o diagramas de caja. Los box-plots [21] ofrecen una representación sencilla, sintética y visual de una distribución de números, en términos matemáticos una PDF discreta. Hay diversos tipos de box-plots y, aunque nos centraremos en los que se muestran en el ECMWF, el lector interesado no tendrá problema en extrapolar las ideas para la gran variedad de EPSgramas en los diversos portales meteorológicos en Internet. El box-plot consta de una caja rectangular principal cuyo borde inferior es el percentil 25 de la distribución, el borde superior es el percentil 75 y una línea horizontal interior marca la mediana o percentil 50. Las líneas verticales inferior y superior tienen unos extremos que marcan el mínimo y el máximo respectivamente. Además, pueden incluirse dos pequeñas subcajitas, inferior y superior, que marcarían los percentiles 10 y 90 respectivamente. En la Figura 27.35 podemos ver los cuatro box-plots correspondientes a un día.

Epsgrama de 10 días. Las cuatro variables de tiempo sensible presentes en los EPSgramas usuales se distinguen con claridad. Como ejemplo de EPSgrama completo para ilustrar las diferencias y aspectos, la Figura 27.36 muestra un EPSgrama completo de

10 días para Huelva. La localización de Huelva es, aproximadamente, 37.26 °N 6.95 °W. El punto de malla más cercano del ECENS está localizado en 37.18°N 7.11°W, 16 km al suroeste de Huelva, en pleno océano.

Nubosidad. El valor mostrado es el valor instantáneo previsto para esa hora, medido en *octas*, fracciones múltiplo de 1/8 de cielo cubierto de nubes. Es la variable que con más dificultad resuelven los modelos y SPC y la que más dispersión presenta, con diferencia. A menudo puede observarse un box-plot de nubosidad que abarca todo el rango desde cielo despejado hasta cielo cubierto. Todo un reto para un predictor.

Precipitación. Se muestra la precipitación acumulada, en mm, en las 6 h precedentes a la hora marcada. Cuando hay valores muy altos en algún día, estos valores se indican con números pequeños, pues los box-plot se salen mucho de la escala. Matemáticamente hablando, la precipitación es definida positiva: sólo presenta valores positivos y por tanto en el EPSgrama su eje Y siempre va desde cero hasta algún otro valor. Para cada marca horaria (00, 06, 12, 18 UTC) la distribución de valores previstos de precipitación del SPC tiene casi siempre un valor mínimo de cero, su percentil 25 también puede ser cero, o, más allá, los percentiles 50 (mediana), 75 o, incluso, el máximo de la distribución pueden ser cero. Por tanto, el box-plot suele tocar fondo, está apoyado en el eje X y, a menudo, colapsa en el eje X. Aunque ésto no siempre es así.

Viento. Se presenta el valor instantáneo previsto para la hora marcada medido en m/s (1 m/s son 3.6 km/h). Presenta un ciclo diurno no tan marcado como el de la temperatura, pero sí claramente influido por factores climatológicos: en lugares con régimen de brisas y, en ausencia de perturbaciones sinópticas tales como borrascas, la brisa de mar, o virazón, suele ser tener más intensidad que la de tierra o terral y marca sus intensidades máximas en horas centrales del día.

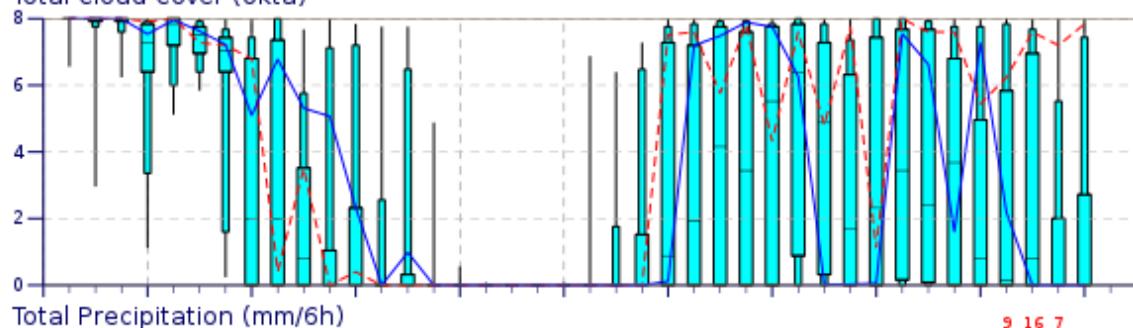
Temperatura. Se presenta el valor instantáneo previsto para la hora marcada medido en °C. Presenta, como se ha explicado previamente, un ciclo diurno y un ciclo estacional bastante claros, así como mayor dispersión durante el día que durante la noche.

ENS Meteogram

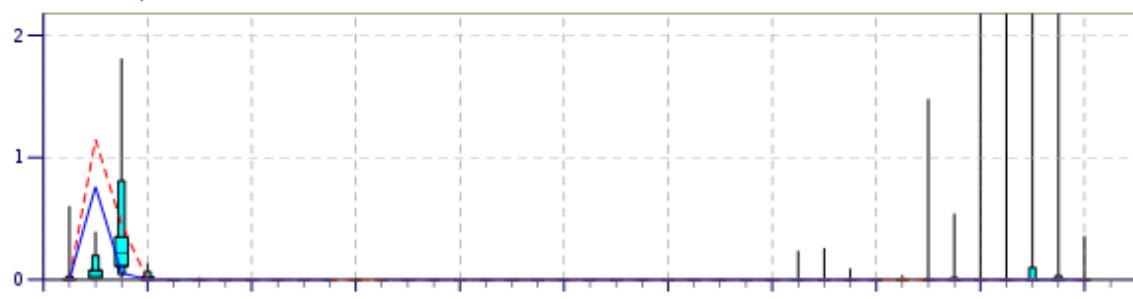
Huelva, Spain 37.32°N 6.89°W (ENS land point) 1 m

High Resolution Forecast and ENS Distribution Sunday 5 March 2017 00 UTC

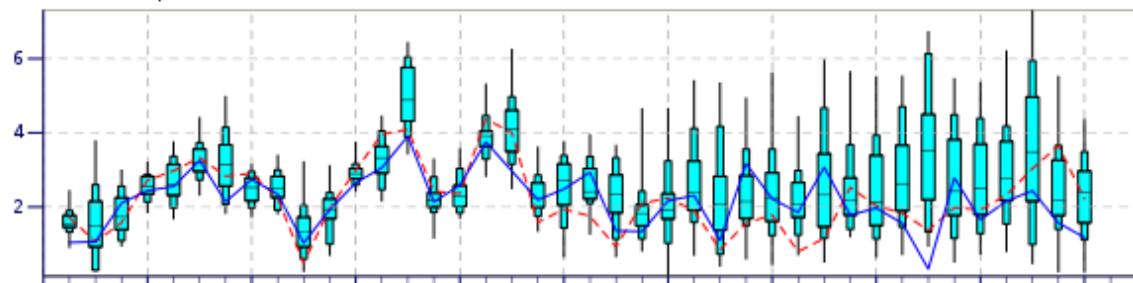
Total Cloud Cover (okta)



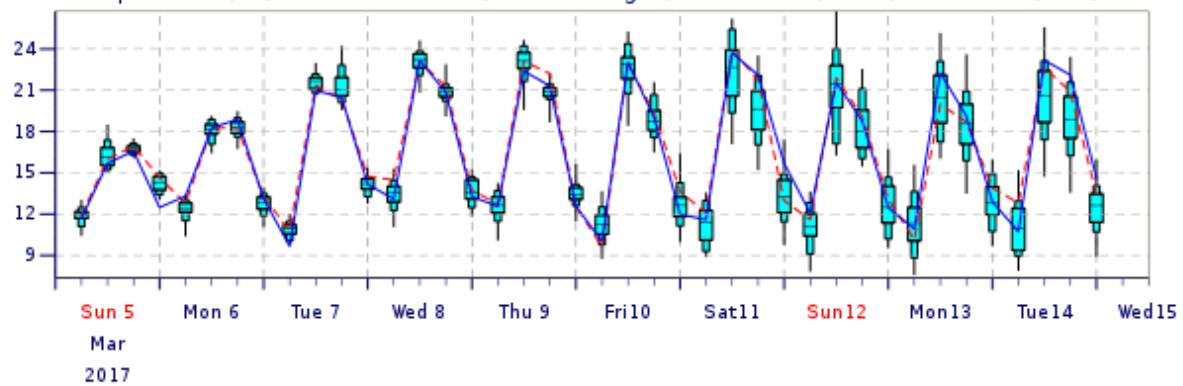
Total Precipitation (mm/6h)



10m Wind Speed (m/s)



2m Temperature(°C) reduced to 1 m (station height) from 25 m (HRES) and 54 m (ENS)



max
90%
75%
median
25%
10%
min

ENS Control(16 km)

High Resolution (8 km)

Figura 27.36: Epsgrama de 10 días para Huelva (ver texto). ECMWF 2017.

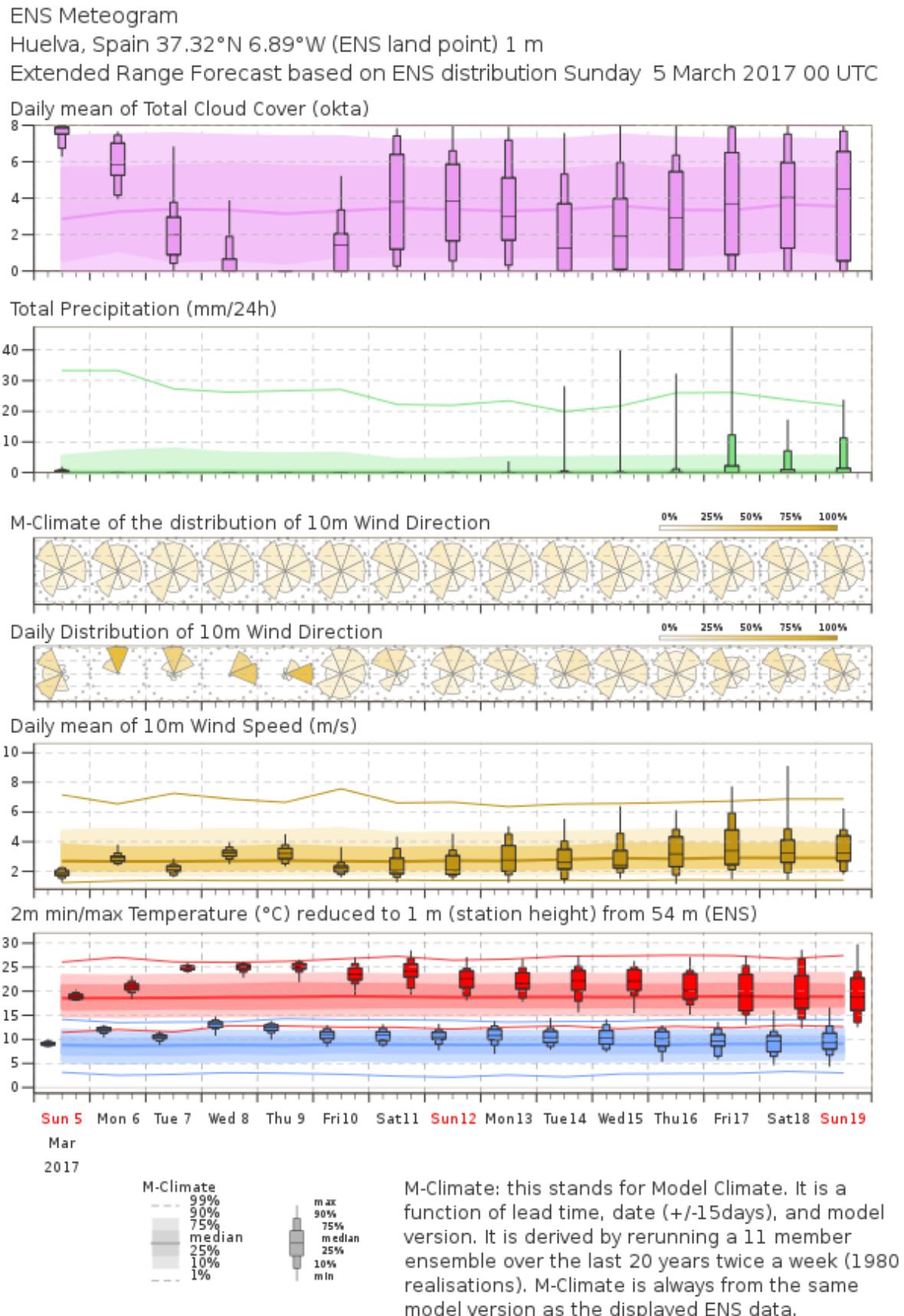


Figura 27.37: Epsgrama de 15 días con clima para Huelva (ver texto). ECMWF 2017.

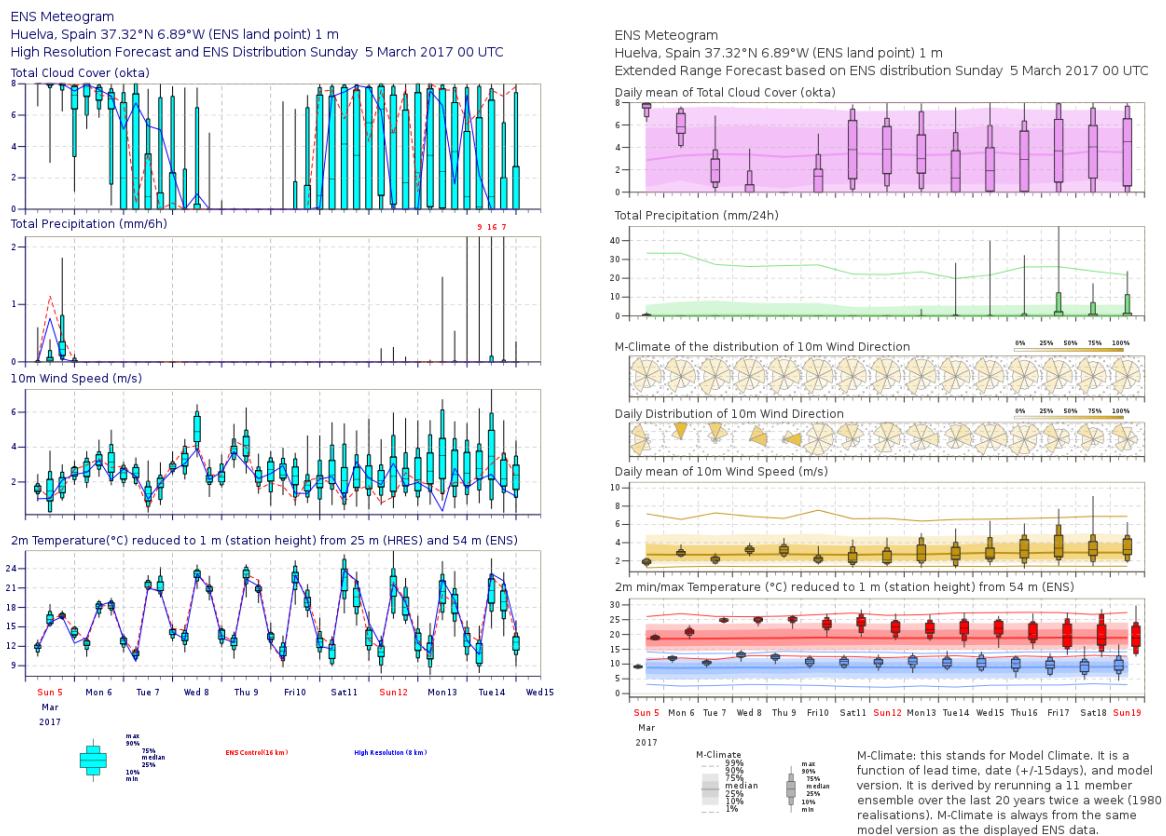


Figura 27.38: Epsgramas para Huelva, 10 días y 15 días con clima. ECMWF 2017.

Epsgramas de 15 días y de 15 días con clima. La Figura 27.37 muestra un EPSgramas de 15 días con clima para Huelva, para contrastar con el de 10 días. Los EPSgramas de 15 días presentan diferencias con respecto a los de 10 días:

- Amplían el alcance de predicción hasta el D+15.
- Muestran las cuatro variables antes mencionadas de tiempo sensible y añaden la dirección del viento.
- La nubosidad que se muestra es la media diaria (media de los valores en 00, 06, 12 y 18 UTC).
- La precipitación es acumulada en 24 h.
- La dirección del viento es la distribución diaria repartida en octantes (N, NE, E, SE, S, SO, O, NO). El área relativa de cada octante es proporcional a la probabilidad de esa dirección de viento. Cada día se reescaliza de acuerdo al octante más poblado, por tanto cada día tiene una escala diferente.
- La velocidad del viento son valores medios diarios.
- La temperatura se descompone en dos partes en la misma gráfica: temperatura máxima en rojo y temperatura mínima en azul, en °C.

Los EPSgramas con clima, además, añaden para cada gráfica unas áreas sombreadas que representan la climatología del ECENS de los últimos 20 años (sec-

ción 27.7.1 en la página 419), proveen así de información de normalidad climatológica, que permite estimar si los valores previstos son normales, superiores o inferiores a lo normal, además de poder inferir cuán significativa sería la diferencia. La línea gruesa central es la mediana, mientras que las líneas finas inferior y superior son los percentiles 1 y 99 respectivamente (es más útil que pintar los valores mínimo y máximo que a veces son muy extremos). La sombra más oscura es el IQR y la sombra más clara es el intervalo entre percentiles 10 y 90.

Aspectos importantes. En todo EPSgrama deben resaltarse los siguientes aspectos, potencialmente útiles para el proceso de predicción:

Interpolaciones. Todos los valores previstos representados, ECHRES, control y box-plot, son interpolaciones espaciales desde los puntos de malla hasta la localización en cuestión, por lo que deben tenerse en cuenta las limitaciones descritas en 27.8. Nótese que, en el título del EPSgrama se indica la altitud de la localización, la altitud del ECHRES y la altitud del ECENS para tener en cuenta este aspecto.

Dispersión creciente con el alcance de predicción.

La dispersión de las variables atmosféricas crece de forma natural con el alcance de la predicción (sec. 13.6.5 en la página 181 y sec. 27.5.1 en la página 412). Esto se refleja, como no podía ser de otra manera, en el EPSgrama. La dispersión está relacionada con el tamaño del box-plot, de su IQR, de su 10-90 y de su min-max. Podemos comprobar en el ejemplo que el tamaño de los box-plot crece gradualmente con el alcance de la predicción. No siempre la dispersión crece: en ocasiones un día con atmósfera estable sucede a otro día más complejo y, en esos casos, la dispersión puede decrecer. Se trata de una pauta general que tiene que ver con la predecibilidad: a más dispersión, menos predecibilidad y, aunque en general la dispersión crece con el alcance de la predicción, cuando no es así tenemos muchas pistas sobre el aumento de predecibilidad de un día para otro.

Dependencia estacional y horaria de la dispersión.

En el EPSgrama de la Figura 27.34 podemos ver cómo la dispersión correspondiente a las 00 y 06 UTC es, en general, inferior a la de las 12 y las 18 UTC. Se trata de otra pauta general que revela aspectos meteorológicos en el EPSgrama: la temperatura a 2 m es más predecible durante la noche en condiciones de atmósfera relativamente estable. Durante el día, el calentamiento solar induce actividad convec-

tiva, más inestabilidad, temperaturas más altas, etc. Por otro lado, también pueden observarse pautas estacionales en la dispersión, aunque más complejas pues tienen también que ver con el paso de perturbaciones.

Ciclos diurno y anual. Variables como la temperatura a 2 m muestran en el EPSgrama un ciclo diurno muy claro en el que las temperaturas, en general, son más bajas durante las horas nocturnas, en los entornos de 00 y 06 UTC. La temperatura mínima suele alcanzarse *circa* las 06 UTC y la máxima entre las 12 UTC y las 18 UTC. Así mismo muestran el ciclo anual, con temperaturas más altas en verano y primavera.

Factores del clima. los condicionamientos climáticos ambientales correspondientes a la localización en cuestión son decisivos a la hora de entender el EPSgrama: altitud, continentalidad, orientación de las laderas, etc. Por ejemplo, en un lugar costero (e. g. A Coruña) la oscilación térmica es suave en comparación con un emplazamiento continental (e. g. Granada); también puede haber diferencias entre un ambiente marino atlántico y uno mediterráneo (e. g. Tarragona). La Figura 27.39 ilustra este ejemplo.

Naturaleza de la variable. Cada variable meteorológica tiene su naturaleza propia: variabilidad, ciclos, dispersión, etc., que se analizan en otro apartado.

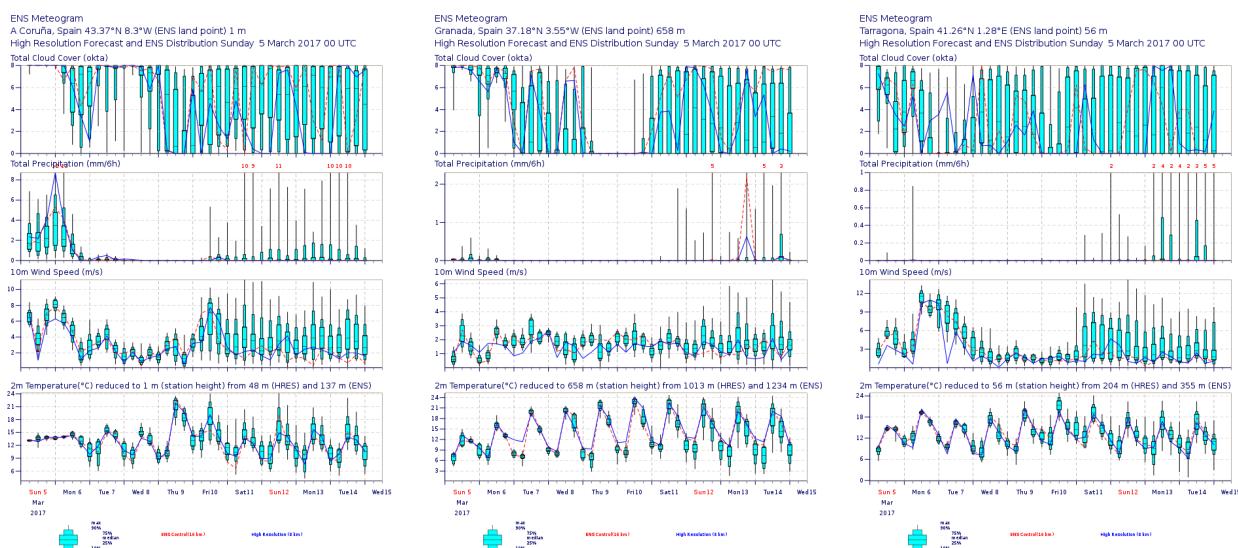


Figura 27.39: Epsgramas para A Coruña, Granada y Tarragona. ECMWF 2017.

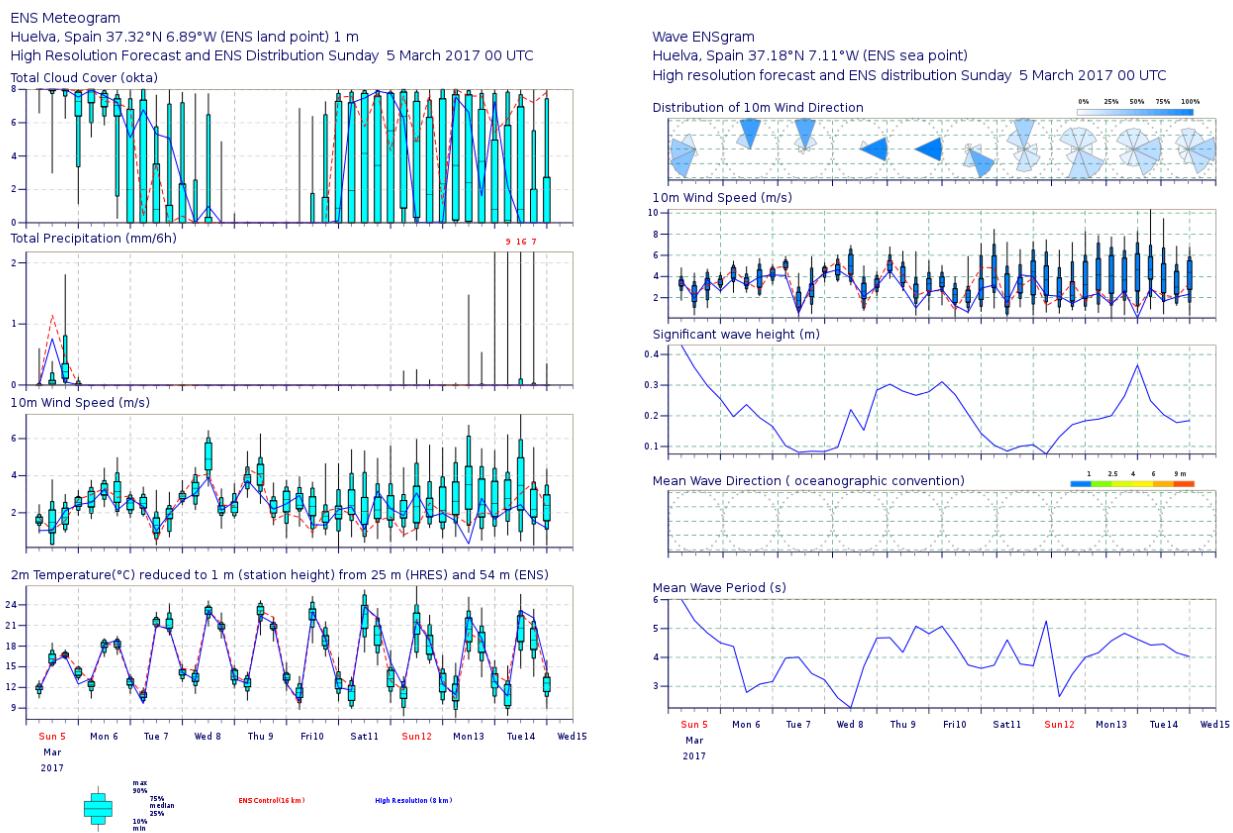


Figura 27.40: Epsgramas para Huelva: 10 días y oleaje. ECMWF 2017.

Epsgramas de 10 días de oleaje. También llamados ola-gramas (wave-grams) muestran variables de oleaje basadas en los WAM ENS y WAM HRES del ECMWF, con resoluciones de unos 28 km y unos 14 km respectivamente (2017), así como dirección y velocidad del viento basadas en el ECENS, a unos 18 km de resolución. Las variables específicas de oleaje son:

Altura significativa de ola. Se muestra el valor previsto para la hora marcada, estimado a partir del momento de orden cero del espectro de frecuencia del oleaje, representando la altura media del tercio superior del oleaje.

Dirección media de ola. Promedio pesado de la dirección de propagación en el espectro de oleaje, con notación oceanográfica, es decir, hacia dónde van las olas y no de dónde vienen, como es el caso del viento. La rosa se construye similarmente a la de viento, pero los colores denotan *altura significativa* de ola.

Periodo medio de ola. Derivado del momento de orden -1 del espectro de frecuencia del oleaje, medido en segundos. Se denomina a veces periodo de energía pues da más peso al mar de fondo de baja frecuencia, rico en energía, que al oleaje de más alta frecuencia.

27.8.2 Penachos (en extinción)

Los penachos son la representación gráfica de algunas variables meteorológicas, normalmente de tiempo sensible, en función del rango de predicción para un punto geográfico determinado. En cierto sentido, son una evolución natural de los clásicos meteogramas, incluyendo toda la avalancha de información que proporciona un SPC.

En una localización geográfica determinada dispondremos normalmente de N predicciones de una variable meteorológica como la temperatura, la precipitación, el viento, etc. Estos N valores, como ya se ha descrito (sec. 27.8 en la página 424) suelen calcularse a partir de los puntos de malla más cercanos, mediante interpolación horizontal y, en el mejor de los casos, teniendo en cuenta la variabilidad con la altitud. Esas N predicciones, correspondientes a N miembros de un SPC, evolucionarán con el alcance de la predicción, de modo que si dibujamos una curva de evolución para cada predicción, tendremos N curvas superpuestas en el gráfico. El miembro de control, si lo hay, suele dibujarse en otro color y, si existe un modelo determinista de mayor resolución, puede añadirse su curva usualmente en otro color distinto al de los miembros y al del control. Igual que en los sellos, para interpretar los penachos hay que considerar gran cantidad de información, y ahí está la dificultad para el predictor. Para localidades específicas y variables meteorológicas de especial interés, pueden servir para profundizar algunos aspectos de la predicción.

El ECMWF muestra productos de penachos para ciudades seleccionadas, con la evolución para 10 días de 3 variables meteorológicas: temperatura en 850 hPa, precipitación acumulada en las 6 horas previas a la marcada y altitud geopotencial en 500 hPa. Las curvas muestran la evolución de ECHRES y del miembro de control del ECENS, mientras que los penachos representan el conjunto de 50 miembros perturbados del ECENS. Las áreas sombreadas, para temperatura y altitud geopotencial solamente indican la probabilidad de que el valor caiga dentro de un cierto rango de valores. Por ejemplo, para la temperatura en 850 hPa la sombra verde clara informa de que la probabilidad de tener una temperatura entre $+0.5^{\circ}\text{C}$ y -0.5°C del valor indicado está entre 0.5 y 10%.

Los penachos ofrecen información sobre la incertidumbre en términos de tendencias generales proyectadas en la localización de interés. Las N curvas forman una nube de posibles valores que, a menudo, no necesitamos con detalle, sino más bien los rasgos generales. Normalmente esa nube de valores, para cada variable, se ensanchará con el aumento del alcance de predicción, correspondiendo a un aumento natural de la dispersión, indicando una disminución de la predecibilidad de la atmósfera para ese punto y rango de predicción.

Si bien esa es la tendencia usual, nada impide que la nube de puntos se estreche de un alcance de predicción al siguiente; es decir, la dispersión puede disminuir si el ambiente meteorológico favorece una disminución de incertidumbre y por ende un aumento de predecibilidad.

El ECMWF no tiene entre sus planes seguir adelante con este producto.

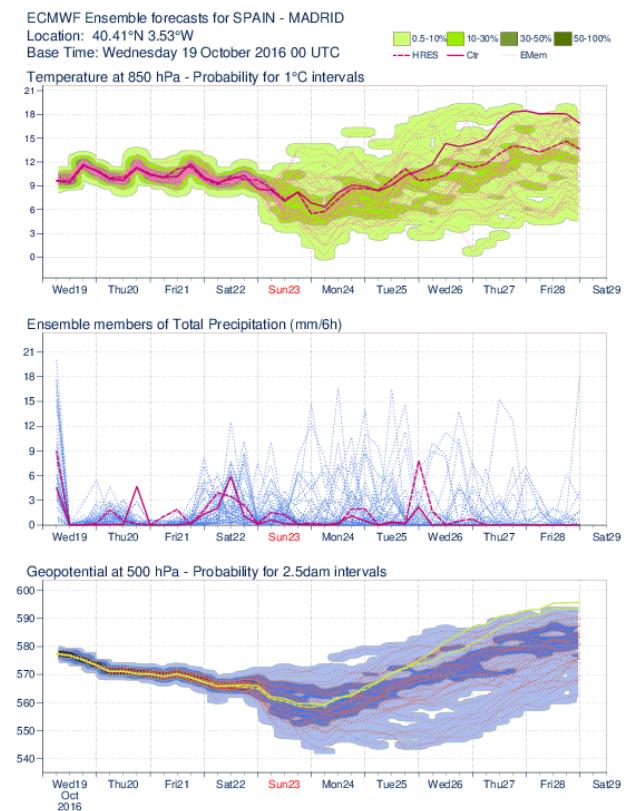


Figura 27.41: Penachos de temperatura, precipitación y geopotencial para un rango de 10 días vista. ECMWF 2017

27.9 Productos experimentales

Los productos existentes disponibles para la predicción operativa surgen de un complejo proceso, en el que se combinan la necesidad, la imaginación y el propio proceso evolutivo de la predicción operativa a medida que va asimilando en su concepción y en sus procedimientos la naturaleza probabilista. Los productos suelen pasar por una fase experimental en la que se va viendo su utilidad, interpretación, consistencia, etc. Un ejemplo de productos en fase experimental son aquellos que buscan de forma objetiva características atmosféricas y las plasman en mapas meteorológicos,

denominados genéricamente *productos orientados a características* (*feature oriented products*). Dentro de esa familia de productos, un ejemplo específico son los *mapas de frentes*, que dibujan en un mapa frentes (frios, cálidos, oclusiones, etc.) y otras zonas con alto gradiente térmico. Cuando se aplica esta técnica con SPC, tendremos «*espaguetis de frentes*», como podemos observar en la Figura 27.42. A medida que avanza el alcance de predicción, los frentes se van desdibujando y puede parecer una auténtica explosión en una fábrica de espaguetis. Podemos observar un frente pasando por el noroeste de la península ibérica en los mapas de la cuarta fila de la figura.

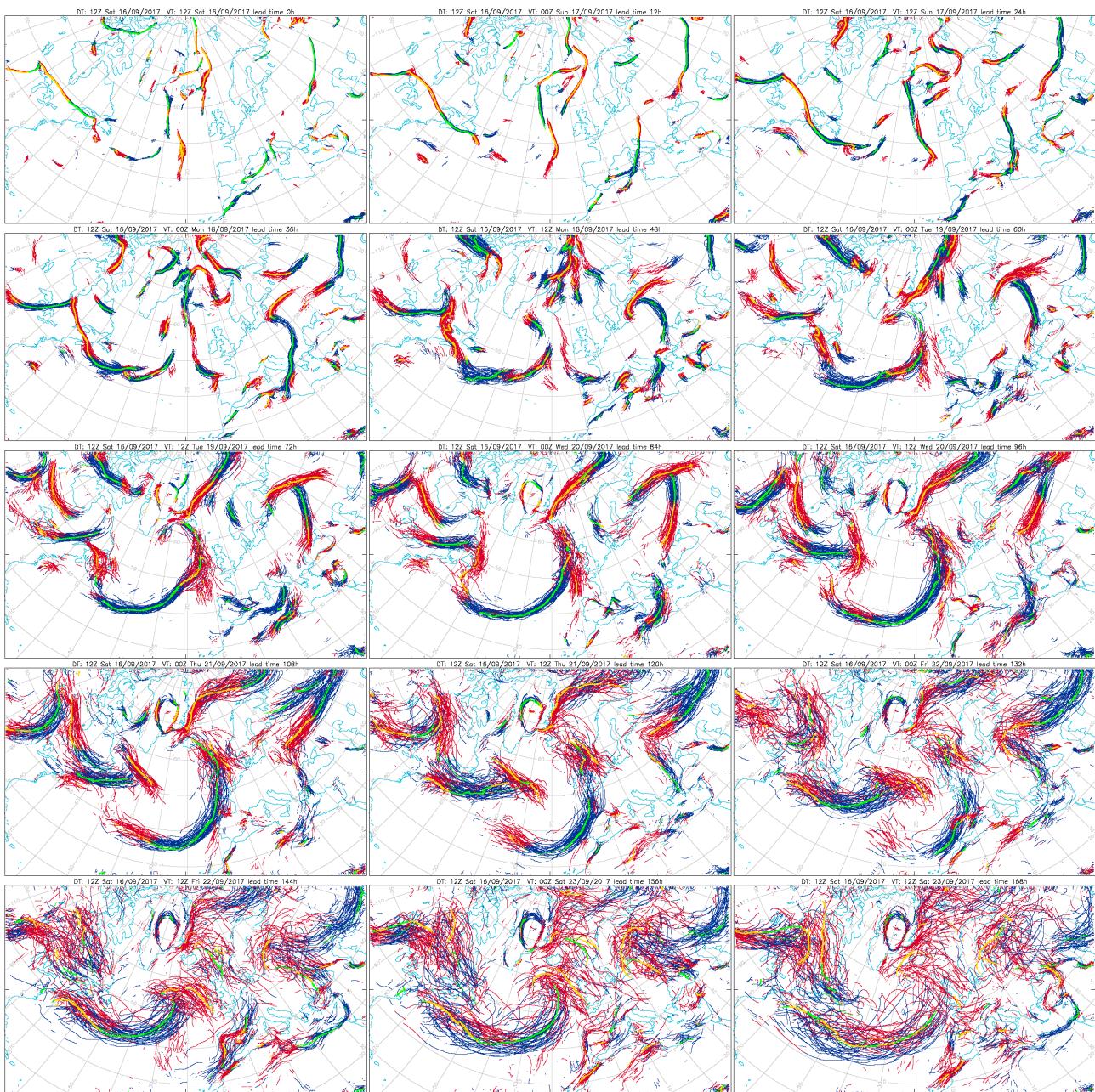


Figura 27.42: *Espaguetis de frentes*, producto experimental en 2017 (ver texto). ECMWF 2017.

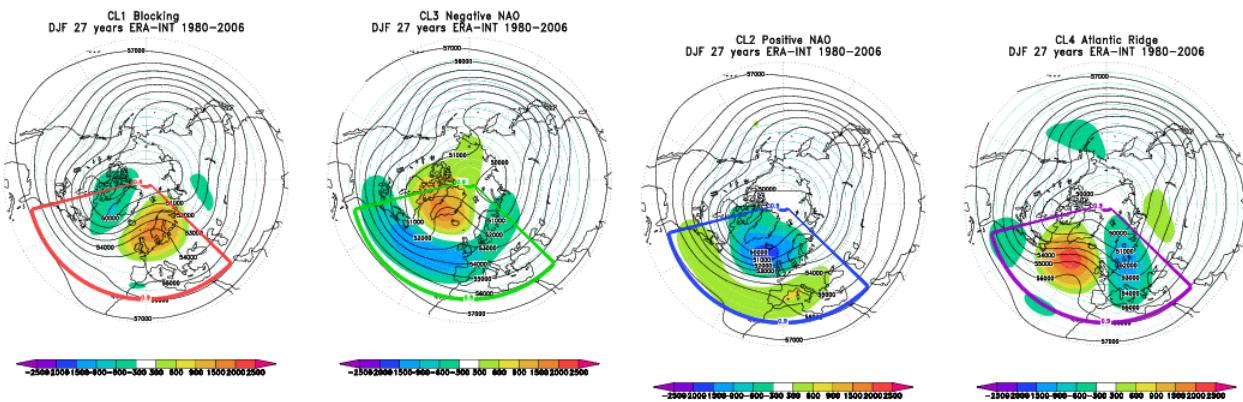


Figura 27.43: Predecibilidad dependiente del flujo (ver texto), ECMWF.

27.10 Predicción de la predecibilidad: predecibilidad dependiente del flujo

Tanto en el capítulo de predecibilidad (cap. 12 en la página 155) como en el de sistemas de predicción por conjuntos (cap. 13 en la página 165) o en este mismo capítulo se intenta enfatizar la idea de que la predecibilidad depende de la situación atmosférica: *predecibilidad dependiente del flujo*. Los SPC intentan muestrear convenientemente las diversas incertidumbres asociadas a la atmósfera y, de este modo, la dispersión que presente la PDF del SPC correspondiente ofrecerá una información cuantitativa, detallada y explícita de incertidumbre atmosférica y, por ende, de predecibilidad. Los sistemas de predicción modernos tienden a incorporar la información sobre predecibilidad de modo natural en las predicciones. En los boletines meteorológicos que AEMET proporciona a los medios de comunicación e Internet, en los portales web de información meteorológica, etc., se implanta poco a poco la información sobre predecibilidad complementaria a la propia predicción, de tal modo que podemos hablar, a día de hoy, de predecir la predecibilidad. Indirectamente, predecir la predecibilidad es predecir el riesgo potencial de fenómenos adversos. El ECMWF, en el apartado de clustering (agrupamiento) en su sitio web, ya incorpora agrupaciones de situaciones atmosféricas y su atribución a uno de cuatro regímenes atlánticos predefinidos.

27.11 Valor de los SPC

En el capítulo de verificación (cap. 15 en la página 207) se presenta con detalle una medida llamada *valor relativo* de un sistema predictivo que permite

medir, en términos relativos, el valor del uso de un sistema predictivo particular con respecto a otro de referencia que puede ser, por ejemplo, la climatología (sec. 15.10 en la página 234). Mediante esta medida se ha podido comprobar, sistemáticamente, que la predicción probabilista resulta enormemente rentable. A la hora de tomar decisiones, ofrecer un rendimiento favorable a la larga, asesorar convenientemente al usuario, etc., la predicción probabilista es un serio candidato como sistema de predicción, no en vano el ECMWF lo pone en lugar preferente en su estrategia 2016-2025 (sec. 19.4 en la página 297). Una buena parte de la sociedad ofrece una resistencia natural a la información dada en términos de probabilidades. Poco a poco estas resistencias irán bajando.

27.11.1 ¿Pueden usarse los SPC sin probabilidades?

Los SPC son intrínsecamente probabilistas y el mejor modo de sacar un rendimiento óptimo de los mismos es como tales sistemas probabilistas. Cualquier intento de utilización determinista sería, en el mejor de los casos, una simplificación que haría perder valor a la información, cuando no una degradación de la misma o algo peor. Por ejemplo, el promedio del SPC no es necesariamente una atmósfera plausible (sec. 13.6.3 en la página 178), aunque estadísticamente obtenga buenas medidas de calidad (sec. 15.7 en la página 221). En los albores de los SPC, un sector profesional de la meteorología, ante la gran cantidad inabordable de información de los mismos, proponía «hacer la media y ya está».

Los box-plot, las barras de error o incertidumbre en las gráficas, los penachos, etc. son una manera sencilla de acercarse a la probabilidad. Las barras de error se

definen usualmente por estimaciones de probabilidad dadas (e. g. intervalos del 70%).

Las probabilidades, en principio, pueden ser un puente entre predictor y usuario de la predicción. Diferentes aplicaciones tendrán diferente sensibilidad a las falsas alarmas o a la tasa de aciertos. Los SPC facilitan al predictor una *modulación* que él puede añadir subjetivamente en la predicción, transformando las probabilidades brutas en terminología común. Por ejemplo: posibles nieblas (probabilidad entre 10 y 40 %), probables nieblas (40-70 %), muy probables (70-90 %), nieblas (90-100 %).

27.11.2 ¿Qué es realmente nuevo?

Antes de la llegada de los SPC los profesionales de la predicción elaboraban de alguna forma un contraste de predecibilidad. Sobre todo en situaciones delicadas, se miraba más de un modelo, se evaluaba la consistencia entre pasadas sucesivas del modelo principal, etc. Por tanto, no toda la información que brindan los SPC es realmente una novedad, al menos en parte. Los aspectos que no son realmente novedosos son:

- Los predictores, tradicionalmente, evaluaban la predecibilidad de acuerdo con su experiencia de la calidad del modelo numérico determinista, errores dependientes del flujo y de las escalas temporales y espaciales de la predicción.
- La inconsistencia de diferentes modelos deterministas se ha usado tradicionalmente como una muestra de la incertidumbre de la predicción.

Por otro lado, son realmente novedosos los siguientes aspectos:

- La dispersión (*spread*) de un SPC representa la incertidumbre de la predicción o la predecibilidad. Este tipo de información no puede extraerse fácilmente de un modelo determinista. Los SPC brindan una representación detallada, explícita y cuantitativa de la incertidumbre y, por ende, de la predecibilidad.
- Los SPC proporcionan una PDF de la predicción de la situación meteorológica, con la que pueden construirse numerosos productos de utilidades específicas para cada usuario o contingencia.
- Especialmente en situaciones de fenómenos adversos potenciales, los SPC son una herramienta mucho más útil. Al brindar numerosos escenarios, es mucho más fácil que alguno de esos escenarios pueda capturar una situación atmosférica crítica.

27.12 Conclusiones

Como herramienta para la predicción o productos para el usuario final, los SPC aportan numerosas ventajas con respecto a los modelos deterministas. Los productos de los SPC representan de forma natural el carácter probabilista de la atmósfera, que es precisamente lo que no aportan los modelos deterministas. Los SPC proporcionan una representación cuantitativa, detallada y explícita de la incertidumbre y, por ende, de la predecibilidad atmosférica. A su vez, ofrecen herramientas para evaluar la plausibilidad de potenciales fenómenos adversos y, por tanto, son herramientas importantísimas para ayudar a la toma de decisiones en la emisión de avisos meteorológicos.

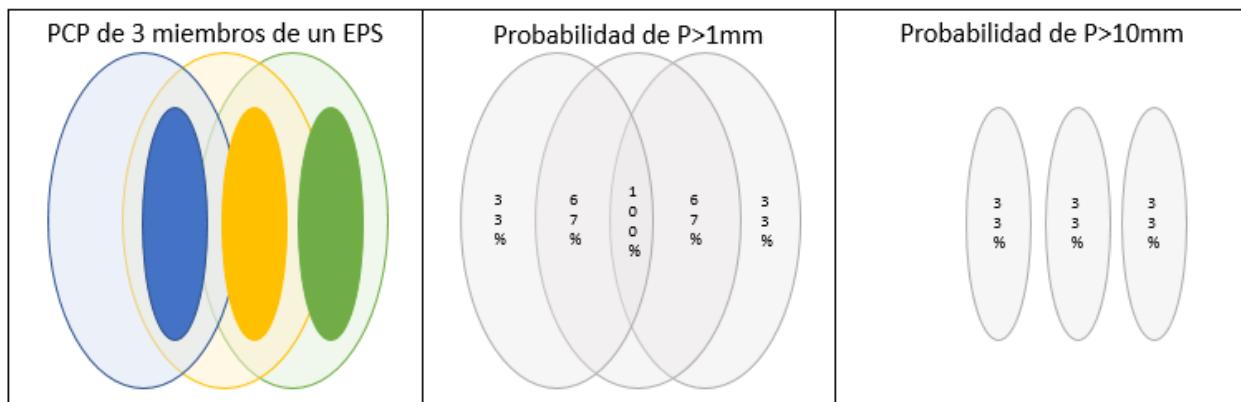


Figura 27.44: Ejercicio teórico con un SPC de tres miembros. a) Precipitación de cada uno de los miembros. Cada color representa un miembro del SPC. El área sombreada en colores claros indica las zonas donde el miembro muestra $PCP > 1 \text{ mm}$. El área sombreada en colores intensos indica las zonas donde el miembro muestra $PCP > 10 \text{ mm}$. b) Probabilidad de $PCP > 1 \text{ mm}$ derivada de a). c) Probabilidad de $PCP > 10 \text{ mm}$ derivada de a) (ver texto).

27.13 Anexo: Hacia un cambio de paradigma en la interpretación de la precipitación de los SPC

SERGI GONZÁLEZ HERRERO
DT en Cataluña, AEMET

En el caso de la precipitación, los SPC de resolución kilométrica plantean una problemática específica en predicción probabilista. A medida que aumenten la resolución y vayan entrando en la cadena de predicción los SPC de resolución kilométrica y *convection-permitting* (cap. 21 en la página 313), va a ser necesario aplicar un cambio de paradigma en la interpretación clásica punto a punto de malla (grid) de la precipitación (PCP). Esto será necesario especialmente en los casos donde esté envuelta la convección, casos que generalmente dejan grandes acumulaciones de precipitación en franjas espaciales y temporales estrechas.

Hagamos un pequeño ejercicio teórico para visualizarlo (Figura 27.44). Tenemos que elaborar una predicción y como herramienta para realizarla tenemos un SPC de mesoescala con sólo 3 miembros. Estos 3 miembros muestran un patrón de precipitación muy semejante, con una amplia zona que indica acumulaciones superiores a 1 mm, y dentro una zona más estrecha que indica una precipitación mayor a 10 mm. Todos los miembros muestran un patrón parecido, pero con una diferencia: este patrón no lo sitúan en el mismo lugar. En la Figura 27.44, parte (a) se muestra el ejercicio con este esquema, donde los tres miembros

muestran el mismo patrón pero con una separación evidente en el eje este-oeste.

Entre los productos más usados para condensar la información del conjunto de miembros de un SPC están los mapas de probabilidad de que se supere un cierto umbral. Para ello se calcula para cada punto de malla el porcentaje de miembros del SPC que superan dicho umbral (sec. 27.6.1 en la página 416). Si lo aplicamos a nuestro ejercicio con un umbral de 1 mm obtenemos un patrón como el que se muestra en la Figura 27.44, parte (b). Ahí se muestra una zona central que indica una gran probabilidad (100 %) de que lloverá más de 1 mm, allá donde los 3 miembros del SPC se solapan. Hacia el este y el oeste esa probabilidad disminuye, dando un 67 % allá donde dos miembros se solapan y un 33 % allá donde ningún miembro se solapa con otro.

Construyamos ahora un mapa de probabilidad pero con un umbral de 10 mm. Como ninguna de las zonas de precipitación mayor a 10 mm se solapa con otras, la probabilidad máxima que obtendremos será de 1/3 (33 % aproximadamente). Esta situación se muestra en la Figura 27.44, parte (c) donde se observa que hay 3 áreas sombreadas con una probabilidad del 33 %.

Normalmente, cuando se usa un SPC no se analizan los datos de cada uno de los miembros por separado ya que en conjunto contienen demasiada información como para ser usados en la rutina operativa (sec. 27.2.1 en la página 402). Por lo tanto, si este ejercicio fuera real, no veríamos nunca la Figura 27.44, parte (a) sino que pasaríamos directamente a analizar las partes (b) y (c). Con esta información tratada de forma clásica y suponiendo que el SPC está bien calibrado, un

predictor posiblemente escribiría un pronóstico del estilo «*En una amplia zona, hay una probabilidad baja (33 %) de que se alcancen los 10 mm*». Pero, ¿es ésta la forma más acertada de decirlo?, viendo la Figura 27.44 en la página anterior, parte (a) pensaríamos que es más adecuado pronosticar que «*En alguna zona estrecha dentro del área de predicción se esperan (con un 100 % de probabilidad) más de 10 mm*».

Nótese que este comportamiento refleja de algún modo un aspecto que se describe en detalle en el capítulo dedicado a verificación (sec. 15.5 en la página 219): los modelos y SPC tienen diferentes rendimientos según las diferentes escalas espaciales en que los podemos usar y, en general, cuanto más fina la escala, más exigencia requiere y peor resultado dan los modelos, como puede verse en el caso de la Doble Penalización (sec. 15.5.2 en la página 220). A veces, cambiar de escala un modelo, por ejemplo, haciendo su resolución un poco más gruesa, puede simplificar información

innececesaria y menos fiable de escala muy fina, sintetizando la información en una escala algo más gruesa y aportando así información útil al predictor.

Situaciones análogas a ésta se producen a diario en los SPC clásicos, pero a menudo pasan desapercibidas ya que como las áreas de precipitación son a menudo grandes, sus solapes son mayores y por lo tanto los errores de interpretación por desplazamiento de las zonas de precipitación entre los distintos miembros del SPC son menores. Pero a medida que la resolución de los SPC aumente y se vayan incorporando al proceso de predicción los SPC con modelos de resolución kilométrica convection-permitting, como el AEMET-γSREPS (cap. 21 en la página 313), éste problema se hará más y más evidente, en especial en situaciones de convección profunda. Estas diferencias entre miembros no se producen únicamente en el patrón espacial, sino que también pueden producirse diferencias temporales.

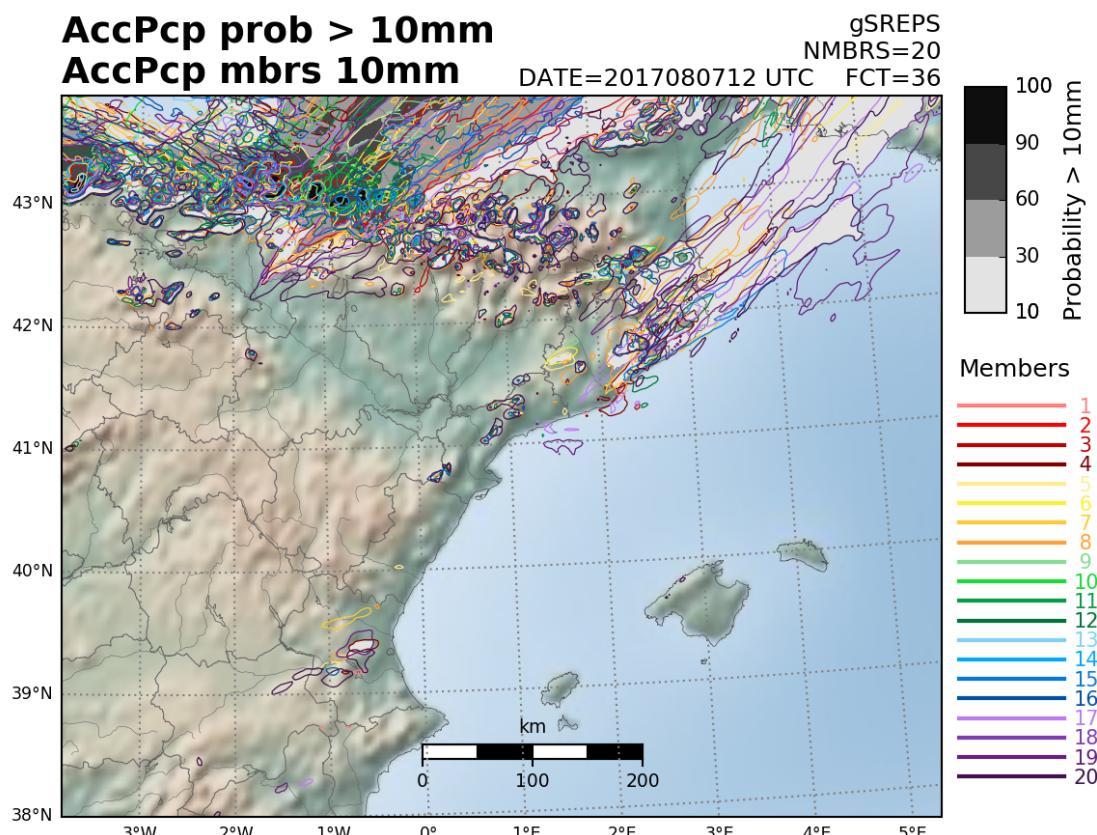


Figura 27.45: Probabilidad de PCP > 10 mm entre las 00 UTC del día 7 y las 12 UTC del día 8 de agosto de 2017 (en escala de grises) según el SPC convection-permitting AEMET-γSREPS. Contorneadas, se muestran las áreas donde cada uno de los miembros por separado (en distintos colores) muestra una PCP > 10 mm. Como hay 20 miembros cada uno de los miembros representa un 5 % de contribución a la probabilidad de precipitación (ver texto).

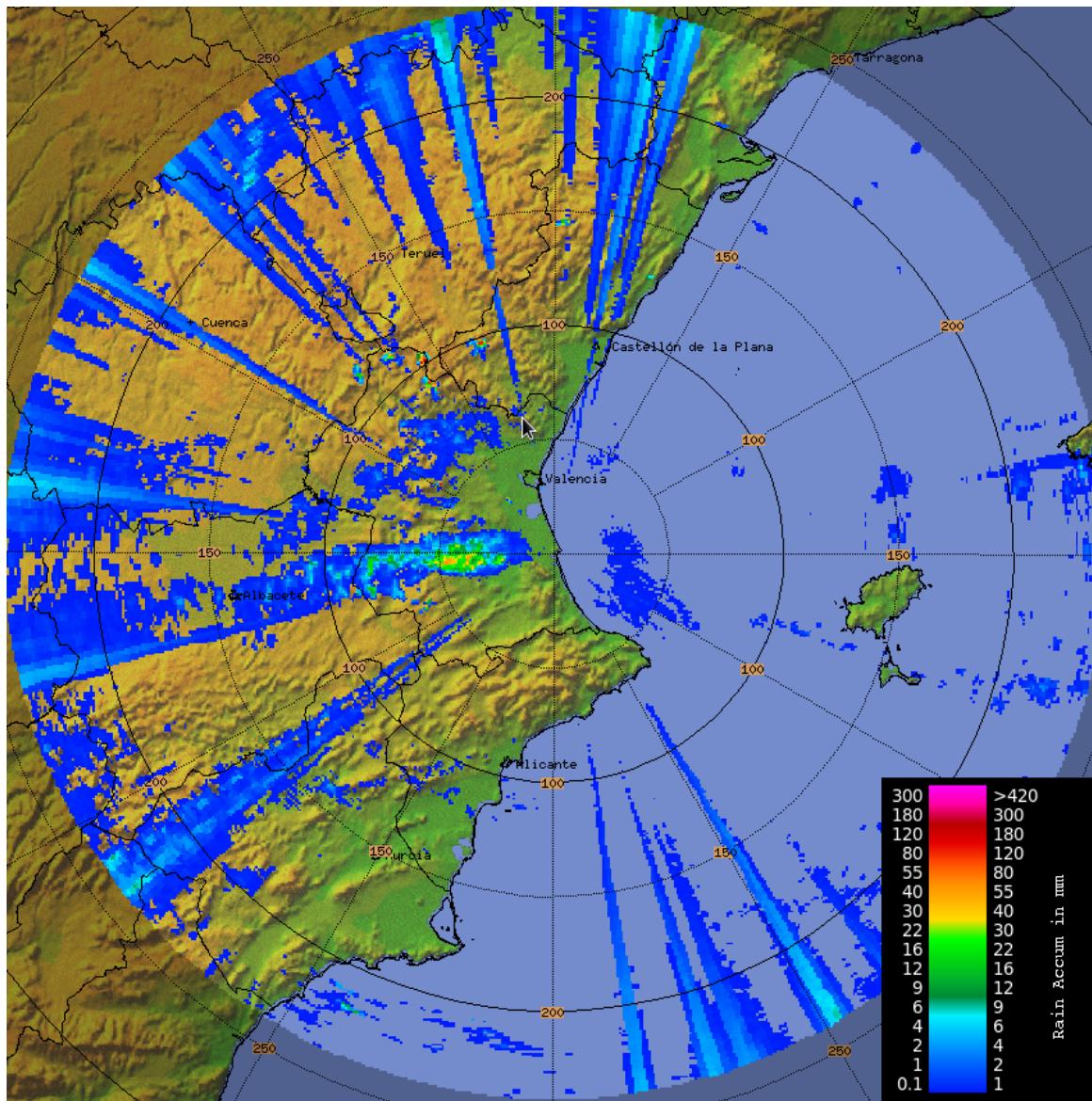


Figura 27.46: Precipitación acumulada durante el día 7 de agosto de 2017 derivado a partir de los datos de radar, considerado como una buena estimación de la realidad (ver texto).

Así pues, los predictores del tiempo debemos replantearnos nuevas formas de analizar los sistemas de predicción por conjuntos, para que el uso inadecuado y no renovado de la tecnología no degrada nuestra capacidad interpretativa (sec. 4.4 en la página 43).

La Figura 27.45 en la página anterior muestra un ejemplo real de lo expuesto anteriormente en el ejercicio teórico, para comprobar cómo ocurre este problema en la realidad. Esta imagen corresponde a la primera predicción en que se usó operativamente, aunque no todavía de forma oficial, el AEMET-γSREPS y muestra la probabilidad de precipitación mayor a 10 mm (sombreado en grises) entre las 00 UTC del día 7 y las 12 UTC del 8 de agosto de 2017. Sobre ella se han contorneado, en distintos colores para cada miembro, las zonas donde los distintos miembros mostraban

precipitaciones por encima de 10 mm. Si nos centramos solamente en la provincia de Valencia podemos observar que, aunque la probabilidad de precipitación superior a 10 mm no llega al 10% (no hay ninguna zona sombreada), hasta 15 miembros individuales muestran la superación de ese umbral en distintas partes de la provincia, casi ninguna solapándose. Ello nos indicaba que la probabilidad de que se produjera una tormenta era alta, pero que sería en todo caso muy aislada. Además, basados en el patrón de precipitación elongado que mostraban varios miembros se podía deducir que tendría un desplazamiento de oeste a este. Efectivamente, así se produjo durante la tarde del día 7, tal y como se observa en la Figura 27.46 (ver también sec. 22.8.1 en la página 354).

Este ejemplo muestra cómo debemos ir hacia un cam-

bio de paradigma a la hora de interpretar los SPC de resolución kilométrica en situaciones de convección. Este cambio de paradigma ha de conducir a crear nuevos productos que nos permitan entender mejor la probabilidad de que se produzca convección y cómo esta se distribuye espacial y temporalmente. Una primera aplicación puede ser la que hemos mostrado para el caso de Valencia, que supondría buscar un umbral suficientemente alto como para separar los miembros de mayor precipitación visualmente y poderlos contar

uno a uno. Sin embargo, en el futuro serán necesarios productos más elaborados, probablemente con visualizaciones interactivas que permitan, por ejemplo, seleccionar un área y decir cuántos miembros dan en algún punto de éste área una precipitación superior a un umbral. En cualquier caso, la predicción probabilista a partir de los SPC de resolución kilométrica está imponiendo una modificación en nuestra forma de interpretar los SPC a la que nos deberemos adaptar en un futuro próximo.

27.14 Anexo: adaptación de la técnica de ECMWF de agrupamiento de SPC en AEMET

ÁLVARO SUBÍAS DÍAZ-BLANCO Y ALEJANDRO ROA ALONSO
ATAP, AEMET

Este anexo se presentó en la Asociación Meteorológica Española en 2018.

Desde Marzo de 2018 se ha establecido en **AEMET** un nuevo procedimiento de agrupamiento de los miembros del ensemble del ECMWF, el ECENS. Dicho procedimiento ha sido desarrollado en ECMWF por L. FERRANTI y S. CORTI (2011 [5]) y adaptado en **AEMET** por los autores, a partir de un estudio previo, para campos de superficie en **AEMET**, por parte de A. RODRÍGUEZ (2014 [18]).

Para la realización del agrupamiento se trabaja en el espacio de componentes principales de la altura geopotencial en 500 hPa. Se propone una varianza explicada del 80% para establecer un truncamiento de la dimensión de dicho espacio. Es en ese espacio, típicamente de dimensión de 2 a 3, en el que se realiza un agrupamiento basado en la técnica de *K-medias*. Fijado un número de grupos la mejor elección se realiza para aquel grupo que maximice el cociente r_n entre la varianza entre centroides y el promedio de la varianza interna de cada grupo. Se obtiene la mejor agrupación para cada número de grupos, desde 1 hasta 6. Este procedimiento permite obtener la elección óptima del número de grupos basada en un criterio de *significación estadística*. Fijado el número de grupos, dicha significación está dada por el porcentaje de grupos calculados a partir de una perturbación gaussiana de la agrupación óptima asociada a dicho número, con coeficiente r menor que r_n . La elección del número de grupos es aquella que maximiza dicha significación junto con algunos criterios dependientes del corto, medio o largo plazo (*hiperheurística*).

En la actualidad este es el procedimiento usado en **AEMET** para el agrupamiento de los distintos miembros del ECENS, efectuado en dos dominios: (1) península ibérica y Baleares (2) islas Canarias.

La técnica de agrupamiento permite la separación del conjunto de 51 miembros en subgrupos, de los cuales

se puede extraer el mismo tipo de información que la que se extrae para el total del ensemble. En particular una vez obtenidos los grupos se pueden obtener, por ejemplo, probabilidades de superación de umbrales para la precipitación.

Se elige el representante de cada grupo como el miembro más cercano a la media, la cual se desecha por carecer de coherencia meteorológica, dado el carácter no lineal de las ecuaciones que gobiernan el estado de la atmósfera (ver sec. 12.1.1 en la página 156).

27.14.1 Método

Un agrupamiento de miembros de un ensemble que se realice de forma automática requiere la elección de al menos un campo de salida del modelo sobre el que realizar una comparación. Típicamente se elige la altura geopotencial en 500 hPa (z500) al tratarse de un campo suave (ver Figura 27.47) a la vez que sinópticamente representativo del estado de la atmósfera.

El anterior método de agrupamiento utilizado en **AEMET**, desarrollado BENITO ELVIRA y AURORA MARTÍN (1999 [3]) realiza un agrupamiento en el campo de z500. El algoritmo trata de minimizar la pérdida de varianza al fusionar dos grupos utilizando la distancia de WARD entre ellos que consiste en la distancia L_2 entre sus centroides, pesada por un factor que tiene en cuenta las poblaciones de ambos, de modo que se favorece la fusión con pocos miembros.

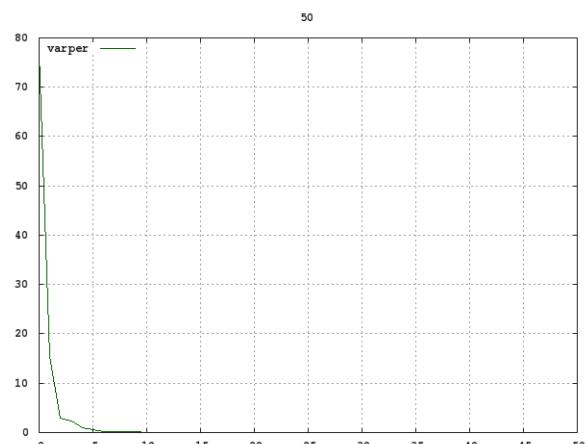


Figura 27.47: Varianza explicada de cada una de las componentes principales presentadas en orden decreciente. Los dos primeros modos explican más del 80% de variabilidad.

Desde 2014, en el ECMWF se utiliza una nueva técnica de agrupamiento desarrollada por L. FERRANTI y S. CORTI (2011 [5]). En ella se comparan campos de z500 en el espacio de componentes principales en vez del espacio de puntos de malla. Fijado un alcance en el que se desee realizar un agrupamiento se calculan las funciones empíricas ortogonales (EOF [14]) a partir de la variabilidad de los 50 miembros perturbados del ECENS junto con el miembro de control (sin perturbar); estas funciones constituyen los modos que diagonalizan la matriz de covarianza del campo analizado. La varianza explicada de los distintos modos decrece muy rápidamente cuando se ordenan de forma decreciente (Figura 27.47 en la página anterior) de modo que, para explicar una varianza de al menos

el 80 %, basta escoger simplemente muy pocos modos (típicamente 2 ó 3). Los campos en los que se realiza la presente técnica han de ser suaves, es decir, que presenten un decaimiento rápido o, dicho de otra forma, que pocos modos expliquen mucha variabilidad.

Una vez que se tienen los modos se proyectan los 51 campos en las funciones empíricas ortogonales y se obtienen las componentes principales de cada uno de ellos. En una situación en la que nos quedamos con 2 modos, basta considerar dos parámetros para identificar cada uno de los campos. Por tanto, es posible hacer una representación en un plano en la que tenemos 50+1 puntos. Desde un punto de vista geométrico, el objetivo consiste en la identificación de grupos o particiones en dicho plano (Figura 27.48).

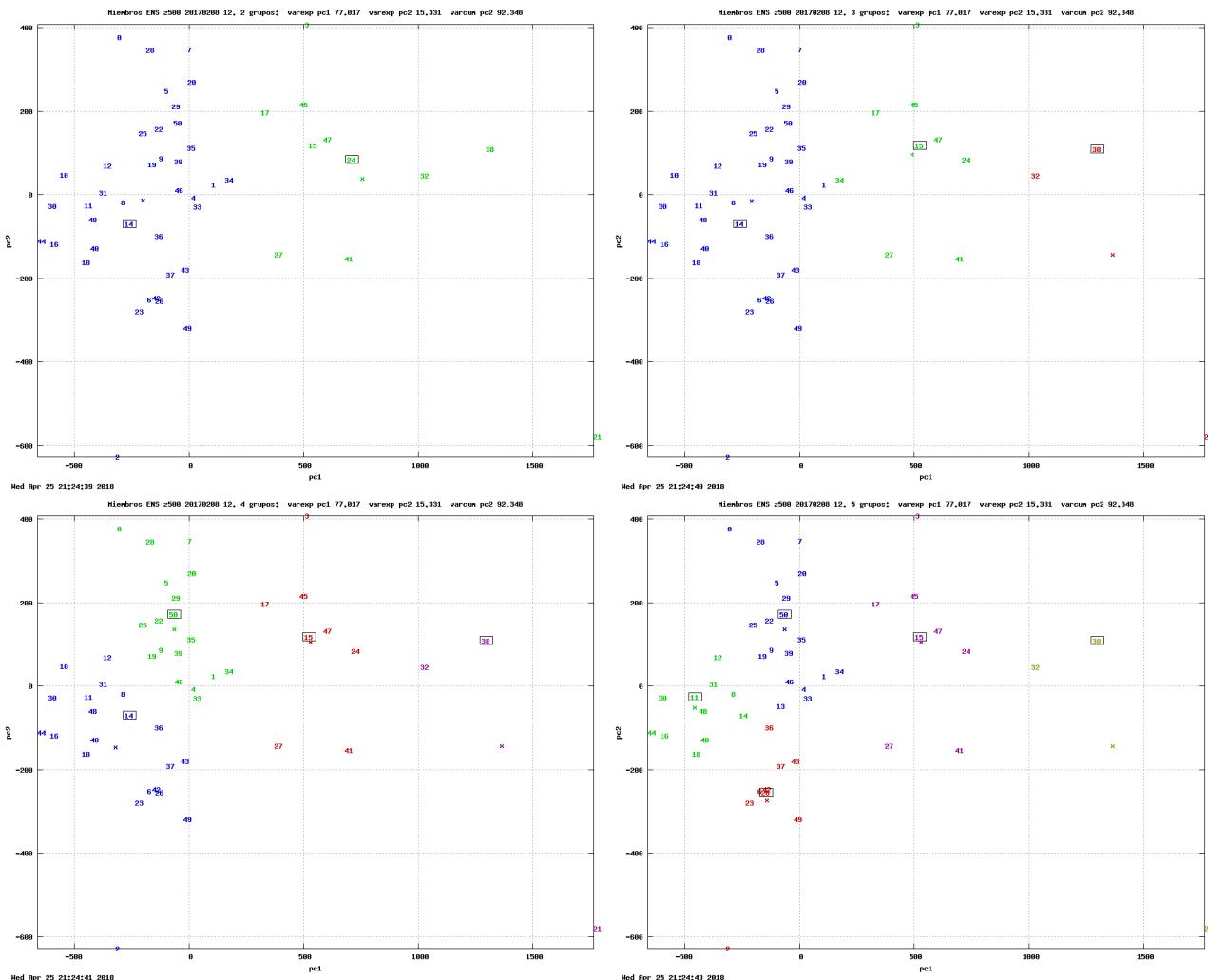


Figura 27.48: Agrupaciones Q2, Q3, Q4, Q5 de los 51 miembros en el espacio de componentes principales (figuras a, b, c, d, respectivamente). En las imágenes se muestra la pertenencia de un determinado miembro a un grupo según su color recuadrando el representante. La letra x representa la media de cada grupo (ensemble mean).

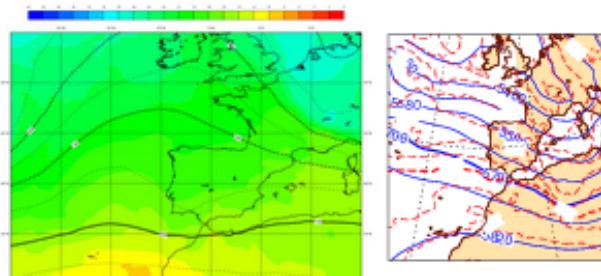


Figura 27.49: Grupo formado por tres miembros (15, 27, 38) en el que se muestra el representante (izquierda) y la media del ensemble (derecha). Pese a representar el mismo grupo pueden apreciarse las diferencias en las isolíneas del campo de geopotencial.

El objetivo es conseguir particiones en las que los *centrodes* estén lo más separados entre sí y, dentro de cada grupo, los miembros muy próximos al centroide. Para ello se define el valor r_Q como el cociente entre la varianza entre centrodes y la media de las varianzas internas de cada grupo. Lo deseable es que este coeficiente sea lo mayor posible.

A partir de una partición Q se realiza un procedimiento de reasignación en el que dos grupos se intercambian un miembro. El procedimiento de reasignación de elementos a grupos más próximos devuelve una nueva partición en la que el procedimiento se estabiliza, es decir, no hace crecer el coeficiente r . De la partición obtenida consideramos el miembro representante, que es el más cercano a la media (ensemble mean). El representante escogido de esta manera es un miembro del ECENS coherente en todos sus campos, eso no ocurre necesariamente para el ensemble mean, en el que la media de determinados campos puede no representar un campo físico real (Figura 27.49).

Una vez obtenida una partición Q , se compara la fuerza del agrupamiento con una función de distribución de referencia. Para ello, se generan particiones Q_m con el mismo número de miembros que Q . Estas muestras se generan a partir de *cadenas de MARKOV* cuya media, varianza y autocorrelación a primer orden se obtienen a partir de Q [5]. Para cada conjunto inicial de datos Q_m se realiza un agrupamiento, obteniéndose el coeficiente r_{Q_m} . La significación estadística estará dada por el porcentaje de muestras sintéticas Q_m tales que $r_{Q_m} < r_Q$.

Inicialmente, se parte de un número n de grupos prefijado, se toma al azar un conjunto de particiones

iniciales y se elige la partición Q_n que maximiza el cociente r_{Q_n} . Una vez hecho esto se compara la significación de las particiones Q_2, \dots, Q_6 y se elige como número de grupos óptimo aquel que maximiza dicha significación (Figura 27.50). Se tiene en cuenta una hiperheurística si el alcance se sitúa en el corto plazo, de modo que se inhibe el número de grupos o, si se sitúa en el largo plazo, de modo que éste se favorece.

27.14.2 Resultados

Una vez identificados los miembros que pertenecen a un determinado grupo, se procede a su representación gráfica. Debido a que un grupo constituye un subconjunto de un sistema(s) de predicción por conjuntos (SPC), se puede obtener el mismo tipo de información. Además de mostrar el miembro representante, podemos obtener probabilidades, como por ejemplo la superación de umbrales de precipitación o la probabilidad de que esté el cielo cubierto.

Esta técnica de agrupamiento se realiza en pasadas de 00 y 12 UTC del ECENS del ECMWF para los alcances cuya validez sea las 12 UTC, desde el día D hasta el día D+14. El cálculo de probabilidades tiene en cuenta las 24 horas del día en el que está centrado el alcance antes mencionado.

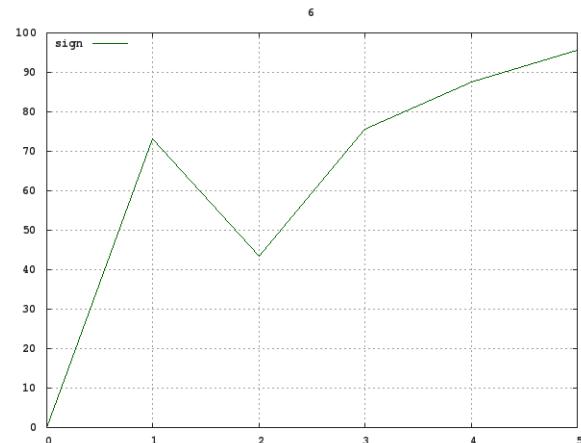


Figura 27.50: Significación estadística según el número de grupos, permite establecer el número de grupos óptimo. En este caso el número óptimo es de dos grupos en el que se sitúa en primer máximo local de significación (hiperheurística de corto plazo). El eje horizontal representa el número de grupos (menos 1).

Agrupamiento basado en z500. Número de clusters: 2. Varianza explicada: 80.0

Validez: mar 14 feb 2017 12 UTC (D+6)

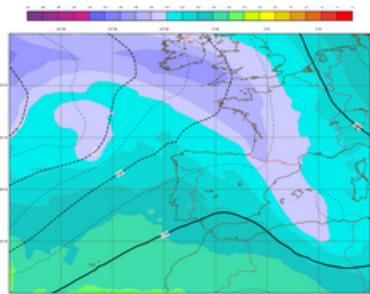
20170208 a 12 UTC HH+144 (rango:132-156)

EPS

control

miembros: 51

Temperatura (C) y Altura Geopotencial (m) en 500 hPa



grupo: 1

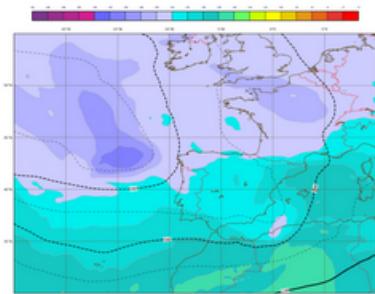
representante: 14

miembros: 40 control

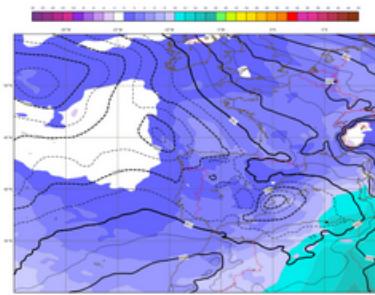
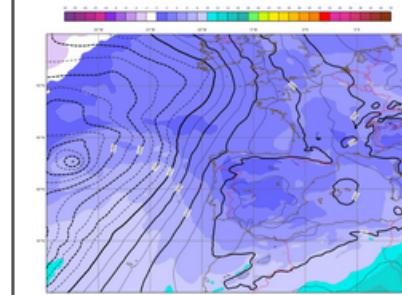
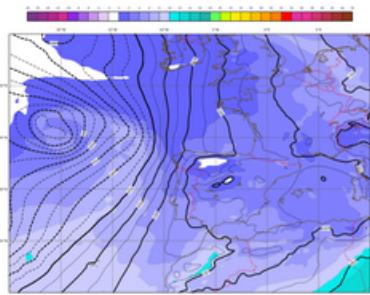
grupo: 2

representante: 24

miembros: 11



Temperatura en 850 hPa (C) y Presión en Superficie (hPa)



Probabilidad precipitación de 0 a 24 UTC superior a 1 mm

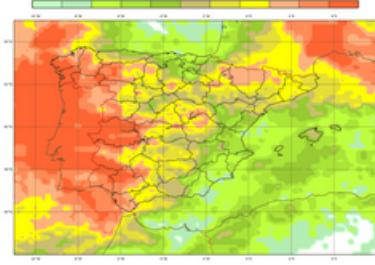
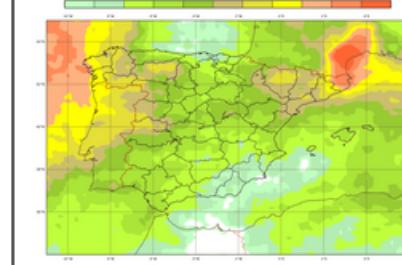
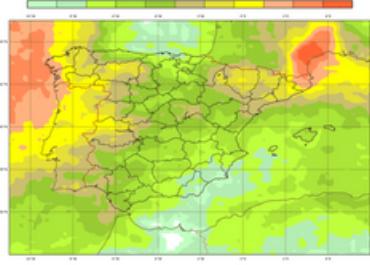


Figura 27.51: Visualización principal del agrupamiento. En este caso se muestran 2 grupos; pueden observarse las diferencias de probabilidad de precipitación entre ambos (ver texto).

Como visualización principal se muestran varias columnas, siendo la primera de ellas el ECENS conjunto con los 51 miembros y como campo de referencia el miembro de control. Las siguientes columnas están dadas por los distintos grupos; para cada grupo, además, se indican tanto el número de miembros que lo forman, como el representante que se utiliza para la visualización (Figura 27.51).

La técnica de Análisis de Componentes Principales (Principal Component Analysis, PCA [8, 9]) separa modos de máxima variabilidad ortogonales entre sí,

es decir, con covarianza nula entre ellos. Lo que intenta el presente agrupamiento es desglosar, en el mapa de desviación estándar del SPC conjunto, distintos polos de variabilidad en distintos grupos. Esto puede contemplarse en el mapa de desviación estándar de z500 para el ECENS y para cada uno de los grupos (Figura 27.52 en la página siguiente, parte superior). Así mismo, en la Figura 27.52 en la página siguiente, parte inferior se muestran, para el ECENS conjunto y para los dos grupos, probabilidades de superación de diferentes umbrales .

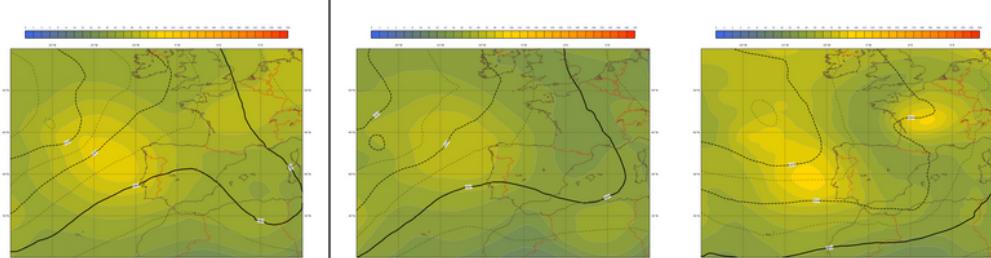
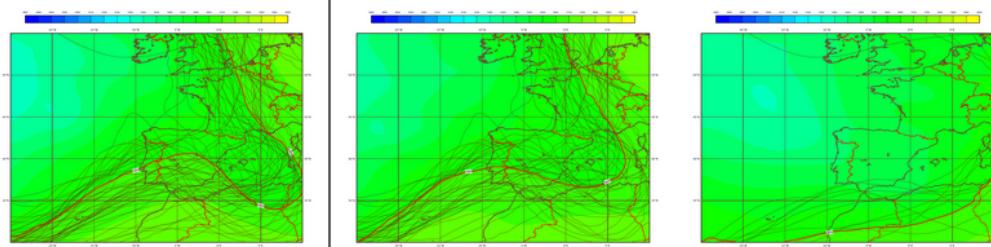
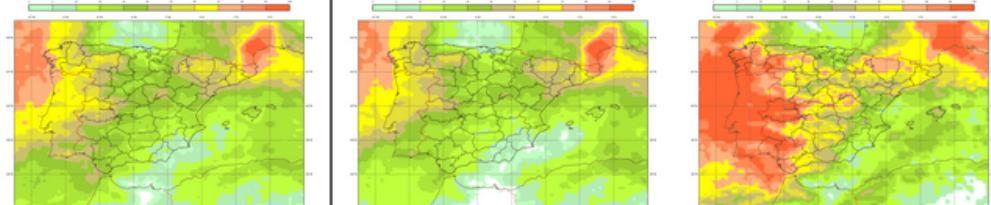
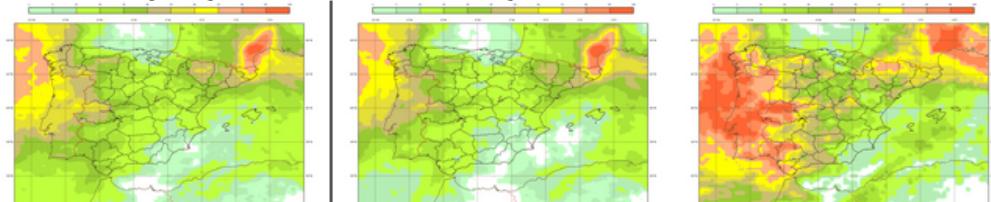
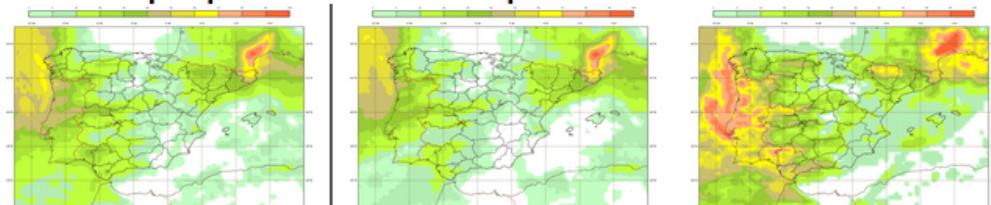
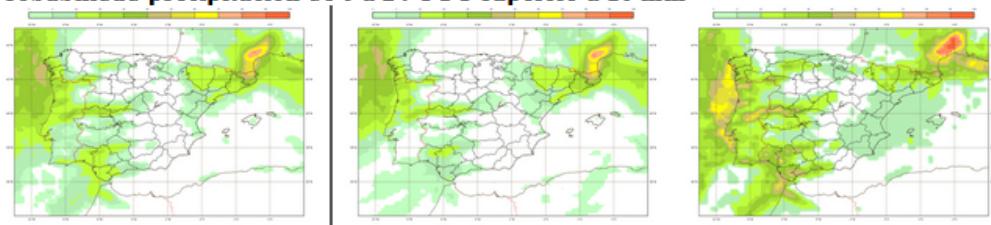
Desviación Estandar e Isohipses de Altura Geopotencial (m) en 500 hPa**Spaghetti de Altura Geopotencial (m) en 500 hPa****Probabilidad precipitacion de 0 a 24 UTC superior a 1 mm****Probabilidad precipitacion de 0 a 24 UTC superior a 2 mm****Probabilidad precipitacion de 0 a 24 UTC superior a 5 mm****Probabilidad precipitacion de 0 a 24 UTC superior a 10 mm**

Figura 27.52: Arriba (dos primeras filas): desviación estándar (sec. 13.6.5 en la página 181 y sec. 27.5.1 en la página 412) de z500 y mapa con spaghetti (sec. 27.4.1 en la página 406) en la isohipsa cuyo miembro de control pasa por Madrid. Los grupos representan la misma situación que en la Figura 27.51 en la página anterior. Abajo (cuatro últimas filas): probabilidad de superación de umbrales de precipitación de 1, 2, 5 y 10 mm para el ECENS conjunto y para los distintos grupos.

27.14.3 Conclusiones

La presente técnica de agrupamiento establece una comparación en el espacio de componentes principales, en vez de hacerlo en el espacio de puntos de malla, lo que permite que cada campo sea etiquetável con un reducido número de valores. Este conjunto es de dos o tres valores, según la situación meteorológica, mientras que para el cálculo en el espacio de puntos de malla se necesitan todos los puntos para su completa descripción (en nuestro ejemplo 176×126). Este hecho permite el desarrollo de algoritmos que sólo son posibles debido a la baja *dimensionalidad* del problema.

A su vez, la nueva representación gráfica de campos permite la intercomparación entre los distintos grupos y el ECENS conjunto, dado que los grupos constituyen un subconjunto del ensemble y por tanto se puede extraer información análoga.

Una posible mejora de esta técnica consistiría en establecer un peso en la zona de mayor interés a la hora de calcular covarianzas. De modo que, por ejemplo, si nos interesa la variabilidad sobre la península ibérica, entonces una gran variabilidad en una región como el sur de las islas británicas no debería de ser causante de consideración de nuevos grupos.

27.15 Referencias

- [1] BUIZZA, Roberto. "Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system". En: *Monthly Weather Review* 125.1 (1997), páginas 99-119. ISSN: 0027-0644. DOI: [10.1175/1520-0493\(1997\)125<0099:PFSOEP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0099:PFSOEP>2.0.CO;2) (citado en páginas 412, 413).
- [2] DEE, Dick P y col. "The ERA-Interim reanalysis: Configuration and performance of the data assimilation system". En: *Quarterly Journal of the royal meteorological society* 137.656 (2011), páginas 553-597 (citado en página 411).
- [3] ELVIRA MONTEJO, Benito y MARTÍN SANZ, Aurora. "Verificación del EPS a través de los grupos y de los supergrupos". En: *Nota Técnica del STAP* 36 (1999). URL: <http://hdl.handle.net/20.500.11765/5114> (citado en página 440).
- [4] ESCRIBÀ, Pau y col. "Probabilistic prediction of raw and BMA calibrated AEMET-SREPS: The 24 of January 2009 extreme wind event in Catalunya". En: *Advances in Geosciences* 26 (nov. de 2010), páginas 119-124. ISSN: 16807340. DOI: [10.5194/adgeo-26-119-2010](https://doi.org/10.5194/adgeo-26-119-2010) (citado en página 419).
- [5] FERRANTI, L y CORTI, S. "New clustering products". En: *ECMWF Newsletter* 127.6-11 (2011), páginas 1-2 (citado en páginas 440-442).
- [6] GARCÍA-MOYA, José Antonio y col. "Predictability of short-range forecasting: A multimodel approach". En: *Tellus, Series A: Dynamic Meteorology and Oceanography* 63.3 (mayo de 2011), páginas 550-563. ISSN: 02806495. DOI: [10.1111/j.1600-0870.2010.00506.x](https://doi.org/10.1111/j.1600-0870.2010.00506.x) (citado en páginas 402, 412, 413).
- [7] HARTIGAN, John A y HARTIGAN, J A. *Clustering algorithms*. Volumen 209. Wiley New York, 1975 (citado en página 407).
- [8] HOTELLING, Harold. "Analysis of a complex of statistical variables into principal components." En: *Journal of educational psychology* 24.6 (1933), página 417 (citado en página 443).
- [9] JOLLIFFE, Ian T. "Principal component analysis and factor analysis". En: *Principal component analysis*. Springer, 1986, páginas 115-128 (citado en página 443).
- [10] LALAURETTE, F. "Two proposals to enhance the EFI response near the tails of the climate distribution". En: *American Meteorol. Soc* (2003), páginas 1-8 (citado en página 421).
- [11] LALAURETTE, François. "Early detection of abnormal weather conditions using a probabilistic extreme forecast index". En: *Quarterly Journal of the Royal Meteorological Society* 129.594 (oct. de 2003), páginas 3037-3057. ISSN: 00359009. DOI: [10.1256/qj.02.152](https://doi.org/10.1256/qj.02.152) (citado en páginas 419, 420).
- [12] LAVERS, David A, PAPPENBERGER, Florian y ZSOTER, Ervin. "Extending medium-range predictability of extreme hydrological events in Europe". En: *Nature communications* 5 (2014) (citado en página 419).
- [13] LAVERS, David A. y col. "ECMWF Extreme Forecast Index for water vapor transport: A forecast tool for atmospheric rivers and extreme precipitation". En: *Geophysical Research Letters* 43.22 (nov. de 2016), páginas 11, 811-852, 858. ISSN: 00948276. DOI: [10.1002/2016GL071320](https://doi.org/10.1002/2016GL071320) (citado en página 419).
- [14] LORENZ, Edward N. "Empirical orthogonal functions and statistical weather prediction". En: (1956) (citado en página 441).
- [15] MARZBAN, Caren y SANDGATHE, Scott. "Cluster Analysis for Verification of Precipitation Fields". En: *Weather and Forecasting* 21.5 (2006), páginas 824-838. ISSN: 0882-8156. DOI: [10.1175/WAF948.1](https://doi.org/10.1175/WAF948.1) (citado en página 407).

- [16] MOLTENI, F. y col. “The ECMWF ensemble prediction system: Methodology and validation”. En: *Quarterly Journal of the Royal Meteorological Society* 122.529 (ene. de 1996), páginas 73-119. ISSN: 1477-870X. DOI: [10.1002/qj.49712252905](https://doi.org/10.1002/qj.49712252905) (citado en página 407).
- [17] PERSSON, Anders. “User guide to ECMWF forecast products”. En: *Ecmwf* March (2011), página 127 (citado en página 402).
- [18] RODRÍGUEZ, A. “Nuevas técnicas de agrupamiento de predicciones por conjuntos del ECMWF aplicadas a campos de superficie”. En: *Nota Técnica del Área de Innovación 1* (2014) (citado en página 440).
- [19] UPPALA, Sakari M y col. “The ERA-40 reanalysis”. En: *Quarterly Journal of the royal meteorological society* 131.612 (2005), páginas 2961-3012. DOI: [10.1256/qj.04.176](https://doi.org/10.1256/qj.04.176) (citado en página 411).
- [20] VITART, F. y col. “Simulation of Interannual Variability of Tropical Storm Frequency in an Ensemble of GCM Integrations”. En: *Journal of Climate* 10.4 (abr. de 1997), páginas 745-760. ISSN: 0894-8755. DOI: [10.1175/1520-0442\(1997\)010<0745:SOIVOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0745:SOIVOT>2.0.CO;2) (citado en página 422).
- [21] WILKS, Daniel S. *Statistical Methods in the Atmospheric Sciences*. Academic Press, 2011, página 676. ISBN: 9780123850225. URL: <https://www.sciencedirect.com/bookseries/international-geophysics/vol/100> (citado en página 426).
- [22] ZSÓTÉR, Ervin. “Recent developments in extreme weather forecasting”. En: *ECMWF Newsletter* 107.107 (2006), páginas 8-17 (citado en página 420).
- [23] ZSOTER, Ervin, PAPPENBERGER, Florian y RICHARDSON, David. “Sensitivity of model climate to sampling configurations and the impact on the Extreme Forecast Index”. En: *Meteorological Applications* 22.2 (abr. de 2015), páginas 236-247. ISSN: 13504827. DOI: [10.1002/met.1447](https://doi.org/10.1002/met.1447) (citado en página 419).

