



Sistema de Predicción por Conjuntos hecho en España: AEMET-SREPS

21

DOI: [10.31978/014-18-009-X.21](https://doi.org/10.31978/014-18-009-X.21)

CARLOS SANTOS BURGUETE

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

ALFONS CALLADO I PALLARÉS Y PAU ESCRIBÀ AYERBE

Delegación Territorial en Cataluña, AEMET

JOSÉ ANTONIO GARCÍA-MOYA ZAPATA
Area de Innovación, AEMET

JESÚS MONTERO GARRIDO

Departamento de Producción, AEMET

JUANMA SANCHO ÁVILA

CNP, AEMET

DANIEL SANTOS-MUÑOZ

Project Leader for System, HIRLAM, AEMET

JUAN PABLO SIMARRO GRANDE

Delegación Territorial en Valencia, AEMET

ROMÁN.- Inventen, pues, ellos y nosotros nos aprovecharemos de sus invenciones. Pues confío y espero en que estarás convencido, como yo lo estoy, de que la luz eléctrica alumbrará aquí tan bien como allí donde se inventó.

SABINO.- Acaso mejor.

El pórtico del templo – MIGUEL DE UNAMUNO

La predecibilidad atmosférica es dependiente del flujo: según el tipo de situación, el comportamiento atmosférico es más o menos predecible. Los modelos atmosféricos utilizados en predicción numérica del tiempo, con escalas espaciales y temporales cada vez más finas, encuentran a menudo limitaciones para predecir los fenómenos adversos y los fenómenos extremos. En España el comportamiento meteorológico está influido por la interacción entre el flujo sinóptico y las características del área mediterránea, con orografía compleja de pequeña escala y el calor latente del mar. Esta interacción produce a menudo sistemas convectivos de mesoescala, altamente caóticos y muy difíciles de predecir, que con cierta frecuencia producen precipitaciones de varios cientos de litros por metro cuadrado en menos de 24 horas. En este contexto, el enfoque probabilista se hace imprescindible. En la primera década del siglo XX, AEMET fue pionera en Europa en la implementación de un sistema de predicción por conjuntos para el corto plazo basado en la técnica llamada *multimodelo*. Presentamos en este capítulo este sistema, denominado *AEMET-SREPS*.

Palabras clave: sistema de predicción por conjuntos AEMET-SREPS, proyecto AEMET-SREPS, multimodelo-multicontorno, proyecto COASTEPS, proyecto PREDIMED, proyecto MEDICANES, proyecto ENSEMBLE.

Imagen parte superior: *cumulonimbus mamma*. Villaseca de Uceda (Guadalajara), hacia el sureste, 25 de junio de 2006, a las 21:28. Fotografía de JOSÉ ANTONIO QUIRANTES CALVO.

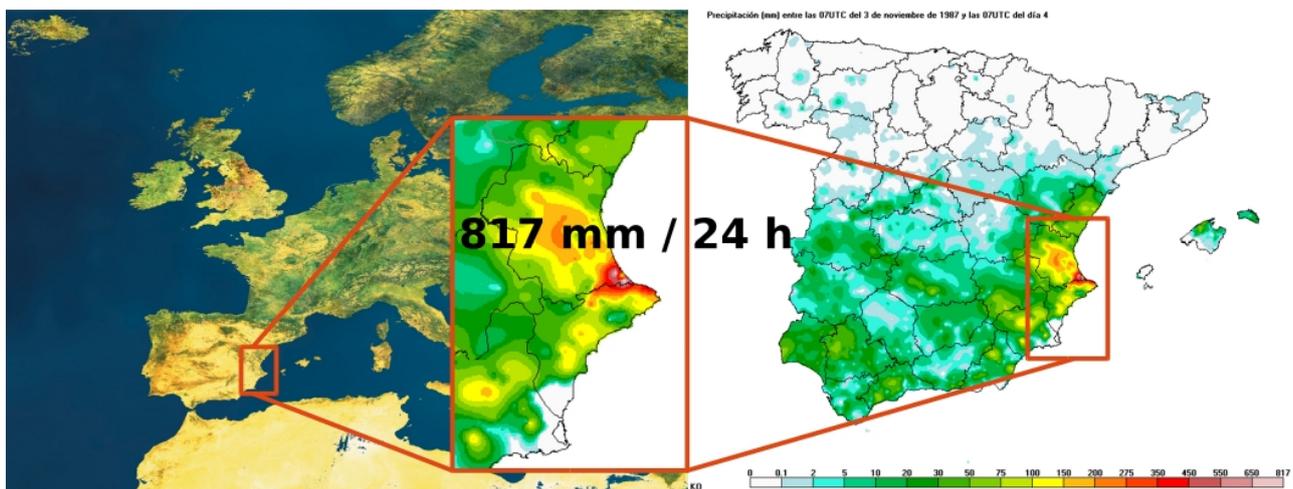


Figura 21.1: Los fenómenos convectivos, altamente caóticos desde el punto de vista dinámico, producen a menudo precipitaciones intensas en el área mediterránea española. En este caso se ilustra el episodio de precipitación de 817 mm registrada en Oliva el 04-11-1987. Imagen: composición, incluyendo imagen de satélite ESA 2003 y mapas pluviométricos de CÉSAR RODRÍGUEZ BALLESTEROS, *Agencia Estatal de Meteorología (AEMET) 2017*.

21.1 El proyecto AEMET-SREPS

Durante los años 2002-2014 se desarrolló en AEMET (INM anteriormente) un sistema de predicción por conjuntos (SPC, cap. 13 en la página 165) para el corto plazo que vino a denominarse AEMET-SREPS. Para la elaboración de predicciones a medio y largo plazo, desde 3-5 días hasta la escala quincenal, AEMET utiliza las salidas directas o posprocesadas del modelo global determinista y del sistema global de predicción por conjuntos (Ensemble Prediction System, ECENS, sec. 19.3 en la página 293), del Centro Europeo de Predicción a Plazo Medio (ECMWF, cap. 19 en la página 289), denominados en este volumen ECHRES (con resoluciones en aquellos tiempos de 50, 25 y 16 m) y ECENS (100, 50 y 32 km), respectivamente. Para la predicción en el corto plazo, en el intervalo 0-48 horas, AEMET ha venido utilizando, según la época, el modelo determinista HIRLAM HNR (5 km) y, más recientemente y hasta la actualidad (2018), el modelo determinista de mesoescala HARMONIE-AROME (2.5 km). Para la predicción probabilista en el corto plazo, AEMET es pionera en el desarrollo y ejecución experimental a escala diaria del SPC AEMET-SREPS, de resolución 25 km y con 20-25 miembros basado, como se describe más adelante, en la integración de 5 modelos numéricos de área limitada diferentes forzados con las predicciones de 5 modelos globales distintos.

El sistema AEMET-SREPS se basó característicamente en la técnica llamada multimodelo (sec. 13.5 en la página 174), siendo AEMET pionero en Europa y uno de los primeros centros en el mundo en desarrollar

un SPC de este estilo. Como sistema focalizado en predicciones a corto plazo, hasta 72 horas de alcance predictivo, se pretendía muestrear las incertidumbres y errores de los modelos, las condiciones iniciales y de contorno, demostrando AEMET-SREPS ser capaz de generar una dispersión muy realista. Se esperaba entonces que AEMET-SREPS fuese una herramienta complementaria a los modelos deterministas de mesoescala, emergentes en aquella época, como ya se ha dicho, e. g. HIRLAM HNR (5 km) y HARMONIE-AROME (2.5 km), focalizado en la predicción en el corto plazo de fenómenos extremos y fenómenos adversos (secs. 27.7 en la página 419 y 42.2 en la página 637), e. g. temporales, precipitaciones intensas, tormentas de nieve, etc.

Según la época, el sistema usaba 4-5 modelos deterministas regionales o de área limitada (sec. 10.1 en la página 130): HIRLAM (del consorcio HIRLAM, cap. 20 en la página 301), HRM (del servicio meteorológico alemán DWD), UM (del servicio meteorológico británico UKMO), MM5 (de la Universidad de Pennsylvania y el Centro Nacional americano para Investigaciones Atmosféricas PSU/NCAR) y LM COSMO (del consorcio COSMO, cap. 20 en la página 301). Estos modelos, además, trabajaban con condiciones iniciales y de contorno provistas por 4-5 modelos deterministas globales (sec. 10.1 en la página 130): ECHRES o HRES-IFS (del Centro Europeo de Predicción a Medio Plazo ECMWF, sec. 19.2 en la página 291), UM (del servicio meteorológico británico UKMO), GME (del servicio meteorológico alemán DWD), GFS (del Centro Nacional americano para predicciones ambientales NCEP), CMC (del ser-

vicio meteorológico canadiense MSC) y GSM (de la agencia meteorológica japonesa JMA). De este modo, tanto los errores en las condiciones iniciales como los errores del modelo se simulaban en el sistema (sec. 13.8 en la página 172).

La verificación del sistema AEMET-SREPS para contrastar su reproducción del flujo atmosférico a gran escala (sinóptico), realizada usando el análisis del ECMWF, mostraba que se trata de un sistema consistente y ligeramente subdispersivo[21] (como la mayoría de los sistemas en esa época). Para los parámetros de tiempo sensible (precipitación, viento, temperatura, etc.) que son los más importantes en relación al alto impacto social, el sistema mostraba una alta calidad (calidad, destreza o pericia, sec. 15.2.3 en la página 211) en su respuesta frente a eventos binarios (normalmente superación de umbrales), mostrando buenos resultados en propiedades como la fiabilidad (sec. 15.9.4 en la página 226), la resolución probabilista (sec. 15.9.5 en la página 229) o la discriminación (sec. 15.9.6 en la página 231). Para verificar las predicciones de precipitación se utilizaron observaciones de las redes climatológicas nacionales europeas, mallas (*gridding*) convenientemente y se comparó con el SPC del ECMWF (ECENS, sec. 19.3 en la página 293). El periodo verificado que se muestra en este capítulo cubre los años 2007–2008 y corresponde a la pasada de 00 UTC con 20 miembros a 0.25° de resolución horizontal.

El proyecto AEMET-SREPS (A-S abreviadamente para el resto del capítulo) comenzó en 2001-2002 y tuvo carácter oficial desde 2005. En 2006 corría todos los días a las 00 UTC el primer conjunto de modelos con diferentes condiciones iniciales y de contorno. Durante 2007 alcanzó un punto de madurez, corriendo

20 miembros con buen rendimiento a una resolución horizontal de 0.25° y 40 niveles en la vertical. En 2008 se añadió una segunda pasada a las 12 UTC. En 2009 se incluyó el modelo global CMC, incrementando el número de miembros a 25. A partir de 2010 y hasta 2014 el sistema corría regularmente dos veces al día, 00 y 12 UTC hasta un alcance predictivo de 72 horas con 25 miembros. La última pasada realizada en el Cray-X1E tuvo lugar en noviembre de 2014 por la terminación operativa de aquel supercomputador en AEMET.

21.2 Metodología

Siguiendo los resultados del Experimento de Tempestades y Ensembles de Mesoescala (Storm and Mesoscale Ensemble Experiment, SAMEX) llevado a cabo en EE. UU. en 1998 [29] y con la inestimable orientación de la Doctora EUGENIA KALNAY, del NCEP, experta y pionera mundial en sistemas de predicción por conjuntos (SPC, cap. 13 en la página 165), AEMET eligió combinar las técnicas multimodelo, multi-análisis y multicontorno para el desarrollo de un SPC para el corto plazo que se llamaría AEMET-SREPS, pensado para proporcionar predicciones probabilistas. Este tipo de predicciones ofrecerían información complementaria a los modelos deterministas operativos en AEMET y permitirían mejorar las predicciones de fenómenos adversos y fenómenos extremos como temporales, precipitaciones intensas, tormentas de nieve u olas de calor. La emisión de avisos de fenómenos adversos es una de las tareas más importantes de los servicios meteorológicos, por el impacto y daño potencial que estos eventos pueden provocar.

Característica	Configuración
Muestreo de error del modelo	Multimodelo (LAM)
Muestreo de error de condiciones iniciales	Multianálisis (GCM)
Muestreo de error de condiciones de contorno	Multicontorno (GCM)
Número de modelos regionales y globales y de miembros	5 LAM x 5 GCM = 25 miembros
LAM para multimodelo	HIRLAM, HRM, LM, MM5, UM
GCM para multianálisis y multicontorno	GEM, GME, HRES-IFS, GSM, GFS
Pasadas	2 al día: 00 y 12 UTC
Rango de alcances predictivos	Hasta 72 horas (3 días)
Resolución horizontal	0.25° ~ 25 km
Resolución vertical	40 niveles

Tabla 21.1: Configuración de AEMET-SREPS en su época de madurez 2009-2014.

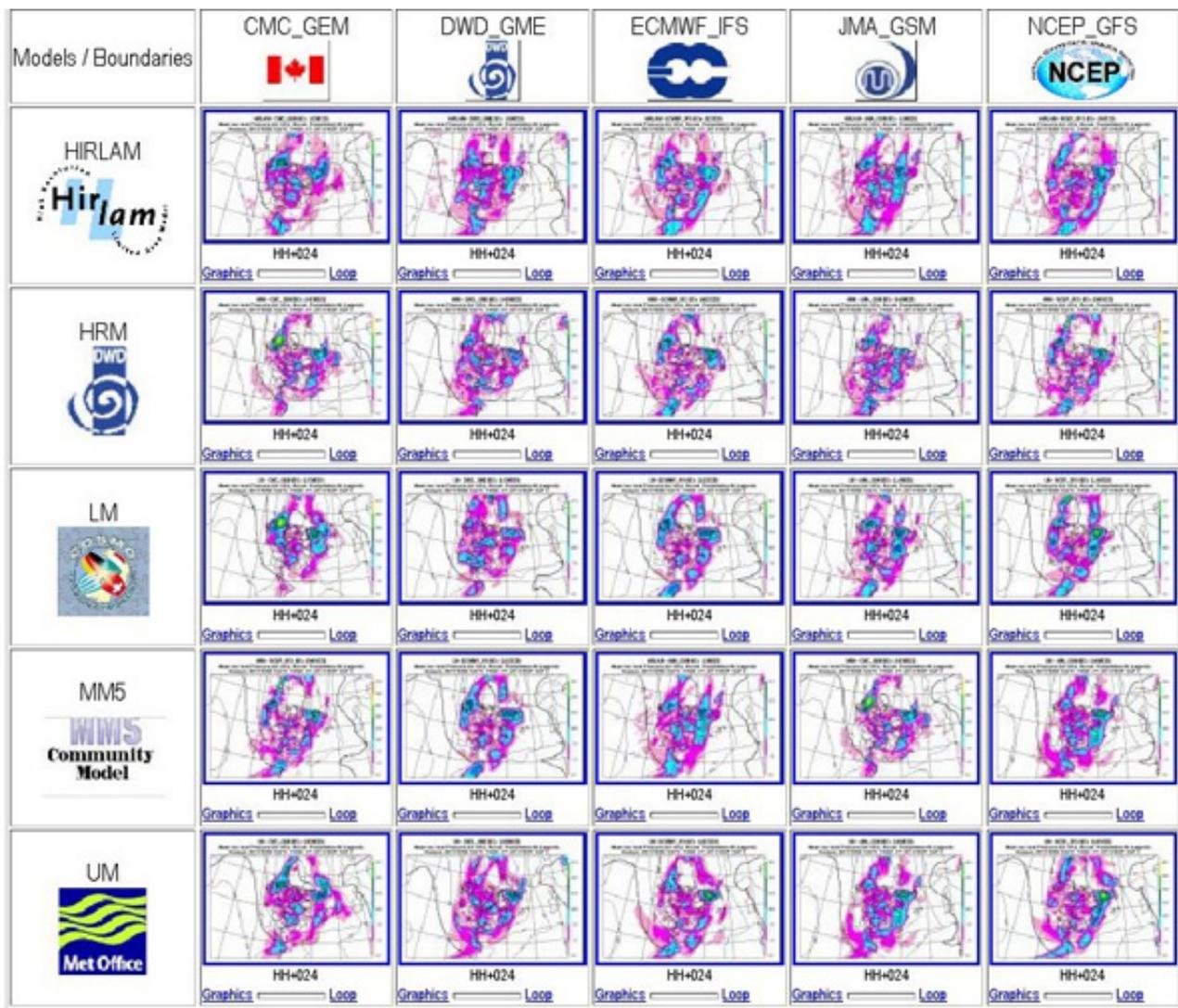


Figura 21.2: Predicciones individuales generadas por los 25 miembros de AEMET-SREPS, en este caso podemos ver predicciones de precipitación. En filas se despliegan los diferentes modelos regionales integrados en el supercomputador y en columnas los diferentes modelos globales usados como condiciones iniciales y de contorno.

21.2.1 Multimodelo

La técnica multimodelo elegida para A-S requirió instalar y hacer funcionar una serie de modelos regionales o de área limitada (sec. 10.1 en la página 130) en el supercomputador de la época en AEMET, un Cray-X1E (sec. 11.4 en la página 150). Los modelos utilizados fueron los siguientes:

- High-Resolution Limited Area Model, HIRLAM [42, 56], del consorcio del mismo nombre (cap. 20.1 en la página 302).
- High-Resolution Regional Model, HRM [37, 38], del servicio meteorológico alemán Deutsche Wetterdienst (DWD).
- Mesoscale Model version 5, MM5 [17, 24], de la Pennsylvania State University y el Centro Nacio-

nal americano para la Investigación Atmosférica NCAR.

- Unified Model, UM [14], del servicio meteorológico británico UKMO.
- LM COSMO Model, LM [15], del consorcio COSMO (sec. 20.1 en la página 302).

Cada uno de estos modelos tenía unas características diferentes, tales como diferentes núcleos dinámicos (sec. 10.4 en la página 135), diferentes parametrizaciones físicas (sec. 10.5 en la página 137), etc. Para poder inicializar y conducir a cada uno de ellos con diferentes modelos globales se requirió desarrollar una serie de herramientas adicionales de preproceso, tales como interpolación vertical para distintos tipos de niveles, interpolación horizontal en diferentes mallas, topografías diferentes, codificación para diferentes variedades de ficheros de datos (diferentes estilos de

Gridded Binary, GRIB) provistos por los diferentes centros emisores de modelos globales, variables especiales que no todos los modelos proveen (un típico caso es la humedad relativa a partir de la humedad específica, la presión y la temperatura), intentando siempre minimizar el ruido que las diferentes etapas de preproceso podían introducir en los datos originales. Una vez aplicadas todas las etapas necesarias de preproceso se disponía de condiciones iniciales y de contorno en formato, número y tipo homogéneos y adecuados para correr cada modelo regional. Un sistema de esta envergadura y heterogeneidad interna implica una enorme cantidad de trabajo y de recursos.

21.2.2 Multianálisis y multicontorno

Las condiciones iniciales y de contorno se tomaban de predicciones de 4-5 modelos deterministas globales distintos. La hora del análisis de estos modelos globales era 12 horas antes del tiempo de partida del A-S, de modo que si el análisis de los modelos globales era a las HH UTC, los modelos regionales miembros de A-S se inicializaban con predicciones de estos modelos globales válidas para HH+12 UTC. Esta estrategia podía devaluar el rendimiento determinista de cada miembro, dado que no se usaban las observaciones más recientes, sino observaciones con una antigüedad de 12 horas. Sin embargo, al no realizar asimilación de las observaciones cada modelo regional (arranque en frío), se ganaba un cierto tiempo y se podía disponer de las salidas del A-S con antelación. Es decir, se asumía esta ligera devaluación en rendimiento a cambio

de disponer de las predicciones con cierta antelación y, de paso, liberaba al A-S de posibles problemas derivados del propio *spin-up* de las condiciones iniciales de los modelos globales.

Las condiciones iniciales y de contorno se obtenían de los siguientes modelos deterministas globales:

- ECHRES o HRES Integrated Forecast System [30, 52], del Centro Europeo de Predicción a Medio Plazo, European Centre of Medium Range Weather Forecasts, ECMWF (sec. 19.2 en la página 291).
- Global Unified Model, UM [14], del servicio meteorológico británico, UKMO
- Global Forecast System, GFS [49, 50], del centro nacional americano para Predicciones Ambientales, NCEP.
- Global Model, GME [39], del servicio meteorológico alemán, DWD.
- Global Canadian Model, CMC [13], del servicio meteorológico de Canadá, MSC.
- Global Spectral Model, GSM [53], de la agencia meteorológica japonesa, JMA, que sustituyó a UM desde 2011.

Todos estos modelos globales incluyen esquemas de asimilación variacional: IFS, UM y CMC usaban 4DVAR, GFS y GME usaban 3DVAR (sec. 16.6 en la página 247) y parametrizaciones físicas de reciente desarrollo (sec. 17.4.4 en la página 269). El uso de estos modelos como condiciones iniciales y de contorno garantizaba una parte importante de la calidad del A-S.

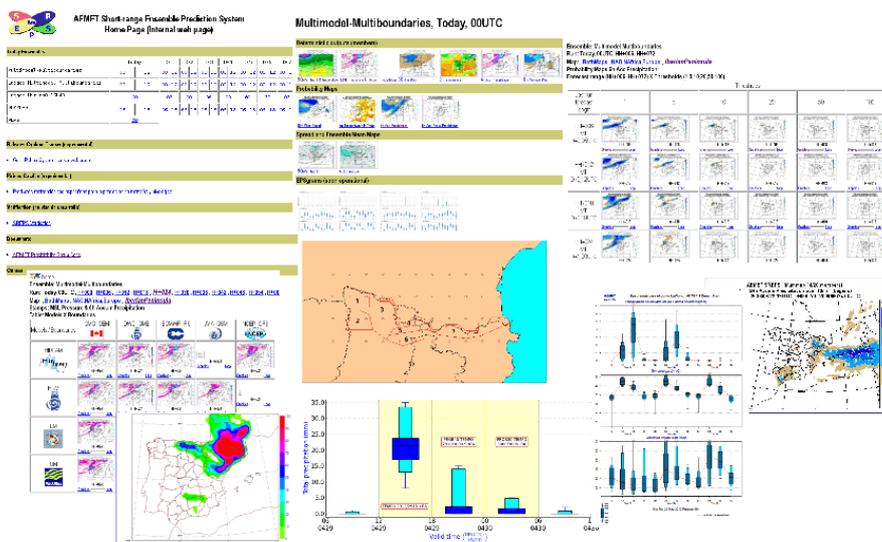


Figura 21.3: Monitorización del sistema AEMET-SREPS en un sitio web interno de AEMET.

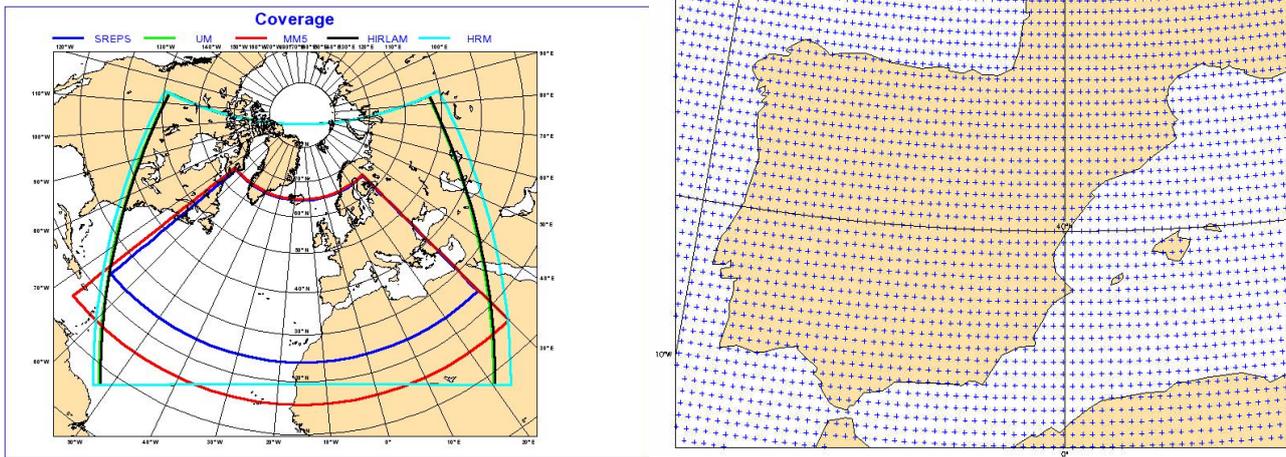


Figura 21.4: Izquierda: Dominios de integración de los diferentes modelos regionales del AEMET-SREPS: HRM y COSMO (cyan), MMS (rojo), HIRLAM (negro) y UM (verde). Dominio de posprocesado y verificación (azul) con malla regular lat-lon de 0.25° , que cubre parte del Atlántico Norte, África septentrional, Europa y el mar Mediterráneo. Derecha: malla común de 0.25° , 25 km aproximadamente, sobre la península ibérica y Baleares.

21.2.3 Posprocesado

Los SPC proporcionan una ingente cantidad de información que, sobre todo en un entorno operativo, es preciso procesar y sintetizar para ofrecer al predictor o usuario final una información condensada y útil. Por otro lado, las salidas directas del SPC pueden mejorarse mediante métodos estadísticos (cap. 14 en la página 193), de modo que tenemos distintos tipos de posprocesado. Los métodos de posprocesado son, por tanto, imprescindibles.

El primer paso en el posprocesado de A-S consiste en la interpolación espacial horizontal de los diferentes dominios de integración a un área y malla comunes, Figura 21.4. El área es la intersección máxima de las áreas de los modelos. Dado que las mallas de los modelos eran diferentes, algunas regulares latitud-longitud, otras rotadas y otras Lambert conformes, la intersección de las áreas era sensiblemente menor que cada una de las mismas. Por tanto, se decidió ampliar las diferentes áreas de integración para maximizar este área común. Finalmente, este área común (Figura 21.4, izquierda) cubría parte del Atlántico Norte, parte de África septentrional, Europa y el mar Mediterráneo. La malla era regular latitud-longitud a 0.25° , aproximadamente 25 km en estas latitudes medias (Figura 21.4, derecha). Se computaban en esta área y malla predicciones o productos probabilistas de distintos parámetros meteorológicos, partiendo de la base de *equiprobabilidad* de los miembros (sec. 13.7.2 en la página 191). Estas predicciones precisaban de una verificación regular adecuada (cap. 15 en la página 207). El sistema se monitorizaba

en un sitio web interno de AEMET en tiempo real. En la Figura 21.2 pueden verse como ejemplo unos sellos de precipitación acumulada y, en la sec. 27.2.1 en la página 402, unos sellos de altitud geopotencial y temperatura en 500 hPa.

21.3 Otras investigaciones

Durante toda la etapa de desarrollo y funcionamiento diario del A-S se probaron diversas configuraciones experimentales para probar su rendimiento, calidad, plausibilidad y conveniencia. Se probaron también otras metodologías totalmente diferentes para la generación de predicciones probabilistas o para la mejora de las salidas de las ya existentes. Mencionaremos aquí brevemente algunas de las investigaciones más relevantes.

21.3.1 Calibración estadística

Mediante el posprocesado de calibración estadística pueden mejorarse las salidas de un modelo o de un sistema de predicción por conjuntos (cap. 14 en la página 193). En el grupo de predicibilidad se realizaron diversas investigaciones al respecto, probando fundamentalmente la técnica de Promediado Bayesiano del Modelo, Bayesian Model Averaging, BMA [46] para calibrar la PDF del SPC mediante las observaciones (sec. 13.23 en la página 190).

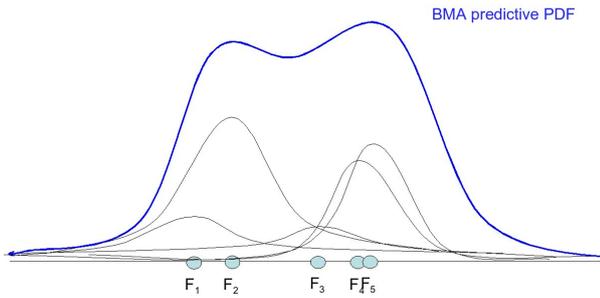


Figura 21.5: Fundamento de la técnica de calibrado bayesiano BMA. Dadas las predicciones de cinco miembros F_i , previamente se les quita el sesgo determinista, después se asigna una PDF continua a cada una de ellas. Después, con el método E-M que optimiza el CRPS[28] se construye una PDF conjunta en dos pasos: la PDF del AEMET-SREPS calibrado (ver texto).

El fundamento de la técnica de calibrado bayesiano BMA, una de cuyas partes se muestra en la Figura 21.5 izquierda, se realiza en dos pasos:

1. Dadas las predicciones de, por ejemplo, cinco miembros F_i , se hace una corrección de sesgo individualmente.

2. Se atribuye una PDF continua, por ejemplo gaussiana, a cada una de las predicciones corregidas de sesgo y se construye una PDF conjunta por suma ponderada de las PDF individuales, aplicando el algoritmo *estimation-maximization*, E-M[23, 43], que asigna pesos a las PDF individuales optimizando una medida probabilística denominada Continuous Ranked Probability Score [28, 58] (sec. 15.9.9 en la página 232).

Este método se aplicó y estudió en varias ocasiones [18, 48]. Los resultados han venido mostrando con claridad [18] que, aunque las medidas estadísticas mejoraban notablemente, se perdía capacidad predictiva en los fenómenos extremos. En la Figura 21.6 derecha, se muestra para el episodio de vientos fuertes del 23-01-2009 en Cataluña cómo el A-S calibrado no mejoraba en capacidad predictiva, aunque su rendimiento medido estadísticamente para varios meses fuese mejor. En el caso particular de vientos fuertes, con un sesgo negativo al ser simulados más débiles por el SPC, eran corregidos erróneamente de un sesgo positivo general, debido a que los vientos flojos mayoritarios en las observaciones usadas en la calibración eran superiores a los observados. Los fenómenos extremos siguen siendo uno de los caballos de batalla del mundo de la predicción.

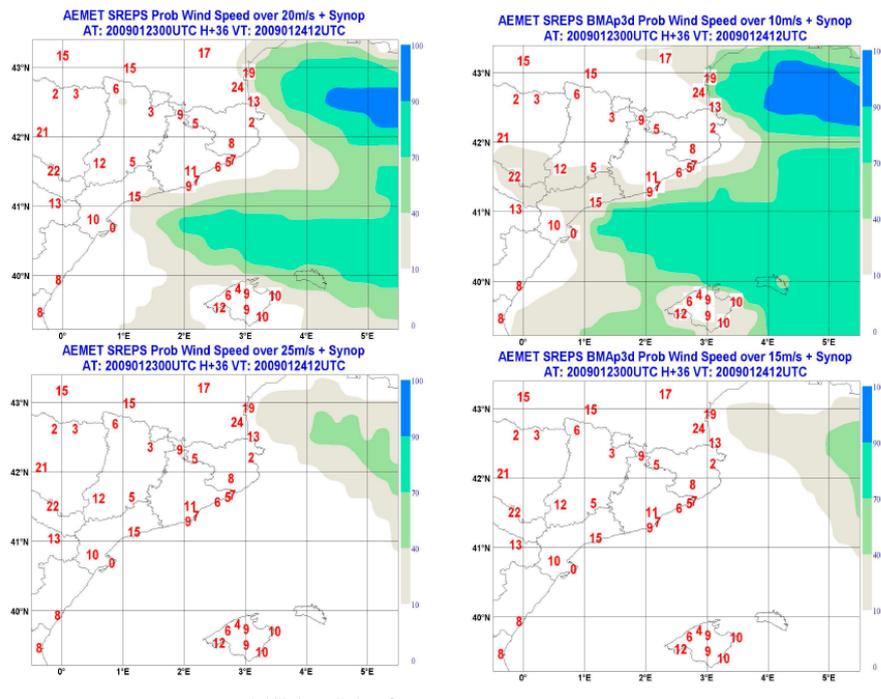


Figura 21.6: Mosaico de cuatro mapas de probabilidad de viento por encima de distintos umbrales para el episodio del 23-01-2009 en Cataluña. A la izquierda el A-S crudo y a la derecha el A-S calibrado con BMA pero mostrando umbrales algo más bajos. Aunque la calibración mejora, estadísticamente, el rendimiento del SPC para eventos extremos como éste se pierde capacidad predictiva. Los fenómenos extremos siguen siendo un problema (ver texto).

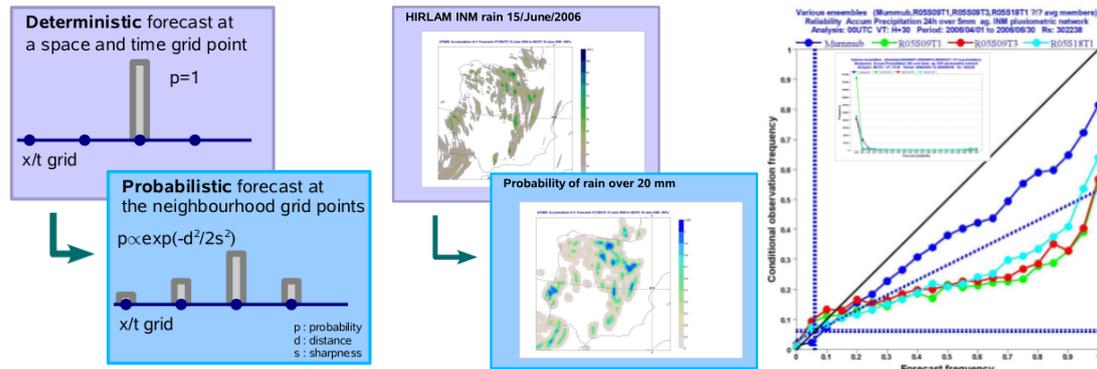


Figura 21.7: Método del SPC difusivo para generar predicciones probabilistas a partir de un modelo determinista [55] aplicado en AEMET [51] sobre el modelo HIRLAM. A la izquierda, el fundamento del método consiste en dispersar la información determinista en el espacio y en el tiempo. En el centro, precipitación acumulada en 24 h de HIRLAM 0.05° para el 15-06-2006 y la probabilidad de precipitación mayor que 20 mm en 24 h correspondiente a la misma fecha, con radio de vecindario de 35 km en este caso. Finalmente, a la derecha, un diagrama de atributos o fiabilidad donde se comparan tres métodos difusivos con distintos parámetros comparados con el A-S completo de aquella época. El A-S daba un rendimiento superior, tanto en fiabilidad, como en resolución y discriminación (ver texto).

21.3.2 SPC difusivo

Mediante las técnicas denominadas de *vecindario*, puede construirse un SPC de bajo coste usando las salidas de un modelo determinista [55]. Éstos tienen al menos algunos errores espaciales y temporales que pueden interpretarse como un desfase en la descripción de los eventos. El *método difusivo* dispersa las predicciones deterministas en el espacio y en el tiempo usando alguna distribución de probabilidad a priori y proporcionando predicciones probabilistas. El *tamaño de vecindario* es un parámetro clave: a medida que se hace pequeño el resultado tiende a las predicciones deterministas y si se hace demasiado grande se dispersa en exceso la información, perdiendo calidad. El método se aplicó a salidas de precipitación y de viento del modelo HIRLAM operativo en AEMET, en sus distintas resoluciones de 0.16° y 0.05°, en el periodo AMJ2006 [51]. Se verificó contra las observaciones de la red pluviométrica de AEMET (INM en esa época). Se utilizó la distribución normal para la difusión, generando una PDF discreta. En la Figura 21.7 se muestra, a la izquierda, el fundamento del método; en el centro arriba, precipitación acumulada en 24 h de HIRLAM 0.05° para el 15-06-2006 y, abajo, la probabilidad de precipitación mayor que 20 mm en 24 h correspondiente, con radio de vecindario de 35 km y, finalmente, a la derecha, un diagrama de atributos (sec. 15.4 en la página 226) donde se comparan tres métodos difusivos con distintos parámetros comparados con el A-S completo de aquella época. El A-S daba un rendimiento superior, tanto en fiabilidad, como en resolución y discriminación.

21.3.3 SPC híbrido

Cuando se combinan un SPC con un modelo determinista de mayor resolución para formar una predicción probabilista híbrida, se está intentando combinar la dispersión e información de incertidumbre del SPC con la información de mayor detalle y precisión dada por el modelo determinista debido a su mayor resolución. Esta configuración o diseño vino a llamarse *ensemble híbrido* [16]. Entrando un poco más en detalle, cada miembro del SPC se considera como suma de una base, a menudo el miembro sin perturbar y una perturbación. El modelo determinista de mayor resolución podría considerarse como una mejor base de modo que se construía cada miembro del SPC híbrido sumando y restando la perturbación al modelo determinista. Pueden surgir numerosos problemas, debido principalmente a las diferencias en resolución: se requieren interpolaciones horizontales y verticales que introducen ruido en el proceso y pueden por tanto reducir el rendimiento. El método del SPC híbrido se probó usando una combinación del HIRLAM operativo en AEMET en aquella época como modelo determinista y de algunos subconjuntos de AEMET-SREPS [7], Figura 21.8 en la página siguiente. En principio, el SPC híbrido no mostraba mejores medidas de verificación que A-S, salvo en casos puntuales, frecuencias muy bajas y siempre con una elección cuidadosa del subconjunto adecuado del A-S, siguiendo las indicaciones del estudio original [16].

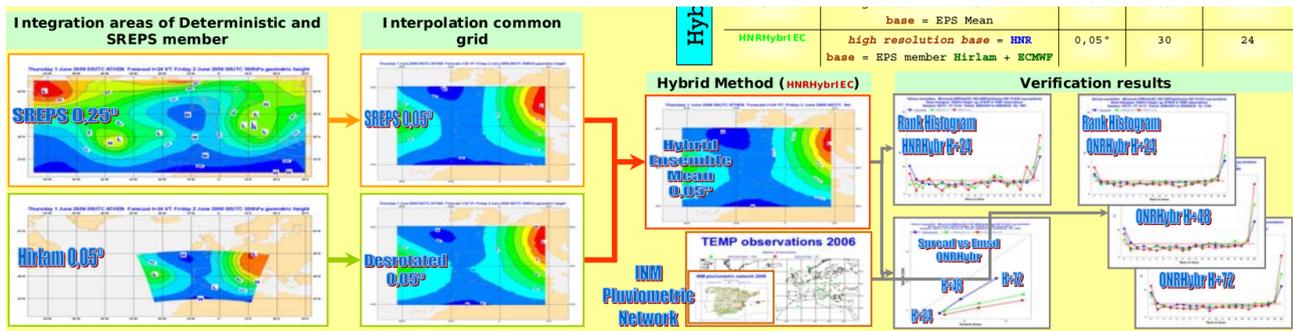


Figura 21.8: Método del SPC híbrido [16] para combinar un modelo determinista que aporte su mayor resolución y un SPC que aporte su dispersión y por tanto su información sobre incertidumbre, aplicado en AEMET sobre HIRLAM y AEMET-SREPS [7]. A la izquierda, combinación de los diferentes dominios en un dominio común. En el centro arriba, la base constructiva del SPC híbrido y, abajo, las observaciones usadas para verificar. A la derecha, gráficas de verificación que muestran cómo A-S rinde mejor en casi cualquier circunstancia que las diferentes configuraciones de SPC híbrido investigadas (ver texto).

21.3.4 Error en las observaciones

En verificación de la predicción se asumía tradicionalmente que el error observacional es despreciable comparado con el error en la predicción. Aunque esta asunción puede ser cierta para plazos de predicción medios y largos no tiene por qué serlo en general; más aún, con las mejoras de los modelos, el error relativo de las predicciones con respecto al de las observaciones va disminuyendo. En un marco de verificación más amplio y general las observaciones, al igual que las predicciones, pueden ser descritas junto con su error o incertidumbre asociados. Normalmente, cuando no se considera el error observacional, el rendimiento de los SPC tiende a subestimarse. Por tanto, adquiere relevancia creciente el hecho de incorporar esa incertidumbre en la verificación. El método general de CANDILLE y TALAGRAND denominado *probabilidad observacional* [9] se estudió en profundidad para la precipitación en un estudio de SANTOS y GHELLI [47], del que se muestra un resumen en la Figura 21.9 en la página siguiente. Cuando se tiene en cuenta el error observacional se evita subestimar el rendimiento de los SPC, que resulta más alto de lo acostumbrado. La Figura 21.9 en la página siguiente izquierda muestra el concepto de probabilidad observacional. En cada caja de malla se toman las observaciones presentes y se construye una distribución de probabilidad que representa la precipitación estimada en esa caja, a partir de la que se infieren las probabilidades. La Figura 21.9 en la página siguiente derecha muestra una parte del ejercicio de verificación realizado para probar esta teoría. Pueden verse series temporales de BSS de precipitación acumulada en 24 h superior a 10 mm y sus componentes para AEMET-SREPS en la Península y Baleares para el periodo usual en este capítulo

y alcance T+54. El BSS, orientado positivamente, en trazo grueso continuo para el método tradicional y en trazo grueso discontinuo para el método de probabilidad observacional que tiene en cuenta el error en las observaciones. Las componentes del BSS, orientadas negativamente, se muestran con símbolos intercalados, fiabilidades con triángulos y resoluciones con círculos. Puede comprobarse cómo el rendimiento medido teniendo en cuenta el error observacional es superior.

21.3.5 Time lagged ensemble con AEMET-SREPS

Una de las configuraciones que se probó en diversas ocasiones y durante largos periodos de tiempo es del tipo *time lagged ensemble* [36] (sec. 13.5 en la página 174), consistente en juntar salidas de miembros de pasadas consecutivas para formar un superensemble. Por ejemplo, si cada pasada de las dos (00 y 12 UTC) produce un ensemble de 20 miembros, pueden tomarse los 20 miembros de la última pasada y los 20 de la penúltima y formar un superensemble de 40 miembros. Este tipo de sistemas no implica coste computacional adicional en lo referente a los modelos en sí, aunque sí que lo requiere en el posproceso para la computación de productos probabilistas que, en cualquier caso, es un coste relativamente pequeño. Los ensemble time lagged son baratos en ese sentido y suelen dar resultados positivos. Se proyectó aumentar el número de pasadas a cuatro (00, 06, 12 y 18 UTC), existiendo en ese caso la posibilidad de construir un superensemble de 80 miembros combinando los 4 ensembles más recientes (HH, HH-6, HH-12, y HH-18), aunque no se llegó a poner en práctica esta posibilidad.

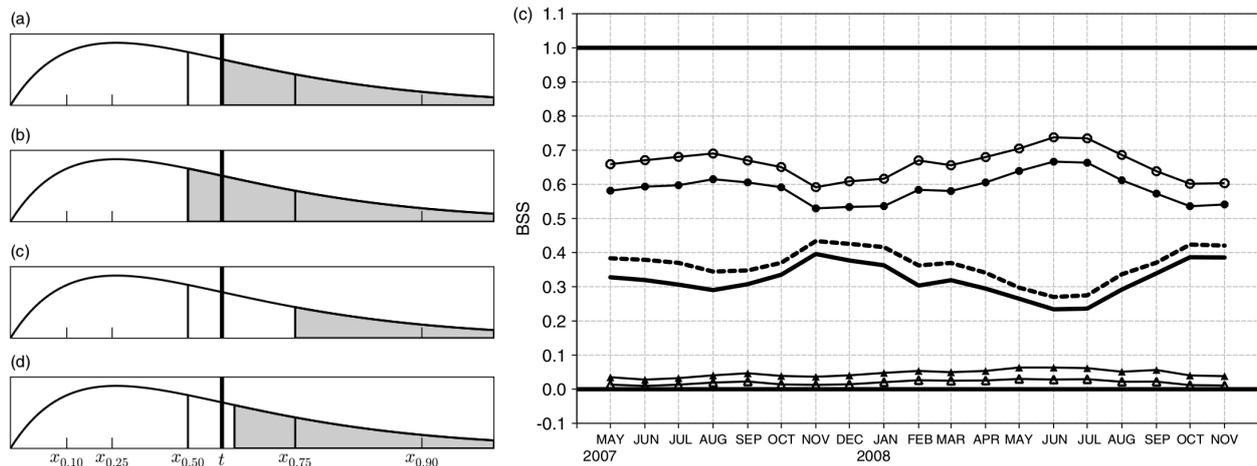


Figura 21.9: Método de probabilidad observacional de CANDILLE y TALAGRAND [9] aplicado aquí a la precipitación por SANTOS y GHELLI [47]. Al tener en cuenta el error en las observaciones, el rendimiento medido del SPC es superior; en este caso AEMET-SREPS (ver texto). Figura de SANTOS Y GHELLI, 2012, QJRMMS [47].

21.4 Verificación: consistencia del flujo a gran escala

La verificación de un ensemble trata de evaluar en qué grado la predicción probabilista es buena dado que exacta no es por naturaleza, dado que es probabilista. Técnicamente, mediante diferentes medidas se evalúan diversas propiedades que presenta el SPC al compararlo con las observaciones o con el análisis correspondiente: la calidad mediante medidas como la fiabilidad o la resolución y el valor con medidas como el valor económico relativo de la predicción, es decir, el beneficio relativo que aportaría a un usuario la utilización de la predicción probabilista.

Para contrastar la consistencia general del sistema a gran escala, digamos escala sinóptica, se comparan campos dinámicos previstos con los correspondientes campos del análisis. La ventaja de usar un análisis es que este cubre todo el dominio de integración (no así las observaciones) con el mismo peso, lo que evita dar prioridades a las zonas de tierra donde la densidad de observaciones es bastante superior. La verificación frente a observaciones (surface synoptic observations (SYNOP), TEMP) da, como es natural, peores resultados pero cualitativamente similares. Los campos dinámicos son aquellos que reflejan la dinámica atmosférica a gran escala sin sufrir influencias de menor escala como, por ejemplo, la orográfica. Se tomaron campos de presión reducida al nivel del mar y altura geopotencial en 500 hPa y se compararon con el análisis operativo del ECMWF.

El periodo de verificación abarca desde abril de 2007

hasta diciembre de 2008, el mismo que se muestra en la sección 21.5.1 en la página 325 para las variables de tiempo sensible. Dada la problemática usual en fases de desarrollo, no puede disponerse del sistema todos los días al 100%, de modo que del periodo señalado se seleccionaron 614 pasadas de 00 UTC. La malla de verificación consta de $380 \times 164 = 62\,320$ puntos. De este modo, el tamaño muestral para evaluar la significación de cada medida en cada alcance de predicción es $62\,320 \times 614 \sim 3,8 \times 10^7$. El horizonte predictivo alcanza hasta 72 h y se verifican alcances cada 6 horas desde T+06 hasta T+72, es decir 12 alcances. Se estudiaron diferentes aspectos del rendimiento de las predicciones que mostramos en las siguientes subsecciones: la calidad individual de los miembros, la consistencia estadística con el análisis y, por último, la dispersión del ensemble.

21.4.1 Rendimiento de los miembros individuales

En los SPC multimodelo, e. g. AEMET-SREPS, existen tareas extra en comparación con los SPC basados en un solo modelo, e. g. ECENS. La calidad o rendimiento individual de cada miembro como modelo determinista es condición necesaria, pero no suficiente, para garantizar la calidad del SPC en conjunto. Hay que tener en cuenta que en el posproceso del sistema se computan productos probabilistas basados en la equiprobabilidad de los miembros. Por tanto, el primer requisito es garantizar una calidad individual cuanto menos equiparable.

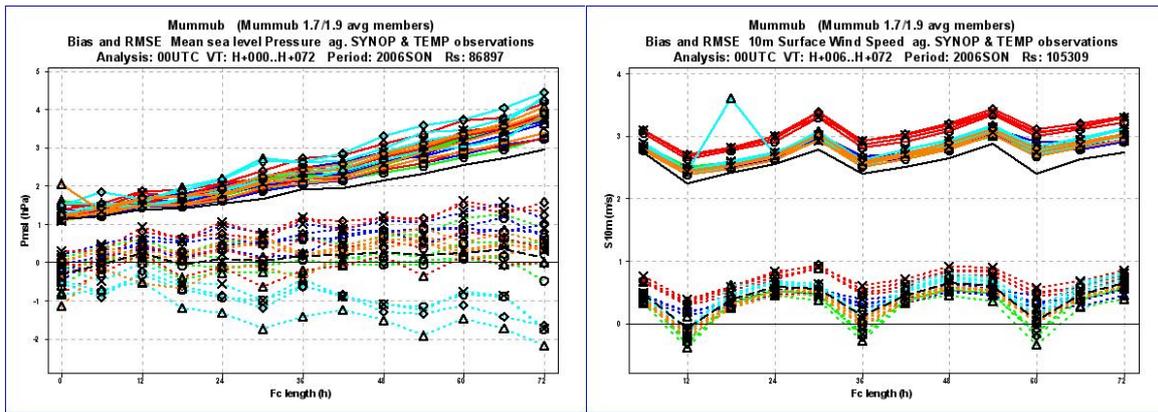


Figura 21.10: Evolución con el alcance de la predicción (eje horizontal) del error medio (trazo discontinuo) y la raíz cuadrada del error cuadrático medio (trazo continuo) de predicciones de presión reducida al nivel del mar (MSLP) computados para cada miembro del SPC (colores) y para el promedio del SPC (negro, trazo grueso). Izquierda: presión reducida al nivel del mar. Derecha: viento a 10 m.

También se espera que se cumpla una propiedad estadística conocida: el error cuadrático medio (MSE, sec. 15.2.3 en la página 211) del promedio del SPC ha de ser más pequeño que el MSE de cualquiera de los miembros [35, 44, 57, 60]. Para comprobar este aspecto, se computan el error medio (ME, sección 15.3 en la página 213) y la raíz cuadrada del error cuadrático medio (RMSE) de predicciones de variables dinámicas (Z500, T500, MSLP) para cada miembro y para el promedio del SPC. La Figura 21.10 muestra los resultados para MSLP. Tanto el ME como el RMSE del promedio del SPC están dibujados en trazo grueso para distinguirlos. Puede comprobarse el rendimiento similar de los miembros que garantiza una calidad equiparable, así como el menor RMSE del promedio. Como ya se ha insistido en otros capítulos, es una propiedad estadística solamente: el promedio del SPC no siempre es una situación meteorológica plausible, pero su rendimiento medido estadísticamente es superior al de los miembros individuales (sec. 13.6.3 en la página 178).

21.4.2 Consistencia estadística con las observaciones en el flujo a gran escala

En tanto que propiamente probabilista, un primer paso para verificar el sistema es el de comprobar la *consistencia* de las predicciones con las observaciones (o el análisis) en el flujo a gran escala. El llamado *histograma de rango* [1, 8, 25, 26, 27] puede usarse para comprobar si, en cada punto, el valor observado es estadísticamente indistinguible del conjunto de los valores previstos por el SPC. Equivalentemente, si tanto el valor observado como los previstos son igualmente probables como candidatos a la *verdad*. Si es así, el sistema se dice que es *estadísticamente consisten-*

te con las observaciones o *fiable*. Un sistema fiable presenta un histograma de rango con forma aproximadamente plana. En la Figura 21.11 izquierda podemos ver el histograma de rango correspondiente al periodo mencionado abril 2007 - diciembre 2008 para Z500 en el alcance T+24 y muestra consistencia a grandes rasgos, con una cierta cantidad de valores atípicos en los extremos, indicando una ligera subdispersión típica de los SPC operacionales de esa época. El hecho de estar promediando 21 meses seguidos podría esconder alguna variabilidad estacional. Para ello, en la Figura 21.11 derecha estratificamos la muestra en las diferentes estaciones, apareciendo tendencias claras: subdispersión en invierno (forma de U), fiabilidad en primavera, sobredispersión en verano (forma de campana) y una ligera subdispersión en otoño. Para otras variables y alcances predictivos los histogramas muestran tendencias similares.

21.4.3 Dispersión

La consistencia de las predicciones con las observaciones (o el análisis) está relacionada con la dispersión del SPC. Se espera que un SPC realice un muestreo adecuado de las incertidumbres subyacentes en la predicción, representada en cada punto por la dispersión de las predicciones (con respecto al promedio o con respecto al miembro de control si lo hay), así como brindar información cuantitativa y explícita de la predicibilidad de la atmósfera, que a posteriori puede ser representada en ese mismo punto por el error (e. g. el RMSE) del promedio de las predicciones comparado con la observación. Un SPC consistente muestra una relación aproximadamente lineal entre ambas cantidades, la dispersión y el error del promedio [5, 6, 57].

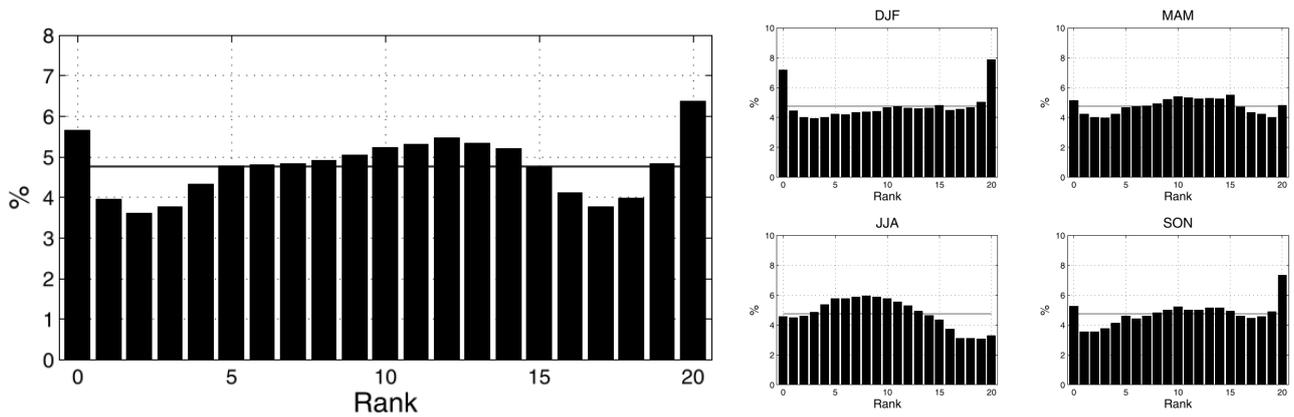


Figura 21.11: Histogramas de rango correspondientes al periodo abril 2007 - diciembre 2008 para Z500 en el alcance $T+24$, tomando como referencia el análisis del ECHRES. A la izquierda, el total. A la derecha, de arriba abajo y de izquierda a derecha, la estratificación estacional: invierno D2007-EF2008, primavera MAM2008, verano JJA2008 y otoño SON2008. [21] TellusA, CC BY 4.0.

Dado que A-S se compone de un producto cartesiano de modelos LAM y modelos GCM surge un aspecto que se ha estudiado por su particular interés: ¿qué parte de la dispersión de A-S proviene de los múltiples modelos LAM y qué parte de las múltiples condiciones iniciales y de contorno? Para responder a esta pregunta, basta con considerar la relación dispersión-error correspondiente a los diferentes subensambles que pueden formarse dentro de A-S: por un lado cinco subensambles multicontorno formados por un modelo LAM y cuatro globales y, por otro lado, cuatro subensambles multimodelo formados por cinco modelos LAM y un modelo global. Una vez formados estos subensambles, puede compararse la dispersión del A-S completo con la dispersión de los distintos subensambles. El resultado se muestra en la Figura 21.12

izquierda y centro, donde no se ha tenido en cuenta el posible impacto del tamaño de cada SPC [6, 19, 20], 20 para el A-S completo y 4 ó 5 para los subSPC. La Figura 21.12 muestra la relación dispersión-error para la variable MSLP en los alcances predictivos $T+6$ hasta $T+72$ cada 6 horas, tanto para el A-S completo (trazo grueso) como para los distintos subSPC (trazos finos), multicontorno a la izquierda y multimodelo en el centro. Como primer resultado, el sistema A-S completo muestra una relación dispersión-error claramente lineal, por lo que el sistema es consistente con el análisis. En segundo lugar, todos los subensambles muestran menos dispersión que el A-S completo, conclusión alcanzada en otros estudios como el proyecto SAMEX [29], el estudio de Arribas y col. [2] o el proyecto DEMETER [45] de predicción estacional.

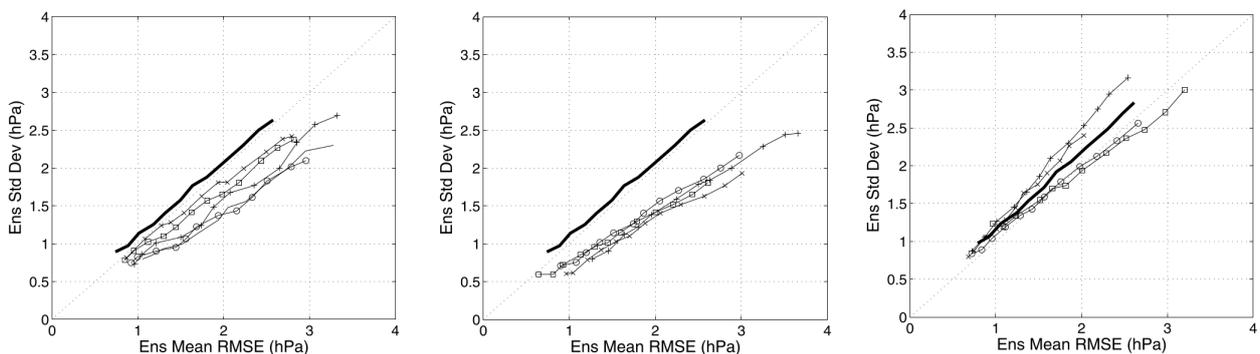


Figura 21.12: Diagramas dispersión-error para la variable MSLP en los alcances $T+6$ hasta $T+72$ cada 6 horas en el periodo mencionado en figuras anteriores. Para el sistema completo, multimodelo y multicontorno con 20 miembros, en trazo grueso en los tres diagramas. Izquierda: para los cinco subsistemas multicontorno en trazo fino. Centro: para los cuatro subsistemas multi-modelo en trazo fino. Derecha: para los cuatro estaciones del año en trazo fino. El sistema completo muestra mejor relación dispersión-error que cualquiera de sus subsistemas y el periodo completo mejor que cualquiera de las estaciones (ver texto). [21] TellusA, CC BY 4.0.

Como tercer resultado, los subSPC multimodelo son más subdispersivos que los subSPC multicontorno según avanzamos en el alcance de predicción, lo que nos lleva a concluir que la contribución de los modelos globales a la dispersión del sistema es algo superior a la contribución de los diferentes LAM, respondiendo a la pregunta que nos planteábamos anteriormente.

Para incluir tendencias estacionales que podrían quedar enmascaradas en el resultado de todo el periodo, mostramos también la estratificación estacional, esta vez sólo para el sistema A-S completo, tal como se hizo anteriormente para los histogramas de rango (sec. 21.11 en la página anterior). Los resultados son consistentes con los que para el alcance T+24 mostraban los histogramas de rango. En invierno el sistema se torna subdispersivo para alcances altos, en primavera muestra una clara sobredispersión ligeramente creciente con el alcance predictivo, en verano sobredispersión pero con menores valores tanto de dispersión como de error y, finalmente, en otoño el sistema muestra una relación dispersión-error muy lineal, claramente cerca de la diagonal, con valores menores a los de invierno-primavera y mayores que los de verano. Estos resultados son consistentes con la mayor predecibilidad esperable en verano y menor predecibilidad acompañada de mayor dispersión en primavera. Las tendencias estacionales se compensan parcialmente unas a otras para producir la curva más equilibrada del periodo completo.

21.5 Verificación: tiempo sensible

21.5.1 Tiempo sensible: respuesta frente a eventos binarios

Un SPC de corto plazo está enfocado principalmente en las variables de tiempo sensible (precipitación, viento, temperatura). La verificación correspondiente más adecuada mide la respuesta del sistema frente a eventos binarios, del tipo llueve/no-llueve, llueve por encima de 5 mm / no llueve por encima de 5 mm, etc., seleccionando un conjunto también adecuado de umbrales meteorológicos, umbrales característicos desde el punto de vista climatológico o predictivo.

21.5.2 Periodo, predicciones y observaciones

El periodo elegido cubre 21 meses consecutivos, desde abril 2007 hasta diciembre 2008. La verificación de

probabilidades tiene una base estadística y el ejercicio de verificación de un SPC es más significativo cuanto mayor es la muestra de datos; el período verificado aquí tiene una significación dentro de lo acostumbrado, tanto menor cuanto más alto el umbral, que es cuando el evento binario tiene lugar en un número de ocasiones menor y en esos casos aparece ruido en las curvas (fenómeno llamado *undersampling* [58]).

Para comparar con otro sistema predictivo de referencia, se seleccionó el SPC del ECMWF, ECENS (sec. 19.3 en la página 293). Al tratarse ECENS de un SPC óptimo para el medio plazo y AEMET-SREPS de un sistema óptimo para el corto plazo la comparación no es idónea; sin embargo, se eligió desde un punto de vista práctico al no tener que comparar A-S con otros sistemas de servicios meteorológicos nacionales. Se escogieron alcances de predicción comunes, con el propósito de minimizar cualquier posible ventaja de algún SPC sobre otro y, por tanto, alcances un poco tempranos para ECENS y algo tardíos para A-S, por la naturaleza de cada sistema. Por simplicidad, mostraremos detalladamente los resultados para precipitación, dejando sólo un breve resumen de la temperatura y el viento.

Para tener una representación adecuada en densidad y cobertura se usaron datos de las redes pluviométricas europeas de alta resolución, en la Figura 21.13 se muestra el ejemplo de cobertura de estas estaciones en España. Los métodos basados en interpolación de los valores previstos en los puntos de malla a los puntos de observación (*pointwise*, sec. 15.5 en la página 219), pueden adolecer por un lado, del impacto de la heterogénea densidad espacial de las observaciones y, por otro lado, de la potencial falta de significación debido a la dependencia entre estaciones cercanas.

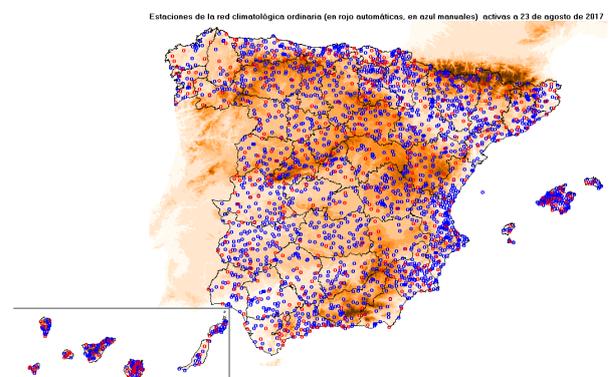


Figura 21.13: Red pluviométrica de estaciones en España. Las estaciones azules son pluviométricas puras y las rojas son termo-pluviométricas. AEMET 2017.

Para evitar ambos problemas se aplicó un método de *gridding* o *upscaling* [12] a las observaciones de precipitación. En este método se toman, para cada caja de malla (*grid box* en inglés), los valores observados dentro de la misma y se computa un representante, normalmente una ponderación adecuada con la distancia para evitar tomar el promedio puro. Sólo se consideran aquellas cajas con un número mínimo de 5 estaciones, para garantizar cierta consistencia. El ECMWF proporcionó estos datos ya mallados a $0.25^\circ \times 0.25^\circ$, malla de verificación del A-S.

Se comparan, así, predicciones de AEMET-SREPS (A-S) y de ECENS (E-E). Dado que las redes climatológicas proporcionan precipitación acumulada de 07 a 07 UTC, que la pasada disponible de A-S era la de la 00 UTC y que se disponía de salidas cada 6 h, se eligieron los alcances de T+30 y T+54 para computar la precipitación acumulada entre 06 y 06 UTC del día siguiente y poder comparar con E-E en alcances no demasiado tempranos (la otra opción era T+6 y T+30). No se tiene en cuenta la diferencia de 1 h entre la ventana de acumulación 06-06 prevista y la ventana 07-07 observada. El número medio de parejas predicción-observación comparadas es de 1 000 por día, mientras que para periodos de tres meses es de aproximadamente 76 000. Por su relevancia en predicción operativa se eligieron los umbrales de precipitación 1, 5, 10 y 20 mm.

21.5.3 Medidas de verificación y estrategia

En el contexto de la respuesta frente a eventos binarios, la calidad del sistema puede comprobarse midiendo diferentes propiedades [8, 10]: *fiabilidad*, correspondencia entre probabilidad del evento y frecuencia condicionada de observación del mismo, representada en diagramas de fiabilidad, tanto mejores cuanto más cercanos a la diagonal; *resolución*, capacidad del sistema para dar valores predictivos probabilistas con más resolución, también probabilista (valga la redundancia), que la propia climatología muestral y *discriminación*, capacidad del sistema para discernir entre la ocurrencia y no ocurrencia del evento, representada en curvas ROC (Relative Operating Characteristics [34]), tanto mejores cuanto más área presenten por encima de la diagonal. El *valor relativo* del sistema suele representarse en curvas de valor económico relativo, tanto mejores cuanto más área y cuanto más centrado esté el máximo (sección 15.10 en la página 234). Este conjunto de propiedades se describen detalladamente en el

capítulo 15 en la página 207 y en los textos fundamentales de verificación de predicciones [31, 32, 33, 58]. Se usó el índice de habilidad de Brier o Brier Skill Score (BSS) [4] y su descomposición en fiabilidad y resolución [32] como medidas netas de fiabilidad, resolución y habilidad. Un sistema con BSS mayor que 0 es más hábil que la climatología muestral y uno con BSS=1 indica un sistema perfecto y determinista. Se usó el índice de habilidad del área ROC o Roc Skill Area (RSA), igual a $2A - 1$, donde A es el área bajo la curva ROC, para medir la discriminación. Las áreas ROC se han computado por el método empírico del trapecioide. Una estimación paramétrica daría resultados con menos sesgo [32, 54, 59], mientras que el método empírico tiende a subestimar el área, especialmente para eventos correspondientes a umbrales altos, así como favorecer a SPC de tamaños superiores. La discriminación es una medida complementaria a la que proporciona el BSS. Mientras que el BSS es relativamente insensible a eventos extremos, el RSA no lo es [32]. En contrapartida, RSA puede ser insensible a algunos tipos de sesgos predictivos [34].

Para cada evento binario de exceder el umbral correspondiente, se computa la distribución conjunta de predicciones y observaciones, en los diferentes puntos de malla y distintos días del periodo, puntos \times días = casos. En cada caso, a las observaciones se les da el valor 0 ó 1 según se haya sobrepasado el umbral y dado el conjunto de predicciones del SPC, se asigna una probabilidad de exceder el umbral siguiendo la hipótesis de equiprobabilidad de LAPLACE: número de miembros que superan el umbral dividido entre número de miembros totales. Para SPC de N miembros, se hace una partición del espacio de probabilidades en N+1 intervalos o clases, para obtener medidas más precisas [60], es decir 52 clases para E-E y 21 clases para A-E. No se ha tenido en cuenta la diferencia de tamaños entre ambos SPC [6, 19, 20], aunque puede suponerse que favorece en todo caso a E-E y tener este hecho en cuenta. A partir de la distribución conjunta, plasmada en diversas tablas de contingencia, se obtiene una serie de medidas. Se estratificó en periodos de tres meses, utilizando la climatología muestral como referencia en cada estrato [40], sobre el dominio completo, construyendo así una serie temporal mensual con media móvil [58] de tres meses. No se calcularon intervalos de confianza para las medidas, aunque es conveniente hacerlo [11, 22, 31, 41] y se ha hecho en otros estudios posteriores.

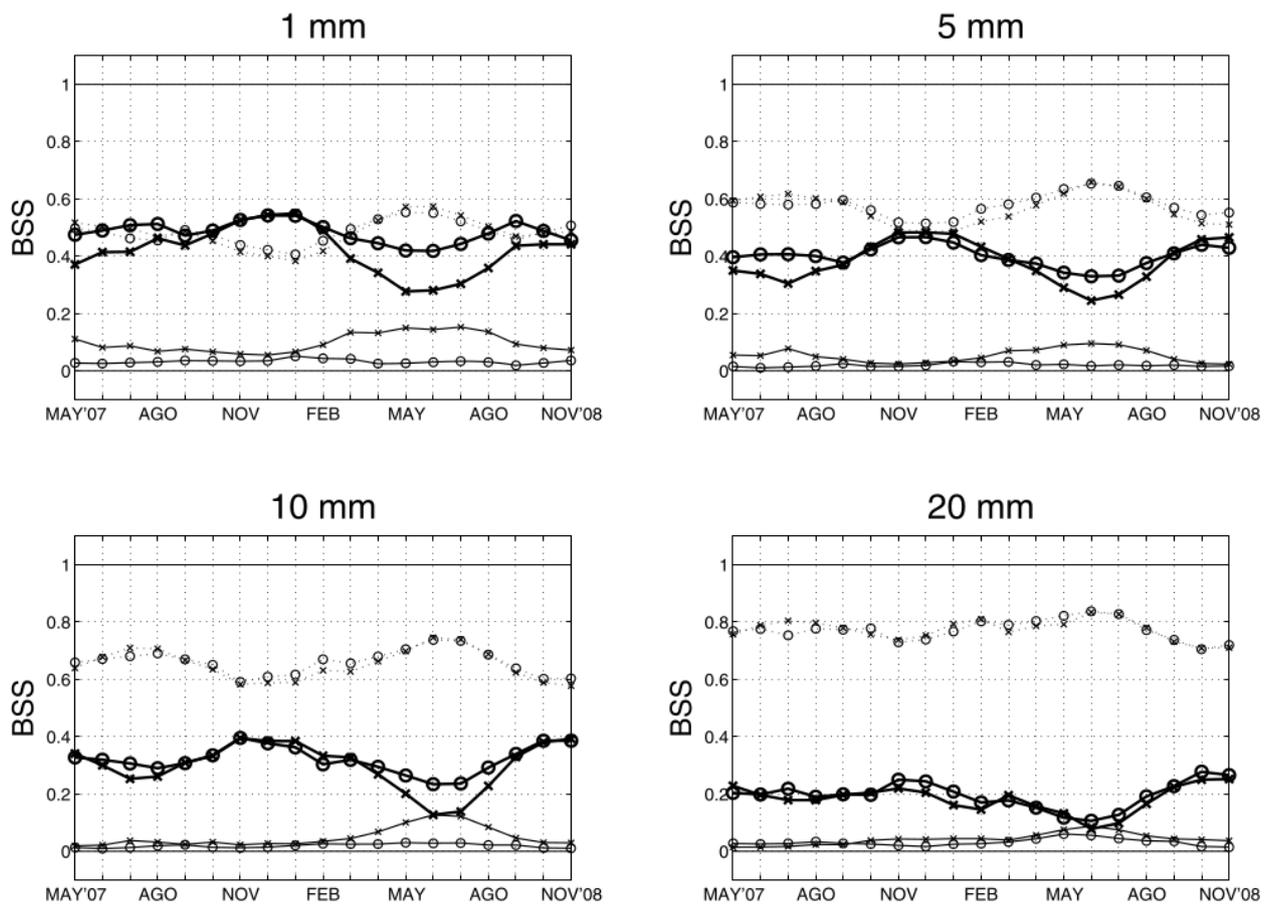


Figura 21.14: Series temporales de BSS (trazo grueso) y de sus componentes de fiabilidad (trazo fino) y resolución (trazo fino discontinuo), utilizando media móvil de tres meses, con centro desde mayo 2007 hasta noviembre 2008. Se muestran resultados de arriba abajo y de izquierda a derecha para los umbrales de 1, 5, 10 y 20 mm de precipitación acumulada en 24 h respectivamente. El alcance predictivo es de $T+54$ (acumulación de $T+30$ a $T+54$), AEMET-SREPS en círculos y ECENS en aspás. [21] TellusA, CC BY 4.0.

21.5.4 Resumen de resultados

La Figura 21.14 muestra series temporales de BSS (trazo grueso) y de sus componentes de fiabilidad (trazo fino) y resolución (trazo fino discontinuo), utilizando media móvil de tres meses, con centro desde mayo 2007 hasta noviembre 2008. Se muestran resultados de arriba abajo y de izquierda a derecha para los umbrales de 1, 5, 10 y 20 mm de precipitación acumulada en 24 h respectivamente. El alcance predictivo es de $T+54$ (acumulación de $T+30$ a $T+54$), AEMET-SREPS en círculos y ECENS en aspás. El BSS tiene orientación positiva, cuanto más alto mejor, cero indicando habilidad igual a la climatología muestral y uno indicando predicción perfecta. Las componentes de fiabilidad y resolución tienen orientación negativa, es decir, cuanto más abajo más fiabilidad y resolución. Como pautas generales comunes, los dos sistemas A-S y E-E presentan su mayor habilidad en invierno, la habilidad se degrada según aumenta el umbral y para

20 mm la variabilidad estacional no es significativa. La primera pauta es consistente con la predominancia de precipitación a gran escala, más bien estratiforme, en invierno. A-S en círculos supera en rendimiento a E-E en primavera y verano de forma clara para los umbrales 1, 5 y 10 mm, mientras que para 20 mm lo supera solo en invierno. Esta diferencia en habilidad se explica en términos de resolución: A-S a 0.25° puede resolver mejor los sistemas convectivos que E-E a 0.50° . Esta diferencia puede esperarse también en otoño, pero no se midió: efectivamente, la convección otoñal se sabía necesitaba todavía mejorar en A-S. Entrando en detalle en las componentes del BSS, A-S muestra más fiabilidad que E-E, especialmente en primavera y verano, diferencia decreciente con el umbral. Por el contrario, E-E muestra más resolución en el espacio de las probabilidades en primavera, otoño e invierno, excepto para 20 mm, lo cual puede deberse a su mayor tamaño con 51 miembros. La resolución no puede mejorarse mediante un proceso de calibración del espacio de probabilidades; por el contrario,

la fiabilidad puede mejorarse mediante calibración a costa de la resolución, aunque los beneficios de tal calibración son difíciles de conseguir en un entorno operativo [3]. La mayor diferencia en fiabilidad que en resolución en primavera y verano da al sistema A-S mayor BSS que E-E en estas estaciones.

La Figura 21.15 muestra series temporales de RSA con las mismas características que la Figura 21.14 en la página anterior en lo relativo a umbrales, periodo, etc. La medida RSA tiene orientación positiva, con un valor de cero indicando discriminación igual que la de la climatología muestral y un valor de uno correspondiendo a discriminación perfecta. Para 1, 5 y 10 mm ambos sistemas, A-S en círculos y E-E con aspas, muestran valores altos de RSA, ligeramente superiores en invierno, cuando esperamos que predomine la precipitación estratiforme, sin diferencias relevantes entre ambos sistemas. Para 20 mm hay diferencia más clara, con A-S mostrando mayor RSA en otoño e invierno, posiblemente debido a la mayor sensibilidad de A-S a eventos más raros, pues en este caso no está tan claro el papel de la resolución horizontal y su relación con la capacidad de describir correctamente la actividad convectiva.

En resumen, A-S muestra en general mejor rendimiento que E-E en términos de habilidad medida con BSS y sus componentes, aunque E-E muestra a veces más resolución, A-S es más fiable y con mayor diferencia en general. Esta diferencia es aún mayor en aquellas estaciones en las que la precipitación de tipo convectivo es más frecuente, mientras que cuando se espera predominio de precipitación de gran escala o estratiforme, no se observa apenas diferencia de habilidad entre ambos sistemas. Estos resultados son consistentes con la diferencia en resolución horizontal de ambos SPC con A-S 0.25° sobre 0.50° del E-E, a pesar de la ventaja en tamaño del E-E con 51 miembros sobre 20, ventaja que explica alguna diferencia puntual. No se encontraron grandes diferencias en discriminación medida con RSA, excepto para el umbral de 20 mm donde A-S muestra valores superiores, posiblemente debido a su mayor sensibilidad a eventos más raros. Si se tuviera en cuenta la diferencia por tamaño del SPC, podríamos esperar aún mayor diferencia de rendimiento a favor de A-S. Para explicar estos resultados, A-S mejora las predicciones de E-E debido probablemente a su mayor resolución horizontal (25-50 km) y quizá debido también a su naturaleza multimodelo, multianálisis y multicontorno.

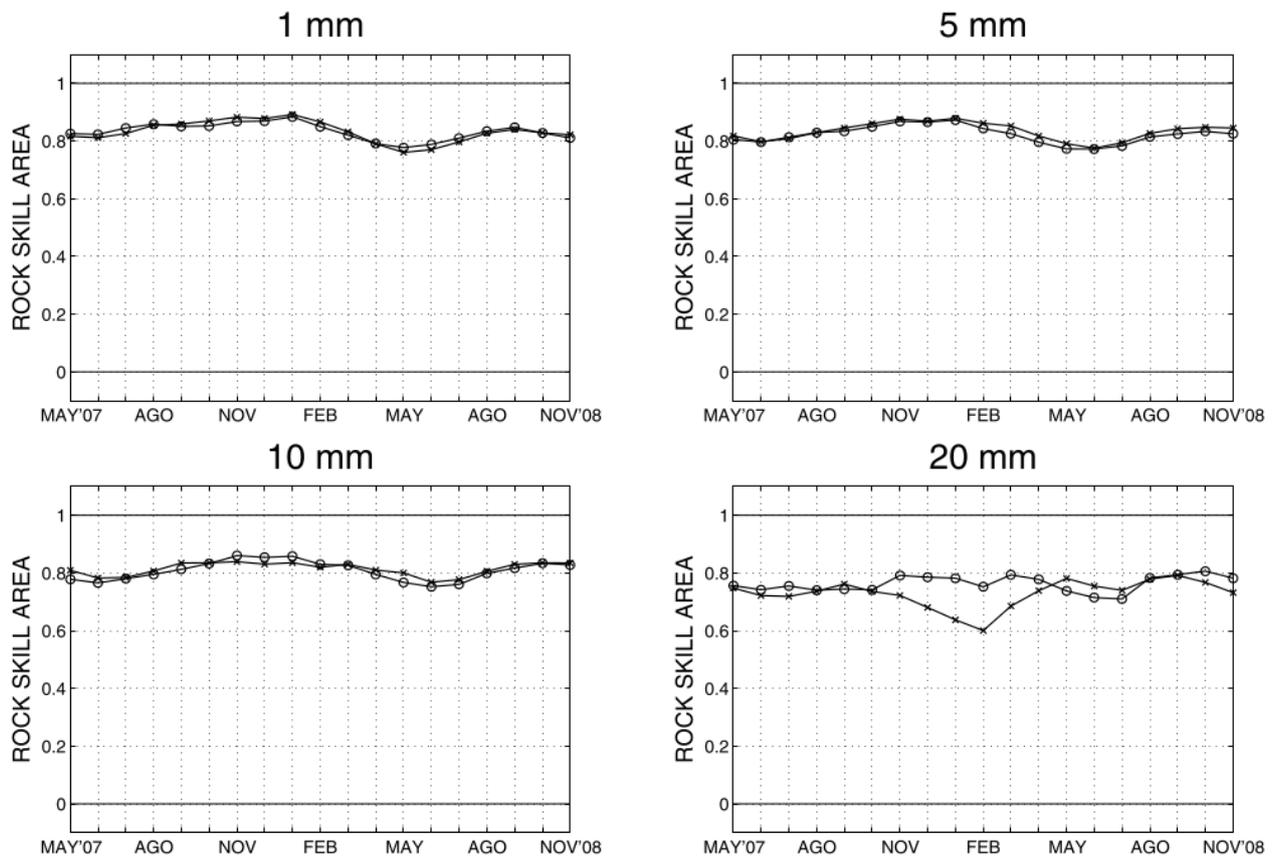


Figura 21.15: Como en la Figura 21.14 en la página anterior, pero esta vez el RSA. [21] TellusA, CC BY 4.0.

21.6 Usuarios

A partir del año 2007 se comenzaron a diseminar en tiempo real tanto las salidas directas de los miembros como los productos resultantes del A-S a diversos servicios meteorológicos, así como otras entidades europeas, que así lo demandaron para sus propios fines operativos o de investigación, a la par que colaboraban en explorar las proyecciones del sistema. Mencionamos aquí algunos ejemplos relevantes.



Figura 21.16: Usuarios destacados de AEMET-SREPS en Europa: consorcio COSMO.

Entre los consorcios europeos que más activamente solicitaron A-S para fines de investigación, podemos mencionar **Consortium for Small-scale Modeling (COSMO)**. Durante un tiempo se le suministraban cuatro miembros de A-S a la resolución de 0.25°. Después estuvieron investigando y operando el SPC de mesoescala COSMO-SREPS, en el ARPA-SIM de Bolonia, para lo que se les suministraban 16 miembros de A-E para producir perturbaciones en algunos parámetros usando A-S como condiciones de contorno. Por último, para anidamiento del SPC de mesoescala COSMO-DE-SREPS usando AEMET-SREPS como contorno, se les suministraban 16 miembros de A-S. Dentro de AEMET, en el **Área de Técnicas y Aplicaciones de Predicción (ATAP)** se abrió una línea de generación de productos probabilistas a partir de A-S. En la D. T. en Baleares se estuvieron usando las salidas de A-S con un algoritmo de detección de *medicane* (capítulo 33 en la página 551).



Figura 21.17: Usuarios destacados de AEMET-SREPS en Europa: Servizio Meteorologico Regionale del ARPA-SIM de Lombardia.

Por otro lado, en el campo de la predicción operativa, A-S fue popular especialmente en aquellos centros operativos donde la convección profunda del área mediterránea es un elemento clave. Además de los GPV de Barcelona y Málaga y, en menor medida Valencia y

Baleares, citamos por su relevancia y amabilidad especial al Servizio Meteorologico Regionale del ARPA-SIM de Lombardia, donde utilizaron durante años A-S con espectaculares resultados, siendo utilísimo para los casos de sistemas convectivos de mesoescala y precipitaciones torrenciales. Se recibían avisos de los lombardos los (pocos) días que recibían A-S con retraso o no lo recibían, debido a que no podían disponer de las predicciones correspondientes.

21.7 Proyectos

Algunos de los proyectos más relevantes en los que se ha involucrado el grupo de predecibilidad de AEMET se describen a continuación.

21.7.1 Proyectos internacionales de gran escala

2008-2012 TIGGE-LAM: THORPEX Interactive Grand Global Ensemble - Limited Area Models. <https://software.ecmwf.int/wiki/display/TIGL>

2006-2012 GLAMEPS: Grand Limited Area Model Ensemble Prediction System, Consorcio HIRLAM. <http://hirlam.org/index.php/hirlam-programme-53/general-model-description/glameps>

21.7.2 Energía eólica

2009-2010 SAFEWIND: Multi-scale data assimilation, advanced wind modelling & forecasting with emphasis to extreme weather situations for a safe large-scale wind power integration. Collaborative project funded by the European Commission under the 7th Framework Program, Theme 2007-2.3.2: Energy. <http://www.safewind.eu/>

2003-2005 HONEYMOON: A High resolution Numerical wind Energy Model for On- and Offshore using eNsemble predictions, DMI, INM, Cork Univ., Oldenburg Univ., EED, etc. http://cordis.europa.eu/project/rcn/67930_en.html

21.7.3 UIB (fenómenos adversos en el mediterráneo)

2018- COASTEPS: Fenómenos meteorológicos severos en zonas costeras mediterráneas: retos de predicibilidad y análisis climático (CGL2017-82868-R), CICYT, UIB. <http://meteo.uib.eu/coasteps> (en construcción en mayo 2018).

2011-2014 PREDIMED: Mejora de las predicciones de tiempo severo Mediterráneo por medio de observaciones adaptativas y métodos avanzados de predicción por conjuntos en el marco de los proyectos MEDEX Fase II y HyMeX (CGL2011-24458), CICYT, UIB. <http://meteo.uib.eu/predimed/PREDIMED.cid242657>

2008-2011 MEDICANES: Ambientes Meteorológicos, Predicibilidad Numérica y Evaluación del Riesgo en el Clima Presente y Futuro (CGL2008-01271), CICYT, UIB.

2005-2008 ENSEMBLE: Aplicación de Técnicas de Predicción por Conjuntos a episodios meteorológicos de gran impacto en el Mediterráneo occidental (CGL2005-05681/CLI), CICYT, UIB.

21.8 Resumen y conclusiones

En el periodo 2002-2014 se desarrolló en INM/AEMET un SPC para el corto plazo multimodelo, multianálisis y multicontorno denominado AEMET-SREPS (A-S). En su etapa de madurez el sistema funcionaba integrando cinco modelos deterministas de área limitada (**Limited Area Model (LAM)** o **Regional Circulation Model (RCM)**) en el superordenador de AEMET usando cinco modelos deterministas de globales o de circulación general (**Global Circulation Model (GCM)**) como condiciones iniciales y de contorno. El sistema se componía, haciendo el producto cartesiano correspondiente, de 25 miembros que se integraban dos veces al día, pasadas de 00 y 12 UTC, produciendo predicciones hasta un alcance predictivo de 72 horas sobre un dominio que cubría parte del océano Atlántico Norte, parte del norte de África y casi toda Europa. La resolución horizontal de los miembros era de 0.25° , aproximadamente 25 km en nuestras latitudes y la vertical de 40 niveles. En este capítulo se ha mostrado un ejercicio de verificación completo, tanto del flujo a gran escala como de las variables de tiempo sensible,

en un periodo de casi dos años, desde abril de 2007 hasta diciembre de 2008 (21 meses), primer periodo de madurez del sistema,

En la representación del flujo a gran escala, sinóptica principalmente, A-S era estadísticamente consistente con el análisis del modelo determinista del ECMWF, el ECHRES, con una ligera subdispersión típica en los SPC operativos más comunes. La dispersión del A-S muestra una relación razonablemente lineal con el error del promedio del SPC según crece el alcance de predicción, una evidencia clara de consistencia, relación mejor que la mostrada por cualquiera de los subSPC que se podían formar con los miembros del sistema, e. g. subSPC multimodelo o subSPC multicontorno. Estos últimos revelaron ser menos subdispersivos que los multimodelo, haciendo sospechar que el ingrediente multicontorno en el sistema tenía algo más de peso sobre la dispersión neta del mismo.

En la respuesta frente a eventos binarios para las variables de tiempo sensible, temperatura, viento y precipitación acumulada, que resultan las más importantes para los usuarios de un SPC para el corto plazo, el sistema presentaba una habilidad, destreza o pericia (skill) comparable o superior a la de SPC operativos destacables. Se muestran en este capítulo, por brevedad pero con cierto detalle, solamente los resultados para la precipitación acumulada en 24 horas. Se utilizaron observaciones de las redes pluviométricas europeas convenientemente malladas (gridding o upscaling) a 25 km, unas 1 000 observaciones diarias. Se compararon predicciones probabilistas del SPC del ECMWF, ECENS (E-E) y el A-S, seleccionando un alcance predictivo, relativamente justo para la comparativa, de T+54. Por su importancia en predicción operativa se seleccionaron los umbrales de precipitación 1, 5, 10 y 20 mm. Usando la medida Brier Skill Score (BSS [8, 32]) y sus componentes, A-S resulta tener más fiabilidad y peor resolución, con una habilidad neta superior en general a la de E-E, diferencia más clara en estaciones del año con actividad convectiva más alta. Usando la medida Roc Skill Area (RSA [16]) para medir la discriminación, solo se encontró diferencia por el umbral de 20 mm, en el que la discriminación de A-S fue superior en otoño-invierno. La más alta resolución horizontal de A-S, 0.25° sobre 0.50° juega un papel importante en estas diferencias de rendimiento, incluso a pesar de su tamaño más pequeño en número de miembros, 20 sobre 50, aspecto no estudiado en este ejercicio [6, 20]. Las situaciones convectivas se describen mejor con mayor resolución, conclusión esperada de antemano. De ahí que en las

situaciones donde predomina la precipitación estratiforme, de gran escala, apenas hay diferencias en el rendimiento de ambos sistemas. La diferencia de naturalezas en su diseño sin duda juega un papel, pero es más difícil de contrastar. A-S es multimodelo, multianálisis y multicontorno, mientras que E-E es un SPC monomodelo construido con perturbaciones en las condiciones iniciales y parametrizaciones estocásticas para tener en cuenta el error del modelo. En este sentido hace falta investigación más profunda [60], fuera del alcance del proyecto.

El reconocimiento es muy claro sobre la estrategia de diseño elegida: las configuraciones seleccionadas, multimodelo, multianálisis y multicontorno, quedan confirmadas como técnicas no solo plausibles sino también eficaces para muestrear convenientemente el error y las incertidumbres de condiciones iniciales, modelo y contorno.

Las predicciones probabilistas proporcionadas por A-S de variables de tiempo sensible pudieron ayudar en la predicción operativa como herramienta complementaria a los modelos deterministas de mayor resolución. A-S brinda información cuantitativa, explícita y detallada sobre la incertidumbre en la predicción y, por ende, sobre la predicibilidad atmosférica en cada área y alcance predictivo. Así mismo, proporciona diferentes escenarios, algunos de ellos de riesgo potencial, por lo que puede ser una poderosa herramienta de toma de decisiones a la hora de emitir avisos de fenómenos adversos, de gran relevancia en España, como son lluvias torrenciales o inundaciones súbitas en el área mediterránea, temporales de viento en el Cantábrico u olas de calor en el cada vez más intenso y extenso verano.

Se dieron varias líneas de investigación y desarrollo interesantes, tanto para mejorar este ensemble como para compararlo con otras alternativas de predicción probabilista, que han aportado conocimiento y experiencia para desarrollos actuales o futuros como el AEMET- γ SREPS (cap. 22 en la página 333). La aplicación de física estocástica en el SPC, utilizada en el ECMWF, puede aportar mejoras en el equilibrio energético en diferentes escalas. Las técnicas estadísticas avanzadas de calibración en el posproceso, aunque muestran resultados excepcionales corrigiendo la dispersión del ensemble en sentido estadístico, todavía fallan a la hora de afrontar fenómenos extremos. La

técnica *time lagged* para construir un superensemble agrupando varias pasadas también ha dado buenos resultados. Se ha probado también un ensemble híbrido, que rescata la parte de la información de la incertidumbre de un ensemble, para agregarla a un modelo determinista de más alta resolución, y por tanto, a priori más exacto y sobretodo detallado, obteniendo así un ensemble de más resolución, es decir, con más detalle. Las técnicas de generación de SPC difusivos, también probadas, pueden mejorar la predicción probabilista de la precipitación, una de las más importantes.

Un aspecto importante que aportó la experiencia con A-S tiene que ver con la oportunidad de establecer el sistema como operativo oficialmente. Con A-S surgieron problemas de licencia para el uso de los modelos, uso del que AEMET no tenía plena libertad para explotación y, en ocasiones, el uso se restringió al contexto de investigación pero no el operativo. Otro aspecto importante tiene que ver con la publicidad interna que necesita un sistema puntero tanto en su concepción teórica como en la práctica de su uso para predicción operativa. En la época de A-S un buen número de predictores eran reacios a la hora de aprender las técnicas elementales de predicción probabilista, por lo que el valor de una herramienta como A-S, que ya era un SPC añadido al ECENS, se ponía en tela de juicio. Añadiendo un largo etcétera para el que no hay espacio en este capítulo, el resultado, en el fondo nada insólito, es que A-S nunca fue un SPC operativo oficialmente y tampoco se utilizó, al menos oficialmente, en los grupos de predicción. Este aprendizaje es útil a la hora de afrontar el desarrollo del actual AEMET- γ SREPS (cap. 22 en la página 333), aunque muchos de los condicionantes de aquella época ya no son tales: los predictores, en la actualidad, aunque cada vez menos, están mejor preparados y con actitud positiva para el aprendizaje de nuevas herramientas. Queda la memoria de AEMET-SREPS un sistema de naturaleza multimodelo, técnica en la que España fue pionera europea.

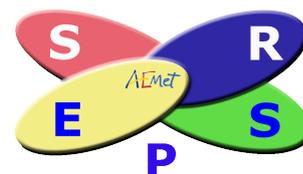


Figura 21.18: Logo de AEMET-SREPS, idea original de DANIEL SANTOS-MUÑOZ.

21.9 Referencias

- [1] ANDERSON, J.L. L. “A Method for Producing and Evaluating Probabilistic Forecast from Ensemble Model Integration”. En: *Journal of climate* 9.7 (1995), páginas 1518-1530. ISSN: 0894-8755. DOI: [10.1175 / 1520 - 0442\(1996 \) 009<1518 : AMFPAE>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2) (citado en página 323).
- [2] ARRIBAS, A, ROBERTSON, K B y MYLNE, K R. “Test of a poor man’s ensemble prediction system for short-range probability forecasting”. En: *Monthly Weather Review* 133.7 (2005), páginas 1825-1839 (citado en página 324).
- [3] ATGER, Frédéric. “Spatial and interannual variability of the reliability of ensemble-based probabilistic forecasts: Consequences for calibration.” En: *Monthly Weather Review* 131.8 (2003) (citado en página 328).
- [4] BRIER, Glenn W. “Verification of forecasts expressed in terms of probability”. En: *Monthly Weather Review* 78.1 (1950), páginas 1-3. DOI: [10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2) (citado en página 326).
- [5] 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 página 323).
- [6] BUIZZA, Roberto y col. “Impact of model resolution and ensemble size on the performance of an Ensemble Prediction System”. En: *Quarterly Journal of the Royal Meteorological Society* 124.550 (1998), páginas 1935-1960. ISSN: 1477-870X. DOI: [10.1002/qj.49712455008](https://doi.org/10.1002/qj.49712455008) (citado en páginas 323, 324, 326, 330).
- [7] CALLADO, A. y col. “Verification of an Hybrid Short - Range EPS”. En: *3rd International Verification Methods Workshop*. Reading: ECMWF, 2007 (citado en páginas 320, 321).
- [8] CANDILLE, Guillem y TALAGRAND, O. “Evaluation of probabilistic prediction systems for a scalar variable”. En: *Quarterly Journal of the Royal Meteorological Society* 131.609 (2005), páginas 2131-2150. ISSN: 00359009. DOI: [10.1256/qj.04.71](https://doi.org/10.1256/qj.04.71) (citado en páginas 323, 326, 330).
- [9] CANDILLE, Guillem y TALAGRAND, O. “Impact of observational error on the validation of ensemble prediction systems”. En: *Quarterly Journal of the Royal Meteorological Society* 134.633 B (abr. de 2008), páginas 959-971. ISSN: 00359009. DOI: [10.1002/qj.268](https://doi.org/10.1002/qj.268) (citado en páginas 321, 322).
- [10] CANDILLE, Guillem y col. “Verification of an Ensemble Prediction System against Observations”. En: *Monthly Weather Review* 135.7 (2007), páginas 2688-2699. ISSN: 0027-0644. DOI: [10.1175/MWR3414.1](https://doi.org/10.1175/MWR3414.1) (citado en página 326).
- [11] CASATI, B. y col. “Forecast verification: current status and future directions”. En: *Meteorological Applications* 15.1 (mar. de 2008), páginas 3-18. ISSN: 13504827. DOI: [10.1002/met.52](https://doi.org/10.1002/met.52) (citado en página 326).
- [12] CHERUBINI, Tiziana, GHELLI, Anna y LAURETTE, François. “Verification of Precipitation Forecasts over the Alpine Region Using a High-Density Observing Network”. En: *Weather and Forecasting* 17.2 (abr. de 2002), páginas 238-249. ISSN: 0882-8156. DOI: [10.1175 / 1520 - 0434\(2002\) 017<0238 : VOPFOT>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0238:VOPFOT>2.0.CO;2) (citado en página 326).
- [13] CÔTÉ, Jean y col. “The operational CMC-MRB global environmental multiscale (GEM) model. Part II: results”. En: *Monthly Weather Review* 126.6 (1998), páginas 1397-1418 (citado en página 317).
- [14] CULLEN, M J P. “The unified forecast/climate model”. En: *Meteorological Magazine* 122.1449 (1993), páginas 81-94 (citado en páginas 316, 317).
- [15] DOMS, G y SCHÄTTLER, U. “The nonhydrostatic limited-area model LM (Lokal-

- Modell) of DWD”. En: *Part I: Scientific documentation, Deutscher Wetterdienst (DWD)* 10 (1999) (citado en página 316).
- [16] DU, Jun. “Hybrid ensemble prediction system: a new ensembling approach”. En: *Preprints, Symposium on the 50th Anniversary of Operational Numerical Weather Prediction*. 2004, páginas 14-17 (citado en páginas 320, 321, 330).
- [17] DUDHIA, Jimmy. “A nonhydrostatic version of the Penn State–NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front”. En: *Monthly Weather Review* 121.5 (1993), páginas 1493-1513 (citado en página 316).
- [18] 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 319).
- [19] FERRO, Christopher A. T. “Comparing Probabilistic Forecasting Systems with the Brier Score”. En: *Weather and Forecasting* 22.5 (2007), páginas 1076-1088. ISSN: 0882-8156. DOI: [10.1175/WAF1034.1](https://doi.org/10.1175/WAF1034.1) (citado en páginas 324, 326).
- [20] FERRO, Christopher A. T., RICHARDSON, David S. y WEIGEL, Andreas P. “On the effect of ensemble size on the discrete and continuous ranked probability scores”. En: *Meteorological Applications* 15.1 (mar. de 2008), páginas 19-24. ISSN: 13504827. DOI: [10.1002/met.45](https://doi.org/10.1002/met.45) (citado en páginas 324, 326, 330).
- [21] 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 315, 324, 327, 328).
- [22] GILLELAND, Eric. “Confidence intervals for forecast verification”. En: *NCAR Technical Note NCAR/TN-479+STR* (2010). DOI: [10.5065/D6WD3XJM](https://doi.org/10.5065/D6WD3XJM) (citado en página 326).
- [23] GNEITING, Tilmann y col. “Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation”. En: *Monthly Weather Review* 133.5 (2005), páginas 1098-1118. ISSN: 0027-0644. DOI: [10.1175/MWR2904.1](https://doi.org/10.1175/MWR2904.1) (citado en página 319).
- [24] GRELL, Georg A, DUDHIA, Jimmy, STAUFFER, David R y col. “A description of the fifth-generation Penn State/NCAR mesoscale model (MM5)”. En: (1994) (citado en página 316).
- [25] HAMILL, Thomas M. “Reliability Diagrams for Multicategory Probabilistic Forecasts”. En: *Weather and Forecasting* 12.4 (1997), páginas 736-741. ISSN: 0882-8156. DOI: [10.1175/1520-0434\(1997\)012<0736:RDFMPF>2.0.CO;2](https://doi.org/10.1175/1520-0434(1997)012<0736:RDFMPF>2.0.CO;2) (citado en página 323).
- [26] HAMILL, Thomas M. “Interpretation of Rank Histograms for Verifying Ensemble Forecasts”. En: *Monthly Weather Review* 129.3 (2001), páginas 550-560. ISSN: 0027-0644. DOI: [10.1175/1520-0493\(2001\)129<0550:IORHFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2) (citado en página 323).
- [27] HAMILL, Thomas M. y COLUCCI, Stephen J. “Verification of Eta-RSM short-range ensemble forecasts”. En: *Monthly Weather Review* 125.6 (1997), páginas 1312-1327 (citado en página 323).
- [28] HERBACH, Hans. “Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems”. En: *Weather and Forecasting* 15.5 (oct. de 2000), páginas 559-570. ISSN: 0882-8156. DOI: [10.1175/1520-0434\(2000\)015<0559:DOTCRP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2) (citado en página 319).
- [29] HOU, Dingchen, KALNAY, Eugenia y DROEGEMEIER, Kelvin K. “Objective verification of the SAMEX’98 ensemble forecasts”. En:

- Monthly Weather Review* 129.1 (2001), páginas 73-91 (citado en páginas 315, 324).
- [30] JAKOB, C. y col. "The IFS cycle CY21r4 made operational in October 1999". En: *Spring* 87.87 (2000), páginas 2-9 (citado en página 317).
- [31] JOLLIFFE, Ian T. "Uncertainty and Inference for Verification Measures". En: *Weather and Forecasting* 22 (2007), páginas 637-650. ISSN: 0882-8156. DOI: [10.1175/WAF989.1](https://doi.org/10.1175/WAF989.1) (citado en página 326).
- [32] JOLLIFFE, Ian T. y STEPHENSON, David B. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. 2003, página 254. ISBN: 0470864419. DOI: [10.1016/j.ijforecast.2005.11.002](https://doi.org/10.1016/j.ijforecast.2005.11.002) (citado en páginas 326, 330).
- [33] JOLLIFFE, Ian T. y STEPHENSON, David B. "Introduction". En: *Forecast Verification*. John Wiley & Sons, Ltd, 2011, páginas 1-9. ISBN: 9781119960003. DOI: [10.1002/9781119960003.ch1](https://doi.org/10.1002/9781119960003.ch1) (citado en página 326).
- [34] KHARIN, Viatcheslav V. y ZWIERS, Francis W. "On the ROC score of probability forecasts". En: *Journal of Climate* 16.24 (2003), páginas 4145-4150. ISSN: 08948755. DOI: [10.1175/1520-0442\(2003\)016<4145:OTRSOP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2) (citado en página 326).
- [35] LEITH, C E. "Theoretical skill of Monte Carlo forecasts". En: *Monthly Weather Review* 102.6 (1974), páginas 409-418 (citado en página 323).
- [36] LU, Chungu y col. "Short-Range Numerical Weather Prediction Using Time-Lagged Ensembles". En: *Weather and Forecasting* 22.3 (jun. de 2007), páginas 580-595. ISSN: 0882-8156. DOI: [10.1175/WAF999.1](https://doi.org/10.1175/WAF999.1) (citado en página 321).
- [37] MAJEWSKI, D y SCHRODIN, P. "Short description of the Europa-Modell (EM) and Deutschland-Modell (DM) of the Deutscher Wetterdienst (DWD)". En: *Quarterly Bulletin (April)* (1994), páginas 1-31 (citado en página 316).
- [38] MAJEWSKI, Detlev. "HRM-user's guide". En: *Deutsche Wetter Dienst (DWD). Offenbach, Germany* 124 (2009) (citado en página 316).
- [39] MAJEWSKI, Detlev y col. "The operational global icosahedral-hexagonal gridpoint model GME: description and high-resolution tests". En: *Monthly Weather Review* 130.2 (2002), páginas 319-338 (citado en página 317).
- [40] MASON, Simon J. "On Using "Climatology" as a Reference Strategy in the Brier and Ranked Probability Skill Scores". En: *Monthly Weather Review* 132.7 (2004), páginas 1891-1895. ISSN: 0027-0644. DOI: [10.1175/1520-0493\(2004\)132<1891:OUCAAR>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1891:OUCAAR>2.0.CO;2) (citado en página 326).
- [41] MASON, Simon J. y WEIGEL, Andreas P. "A Generic Forecast Verification Framework for Administrative Purposes". En: *Monthly Weather Review* 137.1 (2009), páginas 331-349. ISSN: 0027-0644. DOI: [10.1175/2008MWR2553.1](https://doi.org/10.1175/2008MWR2553.1) (citado en página 326).
- [42] McDONALD, A y HAUGEN, Janerik. "A two-time-level, three-dimensional semi-Lagrangian, semi-implicit, limited-area grid-point model of the primitive equations". En: *Monthly weather review* 120.11 (1992), páginas 2603-2621 (citado en página 316).
- [43] McLACHLAN, Geoffrey J y KRISHNAN, Thiriyambakam. *The EM algorithm and extensions*. Volumen 382. John Wiley & Sons, 2007, página 359. ISBN: 0471123587. DOI: [10.2307/1271189](https://doi.org/10.2307/1271189) (citado en página 319).
- [44] MURPHY, James M. "The impact of ensemble forecasts on predictability". En: *Quarterly Journal of the Royal Meteorological Society* 114.480 (1988), páginas 463-493. DOI: [10.1002/qj.49711448010](https://doi.org/10.1002/qj.49711448010) (citado en página 323).

- [45] PALMER, Tim N. y col. “Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER)”. En: *Bulletin of the American Meteorological Society* 85.6 (2004), páginas 853-872 (citado en página 324).
- [46] RAFTERY, Adrian E. y col. “Using Bayesian model averaging to calibrate forecast ensembles”. En: *Monthly Weather Review* 133.5 (2005), páginas 1155-1174. ISSN: 0027-0644. DOI: [10.1175/MWR2906.1](https://doi.org/10.1175/MWR2906.1) (citado en página 318).
- [47] SANTOS, Carlos y GHELLI, Anna. “Observational probability method to assess ensemble precipitation forecasts”. En: *Quarterly Journal of the Royal Meteorological Society* 138.662 (ene. de 2012), páginas 209-221. ISSN: 00359009. DOI: [10.1002/qj.895](https://doi.org/10.1002/qj.895) (citado en páginas 321, 322).
- [48] SANTOS-MUÑOZ, D y col. “Using Bayesian Model Averaging to calibrate short-range forecasts from a multi-model ensemble”. En: *3rd International Verification Methods Workshop*. Reading: ECMWF, 2007 (citado en página 319).
- [49] SELA, Joseph G. “Spectral modeling at the national meteorological center”. En: *Monthly Weather Review* 108.9 (1980), páginas 1279-1292 (citado en página 317).
- [50] SELA, Joseph G. “The NMC spectral model”. En: *NOAA Technical Report, NWS-30* 108 (1982) (citado en página 317).
- [51] SIMARRO, J y col. “Probabilistic forecasts based on the HIRLAM INM deterministic model”. En: *3rd International Verification Methods Workshop*. Reading: ECMWF, 2007 (citado en página 320).
- [52] SIMMONS, A. J. y col. “The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution”. En: *Meteorology and Atmospheric Physics* 40.1-3 (1989), páginas 28-60. ISSN: 01777971. DOI: [10.1007/BF01027467](https://doi.org/10.1007/BF01027467) (citado en página 317).
- [53] SUGI, Masato. “Description and performance of the JMA operational global spectral model (JMA-GSM88)”. En: *Geophys. Magazine* 43 (1990), páginas 105-130 (citado en página 317).
- [54] SWETS, John A y col. “Measuring the accuracy of diagnostic systems”. En: *Science* 240.4857 (1988), páginas 1285-1293 (citado en página 326).
- [55] THEIS, S E, HENSE, A y DAMRATH, U. “Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach”. En: *Meteorological Applications* 12.3 (2005), páginas 257-268 (citado en página 320).
- [56] UNDÉN, Per y col. “HIRLAM-5 scientific documentation”. En: (2002) (citado en página 316).
- [57] WHITAKER, Jeffrey S. y LOUGHE, Andrew F. “The Relationship between Ensemble Spread and Ensemble Mean Skill”. En: *Monthly Weather Review* 126.12 (1998), páginas 3292-3302. ISSN: 0027-0644. DOI: [10.1175/1520-0493\(1998\)126<3292:TRBESA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2) (citado en página 323).
- [58] 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 páginas 319, 325, 326).
- [59] WILSON, Laurence J. “Comments on “Probabilistic predictions of precipitation using the ECMWF ensemble prediction system””. En: *Weather and Forecasting* 15.3 (2000), páginas 361-364 (citado en página 326).
- [60] ZIEHMANN, Christine. “Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models”. En: *Tellus, Series A: Dynamic Meteorology and Oceanography* 52.3 (mayo de 2000), páginas 280-299. ISSN: 02806495. DOI: [10.3402/tellusa.v52i3.12266](https://doi.org/10.3402/tellusa.v52i3.12266) (citado en páginas 323, 326, 331).

