

# Nuevos métodos de verificación de las predicciones numéricas de precipitación: el método SAL aplicado en la mesoescala

M. A. Amo-Baladrón<sup>(1)</sup> y C. Santos<sup>(1)</sup>

(1) Agencia Estatal de Meteorología (AEMET), C/Leonardo Prieto Castro, 8, 28040, Madrid, mamob@aemet.es, csantosb@aemet.es

## Resumen

Los modelos de predicción numérica del tiempo que se encuentran actualmente en la transición a la mesoescala, lo cual requiere de nuevos métodos de verificación que solventen los problemas con los que se encuentran los métodos tradicionales al enfrentarse con las nuevas escalas espaciales, en particular al verificar fenómenos episódicos y altamente localizados como la precipitación (e.g.: el problema de la doble penalización).

El nuevo método de verificación SAL orientado a características permite obtener información cuantitativa y detallada sobre la calidad de los modelos desde una perspectiva más cercana a la verificación subjetiva, aportando información diagnóstica que ayuda a los desarrolladores a mejorar los modelos y a los predictores a entender en qué aspectos fallan.

En este artículo, exploramos las posibilidades del método SAL aplicándolo a modelos de diferentes resoluciones espaciales (HIRLAM, ECMWF, y HARMONIE) en periodos seleccionados para la precipitación acumulada en 24h, verificando con observaciones procedentes de la red pluviométrica de alta resolución de AEMET. Más en concreto, se estudia la variabilidad estacional de los modelos y la influencia que el tamaño del dominio a verificar puede tener en los valores de los parámetros SAL.

## 1. Introducción.

Los modelos de predicción numérica del tiempo mejoran rápidamente debido al rápido progreso en capacidad de computación, datos de observación (especialmente satélite), asimilación de datos, dinámica y parametrizaciones de procesos físicos sub-rejilla. Uno de los desafíos actuales es la transición a la mesoescala, en el que aparecen dificultades nuevas, no sólo para el desarrollo de los modelos, sino también para los métodos de verificación. Recientemente están surgiendo nuevos métodos que aportan una perspectiva diferente a la de los tradicionales, más cercana a la verificación subjetiva (Casati et al., 2008). Entre ellos destaca SAL (Wernli et al., 2008), un método orientado a características u objetos, que permite obtener información cuantitativa y detallada sobre la calidad de los modelos sin penalizarlos por su escala espacial.

Los problemas de la verificación tradicional (Jolliffe and Stephenson, 2011; Wilks, 2011), orientada a las medidas estadísticas (*scores*), están relacionados en parte con las escalas espaciales de modelos y observaciones: representatividad, impacto de los métodos de interpolación, pérdida relativa de significación estadística o independencia debido a correlaciones espaciales o temporales, e interpretación de distintas medidas de acuerdo con la escala espacial, e.g., el problema de la doble penalización.

El problema de la **doble penalización** (Davis et al. 2006; Ebert and Gallus, 2009) surge cuando una

estructura de precipitación observada se encuentra ligeramente desplazada en la predicción. Una medida de error tradicional, como puede ser el error cuadrático medio (*MSE*), penaliza al modelo por no predecir la estructura en los puntos de rejilla con precipitación observada, así como por predecirla en los puntos de rejilla cercanos donde la precipitación no ha sido observada. Sin embargo, un modelo de menor resolución y utilidad que pierda completamente el evento tan sólo es penalizado una vez, obteniendo mejor puntuación según el *score MSE*.

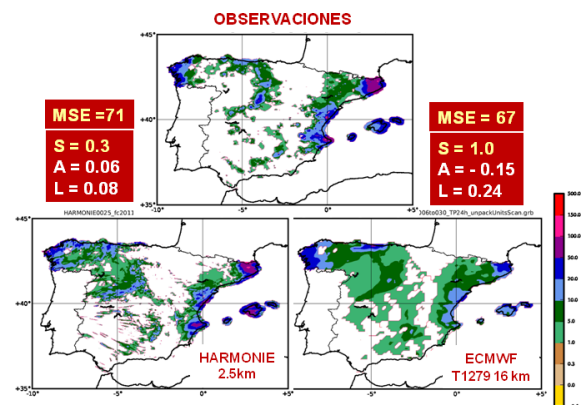


Fig. 1.- Problema de la doble penalización. Comparación de la predicción de precipitación acumulada del modelo de alta resolución HARMONIE (2.5 km; panel inf-izq) con la predicción del modelo de baja resolución ECMWF T1279 (16 km; panel inf-der), según el *score* tradicional *MSE* y según el método de verificación orientado a características SAL. El panel superior muestra el campo observado.

La Fig. 1 ilustra claramente dicho problema. En ella se compara la predicción de precipitación acumulada en 24 h del modelo de alta resolución HARMONIE (2.5 km; panel inf-izq) con la predicción del modelo de baja resolución ECMWF T1279 (16 km; panel inf-der), según el *score* tradicional *MSE* y según el nuevo método de verificación orientado a características SAL. El valor de *MSE* es ligeramente superior en el modelo de alta resolución (71 vs. 67), esto es, puntúa peor que el modelo de baja resolución (un *MSE* = 0 equivale a una predicción perfecta), mientras que en el caso del método de verificación SAL sucede lo contrario (*S=A=L=0* equivale a una predicción perfecta), siendo para el modelo de alta resolución *S,A,L* = (0.3, 0.06, 0.08); y para el de baja resolución *S,A,L* = (1.0, -0.15, 0.24). Comparando visualmente con el campo observado, se aprecia cómo el modelo de alta resolución resuelve con mayor precisión la localización de las estructuras de precipitación así como sus máximos.

Por lo tanto, cada vez son más necesarios nuevos métodos de verificación que traten con patrones y estructuras espaciales para contrastar la calidad de los modelos según aumenta su resolución. Dichos métodos deben ser más cercanos a la verificación subjetiva en el sentido de que proporcionen información diagnóstica, que ayude a los desarrolladores a la mejora de los modelos y a los predictores a entender qué aspectos de éstos fallan. En este escenario surge recientemente el Proyecto de Intercomparación de Métodos Espaciales de Verificación (*Spatial Verification Methods Inter Comparison Project, IPC*; Ahijevich et al. 2009; Gilleland et al., 2009). Dicho proyecto estudia y compara los diferentes métodos de verificación espacial aparecidos recientemente clasificándolos en cuatro grupos diferentes dependiendo de la aproximación utilizada: (1) métodos vecinales (*neighborhood*; Ebert, 2009), (2) métodos de separación de escalas (*scale separation*; Casati, 2010), (3) métodos orientados a objetos o características (*object or feature oriented*; Ebert and Gallus, 2009), y (4) métodos de deformación de campo (*field deformation*).

El método que aquí se aplica, SAL, perteneciente a la familia de los métodos orientados a objetos, permite obtener información cuantitativa y detallada sobre la calidad de los modelos desde una perspectiva completamente diferente a la tradicional. El artículo se estructura de la siguiente forma: en la Sección 2 se explica brevemente el método SAL. Las observaciones utilizadas en la verificación y su postproceso son descritas en la Sección 3. En la Sección 4 se ilustra, mediante varios ejemplos, el significado físico de los parámetros SAL, se describen los experimentos realizados, y se

comentan los resultados obtenidos. Finalmente, las conclusiones son expuestas en la Sección 5.

## 2. El método SAL.

En el contexto de la predicción cuantitativa de la precipitación (*Quantitative Precipitation Forecast*; QPF) el método *Structure-Amplitude-Location* (SAL; Wernli et al., 2008) es una medida orientada a características que describe la calidad de las predicciones de precipitación tomando como referencia observaciones en rejilla (lo cual hace necesario aplicar previamente el proceso de *upsampling*; Cherubini et al., 2002, Ahrens & Beck, 2008), que brinda información cuantitativa, detallada y explícita en términos de tres componentes: estructura, amplitud y localización. Esta información se aproxima más en su concepción a una verificación llamada diagnóstica (más parecida a la subjetiva), que permite contrastar la calidad del modelo desde una perspectiva diferente, así como ser usada como ayuda en la predicción operativa.

Básicamente, el objetivo del método consiste en cuantificar la calidad de la predicción a través de tres parámetros (estructura, amplitud y localización) en una región de interés o dominio. Para los parámetros de estructura y localización es necesario identificar previamente áreas de precipitación contiguas (objetos) en el campo modelo y en el campo observado por separado. A diferencia de otros métodos basados en identificación de objetos (e.g.: MODE; Davis et al. 2006), SAL no busca correspondencias del campo modelo en el campo observado.

Los pasos que se deben dar para calcular los parámetros SAL son (para un mayor detalle ver Wernli et al., 2008):

- 1) Especificar la región de interés o **dominio**. En nuestro caso la península Ibérica y Baleares, y el cuadrante noreste como ejemplo de un dominio más pequeño.
- 2) Seleccionar un **valor umbral de precipitación** ( $R^*$ ) a partir del cual definir los objetos. En nuestro caso tomamos el percentil 95 ( $R^{95}$ ) de todos los puntos de rejilla con valores de precipitación significativos ( $> 0.1$  mm) multiplicado por un factor  $f$ . La elección de este factor es subjetiva y se basa en un proceso de *tuning* ( $f=1/15$ ; Wernli et al., 2008).

$$R^* = f \cdot R^{95} \quad (1)$$

- 3) **Identificación de los objetos** en el campo modelo y observado por separado.

- 4) Cálculo de los **atributos de los objetos** individualmente. Dichos atributos son: precipitación total del objeto, su centro de masas (CM) y volumen escalado (la precipitación total del objeto dividida entre el valor máximo de precipitación dentro del objeto).
- 5) Cálculo de las **propiedades de los campos** por separado: promedio de la precipitación en el campo, su CM, y la distancia máxima entre dos puntos de rejilla del campo. Otras propiedades del campo se derivan realizando promedios de los valores de los atributos de los objetos pesados por la cantidad de precipitación total de dichos objetos: el promedio de los volúmenes escalados, y el promedio de las distancias entre los CMs de los objetos y el CM del campo.
- 6) Cálculo de los **parámetros SAL**: dichos parámetros se calculan a partir de las diferencias normalizadas de las propiedades de los campos:
  - **Estructura**: diferencia normalizada de los promedios de los volúmenes escalados. Informa sobre el **tamaño y/o forma** de las estructuras de precipitación del modelo con respecto a las observaciones. Puede tomar valores  $[-2, +2]$ , siendo  $S=0$  una predicción perfecta, mientras que  $S > 0$  ( $< 0$ ) indica que el modelo predice objetos más extensos y/o planos (más pequeños y/o picudos) que los observados.
  - **Amplitud**: diferencia normalizada de los valores promedio de precipitación de los campos, por lo tanto, informa sobre la **exactitud** alcanzada por el modelo en la **cantidad total de precipitación**. Puede tomar valores  $[-2, +2]$ , siendo  $A=0$  una predicción perfecta, mientras que  $A=1$  ( $=-1$ ) indica que el modelo sobrestima (subestima) en un factor 3 la cantidad total de precipitación observada.
  - **Localización**: se divide en dos partes,  $L1$  y  $L2$ .  $L1$  mide la distancia normalizada entre los CMs de los campos modelo y observado respecto a la distancia máxima entre dos puntos de rejilla del dominio. Puede tomar valores  $[0, +1]$ , siendo  $L1=0$  una predicción donde el CM del campo modelo coincide con el del campo observado.  $L1$  indica, en primera aproximación, la **localización de la distribución de la precipitación** del modelo con respecto a las observaciones. La segunda parte del parámetro

localización,  $L2$ , es la diferencia normalizada de los promedios de las distancias entre los CMs de los objetos y el CM del campo. Informa sobre la **posición relativa de los objetos dentro del campo**, pero es insensible a la rotación. Se encuentra escalado de tal manera que toma valores  $[0, +1]$ , por lo tanto,  $L=L1+L2$  puede tomar valores  $[0, +2]$ .

La información obtenida tras aplicar el método SAL a un grupo de pares de campos modelo-observación, puede resumirse en un diagrama SAL, donde en el eje de abscisas (ordenadas) se representa el parámetro de estructura (amplitud), y en escala de colores el valor del parámetro de localización (e.g.: Fig. 2). Cada punto de este diagrama representa una verificación de un campo modelo con un campo observado. Se muestra también en este diagrama la mediana de los parámetros (líneas discontinuas), así como sus percentiles 25 y 75 (cuadrado sombreado y líneas negras sobre la escala de colores de la localización). Dichas medianas serán utilizadas para cuantificar la calidad de los modelos en diferentes periodos.

### 3. Observaciones.

Para la verificación de los modelos se ha seleccionado la red pluviométrica de alta resolución de AEMET. La precipitación es medida diariamente por voluntarios repartidos por todo el territorio español, realizándose a las 07 UTC. Por lo tanto, se obtienen precipitaciones acumuladas en 24h de 07 UTC a 07 UTC. Dichos voluntarios envían los datos mensualmente a AEMET, donde pasan un control de calidad y son almacenados en una base de datos. En total, la red pluviométrica de AEMET dispone de alrededor de 3000 estaciones. Para este ejercicio se disponen de observaciones del periodo 2007 a 2011.

El método SAL de verificación necesita el campo de observación y el campo modelo en la misma rejilla regular. Por lo tanto, es necesario realizar el proceso de *upscaling* (Cherubini et al., 2002, Ahrens & Beck, 2008) del campo observado a la rejilla del modelo para obtener el campo de precipitación estimado (*Quantitative Precipitation Estimate*; QPE). Se ha aplicado el programa de upscalado proporcionado por C. Santos e I. Guerrero. Las observaciones han sido upscaladas a la rejilla natural de cada modelo. Para dicho proceso se ha establecido un valor umbral de estaciones a promediar de 2 en un radio de búsqueda de estaciones en torno al punto de rejilla  $d = 0.20^\circ$  para todos los modelos, salvo en el caso del Modelo de Circulación General (GCM) ECMWF T799, donde se ha seleccionado un radio de búsqueda de  $0.25^\circ$ . Esto es, los puntos de rejilla asociados a una o

ninguna estación no son considerados y se marcan como *missing values*. El radio de búsqueda se debe elegir de tal manera que se alcance la máxima cobertura posible, teniendo en cuenta la resolución de la rejilla a la cual se quieren upscalar las observaciones. Se ha utilizado la función de estructura inversa cuadrada (*Inverse Distance Weighting*; IDW; Sodoudi et al., 2010) con parámetro  $\alpha = 2$  para promediar la distribución de estaciones en torno a cada punto de rejilla, y así asignar un valor a dichos puntos:

$$QPE_{IDW} = \frac{\sum_d r^{-\alpha} \cdot R_i}{\sum_d r^{-\alpha}} \quad (2)$$

#### 4. Aplicación del método SAL.

##### 4.1. Modelos.

La verificación SAL se ha realizado sobre los modelos recogidos en la tabla 1: modelos del Centro Europeo (ECMWF; Simmons et al., 1989; Jakob et al., 1999), y modelos HIRLAM y HARMONIE (Navascués et al., 2012; <http://www.hirlam.org/>).

Modelo	Resolución	Disponibilidad
ECMWF T799	0.225° (25 km)	1 Ene. 2007 Ene. 2010
ECMWF T1279	0.141° (16 km)	17 Mar. 2011 31 Dic. 2011
HIRLAM ONR	0.16°	17 Abr. 2007 31 Dic. 2011
HIRLAM HNR	0.05°	17 Abr. 2007 31 Dic. 2011
HARMONIE	2.5 km	1 Oct. 2011 31 Dic. 2011

Tabla 1.- Modelos utilizados en la verificación, con su correspondiente resolución y las fechas con disponibilidad de datos.

El Centro Europeo incrementó la resolución horizontal de su modelo a finales de Enero de 2010, pasando de T799 a T1279, por lo que se ha realizado la verificación en dos periodos diferentes (en los cuales la resolución horizontal y vertical de los modelos a considerar se mantuvo sin cambios):

- **Periodo 2007-2010:** se comparan los modelos ECMWF T799, ONR, y HNR en el dominio de la península Ibérica y Baleares. Se realiza una comparación más detallada en el cuadrante noreste (el cual contiene la cuenca del Ebro) para los modelos ECMWF T799 y HNR. En este periodo, el número de días disponibles permite realizar también un estudio estratificado por estaciones.

- **Periodo 2011:** se realiza un primer estudio de los meses Oct-Dic comparando los modelos HIRLAM, ECMWF T1279, y HARMONIE. Se utiliza el diagrama SAL obtenido en la verificación del modelo HARMONIE para ilustrar el significado físico de los parámetros SAL.

##### 4.2. Significado físico de los parámetros SAL.

La Fig. 2 recoge cuatro ejemplos que muestran el significado físico de los parámetros SAL sobre el diagrama correspondiente a la verificación del modelo HARMONIE. En ella se presentan los mapas de los objetos identificados en el campo precipitación y observación para tres días extremos y uno en el que el modelo realizó una buena predicción según SAL.

Los mapas de la cornisa cantábrica muestran un ejemplo de predicción en la que no se localizaron correctamente los objetos de precipitación ( $L = 0.7$ ). Mientras que el modelo predecía precipitación a lo largo de la cornisa cantábrica, las observaciones mostraron precipitación concentrada en el interior de Galicia. Mas aún, la cantidad total de precipitación se sobrestimó en gran medida ( $A = 1.7$ ), pudiendo considerarse como una falsa alarma.

Por el contrario, el ejemplo del archipiélago balear muestra muy buena localización de los objetos ( $L=0.04$ ), pero subestima la precipitación total ( $A = -1.4$ ).

Un claro ejemplo del significado físico del parámetro estructura se puede ver en el mapa de la península situado en la esquina sup-der. El modelo predice estructuras de precipitación extensas, mientras que las observaciones muestran áreas de precipitación más concentradas ( $S = 1.3$ ).

Por último, los mapas de la esquina inf-der muestran un ejemplo de buena predicción, con valores de  $A$  y  $L$  muy próximos a cero y  $S$  ligeramente negativa.

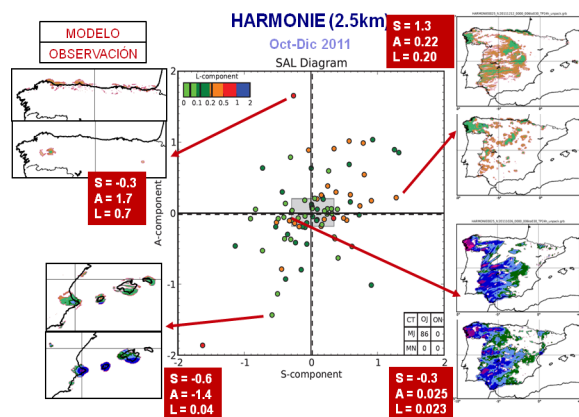


Fig. 2.- Significado físico de los parámetros SAL.

##### 4.3. Periodo 2007-2010.

En este periodo se ha estratificado por estaciones, esto es, por intervalos temporales de 3 meses comenzando por Marzo-Mayo (MAM) y terminando por Diciembre-Febrero (DEF). En la tabla 2 se recogen los resultados obtenidos y en la Fig. 3 se muestra, a modo de ejemplo, la verificación de HNR. El parámetro de estructura no presenta una variación significativa con el periodo estacional ( $S = 0.4-0.6$ ), sin embargo, como se verá en la Sección 4.5, dicha variabilidad podría quedar enmascarada por el tamaño del dominio. El valor del parámetro amplitud sí parece estar relacionado con la estación, siendo máximo en los meses de verano ( $A = 0.7$ ), esto es, los modelos tienden a sobrestimar la cantidad total de precipitación respecto a las observaciones. Sin embargo, los modelos predicen con bastante exactitud dicha magnitud durante los meses de invierno ( $A = 0.03$ ). Por otra parte, el parámetro de localización permanece prácticamente constante con las estaciones ( $L = 0.3$ ).

2007-2010		MAM	JJA	SON	DEF
ECMWF T799	S	0.3	0.6	0.4	0.4
	A	0.20	0.4	0.15	0.01
	L	0.25	0.3	0.3	0.3
HIRLAM ONR	S	0.5	0.5	0.5	0.7
	A	0.3	0.5	0.3	0.14
	L	0.14	0.17	0.19	0.17
HIRLAM HNR	S	0.5	0.5	0.4	0.6
	A	0.3	0.7	0.23	0.03
	L	0.25	0.3	0.3	0.25

Tabla 2.- Resultados de la verificación SAL estratificada por estaciones en el periodo 2007-2010.

Fig. 3.- Diagramas SAL para la verificación del modelo HIRLAM HNR en el periodo 2007-2010 estratificado por estaciones: MAM (Marzo a Mayo), JJA (Junio a Agosto), SON (Septiembre a Noviembre), y DEF (Diciembre a Febrero). Dominio: península Ibérica y Baleares

En el trimestre DEF destaca la casi nula presencia de puntos en el cuadrante IV del diagrama SAL (Fig. 3, inf-der), lo que muestra la dificultad que presentan los modelos para sobrestimar la precipitación ( $A > 0$ ) con objetos o estructuras pequeñas y/o picudas ( $S < 0$ ; Wernli et al., 2008). En casi todos los trimestres se aprecia una acumulación de puntos en el extremo superior-derecho del diagrama ( $A > 1$  y  $S > 1$ ), siendo más apreciable en el trimestre JJA. Dichos días pueden considerarse como falsas alarmas (Wernli et al., 2008), ya que los modelos sobrestiman la precipitación total en más de un factor 3 con estructuras más extensas y/o planas que las observadas.

#### 4.4. Periodo 2011.

La Fig. 4 muestra los resultados de la verificación de los modelos HIRLAM, ECMWF T1279, y HARMONIE correspondientes a los meses de Octubre a Diciembre de 2011. El parámetro de estructura presenta una clara mejora con el aumento de la resolución de los modelos. Sin embargo, el parámetro amplitud no muestra relación con la resolución, siendo los modelos de peor y mejor resolución, ECMWF T1279 y HARMONIE, respectivamente, los que menor valor de amplitud presentan. El parámetro localización presenta un valor similar para todos los modelos. El modelo HARMONIE destaca por ser el que mejor puntúa según el método SAL.

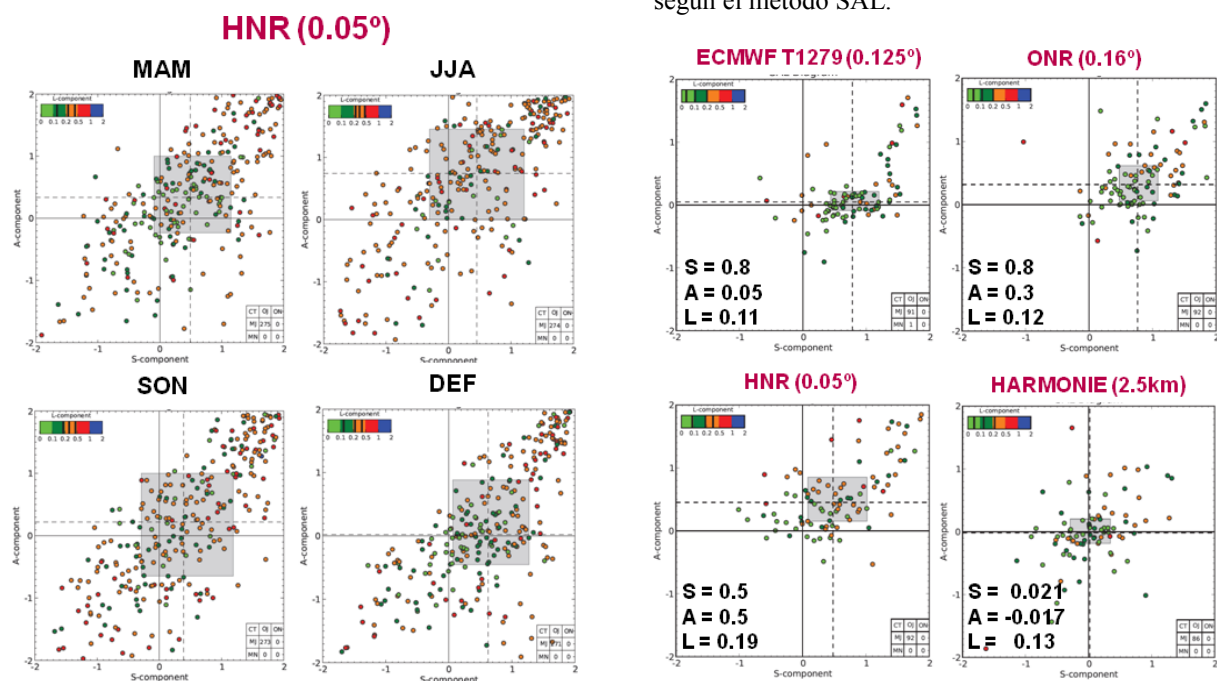


Fig. 4.- Diagramas SAL para la verificación de los modelos ECMWF T1279, ONR, HNR, y HARMONIE



2.5 km, de arriba hacia abajo e izq. a der., correspondientes al periodo Oct-Dic 2010.

#### 4.5. Influencia del tamaño del dominio.

Para comprobar si el tamaño del dominio puede influenciar las dependencias encontradas en los parámetros SAL, se repite el experimento en el periodo 2007-2010 para un modelo típico de circulación general GCM ECMWF T799 (~25km) y un modelo típico de mesoscala HIRLAM HNR (~5km) en el cuadrante noreste de la península Ibérica, que contiene la cuenca del Ebro (Fig. 5). Wernli et al. (2009) sugieren tamaños de dominio de 500 x 500 km<sup>2</sup> como máximo, ya que dominios mayores dificultan la interpretación de los parámetros SAL.

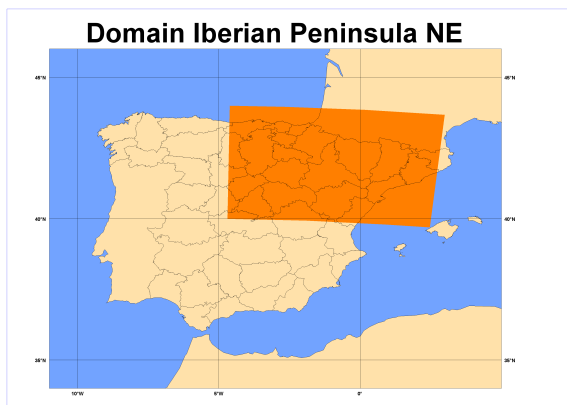


Fig. 5.- Dominio escogido: cuadrante noreste de la península Ibérica, aproximadamente la cuenca del Ebro.

Se construyen diagramas SAL estratificados por estaciones y sin estratificar, para mostrar una comparación más detallada entre ambos tipos de modelos.

La Fig. 6 muestra los diagramas SAL de ambos modelos para el periodo completo **sin estratificar** por estaciones. A simple vista se puede apreciar que la nube de puntos para el modelo HNR se encuentra más centrada que para el modelo T799, indicando un mejor comportamiento del modelo de mesoscala. Ambos modelos presentan desplazamientos o *bias* positivos en la distribución de *A* y *S* (ver los recuadros sombreados que muestran los *IQRs* de ambas distribuciones), indicando que ambos modelos sobrestiman la cantidad total de precipitación y predicen estructuras más extensas y/o planas que las observadas. Sin embargo, HNR muestra una distribución de *S* más centrada, esto es, describe mejor el tamaño y/o forma de las estructuras, como cabría esperar de un modelo de mesoscala. Sin embargo, el modelo T799 presenta un *bias* en *A* menor que HNR.

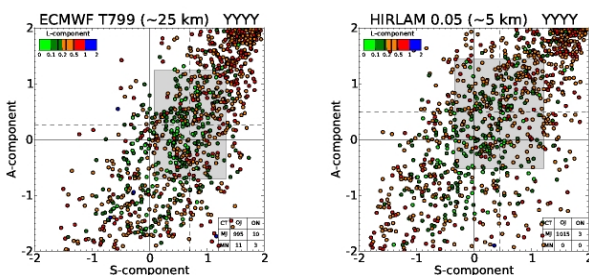


Fig. 6.- Diagramas SAL para la verificación de los modelos ECMWF T799 (izq.) y HIRLAM HNR (der.), correspondientes al periodo 2007 - 2010. Dominio: cuadrante NE de la península Ibérica.

Las **tablas de contingencia** (esquina inf-der en los diagramas SAL) muestran que HNR obtiene 1015 aciertos (objetos de precipitación identificados tanto en el campo observado como en el modelo) y 3 falsas alarmas (sin objetos en el campo observado, pero con al menos un objeto identificado en el campo modelo), sin errores (al menos un objeto identificado en el campo observado, pero sin objetos en el campo modelo) o aciertos negativos (ningún objeto identificado en ambos campos); mientras que T799 obtiene 995 aciertos y 10 falsas alarmas, con 10 errores y 3 negativos correctos. Es decir, HNR muestra mejor respuesta en la captura del carácter lluvioso o no lluvioso de los días.

Se ha realizado una estratificación estacional de la muestra siguiendo el mismo procedimiento que en la Sección 4.3. La tabla 3 muestra los resultados para los modelos T799 y HNR.

2007-2010		MAM	JJA	SON	DEF
ECMWF T799	S	0.5	0.9	0.8	0.6
	A	0.23	0.5	0.21	-0.04
	L	0.3	0.4	0.3	0.3
HIRLAM HNR	S	0.3	0.3	0.5	0.7
	A	0.5	0.9	0.4	0.13
	L	0.3	0.4	0.4	0.25

Tabla 3.- Resultados de la verificación SAL estratificada por estaciones en el periodo 2007-2010 para el dominio cuadrante NE de la península Ibérica.

A continuación comparamos la verificación del modelo HNR en el mismo periodo para el dominio de la península Ibérica y Baleares, y el dominio del cuadrante noreste, esto es, la cuenca del Ebro. La Fig. 7 muestra los diagramas SAL **estratificados** del modelo HNR para el cuadrante noreste.

En el caso del dominio del cuadrante noreste, la componente de estructura presenta sus valores más bajos ( $S = 0.3$ ; mejor puntuación) en los meses de primavera-verano, mientras que en otoño-invierno se alcanzan los valores más elevados ( $S = 0.5-0.7$ ; peor predicción). Esto es similar a lo que sucedía al verificar el dominio de la península Ibérica y Baleares, sin embargo, en este último caso el contraste entre los meses de primavera-verano y otoño-invierno era menor, presentando un valor más constante con las estaciones ( $S = 0.5-0.6$ ). Es decir,

al utilizar un dominio de tamaño mayor que el recomendado, la variabilidad estacional en la componente de estructura se diluye.

En cuanto a la componente de amplitud para el dominio del cuadrante noreste, la sobrestimación en la cantidad total de precipitación es máxima en verano y mínima en invierno. Resultado similar al obtenido al verificar el dominio de la península Ibérica y Baleares. La componente de localización no presenta ninguna tendencia estacional apreciable, mostrando un valor prácticamente constante. El mismo comportamiento se observó para el dominio de la península Ibérica y Baleares.

Por lo tanto, el **tamaño del dominio** no presenta una fuerte influencia en los parámetros de amplitud y localización, pero sí en el de **estructura**, pudiendo **diluir la tendencia estacional** en el caso de dominios con tamaños mayores de los recomendados.

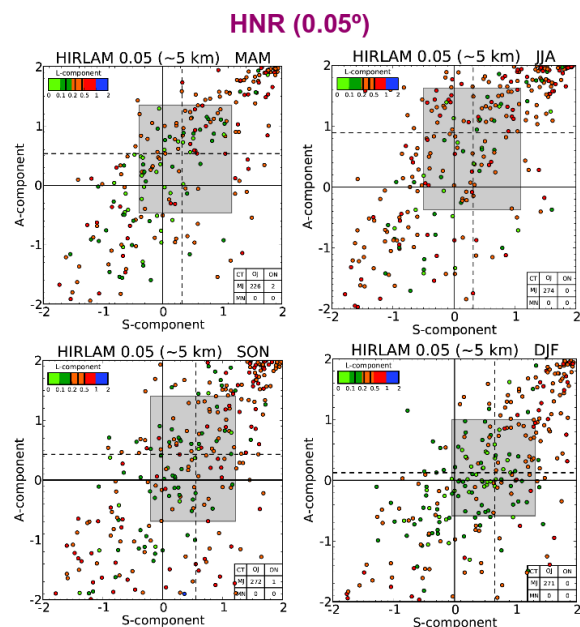


Fig. 7.- Diagramas SAL para la verificación del modelo HIRLAM HNR estratificado por estaciones en el periodo 2007 - 2010. MAM, JJA, SON, y DJF de arriba hacia abajo e izq. a der. Dominio: cuadrante NE de la península Ibérica.

Volviendo a la verificación del modelo HNR en el dominio del cuadrante noreste, la tendencia estacional encontrada en la componente de estructura (mejor predicción del tamaño y/o forma de los objetos en los meses de primavera-verano, y peor predicción durante los meses de otoño-invierno, con objetos más extensos y/o planos que los observados) podría deberse a que la convección de verano se encuentra relacionada con variabilidad de pequeña escala, mientras que la convección de otoño es más dependiente del forzamiento sinóptico, favorable a la predicción de estructuras de mayor escala.

La variabilidad encontrada en la distribución de los parámetros de amplitud y estructura (elevados IQRs) del modelo HNR podría estar relacionado con el hecho de que se trate de un modelo de mesosescala que aún conserva su carácter hidrostático. HNR es un modelo con elevadas fluctuaciones entre buen y mal comportamiento, según la experiencia de muchos predictores en AEMET. Esta variabilidad de comportamiento alcanza su mínimo en invierno (mínimo tamaño de la región sombreada en el diagrama SAL), consistente con los patrones de alta presión y mayor predictibilidad del estado de la atmósfera.

Si comparamos los modelos HNR y T799 estratificados por estaciones (ver tabla 3), vemos como tanto la componente de amplitud como de localización presentan las mismas tendencias, sin embargo, la componente de estructura muestra un comportamiento diferente. Mientras que en el modelo HNR presenta los mejores valores de  $S$  para los meses de primavera-verano ( $S = 0.3$ ), en el caso del modelo T799 las mejores puntuaciones se alcanzan en los meses de primavera-invierno ( $S = 0.5-0.6$ ). Esta diferencia de comportamiento podría estar relacionado con la escala óptima a la que cada modelo puede representar las estructuras. T799 es un modelo de circulación general en teoría óptimo para representar precipitación a gran escala, que predominaría en los meses de primavera-invierno; mientras que HNR es un modelo de mesosescala que podría capturar mejor los procesos de precipitación convectiva relacionados con la orografía.

Por último, señalar que otro tipo de estratificaciones, con mayor significado físico, podrían aportar más información sobre el comportamiento y la calidad de los modelos, e.g., estratificación dependiente del flujo, esto es, estratificar la muestra utilizando técnicas de *clustering* en cuatro regímenes meteorológicos bien conocidos: *NAO+*, *NAO-*, *Blocking*, y *Atlantic Ridge* (Ferranti et al. 1994; Ferranti and Corti, 2010 ).

## 5. Conclusiones.

En este artículo se han explorado las posibilidades del método SAL aplicándolo a modelos de diferentes resoluciones espaciales (HIRLAM, ECMWF, y HARMONIE) en periodos seleccionados para la precipitación acumulada en 24h, verificando con observaciones procedentes de la red pluviométrica de alta resolución de AEMET. Se han obtenido las siguientes conclusiones:

- Se verifican los modelos ECMWF T799, HIRLAM ONR, y HIRLAM HNR estratificando por estaciones para el periodo

2007-2010. En los diagramas SAL observamos una clara tendencia en el parámetro amplitud, con valores máximos en los meses de verano (sobrestimación de la precipitación por parte de los modelos), y mínimos en los de invierno. La componente de estructura no muestra una tendencia significativa debido a que queda enmascarada por el tamaño de dominio escogido. La localización permanece prácticamente constante en todas las estaciones.

- Se muestra la verificación preliminar de los modelos HIRLAM, ECMWF T1279, y HARMONIE para el periodo Octubre-Diciembre de 2011. Se aprecia una clara mejora en el parámetro de estructura (mejor captura por parte del modelo de la forma y/o tamaño de los objetos de precipitación) a medida que se incrementa la resolución de los modelos. El incremento de la resolución no muestra relación con la amplitud, mientras que los valores de la localización se mantienen prácticamente constantes. El modelo HARMONIE (de mayor resolución) destaca por ser el que mejor puntuación recibe en la verificación.
- Se repite la verificación de los modelos ECMWF T799 (modelo de circulación general) y HIRLAM HNR (modelo de mesoscala) para un dominio de menor tamaño, el cuadrante noreste de la península Ibérica (el cual contiene la cuenca del Ebro). Se realiza una comparación detallada de ambos tipos de modelos, y se comprueba como el tamaño del dominio no afecta al comportamiento de los parámetros de amplitud y localización. Sin embargo, la variabilidad estacional, ahora sí observada en el parámetro de estructura, puede quedar enmascarada en el caso de dominios con tamaños mayores de los recomendados. Por lo tanto, se concluye que un tamaño de dominio apropiado es de suma importancia para una correcta interpretación de la verificación realizada con el método SAL.

## 6. Agradecimientos.

Los autores agradecen a Marcus Paulat (DWD), Matthias Zimmer (Univ. Mainz) y Heini Wernli (Univ. Zurich) que amablemente proporcionaran el código fuente original (FORTRAN 77; f77) del programa SAL. Imanol Guerrero y Carlos Santos realizaron posteriormente la traducción del código de f77 a f95, así como diversas mejoras (e.g.:

algoritmo de *clustering*). Los autores también agradecen a Anna Ghelli (ECMWF) las fructíferas discusiones mantenidas que han contribuido de modo esencial al desarrollo de este artículo. Por último, un agradecimiento especial a la labor del personal de AEMET encargado de las plataformas de computación y de la base de datos climatológica, los cuales han proporcionado los medios y datos necesarios para poder realizar este trabajo.

## 7. Referencias.

- Ahrens, B. and A. Beck, 2008. *Meteorol Atmos Phys* **99**, 155
- Ahijevich, D., Gilleland, E., Brown, B. G., and Ebert, E. E., 2009. *Weather and Forecasting*, **24**, 1485
- Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocerich, M., Damrath, U., Ebert, E. E., Browne, B. G., and Mason, S., 2008. *Meteorol. Appl.* **15**, 3
- Casati, B., 2010. *Weather and Forecasting*, **25**, 113
- Cherubini, T., Ghelli, A., and Lalaurette, F., 2002. *Weather and Forecasting*, **17**, 238
- Davis, C., Brown, B., and Bullock, R., 2006. *Mon. Wea. Rev.*, **134**, 1772
- Ebert, E. E. and Gallus, W. A., 2009. *Weather and Forecasting*, **24**, 1401
- Ebert, E. E., 2009. *Weather and Forecasting*, **24**, 1498
- Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., and Ebert, E. E., 2009. *Weather and Forecasting*, **24**, 1416,
- Jakob, C., Andersson, E., Beljaars, A., Buizza, R., Fisher, M., Grard, E., Ghelli, A., Janssen, P., Kelly G., McNally, A. P., Miller, M., Simmons, A., Teixeira, J., and Viterbo, P., 1999: The IFS cycle CY21r4 made operational in October 1999. *ECMWF Newsletter*, **87**, 29
- Ferranti, L., Molteni, F., Branković, Č., Palmer, T. N., 1994. *J. Climate*, **7**, 849–868.
- Ferranti, L. and Corti, S., 2010. Ensemble prediction skill in relation with large scale circulation patterns. 10th EMS Annual Meeting, 10th European Conference on Applications of Meteorology (ECAM) Abstracts, held Sept. 13-17, 2010 in Zurich, Switzerland.



Jolliffe, I. T. and Stephenson, D. B., 2011. Introduction, in Forecast Verification: A Practitioner's Guide in Atmospheric Science, Second Edition (eds I. T. Jolliffe and D. B. Stephenson), John Wiley & Sons, Ltd, Chichester, UK.

Marzban, C. and S. Sandgathe, 2010. Weather and Forecasting, **25**, 1479

Navascués, B., Calvo, J., Morales, G., Santos, C. et al., 2012. History and perspectives of Numerical Weather Prediction at AEMET (accepted). Atmospheric Research.

Simmons, A. J., Burridge, D. M., Jarraud, M., Girard, C., and Wergen, W., 1989: The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution. Meteorology and Atmospheric Physics, **40**, 28

Sodoudi, S., Noorian, A., Geb, M., and Reimer, E. 2010. Theor. Appl. Climatol., **99**, 39

Wernli, H., Paulat, M., Hagen, M., Frei, C., 2008. Mon. Wea. Rev., **136**, 4470-4487

Wernli, H., Hofmann, C., and Zimmer, M., 2009. Weather and Forecasting, **24**, 1472-1484

Wilks, D. S., 2011. Statistical Methods in the Atmospheric Sciences, Vol. **100**, Third Edition (International Geophysics). Academic Press.