

EVALUACIÓN, FRENTE A HIRLAM-02, DE LAS PREDICCIONES DE OCURRENCIA DE LLUVIA DE UN MODELO ESTADÍSTICO DE PRECIPITACIÓN DIARIA

J. Abaurrea ⁽¹⁾, E. Alvarez ⁽²⁾, J. Asín ⁽¹⁾

⁽¹⁾ Dpto. Métodos Estadísticos. Univ. de Zaragoza

⁽²⁾ Instituto Nacional de Meteorología, CMT en Aragón, La Rioja y Navarra

Correo electrónico: abaurrea@unizar.es

Resumen

En este trabajo se compara la predicción de ocurrencia de lluvia apreciable (mayor o igual que 0.1 litros/m²) obtenida con un modelo de regresión logística que usa la información registrada en un observatorio, con las previsiones del modelo Hirlam-02, ambas con un horizonte de predicción de 6 horas.

Se ha construido un modelo útil para cualquier día del año y modelos específicos para las tres épocas (fría, templada y cálida) en las que se ha dividido éste. Los modelos específicos son más simples y fáciles de interpretar pero no producen mejoras notables en la predicción.

La comparación se realiza con los datos y en la celda en la que se ubica el observatorio del aeropuerto de Zaragoza. La predicción que produce directamente HIRLAM-02 no corresponde con el régimen observado, ni en los valores medios ni en el ciclo anual, y la previsión de ocurrencia es peor que la del modelo estadístico. Se analiza también la predicción que produce HIRLAM-02 utilizando umbrales distintos para cada una de las tres épocas consideradas, lo que le permite alcanzar porcentajes de días correctamente predichos superiores a los del modelo estadístico.

Palabras clave

Precipitación diaria local, modelo estadístico de previsión, HIRLAM-02.

Abstract

In this work, the prediction of rainfall occurrence (rainfall higher than 0.1 mm/m²) obtained from a logistic regression model based on data measured in an observatory is compared with the forecast obtained from the Hirlam-02 model, both of them using a 6 hours forecasting horizon.

We have built a global model, useful for any day, and specific models for three periods defined in the year: cold, hot and mild. The specific models are simpler and easier to understand but they do not produce a noticeable improvement of forecasting results.

The comparison is performed using data from the Zaragoza airport observatory and from the HIRLAM cell where it is located. The HIRLAM-02 forecast is not able to reproduce the observed average or the annual cycle and the forecasting results are worse than those obtained from the statistical model. When a different threshold is applied in each of the periods considered, the percentage of success of HIRLAM-02 forecast is highly improved and overcomes those of the statistical model.

Key words

Daily local rainfall, statistical forecasting model, HIRLAM-02.

1. INTRODUCCIÓN

Los métodos estadísticos son herramientas útiles para establecer la previsión de la ocurrencia de lluvia. En particular, el INM ha desarrollado un método de análogos para estimar la probabilidad de que la precipitación supere ciertos umbrales, Del Hoyo y otros (2002); Ribalaygua y otros (1996) trabajaron en una clasificación de situaciones-tipo útil en la predicción de la precipitación y Armengot y otros (2001) desarrollaron, para la Comunidad Valenciana, un modelo de previsión de ocurrencia de lluvias intensas.

En la siguiente sección se presentan los datos disponibles para este trabajo y el modelo ajustado, con una interpretación de sus componentes. La sección 3 se dedica a la discusión de la capacidad predictiva del modelo estadístico. En la sección 4 se analiza la previsión de precipitación obtenida con HIRLAM-02 y en la sección 5 se comparan estos resultados con la predicción estadística; además, se estudia la definición de umbrales que ayuden a mejorar la predicción con HIRLAM. Finalmente la sección 6 presenta las conclusiones del trabajo.

2. MODELOS ESTADÍSTICOS AJUSTADOS

El modelo utilizado es uno de regresión logística en el que la respuesta, $Y(t)$, es una variable indicadora de que la cantidad registrada en ese día alcanza o supera el umbral de 0.1 litros/m². Las covariables disponibles son las medidas en superficie y las procedentes de dos radiosondeos diarios en el aeropuerto de Zaragoza, ver Tabla I. Se utilizan sólo las medidas hechas hasta las 18 horas del día t-1.

Tabla I: Variables disponibles

Superficie: Medidas en 4 horas-tipo	
Presión atmosférica	$P00, P07, P13, P18$
Temperatura	$T00, T07, T13, T18$
Humedad específica	$H00, H07, H13, H18$
Geopotenciales 850, 700, 500 y 300hPa: A las 0 y 12 horas	
Altura geopotencial	$A850^{12}, A700^{12}, A500^{12}, A300^{12}, A850^{00}, A700^{00}, A500^{00}, A300^{00}$
Temperatura	$T850^{12}, T700^{12}, T500^{12}, T300^{12}, T850^{00}, T700^{00}, T500^{00}, T300^{00}$
Punto de rocío	$R850^{12}, R700^{12}, R500^{12}, R300^{12}, R850^{00}, R700^{00}, R500^{00}, R300^{00}$
Diferencia entre geopotenciales alternos	$A85/50^{12}, A70/30^{12}, A85/50^{00}, A70/30^{00}, T85/50^{12}, T70/30^{12}, T85/50^{00}, T70/30^{00}, R85/50^{12}, R70/30^{12}, R85/50^{00}, R70/30^{00}$
Viento en superficie y altura a las 0 y 12 horas	
Dirección del viento en superficie, expresada mediante un factor	<i>Calma</i> <i>Este: entre 45° y 184°,</i> <i>Suroeste: entre 185° y 244°,</i> <i>Noroeste: entre 245° y 344°,</i> <i>Norte: entre 345° y 44°</i>
Dirección en altura, expresada mediante su seno y coseno	$dV850^{12}.cos, dV850^{12}.sen, dV850^{00}.cos, dV850^{00}.sen,$ $dV700^{12}.cos, dV700^{12}.sen$
Velocidad del viento	$vVSUP^{12}, vVSUP^{00}, vV850^{12}, vV850^{00}, vV700^{12}, vV700^{00},$ $vV500^{12}, vV500^{00}, vV300^{12}, vV300^{00}$

Se ha construido un modelo, denominado anual, que puede aplicarse a cualquier día del año; además, se han construido modelos específicos para tres épocas diferenciadas, definidas según la tipología y el predominio de las situaciones atmosféricas generadoras de lluvia, ver Serrano y otros (1999):

- La época Fría que incluye los meses de Noviembre a Marzo.
- La época Cálida que comprende del 16 de Junio al 15 de Septiembre.
- El resto del año constituye la época llamada Templada.

Para la estimación de estos modelos se utilizaron los datos del periodo 1990-1997.

La Tabla II muestra las covariables que tienen un efecto significativo en alguno de los cuatro modelos; para los efectos presentes en dos o más modelos se indica, con fines comparativos, el valor del coeficiente estimado. Para la dirección del viento en 500 hPa se señala, en cada modelo, la dirección correspondiente al mayor efecto positivo en el predictor lineal.

Tabla II: Términos significativos en los modelos estadísticos de previsión; si el coeficiente va acompañado de un paréntesis, indica que el término es significativo sólo en presencia de esa dirección de viento en superficie

Términos	Anual	Fría	Templada	Cálida
Armónicos	2	1	0	0
P18 _{t-1}	0.42x-2.2 10 ⁻⁵ x ² <i>Función decr.</i>	-0.0076x	1.71x-8.8 10 ⁻⁵ x ² <i>Función decr.</i>	-0.011x
ΔP.18.13 _{t-1}	-0.032 <i>Cuadr.(NW)</i>	-0.038 <i>Cuadr.(NW)</i>	-0.043	
P00 _{t-1} - P00 _{t-2}	<i>Crec.(Calma)</i>			
ΔP.7.0 _{t-1}	0.021 (E)		0.036 (E)	
H18 _{t-1}	0.015	0.015	0.017	
H07 _{t-1}	0.030 (E)		0.050 (E)	
T18 _{t-1}	-0.0094 (E)		-0.0083 -0.0030 (SW)	-0.0110 (E)
VVSUP ¹² _{t-1}	-0.037			-0.152
vV500 ¹² _{t-1}	-0.036 (N) 0.014 (NW)			-0.035 (E)
dV500 ¹² _{t-1} (dirección efecto máximo)	140° 270°(Calma) 210° (E) 170° (SW)	170° (SW)	245° (Calma) 205° (E) 250° (NW)	220° (E)
R700 ¹² _{t-1}	-0.0028	-0.0028	-0.0028	-0.0047
A85/50 ¹² _{t-1}	<i>Crec.</i>			
T850 ¹² _{t-1}		<i>Decr.</i>		
T85/50 ¹² _{t-1}		<i>Crec. (E)</i>		
2*T850-R850-2*T500≥520				<i>Crec. (E)</i>

Estos efectos admiten una interpretación física. Así, la influencia de la presión a las 18 horas (decreciente) reflejaría la llegada y estancamiento de una masa fría postfrontal.

El punto de rocío en el geopotencial 700 hPa a las 12h es un efecto común a todos los modelos; en todos ellos tiene un carácter decreciente y un valor comparable; está relacionado con la ocurrencia de episodios lluviosos de un solo día de duración, seguidos de estancamiento seco y habitualmente frío.

En el modelo específico para los días cálidos no interviene la última medida de humedad en superficie, H18_{t-1}, variable que favorece la ocurrencia de tormenta en el día t-1 pero cuyo efecto no alcanza al día t, dado el proceso de renovación atmosférica producido por el movimiento del aire en superficie.

El espesor de la capa 850-500 hPa, tan importante desde el punto de vista dinámico, sólo tiene un efecto significativo en el modelo anual.

La temperatura en superficie a las 18 horas es estadísticamente significativa en los modelos para las épocas templada y cálida, en las que la precipitación está más influida por el calentamiento solar de la capa más superficial; por el contrario en los meses fríos es significativo el efecto de la temperatura en 850 hPa a las 12 horas.

Se aprecia que las covariables asociadas al viento en los distintos niveles tienen una especial relevancia. Existe interacción entre la dirección del viento en superficie y en 500 hPa (ver Figura 1):

- En la situación de calma en superficie los dos picos podrían corresponder a pasos frontales atlánticos diferenciados según la distinta latitud de su trayectoria.
- Las situaciones que corresponden a un flujo en altura de procedencia Sur y Sureste tienen asociado un efecto positivo: la presencia de viento de SW en superficie (ligado generalmente a una situación de baja térmica peninsular que se ha trasladado y ocluido en el valle del Ebro) y de viento SE en altura provoca el mayor incremento en el predictor lineal.
- El efecto positivo asociado al viento Este en superficie combinado con una dirección SW en altura (200°) es menor, debido quizás a su escasa ocurrencia.

En los modelos aparecen términos de interacción entre las covariables y el factor que expresa la dirección de viento en superficie. En la Figura 2 se muestra el efecto de la variación de presión entre dos medidas consecutivas, dependiendo de la dirección del viento en superficie.

Figura 1: Efecto de la dirección del viento en 500 hPa a las 12 hora, modelo anual
Viento en superficie: CALMA

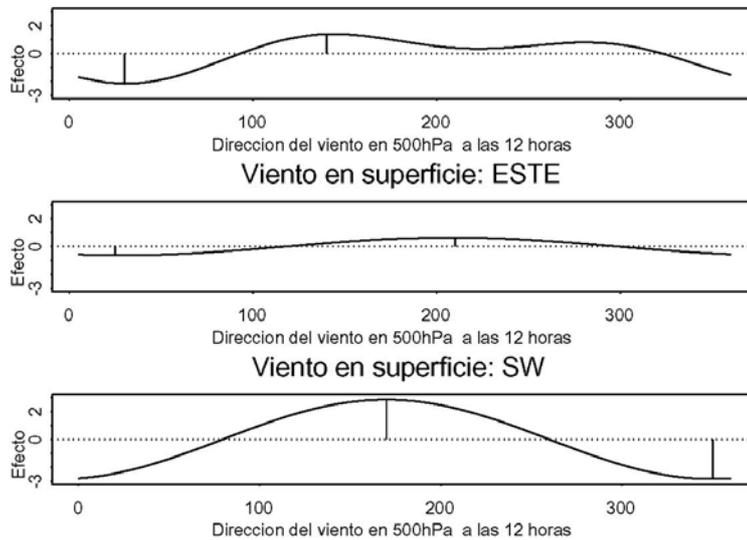
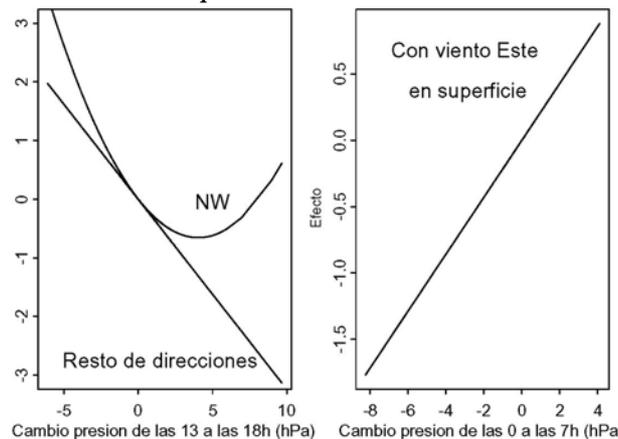


Figura 2: Efecto de la variación de presión entre dos medidas consecutivas, modelo anual



3. CAPACIDAD PREDICTIVA DEL MODELO ESTADÍSTICO

Se analiza la capacidad predictiva del modelo estadístico comparando su previsión de ocurrencia con los datos observados. Un día se predice húmedo cuando su probabilidad de

ocurrencia ajustada es superior a un umbral prefijado; en otro caso se predice seco. Se ha utilizado como umbral el valor de la frecuencia relativa de lluvia en cada día del año, suavizado mediante una regresión que utiliza armónicos hasta el orden 4. Se ha utilizado también el umbral 0.5 concluyendo, tras una comparación, que es menos adecuado.

La capacidad predictiva del modelo se evalúa sobre los datos de un periodo reservado para su validación (1998-2000), lo que se corresponde con el posible uso operativo del modelo. Para medir dicha capacidad se calcula el porcentaje de días bien predichos tanto globalmente como por estaciones y en las diferentes situaciones definidas por $Y(t)$ e $Y(t-1)$. En la Tabla III se aprecia que la situación peor prevista es la de un inicio de racha seca y que los resultados obtenidos son homogéneos en las distintas épocas del año.

Los resultados de los modelos específicos no son claramente superiores a los obtenidos con el modelo global: En los meses fríos el modelo anual predice igual o incluso mejor en todas las situaciones. En los meses templados, el modelo particular es superior al anual, salvo en los días secos precedidos de un día seco, siendo el porcentaje medio de éxito un 10% mayor. En la época cálida, el modelo específico mejora al anual, particularmente en la predicción de los días iniciales de racha seca.

Tabla III: Porcentajes de días bien previstos con el modelo anual y con los modelos estacionales. Se indica entre paréntesis el porcentaje obtenido con el modelo anual en cada una de las épocas

Modelo	Global	$Y_t=0$ $Y_{t-1}=0$	$Y_t=0$ $Y_{t-1}=1$	$Y_t=1$ $Y_{t-1}=0$	$Y_t=1$ $Y_{t-1}=1$	Media 4 situaciones
Anual	69.8	78.0	34.9	58.2	80.3	62.9
Fríos	75.3 (77.4)	82.2 (84.2)	40.6 (43.8)	63.2 (65.8)	83.3 (83.3)	67.3 (69.3)
Templados	67.7 (64.4)	68.2 (78.0)	55.6 (33.3)	60.0 (45.0)	86.5 (75.7)	67.6 (58.0)
Cálidos	66.9 (63.0)	69.8 (66.7)	41.2 (22.2)	63.6 (70.0)	100 (100)	68.7 (64.7)

4. PREDICCIÓN CON HIRLAM-02

Los datos observados se comparan con las predicciones que genera HIRLAM-02, también con un horizonte de 6 horas, para la celda que contiene al aeropuerto de Zaragoza. El periodo de comparación va de Junio de 1997 a Septiembre de 2000.

El modelo HIRLAM proporciona sus resultados para un área que engloba a la Península Ibérica, con una resolución $0.2^\circ \times 0.2^\circ$. El modelo proporciona la previsión de lluvia para cada día, denotada TP, que es suma de dos componentes: la lluvia asociada a la acción de gran escala, LSP, y la generada por fenómenos convectivos, CP.

Periódicamente se incorporan al modelo numérico nuevos refinamientos en la representación de aspectos de los procesos atmosféricos, lo que da lugar a nuevas versiones. En Diciembre de 1999 se puso en marcha una nueva versión en la que se mantiene el esquema Sundqvist para la precipitación de gran escala y se sustituye ese esquema por el de STRACO para la lluvia convectiva. Esto permite representar el desarrollo de la convección somera por encima del estrato estable.

El modelo numérico sobreestima la cantidad de lluvia registrada. En la Figura 3 se muestra la discrepancia entre los suavizados de la cantidad de lluvia observada y la lluvia total prevista por HIRLAM-02. El modelo no reproduce el ciclo anual observado, pues indica la existencia de una época húmeda, inexistente, al inicio del mes de Septiembre.

El suavizado de la lluvia LSP tiene un máximo en el mes de Septiembre que explica, en parte, las discrepancias entre la lluvia total de HIRLAM y la observada. El valor medio de LSP es muy semejante, en bastantes meses, al valor medio observado (con diferencias

menores de 2 décimas de litro) pero es muy discrepante en los meses de Abril, Mayo, Agosto y Septiembre. La previsión de la lluvia CP produce valores medios demasiado grandes en los meses de Mayo, Agosto y Septiembre.

Figura 3: Valores medios suavizados de la cantidad observada (línea continua) y predicha por HIRLAM-02 (a trazos) a lo largo del año para la lluvia total y sus dos componentes

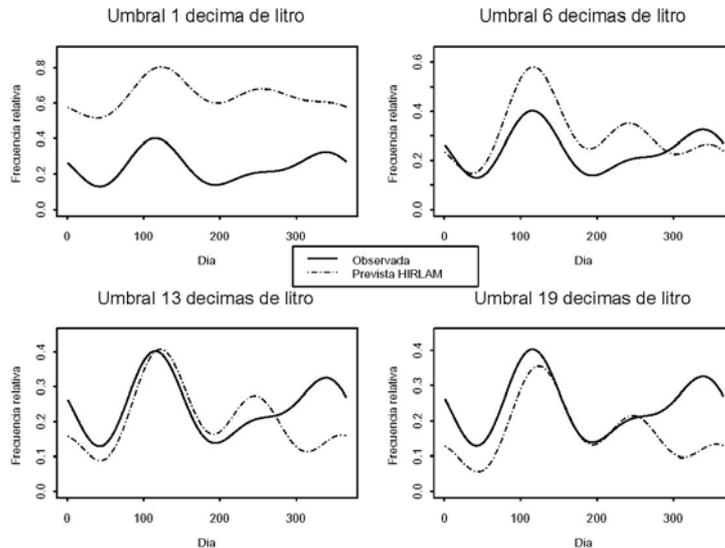


4.1. Evaluación de la predicción HIRLAM ajustando el umbral

La previsión de ocurrencia obtenida con HIRLAM usando el umbral 0.1 litros sobreestima la observada. Para solucionar ese sesgo se ha ensayado ajustar el umbral que determina la previsión de ocurrencia; se ha buscado aproximar la proporción observada de días húmedos con la proporción de días en los que HIRLAM prevé lluvia igual o superior a un cierto umbral.

En la Figura 4 pueden compararse los suavizados de la frecuencia relativa de que la lluvia TP sea igual o mayor que 4 umbrales seleccionados y el suavizado de la frecuencia observada de precipitación. No hay ningún umbral para el que exista una correspondencia a lo largo de todo el año; sin embargo, si se consideran las tres épocas definidas, se observa la similitud entre el suavizado de la frecuencia observada y el correspondiente a la predicha por HIRLAM usando los umbrales 0.6 litros en los meses fríos, 1.9 litros en los meses cálidos y 1.3 litros en los meses templados de Primavera. En los meses de Otoño es imposible acoplar los dos suavizados.

Figura 4: Valores medios suavizados de la frecuencia relativa observada (0.1 litros, en línea continua) y predicha (a trazos) por HIRLAM-02 tomando como umbrales los valores indicados



5. COMPARACIÓN CON LA PREDICCIÓN ESTADÍSTICA

Se ha comparado la predicción de ocurrencia del modelo estadístico anual con la de HIRLAM-02, tomando como umbral para la definición de día húmedo, una cantidad TP mayor o igual que 0.1 litros. En la comparación se utilizan los 820 días para los que se dispone de ambas predicciones, el periodo Junio/1997-Julio/2000, que forma parte, en un 80%, del periodo de validación del modelo estadístico, por lo que éste no tiene especial ventaja.

La Tabla IV muestra que los porcentajes de éxito son mejores en el modelo estadístico, tanto globalmente (un 10% más) como por estaciones y también según el estado de la ocurrencia en el día que se prevé y en el anterior, salvo en los días húmedos. La mayor frecuencia de cantidades TP positivas previstas por HIRLAM hace que los porcentajes de días húmedos bien predichos sean superiores a los que se obtienen con los modelos estadísticos, sobre todo en los comienzos de racha húmeda.

Tabla IV: Porcentajes de días bien previstos por el modelo estadístico anual y por HIRLAM-02, umbral 0.1litros

%aciertos	Estad.	TP
Global	68.0	57.6
Primavera	68.75	60.27
Verano	64.89	52.29
Otoño	61.79	53.66
Invierno	74.88	63.51
si $Y(t-1)=0$	72.2	62.1
si $Y(t-1)=1$	54.4	54.1
$Y(t)=0=Y(t-1)$	74.65	49.01
$Y(t)=1, Y(t-1)=0$	62.30	95.54
$Y(t)=0, Y(t-1)=1$	32.14	26.61
$Y(t)=1=Y(t-1)$	85.19	94.62

Al utilizar los umbrales variables estimados anteriormente, ver Tabla V, destaca el fuerte incremento del porcentaje de éxito obtenido en la predicción HIRLAM. El porcentaje global se incrementa entre un 19% en los meses fríos y un 32% en los cálidos, a pesar de las pérdidas en la predicción de los días húmedos. Los porcentajes de días bien previstos ahora por HIRLAM-02 son siempre superiores al 50%, salvo en los comienzos de racha húmeda en los meses cálidos. El menor porcentaje de éxito global es un 77% y corresponde a los meses templados (recordemos la imposibilidad de ajustar un umbral adecuado en el Otoño). Estos porcentajes de éxito de HIRLAM son superiores a los obtenidos con el modelo estadístico, con la excepción de los correspondientes a los días húmedos precedidos de otro húmedo en las tres épocas y a los días iniciales de racha húmeda en las épocas fría y cálida.

Tabla V: Porcentajes de días bien previstos por el modelo estadístico y por HIRLAM-02 usando umbrales apropiados en cada época

	Meses fríos		Meses templados		Meses cálidos	
	Modelo estad. anual	TP (umbral 6dl)	Modelo estad. espec.	TP (umbral 13dl)	Modelo estad. anual	TP (umbral 19dl)
Global	74.6	82.6	66.7	77.1	63.9	83.3
si $Y(t-1)=0$	79	82.1	65.5	82.8	67.4	90.1
si $Y(t-1)=1$	58.3	83.1	69.2	75.7	48.8	79.5
$Y(t)=0=Y(t-1)$	81.45	89.82	66.92	85.08	65.81	92.36
$Y(t)=1, Y(t-1)=0$	67.39	59.52	60.98	65.71	74.29	48.57
$Y(t)=0, Y(t-1)=1$	35.71	82.05	52.5	70.73	26.67	79.31
$Y(t)=1=Y(t-1)$	90.0	69.05	86.84	66.67	100.0	75.0

6. CONCLUSIONES

1) El tipo de modelo estadístico ajustado, del que se han ensayado ejemplos que usan, relativamente, poca información, puede ser útil en la previsión de la ocurrencia, puesto que ha producido resultados comparables, o mejores, que los obtenidos directamente por HIRLAM-02.

Los modelos estadísticos como el analizado, pueden ayudar en la previsión local, tanto de la ocurrencia de precipitación apreciable como de que la lluvia supere determinados umbrales. Los modelos podrían incorporar variables atmosféricas correspondientes a una malla rodeando al observatorio y utilizar las previsiones de HIRLAM para esas variables a la hora de predecir la ocurrencia. Otros posibles puntos de mejora pasan por definir con un criterio diferente algunas variables, por ejemplo, los rumbos de viento en superficie.

2) Los modelos de predicción estacionales son de interpretación más sencilla pero no muestran una capacidad predictiva sensiblemente superior a la del modelo anual.

3) La previsión estadística hecha a partir del umbral de la frecuencia relativa diaria es preferible a la que se obtiene utilizando el umbral 0.5. De todas formas, interesa estudiar y definir un umbral alternativo que combine las buenas cualidades de ambos valores.

4) HIRLAM-02 pronostica para Zaragoza unas condiciones de humedad que están lejos de las observadas; sus previsiones son, en general, mayores que los datos registrados y no reproduce el ciclo estacional. El uso de umbrales diferenciados para establecer la previsión mejora notablemente sus resultados.

AGRADECIMIENTOS

Agradecemos la colaboración del INM para suministrar los datos que se han utilizado en este trabajo.

REFERENCIAS

- Abaurrea, J. y Asín, J. (2003). A forecasting model for daily rainfall. Presentación oral en las VIII Jornadas Zaragoza-Pau de Matemática Aplicada y Estadística.*
- Armengot, R., Abellán, J. J. y Pérez-Cueva, A. J. (2001). Evaluación de un modelo de regresión logística de lluvias intensas en territorio valenciano con el episodio de octubre de 2000. En Pérez-Cueva, A. J. et al. (ed.) (2001), El Tiempo del Clima, pag. 291-301. Publicaciones de la AEC, Serie A, número 2.*
- Asín, J. (2003). Modelos para la precipitación diaria en el marco de los modelos lineales generalizados. Tesis doctoral sin publicar. Univ. de Zaragoza.*
- Del Hoyo, J., Fernández, A., Mestre, A. y Peral, C. (2002). Sistema de predicción analógica de la precipitación del INM (PAP-INM). Nota técnica del Servicio de Aplicaciones Meteorológicas. Ministerio de Medio Ambiente.*
- Ribalaygua, J., Borén, R. y Balairón, L. (1996). Clasificación de repartos de precipitación diaria sobre la España Peninsular y Baleárica: Aplicación para la generación de clasificaciones atmosféricas de apoyo a la predicción de precipitaciones. Actas IV Simposio Nacional de Predicción. Editado por el INM.*
- Sánchez Rodríguez, J. (1993). Situaciones atmosféricas en España. Editado por el Ministerio de Obras Públicas, Transportes y Medio Ambiente. Dirección Gral. del INM.*
- Serrano, A., García, J. A., Mateos, V. L., Cancillo, M. L. y Garrido, J. (1999). Monthly Modes of Variation of Precipitation over the Iberian Peninsula. Journal of Climate, 12, 2894-2919.*