

METODOLOGÍA PARA EL CONTROL DE CALIDAD Y HOMOGENEIDAD DE UNA BASE DE DATOS DE PRECIPITACIÓN DIARIA

Jesús ABAURREA; Jesús ASÍN; Ana C. CEBRIÁN y Alberto CENTELLES
Departamento de Métodos Estadísticos, Universidad de Zaragoza

RESUMEN

Se propone un proceso para seleccionar una base de datos de calidad de series de lluvia diaria y se consideran los problemas más comunes en el registro de datos. El procedimiento se basa en controles de tipo absoluto y relativo para decidir la calidad y homogeneidad de las series de lluvia diaria. Las herramientas diseñadas se aplican en diferentes observatorios de la Cuenca del Ebro.

Palabras clave: Control de calidad y homogeneidad, precipitación diaria, selección de una base de datos.

ABSTRACT

A procedure for selecting a high quality daily rainfall data set, dealing with the most usual problems in data recording is proposed. The procedure is based on absolute and relative type homogeneity and quality controls. We show the application of the methods proposed with rainfall series from the Ebro river basin.

Key words: Homogeneity and quality data control, daily rainfall series, data set selection.

1. INTRODUCCIÓN

El control de calidad de los datos climáticos es una obligación tanto para los servicios meteorológicos responsables de su registro y custodia, que deberían ofrecer mediciones correctas en un alto porcentaje o, en caso contrario, advertir al usuario de los problemas existentes, identificando los intervalos no utilizables, como para los investigadores del clima que requieren validar esa información antes de usarla, para no introducir señales no climáticas en sus resultados. Cuando se trabaja con datos de lluvia diaria, el control de homogeneidad se convierte en una tarea ardua, debido al volumen de información a tratar, y difícil, por la variabilidad de la lluvia diaria, AGUILAR *et al.* (2004).

Por ello es habitual, al poner a punto las bases de datos diarias para investigación, que la homogeneidad se controle sobre magnitudes agregadas. Así, la calidad de los observatorios que participan en el proyecto ECA (European Climate Assessment), seleccionados por los respectivos servicios nacionales, fue verificada, mensualmente, mediante comparación con los valores de las series de precipitación de Hulme (datos promediados asociados al centro de una celda) y aplicando 4 tests de homogeneidad a la variable, número de días de precipitación anual. LAURSEN *et al.* (2001) seleccionaron como base de precipitación diaria en Dinamarca, las series de la base mensual NACD que resultaron homogéneas en la aplicación estacional de SNHT. Al menos hasta Diciembre de 2002, esta misma estrategia se utilizaba en la base de datos USCHN; la homogeneidad se controlaba en las series mensuales por procedimientos relativos y la base de

datos diaria la formaban las series originales de los observatorios sin desplazamientos, con un registro prolongado y que no habían mostrado alteraciones en los controles mensuales.

En Australia, LAVERY *et al.* (1992) seleccionaron una base de datos de precipitación diaria aplicando el test de Craddock a la lluvia acumulada anual y controles gráficos, basados en el dato diario, para detectar a los observadores poco fiables. En España, ROMERO *et al.* (1998) construyeron una base de datos de precipitación diaria 1964-1993, formada por 410 observatorios situados en la franja mediterránea y Baleares. Su proceso de control analizaba la verosimilitud de cada dato, comparándolo con un valor de referencia obtenido a partir de las observaciones del mismo día en los observatorios cercanos. Para ello, además de los observatorios de la base de datos, consideraron otros 2.842 que, en el periodo seleccionado, tenían al menos 1.000 datos diarios.

El objetivo de este trabajo es presentar una metodología para detectar los problemas asociados al registro de los datos de lluvia diaria. En la siguiente sección se revisan los problemas que se encuentran en este tipo de datos y sus causas; en la sección 3 se presenta la estructura del proceso de control propuesto; en las secciones 4 y 5 se exponen las técnicas de control, de carácter absoluto y relativo, que se utilizan. Estas técnicas se aplican sobre series de la Cuenca del Ebro de buena calidad, Tortosa, Pineta y Flix, y en otras que presentan problemas, Monterde y Pueyo de Jaca. La sección 6 contiene las conclusiones y algunas líneas de trabajo futuras.

2. PROBLEMAS EN EL REGISTRO DE LOS DATOS DIARIOS

La recogida del dato diario de precipitación, útil para investigación, requiere gran constancia para lograr series continuas y rigor en la medición. Las faltas en esa constancia y rigor, junto con eventuales cambios de observador o de aparato de registro, inevitables cuando se estudian periodos amplios, provocan que la calidad de los datos se modifique en el tiempo y que puedan volverse inadecuados. Los controles de calidad deben detectar e identificar los distintos problemas posibles, asumiendo la variabilidad de la lluvia diaria. La búsqueda en el *metadata* de la serie, una labor costosa, permitirá explicar buena parte de los resultados que aparezcan en los controles.

El primer requisito de una serie climática útil es cumplir unas condiciones mínimas de número de datos; otros elementos a valorar son la continuidad del registro, el estado actual, activo o no, del observatorio, la distribución espacial de otras series útiles cercanas, etc.

Existen dos causas principales para la introducción de variabilidad no climática en los registros: los “cambios físicos” (cambios de aparato, de ubicación geográfica o de observador) y los “hábitos del observador”. Cada observador presenta hábitos de medición propios: el umbral que registra con rigor, la precisión de medida, etc., que influyen en la calidad de la serie. El INM establece una precisión de 0,1 mm para la precipitación pero los observadores pueden registrarla con algún redondeo. Una serie con escasa precisión cuyos episodios de lluvia se hayan medido “correctamente”, puede tener calidad suficiente para analizar en ella prácticamente cualquier aspecto de interés. El *umbral de registro regional* (valor mínimo de precipitación a partir del cual, todos los observatorios, en todo el intervalo de estudio, registran, habitualmente, una cantidad no nula) debe seleccionarse cuidadosamente porque servirá para definir, homogéneamente, otras variables, como longitudes de racha, frecuencias de ocurrencia, etc.

Algunos “malos hábitos”, días no observados en los que se registra 0 mm, ausencia sistemática de medición en algunos días de la semana, acumulaciones frecuentes de precipitación en días consecutivos no observados, etc., pueden hacer inservibles los registros. La acumulación de malas observaciones o el cambio de precisión pueden provocar variaciones, inhomogeneidades, en algunas series agregadas.

3. MÉTODOS DE CONTROL DE UNA BASE DE DATOS DIARIA

Para seleccionar una base de datos de lluvia diaria en la Cuenca del Ebro se ha diseñado una estrategia en la que tratamos de identificar los problemas más comunes y detectar su ubicación, utilizando diversos controles. Se propone una sucesión de etapas, en las que se eliminarán series o fragmentos de ellas, que ya no serán analizadas en etapas posteriores, y en las que se identificarán las fechas de cambio en los registros, que deben contrastarse posteriormente usando el *metadata*.

Los problemas que se detecten permitirán clasificar las series, según su calidad, en tres tipos:

- A, series *óptimas o candidatas*, que pasan los controles requeridos en el periodo en que se encuentran activas o en un subperiodo muy amplio;
- B, series *auxiliares*, que pueden servir de contraste y apoyo a las primeras en periodos determinados, y,
- C, series *no útiles*, que presentan problemas graves de homogeneidad, por lo que no deberían ser usadas.

Hay que señalar que en la formación de la base de datos no se va a realizar ningún proceso de corrección de los datos que se consideren anómalos, ya que no se dispone, de momento, de una técnica conveniente de homogeneización.

El orden de aplicación de los procedimientos de control es importante y se ha diseñado buscando el ahorro de tiempo y trabajo del investigador.

- a) En primer lugar, se realizan los controles absolutos “rápidos” que analizan la evolución de la precisión de las medidas, la distribución de frecuencia de las cantidades positivas y la ausencia de medición en días concretos de la semana. Estos indicadores, junto con la descripción gráfica de la serie que muestra la evolución de la lluvia en el observatorio, permiten al investigador alcanzar una visión preliminar del conjunto de series inicial. Tras esta etapa, se descartan aquéllas con indicios de cambios en la frecuencia o con ausencia sistemática de observación. Finalmente, se selecciona el umbral de registro regional necesario para los controles posteriores.
- b) En la segunda etapa, se aplican procedimientos relativos para identificar posibles inhomogeneidades sobre series agregadas:
 1. Se realiza un análisis de los datos mensuales con cantidad nula, para decidir si corresponden a una ausencia de observación.
 2. Se aplica un análisis de homogeneidad, similar al propuesto por ALEXANDERSSON y MOBERG (1997) utilizando SNHT, para detectar cambios no climáticos en la serie de precipitación acumulada anual y en la de frecuencia de días con precipitación mayor o igual que el umbral de registro regional seleccionado. Estos controles deben ser satisfactorios en todas las series que constituyan la base de datos.

- c) En último lugar, debido al esfuerzo que requiere su aplicación, se aplican los controles relativos a las series diarias en las que no se hayan detectado problemas graves de calidad. Si se detectan datos extremos aislados en su entorno geográfico, se utilizará la información disponible en los archivos del INM para establecer la verosimilitud de esos valores que pueden ser influyentes.

En la Cuenca del Ebro se han aplicado ya las técnicas de la primera etapa a todas las series que verifican los requisitos de número de datos exigidos; el umbral de registro seleccionado ha sido 1mm. Actualmente se está completando el análisis de homogeneidad relativa de las series de cantidad agregada anualmente y está pendiente el de la frecuencia de precipitación. Los procedimientos relativos sobre dato diario se han aplicado sobre un banco de datos de ensayo.

4. CONTROLES DE TIPO ABSOLUTO

En este apartado se presentan las herramientas utilizadas para evaluar la precisión de las medidas, seleccionar el umbral de registro e identificar “malos hábitos” del observador.

4.1. Control de la precisión

El índice de precisión real (Real Precision Index, RPI), PETROVIC (2001), es una medida para estudiar los cambios en la precisión de las mediciones y se basa en la frecuencia que en cada año tienen los dígitos 0 a 9 en la cifra de las décimas de litro de las observaciones positivas. Dado que la precisión definida por el INM es 0,1mm, se establece el nivel 0,15 como el valor límite de precisión óptima. En la figura 1 se muestra el RPI de las series de Tortosa, Pineta y Flix.

El RPI distingue épocas con diferente precisión pero no identifica el hábito que provoca esos cambios. Para verlo se ha diseñado un gráfico (Fig. 1, inferior), que dibuja, para cada año, una barra de altura proporcional a la frecuencia de días con lluvia registrada. En cada barra se diferencia, con colores, la frecuencia con la que aparece cada dígito. De este modo se identifican:

1. Los problemas de frecuencia, visibles por el cambio en la altura de las barras (el número de días sin lluvia en cada año se escribe al pie) y,
2. Las cifras más usadas por cada observador. Flix tiene una precisión aceptable en 1941-1948 y 1977-1987, usa exclusivamente 0 y 5 en 1962-1976 y una resolución de 1 mm en 1990-2000.

4.2. Análisis de frecuencias

Se analiza la distribución de las cantidades registradas para valorar cuál es el umbral de registro de cada serie. Para ello se representa la frecuencia acumulada en cada cantidad registrada, hasta 10 mm, diferenciando la correspondiente a cada década (Fig. 2).

Tortosa muestra un comportamiento completamente satisfactorio, concentración de observaciones en los valores pequeños y descenso paulatino de la frecuencia, sin que destaquen barras aisladas. En Pineta, hay una precisión cercana a 0,5 mm, que implica una ausencia de cantidades inferiores, salvo en el periodo 1941-1950. Flix tiene mayor precisión en las dos primeras décadas pero ésta se reduce desde 1961 donde, en la práctica, mide sólo cantidades superiores al litro.

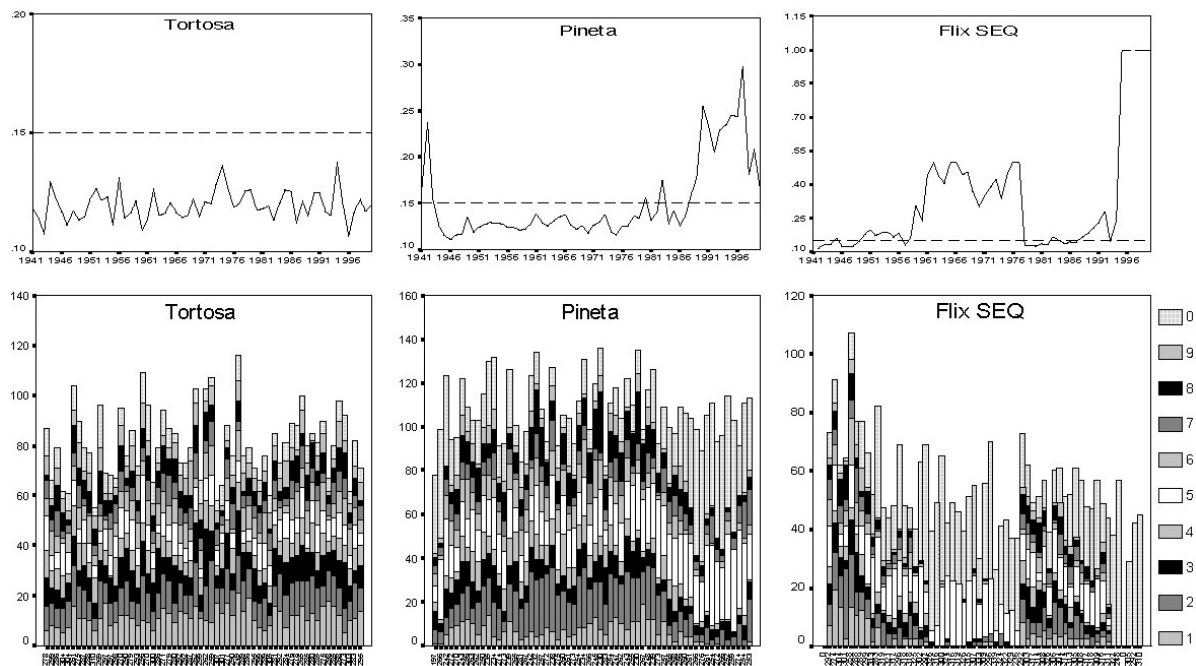


Fig. 1. Real Precision Index (superior) y frecuencia anual de la última cifra significativa (inferior), en Tortosa, Pineta y Flix

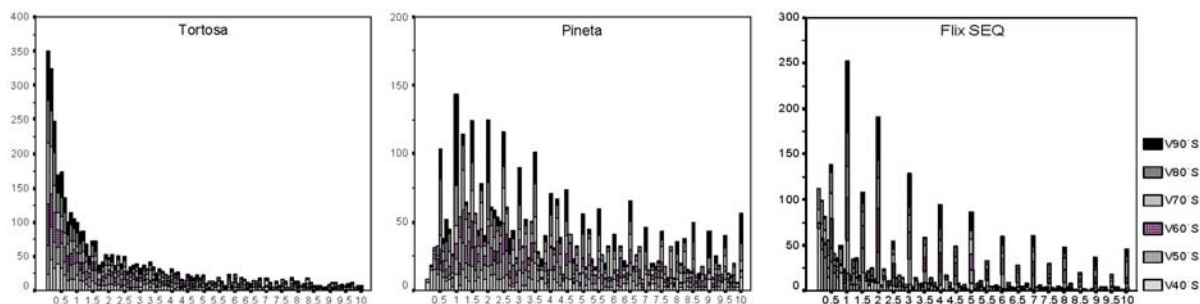


Fig. 2. Frecuencia acumulada en los valores menores que 10 mm, en Tortosa, Pineta y Flix

4.3. Control de malos hábitos

El objetivo de este control es analizar si el observador registra el dato todos los días o si existe una ausencia sistemática de observación en algún día de la semana. Suponiendo, lógicamente, que la ocurrencia de lluvia es uniforme en los días de la semana, comparamos, usando un gráfico de áreas acumuladas, la frecuencia relativa de lluvia en cada día (Fig. 3, izda). Se propone también (Fig. 3, centro), un gráfico que muestra el cociente de la frecuencia en viernes, sábado y domingo entre la frecuencia media en el periodo lunes-jueves.

En la figura 3 se muestran los gráficos de Pueyo de Jaca donde se produce un cambio sistemático a partir de 1993, con un aumento de la frecuencia los domingos y una disminución los viernes y sábados, lo que indicaría que no se mide los fines de semana y, el lunes a las 8h, la lluvia

acumulada se registra asociada al domingo. Esto se confirma en la figura 3, derecha, donde se representa la precipitación anual registrada en domingos, sábados y martes.

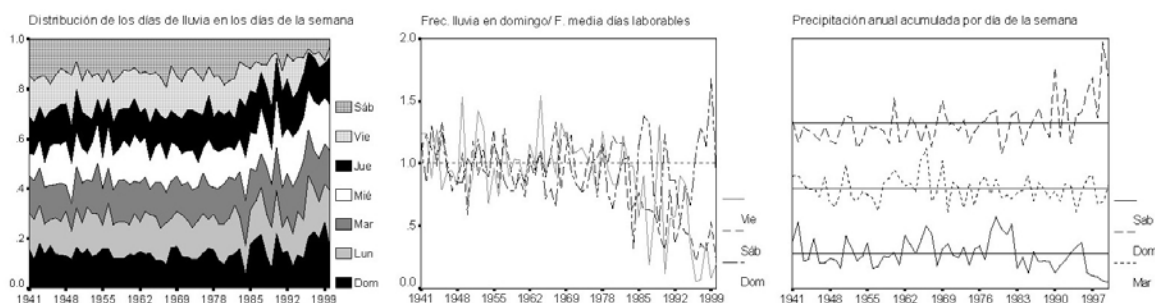


Fig. 3. Control del registro en fin de semana, en Pueyo de Jaca

5. CONTROLES DE TIPO RELATIVO

Estos controles requieren una selección adecuada de las series auxiliares, en la que se debe tener en cuenta la proximidad, la similitud geográfica, la existencia de un periodo de registro común de cierta longitud y un alto grado de correlación. Se han seleccionado como series auxiliares las cuatro más cercanas, a las que se añaden, las más correladas entre las restantes, exigiendo coeficientes mayores que 0,70. Se utiliza, por lo general, un número máximo de 8 series auxiliares, de las que, al menos 4, deben tener el dato analizado. Los periodos más difíciles para cumplir estos requisitos son la década de los noventa en Cataluña y el periodo 1941-1950. En la mayoría de los casos, la serie más cercana se encuentra a menos de 15 km y el conjunto de ellas dentro de un radio de 50 km.

5.1. Registros de lluvia nula en días sin observación

La primera detección de este hábito se hace en la serie agregada mensual. Un mes es un periodo en el que la ausencia de precipitación, salvo en verano, es poco frecuente en la mayor parte de la cuenca; por ello, si ocurre, debe tener reflejo en las series auxiliares.

Consideramos un observador “sospechoso” cuando su serie registra 0 mm y se produce una de estas condiciones:

- Al menos un 75% de las series auxiliares miden una precipitación mensual superior a 15 mm,
- Las dos series más cercanas al observatorio sobrepasan los 10 mm.

Los meses sospechosos se recogen en tablas que incluyen información sobre la distancia y correlación con las series auxiliares. La tabla 1 corresponde a Monterde y recoge los 13 meses en los que hay alarma; de ellos, los 4 en negrita son los que se han considerado sospechosos tras investigar el registro diario.

Tabla 1. Lluvia mensual nula poco verosímil, en Monteverde

Lluvia mensual			Series más cercanas				S. adicionales más correladas			
Año	Mes	Candidata	1ª	2ª	3ª	4ª	1ª	2ª	3ª	4ª
1952	5	0	18,9	18,9	12,9	31,9	27,9	8,5	21,5	40,4
1952	9	0	13,9	14,4	18,1	22,5	17,3	9,9	15	31,8
1954	7	0	24,2	17,6	13,3	22,7	62,7	18,6	2	22,5
1957	8	0	14,5	4	26	17,2	54,1	22,5	40,5	
1961	4	0	33,8	32	37,9	41	59,1	16,9	69	45,2
1964	10	0	16,3	11,5	6	15	31,4		10	23
1966	9	0	9,4	7	15,6	27,4	26,5	30,9	24	20,5
1971	8	0	11,7	10	37,6	25,7	28	2,7	0	8,5
1973	9	0	22,5	5	15,7	19	14,7	3,9	15	11,7
1975	11	0	10,8		10,7	20	14,9	6,5	19,5	13,8
1977	10	0	22,1	23,2	15,4	18,4	16,5	25,4	30	33,8
1980	10	0	15,6	28,3	18,8	25,2	17,2	23,1	28	29,7
1990	12	0	29,5	17	25,9	26,6	42,1	33,8	25	25,8
Distancia (km)			17,05	23,46	27,84	34,35	39,2	45,69	50,07	45,27
Correlación			0,81	0,78	0,79	0,78	0,76	0,73	0,72	0,71

5.2. Desplazamiento en la fecha

El desplazamiento en la fecha de registro puede ser un hecho aislado o un hábito, durante cierto periodo, del observador. Para detectar la existencia de desplazamientos reiterados se ha diseñado un índice que toma valores anualmente y se calcula como el cociente de numerador la suma anual de las diferencias, en valor absoluto, de la precipitación diaria registrada en la serie candidata y en la auxiliar, y denominador la suma de las mismas diferencias cuando la serie candidata se desplaza un día hacia adelante. Este índice será más próximo a cero cuanto mayor sea la semejanza entre la serie candidata y la auxiliar. El procedimiento se aplica de nuevo, retrasando un día las observaciones de la serie candidata. Al tratarse de un control relativo con una única serie de referencia, los problemas detectados pueden corresponder a la serie candidata o a la auxiliar; por ello este control debe aplicarse usando, al menos, como referencias, las dos series más cercanas.

La figura 4 muestra las gráficas para el observatorio de Pineta. Se detectan problemas en 1966 y en los últimos años de registro. El análisis de los ficheros diarios permite identificar un retraso de un día en la mayoría de las observaciones de los seis primeros meses de 1966.

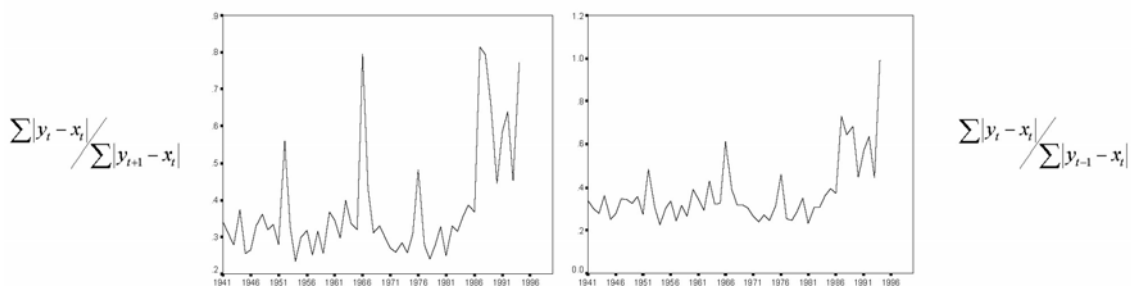


Fig. 4. Gráficos de control de desplazamientos de fecha, en Pineta

5.3. Observaciones anómalas

Observaciones anómalas son aquéllas que no son refrendadas por las mediciones hechas en los observatorios del entorno geográfico; esto no significa siempre que los datos sean falsos dada la gran variabilidad espacial de la precipitación.

Hemos definido dos controles; el primero usa un nivel fijo de lluvia (5mm) y detecta las alarmas denotadas PA (precipitación aislada) y NPA (no-precipitación aislada), que se especifican en la tabla 2. El segundo identifica las alarmas PPA (pequeña precipitación aislada) y GPA (gran precipitación aislada) y está basado en un índice que reexpresa la lluvia tomando como origen el percentil 75 y como escala el rango intercuartílico. Estos estadísticos se han calculado, en cada caso, en una muestra de 15 meses (3x5), formada por el trimestre centrado en el mes correspondiente al dato analizado, en 5 años que cubren el año de la observación. Por cuestiones operativas los quinquenios considerados han sido: 1951-1955, 1956-1960, etc. Las alarmas y reglas de decisión empleadas se exponen en la tabla 2. Cada alarma declarada debe analizarse con la ayuda de mapas e información relevante, para validar el dato o declararlo perdido.

Tabla 2. Controles, alarmas y criterios de decisión

Controles	Alarma	Candidata	Auxiliares
Índice con nivel fijo (5mm)	NPA	Lluvia 0mm	Lluvia superior a 5mm
	PA	Lluvia superior a 5mm	Lluvia 0mm
	Criterio: No discrepa ninguna de las series auxiliares		
Índice con nivel relativo (LL-P75)/IQR	PPA	Índice <0	Índice >=1,5
	GPA	Índice >= 6	
		Índice >= 3	Índice <= 1,5
	Criterio: Confirman el dato todas las series cercanas y se admite una sola discrepancia en las series más correladas		

La alarma NPA detecta cantidades nulas sospechosas, que pueden corresponder a falta de observación o a desplazamiento; la utilidad de PPA ha demostrado ser menor. Las alarmas PA y GPA detectan observaciones que han podido sufrir desplazamiento, que son precipitación acumulada en varios días o mediciones erróneas.

5.3.a. Alarmas con un nivel fijo

El análisis se efectúa creando tablas, por meses y años, que identifican los periodos con mayor concentración de alarmas. No se ha determinado, por ahora, un valor crítico para rechazar una serie o subserie, sino que se compara el número de alarmas en cada año con los resultados obtenidos en series similares.

La coincidencia de alarmas PA y NPA en una época indica un problema de desplazamientos de fecha en la serie. En 55 años de observación, Pineta tiene 25 años (29) sin alarmas NPA (PA). En 1966 (Tabla 3), se acumulan 11 y 9 alarmas de cada tipo, que representan el 15,9% y 13,4% del número total, mientras en el periodo 1991-2000 se observan 25 y 36 alarmas (36,2% y 53,7%);

esto confirmaría lo señalado en la sección 5.2. Dos índices útiles para caracterizar cada periodo son la expresión en % de los cocientes:

- a) Número anual de alarmas NPA entre la frecuencia anual de días sin lluvia y,
- b) Número anual de alarmas PA entre la frecuencia anual de días con lluvia superior de 5 mm.

La Figura 5 (1ª y 2ª por la izquierda) representa la evolución de estos índices en Pineta.

Tabla 3. Resumen de alarmas NPA y PA, en Pineta

Mes	Alarmas NPA												Total	Alarmas PA												Total
	1	2	3	4	5	6	7	8	9	10	11	12		1	2	3	4	5	6	7	8	9	10	11	12	
1965										2		1	3										1		1	
1966	1	5			2	1	1		1				11	2	3			3		1					9	
1967											4	1	5								1			1	2	
1993					1								1							2					2	
1994									1	3	2		6			1	1	1		3	2	1		9		
1995	1	1	1										3	2					1					3		
1997						1					1		2			2	1							3		
1998	1			1	1			1	2				6			1	1		1				1	4		
1999			1	1		1			1				4	1			1			1				3		
Total	4	8	7	5	5	4	2	3	8	5	13	5	69	7	5	1	4	10	8	9	7	6	3	4	3	67

La presencia de alarmas PA y la ausencia de alarmas NPA en un periodo indica la existencia de observaciones que acumulan la lluvia de varios días. En Monterde (Fig. 5, dcha.), el índice PA toma valores superiores al 20% varios años, cuando en Daroca es siempre inferior al 10%.

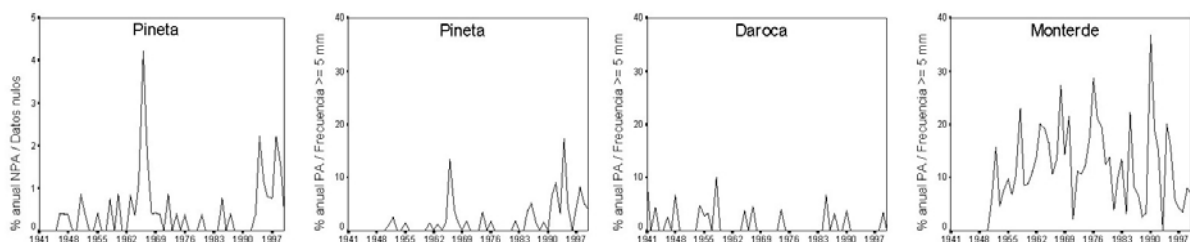


Fig. 5. Índices de alarmas NPA en Pineta y PA en Pineta, Daroca y Monterde

5.3.b. Alarmas basadas en las diferencias de nivel relativo

Estos controles sólo se aplican a series que han pasado los controles anteriores, por lo que, en esta fase final, su número de alarmas no debe ser grande, 30-40, lo que hace su revisión factible.

Pineta presenta 19 alarmas GPA, de las que 11 corresponden a precipitaciones importantes en la serie de Barrosa, observatorio a 4,5 km con alta correlación; cuatro son dos rachas húmedas, 28-29/3/52 y 4-5/11/73, que confirman las series auxiliares. Tras revisar los registros diarios, las 4

alarmas restantes se revelan como datos dudosos sobre los que se debe consultar información adicional.

Daroca tiene 32 alarmas GPA, de las que 21 son congruentes con alguna de las series cercanas, a menos de 30 km. Las 11 alarmas restantes, 4 en 1950, deben validarse en el metadata de la serie.

6. CONCLUSIONES

Se propone un procedimiento para evaluar la calidad y homogeneidad de las series de precipitación diaria que se está aplicando para elaborar una base de datos en la Cuenca del Ebro. Los controles de carácter absoluto y relativo aplicados, permiten verificar cualidades de los registros que no es posible estudiar analizando la lluvia agregada. Además, localizan los grandes valores discrepantes con su entorno, que deben ser estudiados posteriormente para declararlos, o no, datos válidos.

Esta línea de trabajo sigue abierta: se están ensayando alternativas para los controles de tipo relativo, basadas en la reexpresión de las cantidades diarias en percentiles; también se está estudiando la utilidad de gráficos de probabilidad (γ) para analizar las distribuciones de las cantidades e identificar tanto desviaciones sistemáticas como valores anómalos. Finalmente, queremos diseñar un procedimiento similar al de ROMERO *et al.* (1998) que asocie un valor de referencia a cada observación diaria y permita analizar la existencia de sesgo en las mediciones.

7. AGRADECIMIENTOS

Los autores agradecen a la dirección y personal del Centro Territorial del INM en Zaragoza la colaboración prestada con los datos. Este trabajo ha sido financiado por el proyecto REN2002-00009 del Ministerio de Ciencia y Tecnología.

8. REFERENCIAS

- AGUILAR, E. *et al.* (2004). *Guidelines on climate metadata and homogenization*. WMO. WCDMP 53 TD 1186.
- ALEXANDERSSON, H. and MÖBERG, A. (1997). "Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends". *Int. J. Clim.*, 17, pp. 25-34.
- KLEIN TANK, M.G., *et al.* (2002). "Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment". *Int. J. Climatol.*, 22, pp. 1441-1453.
- LAURSEN, E.V. *et al.* (2001). *Observed daily precipitation, temperature and cloud cover from seven Danish sites, 1874-2000*. Technical report 01-10, DMI, Copenhagen.
- LAVERY, B.; KARIKO, A. and NICHOLLS, N. (1992). "A historical rainfall data set for Australia". *Aust. Met. Mag.*, 40, pp. 33-39.
- PETROVIC, P. (2001). Selection of data sets by quality and its role in climate research. En: LÓPEZ, D. y BRUNET, M. (Eds.). *Detecting and Modelling Regional Climate Change*. Springer, pp. 21-29.
- ROMERO, R. *et al.* (1998). "A 30-year (1964-1993) daily rainfall data base for the Spanish Mediterranean regions: First exploratory study". *Int. J. Climatol.*, 18, pp. 541-560.