

# Rescate de datos climáticos antiguos: digitalización de bandas de pluviógrafo del observatorio de A Coruña

Rafael SÁNCHEZ GONZÁLEZ, Pablo GÓMEZ VIÑAS

*AEMET, Delegación Territorial en Galicia*

*rsanchezg@aemet.es, pgomezv@aemet.es*

**Resumen:** El observatorio de A Coruña dispone de registros de pluviógrafo desde 1918. El archivo de las bandas de papel generadas por estos equipos registradores presenta alguna laguna y hay periodos en que la calidad de las mismas no permite obtener valores fiables de precipitación. No obstante, su digitalización facilitaría una serie de datos con resolución temporal del orden de minutos, como la que obtienen actualmente los pluviógrafos de estaciones meteorológicas automáticas. En la delegación de AEMET en Galicia se está llevando a cabo una digitalización de estos registros para los años anteriores a 1984, periodo que permite enlazar con los valores horarios de lluvia que de forma manual se extrajeron de estas mismas bandas desde ese año. Este trabajo describe el proceso de digitalización empleado, muestra las dificultades que se han encontrado y presenta algunas aplicaciones de los datos obtenidos.

**Palabras clave:** pluviógrafo de sifón; pluviómetro de balancín; rescate de datos; digitalización; intensidad de precipitación.

## 1. INTRODUCCIÓN

La información climática actual se registra, en general, con gran resolución temporal, normalmente cada diez minutos, mientras que los datos más antiguos corresponden a intervalos más amplios, habitualmente diarios o hexahorarios. Resulta necesario con frecuencia ampliar los periodos de integración de las medidas actuales para que las series resulten comparables con los registros históricos. En el caso de la precipitación, por ejemplo, estos datos de resolución superior a un día son esenciales para el análisis de riesgos meteorológicos y de sus impactos (BONI *et al.*, 2006). En muchas estaciones la información base para la elaboración de estas series de datos está disponible, pero no digitalizada. Este proceso de recuperación de datos antiguos en distintos soportes se conoce habitualmente como «rescate de datos» y se ha desarrollado considerablemente en las últimas décadas (BRÖNNIMANN *et al.*, 2018). Aquí se entiende por rescate de datos los relativos a la era preinformática, independientemente de que algunos aún se registren de forma manual o de que otros se encuentren en soportes informáticos antiguos que requieran algún tipo de tratamiento para hacerlos más fácilmente accesibles. En concreto, este trabajo se refiere a un tipo de registro analógico en papel: las bandas de pluviógrafo.

La Organización Meteorológica Mundial cuenta con un portal dedicado a esta cuestión, IDARE, <https://www.idare-portal.org/es>, y recopila guías para la digitalización (WILKINSON *et al.*, 2019). También colabora en la puesta en marcha del Copernicus Climate Change Data Rescue Service: <https://climate.copernicus.eu/data-rescue-service>.

## 2. MÉTODO DE DIGITALIZACIÓN

Los archivos climatológicos de la delegación de AEMET en Galicia guardan una gran cantidad de información en papel. Buena parte de ella se conserva en formato numérico, como cuadernos de observación o sondeos con globo piloto, disponibles ambos desde 1914 para el observatorio de A Coruña, junto con tarjetas de colaboradores u hojas climatológicas de diversos observatorios

de Galicia. A grandes rasgos, puede decirse que la información climatológica disponible en formato numérico está digitalizada por completo desde 1961 y de forma más resumida en los periodos anteriores, donde los datos son más escasos y con frecuencia necesitan una interpretación y depuración previas para asegurar que resulten comparables con los registrados actualmente. Desde finales del siglo XX se cuenta con valores horarios de las principales variables (temperatura, precipitación y viento). También en esas fechas parte de la red de observación se automatiza y se conservan registros diezminutales de distintos parámetros. En años aún más recientes la automatización de la red abarca, además de los observatorios principales, un buen número de estaciones colaboradoras. Contrasta la situación actual, donde prácticamente todas las variables disponen de un registro diezminutal, con la anterior a 1960, en que apenas se dispone más que de valores medios diarios o mensuales.

La información climatológica contenida en las bandas de los equipos registradores arranca a comienzos del siglo XX y convenientemente tratada puede extender hacia atrás las series de gran resolución temporal que se están construyendo actualmente. Quizá sean las bandas de pluviógrafo, por las especiales características de la variable que miden, las que resulte más útil digitalizar. Mientras que los valores extremos de otros parámetros fundamentales como la temperatura quedaron registrados diariamente por observación directa y la digitalización de las bandas de un termohigrógrafo no suele aportar más información que una mayor resolución temporal de valores intermedios, de la precipitación quedan ocultos los valores extremos, que son los fundamentales a la hora de describir posibles cambios en el régimen pluviométrico de una estación y de hallar periodos de retorno de las lluvias extraordinarias. En particular, los avisos por lluvias intensas se elaboran para precipitaciones máximas en doce horas (AEMET, 2022) y se da la circunstancia de que se dispone de series muy cortas de esta variable que puedan servir de apoyo en la definición de los umbrales y, además, únicamente para estaciones principales.

En la delegación de AEMET en Galicia se está llevando a cabo la digitalización de las bandas históricas de pluviógrafo del observatorio de A Coruña. El método seguido para ello no es novedoso. Se basa en considerar el trazo de la plumilla sobre la banda como una línea quebrada e introducir en un programa informático los puntos donde se producen cambios significativos de pendiente. Se inició la digitalización con las bandas de este observatorio en el año 1984 y se avanza hacia años más antiguos, pues en los posteriores a 1984 se cuenta, al menos, con registros horarios de lluvia y también de intensidades máximas en ciertos periodos del orden de las decenas de minutos. El de A Coruña no es el primer observatorio de Galicia en el que se aplica esta técnica. El de Ourense informatizó sus registros de pluviógrafo de forma totalmente manual en el periodo 1986 a 2009. En el caso de Ourense se introdujeron desde el teclado los puntos significativos, lo que supone que el procedimiento es lento y requiere una gran dedicación. Sus resultados, sin embargo, son muy buenos.

Existen diversos métodos automáticos basados en el escaneo de las bandas y su tratamiento informático posterior, mediante técnicas de segmentación para aislar del ruido de la gráfica (líneas impresas en la misma, anotaciones manuales sobre la banda, manchas de tinta, etc.). Estos procedimientos ofrecen la ventaja de la rapidez, aunque inevitablemente implican una revisión manual para corregir fallos del programa en situaciones anómalas: (JAKLIČ *et al.*, 2016; SUŠIN y PEER, 2018). En el caso objeto de este estudio, el método empleado requiere el tratamiento, una a una, de las bandas. Después de su escaneo (figura 1a), se utiliza un programa de digitalización para el que ni siquiera es necesario una tableta digitalizadora, basta con un simple ratón para marcar los puntos significativos (figura 1b). Información sobre el programa utilizado, Graph-Grabber, puede encontrarse en [www.quintessa.org/graph-grabber](http://www.quintessa.org/graph-grabber). Obviamente, en este proceso se corrigen defectos en la gráfica a la vez que se digitaliza, eliminando la necesidad de una revisión posterior de la misma. El proceso es rápido, una vez adquirida práctica no lleva más de dos o tres minutos por banda. El resultado es un fichero de texto con coordenadas del tipo (tiempo, precipitación). Finalmente, para evitar posibles anomalías antes de archivar los valores, se dibuja

y se compara con la original mediante una aplicación desarrollada en R (figura 1c). También se compara con los datos diarios de precipitación almacenados en el archivo climatológico, de modo que se descartan bandas muy alejadas de la realidad.

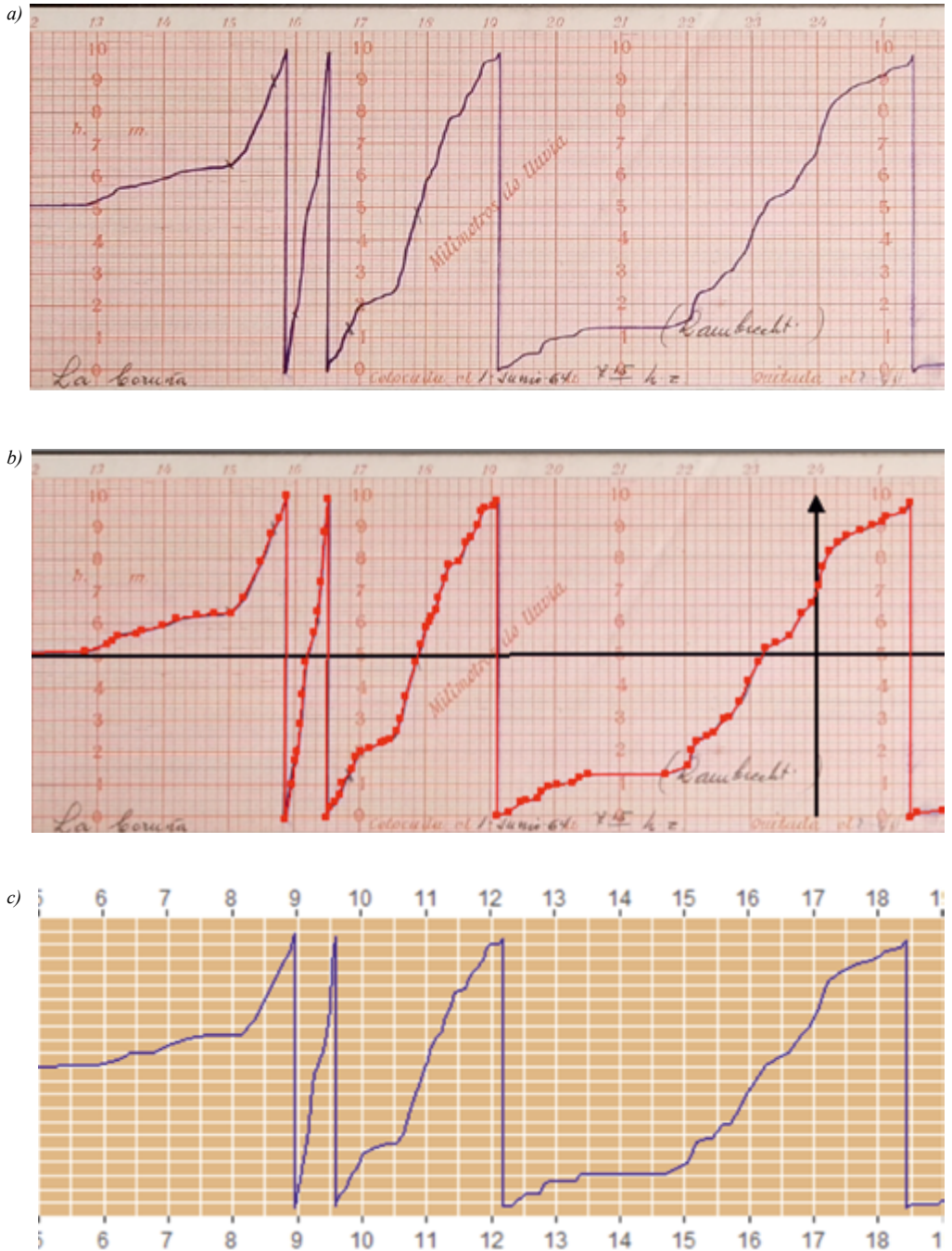


Figura 1. Proceso de digitalización de una banda de pluviógrafo. a) Imagen escaneada de la banda, b) digitalización, c) reproducción de la banda en R.

### 3. DIFICULTADES EN EL PROCESO DE DIGITALIZACIÓN

Las tareas de rescate de datos antiguos suelen tropezar con múltiples problemas que, en la práctica, limitan su utilidad. En este trabajo se pone de manifiesto (figura 2) que los valores recogidos a partir de los registros del pluviógrafo subestiman los verdaderos valores de precipitación. En algunos casos excepcionales, como el año 1978, la precipitación contabilizada mediante las bandas del pluviógrafo apenas supera el 60 % de la precipitación real anual. Esto se debe a que a veces falta alguna banda, pero lo más común es que exista y no haya medido adecuadamente. De forma relativamente frecuente, llegado al tope de 10 mm, la plumilla no retorna al cero de la escala. Se pierde así una parte importante de la precipitación, más si se tiene en cuenta que eso suele ocurrir precisamente los días más lluviosos y en los que serían de esperar las mayores intensidades en intervalos cortos de tiempo. Las lagunas existentes van a impedir que buena parte de los años sean incluidos en las series de valores extremos para el cálculo, por ejemplo, de periodos de retorno de intensidades máximas de precipitación (curvas de intensidad-duración-frecuencia). Solo podrá hacerse un uso estadístico de los resultados, tomados como un muestreo de las precipitaciones que, por las razones explicadas, estará sesgado hacia los valores pequeños de precipitación e intensidad y obligará a introducir algún tipo de corrección que asegure la homogeneidad de la serie.

No están documentados con precisión los cambios en el modelo de pluviógrafo, lo que impide localizar claramente los saltos que estos cambios pudieran haber introducido en las series. A comienzos de la década de 1960 se encuentran bandas pertenecientes a un pluviógrafo Fuess que presenta algunos problemas de funcionamiento. En mayo de 1962 comienza a utilizarse, en paralelo con el anterior, un pluviógrafo Lambretch que parece funcionar correctamente, pero del que faltan ocasionalmente bandas, lo que obliga a recurrir a registros anómalos del Fuess. A partir de 1965 se utiliza ya exclusivamente el pluviógrafo Lambretch, que más adelante será sustituido por un pluviógrafo Thies sin que se haya podido determinar el momento exacto.

La tabla 1 muestra para el periodo 1961-1984 el coeficiente de correlación entre la precipitación total diaria obtenida de las bandas y la que consta en el banco de datos. Se han tenido en cuenta únicamente aquellos días en que se registró precipitación. A grandes rasgos, los años con peores resultados en la figura 2 son también los años con peores resultados en esta tabla. Sin embargo, apenas hay casos en los que sea necesario descartar por completo una banda y en el periodo estudiado los datos disponibles cubren un 98,5 % de las horas, teniendo en cuenta los fallos parciales en el registro del pluviógrafo. Dado que el número medio anual de días con precipitación apreciable en el observatorio de A Coruña es de 172, las lagunas en los registros de pluviógrafo representan realmente alrededor del 3 % del tiempo en que se produjeron precipitaciones. Como casos extremos entre los años analizados hasta el momento, se tiene el de 1977, en el que no hay ninguna laguna, y el año 1974, en el que el 3,5 % de los datos horarios faltan.

El propio proceso de digitalización lleva consigo una dificultad, aunque de mucha menor relevancia que la anterior. Al marcar los puntos sobre el gráfico existe siempre una cierta imprecisión. Incluso en un trazo recto de la plumilla, se marcan puntos aleatoriamente un poco por encima o por debajo de línea. Cuando estas oscilaciones son inferiores a una décima, se asigna a ese punto el valor anterior asumiendo que el tramo es recto. Sin embargo, podrían despreciarse así valores que corresponden a incrementos positivos reales, que pueden ser sucesivamente inferiores a una décima y que, acumulados, representan una cantidad relevante de lluvia. La mejor solución parece estar en no tratar de representar valores inferiores a una décima de milímetro, de modo que cualquier incremento (o disminución) en el valor de la ordenada inferior a esa cantidad pueda ser atribuido sin duda a un mal posicionamiento del cursor y eliminado. Este criterio supone no contabilizar las lloviznas persistentes recogidas por las bandas y debería aplicarse

o no en función del objetivo de la grabación. Si se quiere reproducir fielmente el régimen de precipitación de la estación convendría ser más cuidadoso. Si hay interés únicamente en los valores extremos de lluvia, que suele ser lo más común, puede aplicarse este criterio sin alterar significativamente los resultados finales.

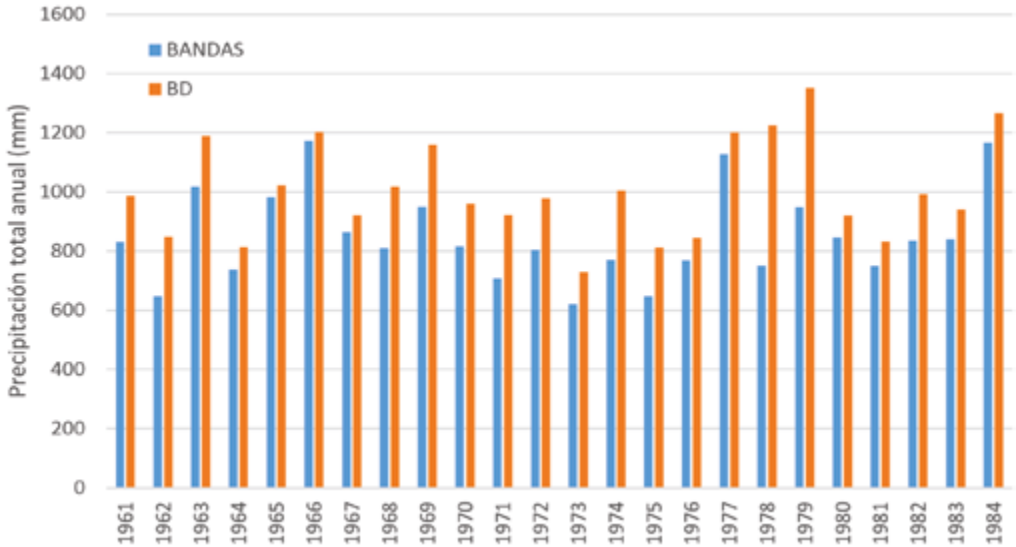


Figura 2. Comparación de los valores de precipitación total anual obtenidos de las bandas y del banco de datos climatológicos de AEMET (BD).

Año	Correlación precipitación > 0	Año	Correlación precipitación > 0	Año	Correlación precipitación > 0
1961	0,93	1969	0,90	1977	0,98
1962	0,82	1970	0,94	1978	0,84
1963	0,88	1971	0,83	1979	0,82
1964	0,90	1972	0,91	1980	0,97
1965	0,99	1973	0,87	1981	0,98
1966	0,99	1974	0,93	1982	0,92
1967	0,98	1975	0,80	1983	0,93
1968	0,75	1976	0,89	1984	0,98

Tabla 1. Valores del coeficiente de correlación entre los datos diarios de precipitación obtenidos de la digitalización de las bandas y los registrados en el banco de datos climatológicos de AEMET.

Las imprecisiones al marcar puntos sobre la banda no solo afectan a las cantidades de lluvia, también al tiempo. Si allí el margen de error se podía acotar en una décima, aquí la experiencia sugiere, de forma bastante subjetiva, estimarlo en cinco minutos. Esa imprecisión no se debe únicamente a errores en la selección de puntos. Las bandas de A Coruña estaban encuadradas y se decidió no desencuadrarlas para no dañarlas. En lugar de escanearlas, se fotografiaron. Las imágenes obtenidas así no son perfectamente planas y puntos separados por igual intervalo temporal sobre la banda pueden estar a distancias ligeramente distintas sobre la imagen. Aunque las diferencias son muy pequeñas, su acumulación en diferentes zonas de la banda provoca que ciertos puntos relevantes, como los de inicio de una descarga, estén en la reproducción en R ligeramente desplazados de su posición real. Este defecto, que puede apreciarse observando



atentamente la figura 1, no afecta a la cantidad de precipitación total ni a las máximas en intervalos largos de tiempo, donde pequeñas oscilaciones en los márgenes del intervalo no suelen tener relevancia. Por fortuna, tampoco es importante en los periodos cortos de tiempo, pues la diferencia entre los intervalos temporales reales y los reproducidos es insignificante y solo su acumulación a lo largo de horas resulta apreciable. Se trata de errores, que, según lo dicho, pueden alcanzar los cinco minutos a la hora de fijar el instante de un cierto evento, por ejemplo el momento de la máxima intensidad diaria, pero que tienen poca o ninguna repercusión al calcular esa intensidad.

Otro problema aparece en las zonas de solape de las bandas. Rara vez coincide el momento en que se retira una de las bandas con el momento en que se coloca la siguiente. Si bien las diferencias normalmente no van más allá de algunos minutos, hay que rellenar los huecos resultantes para evitar que el número de días con registro incompleto se multiplique. El criterio que se ha seguido es el de que, si hay precipitación al final de una banda y al comienzo de la siguiente, se extienden los valores minutales de ambas hacia adelante y hacia atrás repartiendo equitativamente la laguna. Si alguna de las dos bandas no cuenta con precipitación, el hueco se rellena con ceros.

#### 4. ALGUNOS RESULTADOS

En este apartado se va a presentar algún resultado sobre las características de la precipitación en A Coruña en el que se ha podido utilizar una serie relativamente larga de precipitación en intervalos horarios y diezminutales gracias a la digitalización de varias décadas de bandas de pluviógrafo. Para el análisis de la duración de los episodios de precipitación se van a considerar únicamente valores horarios de lluvia, añadiendo a los ya disponibles en el periodo posterior a 1984 los que resultan de la digitalización de las bandas hasta 1961. Los intervalos de duración menor presentan algunas dificultades derivadas en parte de los problemas que se han mencionado en el apartado anterior y, especialmente, de los distintos tipos de pluviógrafo utilizados, desde pluviógrafos de sifón hasta pluviógrafos de balancín más recientemente en los equipos automáticos. La banda del primero de ellos facilita un registro continuo, mientras que las estaciones automáticas proporcionan realmente un muestreo diezminutal. Otra importante diferencia entre ambos métodos se produce durante las lluvias continuas de baja intensidad, que la banda recoge y que en este cómputo aparecen como periodos, a veces prolongados, de precipitación, mientras que la estación automática no la registra hasta que no se completa una décima en el depósito del balancín. Como consecuencia de ello, lo que en la banda es un largo periodo de precipitación, en la medida automática es una sucesión de cortos intervalos secos y lluviosos. Para tratar de salvar estas limitaciones, se considerará un evento de lluvia aquel en el que hubo precipitación en algún momento dentro de cada intervalo horario considerado. Se analizará la duración de los intervalos secos y lluviosos y su evolución a lo largo de los años, no tanto por el interés climatológico de estos valores como para poner de manifiesto las dificultades de mezclar datos procedentes de diferentes equipos y periodos. Por último se estudiará la relación entre la precipitación máxima en intervalos de 10 minutos reales y la máxima en intervalos fijos de 10 minutos. Es este el periodo que se utiliza para los cálculos de intensidades máximas de precipitación y los datos que se manejan se han obtenido, en función del equipo y de la época, con uno o con otro criterio. Rara vez se dispone de ambos datos a la vez y la digitalización de las bandas permite hallar alguna relación estadística entre ambos.

##### 4.1. Duración de las precipitaciones

Para el análisis de la duración de la precipitación, teniendo en cuenta la resolución horaria con la que se está trabajando, se contabilizará como hora con precipitación aquella en la que se haya

producido alguna, independientemente de su duración real. Con este criterio, la duración media de un episodio de lluvia es de 2,6 horas, con la distribución mensual que se aprecia en la parte superior de la figura 3 y que no muestra una gran diferencia entre los distintos meses. La duración máxima media corresponde a diciembre y enero, con cerca de 3 horas de lluvia consecutivas, mientras que en julio y agosto la duración media no llega a las dos horas y media. La mayor longitud de un periodo con lluvia tuvo lugar el 31 de marzo de 2016 y fue de 37 horas ininterrumpidas de precipitación. No se observa una tendencia clara en cuanto a la duración media de las precipitaciones (figura 4) y los valores relativamente bajos de los primeros años de la serie pueden deberse, como ejemplo de las limitaciones de la recuperación de datos, a la interrupción de los periodos más largos de lluvia por los fallos del pluviógrafo que se han mencionado en el apartado 3.

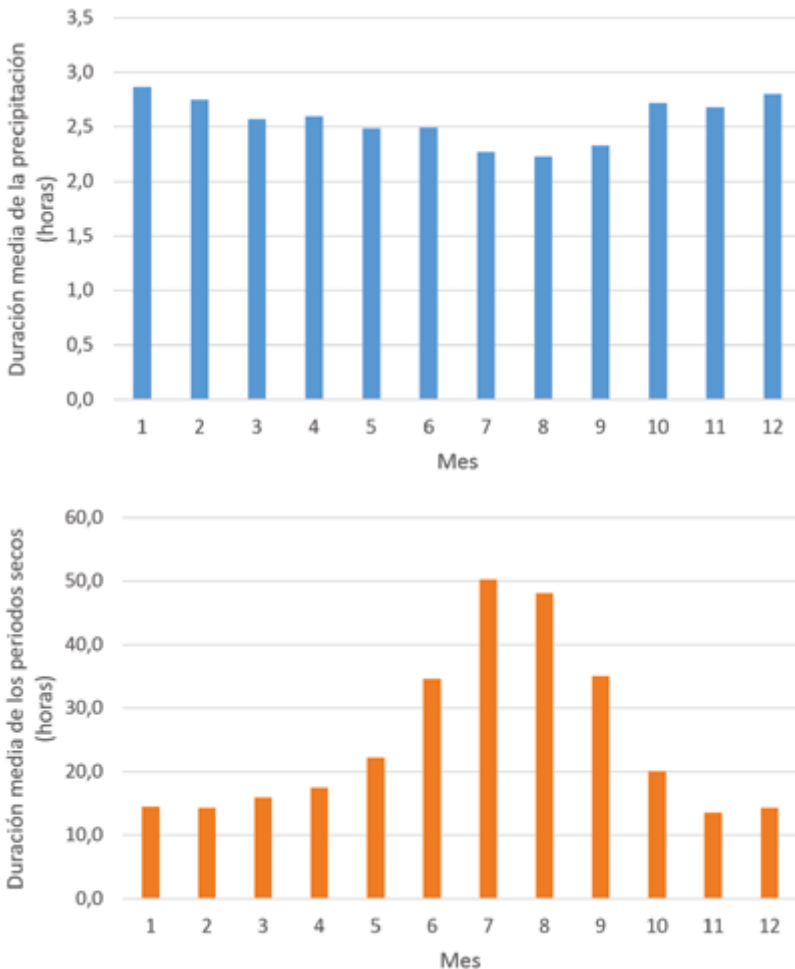


Figura 3. Promedio mensual de la duración de los episodios con precipitación y de los periodos secos (1961-2020).

Por lo que respecta a los periodos sin lluvia, su duración media es de 20,7 horas, con un máximo de 983 horas que terminó el 8 de abril de 1997. Como es de esperar, los valores más largos de periodos sin precipitación corresponden al verano (figura 3), con una duración media

de 50 horas en julio y los más cortos al mes de noviembre, que no alcanzan las 14 horas de media. En cuanto a la evolución a lo largo del periodo (figura 4), se observa una leve tendencia decreciente que probablemente tampoco es real y tiene más que ver con los cambios en los equipos y procedimientos de medida que con variaciones reales de la duración.

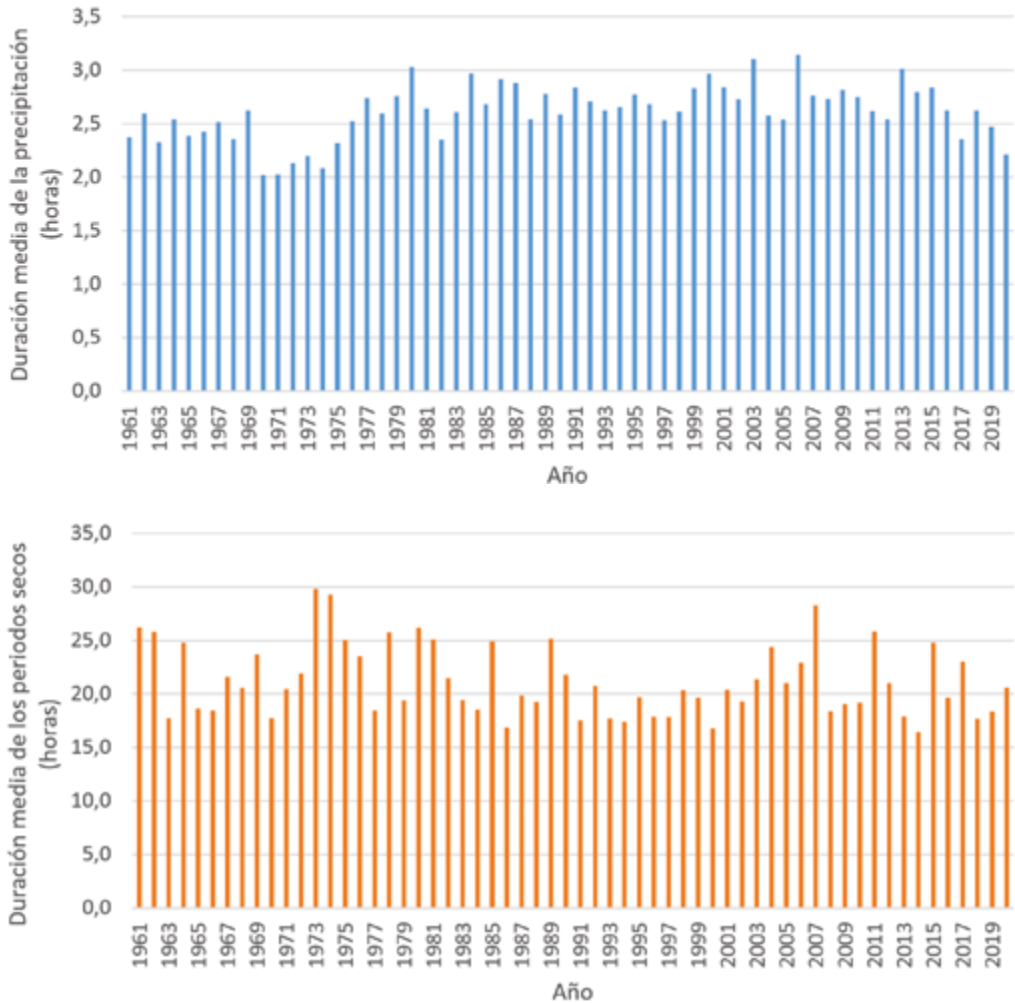


Figura 4. Promedio anual de la duración de los episodios con precipitación y de los periodos secos.

#### 4.2. Valores máximos en intervalos fijos y móviles de diez minutos

Las estimaciones de periodos de retorno de intensidades máximas de precipitación realizados a partir de datos recogidos en intervalos fijos de diez minutos (de 00:00 a 00:10, de 00:10 a 00:20...) subestiman el verdadero valor, que debería obtenerse con intervalos móviles de 10 minutos. Distintos equipos e incluso la misma técnica de medida aplicada en periodos distintos, proporcionan uno u otro valor, de forma que en la misma serie de datos es posible encontrar registros obtenidos con uno u otro criterio. Cuando se facilita a un usuario el dato de intensidad registrado



en una fecha determinada se puede estar subestimando el valor real si se ha medido en intervalos fijos. Puede ocurrir también que se facilite una intensidad real, pero que para evaluarla haya que recurrir a frecuencias obtenidas a partir de intervalos regulares, con lo que se está sobreestimando su rareza.

Es un problema que afecta no solo a intervalos de corta duración, también a precipitaciones en intervalos de horas o días. Hershfield estableció una relación de 1,13 entre los valores en intervalos fijos y móviles asociados a un periodo de retorno dado para cantidades máximas en 2 y 24 horas (HERSHFIELD, 1961). La digitalización de las bandas permite disponer de medidas en intervalos de tiempo más cortos para estudiar si el coeficiente de Hershfield es también aplicable a este caso. Los resultados se muestran en la tabla 2.

Periodo de retorno (años)	Intervalo fijo	Intervalo móvil	Relación
2	33,2	37,5	1,13
5	44,7	51,3	1,15
10	52,4	60,4	1,15
15	56,7	65,6	1,16
25	62,0	72,0	1,16
50	69,2	80,5	1,16
75	73,4	85,5	1,17
100	76,3	89,0	1,17
200	83,4	97,5	1,17
500	92,7	108,7	1,17

*Tabla 2. Valores de intensidad de precipitación (mm/h) correspondientes a diversos periodos de retorno obtenidos por ajuste a una distribución de Gumbel de los máximos anuales en 10 minutos medidos en periodos fijos y móviles (1961-1984).*

Para el periodo de digitalización de las bandas (1961-1984) se hallaron los máximos anuales en 10 minutos por el método de los intervalos fijos y por el método de los intervalos móviles. A partir de estos datos se calculó la intensidad máxima en mm/h y, mediante un ajuste a la distribución de Gumbel, los valores asociados a diversos periodos de retorno. Puesto que, como ya se ha señalado, algunos años se encuentran muy incompletos, los valores obtenidos no pueden tomarse como representativos de la estación de A Coruña, pero eso no impide que se pueda estudiar la relación entre los resultados obtenidos con ambos criterios. La primera conclusión es que existe una notable diferencia entre los valores que resultan de ambos métodos de medida, de modo que es un problema que debe ser tenido en cuenta. La relación entre los valores oscila entre 1,13 y 1,17 y crece rápidamente para valores bajos de periodo de retorno para tender después asintóticamente a 1,17. Solo el periodo de retorno de 2 años reproduce el valor de 1,13 de Hershfield, aunque la desviación no supera el 4 % en ningún caso. Hay que tener en cuenta que los valores de la tabla 2 son el resultado de un ajuste de los datos a una distribución teórica. Si se calcula la relación entre máximos año a año y se promedian los valores resultantes, el resultado es justamente 1,133 y resulta levemente superior, 1,136, si se ajustan los valores a una recta de regresión que pase por el origen.

## 5. CONCLUSIONES

El método propuesto para la digitalización de las bandas de pluviógrafo que se ha aplicado en la estación de A Coruña entre 1961 y 1984 es relativamente rápido y resulta fiable para almacenar en soporte informático una información sobre pluviometría que resultaría muy útil para el estudio,

especialmente, de eventos extremos. Se ha podido comprobar que, cuando la calidad de las bandas es adecuada, los valores obtenidos se asemejan mucho a los medidos con pluviómetro. Sin embargo, no siempre se cumple esa condición y de este modo se limitan las posibilidades de obtener unos registros fiables de forma continua sobre la distribución de la precipitación. Ello no supone que la información conseguida no resulte útil. Al contrario, convenientemente tratada para obtener una serie homogénea de datos, completa los valores medidos en épocas recientes en cortos intervalos de tiempo y contribuye a conocer con más precisión el régimen de precipitaciones y su evolución temporal.

## REFERENCIAS

- BONI, G, PARODI, A. y RUDARI, R., 2006. Extreme rainfall events: learning from raingauge time series. *J. Hydrol.*, 327(3-4): 304-314.
- WILKINSON, C., BRÖNNIMANN, S., JOURDAIN, S., ROUCAUTE, E., CROUTHAMEL, R., BROHAN, P., VALENTE, A., BRUGNARA, Y., BRUNET, M., COMPO, G. y GILABERT, A., 2019. Best Practice Guidelines for Climate Data Rescue. Copernicus Climate Change Service.
- AEMET, 2022. Plan Nacional de Predicción y Vigilancia de Fenómenos Adversos. [https://www.AEMET.es/documentos/es/eltiempo/prediccion/avisos/plan\\_meteoalerta/plan\\_meteoalerta.pdf](https://www.AEMET.es/documentos/es/eltiempo/prediccion/avisos/plan_meteoalerta/plan_meteoalerta.pdf).
- JAKLIČ, A., ŠAJN, L., DERGANČ, G. y PEER, P., 2016. Automatic digitization of pluviograph strip charts. *Met. Apps.*, 23: 57-64. <https://doi.org/10.1002/met.1522>.
- SUŠIN, N. y PEER, P., 2018. Open-source tool for interactive digitisation of pluviograph strip charts. *Weather*, 73: 222-226. <https://doi.org/10.1002/wea.3001>.
- HERSHFIELD, D. M., 1961. Rainfall Frequency Atlas of the United States, Technical Paper No. 40. Weather Bureau, US Department of Commerce, Washington, DC.