

# La cocina de los datos meteorológicos

**Domingo F. Rasilla Álvarez**

*Profesor Titular de Geografía Física. Dpto. de Geografía, Urbanismo y Ordenación del Territorio. Universidad de Cantabria*

Cuando la redacción de la revista Índice me pidió una colaboración, afronté cierto miedo escénico ante el reto de hacer accesible a profesionales de las Ciencias Estadísticas, con un fuerte bagaje matemático, los entresijos de una disciplina como la Climatología, en la que, amén de una fuerte componente estadística, confluyen aspectos de las denominadas Ciencias Naturales y Sociales, y sometida a intereses y polémicas que desbordan el mero conocimiento científico. Afortunadamente, las musas vinieron en mi auxilio, recordándome los respingos con los que afrontamos, con ocasión de algún fenómeno meteorológico señalado, las preguntas de los “plumillas” (particularmente los medios de comunicación regionales). Muchos compañeros convendrán conmigo que frecuentemente han sido interpelados acerca de si “¿es verdad que anteayer fue el día más...en los últimos X años?, ¿es eso normal?, ¿qué relación tiene con el cambio climático?...”. Los lectores comprenderán que estamos abocados a simplificar una respuesta que desgraciadamente, oculta al público un proceso previo de depuración y tratamiento de los datos, poco lucido y en muchos casos tedioso, pero esencial. Sirvan estas líneas para sacar a la luz todo aquello que nos gustaría contar, pero que rara vez ocurre fuera de los circuitos científicos.

**V**olviendo a nuestro interlocutor, la respuesta debería comenzar con el encuadre del fenómeno meteorológico en cuestión dentro de la evolución climática de la región en la que se inserta. Su excepcionalidad podría establecerse, bien por la magnitud alcanzada por una variable meteorológica que se inserta en un fenómeno bien conocido en esa región (p.e. las temperaturas durante la ola de calor de 2003), bien por ser el resultado de un proceso atmosférico ajeno a su dinámica habitual (p.e. la llegada del huracán Vince al Sur de España en 2005). Si en el primer caso no es difícil calcular su periodo de retorno a través de alguno de los conocidos procedimientos estadísticos sobre análisis

de extremos, en el segundo caso debería recurrirse a un análisis algo más complejo. En ambos casos, un requisito previo esencial es disponer de series meteorológicas extendidas en el tiempo, con una resolución temporal adecuada y de calidad.

Para obtener estas series, deberemos comenzar recopilando esa información meteorológica en forma del mayor número de series disponibles en nuestro ámbito de trabajo. Llegados a este punto, deberíamos alertar a nuestro interlocutor de que el análisis de una única serie podría conducirnos a afirmaciones erróneas, mientras que la detección de un mismo fenómeno o tendencia simultáneamente en varios observatorios



refuerza la validez de nuestros análisis. Hasta hace aproximadamente un par de décadas, la mayor parte de la información de carácter meteorológico procedía del antiguo Instituto Nacional de Meteorología (INM), hoy Agencia Estatal de Meteorología, sostenida por los presupuestos del estado. Por ley, centralizaba toda la información, tanto generada por ella como por otras instituciones (p.e. Cuencas Hidrográficas). Durante el último decenio, el acceso a dicha información estuvo limitado por una suerte de “copago”, es decir, no era gratuita. La conocida capacidad de improvisación de los investigadores españoles solventó este obstáculo recurriendo en unos casos al intercambio de datos entre colegas, o en otros a la búsqueda de fuentes alternativas internacionales, como el National Climatic Data Center (NCDC) de EE.UU., United Kingdom Meteorological Office (UKMO) de Reino Unido, etc. Afortunadamente, la política de AEMET ha cambiado en los dos últimos años, permitiendo el acceso gratuito a numerosos datos, aunque nunca se sabe... A partir de la década de los noventa han proliferado otras redes meteorológicas a cargo de organismos públicos (Comunidades Autónomas, Ayuntamientos, etc.), instituciones privadas o incluso particulares. Aunque han venido a cubrir la carencia de información en ámbitos geográficos de gran interés ambiental, por ejemplo las áreas de montaña, y ofrecen un elevado grado de accesibilidad, la falta de una política global de estandarización y calibración del equipamiento hace difícil su comparación con los datos de AEMET.

La gran mayoría de los datos meteorológicos, actualmente en soporte digital, no suelen remontarse más allá de comienzos del s. XX, a diferencia de otros países, más cuidadosos con la información estadística en general, y la meteorológica en particular. No obstante, existen datos antiguos, incluso de mediados del s. XVIII, que proceden de particulares, muchas veces profesionales liberales (p.e. médicos) movidos por una curiosidad intelectual notable. Estas series sólo son accesibles a través de una labor de recopilación bibliográfica en archivos y bibliotecas, pero su explotación es aún problemática a causa de la pobreza de las fuentes de información (coloquialmente conocidas como “metadatos”) que describen su emplazamiento, sus unidades de medida (“pies castellanos”, pulgadas, grados Reamur, etc.), instrumental, etc. Para remontarse aún más atrás en el tiempo también se recurre a fuentes de información alternativas (“paraclimáticas”), entre las que destacaremos los anillos de los árboles (“Dendroclimatología”), documentos públicos y privados, como las Rogativas Eclesiásticas, etc. (“Climatología Histórica”). Estas fuentes exigen un proceso de filtrado y depuración de información no climática muy estricto, que podría hacer pensar al profano que la información extraída es poco representativa. No obstante, esta “arqueología de datos”, objeto de numerosos proyectos de investigación en el último decenio, es esencial, puesto que, gracias a ellos podremos cuantificar el estado del Sistema Climático en un momento en el que el impacto de las actividades humanas habría sido mínimo.

Si para nuestro estudio hemos decidido limitarnos al análisis de las series meteorológicas que llamaremos “oficiales” (p.e. las procedentes de AEMET), no crea el lector que se han acabado los problemas. En primer lugar, incluso las series oficiales muestran valores “sospechosos”, resultado del proceso de medición, manipulación, formateado, transmisión y archivo de los datos, y cuya validez deberá ser contrastada. Por otro lado, dichas series no siempre son continuas, sino que presentan “lagunas”, es decir, periodos más o menos largos en los que no se hicieron mediciones, o éstas se han perdido. Las causas son múltiples, pero en su mayoría responden a los periodos de inestabilidad social o política, como la Guerra Civil española o la Segunda Guerra Mundial en Europa, o al fallecimiento de la persona que estaba a cargo de la estación. Para la detección de casos sospechosos y en su caso eliminación, al igual que para el relleno de las series originales, se han diseñado diferentes procedimientos. Éstos están basados en su comparación con observaciones meteorológicas realizadas en ámbitos próximos (aunque a medida que nos remontamos en el tiempo el radio de búsqueda se amplía lo cual forzosamente reduce su fiabilidad), o, alternativamente, en la búsqueda de información sobre sus efectos (p.e. hemerotecas). Si los valores sospechosos son considerados reales, deben ser mantenidos en las series pero tratados estadísticamente como valores extremos (“outliers”) y sometidos a las técnicas adecuadas para el tratamiento de estos fenómenos.

“ **La gran mayoría de los datos meteorológicos, actualmente en soporte digital, no suelen remontarse más allá de comienzos del s. XX** ”

El lector podrá imaginarse que llegados a este punto, la resistencia del climatólogo es digna de encomio, y en algunos casos, justificaría un rictus irónico ante el entrevistador. Otra fuente adicional de potenciales errores son los cambios sufridos por el emplazamiento de las estaciones. La mayor parte de las observaciones meteorológicas españolas comenzaron allá por mediados del s.XIX o comienzos del XX, a cargo de catedráticos de Universidad o Instituto de Ense-

ñanzas Medias, en el edificio de la propia institución, sita en el casco antiguo. En la década de los años 20-30 algunos se trasladaron a edificios creados ex profeso, en barrios conocidos como “Ciudad-Jardín”, entonces en la periferia de las ciudades. Pero el gran desarrollo urbano acaecido entre los años 50 y 70 los engulló literalmente, de tal suerte que algunos han vuelto a cambiar de emplazamiento en los últimos 20 años. Estos desplazamientos producen discontinuidades que no responden a un factor climático, pero que inadvertidas dan lugar a variopintas interpretaciones. Las series meteorológicas que no han sufrido cambios de emplazamiento están afectadas, sin embargo, por un fenómeno adicional, el crecimiento de las ciudades a su alrededor (p.e. El Retiro en Madrid). La ciudad genera un microclima característico, resultado de la alteración de los balances energéticos que tiene su mayor impacto en las temperaturas, sobre todo las mínimas nocturnas durante noches despejadas (lo que se conoce como “isla de calor urbana”) y en el viento, notablemente ralentizado. Diferencias de 6°C durante las noches invernales, y reducciones del viento de hasta 50 km/h no son desconocidas, y su persistencia en el tiempo provoca también tendencias anómalas en las series meteorológicas. Esta alteración se detecta incluso en aeropuertos, fuera de las ciudades, dado que en las últimas décadas han sido rodeados de edificaciones que también perturban la fiabilidad de algunos parámetros, como el viento. Para no aburrir más al lector, simplemente señalaré que a todo lo anteriormente expuesto, habría que añadir también como fuentes de incertidumbre los cambios en los instrumentos de medida (anemómetros, garita meteorológica), las unidades de medida (mm y hPa), así como las horas de observación.

Pese a todos estos inconvenientes, existen procedimientos muy robustos tanto para el control de calidad como para la identificación y corrección de potenciales discontinuidades. Los primeros detectan y etiquetan errores obvios (datos anómalos –p.e. temperatura > 100°C–, desplazamientos de coma, repetición del mismo valor varias veces consecutivas, valores negativos de lluvia, errores en las unidades de medida –el viento se mide en m/s en los mensajes meteorológicos, km/h o nudos en los aeronáuticos–), establecen los umbrales e identifican los casos fuera de rango (test de tolerancia) o verifican la consistencia interna de cada parámetro (p.e. que la temperatura mínima no sea superior a la máxima), etc. Tras ellos, se aplican además diferentes test de homogeneidad interna cuya finalidad es la detección y corrección de discontinuidades en las series.

Para concluir, si hemos realizado bien nuestra tarea, nuestras series habrán sido depuradas de toda aquella información potencialmente perturbadora y podremos contestar a nuestro interlocutor, tras analizar si el fenómeno descrito se enmarca o no con una tendencia conocida, si responde a una situación meteorológica típica de nuestra región o si la magnitud alcanzada por un parámetro presenta un periodo de recurrencia elevado, o bien al contrario, si nuestra memoria meteorológica no sólo es corta, sino también selectiva. De hecho, even-

## “ Pese a todos estos inconvenientes, existen procedimientos muy robustos tanto para el control de calidad como para la identificación y corrección de potenciales discontinuidades ”

tos tachados de excepcionales hoy no lo son tanto si echamos la mirada atrás y analizamos el mero fenómeno meteorológico; cosa bien distinta es la mayor vulnerabilidad de la sociedad actual (tenemos más que perder porque tenemos más), a pesar de nuestra sofisticación tecnológica. Además, nuestra capacidad de adaptación ha disminuido, no por que seamos menos resistentes (la estatura de las nuevas generaciones indica una mejora sustancial), sino porque no toleramos que la naturaleza se cruce en nuestro camino y nos altere un programa de actividades diseñado al milímetro y poco dado a las improvisaciones. De ahí que el estudio del Cambio Climático y su impacto tanto en el medio natural como en las actividades humanas requiere no sólo que la comunidad científica trabaje con el máximo rigor posible, sino también que sea capaz de trasladar al gran público todo lo que se esconde detrás de un párrafo o una simple cifra.

### Para saber más...

- COST ACTION ES0601-Advances in homogenisation methods of climate series: an integrated approach HOME  
<http://www.homogenisation.org/>.
- The MEditerranean climate DATA REscue:  
<http://www.omm.urv.cat/MEDARE/index.html>
- Proyecto Salvá-Sinobas:  
<http://salva-sinobas.uvigo.es/>
- RODRÍGUEZ, R.; BARRIENDOS, M.; JONES, P.D.; MARTÍN-VIDE, J. y PEÑA, J.C. (2001): “Long pressure series for Barcelona (Spain). Daily reconstruction and monthly homogenization”. *International Journal of Climatology*, 21, 1693-1704. Disponible en:  
[http://www.ub.edu/gc/Documentos/IJC\\_Roberto\\_etal.pdf](http://www.ub.edu/gc/Documentos/IJC_Roberto_etal.pdf)