

# CARACTERIZACIÓN DE SONDEOS ATMOSFÉRICOS ASOCIADOS A LA OCURRENCIA DE TORMENTAS EN LA CIUDAD DE MURCIA

Victoria GALLARDO FERNÁNDEZ<sup>1</sup>, Juan Andrés GARCÍA VALERO<sup>2</sup>,  
Carlos LIZONDO OSSET<sup>3</sup>

<sup>1</sup>*Universidad de Murcia.*

<sup>2,3</sup>*Centro meteorológico de Guadalupe. AEMET.*

**victoria.gallardo@um.es, jugarciav@aemet.es, clizoss@gmail.com**

## RESUMEN

En este trabajo se presenta un estudio de caracterización de sondeos atmosféricos de 333 situaciones de tormentas observadas en la ciudad de Murcia en el período 1984-2007. Para ello, se ha empleado datos relativos a gradientes de temperatura entre diferentes niveles estándar de presión, así como de humedad relativa derivados de los sondeos realizados desde el observatorio de AEMET en Guadalupe (Murcia). La caracterización ha consistido en una agrupación tipo K-medias de aquellos sondeos con un perfil similar en cuanto a las variables mencionadas. Para facilitar esta tarea, se aplicó previamente un Análisis de Componentes Principales a las variables de entrada al método de agrupación. De este modo, se ha obtenido 11 perfiles verticales que explican aproximadamente el 70% de la variabilidad total.

Los resultados muestran la existencia de diferencias importantes entre los distintos tipos de tormentas, así como entre éstos y los sondeos en los que no hubo tormenta.

**Palabras clave:** agrupación, sondeo meteorológico, tormenta, gradiente de temperatura, humedad relativa.

## ABSTRACT

In this work, we present a characterization study of atmospheric soundings of 333 situations of storms observed in the city of Murcia in the period 1984-2007. For this, we've treated temperature gradients between different standard pressure levels, as well as relative humidity data recovered from meteorological soundings at the AEMET observatory in Guadalupe (Murcia). The characterization has consisted of an application of K-means clustering algorithm on those soundings with a similar profile in terms of the mentioned variables. To make it easier, a Principal Component Analysis was previously applied to the input variables. Thereby we've obtained 11 vertical profiles which explain around 70% of the total variability.

The results show important inner differences between the types of storms, as well as between them and the soundings in which there was no storm.

**Key words:** clustering, meteorological sounding, storm, temperature gradient, relative humidity.

## 1. INTRODUCCIÓN

Las tormentas son el resultado de un fuerte desequilibrio hidrostático en la vertical atmosférica. Las grandes repercusiones en numerosos sectores socioeconómicos a consecuencia de las intensas precipitaciones, rachas de viento y granizadas que suelen acompañarlas las convierte en objeto de gran interés. La mitad oriental de la Península y, en especial, el Levante español tienen un régimen muy irregular de precipitaciones, muchas de ellas de carácter convectivo. Los episodios de lluvias torrenciales, propios de otoño en el área mediterránea, están asociados a régimen de levante en superficie y a depresiones aisladas en niveles altos (DANAs), que dan lugar a fuertes tormentas. La mayoría de trabajos encaminados al estudio de tormentas realizados sobre la Península Ibérica se basa en la caracterización de entornos sinópticos favorables a su desarrollo, como es el caso de *Climatología de tormentas veraniegas en Cantabria, Asturias y País Vasco* (Sanz Moral et al., 2004), pero apenas hay estudios que aborden su caracterización desde un análisis profundo de su perfil vertical termodinámico.

En este trabajo se propone una metodología para la clasificación y caracterización de sondeos asociados a situaciones de tormenta, fundamentada en los algoritmos de agrupamiento y en árboles de decisión, como paso previo a la creación de un modelo de predicción de este fenómeno. Es por ello que, aunque el estudio se ha realizado a nivel local, lo aquí expuesto es extrapolable a otras regiones.

Este documento consta de tres secciones aparte de la introducción: en la sección 2 se especifica los datos utilizados y su fuente, y se explica el procedimiento seguido en el análisis, la tercera sección contiene los resultados obtenidos, y en la última se discute la posible implementación de este método al pronóstico de tormentas, así como las limitaciones del estudio y las curiosidades que suscita como posibles líneas de investigación futuras.

## 2. MÉTODOS

El estudio se ha llevado a cabo sobre los datos de sondeos meteorológicos de las 00 y las 12 UTC en la ciudad de Murcia, recabados en el observatorio de Guadalupe (AEMET) y descargados desde la página web de la Administración Nacional Oceánica y Atmosférica (NOAA) <https://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive>. El período estudiado de ocurrencia de tormenta comprende desde el mes de agosto de 1984 hasta octubre de 2007.

Se ha realizado una clasificación de los sondeos de tormenta, para lo cual, dadas las fechas de los días en que se dio este fenómeno y teniendo en cuenta la hora a la que tuvo lugar según los libros de observaciones, se ha retenido el sondeo más cercano al suceso, sin distinguir si fue previo o posterior al mismo.

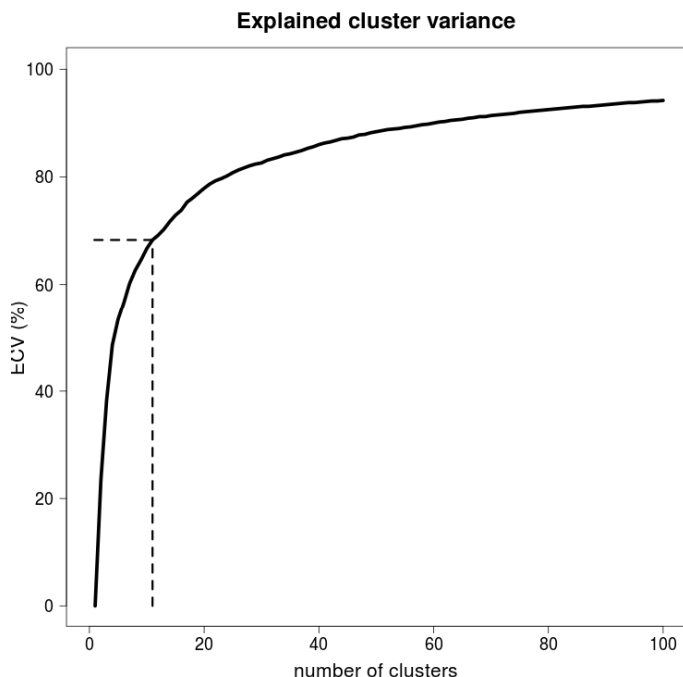
En primer lugar, a partir de las variables temperatura, punto de rocío y presión, se ha calculado las humedades relativas en los niveles de presión estándar de 1000, 850, 700, 500 y 300 hPa, así como por los gradientes de temperatura entre los mismos. De este modo, cada sondeo de tormenta viene caracterizado por 15 variables.

A continuación, se ha hecho un Análisis de las Componentes Principales (ACP) (von Storch & Zwiers, 1999) de los sondeos seleccionados con el fin de eliminar la información redundante. Como resultado del mismo, se ha obtenido que con 5 componentes principales (CP) es posible retener el 90% de la varianza.

Por último, se ha efectuado un agrupamiento tipo K-medias (Wilks, 2011) de los datos descritos en la nueva base de 5 CP. Para determinar el número óptimo de grupos a considerar, se ha analizado la dependencia de la varianza explicada por los grupos, ECV, en función del número de grupos, con

$$ECV = 1 - WSS/TSS,$$

donde WSS es la suma de las varianzas internas de los grupos (Within Sum of Squares) y TSS es la varianza global (Total Sum of Squares). El resultado se muestra en la figura 1. Así, se ha establecido en  $n = 11$  el número adecuado de grupos, teniendo en cuenta la coincidencia de un valor elevado del factor de calidad ECV, de casi el 70%, con un número "manejable" de grupos.



*Fig. 1: Gráfica de la varianza explicada por los grupos en función del número de grupos.*

Fijado el número de grupos o tipos de tormentas (TTs), se ha caracterizado cada uno de ellos. Para esto, se ha hallado la proporción del total de las situaciones de tormenta que reúnen, representado sus sondeos medios y establecido sus histogramas de estacionalidad y las precipitaciones asociadas.

Para profundizar en la diferenciación entre sondeos de tormenta y aquéllos no asociados a este evento, así como en la distinción interna entre grupos, se ha recurrido al algoritmo de árboles de decisión (Hothorn et al., 2006). También se ha usado este método para hallar las variables determinantes en la ocurrencia de cada TT con respecto al resto de situaciones meteorológicas (tanto sondeos en los que hubo

tormentas de otro tipo como sondeos en lo que no hubo tormentas para el período estudiado). Así, se ha entrenado un árbol de decisión a partir de la muestra requerida en cada caso. Se define como variable dominante aquella que prima en la decisión tomada por el algoritmo, y como variable o variables secundarias aquellas que aparecen inmediatamente supeditadas a la primera.

La figura 2 ofrece un ejemplo de estos diagramas, se trata del árbol que compara los sondeos de tipo 1 con todos los demás. En ella puede verse que, bajo la condición de  $HR850 \leq 74\%$ , si esta misma variable en el nivel superior de 700 hPa no sobrepasó el 63%, ninguno de los 18731 sondeos con estas características presentó tormenta de este tipo ( $y = 0$ ). Por otro lado, casi el 14% de los sondeos que presentaron  $HR850 > 74\%$ , más un gradiente de temperatura entre los niveles de 1000 y 700 hPa por encima de casi  $19^\circ\text{C}$ , y en los cuales la humedad relativa en el nivel de 1000 hPa excedió el 63%, pertenecen a dicho grupo de tormentas.

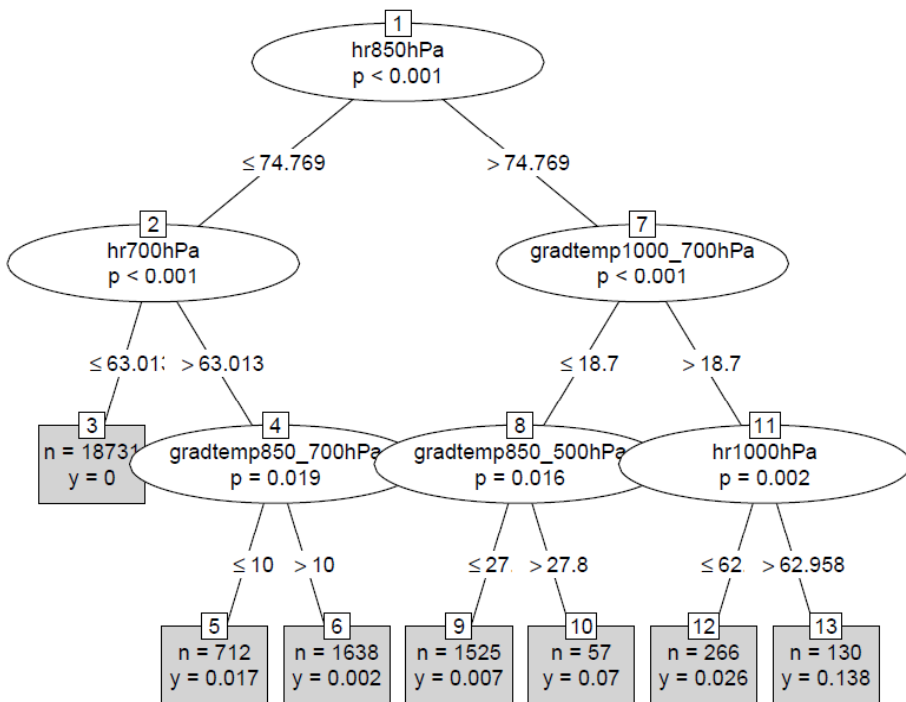


Fig. 2: Ejemplo de árbol de decisión. Compara los sondeos del grupo 1 con el resto de sondeos, los no asociados a tormenta junto con los incluidos en un grupo diferente al primero.

### 3. RESULTADOS

#### 3.1. Sondeos tipo. Estacionalidad, perfiles medios y precipitaciones asociadas

En la tabla 1 se expone el porcentaje representativo de cada grupo sobre el total de las situaciones de tormenta. En ella aparecen los grupos ordenados de mayor a menor repetición. También se informa sobre la estacionalidad de cada uno, siendo el mes principal aquél en el que el tipo en cuestión tuvo mayor ocurrencia, y el mes o los meses secundarios, aquéllos en los que la acumulación de casos superó el 50% de los contabilizados en el primero.

En general, el mes con mayor ocurrencia de tormentas en Murcia es septiembre, seguido de junio, mayo, octubre y agosto.

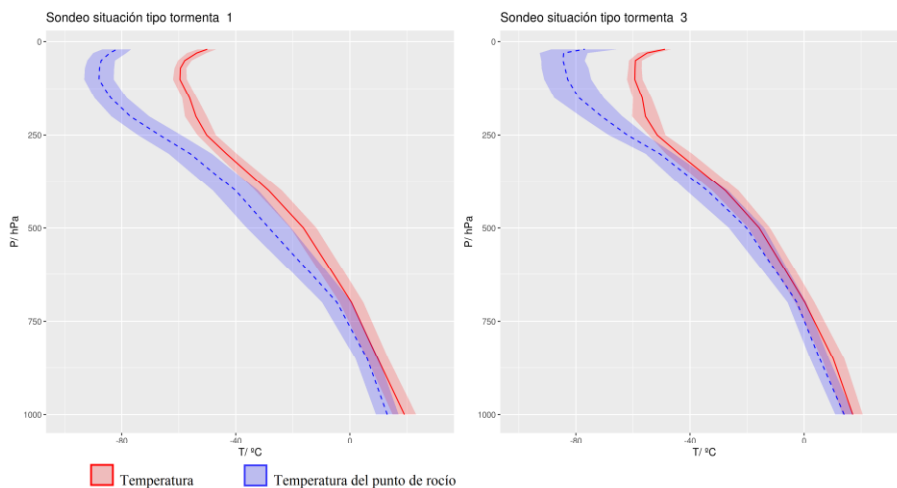
<b>Tipo de tormenta</b>	<b>Frecuencia relativa (en %)</b>	<b>Mes principal</b>	<b>Mes(es) secundario(s)</b>
1	17	octubre	septiembre, mayo
2	15	septiembre	mayo, junio
3	14	octubre	ninguno
4	11	septiembre	junio, octubre
5	10	junio	septiembre
6	8	junio	agosto
7	6	agosto	septiembre
8	6	junio, septiembre	ninguno
9	5	mayo	ninguno
10	4	enero	febrero, abril
11	4	mayo	septiembre

*Tabla 1: Frecuencia relativa y estacionalidad de los distintos tipos de tormenta.*

Por motivo de la limitación en la extensión de este documento, no es posible incluir todas las representaciones gráficas que caracterizan los distintos perfiles de tormenta, ni entrar en un análisis detallado de todos ellos, por lo que sólo se comentará, a modo de ejemplo, un par de estas situaciones tipo. Los TTs escogidos han sido el 1 y el 3, cuyos sondeos medios son presentados en la figura 3. La decisión se ha tomado en base a la conjunción de un gran porcentaje representativo con las mayores precipitaciones asociadas.

En el sondeo de tipo 1 se observa que hay una fuerte mezcla en la capa que va desde los 1000 hasta los 700 hPa, ya que el gradiente de la temperatura es constante. Si se compara con la misma curva de estado del sondeo de tipo 3 a los mismos niveles, puede decirse que el primero es más inestable, pues el gradiente es mayor. De la gráfica también se deduce que el segundo sondeo es más húmedo, dado que la curva de la temperatura del punto de rocío permanece unida a la de la temperatura hasta casi

el nivel de 250 hPa, frente a los 500 hPa del primer ejemplo. No hay diferencias significativas de temperatura por niveles entre las curvas de sendos grupos pues, como se puede ver en la tabla 1, los dos son propios principalmente de octubre.



*Fig. 3: Gráficas de los sondeos característicos de los tipos 1 y 3 de tormenta, a izquierda y derecha, respectivamente. La curva roja representa la temperatura, y la azul es la temperatura del punto de rocío. Las zonas sombreadas en azul y rojo indican la amplitud de la desviación estándar de cada variable.*

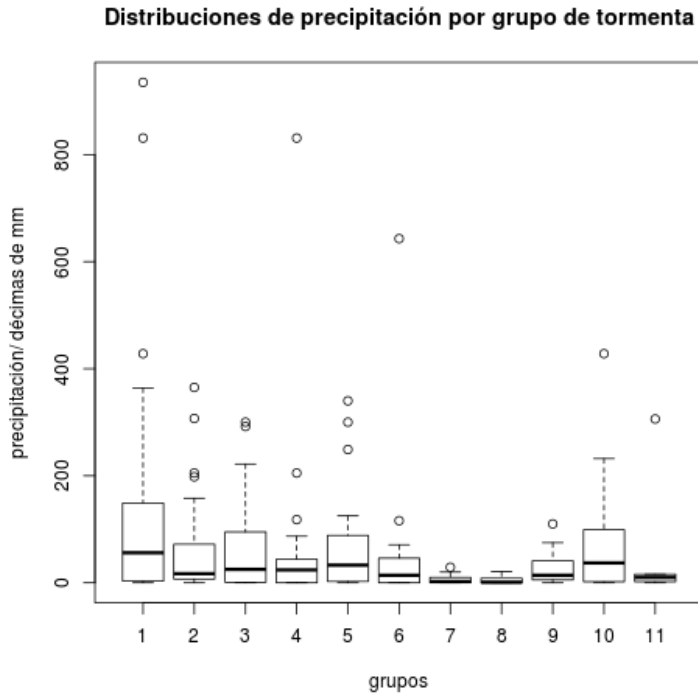
En lo que respecta a la precipitación, tanto el grupo 1 como el 3 la tienen asociada, como ya se ha mencionado, siendo más copiosas en el tipo 1, llegando al extremo de casi 100 mm por suceso (véase figura 4). Sin embargo, la mediana está situada alrededor de los 5 mm en el primer caso, y en los 2.5 mm en el otro. De igual modo, obsérvese que, mientras el tipo de tormenta con mayor precipitación asociada es el 1, y los secundarios son los grupos 10, 5, 3 y 4 (en orden decreciente), las situaciones de los grupos 7, 8 y 11 corresponden a tormentas secas.

### 3.2. Variables determinantes de cada sondeo tipo

La tabla 2 recoge el resultado de entrenar los árboles de decisión que comparan los sondeos de cada TT con el resto de sondeos atmosféricos disponibles, esto es, el total de sondeos con no ocurrencia de tormenta u ocurrencia de otro tipo. Por otro lado, en la tabla 3 se muestra los resultados obtenidos de comparar los sondeos de cada grupo con los de todas las demás situaciones de tormenta.

De contrastar la tabla 2 con la 3 se extrae que, en general, las variables más influyentes en cada TT y las variables que definen sus diferencias internas con el resto de grupos de tormentas, son las mismas. No obstante, éstas no aparecen necesariamente siempre en el mismo orden de prioridad en la decisión, sino que variables dominantes y secundarias pueden encontrarse alternadas según el caso. Cabe destacar que para los

grupos 4 y 5 no se ha hallado coincidencias.



*Fig. 4: Representación tipo boxplot de las distribuciones de la precipitación asociada a los grupos de tormentas.*

Las variables decisivas de la tabla 2 que más se repiten (hasta en cuatro ocasiones cada una de ellas) son la HR1000 y la HR700, y el gradiente de temperatura entre estos dos niveles. Por otro lado, en la tabla 3, las variables más reiteradas (en este caso con tres repeticiones cada una) son tres gradientes: T1000-T850, T700-T500 y T500-T300. Nótese que no hay coincidencias entre estas dos enumeraciones.

Además, del árbol de decisión que enfrenta a todos los sondeos de tormenta con aquéllos en los que no la hubo, se encuentra que la variable que marca la diferencia de forma dominante es la HR700, y la secundaria es el gradiente de temperatura T1000-T500.

#### 4. DISCUSIÓN

De la clasificación y caracterización elaboradas en este trabajo se espera que puedan sentar las bases para la constitución de un nuevo modelo de predicción de tormentas. Una vez entrenados los árboles de decisión, es posible establecer la probabilidad de que un sondeo conocido evolucione a una situación de tormenta, incluso de qué tipo

sería según nuestra clasificación. De hecho, el objeto de haber dejado la década más reciente de sondeos sin analizar responde al interés por probar este nuevo método de pronóstico. No obstante, no se descarta la necesidad futura de recalibrar la clasificación, si se percibe un aumento de la varianza en el tiempo de las situaciones de tormenta englobadas en cada grupo con respecto a sus situaciones medias.

Sería interesante comprobar si, para cada tipo de tormenta, las variables señaladas como determinantes por los árboles de decisión presentan valores anómalos en sus respectivas distribuciones de densidad de probabilidad.

Además, para comprender mejor cada perfil atmosférico de tormenta, se propone el estudio del mapa sinóptico medio asociado a su sondeo característico, así como los de las situaciones anterior y posterior, para observar la evolución en cada caso. Esta información puede complementarse con el análisis de la cizalladura del viento.

<b>Tipo de tormenta</b>	<b>Variable dominante</b>	<b>Variable(s) secundaria(s)</b>
1	HR850	HR700, T1000 - T700
2	T1000 - T700	T1000 - T500, HR1000
3	HR500	HR300, HR1000
4	HR700	T850 - T500
5	T1000 - T700	HR1000
6	T1000 - T300	T700 - T500, T850 - T700
7	T850 - T500	T1000 - T500
8	T700 - T500	HR700
9	T500 - T300	T1000 - T700
10	T500 - T300	HR1000
11	T500 - T300	HR700

*Tabla 2: Variables determinantes en la ocurrencia de cada TT con respecto al total de sondeos con no ocurrencia de tormenta u ocurrencia de otro tipo.*

Tipo de tormenta	Variable dominante	Variable(s) secundaria(s)
1	HR850	T850 - T500
2	T1000 - T700	T1000 - T300
3	T1000 - T500	HR300, HR500
4	T850 - T700	HR850, HR500
5	T700 - T500	T1000 - T300
6	HR700	T700 - T500, T850 - T700
7	T850 - T500	ninguna
8	T1000 - T850	T700 - T500
9	T500 - T300	T1000 - T850
10	T500 - T300	HR1000, T850 - T300
11	T700 - T300	T1000 - T850, T500 - T300

Tabla 3: Variables determinantes en la ocurrencia de cada TT con respecto al resto de grupos.

### AGRADECIMIENTOS

Deseamos reconocer, en primer lugar, la labor de la Unidad de Climatología y Atención a Usuarios de la Delegación de AEMET en la Región de Murcia por la facilitación de los datos necesarios para este estudio. Del mismo modo, nos gustaría agradecer a la *R Foundation* la proporción del paquete *party* con el que poder realizar el análisis estadístico con árboles de decisión.

Por último, queremos mostrar nuestra gratitud a Juan Pedro Montávez por su apoyo durante el trabajo.

### REFERENCIAS

- Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics*, 15(3), 651-674.
- Sanz Moral, M. P., Arteche García, J. L., & Ortiz Berenguer, F. J. (2004). Climatología de tormentas veraniegas en Cantabria, Asturias y País Vasco. J.C. García Codron, C. Diego Liaño, P. Fdez. de Arróyabe Hernández, C.

- Garmendia Pedraja y D. Rasilla Álvarez (Eds.), El clima entre el mar y la montaña. Servicio de Publicaciones de la Universidad de Cantabria: Asociación Española de Climatología (AEC), Serie A, nº 4.
- von Storch, H. & Zwiers, F. W. (1999). Empirical orthogonal functions. Statistical analysis in climate research (pp. 135-192). Cambridge: Cambridge University Press.
- Wilks, D. S. (2011). Statistical methods in the atmospheric sciences (Vol. 100). Academic press.