

Predicción probabilística de tormentas mediante “Machine Learning”

Probabilistic forecast of convective storms using Machine Learning

J. Castro-Arias (1), A. García-Abenza (1), J. Palacios (1), F.J. Bello-Millán (1,2)

(1) Agencia Estatal de Meteorología (AEMET). 2) Universidad de Málaga

RESUMEN

La predicción de tormentas constituye uno de los principales desafíos de la meteorología operativa, especialmente durante el verano, cuando estos fenómenos pueden dar lugar a impactos relevantes. El uso del “graupel” integrado en la vertical, derivado de modelos de modelos meteorológicos, es uno de los métodos más empleados para estimar la probabilidad de tormentas. A pesar de su utilidad, presenta limitaciones para representar adecuadamente la complejidad de los procesos físicos implicados en el desarrollo de la convección profunda. Las técnicas de *Machine Learning* (ML) surgen como alternativa, siendo el objetivo principal de este trabajo el desarrollo de un modelo de ML capaz de estimar la probabilidad de ocurrencia de tormentas en el área de la península Ibérica y el archipiélago balear. De esta forma se persigue, tanto mejorar el método operativo de referencia basado en el “graupel” integrado en la vertical como analizar la capacidad de las técnicas de ML para identificar automáticamente las variables más relevantes asociadas a la formación de convección profunda. El estudio analiza cuantitativamente los resultados obtenidos durante la temporada estival, incluyendo ejemplos representativos que permiten ilustrar el comportamiento del modelo en situaciones concretas. Para el entrenamiento se utilizan datos de los veranos meteorológicos de 2022 y 2023, del verano de 2024 para la fase de validación y de 2025 para la evaluación, asegurando un conjunto de test independiente. Las variables de entrada proceden de las salidas del modelo meteorológico HARMONIE-AROME, permitiendo así analizar el rendimiento del modelo de ML en un contexto próximo al operativo. Para definir de manera objetiva la ocurrencia de tormentas, se considera que un evento convectivo ha tenido lugar cuando se detecta al menos un rayo, tanto nube-tierra como nube-nube, a una distancia inferior a 15 km de un punto determinado. En consecuencia, la evaluación del modelo frente a observaciones permite contrastar directamente las probabilidades estimadas con la ocurrencia real de actividad eléctrica asociada a la convección. Los resultados muestran que el modelo de ML mejora de forma consistente el método de referencia basado exclusivamente en el “graupel” integrado. Asimismo, el análisis de la importancia de las variables identifica la temperatura y la humedad como los predictores más relevantes, lo cual presenta un claro sentido físico, dado su papel fundamental en la inestabilidad atmosférica y en los procesos de génesis y evolución de la convección. Esta coherencia del modelo con el conocimiento meteorológico previo pone de manifiesto el potencial de las técnicas de ML para mejorar la predicción probabilística de tormentas. El enfoque propuesto no solo supera al método operativo de referencia, sino que también aporta información adicional sobre los factores clave implicados en la formación de convección. Por todo ello, el sistema desarrollado presenta un elevado potencial para su implementación en un entorno operativo, contribuyendo a una mejor anticipación de episodios de convección profunda durante la temporada de verano.