

LA DEPENDENCIA DE LA CALIDAD DE LOS AVISOS CON EL TAMAÑO DEL ÁREA PARA EL QUE SE EMITEN

Ricardo Torrijo Murciano⁽¹⁾, Alejandro Lomas González⁽²⁾

⁽¹⁾ AEMET, Paseo del Uruguay, 2, 28071 Madrid, rtorrijom@aemet.es

⁽²⁾ AEMET, c/Leonardo Prieto Castro, 40, 28071 Madrid, alomas@aemet.es

1.-Introducción

En AEMET se controla rutinariamente la calidad de los avisos, como un método imprescindible para evaluar el buen funcionamiento del Sistema Nacional de Predicción. Téngase en cuenta que, por sus grandes repercusiones sociales y económicas, los avisos son una de las principales ocupaciones que justifican la existencia de los Servicios Públicos de Meteorología cuya labor es “uno de los retornos más visibles que los contribuyentes obtienen de los impuestos” y, aunque “es difícil cuantificar dicho retorno en términos económicos”, es esencial llevar un control de su funcionamiento (Gordon y Shaykewich, 2000).

Los avisos se emiten a nivel zonal, para cada una de las comarcas meteorológicas en que se divide una provincia, cuando se considera que se van a superar una serie de valores en las variables meteorológicas. Dichas comarcas no son las mismas que las correspondientes a otras clasificaciones meteorológicas, sino que son propias de una división con la que trabaja AEMET. Una vez que los avisos se han emitido, se verifican siguiendo una serie de criterios. Se considera que han sido emitidos correctamente cuando se observa que afectan al menos a un 30% de la población o a un porcentaje del territorio del 30% ó 40%. Hay otras muchas consideraciones técnicas y prácticas habituales que requerirían mucho más detalle, pero, a grandes rasgos lo dicho es lo fundamental. Por otro lado, en la web de la AEMET (http://www.aemet.es/es/lineas_de_interes/meteoalerta), se puede consultar la última versión del plan METEOALERTA (AEMET, 2015), con documentación sobre los límites de cada comarca y los criterios de emisión de avisos en base a la superación de determinados umbrales. Para aquellos interesados en el tema de este artículo y que no estén familiarizados con los avisos de AEMET, se recomienda revisar la documentación contenida en el enlace.

El área de las comarcas meteorológicas en las que se emiten los avisos es muy variables, así por ejemplo en Canarias y Baleares el tamaño medio es en torno a 15 ó 20 veces inferior a la media y en las Castillas 3 o 4 veces superior. Por eso se plantea, al analizar la calidad de los avisos, si el tamaño de las zonas de aviso influye significativamente. Es lógico considerar que, cuanto menor es la escala espacial objeto de la verificación, mayor dificultad. Para evaluar esta dificultad, aunque sea grosso modo, se ha propuesto un pequeño estudio basado en la diferencia de proporción entre las áreas medias de las zonas de avisos de las diversas comunidades autónomas y los índices de calidad de la predicción de avisos. Ese estudio se centrará en aquellos de nivel naranja y rojo, los de mayor impacto, emitidos para el día siguiente.

En nuestro estudio partiremos de los avisos de rachas máximas de viento, precipitación (que agrupa superación de umbrales de lluvia en una hora y en doce horas), nieve y tormentas. Emplearemos los clásicos índices PD, PFA y CSI. Son adecuados para realizar estadísticas cuando cada una de las predicciones se evalúan de forma categórica, como buenas o malas. La definición de estos índices es la siguiente:

El coeficiente PD (Panofsky y Brier, 1965), es la probabilidad de detección. Se calcula realizando el cociente entre los avisos que han sido correctamente previstos (X) y el número total de avisos observados (tanto los que habían sido previstos X, como los que no lo habían sido, Y):

$$PD = \frac{X}{X + Y} \quad (1)$$

La PFA es la probabilidad de falsas alarmas, utilizando el acrónimo empelado en AEMET y la denominación sugerida por Barnes et al. (2009). Se calcula realizando el cociente entre el número de avisos previstos y no verificados (Z) y el número total de avisos emitidos (tanto los mal previstos, Z, como los bien previstos, X):

$$PFA = \frac{Z}{X + Z} \quad (2)$$

El CSI (Donaldson et al., 1975), es el índice crítico de aciertos (acrónimo del inglés Critical Success Index). Se calcula realizando el cociente entre el número de avisos previstos y verificados entre la suma del número total de avisos emitidos y el número de avisos observados que no habían sido previstos. El CSI se puede calcular a partir de los anteriores con la siguiente fórmula, fácilmente deducible:

$$CSI = \frac{X}{X + Y + Z} = \frac{PD(1 - PFA)}{1 + PFA(PD - 1)} \quad (3)$$

2.-Comentarios de la calidad de los avisos a nivel zonal

A nivel zonal, las estadísticas de avisos emitidos el día anterior a su verificación, para los avisos naranjas y rojos, que son los más difíciles de predecir, dan resultados de una PD de un 68% y de una PFA del 44%. Los avisos en estudio, naranjas y rojos, son los de mayor impacto socio-económico, pero también los de mayor dificultad, ya que los fenómenos extremos son los peor predichos por los modelos numéricos que sirven de apoyo a la predicción y requieren la intervención activa de un predictor experimentado para ponderarlos. Por todo ello, las índices de calidad de verificación de estos avisos son necesariamente inferiores a los de otras predicciones más sencillas. Además de esto, se deben matizar varios aspectos a la hora de interpretar estos resultados:

El primero es que los avisos se emiten de forma probabilista, ya que la dificultad de predecir fenómenos extremos es tan grande que no se puede hacer de otro modo, pero, por otro lado, la verificación de los mismos es categórica, bien o mal, lo que introduce cierta “injusticia” en la calibración de su calidad. Si se tiene en cuenta que

los avisos se emiten en un rango de probabilidades, normalmente se puntualiza que entre el 40 y el 70%, la tasa de falsas alarmas es la que cabría esperar.

También hay que matizar que los avisos se verifican si afectan a un tercio del territorio o población, por lo cual la percepción del público sobre si un aviso se ha cumplido puede ser diferente del criterio técnico que aquí empleamos. Es decir si los avisos solo afectan al 40% de la población se dan por buenos, pero puede que el 60% perciba que han fallado.

Si ponemos en conjunto los dos factores anteriores, se observa, en primer lugar, que la verificación categórica da una visión más negativa que la realidad de las predicciones probabilistas y que, por otro lado, debido a los criterios empleados, se favorece una mejor verificación. Es decir, globalmente contamos con dos factores que se contrarrestan de alguna manera, pero siempre se debe ser muy cuidadoso con la interpretación de los resultados.

En cualquier caso, se observa que, a nivel zonal, la mayoría de los avisos observados, cerca del 70%, estaban bien previstos el día anterior y la mayoría de los emitidos, cerca del 60%, no resultan ser falsas alarmas. Se plantea a veces aumentar el número de avisos para reducir la tasa de no-detectados. Los partidarios de esta medida alegan que es mucho menos grave una falsa alarma que un fenómeno no-detectado, pero otros opinan que este tipo de acciones puede llevar a un significativo incremento del número de falsas alarmas, algo que también tiene repercusiones negativas, sin que, por la impredecibilidad de la atmósfera en situaciones extremas, se reduzca significativamente la tasa de no-detectados.

3.-Relación entre el área de la zona y la calidad del aviso: Desarrollo y resultados

A partir de los avisos naranjas y rojos, emitidos con un día de antelación, de viento nieve, precipitación y tormentas, se ha procedido a realizar una pequeña regresión lineal entre los diversos coeficientes PD y PFA y el tamaño relativo medio del área de las zonas de aviso, para las que se emiten, en cada comunidad autónoma. La razón de escoger los resultados con un día de antelación es que se trata de un período en el que la calidad de la predicción mejora bastante, respecto a los avisos emitidos con dos días de antelación, y a su vez es suficientemente grande para que las autoridades a las que se informa puedan tomar medidas eficientes. Además, y por las mismas razones, en ese alcance se dispone en AEMET de estadísticas y procedimientos de verificación más exhaustivos.

En el estudio se ha partido de las diferencias entre la calidad de la predicción a escala zonal desde enero de 2010 a septiembre de 2017, así como de las diferencias entre las áreas medias de las zonas de aviso de las diferentes comunidades autónomas. Se ha cogido un período corto por mayor facilidad en la disponibilidad de datos, pero escoger un período mayor también hubiera tenido el inconveniente de mayor heterogeneidad en cuanto a la forma y herramientas de verificación y predicción. Naturalmente que en la calidad de la predicción intervienen muchas más variables que el tamaño de la zona a considerar, como la complejidad de la orografía, la ubicación y el tipo y frecuencia de los diferentes tipos de fenómenos adversos más habituales, pero consideramos que un

estudio basado en el tamaño del área puede ser útil. Por otra parte también se realizará un estudio focalizado por fenómenos, ya que no es lo mismo hablar de tormentas, por ejemplo, que de viento o precipitación.

Si se representan los valores globales de los índices, por comunidades autónomas, en función del área media de las zonas de aviso, y sin separar por fenómenos, resulta el gráfico que se muestra a continuación:

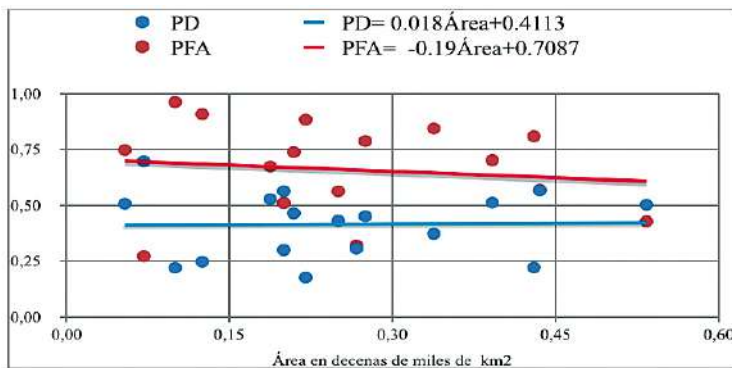


Fig. 1.- Rectas de regresión entre los valores medios para cada comunidad, a escala zonal, de los diferentes índices de calidad de avisos elaborados por AEMET, y las áreas medias zonales, en decenas de miles de km². Para PD el coeficiente R2 es de 0,003 y para PFA de 0,016. Los valores se han obtenido de la intranet de AEMET.

En contra de lo que cabría esperar, las tendencias son muy pequeñas, sobre todo en relación a PD, y las correlaciones son poco significativas estadísticamente. Todo lo anterior requiere una mayor investigación de los resultados. Una de las razones de la falta de correlación, que parece poco intuitiva o esperable, puede tener que ver con que las comarcas más pequeñas no están distribuidas homogéneamente y que en España la orografía y el clima es muy cambiante, lo que hace que intervengan multitud de factores regionales en la calidad de la predicción que pueden enmascarar la dependencia entre el área de las zonas y la calidad de los avisos.

Aparte de lo anterior, también puede haber una dependencia del tipo de fenómeno, lo que se investigará en primer lugar con la siguiente tabla:

Tabla 1. Coeficientes de regresión entre los valores medios para cada comunidad, a escala zonal, de los diferentes índices de calidad de avisos elaborados por AEMET, y las áreas medias zonales, en decenas de miles de km². Los valores se han obtenido de la intranet de AEMET.

	VIENTO	NIEVE	PRECIPITACIÓN	TORMENTAS
COEFICIENTE DE REGRESIÓN ENTRE PD Y EL DEL ÁREA	0,087	0,094	-0,504	0,283
COEFICIENTE DE CORRELACIÓN	0,004	0,007	0,056	0,014
COEFICIENTE DE REGRESIÓN ENTRE PFA Y EL ÁREA	0,129	-0,248	0,073	-0,311
COEFICIENTE DE CORRELACIÓN	0,000	0,038	0,024	0,024

Para la nieve, cuyos índices PD y PFA dan valores en torno a un 50% mejores que la media, al analizar los resultados se observa una tendencia pequeña en el índice PD y mayor tendencia a disminuir el índice PFA. De cualquier forma, hay que ser muy prudente porque la nieve se considera un fenómeno muy difícil de predecir, los resultados son poco significativos y se basan en un número de datos no muy grande, ya que los días de nevadas intensas no son muy frecuentes.

En el caso del viento los índices tienen un valor similar a la media y apenas muestran ni tendencia ni correlación. El estudio parece indicar falta de dependencia entre el tamaño de la zona de predicción y la calidad de la predicción de los avisos de vientos fuertes. Este hecho podría tener que ver con que el viento fuerte muchas veces se produce asociado a situaciones sinópticas que afectan a zonas más amplias y en el que las zonas de transición son más predecibles.

Las lluvias fuertes son el fenómeno con mayor PFA, 0,62, sin embargo la PD está más próxima a la media, 0,58. En el caso de la precipitación las tendencias son de signo contrario a las demás variables. Además la pendiente del índice PD muestra un valor relevante y negativo. Podría esperarse que a mayor área de predicción más fácil es acertar en los avisos, pero se observa una tendencia contraria que requeriría un estudio mucho más detallado. Tal vez tenga que ver con que muchas de las zonas de avisos más pequeñas se concentran en el extremo norte peninsular. Allí la llegada, desde el Atlántico, de las borrascas y frentes más activos pudiera ser más predecible que en zonas más al sur y este, que se ubican en una área de transición o bajo la influencia de tipos de tiempo mediterráneos más impredecibles o en comarcas en que el peso de las tormentas en los días de lluvias intensas es mayor. De hecho, cuando se eliminan de la estadística las comunidades cántabras, la pendiente de la recta PD se hace casi nula y la de PFA pasa a ser -0,310 con un coeficiente R^2 de 0,056.

Las tormentas severas son el fenómeno con la PD más baja, 0,41, si bien la tasa de falsas alarmas es similar a la media. Para las tormentas este estudio demuestra que la probabilidad de detección mejora cuando la predicción se hace para zonas más amplias, sin embargo se reduce algo más la tasa de falsas alarmas. Este hecho está en consonancia con lo que ya era conocido por muchos predictores y jefes de oficinas de predicción y vigilancia, cuando se tienen áreas pequeñas las predicciones son complicadas y que fallan muchas veces porque las tormentas se quedan cerca, pero no dentro del perímetro donde el aviso se había realizado. En todo caso, lo comentando anteriormente, se observa mejor cuando se comparan las estadísticas a nivel provincial y zonal (por ejemplo, en 2016 y para los avisos amarillos de tormentas, la PD fue de 74,3% a nivel provincial, frente a 69,7% a nivel zonal y la PFA fue de 45,1% a nivel zonal, frente a 29,1% a nivel provincial. Se ha escogido en este caso el nivel amarillo porque de este tipo de avisos hay muchos más fenómenos para valorar las estadísticas).

4.-Comparación de resultados a nivel provincial y zonal

Cuando se comparan los resultados para los índices, en relación al tamaño de las zonas de aviso, las correlaciones no son claras, lo que puede ser atribuible, como se ha dicho, a la inhomogeneidad del número y distribución del tamaño de las zonas de aviso

por comunidades autónomas y a las particularidades de dichas comunidades. Téngase en cuenta que en las islas y comunidades cantábricas, que tienen una meteorología muy particular, el tamaño medio de sus zonas es muy inferior al de otras comunidades como las del interior. Por ello al comparar globalmente la calidad de los avisos, según el tamaño de las zonas, se está comparando también la calidad de los avisos en ubicaciones muy diferentes. Para superar esta dificultad observaremos los resultados que se obtienen a nivel provincial y zonal para cada uno de los meses del año. De esta forma las particularidades locales quedarán enmascaradas y saldrá a la luz, de forma más clara, la dependencia real entre el área y la calidad de los índices, puesto que es razonable suponer mucha mayor homogeneidad, entre las zonas de una misma provincia y el conjunto de dicha provincia, que entre zonas de distinto tamaño de diferentes comunidades autónomas.

Dicho esto hay que observar que, en los informes de avisos en que basaremos nuestro estudio, la verificación a nivel provincial no se hace igual que a nivel zonal, aunque para simplificar la trataremos igualmente. Los avisos a nivel zonal se considera que han sido emitidos correctamente cuando se observa afectan a un 30% de la población y a un porcentaje del territorio del 30% ó 40% del territorio. En el caso de la verificación a nivel provincial un aviso se considera válido si ha sido previsto en la provincia y se ha cumplido, a nivel zonal, en alguna de las comarcas en que está dividida. Así por ejemplo si en una provincia hay dos falsas alarmas en una comarca y ese mismo fenómeno no ha sido detectado en otra, se tendrá un acierto a nivel provincial y tres fallos a nivel zona. Por

otro lado si en una provincia hay, para un mismo fenómeno, cinco aciertos a nivel zonal, se tendrá un único acierto a nivel provincial.

Pero aunque se debe ser cuidadoso al interpretar los resultados, debido a que no se emplea el mismo tipo de verificación a nivel zonal y provincial, los datos apuntan a cierta mejora de los índices de calidad de los avisos na-ranjas y rojos cuando el área para el que se emiten es mayor. En la siguiente tabla se comparan los resultados para 2016, año reciente del que se dispone de la información para todos los meses del año:

	MES	FFA	PD	CSI
NIVEL ZONAL	1	0,48	0,49	0,34
NIVEL ZONAL	2	0,5	0,46	0,32
NIVEL ZONAL	3	0,5	0,47	0,32
NIVEL ZONAL	4	0,5	0,47	0,32
NIVEL ZONAL	5	0,5	0,47	0,32
NIVEL ZONAL	6	0,5	0,48	0,32
NIVEL ZONAL	7	0,49	0,49	0,33
NIVEL ZONAL	8	0,49	0,53	0,35
NIVEL ZONAL	9	0,48	0,53	0,36
NIVEL ZONAL	10	0,47	0,52	0,36
NIVEL ZONAL	11	0,47	0,49	0,34
NIVEL ZONAL	12	0,47	0,51	0,35
NIVEL PROVINCIAL	1	0,36	0,57	0,43
NIVEL PROVINCIAL	2	0,37	0,55	0,42
NIVEL PROVINCIAL	3	0,38	0,56	0,42
NIVEL PROVINCIAL	4	0,37	0,56	0,42
NIVEL PROVINCIAL	5	0,36	0,57	0,43
NIVEL PROVINCIAL	6	0,36	0,57	0,43
NIVEL PROVINCIAL	7	0,35	0,57	0,44
NIVEL PROVINCIAL	8	0,35	0,59	0,45
NIVEL PROVINCIAL	9	0,35	0,58	0,44
NIVEL PROVINCIAL	10	0,38	0,57	0,42
NIVEL PROVINCIAL	11	0,38	0,55	0,41
NIVEL PROVINCIAL	12	0,38	0,57	0,42
IN.ZON./IN. PROV.	-	1,33	0,87	0,79

Tabla 2.- Estadística de calidad de avisos elaborados por AEMET para 2016, a nivel zonal y provincial, promediado para toda España y para cada uno de los meses del año. El área media provincial es 3,7 veces mayor que la zonal. Fuente de datos: AEMET.

Es posible que la forma en que se verifican los avisos provinciales favorezca más a la calidad de su verificación que la que se hace a nivel zonal. Por eso los resultados aquí obtenidos no son del todo extrapolables. De hecho, si se aplicasen los mismos criterios a nivel provincial que zonal, probablemente la calidad de los primeros bajaría algo. De todas maneras, aunque habitualmente los avisos provinciales verifican mejor que los zonales, se observa, como curiosidad, que hay meses en que, en algunas comunidades autónomas, esto no ocurre, lo que es una consecuencia de los criterios utilizados para dicha verificación. En el siguiente ejemplo se pone una situación virtual en que a nivel provincial la verificación es peor que a nivel zonal:

	DIAX	DIAY	DIAZ	DIAX	DIAY	DIAY	ACERTO	FALSAALAFIMA	NODETECTADO	
ZONA A	1	1					2			
ZONA B	1	1					2			
ZONA C	1	1					2			
ZONA D	1	1	1				3			
ZONA E	1	1	1				3			
ZONA F	1	1	1				3			
ZONA G	1	1	1				2	1	1	
ZONA H	1	1		1		1	1	2	2	
SUMAZONAS	7+1	8	3+1+	1+		1	18	4	2	
PROVINCIA	1	1	1	1		1	4	1	1	
ACERTOS							CS	FA	FD	
FALSAALAFIMA							NIVEL ZONAL	0,75	0,18	0,80
NARANJANO DETECTADO							NIVEL PROVINCIAL	0,67	0,20	0,80

Tabla 3. Ejemplo de situación en que la verificación a nivel zonal da mejores resultados que a nivel provincial con los métodos que se usan en la actualidad.

¿Por qué ocurre esto? Cuando se juntan, en ciertos períodos avisos que verifican bien y que afectan a un gran número de zonas, como por ejemplo puede ocurrir con el viento o precipitaciones generalizadas, las estadísticas de calidad de los índices de avisos a nivel zonal mejoran mucho más que a nivel provincial. Sin embargo, en las situaciones más habituales y como cabría esperar, se observa una mayor calidad de los avisos provinciales frente a los zonales.

5.-Conclusiones

A nivel nacional, las estadísticas de avisos emitidos el día anterior a su verificación, para los avisos naranjas y rojos, que son los de mayor dificultad y repercusión, dan resultados, promediados para toda España y todos los tipos de aviso, a nivel zonal, de una PD de un 68% y de una PFA del 44%. En las diferentes comunidades además los resultados son diferentes debido a multitud de factores. Al hacer un estudio focalizado por tipo de aviso también se observan ciertas particularidades y diferentes comportamientos. La interpretación de los resultados es muy compleja y requiere un estudio muy detallado para cada comarca, cada fenómeno y cada tipo de situación meteorológica.

En algunas discusiones se ha planteado la diferencia que puede haber entre la calidad de los índices de verificación, a nivel zonal, en las diferentes comunidades autónomas, según el tamaño medio de sus comarcas. Evaluada esta situación, los resultados ofrecen tendencias pequeñas y poco significativas. Parecería esperarse a priori que las predicciones mejoran cuando la escala espacial es mayor y, sin embargo, esto no parece

observarse. Ello podría ser atribuible a la inhomogeneidad del número y distribución del tamaño de las zonas de aviso por comunidades autónomas y a las peculiaridades de dichas comunidades. Téngase en cuenta que en las islas y comunidades cantábricas, que tienen una meteorología muy particular, el tamaño medio de sus zonas es muy inferior al de otras comunidades como las del interior.

Para superar la anterior dificultad, también se ha estudiado la relación entre la calidad de los avisos a nivel zonal y provincial. Esta comparación, con la que se consigue que muchas particularidades locales no afecten a los resultados, muestra que, a nivel provincial, los índices son claramente mejores que a nivel zonal en todos los meses del año. Aunque se debe ser cuidadoso al interpretar los resultados, debido a que no se emplea el mismo tipo de verificación a nivel zonal y provincial, los datos apuntan a que podría haber cierta mejora de los índices de calidad de los avisos naranjas y rojos cuando el área para el que se emiten es mayor.

Referencias:

- AEMET, 2015: Plan Nacional de Predicción y Vigilancia de Fenómenos Adversos: METEOALERTA, Version 6. Disponible: http://www.aemet.es/documentos/es/eltiempo/prediccion/avisos/plan_meteoalerta_v6.pdf [Consulta: 2017, Abril, 30]
- Donaldson, R.J. Jr, 1958: Analisis of severe convective storms observed by radar. *J. Meteor.*, 14, 44-50.
- Gordon N. And J. Shaykewich ,2000: Guidelines of Performance of Public Weather Services. *World Meteorological Organization. WMO/TD No. 1023.*
- Barnes, L.R., D.M. Schultz, E.C. Grunfest, M.H. Hayden, C.C. Benight, 2009, CORRIGENDUM: False Alarm Rate or False Alarm Ratio?. *Wea.Forecasting*, 24, 1452-1454.
- Panofsky, H.A. y G.W. Brier, 1965: Some applications of statistics to meteorology. *The Pennsylvania State University, University Park, PA*, 224 pp.