

**XXXIV Jornadas Científicas de la Asociación Meteorológica Española
(Teruel, 29 febrero – 2 marzo 2016)
ISBN 978-84-617-5240-9**

**Análisis del uso operativo de los indicadores de verificación
basados en tablas de contingencia en el pronóstico de fenómenos
meteorológicos adversos**

Juan de Dios del Pino Corredera⁽¹⁾

⁽¹⁾ Agencia Estatal de Meteorología, Avda. Américo Vespucio 3, CP41071 Sevilla (España),
jdelpinoc@aemet.es

Resumen

Desde hace algunas décadas en la Agencia Estatal de Meteorología (AEMET) se usan de forma operativa diversos indicadores de calidad basados en tablas de contingencia para evaluar los pronósticos de fenómenos meteorológicos adversos (FMA). Además en los últimos 6 años los indicadores FAR (False Alarm Ratio), POD (*Probability of Detection*) y CSI (Critical Success Index) forman parte del cuadro de mandos para analizar la evolución de los procedimientos operativos de predicción aplicados a la emisión de avisos de FMA.

Puesto que los FMA suelen ser más frecuentes en los meses invernales, se agregaron los episodios con FMA en valores acumulados interanuales, para calcular sobre estos valores interanuales los indicadores de calidad FAR, POD y CSI. De esta forma se palía la dependencia estacional y se suavizan las oscilaciones de menor periodo (recuento diario o mensual de los episodios).

La experiencia acumulada en estos años ha evidenciado la dependencia del valor de los indicadores de calidad FAR, POD y CSI con el número de FMA, incluso cuando se agregan en periodos interanuales.

El presente trabajo presenta modelos conceptuales aplicados a las dos situaciones más frecuentes de desvío en la predicción de las estructuras meteorológicas que ocasionan adversidad. Estos desvíos son el aumento o disminución de la región afecta, debidos a su vez a un aumento o disminución de la intensidad de la estructura meteorológica, y al desplazamiento del centro de la estructura. Ambos modelos muestran que en caso de producirse estos desvíos, se obtienen mejores resultados (baja FAR y altos POD y CSI) cuanto mayor es el tamaño de la

estructura meteorológica que genera el episodio adverso.

Los datos empíricos obtenidos a partir de situaciones reales en España durante el periodo 2010-2015 muestran también que la agregación mensual o interanual de episodios es insuficiente para sustraer la dependencia de los indicadores FAR, POD y CSI con el número de FMA.

La principal conclusión a la que se llegará es que si FAR, POD y CSI dependen del número de FMA, estos indicadores no se deben usar para analizar la evolución temporal de la calidad de un sistema de alerta temprana, pues cualquier variación positiva de los indicadores puede estar ligada a una mejoría en la predicción que lleva a la emisión de avisos o que han aumentado el número de FMA o a ambas circunstancias.

1.- Introducción

El procedimiento de verificación implantado en AEMET tiene como objetivo disponer de indicadores adecuados dentro del cuadro de mandos que permitan un seguimiento de la tendencia en la calidad de los avisos emitidos sobre FMA frente a las observaciones.

Los pares de datos de observación y predicción de un FMA son evaluados día a día. Obviamente esta serie de datos diarios presenta muchas oscilaciones, por lo que no es adecuada para analizar tendencias. Para paliar este problema en principio se agregaron los datos en periodos mensuales obteniéndose series temporales en las que la oscilación es menor, pero aún suficiente para dificultar el análisis de tendencias. Finalmente se ha optado por agregar los eventos en series interanuales. Es decir se agregan los avisos

diarios correspondientes a los 12 últimos meses y se asigna el valor al mes más reciente.

1.1.- Características del procedimiento de verificación.

Aunque los avisos emitidos por AEMET son probabilísticos, en el procedimiento de verificación son considerados como deterministas, asignándose siempre la probabilidad del suceso seguro.

Otra simplificación del procedimiento de verificación se refiere al número de niveles de aviso. El Plan Nacional de Predicción y Vigilancia de FMA, conocido como Meteoaleta, contempla cuatro categorías para un FMA: verde (no se emite aviso), amarillo, naranja y rojo. Sin embargo el procedimiento de verificación es dicotómico, es decir, solo contempla dos categorías: no aviso (incluye los niveles verde y amarillo) y si aviso (incluye los niveles naranja y rojo).

El procedimiento de verificación contempla solo los avisos emitidos el día anterior (D-1) al de validez del aviso y la resolución temporal de un aviso es horaria.

La resolución espacial de los avisos está definida a partir de 180 zonas geográficas, con similitudes climatológicas, de un tamaño aproximado equivalente a un círculo de 30 Km de radio. No obstante, también se analizan los avisos agregando los de las zonas de una misma provincia Esta agregación incluye aproximadamente 3 a 4 zonas por provincia.

Los FMA que se evalúan son la precipitación (máxima horaria o acumulada en 12 horas), la nieve (acumulada en 24 horas), la racha máxima horaria y la tormenta.

1.2.- Tablas de contingencia e indicadores del procedimiento de verificación.

Puesto que los eventos que se evalúan son finalmente dicotómicos: SI/NO aviso, el procedimiento de verificación descansa sobre la evaluación de tablas de contingencias y de los indicadores: FAR, POD y CSI.

		Categoría OBSERVADA		
		SÍ	NO	TOTAL Previstos
Categoría PREVISTA	SÍ	SS	SN	SS+SN
	NO	NS	NN	NS + NN
TOTAL Observados		SS + NS	SN + NN	N = SS+ SN + NS +NN

Tabla 1. Distribución conjunta de predicciones y observaciones (tablas de contingencia).

Siendo,

- SS: número de eventos Si previstos y Si observados (éxitos)
- SN: número de eventos Si previstos y No observados (falsos avisos)
- NS: número de eventos No previstos pero Si observados (fallos)
- NN: número de eventos No previstos y No observados

$$FAR = \frac{SN}{SN + SS} \quad (1)$$

$$POD = \frac{SS}{SS + NS} \quad (2)$$

$$CSI = \frac{SS}{SS + SN + NS} \quad (3)$$

En este trabajo se utilizarán PFA, PD y CSI en lugar de PFA, POD y CSI, respectivamente.

2.- Resultados de verificación de la serie de eventos agregados mensuales

Como primera solución para lograr el objetivo de disponer de indicadores que permitan un seguimiento mes a mes de la evolución de la calidad de los avisos emitidos para FMA, se opta por la agregación en periodos mensuales de los pares observación y predicción emitidos día a día.

Las figuras 1 y 2 presentan los indicadores PFA, PD y CSI desde enero de 2010 hasta diciembre de 2015 acumulando los valores mes a mes. En estas figuras se aprecia la enorme variabilidad mensual de estos datos, que hace que los indicadores sean poco útiles para poder analizar alguna tendencia que informe de la mejora o no del procedimiento de emisión de avisos de FMA de AEMET.

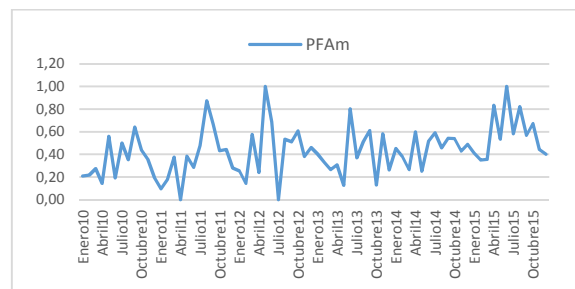


Fig. 1. Proporción e Falsas Alarmas (PFAM) calculadas a partir de datos acumulados en un mes durante el periodo 2010-2015.

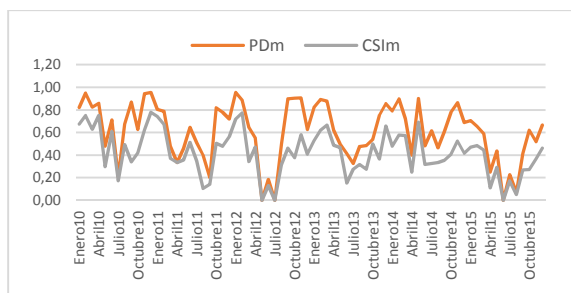


Fig. 2. Fig. 1. Probabilidad de detección (PDm) e Índice de Aceritos Críticos (CSIm) calculados a partir de datos acumulados en un mes, durante el periodo 2010-2015

La figura 3 muestra el número de avisos emitidos (PREm) y el número de FMA observados (OBSm) agregando en ambos casos los datos para un mismo mes. Se observa en esta gráfica una gran oscilación, explicada en parte por el carácter estacional de los FMA en España, que son más frecuentes en los meses invernales.

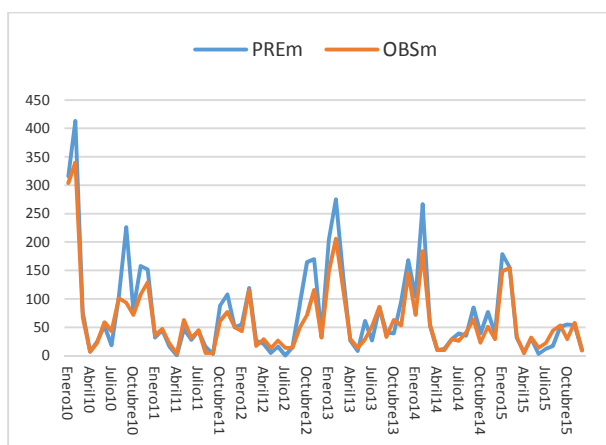


Fig. 3. Número de FMA previstos o avisos (PREm) y observados (OBSm) acumulados en un mes durante el periodo 2010-2015

Destaca también en la figura 3 como los máximos, tanto de las observaciones como de las predicciones, no tienen la misma amplitud, mostrando un máximo muy destacado al comienzo de la serie en los meses de enero y febrero de 2010. La tabla 2 recoge el valor y momento de estos máximos.

Fecha	FMA previstos	FMA observados
2010-feb	413	340
2010-ene	316	304
2013-feb	275	206
2014-feb	267	184

Tabla 2. Valores y fechas de los máximos con mayor número de FMA observados y previstos acumulados en un mes durante el periodo 2010-2015.

Las variabilidades apreciadas en el número de FMA previstos u observados y en los indicadores PFA, PD y CSI, como veremos, están relacionadas.

3.- Resultados de verificación de la serie de eventos agregados interanuales

Debido a las oscilaciones que presentan las series de datos agregados mensuales, se agregaron los datos para periodos interanuales.

La figura 4 muestra los indicadores habituales (PDA, PFAa, CSIA) del cuadro de mandos desde enero de 2010, año a partir del cual AEMET comenzó a evaluarlos, hasta diciembre de 2015.

Estos indicadores son interanuales, es decir, a cada mes se le asignan todos los eventos del mes en cuestión y los 11 precedentes. Sin embargo, debido a que el procedimiento se inicia en enero de 2010, a los 11 meses primeros de 2010 no se le puede calcular un valor interanual, sino que se acumulan solo los datos de los meses precedentes. Esto tendrá un peso importante a la hora de interpretar la tendencia.

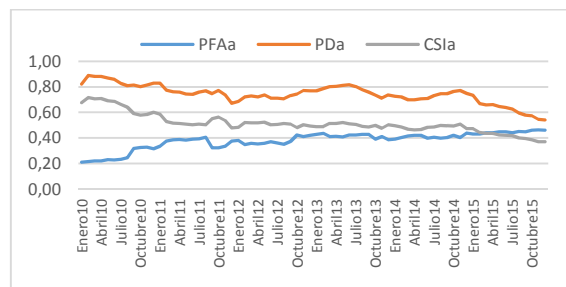


Fig. 4. Proporción e Falsas Alarmas (PFAa), Probabilidad de Detección (PDA) e Índice de Aceritos Críticos (CSIA) calculadas a partir de datos acumulados en 12 meses, durante el periodo 2010-2015.

En la figura 5 se representan los valores agregados interanuales tanto para los FMA previstos (PREa) como los observados (OBSa). En estas gráficas se aprecia un ascenso inicial, que se debe a que aún el año no está completo, pese a que parte del segundo máximo histórico de la serie tanto del número de FMA observados como previstos. A partir de diciembre de 2010, cuando la serie ya incluye 12 meses para cada valor, se aprecia un descenso continuado de los FMA observados.

El número de FMA previstos y observados de la figura 5 vuelve a subir en el comienzo de 2013 debido al tercer máximo histórico mensual de los 206 FMA de febrero de 2013, y que como cualquier valor extremo extiende su inercia en los meses siguientes, dejando después un descenso continuado en el número de FMA acumulados interanual desde septiembre de 2013 hasta la fecha. Un periodo de más de 2 años, que comprende casi 3 otoños-inviernos, que son las estaciones con mayor número de FMA.

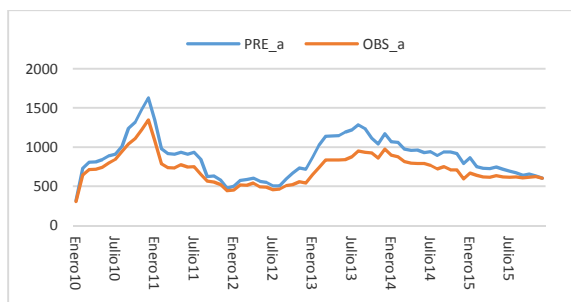


Fig. 5. Serie FMA previstos (PREa) y observados (OBSa) agregados en periodos interanuales para el periodo (2010-2015).

Regresando a la figura 4 de los indicadores interanuales, tras la caída continuada en el valor de PD desde enero de 2011 hasta diciembre de 2011, vuelve a subir con ligeras oscilaciones hasta alcanzar un nuevo máximo en junio de 2013 (avisos desde julio de 2012 ha junio de 2013). A partir de aquí y casi hasta el final de 2015 se aprecia un descenso del valor de PDA, acompañado de un ascenso en la PFA que se traduce en un descenso del CSI. Esto podría interpretarse a priori como un empeoramiento en la calidad de los avisos, pero como veremos está ligado al descenso en el número de FMA previstos (avisos emitidos) u observados que hemos visto en la figura 5.

4.- Análisis de la rareza de los FMA

La sintonía o acoplamiento que parecen apreciarse en la calidad los avisos agregados en periodos mensuales entre los indicadores PFA, PD y CSI y el número de FMA observados en el mismo periodo, más evidente con la agregación interanual de los eventos, parece intuir una estrecha dependencia entre el valor de los indicadores y el número de FMA.

De ser así, una posible tendencia positiva o negativa en la evolución de cualquier indicador podría no significar una mejora o un empeoramiento del procedimiento de emisión de avisos, sino un aumento o una disminución de la frecuencia de avisos, o de ambos.

Para analizar el impacto del número de avisos en el valor de los indicadores, vamos a acudir al concepto de rareza de un FMA. Un evento se considera raro cuando $SS + NS \ll N$. No obstante, debido a que el número de observaciones disponibles es a veces insuficiente o los criterios para considerar que un FMA ha sido observado o no, recurrimos a la serie de avisos emitidos para analizar el concepto de rareza de los FMA que se incluyen en el procedimiento de verificación de AEMET. Es decir, consideramos que un evento es raro si $SS+SN \ll N$ o $(SS+SN)/N \ll 1$.

De los FMA que se consideran, el más frecuente es la racha máxima que comprende el 43% de los avisos

que se emiten, seguidos de la precipitación (33%), la nevada (18%) y la tormenta (7%).

El conjunto de los 4 FMA (precipitación, racha máxima, nevada y tormenta) para el periodo 2010-2015 presenta un cociente $(SS + SN)/ NN$ igual a 0.0029. Restringido a la racha máxima que es el FMA más frecuente el coeficiente sube a 0.0055. Es decir, para el conjunto del periodo 2010-2015 los FMA que se evalúan pueden considerarse como raros.

Sin embargo si se analiza cada episodio meteorológico, es decir, los datos día a día, sin agregar, se observa una gran variabilidad en cuanto a la frecuencia del fenómeno adverso que se considere. En concreto, para el caso de la racha máxima varía desde un máximo de 96 avisos emitidos en un día (para el 27 de febrero de 2010), que representa el 53% del territorio español previsiblemente afectado por avisos de nivel naranja o rojo, a ningún aviso, que sucede en el 89% del periodo analizado (2010-2015).

Por tanto, para la racha máxima el cociente $(SS+SN)/N$ varía entre un 0.53 y 0.

Por lo tanto, el concepto de evento raro es muy variable de una situación a otra en la meteorología española. Esto como veremos dificulta el uso de los indicadores (PD, PFA, CSI,...), para eventos raros cuando se utilizan para analizar la tendencia en un periodo largo (meses, años, etc.)

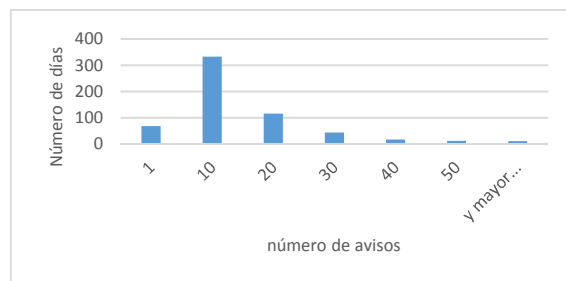


Fig. 6. Histograma con la distribución de frecuencias absolutas del número de días en los que se ha observado un FMA en el periodo 2010-2015.

5.- Estudio de la dependencia de los indicadores con el número de FMA

Para analizar el comportamiento de los indicadores (PFA, PD y CSI) con el número de FMA vamos a simular las dos situaciones más frecuentes de desviación entre la predicción y la observación, ocasionados generalmente por variaciones en los modelos numéricos de simulación de la atmósfera o por el diagnóstico del predictor. Estos casos son:

- Desplazamiento del centro de la región afectada por el FMA respecto a la prevista.

- Aumento o disminución de la zona afectada por el FMA respecto a la prevista, conocido generalmente como sub-predicción y sobre-predicción.

5.1.- Desplazamiento del centro de la región afectada respecto a la prevista

En muchas ocasiones los modelos numéricos pronostican bien la intensidad y extensión de las estructuras meteorológicas que pueden dar lugar a FMA, pero no están igualmente acertados en la posición o centro de dicha estructura. Este desvío es muy frecuente en el caso de Depresiones Aisladas en Niveles Altos (DANA), habituales en primavera y en otoño en el suroeste europeo.

Para modelizar este caso suponemos que sobre el territorio peninsular español (492175 Km² de superficie, equivalente a un círculo de 396 Km de radio), se prevé una situación con FMA para una región de forma circular, cuyo centro resultará desplazado una distancia *d* respecto a la posición prevista. El tamaño de la superficie de la región afectada será variable, permaneciendo constantes la superficie de España peninsular y la separación entre el centro de las regiones prevista y observada.

Utilizaremos la siguiente notación para caracterizar a cada una de estas regiones y distancia:

- Ae = área de España = 492175Km², equivalente a un círculo de 396 Km (constante)
- Ap = área de la región prevista con FMA (variable)
- Ao = área de la región con FMA observados (variable)
- d = 100 Km (constante)



Fig. 7. Modelo conceptual de error en la posición del centro de una estructura meteorológica que genera FMA.

Si llamamos *Ai* a la intersección de las regiones previstas y observadas ($A_i = A_p \cap A_o$), tenemos que los valores de los elementos de la tabla de contingencia (tabla 1) son:

Conceptos	Valor
N	Ae
SS	Ai
SN	Ap-Ai
NS	Ao - Ai
NN	Ae - Ai - (Ap- Ai) - (Ao - Ai) Ae - 2Ao -Ai

Tabla 3. Elementos de la tabla de contingencias e indicadores para el modelo conceptual de error en el pronóstico del centro de la estructura meteorológica que genera FMA.

La intersección de ambas regiones es la variable que dependerá del tamaño de la región afectada.

De la figura 8 se deduce que,

$$A_i = 2 A_c = 2 (A_s - A_t) \quad (4)$$

Siendo *Ac* el área del círculo que representa a la región prevista u observada con FMA (*Ao* o *Ap*), *As* el área del sector circular de radio *R* (radio de *Ao* o *Ap*) y *At* el área del triángulo delimitado por la línea que une la intersección de los círculos *Ao* y *Ap* y los radios que van desde sus extremos al centro de la región prevista con FMA.

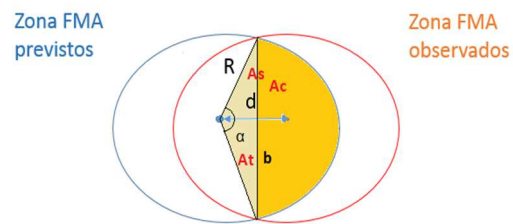


Fig. 8. Relaciones geométricas para determinar la intersección entre la región prevista y la observada, desplazadas una distancia *d*.

De las relaciones geométricas de la figura 8 se deduce las siguientes expresiones:

El área del triángulo

$$A_t = \frac{d}{2} \times b = \frac{db}{2} = \frac{dR}{2} \sin \alpha \quad (5)$$

El área del sector

$$A_s = \frac{\alpha R^2}{2} \quad (6)$$

Siendo

$$\alpha = \arccos d/2R \quad (7)$$

Y finalmente en área de la intersección:

$$A_i = \alpha R^2 - dR \sin \alpha \quad (8)$$

Esto nos permite calcular los elementos de la tabla de contingencia (SS, SN, NS, NN y N) y a partir de éstos

los indicadores PFA, PD y CSI que tienen las expresiones siguientes:

$$PFA = 1 - \frac{\alpha}{\pi} + \frac{d \operatorname{sen} \alpha}{\pi R} \quad (9)$$

En esta expresión, si aumenta el radio de la región prevista y observada, es decir, si R aumenta, $d \operatorname{sen} \alpha / \pi R$ disminuye y PFA disminuye. También si R aumenta, α aumenta, α / π aumenta y PFA disminuye. Es decir, cuanto mayor es la superficie de las regiones prevista y observada con FMA, y por tanto mayor el número de avisos que hay que emitir para esta zona, menor será el número de falsas avisos (PFA). Esta dependencia entre PFA y el número de FMA es lo que se aprecia en la figura 9.

$$PD = \frac{\alpha}{\pi} + \frac{d \operatorname{sen} \alpha}{\pi R} \quad (10)$$

En esta expresión, si R aumenta, $d \operatorname{sen} \alpha / \pi R$ disminuye y por tanto PD aumenta. También si R aumenta, α aumenta, α / π aumenta y PD aumenta. Es decir, cuanto mayor es la superficie de las regiones prevista y observada con FMA, y por tanto mayor el número de avisos que hay que emitir para esta región, mayor será la proporción de FMA que somos capaces de detectar (PD). Esto es lo que se aprecia en la figura 9.

$$CSI = \frac{PD}{2 - PD} \quad (11)$$

En este caso, si R aumenta, PD aumenta y por tanto CSI también aumenta. Es decir, cuanto mayor es la superficie de las regiones prevista y observada con FMA, y por tanto mayor el número de avisos que hay que emitir para esta región, mayor será CSI, coincidente con lo que se aprecia en la figura 9.

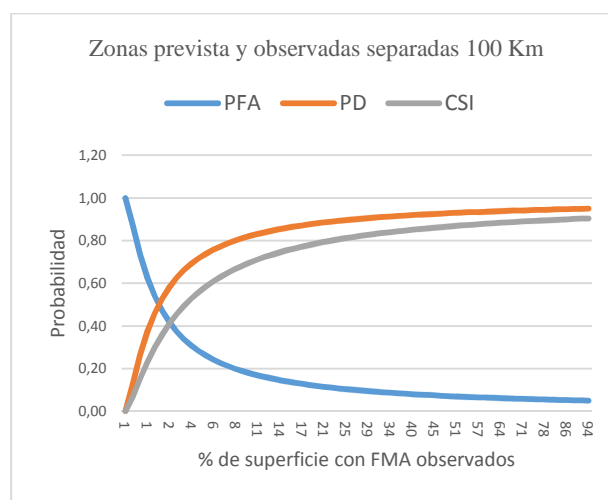


Fig. 9. Comportamiento de PFA, PD y CSI respecto a una variación en el porcentaje de región afectada con FMA en el caso de un error en el centro de la estructura meteorológica con FMA.

Las gráficas de la figura 9 no solo muestran la dependencia de PFA, PD y CSI con el número de FMA observados, sino que esta dependencia bastante

más significativa cuando hay pocos FMA. Es decir, para PD y CSI el aumento es más rápido entre el 0 y 5% de la región afectada. En España peninsular equivale a situaciones con avisos en menos de 9 zonas para un mismo FMA en un día, que suelo ser según se aprecia en la figura 6 el caso más frecuente. A partir del 10% de la región afectada las variaciones son menos importantes, que equivale a unas 18 zonas afectadas

5.2.- Sobre-predicción o sub-predicción de la superficie de la región afectada

En muchas otras ocasiones los modelos numéricos de simulación de la atmósfera pronostican bien la posición de las estructuras meteorológicas que pueden dar lugar a FMA, pero no su intensidad, reduciendo o aumentando así la extensión del área observada respecto a la prevista.

Para modelizar este caso suponemos sobre el territorio peninsular español de 492.175 Km² de superficie (A_e), equivalentes a un círculo de unos 396 Km de radio, se prevé una situación con FMA para una región de forma circular de radio R_p y superficie A_p que será concéntrica con la región en la que se observa el FMA que será también circular de radio R_o y superficie A_o , que podrá ser tanto mayor ($R_o > R_p$) o menor ($R_o < R_p$) y en consecuencia se corresponderá con situaciones de sub-predicción o sobre-predicción respectivamente.

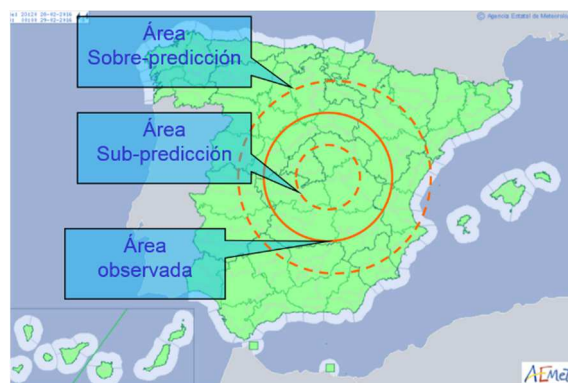


Fig. 10. Modelo conceptual de error en el tamaño de la región prevista sobre la observada debido a un cambio en la intensidad de la estructura meteorológica que genera FMA

Conceptos	$R_p < R_o$ sub-predicción	$R_p = R_o$	$R_o < R_p$ sobre-predicción
SS	A_p	A_p	A_o
SN	0	0	$A_p - A_o$
NS	$A_o - A_p$	0	0
NN	$A_e - A_o + A_p - A_p = A_e - A_o$	$A_e - A_p$	$A_e - A_p + A_o - A_o = A_e - A_p$
Indicadores			
$PD = SS / (SS + SN)$	$(R_p / R_o)^2$	1	1
PFA	0	0	$(R_p^2 - R_o^2) / R_p^2$
$CSI = SS / (SS + NS + SN)$	$(R_p / R_o)^2$		$(R_o / R_p)^2$

Tabla 4. Elementos de la tabla de contingencias e indicadores para el modelo conceptual de error en el pronóstico de la intensidad o tamaño de la región afectada por FMA.

Sub-predicción (muchos FMA no detectados)

Suponemos que las regiones con FMA previstos y observados son dos círculos concéntricos, siendo mayor en 50 Km el radio del círculo con FMA observados. Es decir

$$R_o = R_p + 50 \text{ Km} \quad (12)$$

Obviamente en este caso no hay falsas alarmas, ya que el área prevista siempre es menor que la observada, pero tanto PD como CSI aumentan y por tanto mejoran los resultados cuanto mayor es el radio de la zona afectada. La figura 11 describe este comportamiento de los indicadores a medida que la región con FMA observados y por tanto R_o aumenta. También se aprecia que la dependencia de los indicadores con el número de FMA es mayor cuando es menor la región afectada, que suele ser el caso más frecuente.

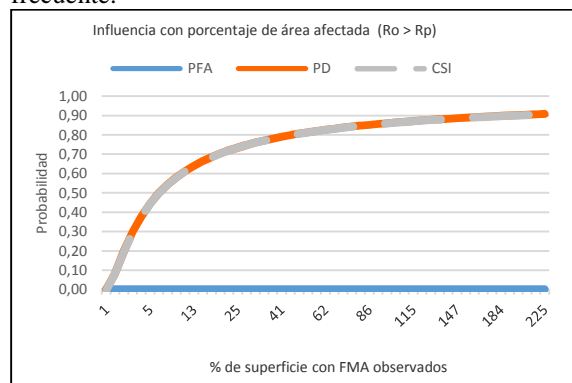


Fig. 11. Comportamiento de PFA, PD y CSI respecto a al porcentaje de región afectada con FMA, debido a una sub-predicción ($R_o > R_p$) de la intensidad de la estructura meteorológica que genera los FMA.

Sobre-predicción (muchas falsas alarmas)

En este caso volvemos a suponer que las regiones con FMA previstos y observados son dos círculos concéntricos, siendo menor en 50 Km el radio del círculo con FMA observados. Es decir

$$R_p = R_o + 50 \text{ Km} \quad (13)$$

Obviamente en este caso todos los FMA observados han sido previamente previstos, es decir, PD es nulo, ya que el área prevista siempre es mayor que la observada. Sin embargo siempre hay falsas alarmas, pero estos falsos avisos son en proporción menor cuanto mayor es el tamaño de la zona afectada. Este comportamiento descendente de la PFA implica un aumento con el tamaño de la zona afectada del CSI y

por tanto mejores resultados. Este comportamiento de los indicadores puede visualizarse en la figura 12. En la que también se aprecia que la dependencia de los indicadores con el número de FMA es mayor cuando es menor la región afectada, que suele ser el caso más frecuente.

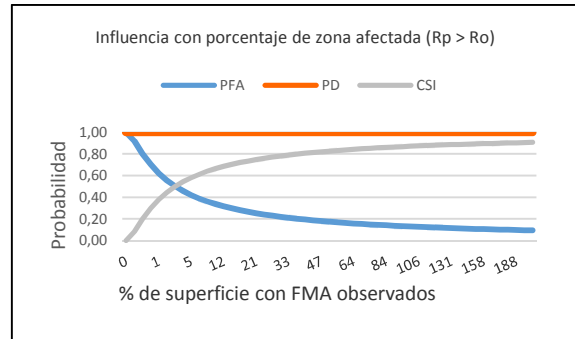


Fig. 12. Comportamiento de PFA, PD y CSI respecto a al porcentaje de región afectada con FMA, debido a una sobre-predicción ($R_p > R_o$) de la intensidad de la estructura meteorológica que genera los FMA

6. Dependencia de PFA, PD y CSI con el número de FMA en el procedimiento de verificación de AEMET.

Los modelos conceptuales del apartado anterior han mostrado como en los casos más habituales de desviación en la predicción de una situación meteorológica que derive en FMA y que por tanto impliquen la emisión de avisos, los resultados de verificación ofrecen mejores resultados cuanto mayor es la región afectada para los indicadores, es decir, PFA bajos, y PD y CSI altos.

Los modelos conceptuales anteriores tratan por separado los desvíos debidos a un error en la posición de la estructura meteorológica y la intensidad de la misma. Los casos reales serán una mezcla de ambos fenómenos, pero la influencia de los resultados con el tamaño de la región afectada será similar ya que se superponen con el mismo signo.

En este apartado se trata de ver si los resultados de los modelos conceptuales son consistentes con los resultados empíricos derivados del procedimiento de verificación establecido en AEMET aplicados al periodo 2010-2015.

6.1- Indicadores de las series de datos derivadas de la agregación mensual.

Para la serie mensual de número de FMA generada por agregación de todos los ocurridos en un mes, en el apartado 2 se destacó que los FMA son generalmente más frecuentes en los meses invernales y bastante menos frecuentes en los meses veraniegos. Esta oscilación en el número de FMA se aprecia

también en los indicadores PFA, PD y CSI calculados a partir de la serie de eventos o pares de niveles previstos y observados generada por agregación mensual de todos los eventos ocurridos en un mes.

Esta dependencia o acoplamiento entre la serie de FMA y las series de PFA, PD y CSI se visualiza mejor si ordenamos los valores de la serie de indicadores no secuencialmente en el tiempo, sino de mayor a menor número de FMA observados y las representamos gráficamente. Por ejemplo en la figura 13 se visualiza que el número de FMA detectados por la predicción (PD) disminuye con el número de FMA. En la misma figura se superpone la tendencia lineal, que muestra un PD medio para meses en los que se han registrado pocos FMA en torno a 0.40 (menos de 10) y el doble para meses más frecuentes en FMA (más de 200). Es decir, un ascenso medio del 2% por cada incremento en 10 FMA. A esta tendencia creciente se superpone una gran oscilación que es mayor en los meses con pocos FMA. En definitiva, la Probabilidad de Detección (PD) aumenta con el número de FMA observados, aunque con grandes oscilaciones.

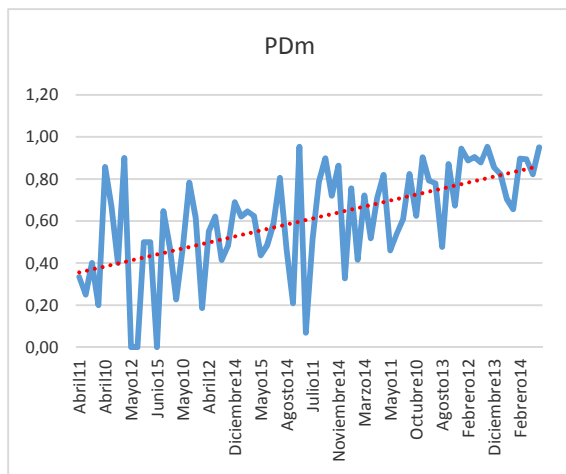


Fig. 13. Serie del indicador Probabilidad de Detección de datos agregados mensuales (PDm) ordenados de menor a mayor número de FMA observados para el periodo 2010-2015.

Por el contrario, en la figura 3b, las falsas alarmas (PFA) aumentan cuando hay menos FMA observados, desde un 30% cuando hay muchos FMA a 60% cuando hay pocos. A esta tendencia también se le superpone una gran oscilación, que es mayor cuando es pequeño el número de FMA.

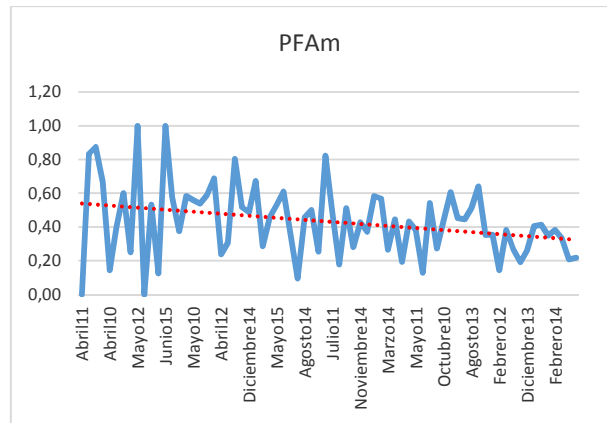


Fig. 14. Serie del indicador Proporción de Falsas Alarmas generada con datos agregados mensuales (PFAm) ordenados de menor a mayor número de FMA observados para el periodo 2010-2015.

En el CSI, que combina tanto las falsas alarmas como los FMA detectados, prevalece el comportamiento de PD, es decir, disminuye cuando aumenta el número de FMA observados. La oscilación también es mayor cuando son pocos los FMA observados.

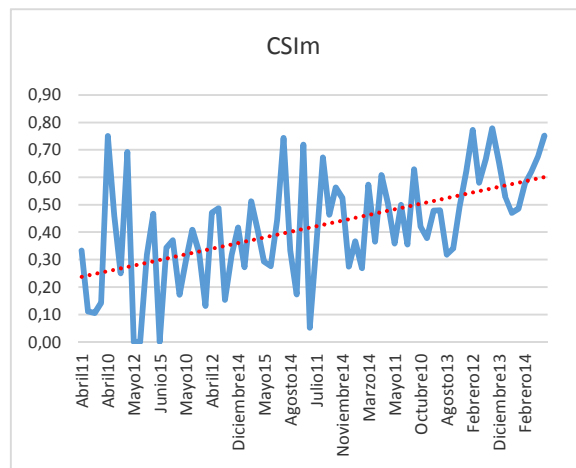


Fig. 15. Serie del Índice de Aciertos Críticos generada con datos agregados mensuales (CSIm) ordenados de menor a mayor número de FMA observados para el periodo 2010-2015.

En resumen, la dependencia manifestada por los modelos conceptuales referentes a un episodio y a un FMA se traslada también cuando los pares predicción-observación se agregan en series mensuales e incluso cuando esta agregación incluye diferentes FMA, siendo evidente que los resultados son mejores (menor PFA y mayores PD y CSI) cuando se registran muchos FMA.

6.2- Indicadores de las series de datos derivadas de la agregación interanual.

Se puede hacer un estudio similar al que se ha expuesto en el apartado anterior para las series

mensuales con el propósito de analizar si la dependencia entre los indicadores de verificación y el número de FMA observados de los modelos conceptuales se traslada a las series de datos interanuales. Este estudio ofrecería parecidas conclusiones. Sin embargo en este caso vamos a analizar sobre la propia serie interanual la influencia con el número de FMA observados.

En la figura 4 pude observarse que existen 3 máximos bien identificados en la serie interanual de PD. El primer máximo (PD=0.89) en febrero de 2010, el segundo (PD=0.52) en mayo de 2013 y el tercero (PD=0.51) en noviembre de 2014, dando a entender que la curva presenta una tendencia decreciente con oscilaciones cada vez menores.

La singularidad del máximo absoluto de febrero de 2010 se debe a que hasta esa fecha solo se disponen de dos meses para poder hacer la acumulación, pues los datos arrancan desde enero de 2010. Además en estos dos meses, enero y febrero de 2010, se registraron los dos máximos absolutos en el número de FMA observados, 304 y 340 respectivamente. En concordancia con las conclusiones de los modelos conceptuales anteriores, corresponde a este momento el mejor valor para PD. Para resaltar el enorme peso que tiene estos máximos de FMA en los primeros datos de las series PDA, PFAa y CSIa, amplificado por la imposibilidad de atenuar sus efectos en un periodo interanual completo, la figura 16 representa el número medio mensual de FMA observados en cada periodo de la serie interanual.

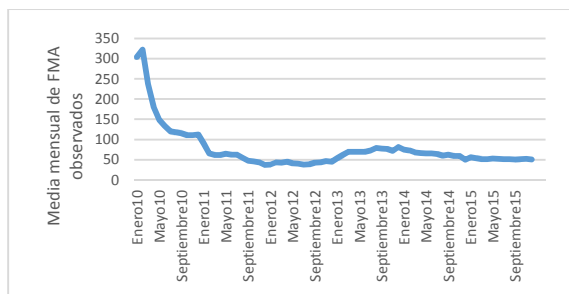


Fig. 16. Media Mensual de FMA observados agregados en periodos interanuales para el periodo 2010-2015. Desde enero a noviembre de 2010 el periodo interanual se reduce a los meses anteriores disponibles.

El máximo de junio de 2013 se produce 4 meses después del tercer máximo histórico de FMA observados en febrero de 2013, incluyendo este y otros meses con bastantes FMA observados. Esta inercia de un valor alto de FMA en los meses siguientes se extiende hasta que aparece un mes con menos FMA que el mismo mes del año anterior.

El máximo de noviembre de 2014 se produce 9 meses después de cuarto máximo de FMA observados, incluyendo también diciembre de 2013 que fue un mes con 145 FMA observados. Decae en un mes

invernal, pues el invierno de 2013, en contra de lo habitual, fue escaso en FMA.

Destaca también el mínimo absoluto de FMA a finales de 2015 respecto a lo habitual en el mismo periodo de los años anteriores. Este menor número de FMA está teniendo un gran peso en el aumento de la PFA, y en la disminución de PD y CSI.

Por tanto, la influencia del número de FMA observados, favoreciendo el comportamiento de los indicadores PFA, PD y CSI cuando son altos, se extiende también a los datos agregados en 12 meses.

Además el arranque de la serie en enero y febrero, coincidiendo con los meses más adversos de la serie histórica junto con el hecho de que en el comienzo la serie interanual está incompleta al no poderse acumular 12 meses, realza la falsa sensación de que los resultados de la verificación son cada vez peores.

7.- Conclusiones

- 1) Los FMA incluidos en el procedimiento de verificación de la AEMET (precipitación, nevada, tormenta y racha máxima) tienen una variabilidad anual, siendo más frecuentes en los meses invernales. A esta variabilidad anual se superpone otra de mayor ciclo, destacando el máximo absoluto de FMA del comienzo de 2010, que supera con creces a los otros máximos. Destaca también el escaso número de FMA de finales de 2015.
- 2) En los casos más habituales de errores en los avisos emitidos para una situación meteorológica adversa (desplazamiento del centro de la estructura meteorológica respecto al previsto y aumento o disminución de la intensidad del fenómeno previsto respecto al observado), modelos conceptuales simples muestran que cuanto mayor es el número de FMA, menor es la PFA y mayor es PD y CSI. Es decir, para situaciones con muchos FMA los resultados de la verificación son mejores. Además esta dependencia entre los indicadores y el número de FMA se acentúa más cuando son pocos los FMA y en definitiva los avisos emitidos, que suelen ser las situaciones más frecuentes.
- 3) Los datos empíricos de la serie de pares de predicción y observación de FMA del procedimiento de verificación de AEMET en el periodo 2010-2015 muestran que la agregación mensual e interanual de episodios mantiene la dependencia de los

indicadores PFA, PD y CSI con el número de FMA.

- 4) La agregación mensual de FMA previstos y observados enmascara la dificultad a priori que tiene una situación con pocos FMA, ya que trata por igual un mes con muchos FMA concentrado en un solo episodio, que muchos episodios con pocos FMA. Este problema se acentúa más en la agregación interanual que es la utilizada en el cuadro de mandos, pues considera todos los FMA ocurridos en un año como si fueran de un único episodio meteorológico.
- 5) El uso de los indicadores PFA, PD y CSI para el análisis de la evolución temporal de un procedimiento de predicción puede llevar a falsas conclusiones, pues por ejemplo al comparar dos situaciones meteorológicas muy diferentes, una con muchos avisos y otra posterior en el tiempo con pocos avisos, mejores resultados en los indicadores en la primera situación pueden deberse a una mejor predicción o la facilidad en el pronóstico debida al mayor número de FMA.
- 6) Los indicadores PFA, PD y CSI son útiles para comparar varios procedimientos de predicción (predictores humanos, modelos numéricos, climatología, etc.) aplicados a una misma situación meteorológica, ya que en este caso el factor número de FMA es el mismo para todas las predicciones.
- 7) Como líneas de trabajo para eliminar el factor número de FMA en el valor de una predicción y poder extraer conclusiones en la evolución temporal de un procedimiento de emisión de avisos se proponen:
 - Calcular la ganancia o valor añadido del procedimiento que quiere evaluar sobre otro de referencia (modelo numérico, etc.).
 - Calcular la ganancia o diferencia entre los valores de los indicadores para una situación concreta y los esperados para situaciones meteorológicas similares. Por ejemplo el valor previsto por la recta de regresión de las gráficas 13, 14 y 15.
 - Hacer un correcto uso de la probabilidad e incorporar indicadores adecuados

Bibliografía

- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.
- Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585.
- Marzban, C., 1998: Scalar measures of performance in rare-event situations. *AMS*, **13**, 75763.
- Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293.
- Peirce, C. S., 1884: The numerical measure of the success of predictions. *Science*, **4**, 453–454.