

Comunicación C-11

VALORACIONES Y RESULTADOS DE UN ESQUEMA DE VERIFICACIÓN DE PREDICCIONES GENERALES

Ramón Garrido Abenza

Sección de Estudios y Desarrollos del CMT de Murcia (INM)

RESUMEN

Se presenta un método de verificación de predicciones generales, basado en índices derivados del Ranked Probability Score y tomando una referencia climatológica para valorar la pericia. Es aplicado a las predicciones diarias para la Región de Murcia durante un semestre, con el objeto de detectar fallos significativos y facilitar un control de calidad. Las predicciones son codificadas y sometidas a una interpretación probabilista, limitándonos, de momento, al valor medio de la nubosidad y a la ocurrencia de precipitación y tormenta. Se comprueba que la predicción de nubosidad es la de mayor porcentaje de fallos, aportando escasa información adicional respecto de la climatología, mientras que la de las otras variables resulta más valiosa.

1. Introducción

El problema de la verificación de las predicciones meteorológicas es complicado, siendo poco probable que se pueda desarrollar una técnica universal de evaluación adecuada a cualquier circunstancia. En España, afortunadamente, la divulgación e implantación de un *Manual de términos meteorológicos* (INM, 1992) ha propiciado la existencia de predicciones generales expresadas objetivamente, reduciendo notablemente el problema de la interpretación, si bien todavía persisten algunos inconvenientes, al no estar específicamente concebido ese texto para la verificación.

Otro tipo de problemas son los relativos a la evolución temporal y a la distribución espacial (Garcías y Guijarro, 1995), los cuales pueden abordarse mediante la utilización de valores medios temporales y la aplicación de la predicción a una pequeña área, de la cual sean representativas las observaciones que se utilicen para la verificación. A continuación podríamos preguntarnos qué índice de verificación adoptar, de entre una gran cantidad de ellos (Murphy y Epstein, 1967), o cómo influye en la actuación del predictor el conocimiento del sistema de verificación que se aplique, incluso a nivel psicológico (Sanders, 1967). Está claro que el

índice adoptado debería aportar información acerca de la dificultad de la predicción, en comparación con predicciones triviales como puedan ser las basadas en la climatología o en la persistencia, pero no lo está tanto entre qué valores podría considerarse acertada o fallada la predicción de las variables meteorológicas, cada una de ellas con características muy peculiares. La elección de los umbrales dependerá de la cantidad de casos que se desee seleccionar.

La complejidad de los problemas someramente expuestos aconseja abordar el trabajo con acotaciones que lo simplifiquen, al menos en las etapas iniciales. En este marco se inscribe este estudio, cuyo objetivo inicial es la detección de días con fallo significativo de la predicción, aunque pueda utilizarse también para evaluar la calidad global en conjuntos temporales más amplios.

2. Metodología

2.1. Índices de verificación

La redacción de las predicciones de manera objetiva, en términos de probabilidad o con equivalencias numéricas, permite el empleo de índices de verificación especialmente diseñados para este tipo de predicciones, como el Ranked Probability Score (*RPS*) establecido por Epstein en 1969, modificado por Murphy en 1971 y adoptado por el Servicio Meteorológico canadiense en 1982. Este índice, diseñado para ser aplicado a variables ordenadas, pondera la distancia entre las categorías predichas y las observadas, penalizando los errores de acuerdo a su magnitud u orden. Está definido, para k categorías ordenadas y mutuamente excluyentes, mediante una expresión (Stanski y otros, 1989) equivalente a la siguiente:

$$RPS = 1 - \frac{\sum_{i=1}^{k-1} \left(\sum_{n=1}^i P_n - O_n \right)^2}{k-1} \quad [1]$$

con P el vector de predicción y O el de observación, ambos de k componentes, con la componente $O_i = 1$ si se observa la categoría i y $O_i = 0$ en caso contrario. El índice *RPS* varía de 0 a 1, estando orientado positivamente, es decir, sus valores mayores están asociados a los mejores pronósticos.

Para valorar la destreza de la predicción podemos comparar con un estándar, como pudiera ser el valor del mismo índice que se derivaría de la climatología o de la persistencia. Así se pretende asignar poco valor a las predicciones que también se hubieran acertado desde el punto de vista climatológico, por ejemplo. Para ello, en algunos textos (Stanski y otros, 1989; García-Moya, Genovés y otros, 1990), se opta por medir la pericia de la predicción mediante el Ranked *Probability Score Skill* (*RPSS*), definido como:

$$RPSS = \frac{RPS_p - RPS_c}{1 - RPS_c} \quad [2]$$

siendo p y c los subíndices correspondientes a la predicción y a la climatología, respectivamente. El índice oscila entre menos infinito y la unidad, adoptando valores positivos cuando la predicción ha sido más adecuada que la referencia climatológica y negativos en caso contrario. Alcanza la unidad cuando también lo hace el RPS_p .

Sin embargo, esa definición no es afortunada en algunos casos y puede conducir a paradojas. Una predicción acertada de helada en enero para Navacerrada, por ejemplo, posee el mismo valor, la unidad, que para Alicante, cuando esta última sería mucho más valiosa. Idéntico resultado obtendríamos si, fijando el lugar, nos moviéramos de un mes a otro. Este índice presenta también el inconveniente de su gran inestabilidad, al no estar acotado inferiormente. Además, es posible encontrar casos en los cuales una predicción obtiene un valor de *RPSS* inferior al que obtendría una predicción peor, con la misma referencia climatológica. Por ejemplo, una predicción de «poco nuboso» (en el sentido que se matizará más adelante) para Alcantarilla en octubre, conduciría a los valores del índice *RPSS* que figuran en la Tabla 1, con un valor de $-2,14$ en una observación de «nuboso» y de $-0,94$ en una de «cubierto».

En cambio, si nos limitamos a manejar la mera diferencia de índices RPS , tendríamos un índice RPS_{pc} definido como:

$$RPS_{pc} = RPS_p - RPS_c \tag{3}$$

Este índice oscila entre -1 y $+1$, siendo la discusión sobre los signos idéntica a la realizada anteriormente para el caso del $RPSS$. Al estar acotado inferiormente posee mayor estabilidad y, al ser las cotas simétricas respecto de cero, adquiere mayor sentido la posibilidad de obtener valores medios. Además, en el ejemplo antes manejado de predicción de helada, este índice puntuaría muy bien el acierto en Aliaante ($RPS_{pc} = 0,98$) y sólo discretamente en Navacerrada ($RPS_{pc} = 0,02$). Por último, en el ejemplo de predicción de nubosidad antes citado, se comprueba, a la visra de la Tabla 1, que el índice RPS_{pc} está orientado plenamente de manera positiva. El inconveniente que presenta frente al $RPSS$ es que, aun en el caso de una predicción perfecta, puede quedarse lejos de la unidad, pues siempre se está restando el RPS_c ; el cual, por definición, es positivo.

Tabla 1

Valores de los índices RPS_p , RPS_c , $RPSS$ y RPS_{pc} , en una predicción de «poco nuboso» para Alcantarilla en octubre, según las categorías observadas

CATEGORIA OBSERVADA	INDICE			
	RPS_p	RPS_c	$RPSS$	RPS_{pc}
Despejado	0,78	0,67	0,35	0,12
P. nuboso	1,00	0,88	0,98	0,12
Nuboso	0,78	0,93	-2,14	-0,15
M. nuboso	0,53	0,84	-1,93	-0,31
Cubierto	0,28	0,63	-0,94	-0,35

En este trabajo simultanearemos ambos índices, extrayendo las consecuencias más importantes que se deriven de cada uno de ellos. También consideraremos el orden de los errores, pero sin abandonar los índices anteriores, para evitar premiar una excesiva tendencia de las predicciones hacia los valores centrales, o de bajo riesgo.

2.2. **Ámbito geográfico y temporal**

El objetivo inicial de la evaluación se limita a dos puntos de la Comunidad de Murcia, como son Alcantarilla (que puede ser representativo del interior no montañoso de la Región) y San Javier (representativo del litoral). Estos observatorios han de considerarse representativos de la variable nubosidad a nivel regional. Para otras variables, como la precipitación o la tormenta, en cuya predicción se expresan probabilidades de ocurrencia, sería necesario tomar una muestra temporal más amplia y estudiar si, a largo plazo, las probabilidades están siendo bien asignadas o, mejor aún, considerar una muestra territorial mayor y valorar si, para un día determinado, el porcentaje de observatorios en el cual se ha registrado el fenómeno se corresponde con la probabilidad expresada en la predicción para un lugar, tomado al azar, del ámbito territorial considerado. Para estas últimas variables, por simplicidad de nuestro esquema de verificación, a la hora de valorar una predicción como fallida, será condición necesaria, pero no suficiente, que ésta haya errado en las estaciones elegidas.

La limitación temporal que vamos a adoptar, para acotar el problema, será la de verificar únicamente la predicción realizada para el día siguiente, a las 09:45 h. l., en valores medios, lo cual simplifica el problema al ignorar la evolución temporal. El mayor inconveniente que presenta esta simplificación es que se pueda dar por buena una predicción que tuviese invertidos los términos de la evolución temporal, pero que acertase el valor medio.

2.3. **Variables verificadas**

El objetivo de una verificación global debería abarcar a todas las variables meteorológicas que figuran en los boletines de predicción, que son, por este orden, nubosidad, precipitación, tormenta, visibilidad, temperatura y viento. En principio, en este trabajo, nos limitaremos a las tres primeras, las cuales coinciden, ca-

sualmente, con las que podría verificar un usuario general de las predicciones que no dispusiese de ningún aparato de medida.

2.3.1. Nubosidad

La expresiones literales de la predicción de la nubosidad que aparecen en los boletines se agrupan, en función del valor medio previsto, en una de las 9 que figuran en la Tabla 2. Este conjunto es suficiente, dado que la precisión con la cual se mide la variable es de ± 1 octavo. Como indica dicha Tabla, las categorías observables son 5, de acuerdo con el *Manual*.

Tabla 2

Expresiones de la predicción de nubosidad y probabilidad asignada a las categorías observables

EXPRESION DE PREDICCIÓN (Nº y descripción literal)	PROBAB. CATEGORIAS OBSERVABLES				
	1: Desp.	2: P nub	3: Nub.	4: M nub	5: Cub.
0: Despejado.	0,8500	0,1500	0,0000	0,0000	0,0000
1: Poco nuboso o/a despejado.	0,4625	0,5000	0,0375	0,0000	0,0000
2: Poco nuboso.	0,0750	0,8500	0,0750	0,0000	0,0000
3: P. nub o/a nub, algunos i.n.	0,0375	0,4625	0,4625	0,0375	0,0000
4: Nuboso, intervalos nubosos.	0,0000	0,0750	0,8500	0,0750	0,0000
5: Nuboso o/a muy nuboso.	0,0000	0,0375	0,4625	0,4625	0,0375
6: Muy nuboso.	0,0000	0,0000	0,0750	0,8500	0,0750
7: Muy nuboso o/a cubierto.	0,0000	0,0000	0,0375	0,5000	0,4625
8: Cubierto.	0,0000	0,0000	0,0000	0,1500	0,8500

Pero la estricta utilización de intervalos para el establecimiento de una predicción de categorías observables puede ser inapropiada. De hecho, aunque en las predicciones de nubosidad no se utilizan términos de probabilidad, es evidente que cuando se pronostica una categoría observable también se está indicando que, en caso de que la categoría observada no sea esa, es más probable que se observe una categoría cercana que una lejana. Resulta así aconsejable establecer un solape entre las expresiones, difuminando sus intervalos de definición, en el sentido de la lógica borrosa. Efectuaremos, de este modo, una interpretación probabilista de las expresiones de predicción en relación con las categorías observables, de acuerdo con lo indicado en la Tabla 2. De acuerdo con ella, cuando se realiza una predicción de «poco nuboso»), por ejemplo, se interpreta que la probabilidad de que la nubosidad inedia esté comprendida entre 0,5 y 3,5 octavos es de un 85%, quedando otro 15% restante a repartir, a partes iguales, entre las categorías adyacentes. El valor del 85% surge por analogía con lo establecido en el *Manual* para otras variables, para las cuales no debe hacerse mención de la probabilidad cuando ésta esté comprendida entre el 70 y el 100%.

Tanto la utilización eventual de expresiones no cuantificadas (nubosidad de evolución, de estancamiento, etc.) como el hecho de que los archivos climatológicos no proporcionen la cantidad de nubes altas (exentas de citación en la predicción cuando suponen una fracción menor que 6 octavos), son detalles que podrían modificar algo los resultados de una verificación. Por otro lado, existe el problema de que las climatologías existentes establecen unas categorías observables que no se corresponden con las adoptadas para las predicciones, lo que obliga a un trabajo previo de tipo climatológico.

2.3.2. *Precipitación y tormenta*

En la verificación de intensidades de precipitación surge el problema de que la intensidad que se predice está referida a la precipitación registrada en una hora, mientras que la que figura de manera rutinaria en los archivos climatológicos se basa en la intensidad máxima en diez minutos, sin que existan tampoco climatologías de intensidades. A falta de un trabajo más concienzudo, utilizando información procedente de estaciones automáticas, nos limitaremos simplemente a la verificación de la ocurrencia del fenómeno.

Cada expresión literal de predicción de la precipitación es clasificada en una de las 10 que figuran en la Tabla 3, la cual nos da también la interpretación probabilista de cada una de ellas en relación con las categorías observables. Para el cálculo de los valores de la probabilidad se han adoptado los criterios recogidos en el *Manual*, multiplicando los valores centrales de los intervalos de probabilidad asignados a las diversas expresiones, ya fuesen de probabilidad o de distribución espacial. Así, por ejemplo, la ((probabilidad de precipitaciones)) (comprendida entre el 40 y el 70%) «aisladas» (entre el 10 y el 30%), se interpreta globalmente como una probabilidad del 11% de que tenga lugar la precipitación en el lugar elegido. Se ha considerado que frases como «se esperan»), «más probables» o «preferentemente» no aportan información evaluable, que «alguno/s» es equivalente a «aislado/s» y que términos como «ocasional» no afectan a la cuantificación, pues describen la evolución temporal, no la distribución geográfica.

Tabla 3

Expresiones de predicción de la precipitación y probabilidad asignada a cada categoría observable

· En el caso de la tormenta también surgen problemas a la hora de verificar la adjetivación, pues no se dispone del número de relámpagos ni del tamaño del granizo en los archivos climatológicos convencionales. También nos limitaremos a verificar la ocurrencia del fenómeno tormentoso, siendo válido lo indicado anteriormente para las precipitaciones, sustituyendo el fenómeno por el de tormenta, con una interpretación probabilista análoga a la de la Tabla 3.

3. Resultados y discusión

El proceso de verificación descrito se ha aplicado, a título experimental, a las predicciones para la provincia de Murcia durante el segundo semestre de 1995, una vez codificadas de acuerdo con todo lo anterior. Como referencia estandarizada se ha recurrido a la climatología, incorporando los valores climatológicos mensuales de cada variable.

Para establecer los días con posible fallo significativo de la predicción, ha sido preciso fijar previamente unos umbrales de los índices de verificación. En el caso de la nubosidad, el valor del índice RPSS no era monótonamente decreciente con el orden del fallo, pero tampoco se ha podido encontrar un umbral adecuado a cualquier época para el índice RPS_{pc} , fijándose un criterio basado en el orden del fallo, adoptándose el orden segundo o superior y completándolo con un seguimiento de los valores medios del RPS_{pc} , para evitar dar un tratamiento favorable a los pronósticos «centrales», sujetos a escaso riesgo de error significativo. Los resultados obtenidos en los observatorios seleccionados podrán considerarse como representativos a nivel regional.

Tabla 4

Expresiones de predicción seleccionadas por los umbrales del índice RPS_{pc} , según las categorías observadas y las probabilidades climatológicas mensuales extremas

PROBAB.	VARIABLE			
	Precip.		Tormenta	
	1. Sí	2. No	1. Sí	2. No
Mínima	0	7-9	0	2-9
Máxima	0-5	8-9	0-1	2-9

En los casos de la precipitación y de la tormenta, tras un análisis de las combinaciones posibles entre las expresiones de predicción y las categorías observables, se llega a la conclusión de que, para el índice RPS_{pc} , los umbrales $-0,09$ y $-0,005$, respectivamente, poseen un interés especial. Con esos umbrales, en la Tabla 4 figuran las expresiones de predicción que serían seleccionadas como posiblemente erróneas, dependiendo de la ocurrencia o no del fenómeno y de las probabilidades climatológicas mensuales extremas. Así, por ejemplo, si se ha registrado precipitación, se seleccionarían sólo los días sin su mención en la predicción (es decir, con la expresión 0) en el mes de menor probabilidad de lluvia y la utilización de expresiones que quizá sean demasiado débiles (expresiones 0 a 5) en el mes de mayor probabilidad. El resto de los meses estaría

comprendido entre estos dos casos extremos. Nótese que la utilización de la expresión 6, no contemplada en los casos precedentes, no llegaría a cuestionar nunca la predicción. Sin embargo, el abuso de dicha expresión tendría repercusión en la valoración de la calidad global en rangos superiores al diario, al manejar valores medios de los índices. Finalmente, la confirmación del fallo a un nivel geográfico más amplio que el local se obtiene tras el análisis de lo sucedido en el resto de las estaciones pluviométricas.

El listado mensual presentado en la Tabla 5 es un ejemplo de las salidas proporcionadas por el esquema de verificación para el observatorio de San Javier en el mes de noviembre. Se indican, para cada día, la

Tabla 5

Salida mensual del esquema de verificación, correspondiente al mes de noviembre, para el observatorio de San Javier

san1195											
DIA	NUBOSIDAD				PRECIPITACION				TORMENTA		
	EP	CO	RPSpc	RPSS	EP	CO	RPSpc	RPSS	EP	CO	RPSS
1	3	2	0.05	0.43	0	2	0.07	1.00	0	2	1.00
2	3	3	0.01	0.16	0	2	0.07	1.00	0	2	1.00
3	3	2	0.05	0.43	0	2	0.07	1.00	0	2	1.00
4	2	2	0.11	0.97	0	2	0.07	1.00	0	2	1.00
5	3	3	0.01	0.16	0	1	-0.45*	-0.83	0	2	1.00
6	3	3	0.01	0.16	0	1	-0.45*	-0.83	0	2	1.00
7	3	5	-0.16*	-0.43	0	2	0.07	1.00	0	2	1.00
30	4	3	0.07	0.96	1	1	-0.35'	-0.65	0	1	-0.05*
UMBRALES RPSS N° de días RPSS > 0.8: 6 RPSS > 0.0: 16 RPSS <= 0.0: 14 MEDIA RPSpc Signo: UMBRALES (*) Nubosidad (Orden fallo >=2): 7, 23, Precipitación (RPSpc <= -0.090): 5, 6, 12, 25, 30, Tormenta (RPSpc <= -0.005): 30, (*): días con posible fallo significativo											

expresión de predicción EP, la categoría observada CO y los valores de los índices RPS_{pc} y $RPSS$. En el resumen final se incluye el número de días en los cuales el índice $RPSS$ rebasa determinados umbrales: 0,8 (para señalar las predicciones prácticamente inmejorables) y 0,0 (para las predicciones que han resultado ser mejor o peor que la referencia climatológica), así como el signo del valor medio mensual de RPS_{pc} , para posibilitar un seguimiento de la calidad a más largo plazo. Por último, aparece una relación de días con posible fallo significativo de la predicción, en base a los umbrales elegidos para el RPS_{pc} . Continuando con el ejemplo del mes de noviembre, en la Tabla 6 se presenta un listado de las fechas con fallo significativo de la predicción a nivel regional, especificando la predicción realizada y la categoría observada.

Tabla 6

Fechas con fallo significativo de la predicción a nivel regional, correspondientes al mes de noviembre

En la Tabla 8 figuran los signos del valor medio del índice RPS_{pc} para cada uno de los meses considerados. Se vuelve a apreciar que la variable más susceptible de mejorar en el pronóstico es la nubosidad, ya que en el interior dicho valor sólo ha sido positivo en el mes de agosto. Por el contrario, destaca como buena predicción la de la precipitación en el litoral, para la cual sólo nos encontramos con valor negativo en el mes de noviembre.

Tabla 8

Signos del valor medio del índice RPS_{pc} para los meses considerados, en el interior-o el litoral, para cada variable verificada

MES	VARIABLE					
	Nubosidad		Precipitación		Tormenta	
	Interior	Litoral	Interior	Litoral	Interior	Litoral
Julio	-	-	-	+	-	-
Agosto	+			+	+	+
Septiembre	-	+	+	+	+	+
Octubre	-	+	-	+		
Noviembre			+		+	
Diciembre			+	+	+	+
TOTAL SIGNOS +	1	2	3	5	4	3

4. Conclusiones

1. Se presenta un método de verificación de predicciones generales basado en índices derivados del RPS. El RPSS presenta algunos inconvenientes a la hora de medir la pericia, por lo que se ha adoptado también otro índice RPS_{pc} , relacionado con el anterior, pero que parece ser más adecuado.

2. El método de verificación se ha limitado a los valores medios diarios de la nubosidad y a la ocurrencia o no de los fenómenos de la precipitación y de la tormenta. Las expresiones de predicción son agrupadas en un conjunto de clases y se equiparan con una probabilidad de ocurrencia de las categorías observables. En el caso de la nubosidad, además, se han difuminado los intervalos de definición, para establecer una interpretación probabilista multicategoría de las expresiones de predicción.

3. Para decidir la existencia de un fallo significativo de la predicción de nubosidad se ha adoptado el criterio de que dicho fallo alcance el segundo orden, mientras que para la precipitación y la tormenta se han adoptado, para el índice RPS_{pc} , los umbrales $-0,09$ y $-0,005$, respectivamente.

4. El proceso de verificación se ha aplicado a las predicciones para la provincia de Murcia para el día siguiente, durante el segundo semestre de 1995, utilizando inicialmente las observaciones de las estaciones de Alcantarilla, en el interior, y de San Javier, en el litoral, suficientemente representativas de la nubosidad. En el caso de la precipitación y la tormenta, la posibilidad de fallo a nivel regional debe ser confirmada posteriormente con información del resto de la red pluviométrica. Mensualmente se genera un listado útil para localizar días con posible fallo significativo y un resumen, así como una relación de fechas con fallo confirmado a nivel regional.

5. En el período de tiempo considerado, se ha comprobado que la predicción de nubosidad es apenas mejor que la referencia climatológica. Por el contrario, en el caso de la precipitación y la tormenta, la predicción es mejor en un 80% de los días, aproximadamente. También el mayor número de días con fallo significativo lo presenta la nubosidad, alcanzando el 10% en el interior, reduciéndose hasta valores del 1% en el caso de la tormenta.

6. Los resultados cuando se maneja el valor medio mensual del índice RPS_{pc} son similares, destacando los ineducados resultados de la nubosidad y los buenos de la precipitación en el litoral.

7. Los objetivos del esquema de verificación son ampliables (más variables, mayor resolución espacio-temporal, mayor discriminación de categorías, etc.), posibilitando la existencia de una nueva información de retorno al GPV y facilitando el control y seguimiento de la calidad de las predicciones.

Referencias

Brier, G. W. y R. A. Allen, 1951: *Verification of weather forecast. Compendium of Meteorology*, pp. 841-848.

García-Moya, J. A.; A. Genovés; B. Garcías y A. Jansá, 1990: *Resultados de la verificación sistemática de las alertas en el GPV de Baleares durante 1990. Segundo Simp. Nac. de Predicción del INM*, pp. 399-423.

Garcías, B. y J. A. Guijarro, 1995: *Problemática de la evaluación de las predicciones meteorológicas expresadas en lenguaje claro. Centro Meteor. Territ. de Baleares. Secc. de Estudios y Desarrollos, NT núm. 1*, 14 pp.

Murphy, A. H. y E. S. Epstein, 1967: *Verification of Probabilistic Predictions: A Brief Review. J. Appl. Meteor.*, 6, pp. 748-755.

Sanders, F., 1967: *The Verification of Probability Forecasts. J. Appl. Meteor.*, 6, pp. 756-761

INM. 1992: *Estilo. Manual de términos meteorológicos. Inst. Nac. de Meteor.*, 29 pp.

Stanski, H. R. y otros, 1989: *Survey of common verification methods in meteorology. WMO World Weather Watch Techn. Rep. no. 8, WMD/TD no. 358*, 113 pp.

Agradecimientos

El autor agradece las sugerencias aportadas por Francisco Martínez, del CMT de Murcia, así como la información bibliográfica proporcionada por José A. García-Moya, del Servicio de Predicción Numérica (INM)