

ENFOQUE PROBABILÍSTICO PARA LA MEJORA DE LAS ESTIMACIONES DE RACHAS MÁXIMAS ALTAS MEDIANTE KRIGEADO UNIVERSAL

M. Rodrigo ^{(1) (2)}, J. A. López ⁽²⁾

⁽¹⁾ Consorcio de Compensación de Seguros (CCS), Paseo de la Castellana, 32, 28046, Madrid (España), ccseguros@aemet.es

⁽²⁾ Agencia Estatal de Meteorología (AEMET), Calle Leonardo Prieto Castro, 8, 28040, Madrid (España), jlopezd@aemet.es

Introducción

El Consorcio de Compensación de Seguros, CCS, compensa los daños producidos a las personas y en los bienes por determinados fenómenos de la naturaleza. Uno de estos fenómenos se refiere a la TCA, Tempestad Ciclónica Atípica, en la que quedan incluidos los tornados y los vientos extraordinarios (rachas superiores a 120 km/h), de acuerdo con el Reglamento del Seguro de Riesgos Extraordinarios en la forma que se definen por el Consorcio de Compensación de Seguros (2012)

Para la delimitación geográfica del área de afectación del fenómeno señalado debido a vientos extremos la Agencia Estatal de Meteorología, AEMET, aplica las técnicas de interpolación geoestadísticas más adecuadas utilizando las mediciones existentes y de forma que se procure la mayor homogeneidad posible en la definición del área. La AEMET es la encargada de elaborar los informes técnicos que concretan dicha zona geográfica afectada por viento extremo, susceptible de cobertura por el CCS.

La delimitación precisa de las zonas en que se cumple la condición de TCA presenta considerables dificultades por diversos motivos como pueden ser la escasez de observaciones y la complejidad de la orografía.

Cuando los datos son abundantes, muchas de las técnicas de interpolación dan resultados similares. Sin embargo, cuando los datos son escasos la elección del método y sus parámetros puede ser crítico y se debe evitar llegar a resultados engañosos.

Para estimar las zonas en que se ha superado la racha de viento de 120 km/h se ha desarrollado una técnica de interpolación geoestadística en AEMET que es un tipo de krigeado en el que la estimación en un punto cualquiera es una combinación lineal de las observaciones en el resto de puntos debidamente ponderadas (se busca que el estimador sea insesgado y que presente varianza mínima).

Los métodos geoestadísticos proporcionan una gran flexibilidad para la interpolación, dentro de estos la técnica del krigeado universal es la elegida para las estimaciones de las áreas de rachas máximas de viento, se basa en un modelo de tendencia lineal en el que entran en juego diferentes variables auxiliares (Burrough y MacDonnell, 1998).

El método operativo usado en AEMET utiliza este estimador que se apoya en los datos de observación de las estaciones de AEMET y de diferentes Servicios Meteorológicos Externos y toma como variables externas la elevación del terreno, la distancia a la costa y las salidas del modelo HIRLAM 0,05 (High Resolution Limited Area Model) de alcance más corto del campo de racha máxima, representadas en la Fig. 1. De esta manera se obtienen efectos de variación suave en la interpolación.



Fig. 1.- Mapas para las variables externas: elevación terreno (izq.), distancia a la costa (centro) y ejemplo de las salidas del modelo HIRLAM del campo de racha máxima (dcha.)

Señalar que la estructura del semivariograma usada se especifica sin “efecto pepita” de tal forma que el estimador del krigeado es exacto, esto es, en los puntos con observación el valor estimado coincide con el observado. Esto es importante para el CCS puesto que es necesario respetar al máximo los valores observados ya que son los que tienen mayor fuerza legal.

Motivación

Los valores estimados en cada punto mediante el método operativo descrito anteriormente se interpretan como el valor medio o esperado de la función de densidad de probabilidad de la racha máxima de viento estimada en ese punto.

Con objeto de valorar las estimaciones que se realizan mediante este método se hicieron validaciones en el rango de velocidades altas de viento con objeto de mejorar la estimación de rachas fuertes.

Las validaciones consisten básicamente en seleccionar una serie de estaciones. Esta selección se desarrolla mediante R, lenguaje y entorno de programación para análisis estadístico y gráfico (R Core Team, 2013). Estas estaciones no se tendrán en cuenta en el momento de realizar la estimación para posteriormente poder comparar el resultado de la estimación en la estación con su valor real.

Para efectuar las validaciones se distinguen 15 situaciones de TCA que pretenden abarcar fenómenos ocurridos en diferentes áreas geográficas y con extensiones, en cuanto a su repercusión, de diversas dimensiones.

Las estaciones que serán validadas se seleccionan de manera aleatoria y sin reemplazo. Se escogen un 20% del total de las estaciones de observación, para cada una de las situaciones. Esta selección aplica el porcentaje por zonas, para asegurar una cobertura espacial y que todas las provincias tengan representatividad en los resultados, y por un límite en el valor de racha máxima (se toma como referencia el valor de 80 km/h), para asegurar un número de estaciones para validar con valores de racha máxima altos, pues son los de nuestro interés principal.

Estas validaciones se realizan utilizando el procedimiento operativo que consiste en realizar un krigeado universal (con las variables externas ya mencionadas) mediante el Sistema de Información Geográfica SAGA (System for Automated Geoscientific Analyses).

Para evaluar los resultados de las validaciones se calculan los sesgos medios (Ecuación 1) y los errores relativos medios (Ecuación 2) teniendo en cuenta que la diferencia existente resta el valor estimado del observado, necesario para la interpretación del sesgo.

$$Sesgo\ medio = RM^{Est} - RM^{Obs} \quad (1)$$

$$Error\ relativo\ medio = \frac{|RM^{Est} - RM^{Obs}|}{RM^{Obs}} \times 100 \quad (2)$$

dónde RM es el valor de racha máxima observado (superíndice *Obs*) o estimado (superíndice *Est*), ambos en km/h.

Los resultados de las primeras validaciones que utilizan el SAGA y se aplican a 15 situaciones de TCA diferentes se muestran en la Tabla 1 y Tabla 2.

TODAS ESTACIONES			
PROCEDIMIENTO OPERATIVO (SAGA)	PROMEDIOS		
	Sesgos (km/h)	E. Relat. (%)	
TCA (fecha)	20100113a15	1	23
	20100227a28	1	18
	20111023a27	-1	18
	20111112a14	-1	18
	20111215a17	-1	18
	20120105a08	0	23
	20120202a05	0	23
	20120206a08	-1	20
	20120415a17	-1	17
	20120423a26	0	16
	20121027a28	0	20
	20121124a25	0	23
	20121213a16	0	20
	20130118a20	1	19
	20130123a24	2	18
PROMEDIO		0 km/h	20%

Tabla 1.- Resultado de las validaciones para el procedimiento operativo (SAGA) utilizando todas las estaciones

ESTACIONES RM ≥ 80 km/h			
PROCEDIMIENTO OPERATIVO (SAGA)	PROMEDIOS		
	Sesgos (km/h)	E. Relat. (%)	
TCA (fecha)	20100113a15	-7	16
	20100227a28	-4	16
	20111023a27	-17	21
	20111112a14	-13	19
	20111215a17	-8	16
	20120105a08	-20	24
	20120202a05	-22	24
	20120206a08	-13	20
	20120415a17	-17	20
	20120423a26	-15	19
	20121027a28	-13	22
	20121124a25	-10	17
	20121213a16	-18	19
	20130118a20	-7	16
	20130123a24	-13	18
PROMEDIO		-13 km/h	19%

Tabla 2.- Resultado de las validaciones para el procedimiento operativo (SAGA) tomando las estaciones con racha máxima superior a 80 km/h

El sesgo que se obtiene al tener en cuenta todas las estaciones de validación es nulo y el error relativo medio presenta un valor bastante elevado que da idea de la complicación que supone el cálculo de esta variable de racha máxima indicando que la técnica tampoco ofrece demasiada precisión.

Sin embargo, para los valores altos de racha máxima, tomando las estaciones de validación que presentan valores superiores a 80 km/h, existe un importante y sistemático sesgo negativo, dando un valor medio, para las 15 situaciones de TCA de unos -13 km/h. El error relativo que se obtiene es del mismo orden que el anterior.

Este sesgo negativo se explica por el hecho de que, como se explicó anteriormente, el krigeado produce un estimador insesgado globalmente, pero sobre un rango particular de observaciones acerca el estimador a la media. Como se seleccionó un rango de valores de las observaciones más alto que la media es de esperar un sesgo negativo.

Por tanto, las validaciones realizadas con este procedimiento muestran un sesgo negativo para velocidades altas de racha máxima. La técnica empleada, por tanto, subestima la racha máxima de viento para el rango de interés (Rodrigo y López, 2013).

Desarrollo

El objetivo de este trabajo consiste en presentar una técnica que corrige el sesgo de las velocidades altas de racha máxima.

Para realizar esta mejora se abre esta línea de investigación basada en un enfoque probabilístico que, de hecho, se presenta ya de forma implícita en la base de la actual técnica operativa puesto que las estimaciones actuales son en realidad las estimaciones de la media del campo de vientos en cada punto.

Esta técnica se basa en la descripción de la distribución de probabilidad mediante los parámetros estimados de dicha distribución, su valor medio y su desviación estándar, para buscar de esta manera los percentiles que corrigen el sesgo en las estimaciones correspondientes a las observaciones con velocidades de rachas máximas superiores a 80 km/h. Esto abre la posibilidad de buscar el percentil que corrija de forma óptima el error sistemático y pasar del percentil 50, que es la media de la distribución y por tanto es la estimación actual, a otro percentil mayor.

El percentil asociado a esta probabilidad dada se calcula mediante la Ecuación 3 siendo $perc0$ el percentil para la misma probabilidad de la variable normal estándar, $PRED$ el valor medio y SE la desviación estándar.

$$genPerc = PRED + (0.01 \times perc0) \times SE \quad (3)$$

Por tanto, para tener un conocimiento más detallado de la distribución de probabilidad no basta con

conocer las medidas de tendencia central sino que también es necesario conocer las medidas de dispersión, para ello se han calculado las desviaciones estándar (Cressie, 1991).

Con el procedimiento utilizado hasta el momento, en el que se utiliza la herramienta SAGA para obtener dichas estimaciones, no existe la posibilidad de obtener estas medidas de dispersión pues sólo realiza el cálculo para el valor esperado, medida de centralidad.

La técnica empleada en este estudio aplica la función krige desarrollada en R que extrae tanto las medidas centrales como las de dispersión.

Todo el desarrollo permitirá encontrar las denominadas probabilidades correctivas que se corresponderán a aquellos percentiles que anulan el sesgo para las validaciones teniendo en cuenta sólo las estaciones con racha máxima superior a 80 km/h.

Estudio

El estudio incluye el análisis de una serie de casos de TCA, en concreto se han seleccionado 15 situaciones que presentan diversas características. Hay que prestar especial interés a aquellas que presentan mayores extensiones puesto que son las de mayor impacto económico para el CCS (del orden de 10 veces mayor). Dentro de estas señalamos:

Floora: 13 al 15 de enero de 2010 (nº 1)
Xynthia: 27 al 28 de febrero de 2010 (nº 2)
15 al 17 de diciembre de 2011 (nº 5)
Gong: 18 al 20 de enero de 2013 (nº 14)

Para desarrollar el trabajo la herramienta fundamental utilizada es el lenguaje y entorno de programación R. Mediante este lenguaje se ha programado la técnica operativa, que usa un krigeado universal y que permite obtener no sólo las estimaciones propias de la variable de interés, la racha máxima, sino una estimación de la desviación estándar, variables necesarias para el cálculo de las probabilidades correctivas.

Además, esta técnica ofrece diferentes posibilidades y se aplica utilizando diferentes parámetros en el semivariograma y opciones en el propio krigeado.

El semivariograma, similar a la covarianza sin centrar, es la función que caracteriza la estructura de momentos de segundo orden mediante un solo argumento (espacial). Para la técnica geostadística considerada se cumple la estacionariedad, es decir, invarianza de los momentos de segundo orden frente a desplazamientos.

El kriging universal en R trabaja con los residuos y estima la matriz de covarianzas (o semivariograma) trabajando con estos residuos mientras que en SAGA no se utilizan.

El estudio incluye 4 casos diferenciados por un lado por la aplicación del semivariograma (incluyendo sólo la racha máxima o añadiendo todas las variables externas) y por otro lado por la función krige (imponiendo o no un número máximo de estaciones a tener en cuenta en la interpolación local). Se obtienen los siguientes casos:

- CASO 1 (C1):
Semivariograma todas las variables
Krige con número máximo de estaciones
- CASO 2 (C2):
Semivariograma sólo racha máxima
Krige con número máximo de estaciones
- CASO 3 (C3):
Semivariograma todas las variables
Krige sin número máximo de estaciones
- CASO 4 (C4):
Semivariograma sólo racha máxima
Krige sin número máximo de estaciones

Para cada una de estas TCA se generan los cálculos de las estimaciones y los errores estándar utilizando la función señalada y fijando los diferentes parámetros dentro de esta.

En la Fig.2 el resultado de los mapas obtenidos para uno de los casos en el que se muestra el mapa final de estimaciones de racha máxima y los errores estándar asociados que se obtienen de manera automática mediante el programa generado (Bivand et al., 2008).

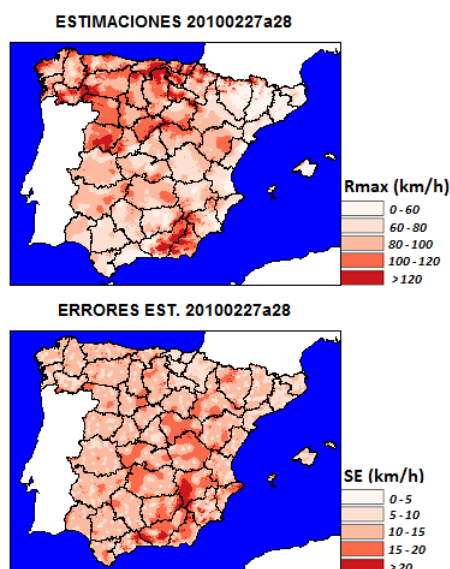


Fig. 2.- Mapa para las estimaciones y errores estándar de rachas máximas de viento de la situación del 27 al 28 de febrero de 2010.

Resultados

Se realizaron unas primeras validaciones para la técnica operativa utilizando R teniendo en cuenta los diferentes casos, expuestos anteriormente, que entran en juego al variar los diferentes parámetros.

Los resultados obtenidos muestran valores casi idénticos para este procedimiento utilizando indistintamente cualquiera de las dos herramientas de trabajo (SAGA y R).

Además, los resultados analizados para los diferentes casos de R dan también resultados similares.

En la Tabla 3 y Tabla 4 se exponen los resultados teniendo en cuenta todas las estaciones.

TODAS ESTACIONES		SESGOS (km/h)				
		SAGA	R C1	R C2	R C3	R C4
TCA (fecha)	20100113a15	1	1	1	1	1
	20100227a28	1	0	0	0	0
	20111023a27	-1	-1	-1	0	0
	20111112a14	-1	-1	-1	-1	-1
	20111215a17	-1	-1	-1	-1	-1
	20120105a08	0	1	0	1	0
	20120202a05	0	0	0	0	0
	20120206a08	-1	-1	-1	-1	-1
	20120415a17	-1	-1	-1	0	0
	20120423a26	0	0	0	0	-1
	20121027a28	0	1	0	0	0
	20121124a25	0	-1	-1	-1	-1
	20121213a16	0	0	0	0	0
	20130118a20	1	1	1	1	1
	20130123a24	2	2	2	1	1
PROMEDIO		0	0	0	0	0

Tabla 3.- Resultado de las validaciones, sesgos medios, para el procedimiento operativo (SAGA vs R) utilizando todas las estaciones y para los diferentes casos

Nuestro interés se centra en las validaciones realizadas para los valores de racha máxima elevadas, Tabla 5 y Tabla 6.

Para cualquiera de los resultados mostrados se aprecia un sesgo sistemático negativo aunque varía según el tipo de TCA, para las más extensas (señaladas anteriormente) se presenta un sesgo menor.

Según el enfoque probabilístico en que se basa nuestro estudio, calculamos los sesgos y los errores relativos para diferentes percentiles a partir de los valores estimados y de su desviación estándar, es decir, utilizando el kriging en R. Para este cálculo se han tenido en cuenta el conjunto de las 15 TCA en estudio (sin diferenciarlas según su extensión), Tabla 7 y Tabla 8.

TODAS ESTACIONES		E. RELATIVOS (%)				
		SAGA	R C1	R C2	R C3	R C4
TCA (fecha)	20100113a15	23	23	23	22	21
	20100227a28	18	18	18	17	17
	20111023a27	18	18	19	17	18
	20111112a14	18	18	18	17	17
	20111215a17	18	18	18	17	18
	20120105a08	23	22	22	22	21
	20120202a05	23	21	22	22	22
	20120206a08	20	19	19	18	18
	20120415a17	17	17	17	17	17
	20120423a26	16	16	17	16	16
	20121027a28	20	19	19	20	19
	20121124a25	23	24	25	23	24
	20121213a16	20	20	20	19	20
	20130118a20	19	18	18	17	17
	20130123a24	18	18	18	17	17
PROMEDIO	20	19	20	19	19	

Tabla 4.- Resultado de las validaciones, errores relativos medios, para el procedimiento operativo (SAGA vs R) utilizando todas las estaciones y para los diferentes casos

ESTACIONES RM ≥ 80 km/h		SESGOS (km/h)				
		SAGA	R C1	R C2	R C3	R C4
TCA (fecha)	20100113a15	-7	-8	-8	-8	-8
	20100227a28	-4	-5	-5	-5	-5
	20111023a27	-17	-17	-15	-16	-14
	20111112a14	-13	-13	-13	-14	-13
	20111215a17	-8	-8	-8	-9	-9
	20120105a08	-20	-18	-18	-20	-17
	20120202a05	-22	-22	-22	-23	-22
	20120206a08	-13	-15	-15	-15	-15
	20120415a17	-17	-18	-18	-16	-16
	20120423a26	-15	-17	-17	-17	-17
	20121027a28	-13	-14	-13	-19	-17
	20121124a25	-10	-12	-12	-14	-14
	20121213a16	-18	-18	-18	-17	-17
	20130118a20	-7	-9	-8	-10	-9
	20130123a24	-13	-13	-13	-15	-15
PROMEDIO	-13	-14	-14	-15	-14	

Tabla 5.- Resultado de las validaciones, sesgos medios, para el procedimiento operativo (SAGA vs R) tomando las estaciones con racha máxima superior a 80 km/h y para los diferentes casos

Fijándonos en las validaciones para rachas máximas elevadas observamos claramente la mejora que se produce en el sesgo a medida que tomamos percentiles mayores (e incluso llega a hacerse positivo). En cambio, al tener en cuenta todas las estaciones este sesgo va aumentando.

A percentiles mayores el aumento de los valores de racha máxima se produce para todo el rango de velocidades.

ESTACIONES RM ≥ 80 km/h		E. RELATIVOS (%)				
		SAGA	R C1	R C2	R C3	R C4
TCA (fecha)	20100113a15	16	16	16	15	15
	20100227a28	16	16	16	16	16
	20111023a27	21	20	19	18	19
	20111112a14	19	18	18	17	17
	20111215a17	16	15	15	14	14
	20120105a08	24	22	22	22	20
	20120202a05	24	23	23	23	23
	20120206a08	20	20	20	19	19
	20120415a17	20	20	20	18	18
	20120423a26	19	20	20	20	21
	20121027a28	22	21	21	23	24
	20121124a25	17	17	17	18	18
	20121213a16	19	19	19	18	19
	20130118a20	16	15	15	13	13
	20130123a24	18	18	18	17	17
PROMEDIO	19	19	19	18	18	

Tabla 6.- Resultado de las validaciones, errores relativos medios, para el procedimiento operativo (SAGA vs R) tomando las estaciones con racha máxima superior a 80 km/h y para los diferentes casos

ESTACIONES RM ≥ 80 km/h		SESGOS (km/h)				
		SAGA	R C1	R C2	R C3	R C4
TCA (fecha)	20100113a15	-7	-8	-8	-8	-8
	20100227a28	-4	-5	-5	-5	-5
	20111023a27	-17	-17	-15	-16	-14
	20111112a14	-13	-13	-13	-14	-13
	20111215a17	-8	-8	-8	-9	-9
	20120105a08	-20	-18	-18	-20	-17
	20120202a05	-22	-22	-22	-23	-22
	20120206a08	-13	-15	-15	-15	-15
	20120415a17	-17	-18	-18	-16	-16
	20120423a26	-15	-17	-17	-17	-17
	20121027a28	-13	-14	-13	-19	-17
	20121124a25	-10	-12	-12	-14	-14
	20121213a16	-18	-18	-18	-17	-17
	20130118a20	-7	-9	-8	-10	-9
	20130123a24	-13	-13	-13	-15	-15
PROMEDIO	-13	-14	-14	-15	-14	

Tabla 7.- Resultado de las validaciones, para el conjunto de todas las TCA, según el enfoque probabilístico en R (para unos percentiles dados) utilizando todas las estaciones y para los diferentes casos

Para tener un conocimiento más detallado de estos resultados, se realizan los cálculos de las validaciones para cada una de las TCA por separado. En la Tabla 9 y Tabla 10 se presentan los sesgos y errores relativos de nuestro interés, las que tienen en cuenta las estaciones con racha máxima superior a 80 km/h. Además, en vista de los resultados para los diferentes casos, que son similares, se toma el primero de ellos puesto que incluye un número máximo de estaciones en el krigeado local (al igual que hace SAGA) e incluye todas las variables en el semivariograma, dando más coherencia a la técnica.

El percentil 50 es la media de la distribución y por tanto coincide con los resultados de las validaciones

de la estimación actual (realizada en R para el caso 1), se corresponde con el valor central.

PROMEDIOS (TODAS TCA)		ESTACIONES RM \geq 80 km/h			
		R-C1	R-C2	R-C3	R-C4
SESGOS (km/h)	perc.50	-12	-12	-13	-12
	perc.55	-10	-10	-11	-11
	perc.60	-8	-9	-9	-9
	perc.65	-7	-7	-8	-8
	perc.70	-5	-5	-6	-6
	perc.80	0	-1	-2	-3
E. RELATIVOS (%)	perc.50	18	18	17	17
	perc.55	17	17	16	17
	perc.60	16	17	16	16
	perc.65	16	16	15	16
	perc.70	16	16	15	15
	perc.80	16	16	15	15
	perc.90	17	17	16	16

Tabla 8.- Resultado de las validaciones, para el conjunto de todas las TCA, según el enfoque probabilístico en R (para unos percentiles dados) tomando las estaciones con racha máxima superior a 80 km/h y para los diferentes casos

EST. \geq 80km/h	SESGOS (km/h)						
	Percentiles						
	50	55	60	65	70	80	90
TCA (n°)							
1	-8	-6	-4	-2	0	5	12
2	-5	-3	-2	0	2	6	12
3	-17	-15	-13	-11	-9	-5	2
4	-13	-12	-11	-10	-8	-5	-1
5	-8	-6	-5	-3	-1	4	10
6	-18	-16	-14	-12	-10	-5	2
7	-22	-19	-17	-15	-12	-6	2
8	-15	-13	-12	-10	-9	-5	0
9	-18	-16	-15	-13	-11	-8	-2
10	-17	-16	-14	-13	-11	-8	-3
11	-14	-12	-10	-8	-6	-1	5
12	-12	-11	-9	-8	-7	-4	0
13	-18	-17	-15	-14	-12	-9	-4
14	-9	-7	-4	-2	0	5	12
15	-13	-11	-10	-8	-6	-2	3
PROM.	-14	-12	-10	-9	-7	-3	3

Tabla 9.- Resultado de las validaciones, sesgos medios, para cada TCA, según el enfoque probabilístico en R (para unos percentiles dados) tomando las estaciones con racha máxima superior a 80 km/h y para el caso 1

Teniendo en cuenta las medias calculadas por TCA, el sesgo se anula entre los percentiles del 80% y 90%. Sin embargo, la anulación del sesgo depende del tipo de TCA de manera que las de mayor extensión se anulan para percentiles menores y viceversa.

El cálculo de los diferentes percentiles muestra como a medida que seleccionamos percentiles con valores más altos el campo de racha máxima se ve

afectado obteniendo a su vez unos valores mayores para todo el rango de velocidades.

EST. \geq 80km/h	E. RELATIVOS (%)						
	Percentiles						
	50	55	60	65	70	80	90
TCA (n°)							
1	16	16	15	15	15	16	19
2	16	16	16	16	16	17	19
3	20	19	18	17	17	16	16
4	18	18	17	17	17	17	17
5	15	15	14	14	14	15	18
6	22	21	19	18	17	14	13
7	23	21	20	19	18	17	19
8	20	19	18	17	17	16	17
9	20	19	18	17	16	14	13
10	20	19	18	17	17	16	17
11	21	20	19	18	17	15	16
12	17	16	16	15	15	14	13
13	19	17	17	16	15	13	11
14	15	14	14	14	15	16	19
15	18	17	17	17	17	17	19
PROM.	19	18	17	16	16	16	16

Tabla 10.- Resultado de las validaciones, errores relativos medios, para cada TCA, según el enfoque probabilístico en R (para unos percentiles dados) tomando las estaciones con racha máxima superior a 80 km/h y para el caso 1

En la Fig. 3, como ejemplo, podemos visualizar la clara diferencia que existe entre el campo de racha máxima dado por el percentil 50, que es la estimación actual, frente al campo obtenido con el percentil 90. Este tipo de mapas se han ido obteniendo con R mediante un proceso automatizado, que no permite SAGA, y dan idea de los campos de viento que se obtienen para cada una de las situaciones y percentiles.

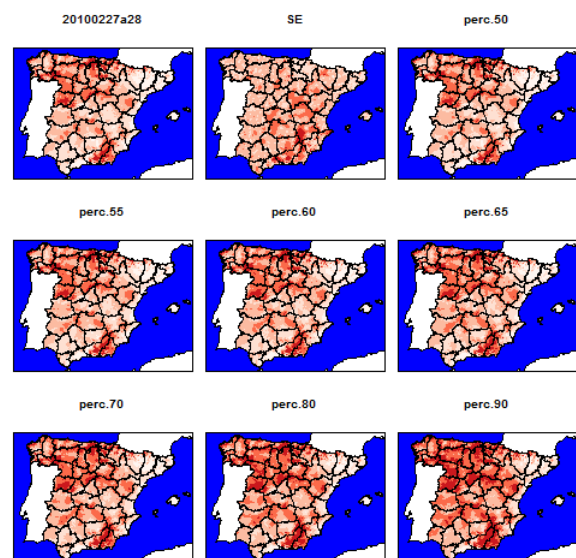


Fig. 3.- Mapa con la estimación central, el error estándar y diferentes percentiles para la racha máxima de viento de la situación del 27 al 28 de febrero de 2010.

Las validaciones realizadas con la función krige sobre el conjunto de casos de TCA muestran que los sesgos, para aquellas estaciones que presentan rachas máximas de viento elevadas, superiores a 80 km/h, mejoran a medida que tomamos percentiles mayores.

De esta manera se calculan las probabilidades correctivas, definidas como aquellas probabilidades en las que se anula el sesgo, Tabla 11.

ESTACIONES RM ≥ 80 km/h		PROBABILIDADES CORRECTIVAS (%)			
		C1	C2	C3	C4
TCA (fecha)	20100113a15	69	70	71	72
	20100227a28	65	66	66	67
	20111023a27	88	93	89	94
	20111112a14	91	98	94	99
	20111215a17	72	77	74	80
	20120105a08	87	93	92	93
	20120202a05	88	88	91	90
	20120206a08	89	86	91	88
	20120415a17	93	92	94	92
	20120423a26	94	99	95	100
	20121027a28	83	84	92	91
	20121124a25	89	97	94	99
	20121213a16	95	98	95	98
	20130118a20	70	70	75	74
	20130123a24	85	83	90	88
PROMEDIO		84	86	87	88

Tabla 11.- Probabilidades correctivas para cada una de las 15 TCA y para los diferentes casos

Si nos limitamos al estudio del caso 1, como hemos hecho anteriormente, podemos encontrar que la probabilidad correctiva mínima para las 15 TCA es de un 65%, la máxima un 95%, la media un 84% y la probabilidad correctiva combinando todas las TCA un 81%, Tabla 12.

PROBABILIDADES CORRECTIVAS (caso 1)			
mínima	máxima	media	conjunta
65%	95%	84%	81%

Tabla 12.- Probabilidades correctivas para 15 TCA, caso 1

La característica principal que presentan estos valores de las probabilidades correctivas es que son menores a medida que la TCA es más extensa.

Dentro de las TCA citadas, por ejemplo para el caso 1, tenemos un 69% para Floora, un 65% para Xynthia, un 72% para el número 5 y un 70% para Gong, todas estas probabilidades se corresponden con los valores mínimos. Del mismo modo ocurre con el sesgo que presenta valores menores cuánto más extensa es la TCA.

La dependencia encontrada entre dichas probabilidades y la extensión de la TCA, de forma

que para las TCA más fuertes los valores de la corrección son menores y viceversa, se muestra en la Fig. 4

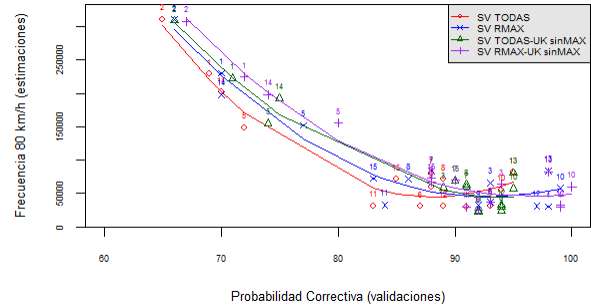


Fig. 4.- Dependencia probabilidad correctiva y extensión TCA para los diferentes casos con ajuste cuadrático

Para evaluar la extensión de la TCA y poder visualizar esta dependencia se han calculado las áreas que resultan de las estimaciones que presenten valores de racha máxima superiores a 80 km/h (representado en el eje de ordenadas).

La numeración de las TCA más extensas que habíamos descrito hasta ahora se corresponde con los números 1, 2, 5 y 14.

Claramente observamos la distinción que existe en dos grupos. Por un lado tenemos las más extensas con probabilidades correctivas menores y por otro lado las pequeñas cuyas probabilidades correctivas son las mayores.

En la Fig. 5 puede verse con más claridad teniendo en cuenta sólo el caso 1. Además, en la Fig. 6 se representa la dependencia con el sesgo en la cual, lógicamente, se observa la misma separación entre TCA.

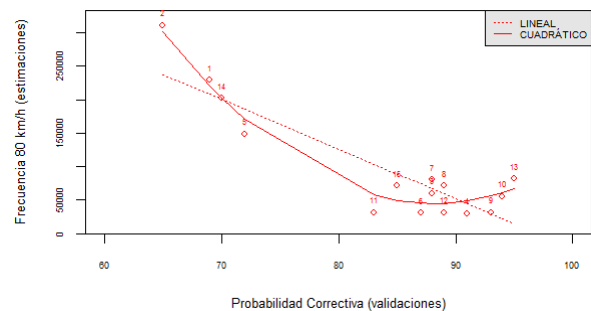


Fig. 5.- Dependencia probabilidad correctiva y extensión TCA para el caso 1 con ajuste cuadrático (línea continua) y ajuste lineal (línea discontinua)

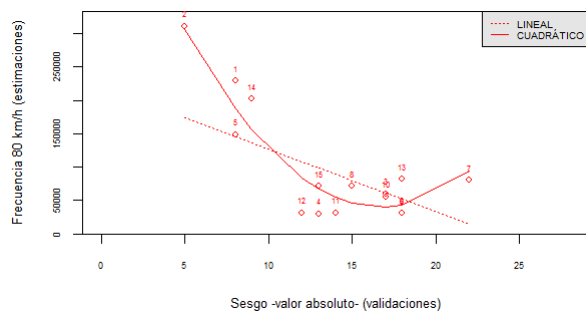


Fig. 6.- Dependencia sesgo medio y extensión TCA para el caso 1 con ajuste cuadrático (línea continua) y ajuste lineal (línea discontinua)

Conclusiones

Una de las principales ventajas que se presentan al realizar las estimaciones de racha máxima con R, en lugar de utilizar el Sistema de Información SAGA como se viene haciendo hasta el momento, es que el proceso queda totalmente automatizado.

Por otro lado permite una mayor flexibilidad en cuanto al establecimiento de los diferentes parámetros que se aplican en la técnica geoestadística, en especial, el hecho de poder incorporar en el semivariograma las variables externas puesto que hasta ahora no eran tenidas en cuenta.

Lo más relevante en cuanto al uso de R es que nos permite obtener las estimaciones de racha máxima mediante el enfoque probabilístico que hemos desarrollado a lo largo de este trabajo. R permite conocer la desviación típica de la distribución del viento en cada punto y por tanto dar una estimación basada, no sólo en la media estimada, sino también en cualquier percentil que se desee.

La técnica probabilística presentada tiene una base teórica sencilla y garantiza la total compatibilidad con la técnica actual de manera que las zonas delimitadas con la técnica probabilística (para la selección de cualquier percentil) engloban las zonas delimitadas con la técnica actual (percentil 50).

El aumento en los percentiles da lugar a un aumento en los valores de racha máxima para todo el rango de velocidades y permite obtener diversos escenarios con gran facilidad, presentando la ventaja, como acabamos de señalar, que los percentiles mayores siempre incluyen las áreas dadas por los percentiles menores.

Durante el estudio se obtienen las llamadas probabilidades correctivas, que son aquellas que anulan el sesgo, que presentan una dependencia con las características de la TCA de manera que cuánto más extensa sea, menor será la corrección.

Este análisis permitirá, por tanto, aplicar un criterio para el cálculo de las estimaciones de racha máxima de viento basándose en las probabilidades correctivas, buscando la probabilidad más óptima según las características que presente la situación de TCA

Agradecimientos. Este trabajo surge gracias a la colaboración entre la Agencia Estatal de Meteorología y el Consorcio de Compensación de Seguros, los cuales, junto a AT Sistemas, patrocinan el proyecto. Los autores desean agradecer el apoyo de muchas personas que pertenecen a estas organizaciones y que hacen posible estas investigaciones. Asimismo, agradecer a la AME (Asociación Meteorológica Española) por la concesión de la bolsa de viaje para jóvenes científicos en la celebración de las XXXIII Jornadas Científicas.

Referencias

Bivand, Roger S., Pebesma, Edzer J., Gómez-Rubio, V., 2008: Applied Spatial Data Analysis with R, Springer, 374 pp.

Burrough, P. A. y McDonnell, R. A., 1998: Principles of Geographical Information Systems, Oxford University Press, 333 pp.

Consorcio de Compensación de Seguros, 2012: Recopilación Legislativa, Edición de febrero de 2012, 202 pp.

Rodrigo, M. y López, J. A., 2013: Mejora del procedimiento operativo de AEMET para la estimación de las áreas con rachas máximas de viento, Tethys 2013, 10, 35-44 pp.

Cressie, Noel A. C., 1991: Statistics for spatial data, A Wiley – Interscience Publication, 900 pp.

R Core Team, 2013: R. A language and environment for statistical computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria.