

Improving the AEMET operational procedure for estimating areas of maximum wind gusts

M. Rodrigo^{1,2} and J. A. López²

¹Consortio de Compensación de Seguros (CCS), Paseo de la Castellana, 32, 28046, Madrid

²Agencia Estatal de Meteorología (AEMET), Calle Leonardo Prieto Castro, 8, 28040, Madrid

Received: 16-V-2013 – Accepted: 23-X-2013 – **Translated version**

Correspondence to: ccseguros@aemet.es

Abstract

The CCS (the Spanish Insurance Compensation Consortium) is the national agency that provides insurance coverage against weather events that involve an extraordinary risk. One of the extraordinary risks covered by the CCS refers to extraordinary wind, defined as wind with gusts exceeding 120 km h⁻¹. For about two years, the operational procedure performed in AEMET (the Spanish Meteorological Agency) for estimating the areas with maximum wind gusts has been using the technique of universal kriging interpolation based on observational data. External variables involved in the interpolation are the ground elevation, distance to the sea and the HIRLAM 0.05 model output of maximum gust field. The aim of the procedure is to delineate areas with maximum wind gusts that exceed 120 km h⁻¹. During previous research focused on the study of the accuracy given by the introduction of the HIRLAM model for this estimation technique, various validation analyses were conducted. These validations show a systematic negative bias for the estimation of high values of maximum gust, which implies an underestimation of the gusts through the operational procedure. This paper presents a new method of interpolation that provides a significant improvement. The bias is reduced by approximately 60% for stations that have maximum wind speeds of more than 80 km h⁻¹. The new methodology combines two interpolation fields. The first is obtained by applying the current operational method and includes all observational data. The second is obtained similarly, but using only the observation values of meteorological stations that have high values of maximum gust. The combination of both fields is based on a weighting given at each grid point, which depends on the overall density of the observations by region.

Key words: maximum gust, atypical cyclonic storm, extraordinary risk, interpolation methodology, validation

1 Introduction

Extreme meteorological episodes sometimes go unnoticed but frequently become apparent due to the scale of the circumstances, causing an impact on society with its corresponding consequences. Moreover, the catastrophic nature of a phenomenon depends not only on the extreme value the climate element takes on, but other characteristics also influence, such as population distribution or geomorphological features, among many others (García-Legaz and Valero, 2003).

The CCS is the national body which aims, as far as Extraordinary Risk Insurance is concerned, to compensate in the form stipulated by the relevant regulation, and as compensation regime, the losses arising from extraordinary events occurring in Spain that affect risks located in it.

For the purpose of covering extraordinary risks, one of the points mentioned refers to the ACS, Atypical Cyclonic Storm, extremely adverse weather caused by violent cyclones of a tropical nature, intense cold cyclones with Arctic air advection, tornadoes and extraordinary winds, as defined by the Consorcio de Compensación de Seguros (2012).

Table 1. Validations (biases and average relative errors) for the maximum gust estimation (MG) with the operational procedure in AEMET and taking only high observation values, equal to or greater than 80 km h^{-1} ($MG^{Obs} \geq 80 \text{ km h}^{-1}$).

ACS (Atypical Cyclonic Storm)	Average	
	Bias (km h^{-1})	Relative errors (%)
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
-13 km h^{-1}		19%

This legislative compilation classifies extraordinary winds as those with gusts exceeding 120 km h^{-1} , understanding a gust to be the greatest value of sustained wind speed for a three-second period.

The CCS asks the AEMET for reports on the possible existence of ACSs, in anticipation of the existence of the circumstances set out in the existing regulation, as extraordinary winds. Estimating the areas affected by this risk corresponds to AEMET, which has been using a geostatistical technique, universal kriging, which is supported, in addition to maximum gust observations, by physiographic variables and by fields of the HIRLAM weather prediction model. In this way, and once the CCS has the AEMET's final report, it will compensate the damage occurred to insured persons and goods, as established, in the areas where the winds showed, or very probably could have shown, gusts exceeding the aforementioned speed threshold.

During previous research, mainly focused on the study of the introduction of the HIRLAM model in this statistical technique, various validation analyses were conducted. To perform these verifications, the data available in the AEMET Climatology Database and in the External Meteorological Services of Catalonia, the Basque Country, Navarre, La Rioja and Galicia were used. The HIRLAM 0.05° outputs were also used, taking the highest value of the studied period from the maximum gust field of the four daily passes with least range. The external variables that will affect the interpolation are ground elevation and distance from the coast.

The validations are performed for specific ACS situations and on a number of randomly selected observation stations. This selection is fixed for each of the selected ACS cases and is used to validate the different procedures,

thus allowing the methods under study to be examined and possible improvements introduced to be evaluated.

The aim of this work is to improve and optimize the procedure that is currently operational at AEMET for estimating extraordinary winds in order to geographically limit the areas affected from the possible existence of ACS by maximum wind gust. The precise delimitation of the areas in which this condition is met shows considerable difficulties in our country, given the lack of wind observations and the complexity of the terrain. Moreover, the situation of our country in the middle latitudes (apart from the Canary Islands), surrounded by an ocean to the west and by the Mediterranean Sea to the east, with a highly complex terrain, makes the description of situations that could produce strong winds a difficult task.

The paper describes a modification of the technique used to estimate areas of extreme wind gusts that significantly improves the bias.

2 Operational methodology

The estimates made by AEMET to delineate areas in which a wind gust of 120 km h^{-1} has been exceeded use a geostatistical interpolation technique, the kriging method, which is based on considering the observations as a realization of a theoretical random field and which assumes that stationarity is met, i.e., invariance of second order momentums versus displacements. This allows characterizing the structure of second-order moments by a single argument function (spatial), which in kriging is the semivariogram, similar to uncentered covariance. The kriging postulates a field estimator at any point as a linear combination of

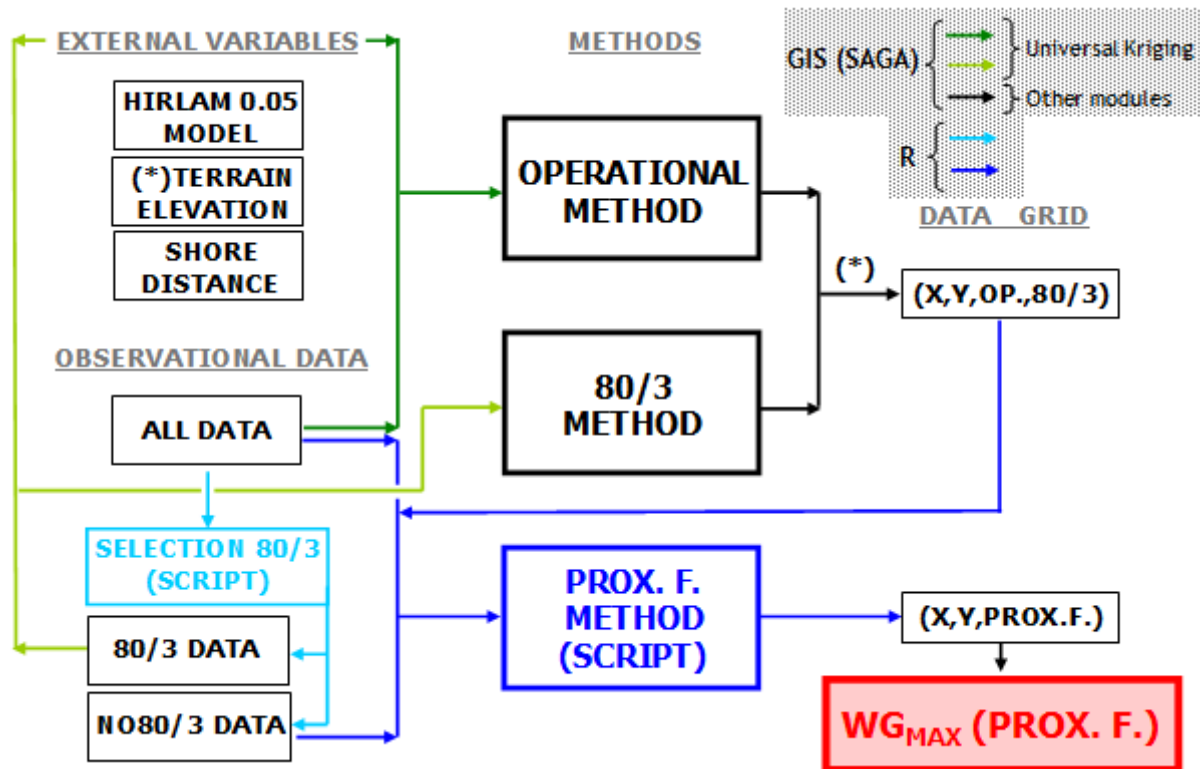


Figure 1. Diagram of the operational nature of the new procedure combined by means of proximity function.

the duly weighted observations in the rest of the points. To determine the weighting the estimator is intended to be unbiased and with a minimal variance (Samper and Carrera, 1990).

There are several types of kriging according to the additional hypotheses admitted. Universal kriging postulates a linear trend model, so that the mathematical expectation of the random field value at any point is expressed as a linear combination of the values that take several auxiliary deterministic functions at that point. This allows the incorporation of mild variation effects in the geostatistical interpolation (Burrough and McDonnell, 1998).

Specifically at AEMET, ground elevation, distance to the sea and the HIRLAM model output of the maximum gust field for the period under study are used as auxiliary deterministic functions in universal kriging. The structure of the semivariogram used is specified without “nugget effect” so that the kriging estimator is exact, that is, at the points with observation the estimated value coincides with that observed.

The idea of developing a modification of the operational technique described above arose from the results obtained in the verification studies.

To perform validations there are 15 ACS situations that try to cover recently occurred phenomena for different geographical areas and with extensions, in terms of their impact, of various dimensions.

Table 2. Mix factor. Proximity Function Method. Values that take the parameters in the previous validations done with the new method for their choice.

$-EXP$	constant	-2
		10
	constant	5
		2,5
$F_{esc} (km)$	$F_{escMax} (km)$	100
		$197,5 \times 10^{-3}$
	variable	$cc (dimensionless)$
		$141,0 \times 10^{-3}$
		84.6×10^{-3}
	$R (km)$	100

Later, a selection is done, randomly and without replacement, of the available observation stations for each of the ACSs. These stations are removed before performing the interpolation and it is therefore possible to analyze the existing differences between the estimated and observed values. 20% of the observation stations are chosen for each situation, so that these values enter in the study. This selection applies the percentage by area, to ensure spatial coverage in the observational data, and by a limit on the maximum gust value (the value of 80 km h^{-1} is taken as reference value) to ensure a number of stations to be

Table 3. Comparison of the validations done with the operational procedure and the new procedure for the different values of the parameters considering high maximum observation gusts ($MG^{Obs} \geq 80 \text{ km h}^{-1}$). Average values for 5 ACSs.

Procedure	Operational	Proximity F.					
		F_{esc} (constant)			F_{esc} (variable)		
		$F_{esc} (km)$			$cc(dimensionless) \times 10^{-3}$		
		10	5	2,5	197,5	141,0	84,6
Average biases (km h ⁻¹)	-10	-7	-5	-3	-5	-4	-2
Average relative errors (%)	17	15	14	14	14	14	13

Table 4. Comparison of the validations (average biases, km h^{-1}) done with the operational procedure and the new procedure of proximity function, considering high maximum gusts observed (left) and all observed maximum gust values (right).

ACS	$(MG^{Obs} \geq 80 \text{ km h}^{-1})$		$(ALL MG^{Obs})$	
	OP.	PROX.	OP.	PROX.
20100113a15	-7	0	1	11
20100227a28	-4	1	1	10
20111023a27	-17	-6	-1	11
20111112a14	-13	-6	-1	12
20111215a17	-8	0	-1	12
20120105a08	-20	-8	0	11
20120202a05	-22	-11	0	14
20120206a08	-13	-6	-1	11
20120415a17	-17	-7	-1	15
20120423a26	-15	-9	0	7
20121027a28	-13	-6	0	13
20121124a25	-10	-3	0	11
20121213a16	-18	-7	0	11
20130118a20	-7	-1	1	13
20130123a24	-13	-4	2	13
AVERAGE	-13 km h^{-1}	-5 km h^{-1}	0 km h^{-1}	11 km h^{-1}

validated with high maximum gust values, since these are our main interest.

All these conditions are programmed in R (R Core Team, 2013), a programming language and environment for statistical and graphical analysis, getting a script that specifies the fixed parameters: a percentage of 20% and maximum gust value limit of 80 km h^{-1} , (Venables and Ripley, 2005).

For the analysis of the selected ACS validations, we calculated the average bias, which shows the difference between the estimated and observed value (Equation 1), and the average relative error, considering the difference given above, but in absolute value (Equation 2).

$$\text{Average bias} = MG^{Est} - MG^{Obs} \quad (1)$$

$$\text{Average relative error} = \left| \frac{MG^{Est} - MG^{Obs}}{MG^{Obs}} \right| \cdot 100 \quad (2)$$

where MG is the maximum gust value observed (superscript Obs) or estimated (superscript Est), both in km h^{-1} .

Verification studies show that in the speed range of maximum wind gusts with high values, which are of interest for

the purposes of CCS coverage, there is a significant negative bias in the estimation of universal kriging. For the specific case of observed maximum gust speeds of more than 80 km h^{-1} , and for a studied set of 15 strong wind situations, we obtained an average value of bias in verification of -13 km h^{-1} (Table 1).

The validation results show a systematic negative bias for all ACS cases, which means that the method tends to underestimate these extreme values.

3 Methodology proposed

The negative bias, obtained for high observation values when using the operational method, is explained by the fact that the kriging produces an unbiased estimator globally, but when applied to a particular range of observations, like any technique in the regression family, it brings the estimator closer to the average. In our case, we selected a range of observation values higher than the average and therefore expected a negative sign.

In this way, we tried to use the same universal kriging technique but using only high rank observations, those

Table 5. Comparison of the validations (average relative errors, %) done with the operational procedure and the new procedure of proximity function, considering high maximum gusts observed (left) and all observed maximum gust values (right).

ACS	$(MG^{Obs} \geq 80 \text{ km h}^{-1})$		$(ALL MG^{Obs})$	
	OP.	PROX.	OP.	PROX.
20100113a15	16	13	23	31
20100227a28	16	15	18	26
20111023a27	21	16	18	30
20111112a14	19	14	18	41
20111215a17	16	12	18	27
20120105a08	24	13	23	37
20120202a05	24	15	23	38
20120206a08	20	14	20	30
20120415a17	20	11	17	37
20120423a26	19	16	16	22
20121027a28	22	14	20	34
20121124a25	17	11	23	43
20121213a16	19	13	20	32
20130118a20	16	12	19	32
20130123a24	18	11	18	31
AVERAGE	19%	13%	20%	33%

Table 6. Statistical treatment done on the 15 ACSs studied that represent the existing relative differences between both methods considering areas with maximum gusts equal to or greater than 120 km h^{-1} , their extension to municipalities and the total population of those municipalities.

Regarding		Area	Municipalities (by number)	Municipalities (by population)
<i>PROX.</i>	OP.	16%	-2%	-11%
<i>PROX. + OP.</i>	OP.	43%	27%	18%
<i>PROX. + OP.</i>	PROX.	23%	30%	32%

greater than or equal to 80 km h^{-1} . In addition, to ensure good spatial coverage in the provinces (or previously fixed territorial areas) in which this condition does not guarantee at least 3 stations, it is complemented with the observation values of the greatest gust, until each province has at least 3 stations involved in the interpolation.

This practice (method 80/3), as expected, shows a substantial improvement in the same checks and on the high range of speeds. The bias is reduced and values are around 2 km h^{-1} on average for each situation.

The problem with this method is that the interpolated field overestimates a great deal in the rest of the range of observations and so the delimited areas of interest, maximum gusts equal to or more than 120 km h^{-1} , could be extended in making the estimates. This is a consequence of ignoring the information given by the observation stations that show maximum gusts with low values, whose information, on the other hand, is essential for making the estimates.

We proceed to implement all of the observational data by combining the two interpolation fields seen so far, on the one hand, that which is currently used, operational, and on the other hand, that which uses the high-range observations. These two interpolations use the same geostatistic technique,

universal kriging, while the combination of both is done by assigning different weighting to one or another field for each point at which the maximum gust value is estimated. Moreover, due to the spatial difference between the number of weather stations, a factor that takes into account the total density of observations used according to the region is introduced.

This is the basis of the new procedure, called Proximity Function, which will be developed below.

3.1 Formula of Proximity Function

The combination of the two interpolations, the operational one and the one restricted to more than 80 km h^{-1} (Equation 3) was performed, assigning different weighting to each of the grid points forming the entire area. We call the new combined method F. PROX., the operational kriging on all observations OP, and the interpolation with observations of over 80 km h^{-1} basically 80/3.

$$F.PROX(\vec{p}) = F_m(\vec{p}) \cdot OP(\vec{p}) + (1 - F_m(\vec{p})) \cdot 80/3(\vec{p}) \quad (3)$$

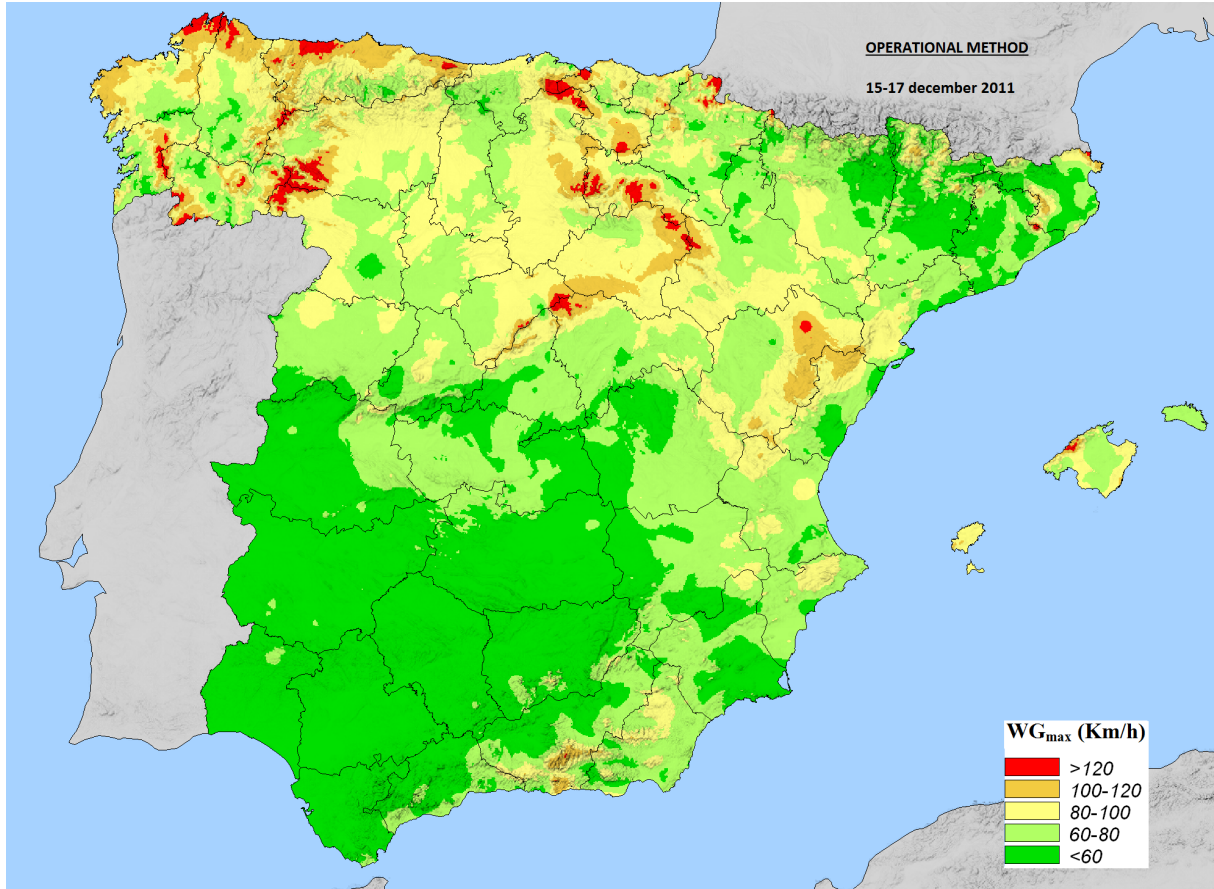


Figure 2. Maximum gust estimations for the ACS situation of 15 to 17 December 2011 using the operational method.

where \vec{p} denotes an arbitrary point and F_m is a blend factor between 0 and 1. This was obtained from other positive F'_m according to the monotonic transformation (Equation 4):

$$0 \leq F_m(\vec{p}) = \frac{F'_m(\vec{p})}{1 + F'_m(\vec{p})} \leq 1 \quad (4)$$

To determine F'_m , it was taken into account that the weighting assignment is done by taking the distance between each problem point \vec{p} and the points \vec{s} that are stations that are not used in the method 80/3, so that the nearer they are, the more weighting the operational method will have and the greater value will be shown by the mix factor, and vice versa.

Moreover, given the differences in the overall density of the observations used in different regions, we thought that we should include a scale factor that took this local density of observations into account (Equation 5).

$$F'_m(\vec{p}) = \sum_{\vec{s} \in \text{NO80/3}} \left(\frac{\text{dist}(\vec{p}, \vec{s})}{F_{esc}(\vec{p})} \right)^{-EXP} \quad (5)$$

where the summation is extended to the stations \vec{s} not used in the 80/3 method, the exponent $-EXP$ was taken equal to

-2 and the scale factor F_{esc} has dimensions of km to give a dimensionless mixing factor (Equation 6).

$$F_{esc}(\vec{p}) = \min \left[F_{escMax}, \frac{cc}{\sqrt{\rho_{local}(\vec{p})}} \right] \quad (6)$$

The maximum value of the scale factor has a predefined limit, with cc being a constant to set, and the local density of the local stations ρ_{local} that is determined locally. The search for stations is performed in a circle $C(\vec{p}, R)$ around the point \vec{p} of a 100 km radius R (Equation 7).

$$\rho_{local}(\vec{p}) = \frac{n_{est}^o \in C(\vec{p}, R)}{\pi \cdot R^2} \quad (7)$$

Thus, on the assumption of uniform density of stations around any \vec{p} point, a homotety around \vec{p} of arbitrary factor r leaves F'_m invariant (if it does not reach the F_{escMax} limit), as distances are multiplied by r and the local density by r^{-2} .

Validations are performed with different values of the above parameters in order to select the most suitable ones (Table 2). The choice of these specific values has been agreed according to the development of several empirical tests that show good performance of the function, noting, in turn, a coherent and logical meaning for its application.

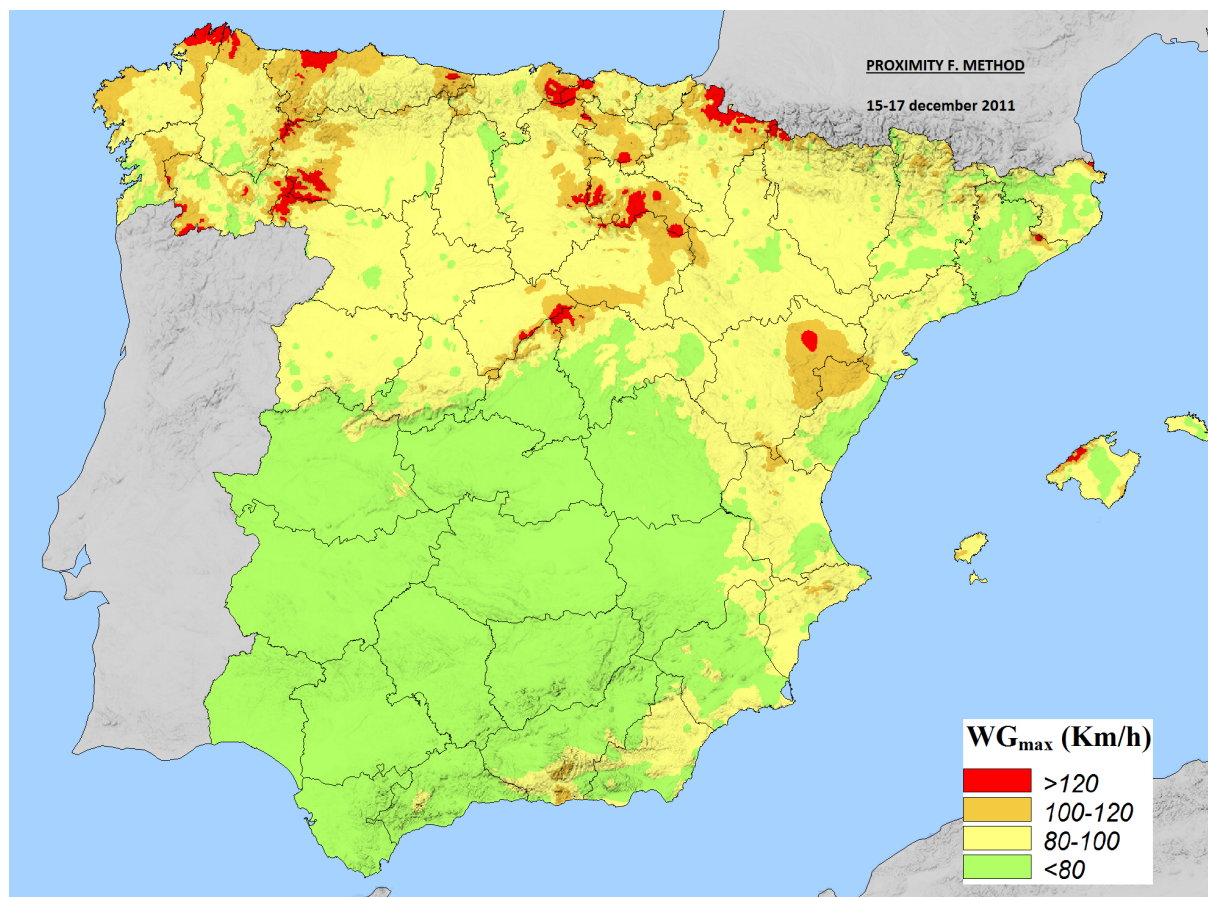


Figure 3. Maximum gust estimations for the ACS situation of 15 to 17 December 2011 using the combined method with proximity function.

Finally, the choice was to take the value of the dimensionless constant $cc = 84.6 \times 10^{-3}$ which is determined by the condition that F_{esc} has a value of 1.5 km when there are 100 stations in a radius of 100 km.

3.2 Operational Proximity F.

To obtain the final estimate using the new method, in order to delimit areas of maximum gust that exceed 120 km h^{-1} , two different interpolations are combined, the operational one and the 80/3 one (Figure 1). These use a geostatistical technique, universal kriging, which is based on observational data and takes other external variables. Subsequently, the estimates are done using the proximity function method according to the formulation given in the previous section.

These formulas are programmed in R (Venables and Smith, 2012) resulting in a single script which is run to get a file with the final information, the maximum gust estimates by means of the new method. To do so it is necessary, on one hand, to have the input files that contain basically the information given by the observational data and, on the other hand, to set the values of the parameters listed above (Elosua, 2011).

Finally, it should be noted that the calculations done by the script are performed in a matricial way through a subgrid belonging to the total 1-km resolution grid. This subgrid is taken every 5 points, in both X direction and Y direction, thus allowing an optimization of the program execution time. The assignation of values, finally, for the original grid is performed using the nearest neighbor technique.

4 Results

The proposed method, combined with the proximity function, introduces significant improvements for the target set, the delimitation of areas with higher maximum gusts.

Initial validations for the different values of the parameters, are performed for five ACS cases (Table 3).

The final choice was to take the value of the dimensionless constant $cc = 84.6 \times 10^{-3}$, as noted above. It was noted that the average bias for the combined method on the 15 analyzed high wind situations and checking with observations above 80 km h^{-1} had dropped by -5 km h^{-1} , as opposed to -13 km h^{-1} of the operational method, a reduction of around 60% (Table 4).

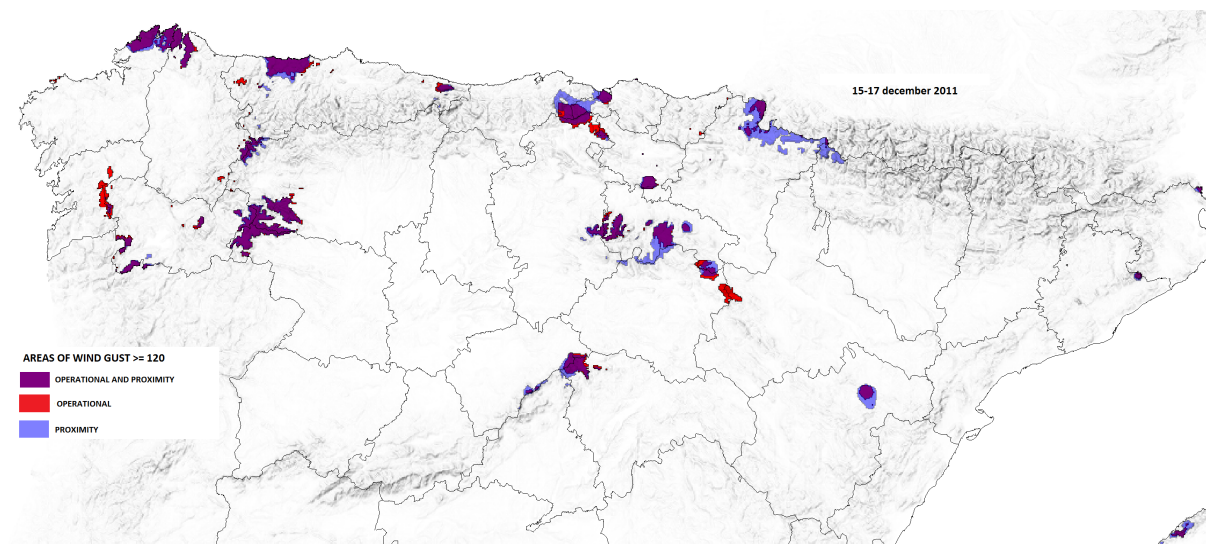


Figure 4. For the ACS situation of 15 to 17 December 2011, delimitation of the estimated maximum wind gusts equal to or higher than 120 km h^{-1} comparing both methods, operational and proximity. Cut of the north-central area of the peninsula.

Moreover, the result of the validations of the entire range of observations shows that the operational technique, which is theoretically unbiased in this range, gives very small biases.

The results of the relative errors derived from the validations (Table 5), calculated as an absolute value (regardless of their sign) are also improved. Although this improvement is not very significant, errors with the new method show similar results for whatever the situation, i.e., with the combined method with proximity function a lower dispersion is obtained, $s_{PROX.} = 2\%$ of the new method compared to $s_{OP.} = 3\%$ of the operational method.

Furthermore, comparing the results obtained by taking only the stations with a maximum gust observed of $\geq 80 \text{ km h}^{-1}$ with the results considering the set of all stations, whatever their maximum gust speed, we observe that given the significant improvement shown by the validation study for the first case, deterioration occurs in the second one. Therefore, the improvement achieved by the new method which focuses on high values of maximum gust is counteracted, as expected, with worse outcomes for relative biases and errors when low speeds are also taken into account.

The analysis is completed with a statistical treatment that encompasses all the ACSs being studied, which consists of comparing the relative differences between both methods. The variables that are treated and that allow these differences to be quantified refer to the extent of the area that presents maximum gusts of $\geq 120 \text{ km h}^{-1}$, to the number of municipalities affected by that condition and, finally, to the number of inhabitants of the resulting municipalities (Table 6).

When comparing the combined method with the proximity function versus the operational method, it is seen that the former has a greater coverage of the area bounding

the maximum gusts greater than or equal to 120 km h^{-1} . To the contrary, by extending these areas to municipalities, an application typically done by the CCS to cover damages caused according to their regulation, there is a 2% decrease. To attempt a more representative analysis of the costs which may result in a situation of ACS, we resort to population of the municipalities in which we obtained a decrease of 11% with the new method compared to the operational one.

These same calculations are applied for the joint coverage of both methods, i.e. taking the maximum gust areas above 120 km h^{-1} considered by any of the procedures, and this is compared to each of the methods separately, although our special interest is focused on the resulting difference regarding the operational method. We proceed in the same way for the number of municipalities and population.

In this case we see that there is a considerably high increase in the areas covered by applying the joint method in comparison to those covered by the operational method, 43%, albeit in practice, it is more representative to consider the population, 18%; which is no longer such a high value. When interpreting these results, it is important to note that the encompassed ACSs are very diverse, with very different extensions.

The ACS situations studied concern the Iberian Peninsula and the area of the Balearic Islands, although the procedure is also active for the Canary Islands. A specific example is shown, from 15 to 17 December of 2011, where the regions of interest (maximum gust speed equal to or greater than 120 km h^{-1}) are marked in red, (Figure 2) and (Figure 3).

Both cases, the operational method and the proximity function, have a similar extension although it is slightly higher for the new method, as we saw in the statistical treatment. Just small changes occur in the position they

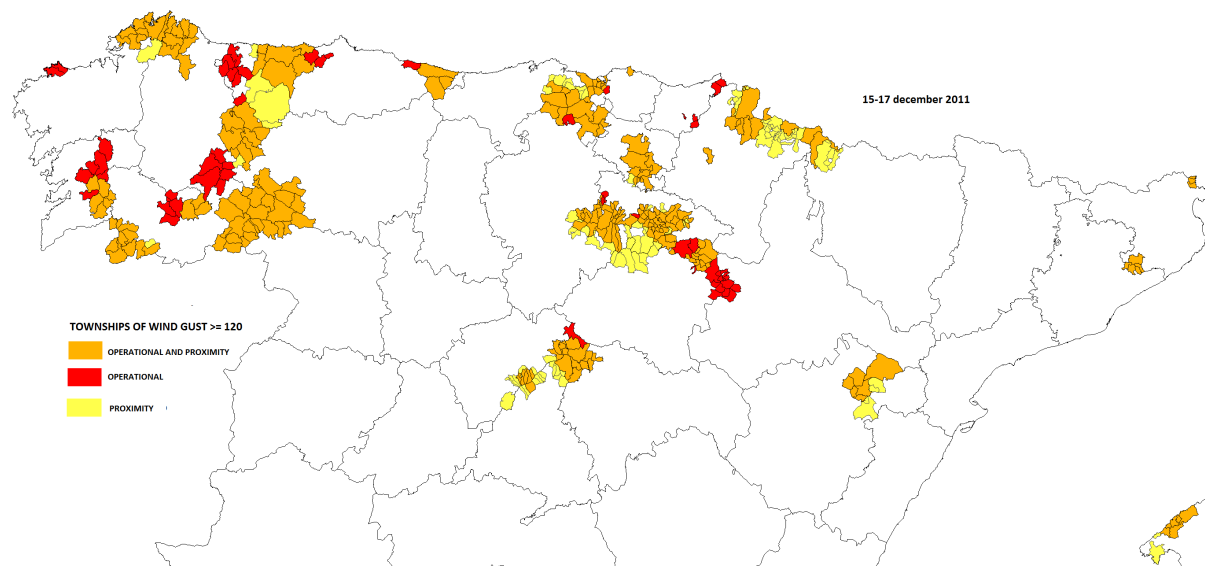


Figure 5. For the ACS situation of 15 to 17 December 2011, application of the above mentioned areas to the municipalities. Cut of the north-central area of the peninsula.

occupy but, overall, they do not incorporate or eliminate large areas that could change the importance of the same ACS situation.

Instead, there are variations for low maximum gust values, as greater accuracy in estimating high values leads to increasing errors for other speeds. As our interest is focused on the areas in which gusts exceeding 120 km h^{-1} occur, a comparison of these areas is shown (Figure 4), and its extension to municipalities (Figure 5), CCS application coverage.

5 Conclusions

The new method presented, combined with the proximity function, introduces significant improvements regarding operational procedures in estimating high maximum gusts, which are those of interest to us, because the current CCS regulation specifies the limit of 120 km h^{-1} for the coverage of extraordinary risks. The main one is that it greatly reduces, by 60%, the negative bias in estimating the maximum gust. It also decreases the relative error and has less dispersion.

Furthermore, as was the case with the operational procedure, the new method is an exact estimator, i.e., in the observation points, the estimated value coincides with the observed value, the actual value given by the station is preserved. This is a consequence of it being based on combining the operational method and the 80/3 one described above, that still use the universal kriging technique without “nugget effect”. The estimate given by the new method at these points is a mixture of these two exact values and, therefore, the result will also be accurate. This exact

interpolator character is important as for the Consortium it is necessary to respect the maximum observed values as they are those with higher legal strength.

The statistical analysis that encompasses all ACSs shows that the areas bounded by the maximum gust $\geq 120 \text{ km h}^{-1}$ are increased to 16% with the new method in comparison to the operational one, while the number of municipalities affected by gusts of $\geq 120 \text{ km h}^{-1}$ is almost the same and the population of all these municipalities has a 11% decrease.

Another feature to be taken into account in the proposed method refers to its effectiveness. It is presented in a very similar way to that used so far using basically the same statistical technique with the same tools, it just introduces the execution of certain scripts that are easy to operate. Furthermore, the preparation of the information is identical and the estimates, using the new method, are carried out in a time that is only slightly above that used to evaluate the operational method.

Acknowledgements. This work has been possible thanks to the collaboration between the Spanish Meteorological Agency and the Insurance Compensation Consortium which, together with AT Sistemas, are sponsoring the project. The authors wish to acknowledge the support of many people who belong to these organizations and have made this research possible.

References

- Burrough, P. A. and McDonnell, R. A., 1998: Principles of Geographical Information Systems, Oxford University Press, 333, ISBN-13 978-0-19-823365-7, ISBN-10 0-19-823365-5.
- Consorcio de Compensación de Seguros, 2012: Recopilación Legislativa, Edición de febrero de 2012, 202.

- Elosua, P., 2011: Introducción al entorno R, Universidad del País Vasco, 102, ISBN: 978-84-9860-497-9.
- García-Legaz, C. and Valero, F., 2003: Riesgos Climáticos e Impacto Ambiental, Ed. Complutense de Madrid, 356, ISBN: 84-7491-711-5.
- R Core Team, 2013: R: A language and environment for statistical computing, Technical report, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>.
- Samper, F. J. and Carrera, J., 1990: Geoestadística. Aplicaciones a la hidrología subterránea, Ed. Complutense de Madrid, 484, ISBN: 84-404-6045-7.
- Venables, W. N. and Ripley, B. D., 2005: Modern Applied Statistics with R, Springer-Verlag, 106, ISBN: 3900051-127.
- Venables, W. N. and Smith, D. M., 2012: An Introduction to R, R Foundation for Statistical Computing, Vienna, Austria, 95, ISBN: 3-900051-12-7.