Wind Storm Risk Management

Alexandre Mornet, Thomas Opitz, Michel Luzi, Stéphane Loisel

▶ **To cite this version:**

Alexandre Mornet, Thomas Opitz, Michel Luzi, Stéphane Loisel. Wind Storm Risk Management : Sensitivity of Return Period Calculations and Spread on the Territory. 2016. hal-01299692

HAL Id: hal-01299692
https://hal.science/hal-01299692

Preprint submitted on 12 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wind Storm Risk Management - Sensitivity of Return Period Calculations and Spread on the Territory

Alexandre Mornet* Thomas Opitz† Michel Luzi‡ Stéphane Loisel§

February 2015

* Université de Lyon, Université Claude Bernard Lyon 1, Institut de Science Financière et d'Assurances, 50 Avenue Tony Garnier, Lyon F-69007, France and Allianz, Coeur Défense, 82 Esplanade du Général de Gaulle, Courbevoie F-92400, France, alexandre.mornet@allianz.fr

† Biostatistics and Spatial Processes Unit, National Institute of Agronomic Research, Avignon, France

‡ Ancien Directeur actuariat non vie and Directeur recherche et développement at Allianz France, Membre certifié de l'Institut des Actuaires, 132, rue du Président Wilson, Levallois-Perret F-92300, France

§ Université de Lyon, Université Claude Bernard Lyon 1, Institut de Science Financière et d'Assurances, 50 Avenue Tony Garnier, Lyon F-69007, France, stephane.loisel@univ-lyon1.fr

Abstract

Models and forecasts of damage from wind storms are a major issue for insurance companies. In this article, we focus on the calculation sensitivity of return periods for extreme events. Numerous elements come into play, such as data quality (location of insured buildings, weather report homogeneity), missing updates (history of insurance portfolios, change of ground roughness, climate change), the evolution of the model after an unprecedented event such as Lothar in Europe and temporal aggregation (events defined through blocks of 2 or 3 days or blocks of one week). Another important aspect concerns storm trajectories, which could change due to global warming or sweep larger areas. We here partition the French territory into 6 storm zones depending on extreme wind correlation to test several scenarios. We use a storm index − defined in [20]− to show the difficulties met to obtain reliable results on extreme events.

Keywords. Storm Index, Wind Speed, Insurance, Extreme Value Theory, Extreme Dependence, Return period, Solvency II

1 Introduction

The study of storms is both a matter for natural phenomena linked to climatic and geographical factors, and on the other hand, for its consequences in terms of catastrophes linked to economical, historical, and human factors. The management of risk during extreme events is something which has an influence on the whole of society [11], far beyond insurance.

First, to understand our approach, it is necessary to consider the problematic of the management and evaluation of storms. The economic and structural information available in an insurance firm show the volatility and intensity of damages. Yet they can only account for historical data that are limited and difficult to update [13]. This is why we have decided to complete our data with meteorological reports. We have taken into account different variables that may be linked to storms and finally to focus on wind speeds ([4], [16] and [22]).

However the evaluation of the impact of a storm is not as simple and direct as we could wish. Several difficulties we have come across will be discussed here in order to account for the variability of the results we have obtained. If it is possible to get data on the disasters and portfolios on relatively precise levels (city or postal address), we generally only have statements on the national or departmental levels, as soon as we use market data or old-time history. Furthermore, if we refer to meteorological data, we have only a partial number of stations, roughly one or two in each department which makes it difficult to work with great precision. We can also note a great variability of recorded speeds according to the geographical situation of the station [24]. The stations at high altitude or on the coast often show speeds that are greater than in the rest on the department and that must be balanced according to their distance from main towns. Then, for the same station, the homogeneity of the recordings on a long period is not necessarily confirmed. The study of the breaks [25] has allowed us to notice discrepancies, perhaps due to a change in equipment, but that have to be connected with other phenomena such as different roughness of soils or climatic changes which can also disturb the data. Considering those difficulties, one possible solution is to regroup the French departments according to the extreme speeds of winds. Paradoxically, one might be tempted to try and work in a more subtle way, at least to avoid comparing wind speeds several dozen miles from risks.

Then here, we choose the opposite solution. Our division is justified by the fact that there are departments with very few important damages (because there are few contracts). We must add to this our limited number of significant events and the heterogeneity of distances between the risk and the station. A station approach by department would not be clear enough .We have chosen a distribution into six zones of risks, which allows us to act on an intermediate scale between the department

(dimension 95) the whole of France. For this, we use the `k-medoids` clustering algorithm [3]. The work on these zones allows us to test various methods of agglomeration of departments (extreme dependence [2], maximal speeds, and geographical distance). It allows us to consider various scenarios of the extent of storms as in [12] and [17]. These amplitudes might change with climatic change (arrival of storms of the Medicane([1]) type for example) or sweep over various zones successively. Such projections are a coherent means to foresee the small number of major storms in our base of observation.

On a more technical level, the building of a storm index implies the choice of a time-pace. Here we must separate the grouping of days in order to stick to the notion of the event retained as a definition by insurers and of clusters that correspond to a period favorable for the starting point of successive events. We have realized that, contrary to the evaluation of changes, which makes necessary the addition of catastrophes during the whole event [20], the maximal daily evaluation of our index is the best information about the relative intensity of storms. Of course, these decision will have an impact on the projection of storms. We have compared the results of statistic approaches according to the use of daily or weekly data. In our models, we have also tried to show the sensitiveness of the calculations of the return period to the inclusion or not of the most important events like Lothar and/or Martin in 1999.

For an insurer, the management of the damages due to the wind depends very much on the return period that is attributed to the most extreme storms. Such events, of course, are very few and the story of claims proves inadequate to give sound answers to questions related to the pricing, to volatility of results, therefore on the need of capital stock. The communication on the uncertainty that goes with the prevision of extreme events is important and is often neglected. Yet, there are technical means to determine that uncertainty, but a sociological reluctance to the aleas limits its diffusion. It was the case during Juno storm in 2015 that was largely overestimated as Winkler [27] explains. The purpose of this article is to bring a new light on the modelling and calculation of these extreme events by basing on the storm index as defined in our previous articles [20].

Here we wish to insist on the great sensitiveness of results to the hypotheses we have returned, while proposing complete concrete results leaning on meteorological data we have retained, at the same time proposing concrete results leaning on the available over the period of abaut40 years, and less disturbed by the actualization problems than insurance data. This sensitiveness will be illustrated both by the calculation of return periods and by the estimation of needs in capital stock.

[1] Medicane is a contraction of Mediterranean and hurricane. This climatic phenomenon has existed for the past two decades. It is due to the conjunction of cold air at altitude and an unusual warming of the Mediterranean sea.

2 About wind speed

2.1 Definition

The reference value of the speed and direction at a given measurement point and moment are the immediate speed and direction averages at long enough intervals (usually 10 minutes) which precede the reported measurement time. Weather reports refer to medium wind, while possibly mentioning the observed measurement or the expected intensity of gusts (whose values can exceed the medium wind measurements by more than half). It is also possible to measure the vertical component of the wind with a three-dimensional anemometer. Thus, we can study the subsiding and rising air movements. However, our data sources are not that exhaustive, and we settle for learning as much as possible from available data.

2.2 Problem of breaks in data series

One issue calling for particular attention is the reliability of meteorological data. Indeed, when we focus on a long period of time, measurement devices and data restitution methods can change and bias our results. The presence of a changepoint can lead to a significant loss of precision [25]. It is consequently important to detect ruptures in the structure of data series and to correct them. Through observation of annual means/medians/variances, for each station, we try to spot a change in the series.

2.3 Break detection in our data

We detect potential breaks in the central tendency of wind speed data using a standard software implementation (`changepoint` package of the `R` software library [23]). Relevant methods are described in Killick et al. [15]. Formally, for a data sequence $(y_1, ..., y_n)$, it is said that a break $\tau \in \{1, ..., n-1\}$ exists if the statistical properties of subsets $\{y_1, ..., y_\tau\}$ and $\{y_{\tau+1}, ..., y_n\}$ are different according to certain criteria. Multiple break detection necessitates minimizing the target function

$$\sum_{i=1}^{m+1} \left[\mathcal{C}\left(y(\tau_{i-1}+1) : \tau_i\right)\right] + \beta f(m) \tag{1}$$

with $\mathcal{C}$ a cost function for each segment and $\beta f(m)$ a penalty to avoid over-segmentation. Here, we use the hierarchical and bisection-based PELT (Pruned Exact Linear Time) algorithm. PELT first does a rupture test on the whole period. If a break is detected, data is divided in two subsets at the break position. Then, detection is conducted on both subsets and iterated until no more breaks are found. To avoid seasonality problems, we here focus on the annual averages of each station. Besides, since our index is built from the highest speeds, detection is here made on the threshold exceedance

of the empirical quantile q at 80%.

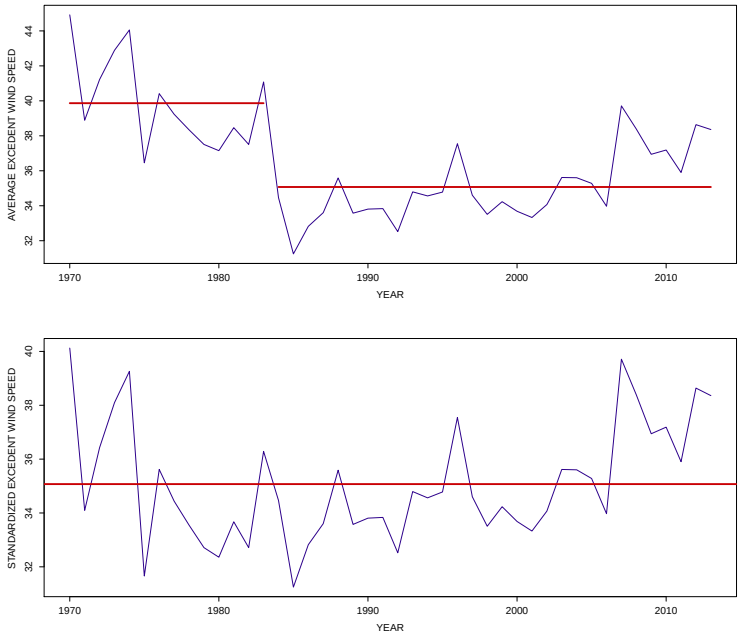

Figure 1: Wind speed exceedances in Nîmes - before (above) and after (below) readjusting

On the graph at the top of Figure 1, we can observe the results of the break detection on an annual average of wind speed exceedances. The period is partitioned into two sequences: from 1970 to 1983 then from 1984 to 2013. For this station, the recorded speeds during the first sequence are higher compared to the second sequence. It is then necessary to homogenize the data for Nîmes. Until 1983, the average exceedance is 39.87 m/s and starting from 1984, it is 35.07 m/s. We standardize the series (at the bottom of Figure 1) by adjusting the exceedances of the first sequence with respect to the second one, more recent, whose reports we assume to be more reliable. We underline that this approach is robust yet relatively simple, so we have to remain cautious when interpreting results which involve data from before 1980. Old data remains less reliable and noisier even after preprocessing.

Several explanations can be provided for the origin of these breaks. Changing the topology around the station following, for example, the construction of a building nearby, or changing the type of anemometer used are sources of disturbance for the measurement. Sacré et al [24] explain that the Papillon anemometers used in the 1970s caused problems, especially for high wind speeds. Figure 2 represents the transmitter of the device in question [21]. Mechanical anemometers are sensitive to frost and humidity, which leads us to underestimate wintry wind speeds in the mountains. Current state-of-the-art devices use transmitters-receivers from directional ultrasounds, but measurements can still be prone to technical problems. The Nice airport station for example presents recurring

errors despite its wind profiler, which turns out to be unable to detect land breeze and mountain breeze from the nearby Var valley.

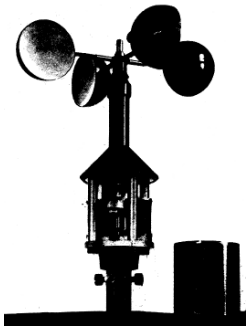

Figure 2: Papillon anemometer transmitter

3 Distribution of stations and risks over the territory

Météo France's Radome network counts 554 professional meteorological stations in mainland France (1 for every 30 kilometers) and 67 overseas. Currently, these stations automatically measure the basic parameters, which are temperature and humidity, rainfall and wind (speed and direction). Some parameters meet more specific needs and are measured only on sensitive areas: airfields, zones with high risk of floods, forest fires or avalanches. Except traditional measurements made by barometers (pressure), hygrometers (air humidity), anemometers (wind) and rain gauges (rain), technical progress allows nowadays to automatically measure a greater and greater number of parameters such as soil moisture, ground state, radiation, visibility, cloud height.

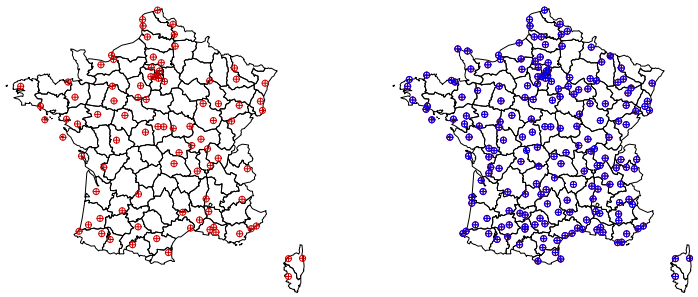

Figure 3: Maps of active stations - from 1963 (on the left) and from 1998 (on the right)

However, the high spatial resolution of the Radome network is only relatively recent. Since we wish to work on a 50-year history, we have to consider data availability for the 1963 to 2013 period. On Figure 3, all active stations in 1963 (the start of our observation period) and 1998 (the start of claim records with daily precision at Allianz) are depicted. The maps show a considerable evolution

of station density on the territory. In 1963, we have the records of 89 stations. Some administrative departments are not available and there are few data for the Limousin and Champagne-Ardennes regions. Thirty years later, the number has more than doubled with 190 active stations and almost no missing department except for two departements in the Paris region.

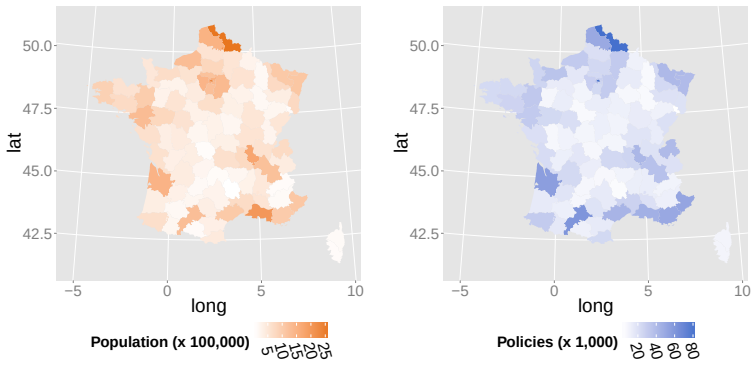

Figure 4: Maps of population (on the left) and number of Allianz' policies (on the right)

This distribution of stations has to be put into relation with population density and the number of policies in the Allianz portfolio for each department (Figure 4). This relation must be incorporated when seeking a connection between meteorological information and insurance claims. The department scale is not necessarily the best option for building a storm index. In the next section, we will see that it is pertinent to work on larger and more homogeneous zones by grouping departments according to their exposition to high wind speeds.

3.1 Classifying departments into homogenous storm risk zones

We use daily wind speed records over 50 years. Our goal is to aggregate the data on larger territories than a department to facilitate the search of a connection between wind speeds and costs. Indeed, working on France as a whole would not take the heterogeneities of the territory into account, whereas an approach by department causes a problem in terms of availability, reliability and heterogeneity of data. Our intermediate approach allows looking at the results for each of these larger zones, which would have been difficult for 95 departments. The French territory shows great variety of relief and climate; therefore, it is appropriate to delineate zones likely to be simultaneously affected. We can further recover information on typical storm trajectories and on regions relatively unscathed by wind damages. Furthermore, statistical uncertainty in department-based results would have been rather strong, whereas grouping together "similar" departements enables us to reduce them. For insurance, this approach also allows us to detect the regions where the risk is the highest or the lowest compared with the national average.

For a fixed value of $k = 2, 3, ...$, the k-medoids algorithm of [14] returns k zones by minimizing the distances between a central department and associated departments. Medoids are the central representatives of a zone, i. e. the central department in our case. The choice of distance is crucial; we here build a distance measure based on a dependence index λ of bivariate extremes, measuring the dependence between distribution tails of two data series. This index represents the probability of simultaneous extreme realizations and therefore its values are between 0 (asymptotic dependence) and 1 (total dependence). For two random variables, X and Y having as distribution functions F_X and F_Y, the index is defined by

$$\lambda = \lim_{v \uparrow 1} \left(\mathbf{P}(Y > F_Y^{-1}(v) | X > F_X^{-1}(v)) \right), \tag{2}$$

provided that this limit exists. We calculated this index from daily wind speeds by fixing as a threshold the quantiles at 80, 90 and 99%. This corresponds to retaining for each department successively the 3700, 370 and 185 highest speeds. For a spatial representation of dependencies, we transform this index into a distance compatible with the k-medoids algorithm. We propose

$$d_1(x, y) = 1 - \lambda. \tag{3}$$

Naveau et al [3] uses a similar distance defined for annual maxima in order to assemble precipitation measuring sites. This method makes it possible to get quite homogenous groups geographically. Nevertheless, some departments can be very far from their medoid, which prevents us from having zones clearly delimited. For this reason, we undertook to slightly change the distance used by adding to it a component dependent on the geometrical Euclidian distance (d_{geo}) between each point considered based on a projection of longitude-latitude coordinates into the plane. Moreover, this component also allows taking into account potential differences in the management of local administrations with respect to storms. Calculated intersite distances have been rescaled in order to make them comparable with $d(x, y)$, and we propose a new distance

$$d_2(x, y) = p \times d_1(x, y) + (1 - p) \times d_{geo}(x, y), \tag{4}$$

where p taken between 0 and 1 allows regulating the influence of geometric distance on our extremes-based distance. The inclusion of a geographic distance criterion makes it possible to counterbalance issues of inaccurate anemometer, at least partly. Based on preliminary analyses, we use distance d_2 with the `k-medoids` algorithm, set $p = 0.7$ and divide mainland France into $k = 6$ zones. The `k-medoids` algorithm results for the three selected thresholds (0.8, 0.98 and 0.99) are presented in the Figures 5, 6 and 7. On the maps, we can see that the affiliation of a department to a group is more or

less marked depending on the dimension of its central point and the line which binds it to the central department (medoid). We determine the strength of each group from the *silhouette* which gathers information in the form of a histogram whose bar size is proportional to the affiliation strength to the group. The silhouette $s(\text{dep}) \in [-1, 1]$ of a department dep is the difference between $a(\text{dep})$, defined as the average of distances with the other departments of its area, and $b(\text{dep})$, defined as the minimum of this average distance calculated by reference to all the other areas. This difference is standardized by the maximum of $a(\text{dep})$ and $b(\text{dep})$ such that $s(\text{dep}) = (a(\text{dep}) - b(\text{dep})/max(a(\text{dep}), b(\text{dep}))$. The algorithm does not directly optimize the silhouette (which would be too heavy in calculation), and negative values may exist. The silhouette is a means to check if the algorithm globally succeeds in delineating zones; ideally, there are no negative values. In front of each histogram, the number of departments which constitute a group is reported.

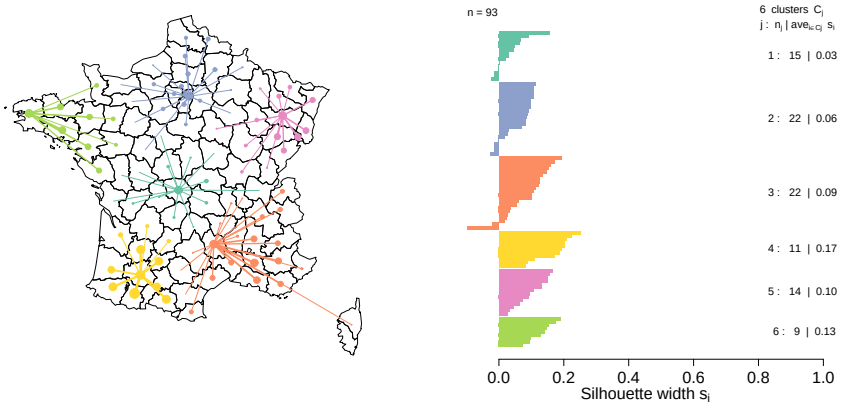

Figure 5: Partition of France and silhouette according to 80% wind speed quantile

The first divisions appear in Figure 5. The groups 1, 2 and 3, which respectively count 15, 22 and 22 departments, are the least solid. Looking at their silhouettes in detail, we note negative values corresponding to departments not closely attached to their medoid. However, groups 4, 5 and 6 present entirely positive silhouettes with the best result for the group 4 (0.17). On the map, this strength is expressed through the dimension of points located at the center of departments.

Figure 6 reports the geographic evolution of the 6 zones when the probability threshold increases from 0.80 to 0.98, shifting focus to extreme and unusually high wind speeds. According to the silhouette, the average strength slightly decreases (from 0.09 to 0.07) and group 5 is the only one with an entirely positive histogram. Group 2, in northern France, is largest one, regrouping 26 departments. Zone 6 henceforth spreads from Brittany to the center. In the south, the division is less homogenous with departments isolated from their medoid.

Finally, Figure 7 was obtained by fixing the quantile to exceed at 99%. The average strength of the silhouette further decreases to 0.05. Groups thus formed are quite different from the previous ones. The departments along the Eastern frontier of France are reunited with Corsica. The only

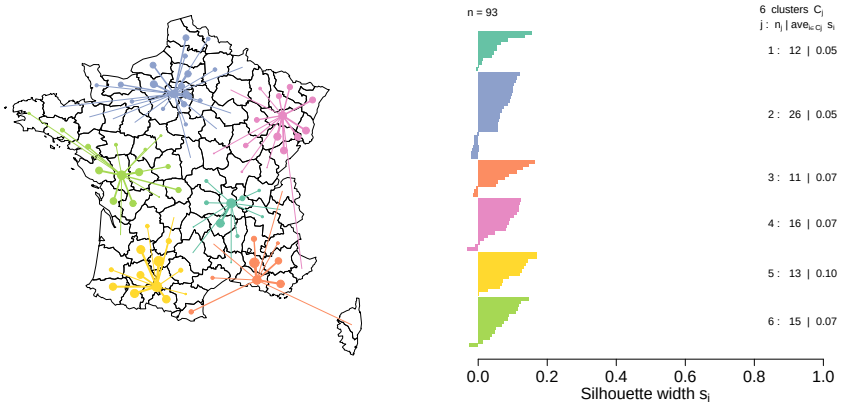

Figure 6: Partition of France and silhouette according to 98% wind speed quantile

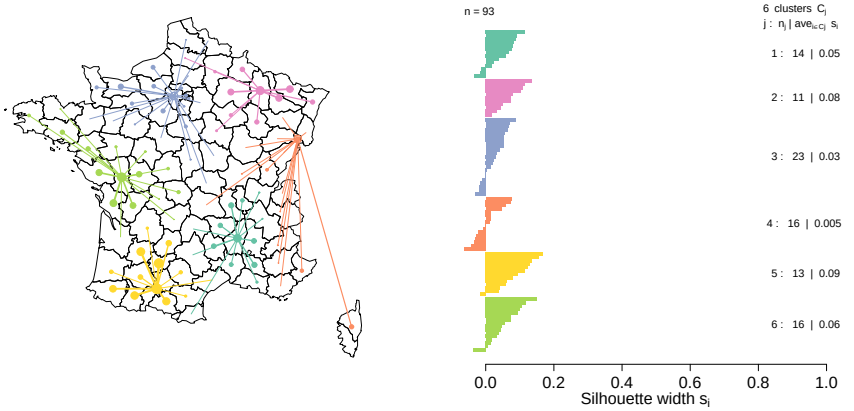

Figure 7: Partition of France and silhouette according to 99% wind speed quantile

entirely positive zone is the second one in the Northeast with a silhouette of 0.08. If the division is less homogenous than before, we can still notice that obtained zones correspond quite well to the French large river basins. These river catchments are fundamental in the study of natural risks [26]. The vulnerability to a storm depends on structural, geographical and seasonal factors at once. Our work on risk zones aims at better exposing the spatial structure of these vulnerabilities.

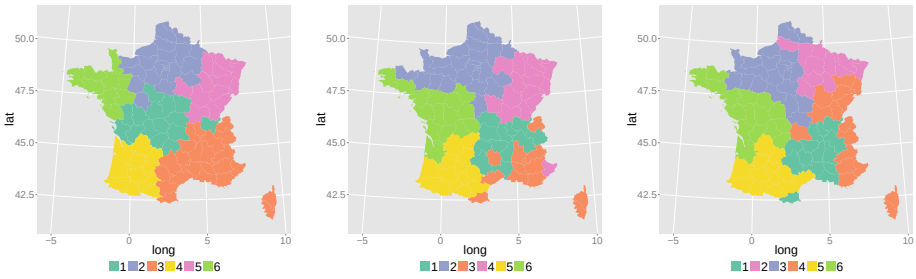

Figure 8: Distribution of the 6 wind zones with d_2 - according to 0.8, 0.98 and 0.99 thresholds (from left to right)

Figure 8 recaps the three different partitions according to the value of chosen wind quantile. The

first map (on the left) presents the most homogeneous zones on the whole and expresses a territorial sensitivity to small-scale national storms. When we focus on the most extreme exceedances, the division becomes less accurate geographically but it is more likely to show possible exceptional storm trajectories. Besides dependence of extremes and geographical distance, we have further added a third criterion: the resemblance of entire series of wind speed by department. We then built d_{tot}, which is the Euclidian norm of the difference of wind series at two stations. Our goal is to enhance the strength of the groups and thus delineate at best, by combining those three approaches, homogenous zones where wind circulation is significantly similar. The values of these different distances have been scaled in order to make them comparable. We arrive at the following definition of our distance:

$$d_3(x,y) = p_1 \times d(x,y) + p_2 \times d_{geo}(x,y) + (1 - p_1 - p_2) \times d_{tot}(x,y), \tag{5}$$

with $p_1 = p_2 = 0.33$ chosen in a preliminary analysis. This new formula is used as before with the three quantiles serving as a threshold in the distance of extremes. We consistently notice an improvement in the strength of zones. For q_{80} the average strength of all groups reaches 0.16. Group 1, which counts 21 departments, is the least solid (0.05). By observing its silhouette in detail, we find negative values, which correspond to departments not closely bound to their medoid. In contrast, other groups present mostly positive silhouettes with the best result for group 2 (0.28) in the region north of Paris. For q_{98} the division is substantially the same. The average strength decreases by 3 points (from 0.16 to 0.13), and groups 2 and 5 stay the most homogenous. Group 1, in Southern France, remains the most important by gathering 19 departments, but also the least solid. Zone 3 gains one department towards the North. Zone 4 settles more distinctly on Eastern France from the North to the South and zone 6 broadens with two additional departments. Finally, for q_{99}, the average strength of the silhouette only decreases by one point to 0.12. Groups thus formed are quite different from the previous approaches (and their numbering also changed). 23 departments along the Eastern frontier of France are gathered, from Alsace to Corsica. The zone around the Paris now counts the most departments (24). The most homogenous zone is the fourth in the South-West with a silhouette at 0.21; it has the same contour as zone 5 in the previous divisions. The new zone 5 ends up centered on Brittany and becomes smaller in terms of surface area. The three maps appear in Figure 9.

Those different approaches lead us to the concept of materiality of perils, expressed through potential drifts caused by storm damages on a geographical and temporal scale. We now investigate how much the morphology of zones depends on few influential events, the most extreme in particular. Therefore, we take the simulation of the k-medoids algorithm but without wind speeds recorded between December 24th and 28th 1999, which are associated to the storms Lothar and Martin. We

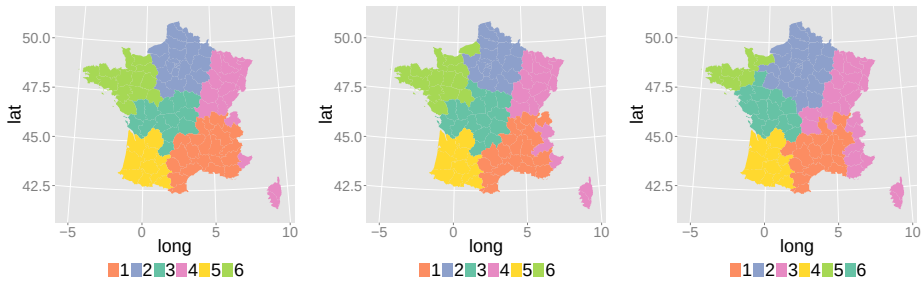

Figure 9: Distribution of the 6 wind zones with d_3 - according to 0.8, 0.98 and 0.99 thresholds (from left to right)

use the last distance defined 5 to establish the resulting zones.

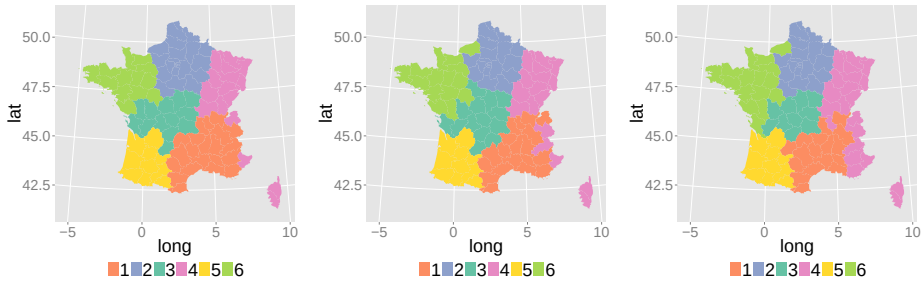

Figure 10: Distribution of the 6 wind zones without Lothar - according to 0.8, 0.98 and 0.99 thresholds (from left to right)

On Figure 10, we have depicted the six storm risk zones obtained after removing the data of Lothar and Martin storms. Average strength remains unchanged (0.16 for the quantile at 80%, 0.13 for the quantile at 98% and 0.11 for the quantile at 99%). The group attribution and hence the geographic shape of zones might change as a consequence. For the first two thresholds (q_{80} and q_{98}), changes are negligible; only for the highest threshold (q_{99}), we observe noticeable differences. The most obvious concerns zone 6 around Brittany which clearly broadens at the expense of zones 2 and 3. Zone 4 in Eastern France still extends from North to South but does not spread to the center. Finally, in the South, zones 1 and 5 remain stable, which can seem normal given that they have been the least seriously affected by the storms Lothar and Martin.

Hazard management is here complicated by the fact that elements outside the scope of variables available to insurers influence the number of claims. A natural phenomenon conveys a field of action (space), intensity (impact) and recurrence (frequency). It is possible to use spatial statistics ([9], [1]) to establish damage simulations adapted to each of the wind zones previously defined. For the rest of the study, the division with the quantile at 99% and all the wind speeds (Figure 11) will be used.

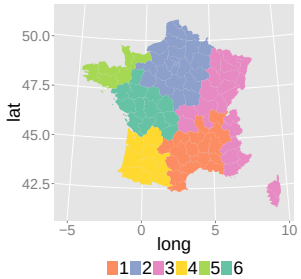

Figure 11: Final division for the 6 wind zones

3.2 Distribution of the portfolio on the zones

To understand the relative weight of each zone in terms of the insurance portfolio, we consider three quantities: the number of policies, the sum of charges and the frequency of claims. Particular interest lies in the average annual sum of charges for each zone. To estimate this average and its variability according to the presence or not of major storms as in 1999, we use the bootstrap method by creating 2500 different samples, each representing the daily charges over the whole estimation period. We must note that in our case with strongly heavy-tailed distributions this non-parametric method has some flaws. A small number of events has a very strong influence on the results here, which impedes formal statistical consistency [10]. However, the method remains interesting since it is based on a minimum number of model hypotheses. We use it for a first empirical study before moving to models with better theoretical support based on the theory of extreme values.

Table I: Distribution of annual portfolio and claim numbers (in 10^6 €) by zone

Zones	1	2	3	4	5	6
Nb. Contracts	155475	265222	248957	147899	52033	130413
Ratio	16%	27%	25%	15%	5%	13%
Nb. Dep.	17	25	24	10	6	13
Average claim number						
Daily sum						
$q_{2.5}$	1539	2806	2689	899	517	821
$q_{97.5}$	6498	12868	11480	7580	3353	7117
Average cost (million euros)						
Daily sum						
$q_{2.5}$	4.8	9.2	8.3	2.4	1.8	2.6
$q_{97.5}$	36.1	55.7	52.7	33.9	20.1	48.4
Weekly sum						
$q_{2.5}$	3.6	6.3	6.7	1.9	1.4	2.2
$q_{97.5}$	42.3	66.2	57.4	35.3	21.2	51.4

Table I provides an overview of the number of policies and claims in each of the six zones. The portfolio distribution is nearly proportional to the number of departments constituting the groups. Zone 2 accounts for the most departments and the biggest part of the portfolio. The distribution of insurance policies between zones is homogenous. The following lines contain quantiles at 2.5% and 97.5% of the costs and of the number of claims for each zone calculated from the bootstrap

sample. Distributions show high variability. Estimated costs are generally proportional to the size of the group, except for zone 6 whose quantile at 97.5% is of similar order as zone 2 and 3 (around 55 million euros). Large interquantile ranges appear also for the number of claims, for instance between 2800 and 13000 claims per year in zone 2. As opposed to what we had noted on costs, zone 6 (with claims ranging between 800 and 7100) no longer appears as overexposed when compared to other zones. Claim frequency is better aligned with exposition in terms of portfolio risk. In summary, we see that the observation period will always have a strong influence on the results of a model including extreme events.

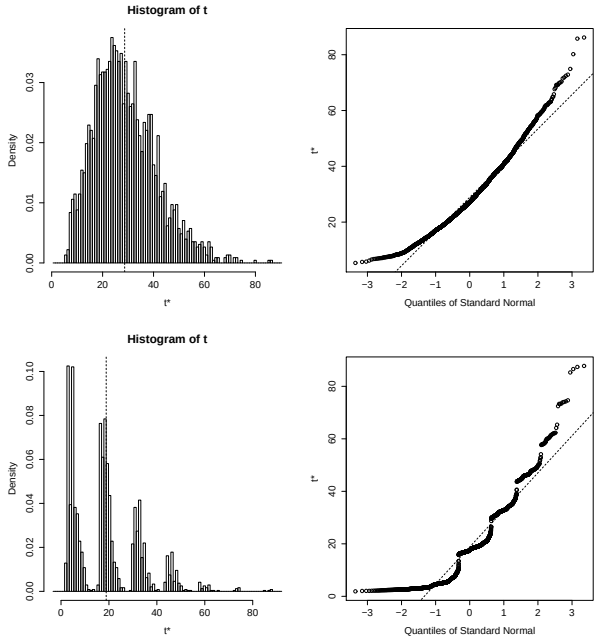

Figure 12: Bootstrap estimates of claims - in zone 2 (up) and zone 6 (down)

Figure 12 reports histograms of the bootstrap distributions on zones 2 (upper display) and 6 (lower display). Zone 2 spreads from the Paris region to the North of France. The histogram on the left represents the average annual estimates expressed in millions of euros. The distribution is asymmetric with a stretched right tail. It varies strongly with respect to the observation period, with a range from 10 to 120 million euros. The QQ-plot on the right opposes empirical quantiles to those from a standard normal distribution. The highly non-linear display indicates heavier tails in the data, which fully confirms the visual finding from the histogram. On the displays below, we once again present the bootstrap distribution for zone 6. This zone is smaller than zone 2. Indeed, it only counts 13 departments versus 25 in the latter. The shape of the distribution is slightly different. We notice several peaks, as if composed of a mixture of distributions. This zone actually illustrates the shortcomings of the bootstrap with heavy-tailed data. The periodicity in the bootstrap distribution may be due to a single huge event, and peaks appear when the event is drawn $0, 1, 2, \ldots$ times. For instance, it is known from elementary bootstrap theory that a particular event does not occur in a

bootstrap sample with a probability of about $\exp(-1) = .37$ when n (the number of observations) is high. Moreover, we can see the importance of geographical scale besides the temporal scale for the stability of results.

This bootstrap approach is based on daily data. It does not convey the notion of event and it overlooks dependencies between days. For comparison, we perform a bootstrap based on weekly blocks of aggregated charges to better capture events spreading over several days. In the last part of Table I, we present the quantiles at 2.5% and 97.5% of charges in average annual and in millions of euros. We can see a higher variability in the distribution of the average due to a grouping of clustered events for several days. The estimations for the low quantiles (2.5%) are lower than the ones in the daily approach. For instance, in zone 2, we have a decrease of more than 30% from 9.2 to 6.3 millions. For an upper quantile, the estimations based on weekly data are higher than the daily ones. Still in zone 2, we move from 55.7 to 66.2 million euros, i.e. an increase of almost 19%.

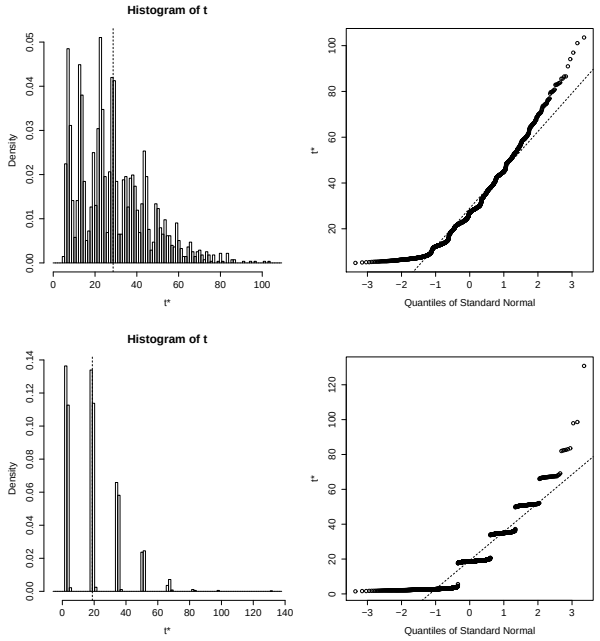

Figure 13: Bootstrap estimates of claims with week aggregation - in zone 2 (up) and zone 6 (down)

On Figure 13, we observe again the distribution of annual charges for zones 2 and 6 but this time from a weekly aggregation. The problems with the bootstrap approach appear amplified. The claim rate varies according to the observation period. We can finally conclude from this approach that we need more refined methods, such as a parametric model (GPD), taking clusters into account.

3.3 Spatial representation of main storms

In a first approach, we give a spatial representation of the distribution of claims during major events in terms of costs and compare them with available climatic variables. The different variables we have checked are temperatures, wind speeds, precipitation and North Atlantic Oscillation (NAO).

A preliminary analysis showed that only wind speeds are geographically correlated with the path of storms. Moreover, a wind index can provide better information because it enables to detect regions where unusually high wind force occurred. Damages are linked to the strength of the storm and the condition of the buildings, whose architecture or construction is more or less resistant to bad weather conditions depending on the region. The wind index formula we use is the one defined in [20]. Based on a series of daily maximum wind speed measurements $w^d(s)$ for a sequence of days $d \in D$ at station s, it is defined as follows:

$$I_w^d(s) = \left([w^d(s) - w_q(s)]_+\right)^\alpha, \tag{6}$$

where w_q is the q%-quantile of $w_d(s)$ for $d \in D$.The use of the exceedances of a quantile in the construction of a risk index allows us to focus on the most extreme values. This method is quite common [19]. We then present for each department the observed event in terms of the sum of costs, maximum wind speeds and wind index over a period of 3 days.

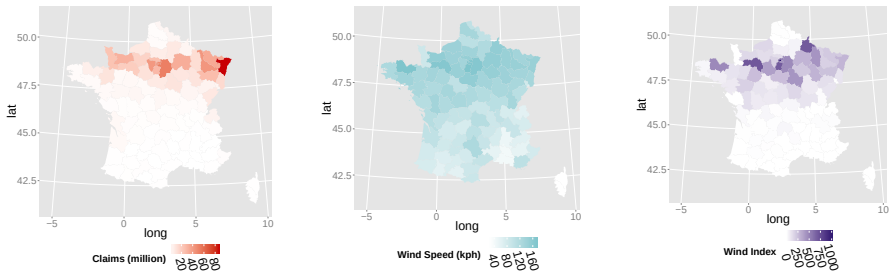

Figure 14: Storm Lothar on 12/1999 in France - claims, wind speed and wind index (from left to right)

The left graph of Figure 14 (representing the updated values in millions of euro) depicts the path of Lothar, which swept through France from West to East on a Brittany (around 4am local time) - Lorraine and Alsace (11am) axis with a front measuring 150 km wide. This storm was not a hurricane (tropical cyclone), albeit sometimes termed as such, but a depression of average latitudes unusually intense for Europe. The graph in the middle depicts maximum wind speeds recorded between December 24th and 26th. In the most impacted zones, we observe from an insurance point of view wind speeds among the highest, but this information does not suffice to precisely locate concerned zones. The graph on the right shows the variations of the wind index. The path of the storm is clearly illustrated in Northern France. The most impacted departments are not exactly the same which be explained though the lack of a notion for risk exposure (the number of policies) in our index formula. We will include this parameter in a refined version of the storm index. The path of storm Martin (Figure 15) mainly hit the southern half of France between December 26th and 27th. Again, zones with high damages correspond to high wind speeds. The wind index eliminates zones little affected by the storm in terms of insurance damage, like the Southeast and Brittany. Nevertheless, a high index does not always coincide with big damages for Allianz as we can see in the

departments situated in the center of France, but this is due to a lower density of urbanisation and risks in this zone.

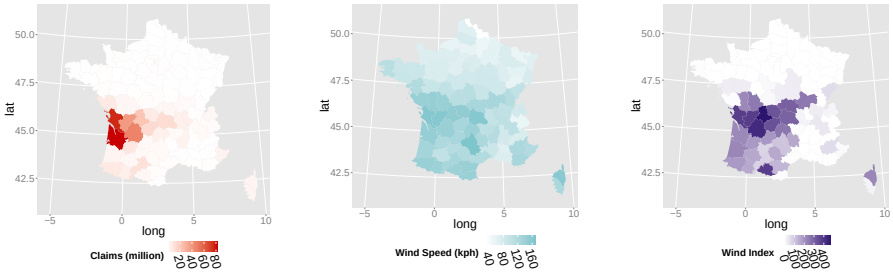

Figure 15: Storm Martin on 12/1999 in France - claims, wind speed and wind index (from left to right)

In conclusion, the variable "wind speed" can explain a large part of the damages, but not enough to fully describe such a complex phenomenon as a storm. This complexity necessitates to complement the estimates with a discussion about uncertainties, specifically for extreme events [27]. The wind originates from the difference of atmospheric pressure, and to understand wind damage it is important to distinguish two types illustrated in Figure 16. Synoptic wind is generated by continental or global phenomena; it is sustained, occurs on a large scale and can be violent. Local wind is due to temperature disparities; it is thermal wind. Indeed, the behavior and evolution of the wind flow will be different according to its type. Synoptic wind is generally slowed down by an obstacle (natural or artificial), whereas thermal wind is strengthened by the relief.

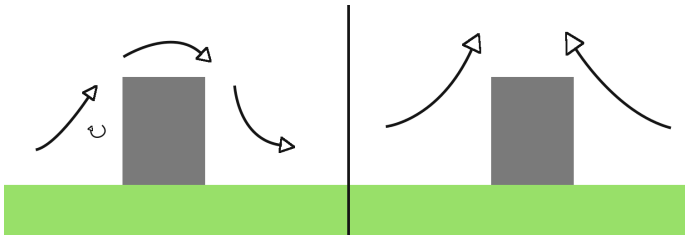

Figure 16: Two types of wind behavior (synoptic on the left, thermal on the right)

Wind direction is also important; combined with speed, they form velocity. Knowing these parameters and the duration of wind gusts could improve the understanding of storms.

4 Modeling

4.1 Definition of a storm index

Building a model from extreme event data that are rare by definition leads to preliminary questions about the choice of variables and their use, since the same raw data can yield very different results. Several approaches can be considered to describe the impact of a storm in space and time. We will focus on three criteria and on their influence on the model. The first one is the length of the observation history. The second concerns temporal aggregation in terms of either daily or weekly charges and the resulting change of perception of storm frequency. The third criterion is related to the existence and knowledge of the event considered as the most extreme in the observation period. We here test the sensitivity of the fitted distribution and of the return period in particular with respect to the presence or not of cyclones like Lothar and Martin in our data.

In this section, we propose an improved and more flexible version of the storm index, defined in a previous article [20], and with its extrapolation beyond the values observed. The index can be studied over a longer period than damage data, it presents less nonstationarities and can exploit a finer spatial resolution. In each of the six risk zones $A(k)$ previously delimited, we build a specific storm index combining wind indexes and exposure for each department in terms of portfolio risk. For $k \in \{1, ..., 6\}$ we define $I_S(k)$ as

$$I_S(k) = \sum_{s \in A(k)} R(s) \times \max_{d \in E} \left(\frac{I_w^d(s)}{N^d} \right), \tag{7}$$

where in each station $I_w^d(s)$ is the wind index of day d, and the number of risks $R(s)$ is weighted by the number of contracts. We also take into account the size of the damaged area (geographic aggregation), the duration of the storm event E [2] and the number of active stations for day d, N^d. We stress that we choose the maximum and not the sum of daily indexes for the temporal clusters associated to each storm. As opposed to insurance costs where are added over the duration of the event, wind peaks better reflect the relative intensity of storms events. We then obtain a global storm index on the whole country by adding up the six storm indexesm weighted both by the portfolio and by a parameter vector $B = \{\beta_1, ..., \beta_6\}$ with $\beta_1 = 1$ fixed. This additional parameter allows us to over- or underweight each zone according to the impact of a hypothetical storm passing through the territory in terms of property damages. The formula becomes

$$I_S = \sum_{k=1}^{6} I_S(k) \times \beta_k, \tag{8}$$

[2] For each event $E \subset D$ we select the daily maximum, for example, during the storm Klaus: $E = \{01/23/09, 01/24/09, 01/25/09\}$.

We have chosen the parameters of this index such that it reflects at best the distribution of the most extreme damages in terms of rank. Our ranking for the most important storms since the 1970s is based on [18]. Based on this ranking, we propose a final modification to the index values such that they become approximately proportional to aggregated claim amounts. Therefore, to make the index value distribution more heavy-tailed, we use the exponential function and a last parameter γ, estimated with a Least Squares method. The link between our storm index and the cost C for the insurer now is

$$C = \exp\left(\gamma \times I_S\right). \tag{9}$$

4.2 Statistical modeling of the storm index

Extreme values theory [6] has spawned two complementary approaches: the block maximum and the threshold exceedance. Here, we use threshold exceedance modeling through the Generalized Pareto Distribution (GPD). With this model and for an appropriate threshold u, the distribution function of $Z = (X - u)$, knowing that $X > u$, is defined by:

$$H(z) = 1 - \exp\left(-\frac{z}{\sigma}\right), \xi = 0,$$

or

$$H(z) = 1 - \left(1 + \frac{\xi z}{\sigma}\right)^{\frac{-1}{\xi}}, \xi \neq 0,$$

according to the value of the shape parameter ξ controlling the tail of the distribution. We have used the `extRemes` package [8] of the software `R` to estimate the different distributions of extremes. We have consider two different observation periods: the first one over twenty years (1993-2013), the second one over forty years (1970-2013), with data for the first period being the most homogenous. Insurance records do not reach back to such long periods. We work with the storm index I_S and fit the parametric GP distribution to its tail. To begin, the threshold u must be fixed, here done by considering two standard visual methods. The first one plots estimates of shape and scale parameters according to different thresholds and seeks a threshold level above which estimates for different thresholds remain within the confidence interval of all estimates above the threshold. The second one is the mean excess plot confidence intervals.

Figure 17 shows the resulting diagnostic plots for threshold selection. Both graphs on the left represent variations of σ and ξ with their confidence intervals at 95% according to the considered threshold values in the range between $u = 5$ and $u = 40$, where the index has been calibrated such that its observed maximum (Lothar storm in December 1999) is 100. By observing the confidence intervals increase in size for larger u, we here choose $u = 20$ as the upper bound above which the variability of estimates becomes too important. The graph on the right depicts the mean excess above

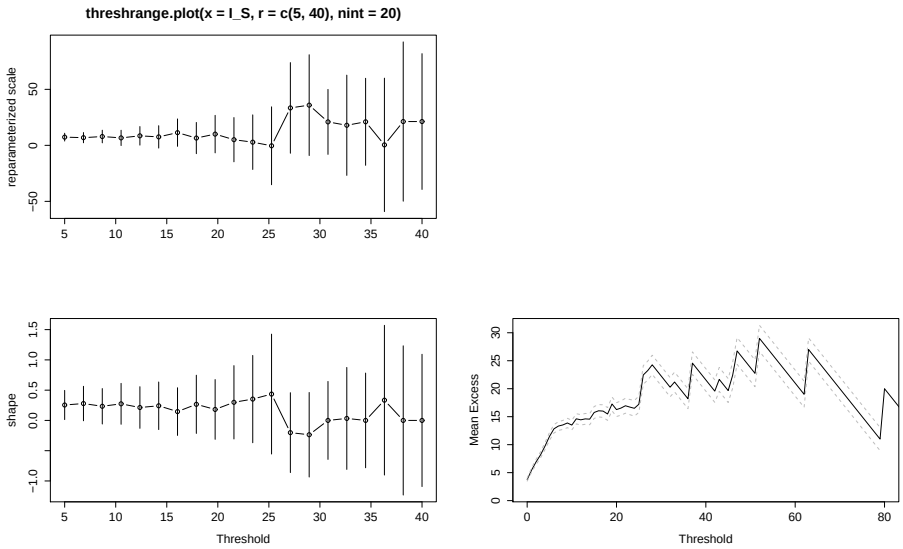

Figure 17: Threshold selection diagnostic plots for the GP fit to the storm index

the threshold with the confidence interval at 95% for thresholds from 0 to 80 (dotted). In coherence with the first results, the curve gets erratic around $u = 20$. We choose to study thresholds $u = 10$ and $u = 20$ in more detail for our models.

We further propose an analysis of sensitivity with respect to Lothar and Martin, the two most important events of our observation period. We observe how return periods change if we remove Lothar and/or Martin from the data. Our overall aim is to understand the robustness of the model (or lack thereof) when quantifying events with return periods of over 70 years or even 200 years based on a data record of less than 50 years.

4.3 u = 10 and period = 1970-2013

We first estimate the parameters of the Generalized Pareto distribution with a threshold $u = 10$ and over the period 1970-2013. The threshold is exceeded on 62 occasions, that is to say around 1.4 times per year over the observation period. The parameters are estimated according to the maximum-likelihood method (MLE). Figure 18 provides a QQ plot of data quantiles against fitted model quantiles, a comparison of quantiles from model-simulated data against the data, a nonparametric density estimate of the data along with the fitted model density and a return level plot.

Table II: Maximum Likelihood Estimates : I_S with u=10 and period=1970-2013

	MLE	Std. Err.
Scale (σ):	10.78	2.13
Shape (ξ):	0.20	0.15
Negative log-likelihood :	222.08	

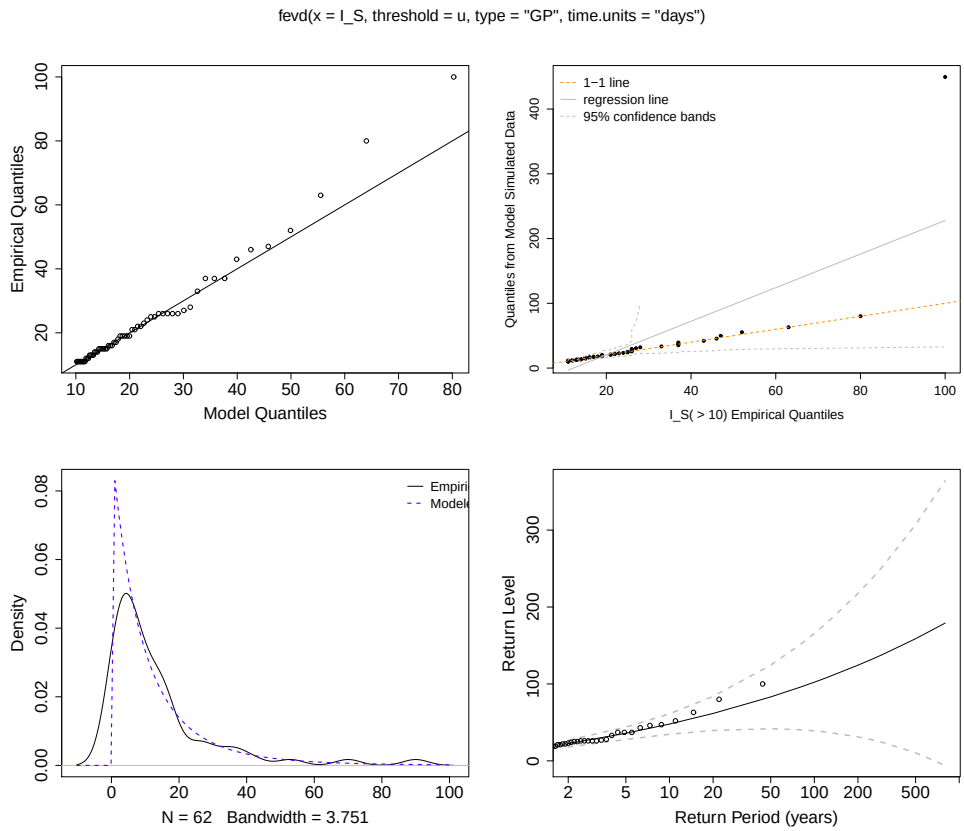

Figure 18: Results of the GPD analysis for the storm index.

The parameter ξ is positive (0.20 in Table II), which corresponds to a heavy tail and a distribution function belonging to the Fréchet maximum domain of attraction. Both graphs on the upper row of 18 are QQ plots. On the left one, we note a slight gap in the strong quantile values but a good global alignment. On the right one, the alignment of the dots is excellent to compare the model probabilities and empirical probabilities, except for a single outlier point. The density function of the model is confronted with a kernel estimate on the bottom left of the graph. We notice a good correspondence away from the lower bound 0 where the density estimate is not reliable. All return level estimates stay within the 95% confidence interval. This is done on the log-scale for the abscissa so that the type of GP distribution can be seen from the shape, where heavy-tailed distributions are concave, exponential-tailed distributions are straight lines, and distributions with finite upper bound are convex). In our case, the estimate seems to indicate the upper-bounded Weibull maximum domain of attraction. The red dot in Figure 19 indicates that storm as strong as Lothar occurs approximately every 93 years. However, when interpreting results beyond the observation period of forty years, uncertainty is high and the lower bound of the confidence interval decreases strongly, calling for cautious interpretation.

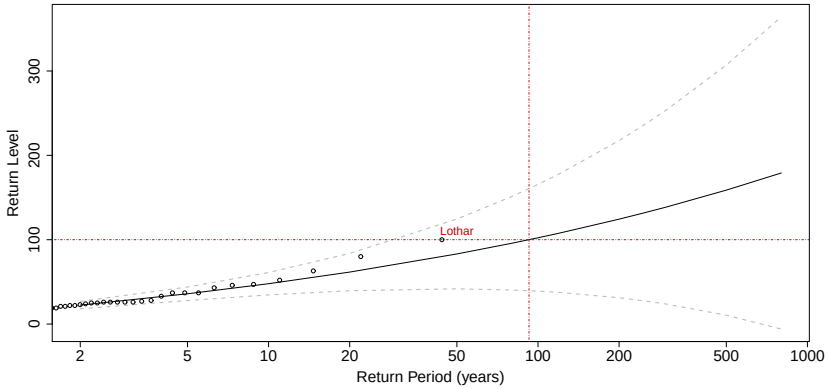

Figure 19: Return period for the storm index.

4.4 u = 10 and period = 1993-2013

Still with the same threshold, we now want to compare previous results with those obtained on a period twice as short, spanning 20 years between 1993 and 2013. During this period, the threshold is exceeded only 20 times, one average once a year. The estimation of the parameters appears on Table III, and we see that standard errors have become higher.

Table III: Maximum Likelihood Estimates : I_S with u=10 and period=1993-2013

	MLE	Std. Err.
Scale (σ):	9.87	4.18
Shape (ξ):	0.50	0.38
Negative log-likelihood :	75.76	

The scale parameter σ descreased from 10.78 to 9.87. The parameter ξ increases to 0.50, therefore we still deal with a heavy-tailed distribution. The return period of an event such as Lothar strongly deminishes, as we can notice on Figure 20. We now obtain that a level 100 storm on our index could occur every 31 years. We note here that the length of the observation period greatly influences the calculation of the return period. With this limited data record, we tend to overestimate the probability that such an event happens. We thus focus for the rest of this study on the longest available data record from 1970 to 2013.

4.5 u = 20 and period = 1970-2013

The following study is based on the 1970 to 2013 period, but with a higher threshold $u = 20$. This threshold is exceeded by our storm index on 26 occasions, on average 0.6 times per year. The estimates of the model parameters are reported on Table IV.

The estimated scale parameter σ increased from 10.78 to 12.96. The shape parameter ξ is estimated as 0.21, a value very close to the one obtained with the threshold $u = 10$ over the same period. However, the standard error increases. In terms of return levels (Figure 21), we get a storm as strong

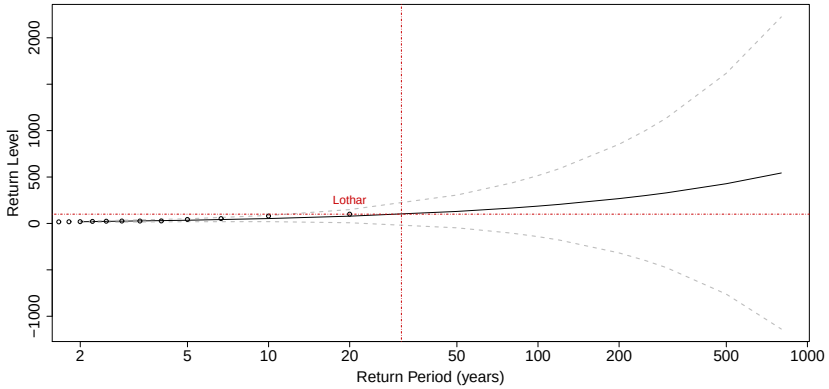

Figure 20: Return period for the storm index.

Table IV: Maximum likelihood estimates : I_S with $u = 20$ and period 1970 to 2013

	MLE	Std. Err.
Scale (σ):	12.96	4.20
Shape (ξ):	0.21	0.26
Negative log-likelihood :		98.09

as Lothar every 88 years. This prediction is slightly lower than the one from the first model, but remains of the same order of magnitude.

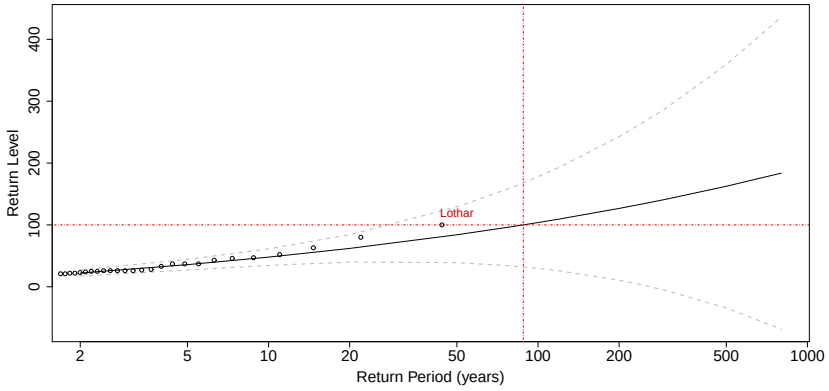

Figure 21: Return period for the storm index.

By modifying the threshold u of the model between 10 and 20, the return period of the major event for our observation period (Lothar) decreases of only 1.06% from 93 to 88 years. To check the sensitivity of results, we reproduce the same estimations by removing Lothar from the database in the following sections.

4.6 u = 10 without Lothar and period = 1970-2013

We replicate the base model ($u = 10$) over the longest period of time (1970-2013), but we now emulate a situation where the storm Lothar did not occur. In other words, the most substantial damage since the 1970s would does not exceed the level of storm Martin, which is two times lower than Lothar in

terms of insurance claims in France. The estimations of the model parameters are given in Table V.

Table V: Maximum likelihood estimates : I_S without Lothar

	MLE	Std. Err.
Scale (σ):	11.02	2.15
Shape (ξ):	0.10	0.15
Negative log-likelihood :		213.54

The estimated scale parameter σ increases from 10.79 to 11.02. The parameter ξ is still positive (0.10) but decreases compared with the Lothar model (0.20). This decline can be explained quite naturally: the distribution tail becomes less heavy because the strongest value has been removed. The standard error also increases. Regarding the return periods of the model (Figure 22), a storm as important as Lothar occurs once about every 280 years. Without Lothar, we multiply the return period of such an event by 3, a drastic change. This example substantiates the fragility of this type of modeling, whose forecasting radically changes when a new record value occurs.

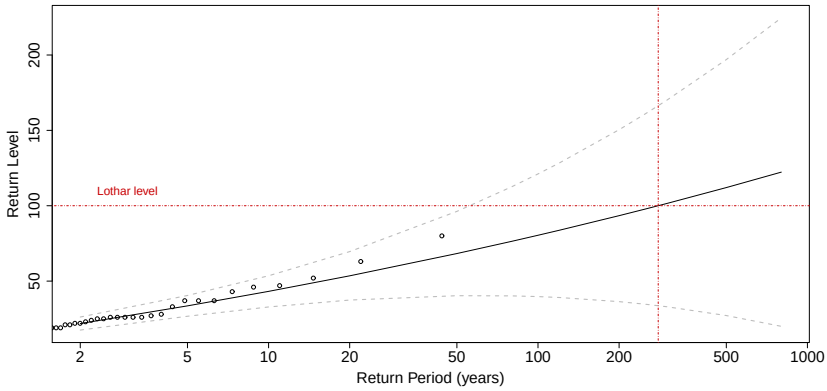

Figure 22: Return period for the storm index sans Lothar.

4.7 u = 10 without Martin and period = 1970-2013

Following the same idea as with Lothar, we propose to reconsider absence of the second most important storm Martin from the data record. The estimations of the model parameters are reported in Table VI.

Table VI: Maximum likelihood estimates : I_S without Martin

	MLE	Std. Err.
Scale (σ):	10.61	2.04
Shape (ξ):	0.16	0.14
Negative log-likelihood :		214.60

The scale parameter σ slightly decreases from 11.02 to 10.61. The parameter ξ is positive (0.16) and barely decreases compared with the complete data model (0.20). In terms of return levels (Figure 23), a storm as important as Lothar occurs once every 160 years. Without including Martin,

the return period is 1.7 times higher for Lothar. In this numerical experiment, the model is again strongly determined by the omission of a major storm. The scale change is not as important as with the omission of Lothar but it remains considerable.

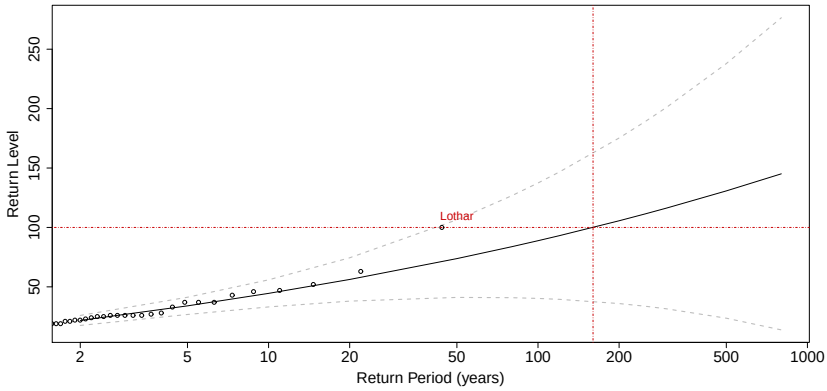

Figure 23: Return period for the storm index without Martin.

4.8 Summary and complement

The GPD model heavily depends on the most extreme value for the estimation of the parameter ξ, which controls the tail distribution. The estimation of ξ decreases in the absence of Lothar from the data record, and the heavy-tailed model comes close to a light-tailed model ($\xi = 0$) if we take the standard error into account. However, the estimates of the scale parameter σ are stable. Therefore, return periods of extreme events strongly increase. The observation period also plays a crucial role; we noticed that a too short duration can entail a strong underestimation of return periods. For astorm as strong as Lothar, this period varies between a minimum of 31 years and a maximum of 280 years in our numerical experiments. The removal of an important intermediary event like Martin also impacts the results of the model through an increase of return periods. Table VII shows a summary of the return period variations for an event similar to Lothar according to the considered methods and scenarios. To complete this analysis, we have added the estimations for a loss event which would exceed Lothar by a factor of 1.5, $L(1.5)$. To determine the value of the index I_S associated to this level of loss, we reproduced the Formula 9.

$$L(1.5) = \frac{\log(1.5)}{\gamma} + 100 \tag{10}$$

Table VII compares the return period variations according to the magnitude of the storm. For the first two estimations (between 1970 and 2013, or all data), there is an almost linear increase of the event frequency according to the induced damages. This increase reaches a maximum of 490 years in the scenarios with the omission of a major storm. The strongest variation is observed on the period

Table VII: Calculations of return period according to different scenarios

Period	1970-2013				1993-2013
Threshold : u	10	20	10wL	10wM	10
Lothar Return Period	93	88	280	160	31
L(1.5) Return Period	136	128	490	252	84

1993-2013 with a multiplication by 2.7, where we move from 31 to 84 years. Such long time intervals may be beyond imagination, but they have direct consequences on the quote and reinsurance costs. Moreover, the perspective recommended by Solvency II to insurance companies imposes to work with quantiles at the 99.5% level, corresponding to vents with a return period of 200 years.

5 Capital requirements

Storms have a major impact on portfolios for property damage. In 1999, Lothar and Martin represented an aggregated claims equivalent to an average year. In a single year, insurance companies therefore had to support the equivalent of around two years of losses, which is not without risks for their own capital. Within the European project Solvency II, supervisory authorities seek to determine rules to limit the risk of financial ruin for insurance companies. The first pillar of this regulation defines the norms for calculating regulatory own capital. The SCR (Solvency Capital Requirement) corresponds to the need in capital required for the insurer to keep his commitments over the span of a year. A confidence interval at 99.5% was retained for annual costs. Here, hedging mechanisms (reinsurance) allow to smooth extreme situations by establishing a form of mutualization in space and time. They are technical and a corresponding market exists. However, the needs in own capital are generally measured after reinsurance.

In view of the complexity of these systems which are specific to each company, we focus on the sensitivity of raw results within Solvency II. We study the target capital needed to absorb the shock caused by a risk resulting from extreme and irregular events. The European Supervision Authority of Insurance and Pension defined a standard formula of the SCR. For an annual variable expense C, the formula is

$$\text{SCR} = \text{VaR}_{99.5}(C) - \mathbf{E}(C),$$

with $\mathbf{E}(C)$ the average annual expense calculated on the observation period and $\text{VaR}_{99.5}(C)$ the value at risk at 99.5% of the distribution of C.

5.1 Calculation of the average annual expense

To understand how much exceptional expenses as in 1999 influence the average annual expense, we propose to include it in our calculations according to the different return periods obtained in the previous section. According to the historical records [5] of the FFSA, 1999 alone represents 30% of the total costs over the period 1982-2012. In the following example, we will consider that the updated total expense of a company over a 30-year period amounts to 5 billion euros. The annual expense of 1999 would have reached 1.5 billion euros. The average annual charge due to 1999 then equals 116.7 millions.

Table VIII: Evaluation of the average annual expense (in millions of euros)

1999 Return Period	31	88	93	160	280
Claims : 1999	48.4	17.0	16.1	9.4	5.4
Average annual claims	165.1	133.7	132.8	126.1	122.1

Table VIII shows that the sum of costs updated for Lothar represents more than two thirds of the total of the period costs. Depending on the chosen return period, the average annual expense varies between 165.1 million (31 years) and 122.1 million euros (280 years). The contribution of year 1999, which includes the two main events of the period, is evaluated at 48.4 million or 5.4 million, depending on two extreme scenarios.

These differences show us that it is difficult to obtain reliable results. Using tools for smoothing and extrapolation may create a false illusion of mastering the problem of rare events. Models may be more or less appropriately chosen, but essentially there always is a lack of precision when evaluating the return period of rare events with high impact.

5.2 Measurement of the dispersion of claims

5.2.1 Can we measure the dispersion of claims?

To tackle the problem of dispersion, we need to bear in mind the context of insurance procedures. Companies generally work on annually aggregated net charges after reinsurance. Calculating net results poses the problem of quantifying the reinsurance regime which depends on technical, strategic and business conditions that are variable over time. Therefore, in our theoretical study we prefer to remain at the level of gross reinsurance. However, the elements at hand should allow the simulation of reinsurance, which necessitates the knowledge of costs per events. Ideally, we should work on this type of measurement to correctly address the issue of dispersion measurement.

5.2.2 Which observations are available to perform measurements?

Regarding the most important insurance events, we can access historical records covering around 40 years [18]. This benchmark of major events for the four last decades is based on the basis of the whole insurance. The update over a relatively long period on a storm insurance scale is difficult. Among other aspects, we need to consider the evolution of construction indexes (FFB or RI), of the relative weight of individual and business segments in the portfolio, of the coverage proportion of the storm guarantee which is only mandatory since July 1990, or of the progress of real estate in France. Figure 24 reports the updated costs of the main storms since 1970 according to market data. The year 1999, exceptional and unprecedented in France, substantiates the high variability of the storm guarantee.

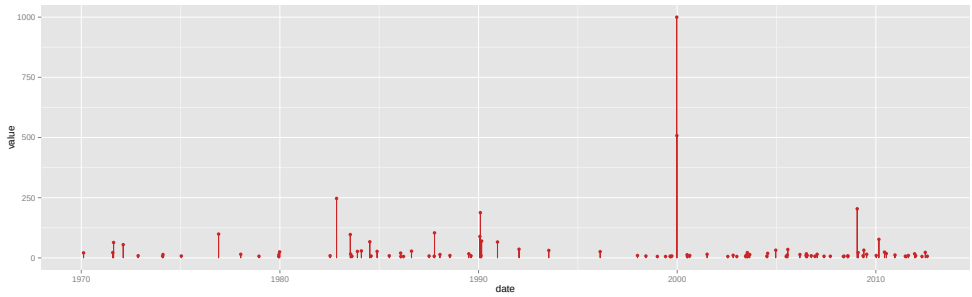

Figure 24: Updated storm costs in France since 1970

We can also work on a *company* basis over shorter periods with a fully detailed claim record which provides more complete information about major events. Since 1998, Allianz lists all the individual wind-related claims with a departmental and daily precision. It represents over 520.000 claims, mostly for individuals (310.000). On a premium and claim basis, Allianz's market-share for the storm guarantee in mainland France is about 10% over the whole period. Moreover, costs carried by this company are highly correlated (0.99) to the market costs. Nevertheless, the update difficulties remain present.

5.2.3 Which assessment can we make?

As shown in section 5.1, data records are too short to derive average expenses, let alone to reconstruct the complete distribution of the annual costs. In general, three groups of claims can be distinguished. A group including the largest number of events (low segment), another one including around a dozen of events per year (intermediary segment), and a last group concerning only the most extreme storms from the observation period (high segment). Concerning the terminology of such events, we talk about frequent, intermediary or extreme/rare events. In practice, these three groups are not modeled through a unique distribution, but according to a mixture of different and adapted distributions. Moreover, uncertainty is not identical the three segments of the overall distribution. Available data are usually allow us to satisfactorily reconstruct the low and intermediary parts of the cost distribu-

tion, which is not the case for the most extreme costs.

5.2.4 Current solutions and a new approach

Various solutions are already used by the market actors. They can be distinguished in terms of the database used for their construction and of the calculation method. The distribution function is sometimes estimated empirically from the insurer's records. In other cases, it is based on models coupled with Monte-Carlo techniques to simulate a great number of years. We distinguish three types of approaches: classic models (annual frequency and intensity function, by level of costs); market models (using event catalogues, damage functions, location information about the portfolio) and our model built from daily costs and storm index statements.

5.2.5 Do these solutions provide uncontestable answers?

Theoretical, technical or material difficulties will always compromise the robustness of results. One should be aware of these difficulties and associate them to the models to better understand their qualities and shortcomings and to be able to interpret their results cautiously.

In classic models, functions are defined from a limited number of observations. The fragility increases with the level of costs. The definition of extreme distributions is derived from a very limited number of observations. We sometimes have to settle for a single annual observation, or even no observation at all. By extrapolation of this poorly known distribution of extremes we can obtain the full distribution.

Market models try to be abstract away from observed data by using external data like catalogues of artificial events and damage functions. Validating a model in this context is difficult since it cannot be compared to observations. For instance, it is difficult to validate a catalogue of 18.000 events, generated for France, based on only 600 observations. How can we validate damage functions when no observation is recorded on the selected criteria? No claim file in France has recorded wind speed and damage rate. These models did not inform us on the data related to damage functions. The work carried out in 2012 and 2013 by [18] showed us how much the results delivered by one of these models were disconnected from reality (in particular for the average on non-extreme costs and on geographical restitutions).

Our choice to work with daily data allows us to preserve a maximum of information in a context of historical data records available only over a relatively short time span. This approach then requires a transition from daily to an annual cost distributions. Since we refute the independence of consecutive daily events, we must find a way to estimate dependence over several days and integrate it into our model.

5.2.6 Our model

We have chosen to model updated daily costs according to the GPD law. We first consider the costs without covariates and then with the storm index in covariates. In Table IX, we present an estimate of settings for the distribution law, first in the entire France, then according to six storm zones previously defined.

Table IX: Maximum Likelihood Estimates : claims (global and areas)

Zone	Global	Negative log-likelihood			1228.63	
Scale e+06 (σ):	1.51	AIC			2461.26	
Shape (ξ):	1.12	BIC			2465.92	
Zone	1	2	3	4	5	6
Scale e+05 (σ):	4.28	5.88	3.26	2.10	0.96	3.29
Shape (ξ):	1.03	1.17	1.12	1.26	1.14	1.09

In all these models, the shape parameter ξ is higher than 1 for the costs, which corresponds to a distribution of very heavy-tailed extremes. The mathematical expectation of such a distribution is infinite, which shows how strong the influence of extremes is on all the categories of claims. Note that the difference between the ξ-estimates of the costs and of the storm index is explained through the exponential function in formula 9: taking the exponential reflects the explosion of costs which takes place for the most extreme storms. By estimating the parameters according to geographical storm zones, the shape parameter varies between 1.03 for zone 1 and 1.26 for zone 4. Regarding estimated scale parameters, variations are approximately proportional to the exposition to risks according to the portfolio. These models can be improved by integrating the storm index as a covariate into the shape parameter ξ or the scale parameter σ. To decide on how to integrate the covariate, we use Figure 25 which confronts the logarithm of costs exceeds with the storm index for the twenty-five most important exceeds. The distribution of points is close to linear, in particular for the most extreme

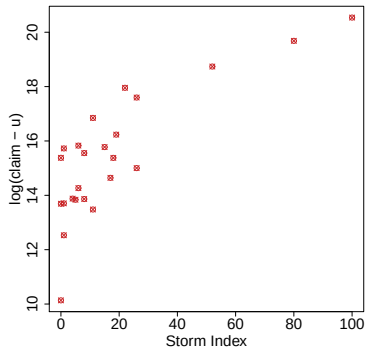

Figure 25: Logarithm of costs according to the storm index

events. Since the logarithm of observations from a GPD with positive shape ξ is approximately exponentially distributed with scale ξ, Figure 25 indicates a good approximation through a GPD

whose shape parameter is linearly related to the storm index. On Table X, we report estimates for the model with the storm index integrated as covariate into the shape parameter via $\xi(I_S) = \xi_0 + \xi_1 I_S$. The covariate leads to a better model according to model selection criteria AIC and BIC, which both improve.

Table X: Maximum likelihood estimates : claims with storm index as covariate in the shape parameter

Zone	Global
Scale (σ):	1.51e+06
Shape (ξ_0):	1.50e-03
Shape (ξ_1):	9.25e-02
Global Negative log-likelihood	1210.78
Global AIC	2427.55
Global BIC	2434.55

We now have fitted an appropriate GPD model to high daily threshold exceedances, yet we aim to build a model for the overall annual costs. Sinc the shape index of the tail distribution is very high (ξ around 1), we can conclude that the lower and intermediary parts of the distribution have little influence on the annual distribution since the few strongest daily values will dominate the sum of daily values. To properly transform the daily distribution into an annual one while respecting temporal dependence among successive days, we need the distribution of the number of exceedances over a year. Under the assumption of independent days, we would have a binomial distribution with parameters $n = 365$ and $p \in (0, 1)$) the probability of exceedance. Since n is high and p is low, we can approximate it through the law of rare events, the Poisson distribution with intensity parameter $\lambda = n \times p$. It has a variance/expectation ratio $\lambda/\lambda = 1$, known as its dispersion index.

Owing to dependence between days, we get clusters of extreme events, which leads to overdispersion with a dispersion index > 1. This stronger variability in the number of exceedances is translating into a stronger variability of annual costs. We here choose the negative binomial distribution as a classical distribution well adapted to overdispersion. It can be seen as an extension to the Poisson distribution with an additional parameter to take into account the degree of overdispersion.

As the main sources of clustering in threshold exceedances of daily aggregated claim amounts due to wind storms, we can consider (in decreasing order of temporal extent) seasonal behavior, atmospheric perturbations stretching over several days or weeks, and extreme storms stretching over at least two days. Of course, these three sources are closely related.

Seasonal behavior in wind speeds and the resulting insurance claims lead to a natural cluster behavior on a relatively large temporal scale, see Figure 26. Most extreme events, here given as exceedances over a fixed high threshold, occur in summer (June to August) and winter (December to February). We present estimates of the extremal index, based on the Ferro-Segers estimator ([7]), on the left of Figure 27. The extremal index is a measure of serial dependence of extremes in stationary time series. We should note that estimation here certainly is biased towards too strong dependence

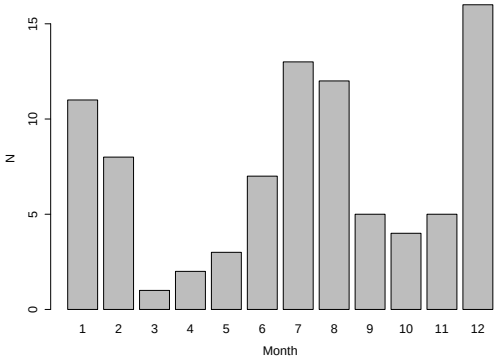

Figure 26: Number of threshold exceedances in claim numbers by month

due to seasonal clustering. Estimated values indicate weaker clustering for the most extreme events. We further propose to estimate the index of dispersion $\mathbb{V}(N)/\mathbb{E}(N)$ based on the overdispersion parameter θ of the negative binomial distribution, here parametrized as $\mathrm{NB}(\mu, \theta)$ with expectation μ and variance $\mu + \mu^2/\theta$. Estimates of the dispersion index for the number N of yearly exceedances above a high threshold is given on the right of Figure 27, letting us conclude that there certainly is some overdispersion since values are between 2 and 3. They are relatively far from 1, the reference value for Poisson distributed N when there is no clustering. The data used for estimating consist of 15 count observations for 15 observation years.

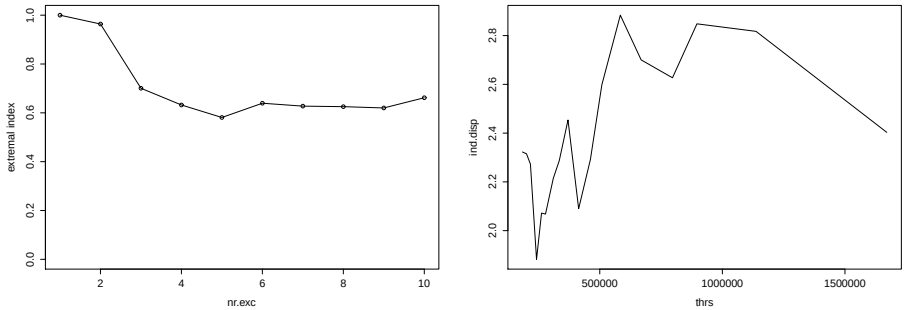

Figure 27: Estimates of the extremal index (left) and of the dispersion index (right)

A principal objective in the insurance context is to predict the yearly aggregated claim amount and its variability. For regulatory reasons, a good prediction of high quantiles of yearly aggregated claim amounts ("Value at Risk") is crucial. Due to estimated tail indices around 1 (see the Hill plot in Figure 28 for daily aggregated claim amounts), tails are very heavy and, at least theoretically, may correspond to distributions where neither expectation nor variance are finite. In practice, though, insurance coverage limits impose a finite upper bound to this distribution. Still, in our context yearly aggregated claim amounts will be strongly dominated by the few most extreme daily events having occurred during the year, and often already the single most extreme daily event may be dominating,

i.e., it will be of comparable size than the sum of all other events.

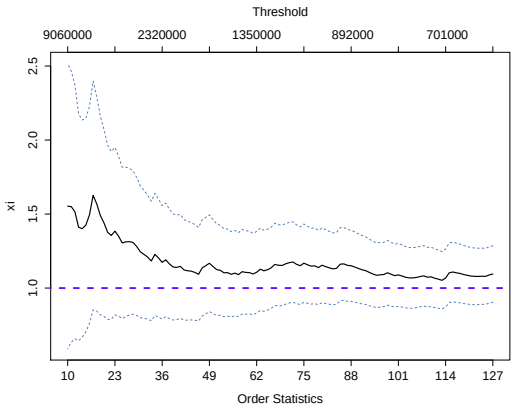

Figure 28: Hill plot of daily claims

We propose an approach based on Monte-Carlo simulations to reconstruct the distribution of yearly aggregated claim amount. We need an appropriate model for the number of yearly exceedances above a fixed high threshold, and we further use a generalized Pareto model for the corresponding daily exceedance distribution. We will not explicitly model the distribution of daily claims below the threshold, but instead use a block bootstrap approach to correctly capture their temporal dependence and their contribution to yearly simulated values. To simulate yearly claim amounts, we will therefore fix a high threshold u and use the following three elements: 1) a model for number of exceedances N, here chosen as the overdispersed negative binomial distribution 2) a GP model for the exceedances $Y \stackrel{d}{=} (X-u)|X>u$ to simulate extreme events $u+Y$ 3) a block bootstrap scheme for non-exceeding daily events, for instance based on monthly blocks with extreme events removed, where the January block is sampled from the 15 available January blocks, and so on. The advantage of this modeling approach is its robustness since we do not have to explicitly model clusters of extreme daily events. For instance, removing the influence of seasonal behavior and taking into account atmospheric perturbations ranging over several weeks would be rather complex.

Figure 29 shows the distribution of the yearly aggregated claim amounts (transformed to the $\log_{10}$-scale), based on 100.000 simulations according to our model. We investigated four threshold choices, corresponding to $3, 6, 9, 12$ exceedances per year on average. Values at Risk for probabilities 0.95 and 0.995 are indicated. The choice of threshold has some impact on quantiles (between $10^{8.78}$ and $10^{9.01}$ for $p = 0.95$, and between $10^{9.11}$ and $10^{9.60}$ for $p = 0.995$), with a systematic shift towards higher quantile estimates when choosing higher thresholds. We further observe that the impact of the resampling scheme below the threshold is negligible for these high quantiles, since the aggregates of resampled values below the threshold rarely exceeded 10^8 over one year. When using only 3 exceedances on average, we can discern a second mode in the distribution that disappears for the higher numbers of exceedances. This second mode is due to the resampling scheme below the

threshold, which yields a less smooth distribution when the threshold for exceedances is set very high.

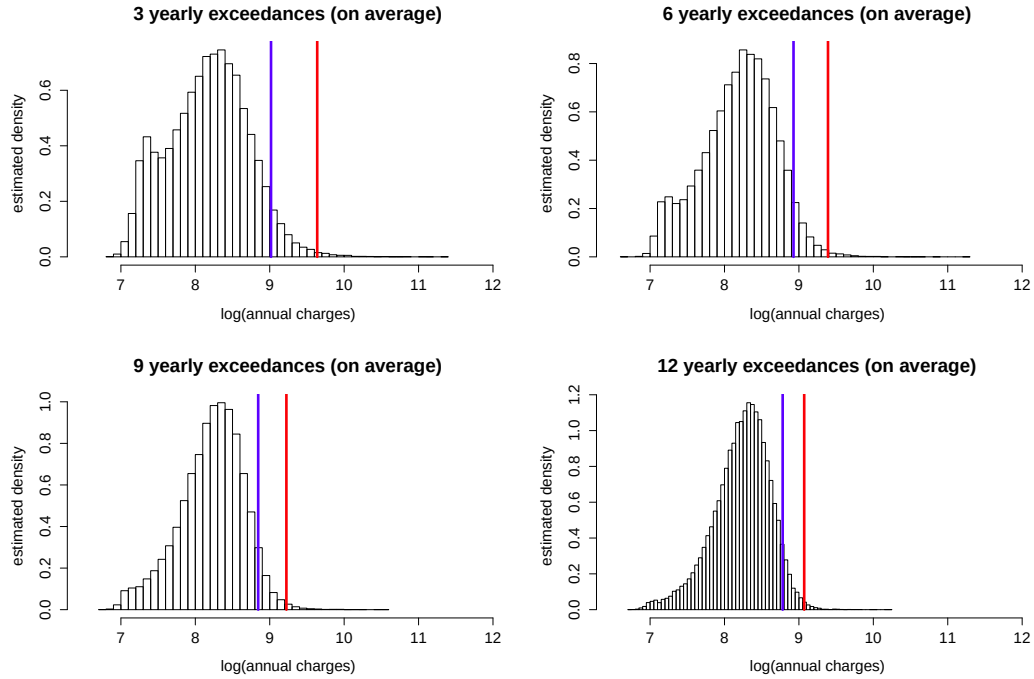

Figure 29: Simulated yearly claim distribution on a log_{10}-scale with VaR at 0.95 (blue) and 0.995 (red)

It is important to analyze these simulated distributions, here based on the Allianz portfolio, from an insurer's practical point of view. Besides the initial constraint imposed by a limited insurance database, we show the volatility inherent to a statistical model which incorporates internal assumptions. The aggregated claim amount for a return period of twenty years (VaR at 95%) spreads out between 605 million (estimation based on 12 exceedances per year : Figure 30) and 1.031 billion euro (based on 3 exceedances per year). The level associated to a 200 year return period (VaR at 99.5%) is estimated between 1.286 billion (12 exceedances per year) and 4.016 billion euros (3 exceedances per year). The considerable range of estimates depending on the number of exceedances illustrates the strong sensitivity of the results to model hypotheses and the amount of available data, particularly for predictions over long periods. It is worth noticing that when we focus on the most extreme data, we tend to have more pessimistic predictions. Beyond these relative variations, it is necessary to underline the discrepancy with regard to the current insurance market. Allianz property global premiums are close to 1.5 billion euro per year; hence, a capital requirement of almost 4 billion is problematic. Our results confirm that the uncertainty in the estimation of quantities required by capital reserve rules like those stated in Solvency II makes application of these rules difficult for certain types of risks.

Moreover, global warming and its potential impact in the coming years increases uncertainty.

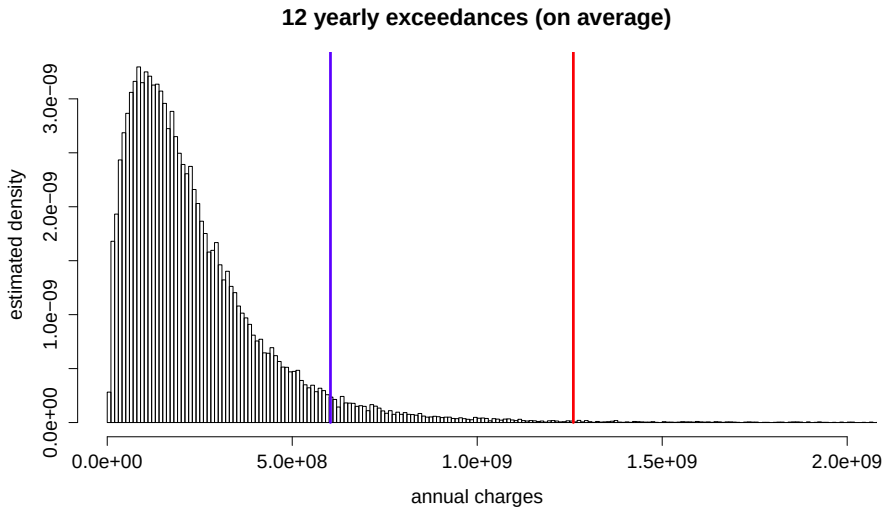

Figure 30: Simulated yearly claim distribution with VaR at 0.95 (blue) and 0.995 (red)

Some studies such as those of the AFA[3] or of the CCR[4] announce a strong increase in costs for all climatic hazards over the next 25 to 35 years. However, the contribution of global warming to these additional costs is not the same according to the studies (between 20 and 30%), and it seems to concern only the damage related to floods and droughts. No clear tendency is therefore identified for storms.

It is also difficult to associate an increase in wind speeds of $X\%$ to an increase of the damage of $Y\%$, as we can see in some models of the market. In the simulations based on our storm index, the consequences of a variation of wind speeds are different depending on the magnitude of the event, the size of the affected area and its location.

Concerning temporal clusters, we state that certain atmospheric conditions can foster the development of storms or prevent them. Historical insurance data are available over a few decades and already implicitly represent these conditions. They contain cluster periods like the spring of 1990 or the month of December of 1999, and more quiet periods. A more explicit model could allow us to better take into account clustering behavior, but the main question remains the validation of involved parameters.

6 Conclusion

Our storm index was designed to estimate and compare the scale of damages caused by wind phenomena during the most extreme events observed over several decades. The interest of this approach lies in the fact that we do not face the same difficulties and needs when measuring variables like wind or damage and their extremes. We underline that for small and medium-sized events insurers

[3] AFA (Association Française de l'Assurance : French Insurance Association)
[4] CCR (Caisse Centrale de Réassurance : Reinsurance Central Fund)

do not have to use complex models instead of raw data since their data record of 15 to 20 years is sufficient to answer questions related to the distribution of claims. Regarding extreme events, insurers can transfer the risks to reinsurers, who determine the exposition and express it through reinsurance premiums. Nevertheless, some issues remain since usually some high-impact events in the portfolio are not entirely reinsured. Moreover, reinsurance might not prevent insurance companies from analyzing their exposition on extreme levels, for instance with a view towards justifying their insurance premiums. In any case, it is helpful to exploit additional information like meteorogical data for major events.

In this work, we have first discussed the nature of wind speeds and problems related with their measurement. Six geographical risk zones have been delimited in France with the `k-medoids` algorithm taking into account extreme dependence, absolute wind speeds and Euclidian distance between the centers of departments, leading to zones that are relatively homogeneous in terms of wind conditions. Due to the heterogeneity of distances between risks and stations and the limited number of significant events and risks, statistical methods lack robustness when working at the finer departmental scale. This approach has allowed us to build a flexible storm index that reflects well the insurance data. We have compared daily and weekly clustering approaches with the bootstrap method. We finally checked different scenarios and methods to model the storm risk.

The Generalized Pareto Distribution (GPD) for threshold exceedances is the main ingredient of our model. Estimated parameters inform us about the nature of the distribution tail, and crucial information is gained from the corresponding return periods associated to major events. They can be used to calculate insurers' average annual costs. It still remains difficult to determine return periods for very extreme events. To meet Solvency II requirements, our experiments show that uncertainty can be almost prohibitively high to extract reliable results. Modeling hypotheses and parameter choices have a great influence on the index value and on the extrapolation of distributions beyond the observed range of values. We analyzed this sensitivity through our results, with a focus on return periods for the most extreme storm events.

Incorporating external data such as wind speeds to establish risk zones remains a good solution for the management of storm risk. However, the volatility of insurance results still depends on the choice of frequency and intensity of the most extreme events, for which also the history of meteorogical records is too short to determine them with satisfactory certainty. We still need to seek for new ways and methods to significantly improve data records concerning extreme events. Even the best models cannot remedy incomplete and relatively short data records. Uncertainty remains high, which is also due to often unverifiable model hypotheses which might create a false illusion of accuracy.

7 Acknowledgements

For this study, the authors have received a research grant which allowed us to work on the records of 192 meteorological stations of Météo France over a period covering 50 years. This work was supported by Allianz France. The authors thank the research chair Actuariat Durable sponsored by Milliman for additional support, as well as Sidney Resnick and Gennady Samorodnitsky for relevant comments.

References

[1] A Bonazzi, S Cusack, C Mitas, and S Jewson. The spatial structure of european wind storms as characterized by bivariate extreme-value copulas. *Natural Hazards and Earth System Science*, 12(5):1769–1782, 2012.

[2] Stuart Coles, Janet Heffernan, and Jonathan Tawn. Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365, 1999.

[3] Dan Cooley, Philippe Naveau, and Paul Poncet. Variograms for spatial max-stable random fields. In *Dependence in probability and statistics*, volume 187 of *Lecture Notes in Statist.*, pages 373–390. Springer, New York, 2006.

[4] M. G. Donat, T. Pardowitz, G. C. Leckebusch, U. Ulbrich, and O. Burghoff. High resolution refinement of a storm loss model and estimation of return periods of loss-intensive storms over Germany. *Nat Hazards and Earth Syst Sci*, 11(10):2821–2833, 2011.

[5] C. Douvillé and E. Burbaud. Tempêtes, grêle et neige : Résultats 2010. Technical report, Association française de l'assurance, FFSA and GEMA, 2012.

[6] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*. Springer, Berlin, 1997.

[7] Christopher AT Ferro and Johan Segers. Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):545–556, 2003.

[8] Eric Gilleland, Mathieu Ribatet, and AlecG. Stephenson. A software review for extreme value analysis. *Extremes*, 16(1):103–119, 2013.

[9] Xavier Guyon. Statistique spatiale. Technical report, Conférence S.A.D.A. 07 - Cotonou - Benin, 2007.

[10] Peter Hall. Asymptotic properties of the bootstrap for heavy-tailed distributions. *The Annals of Probability*, pages 1342–1360, 1990.

[11] S. Hochrainer-Stigler and G. Pflug. Risk management against extremes in a changing environment. *Environmetrics*, 23(8):663–672, 2012.

[12] Brian J Hoskins and Kevin I Hodges. New perspectives on the northern hemisphere winter storm tracks. *Journal of the Atmospheric Sciences*, 59(6):1041–1061, 2002.

[13] R. W. Katz and B. G. Brown. Extreme events in a changing climate: Variability is more important than averages. *Clim Change*, 21:289–302, 1992.

[14] Leonard Kaufman and Peter J Rousseeuw. *Finding groups in data: an introduction to cluster analysis*, volume 344. John Wiley & Sons, 2009.

[15] R. Killick, P. Fearnhead, and I. A. Eckley. Optimal detection of changepoints with a linear computational cost. *J. Amer. Statist. Assoc.*, 107(500):1590–1598, 2012.

[16] M. Klawa and U. Ulbrich. A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Nat Hazards Earth Syst Sci*, 3:725–732, 2003.

[17] MLR Liberato, JG Pinto, RM Trigo, P Ludwig, P Ordóñez, D Yuen, and IF Trigo. Explosive development of winter storm xynthia over the subtropical north atlantic ocean. *Natural Hazards and Earth System Science*, 13(9):2239–2251, 2013.

[18] M. Luzi. Travaux sur les historiques tempêtes. Technical report, 2013.

[19] Cameron A MacKenzie. Summarizing risk using risk measures and risk indices. *Risk Analysis*, 34(12):2143–2162, 2014.

[20] Alexandre Mornet, Thomas Opitz, Michel Luzi, and Stéphane Loisel. Index for predicting insurance claims from wind storms with an application in france. *Risk Analysis*, 35(11):2029–2056, 2015.

[21] J. Papillon. Les appareils de mesure de la vitesse du vent. anémomètres de pression et anémomêtres de vitesse. description des anémomètres actuellement utilisés par l'office national météorologique: Anémomètre électromagnétique papillon; anémomètre magnétique Ã main richard. *Annales Francaises de Chronometrie*, 9:289–303, 1939.

[22] J. G. Pinto, E. L. Fröhlich, G. C. Leckebusch, and U. Ulbrich. Changing European storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM. *Nat Hazards Earth Syst Sci*, 7:165–175, 2007.

[23] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

[24] C. Sacré, J. M. Moisselin, M. Sabre, J. P. Flori, and B. Dubuisson. A new statistical approach to extreme wind speeds in France. *J Wind Eng Ind Aerodyn*, 95:1415–1423, 2007.

[25] P. Thomson, B. Mullan, and S. Stuart. Estimating the slope and standard error of a long-term linear trend fitted to adjusted annual temperatures. *Statistics Research Associates Ltd*, 2014.

[26] Jean-Claude Thouret and Robert d'Ercole. Vulnérabilité aux risques naturels en milieu urbain: effets, facteurs et réponses sociales. *Cahiers des sciences humaines*, 32(2):407–422, 1996.

[27] Robert L. Winkler. The importance of communicating uncertainties in forecasts: Overestimating the risks from winter storm juno. *Risk Analysis*, 35(3):349–353, 2015.