Nat. Hazards Earth Syst. Sci. Discuss., 2, 5247–5285, 2014
www.nat-hazards-earth-syst-sci-discuss.net/2/5247/2014/
doi:10.5194/nhessd-2-5247-2014

Natural Hazards
and Earth System
Sciences
Discussions
Open Access

A framework for modeling clustering in natural hazard catastrophe risk management and the implications for re/insurance loss perspectives

S. Khare, A. Bonazzi, C. Mitas, and S. Jewson

Risk Management Solutions Ltd., 30 Monument Street, London, EC3R 8NB, UK

Received: 19 June 2014 – Accepted: 8 July 2014 – Published: 20 August 2014

Correspondence to: S. Khare (shree.khare@rms.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In this paper, we present a novel framework for modelling clustering in natural hazard risk models. The framework we present is founded on physical principles where large-scale oscillations in the physical system is the source of non-Poissonian (clustered) frequency behaviour. We focus on a particular mathematical implementation of the "Super-Cluster" methodology that we introduce. This mathematical framework has a number of advantages including tunability to the problem at hand, as well as the ability to model cross-event correlation. Using European windstorm data as an example, we provide evidence that historical data show strong evidence of clustering. We then develop Poisson and clustered simulation models for the data, demonstrating clearly the superiority of the clustered model which we have implemented using the Poisson-Mixtures approach. We then discuss the implications of including clustering in models of prices on catXL contracts, one of the most commonly used mechanisms for transferring risk between primary insurers and reinsurers. This paper provides a number of new insights into the impact clustering has on modelled catXL contract prices. The simple model presented in this paper provides an insightful starting point for practicioners of natural hazard risk modelling.

1 Introduction

The broad subject of interest in this paper is natural hazard catastrophe risk modelling. Natural hazard catastrophe risk models are widely used by the insurance industry to address questions related to pricing, capital allocation and risk management. Catastrophe risk models are often based on a timeline simulation of the occurrences of a particular natural phenomenon. The timeline simulation of the natural phenomenon is then translated into a timeline simulation of financial losses on a portfolio of insured risks. The timeline simulation of financial loss on a primary insurance company portfolio is also used to model the price of contracts which are used to transfer risk to reinsurers (which provide insurance for primary insurance companies). This transfer of risk enables improved risk return characteristics of the companies involved in the insurance market, and provides for a more stable insurance industry which is able to meet the obligations to policy holders in the event of a natural hazard catastrophe. Catastrophe models cover natural hazard perils such as North Atlantic hurricane, European windstorm, Asian typhoons, Global earthquakes and floods.

In this paper our focus is on the question of how to model so-called clustered natural hazard phenomena within the context of natural hazard catastrophe risk models. A useful starting point for building a catastrophe model is to generate simulations based on the assumption that the frequency distribution of the underlying process is Poisson. While the Poisson assumption is oftentimes sufficient, for some natural phenomenon, models based on the Poisson assumption fail to accurately model the risk. For example, it is now well known that European windstorms exhibit a strong degree of clustering (Mailier et al., 2006). European windstorm models which are based on a Poisson assumption are indeed useful as a starting point for risk assessment. However, the Poisson frequency distribution is too restrictive in that timeline simulations based on the Poisson assumption do not exhibit the degree of variability that is evident in historical data. As discussed in Mailier et al. (2006), January 1990 and more recently December 1999 were two months in particular that had a high number of very intense windstorms causing considerable losses in the insurance industry (greater than EUR 10 billion). These two years in particular are difficult to explain by European windstorm models that are based on the Poisson assumption, as we will show in this paper. Other phenomenon of interest such as earthquakes, hurricanes, typhoons and severe flooding events may also exhibit some degree of clustering.

In this paper, we make a number of novel contributions to the subject of how to model clustering within the context natural hazard catastrophe modelling. The research presented in this paper is a result of our practical experience in building European windstorm catastrophe models. The contributions of this paper are as follows:

1. In Sect. 2, we present a novel mathematical framework for modelling clustered phenomenon called the Super-Cluster methodology. The Super-Clusters method uses historical data as a starting point for identifying groups of historical events that are strongly associated with oscillations in the physical system, which is the source of clustering.
2. We also provide the mathematical details for a particular implementation of the Super-Clusters method called Poisson-Mixtures. This implementation offers risk modelers a practical and theoretically sound framework for implementing clustering in a meaningful manner.
3. In Sect. 3, using historical data from European windstorms, we demonstrate the Poisson-Mixtures implementation of the Super-Clusters method. In doing so, we show how the historical data is strongly clustered. We develop a clustered model of the data, and demonstrate clearly the superiority of the clustered model compared to a model based on the Poisson assumption. The example we provide allows us to develop a high degree of insight in to modelling clustering, but is simple enough so that it can be used as a starting point for practitioners
4. Finally in Sect. 4, we provide new insights into the impact that modelling clustering has on so-called catastrophe excess of loss contracts, which is one of the important mechanisms for transferring risk between primary insurers and reinsurers. These insights are developed using numerical simulations, and to a certain degree analytic theory.

In Sect. 5 we provide a summary, discussion and draw conclusions from this work.

2 A Super-Cluster framework for modelling clustered frequency processes

We begin this section by defining clearly what we mean by timeline simulation. We then discuss our motivation for developing frameworks for modelling so-called clustered processes from the point of view natural hazard risk modelling. We then introduce a framework for modelling clustered processes which we call the Super-Cluster framework. This Section ends by providing the mathematical details which underlie a particular application of the Super-Cluster methodology called the Poisson-Mixtures formulation.

2.1 Motivation and background: poisson vs. clustered timeline simulation

The starting point for risk models we have in mind is a so-called timeline simulation. A timeline simulation represents event occurrences of phenomena like hurricanes, earthquakes and windstorms. In Fig. 1 we provide a pictorial illustration. In year 1, there are 3 events occurring (represented by the dots), in year 2 there are 2 events, and so on. On the vertical axis is “Loss” which is a representation of financial loss against some insured exposure. Time is on the horizontal axis, discretized into years. Each year represents a draw from the distribution of possible event-loss occurrences. Whilst event time stamps can be created to reflect the seasonality of the phenomenon being considered, in this paper we ignore seasonality and consider loss statistics based on annual time scales. This choice mimics most common insurance industry practice, where the vast majority of risk transfer takes place based on annual contracts.

In practice, there is a computer code that generates the timeline simulation, consistent with some underlying frequency distribution like the Poisson or Negative Binomial distribution. The physics of the system we are trying to model dictates the appropriateness of any particular frequency distribution choice. In this paper, we are interested in physical phenomena that exhibit strong “clustered” behaviour. As we will describe below, clustered simulations impose more “structure” on the timeline compared to Poisson processes. For natural hazards of interest, we assume this structure on the timeline is associated with some underlying physical driver/modulator. From the timeline simulation in Fig. 1, we can readily compute the sample mean annual rate by averaging the sum of annual event occurrences, and we can then compute the (annual) variance (or any other moment). These moments should be consistent with the underlying frequency distribution embedded in the computer code (whether clustered or not).

In what follows, our goal is to understand the fundamental differences between Poisson and clustered simulations. We will understand why Poisson processes can be thought of as “unstructured” and/or random, and provide some relevant physical examples of clustered phenomena.

We begin by developing our understanding of Poisson timeline simulation. We assume that the starting point for our Poisson simulation is a so-called event table comprised of M “events” where each event has a mean annual rate λ_i where $i = 1,...,M$, and each event is simulated using a Poisson distribution. By definition, the average annual number of times event i occurs tends towards λ_i as the length of the timeline simulation grows to infinity. We assume all events in our event table behave independently. We now describe a straightforward way of generating timeline simulations from this event table (using a Poisson assumption): in year 1 we take a random draw from a Poisson distribution with total rate λ, which we label as n_1 (λ is the total rate taken by summing all the λ_i). We then select n_1 events from the event table, by randomly drawing events with replacement, where each event has a probability of being chosen proportional to its rate λ_i. Repeat this process for all subsequent years in the simulation. If one does so, one will find that the variance σ^2 of the annual number of events converges to λ as we increase the length of the simulation. In other words, the frequency distribution has an over-dispersion $\sigma^2/\lambda = 1$. As it turns out, having an over-dispersion exactly equal to 1 leads to timeline simulations that can be described as random and unstructured. We seek to understand this, by considering an alternative and equally valid way to generate Poisson timeline simulations.

Consider the following alternative manner to generate Poisson simulations. Suppose we have an event table consisting of only one event with average annual rate λ. We would like to generate a Ψ years simulation of the occurrence of this event in time. One way of doing this is to is to assume the event occurs $\Psi\lambda$ times (hereafter rounded to the nearest integer) over the Ψ years (a good approximation if Ψ is large). Then, *randomly*

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

assign each of the $\Psi\lambda$ event occurrences to a year from $1,\ldots,\Psi$, independently of one another (each year has probability of $1/\Psi$ of being picked, we are therefore using a uniform distribution on the years, with replacement). For $\Psi \rightarrow \infty$, this prescription is equivalent to the simulation strategy described in the previous paragraph. The key point is that each event is *randomly* assigned to each year in the simulation. The occurrence of one event has no bearing whatsoever on another event. There is therefore no "clustering" in that years with large numbers of events only happen by chance. The same argument can be made for an event set consisting of M events. Poisson simulations have no imposed structure on the timeline implying a lack of clustering.

In this paper, we define clustered processes as those which are over-dispersive so that $\sigma^2/\lambda > 1$, and those which exhibit cross event correlation (which we define in due course). As we will see in this paper, large-scale European windstorms are over-dispersive. As well, European windstorm activity appears to be strongly associated with the north Atlantic oscillation. If we look to years 1990 and 1999 in the recent past (Mailier et al., 2006), it appears as though European windstorms tend to happen in clusters of particularly intense events, at least suggesting event occurrences within particular years are correlated. The mathematical challenge is therefore to develop a framework that allows us to generate over-dispersive simulations which *also* exhibit cross-event correlation.

2.2 Super-Clusters methodology

Our proposed framework for modeling clustering stems from physical principles. The idea is that there exist certain physical drivers which in some years favour the occurrence of certain types of events over others. For example, for European windstorms, if over the course of one year the atmosphere tends to be in a strong positive phase of the North Atlantic Oscillation, the large-scale atmospheric conditions might tend to favour intense windstorms largely propagating in the northeast direction. We assume that the events in our event table can be split into K unique groups of events. Some of these event groups are assumed to be strongly modulated by the physical environment in which the events are embedded. This modulation, as it turns out, is the source of over-dispersion. In this formulation, each of the events in the M event event-table belong to one unique group among the K. We call these groups "Super-Clusters". To model the occurrence of clustering in the Super-Clusters, we will need to use frequency distributions which exhibit over-dispersion and are therefore not Poisson. We also assume that the K Super-Clusters behave independently from one another (although this is not strictly necessary). We focus on the independent case for simplicity. Within each Super-Cluster, the event occurrences are prescribed to be correlated, which is one of the drivers of the over-dispersion that we seek to model.

To split our event set into the K Super-Clusters, we assume the existence of an archive of historical events. The historical events are *not* part of the event table we use for timeline simulation, but are simply used to help use define the composition of the Super-Clusters. A typical archive of historical events for applications we have in mind consists of order 50 years of events. Given an archive of historical events, the idea is to proceed as follows: (1) apply a statistical clustering algorithm to an archive of historical events, this defines the properties of the K Super-Clusters. (2) Use regression analysis to determine the physical drivers of the different Super-Clusters defined on the historical set, to check for physical significance. (3) Check that the K Super-Clusters are behaving independently (this can be relaxed, but in any case it is important to be able to quantify the correlation across Super-Clusters). (4) "Match" the M events in the event table to the K Super-Clusters defined on the historical data using a nearest neighbour algorithm. (5) Determine the target over-dispersion of each of the K Super-Clusters from the historical data (with errors estimates). (6) Apply a mathematical modelling framework (such as the Poisson-Mixtures formulation described in Sect. 2.3) to generate simulations that are in the suggested range of over-dispersions from the historical data, and also exhibit an appropriate degree of within Super-Cluster correlation.

2.3 Poisson-Mixtures formulation

We now describe what we call a Poisson-Mixtures formulation of the Super-Cluster methodology. The Poisson-Mixtures formulation allows us to model over-dispersive Super-Clusters which have cross-event correlation, as required by the Super-Clusters methodology. As before, we assume our stochastic event set is comprised of a total of $i = 1,...,M$ unique events with average annual rates λ_i. Suppose we have divided this stochastic event set into $k = 1,...,K$ Super-Clusters. The kth Super-Cluster has M_k events where each event belongs to one of the Super-Clusters and therefore $\sum_{k=1}^{K} M_k = M$. Given our assumption that the Super-Clusters are mutually independent, it suffices to understand the mathematical behaviour of one of the Super-Clusters before putting all the independent Super-Clusters together in one timeline simulation.

We now focus our attention on the kth Super-Cluster comprised of $j = 1,...,M_k$ events. The important mathematical results underlying the Poisson-mixtures formulation are drawn from Wang (1998). In what follows, we cast the Poisson-mixtures formulation into our context, pointing out important aspects along the way.

Let the discrete random variables $N_1,...,N_{M_k}$ represent the annual number of occurrences of events $j = 1,...,M_k$. We want to generate a timeline simulation. For a Poisson assumption, we would proceed as described in Sect. 2.1. In the Poisson-mixtures formulation, however, we "modulate" the event rates by random draws from a gamma distribution, before simulating any particular year of events. This modulation leads to a simulation with over-dispersion greater than 1, and correlation in the annual occurrences of events within the Super-Cluster. Let Θ_k be a random variable drawn from the univariate gamma distribution $g(\theta_k|\alpha_k,\tau_k)$ where the shape parameter is α_k and the scale parameter is τ_k. In this notation the underscore k represents the kth Super-Cluster. Suppose we are generating a simulation for year 1. The total number of events n_1 to select from the Super-Cluster is,

$$n_1|\Theta_k = \theta_{k,1} \sim \text{Poisson}(\theta_{k,1}\lambda_k) \quad (1)$$

where λ_k is the total rate for Super-Cluster k, and $\theta_{k,1}$ is the particular realization from the gamma distribution for year 1. To determine which events to pick for year 1, we draw, with replacement, events from the Super-Cluster k where the probability of any event being selected is proportional to its rate. This procedure is then repeated for all subsequent years in the simulation. The gamma distribution acts as a "modulator". This modulator of the rates is intended to account for large-scale physical drivers that increase/decrease the annual rate of the particular types of storms in Super-Cluster k. The idea is that historical data will be our guide in determining the amplitude (variance) of this modulator. In this construction, our intention is to not change the mean number of occurrences of all the $j = 1, \ldots, M_k$ events. This can be done by ensuring that the $E[\Theta_k] = 1$. Now, for the gamma distribution, we know that $E[\Theta_k] = \alpha_k \tau_k$. We choose $\alpha_k = 1/\tau_k$ and hence $E[\Theta_k] = 1$. Since $\mathrm{var}[\Theta_k] = \alpha_k \tau_k^2$, this choice implies that $\mathrm{var}[\Theta_k] = \tau_k$. So, the modulator has mean 1 and variance equal to the scale parameter. Notice that in each year of simulation, each member of the Super-Cluster is modulated by the *same* realization of Θ_k.

In Appendix A we provide the expression for the probability generating function of the frequency distribution consistent with the Poisson-mixtures framework as applied in our context, using the results in Wang (1998). The ability to formulate the probability generating function allows us the capability of computing many important loss statistics (explored Sects. 3 and 4) analytically. This is convenient from an implementation point of view.

As discussed in Wang (1998), the Poisson-mixtures probability generating function implies a multivariate Negative Binomial distribution with mean annual rate of $\lambda_k = \sum_{j=1}^{M_k} \lambda_j$ and variance $\lambda_k + \tau_k \lambda_k^2$, which implies that the over-dispersion is $1 + \tau\lambda_k$ for the Super-Cluster as a whole. The over-dispersion is linear in both the variance of the modulation and the overall rate for the Super-Cluster. Finally, as discussed in Wang (1998), the marginal distributions are Negative Binomial with mean λ_j and variance $\lambda_j + \tau_k \lambda_j^2$. Crucially, it can also be shown that the covariance coefficient for the annual occurrences of any two events in the Super-Cluster is equal to $\frac{\mathrm{cov}[N_\mathrm{a},N_\mathrm{b}]}{E[N_\mathrm{a}]E[N_\mathrm{a}]} = \mathrm{var}[\Theta_k] = \tau_k$ (where N_a and N_b are the random variables associated with any two events a and b in the set of $j = 1,...,M_k$ events). The covariance coefficient is identical for all pairs of events in our Super-Cluster, and is equal to the variance of the modulating gamma distribution.

The above formulation is applied to each of the K Super-Clusters where each Super-Cluster has a unique magnitude of modulation given by the scale parameter of the gamma distribution τ. In other words, the model is fully specified by assigning unique τ values for all K Super-Clusters. We now describe what we believe to be the 6 key features/considerations of this formulation:

1. The K Super-Clusters are independent of one another. From an algorithmic/computational point of view this is clearly convenient because we can generate simulations from the different Super-Clusters in parallel. More importantly, this puts some restrictions on how we group the full event set into K Super-Clusters. To apply this formulation, for any particular hazard of interest, one needs to be able to demonstrate that the K Super-Clusters are indeed behaving independently to a reasonable degree. Although outside the scope of this paper, we can in principle use copula methods to rank correlate the K Super-Clusters through their modulating gamma distributions.

2. Each of the $k = 1,...,K$ Super-Clusters has an over-dispersion of $1 + \tau_k\lambda_k$ where λ_k and τ_k are the overall annual rate and gamma distribution variance for Super-Cluster k. Crucially, the over-dispersion can be "calibrated" by choosing a τ_k designed to mimic results from our historical catalogue, if so desired.

3. Within a Super-Cluster, the covariance coefficient is equal to τ_k. As shown above, the modulation of the rates within any particular Super-Cluster is driven by one single parameter. This yields the covariance coefficient which is equal to the variance of the modulating distribution. In a simulation context, this modulation can

be designed to significantly increase the probability of getting many events from any given Super-Cluster in the same simulation year, and conversely can significantly increase the probability of getting few events from any given Super-Cluster in the same year. Our approach is *fundamentally* different than simply putting an over-dispersive Negative Binomial distribution on each event and treating the events *independently*. The Poisson-mixtures formulation is designed to create the “big years” where many events occur in the same year, having dramatic impacts on catastrophe related exposures. Attempts to create “big years” in simulations where the events within a Super-Cluster are treated as independent are very difficult to make work in practice (in our experience). The ability of the Poisson-mixtures approach to create years with many significant events is one of the key reasons we have chosen this formulation.

4. We are able to write down the analytical expression for the joint probability generating function (Appendix A), convenient for analytical computations of key loss statistics.

5. The overall approach is “top down”. By this we mean that the approach to modelling clustering is not derived from first principles physical arguments. Our framework is a *parameterization* of clustering. Our approach has the clear advantage of tunability but comes at the cost, perhaps, of less physical realism. For many phenomena of interest, such as European windstorms, it is unclear whether a bottom up physically based numerical modelling approach would be useful, due to numerical model biases. The more complex the phenomena of interest, the more difficult it becomes to take a fully bottom up approach, and as a result having a well understood top down approach is useful.

6. The Poisson-Mixtures formulation provides even more flexibility/tunability in the following sense: Each of the K Super-Clusters can be divided into a Poisson and over-dispersive part if one chooses to do so. Events which tend to be over-dispersive can be included in the over-dispersive subset of each Super-Cluster.

In European windstorm, our experience suggests that the over-dispersive events tend to be the most severe. Physical thresholds can be used to achieve an appropriate split between the Poisson and over-dispersive subset of each of the k Super-Clusters. By doing so, the over-dispersion of the kth Super-Cluster becomes $1+\tau_k \frac{\lambda_k'^2}{\lambda_k}$ where λ_k' is the sum of the rates of the events in the over-dispersive subset (and clearly $\lambda_k' < \lambda_k$). This subsetting lowers the overall over-dispersion of the kth Super-Cluster. By adding a threshold to split Super-Clusters in to Poisson and Over-Dispersive parts, the simulations may be more physically realistic, and the additional parameter affords us extra flexibility in the calibration process.

3 Application of the Super-Clusters methodology to European windstorm data

In this Section we start by examining whether or not clustering exists in a data set based on 135 of the most intense windstorms to occur in Europe over the period from 1972–2010. As we will show, the data are strongly clustered. We then develop a model of these data by applying the Poisson-Mixtures formulation of the Super-Cluster methodology. Our example not only provides a good demonstration of the importance of clustering, but in using a simple example, we are able to clearly elucidate the steps in applying the Poisson-Mixtures formulation for practitioners interested in more complex applications. We begin by describing the data set.

We use a data set of European windstorm reconstructions discussed extensively in Bonnazi et al. (2012). The reconstructions consist of 135 of the most intense windstorms to occur in Europe over the period from 1972–2010. Using 3 s peak gust data collected from 1972–2010 covering 15 European countries, a set of 135 event reconstructions on a high resolution variable resolution grid was created. The event reconstructions consist of maps of local maximum gusts experienced during the passage of the 135 synoptic events. The high resolution reconstructions were then aggregated to the CRESTA level for Germany, UK, France and Denmark (our chosen domain of interest for demonstration purposes). For each storm, a single number called the storm severity index was produced by computing the following for each storm,

$$\mathrm{SSI} = \frac{\sum (v_i - v_0)^3 A_i}{\sum A_i}^{1/3} \tag{2}$$

where the sum takes place over all affected CRESTA zones for a particular storm, and v_i is the wind gust at CRESTA cell i and A_i is the area of the ith CRESTA cell. We use a threshold of v_0 set to 25 m s^{-1}. The summation takes place over all CRESTA cells in the superset of Germany, UK, France and Denmark. The storm severity index correlates well to aggregated damage/loss due to the passage of the storms, and is therefore particularly appropriate for the risk modelling applications we have in mind. The set of 135 storms we consider contains famous European windstorm events such as 87J, Daria, Vivian, Anatol, Lothar, Martin and Kyrill. For details related to the observational data set, the storm reconstruction method, and how the 135 most severe storms were selected, the reader is referred to Bonazzi et al. (2012).

We now examine the properties of the data relevant to clustering. In Fig. 2, the triangles denote the annual storm counts (derived from the set of 135 storms) but normalized by removal of the mean and division by the sample standard deviation. The red line represents the decadal rolling mean of the storm counts. From these data, it is clear that the 1990s represented a period of relatively high activity. The black dots denote the normalized values for the North Atlantic Oscillation index, and the black line represents the decadal rolling mean for the NAO index. Note that the NAO index is computed for one particular year by taking the mean of the monthly NAO indeces for months January–March and September–December, so our index represents the winter/baroclinic portion of the year. Figure 2 suggests an association between historical event count and the NAO index.

In Fig. 3, we plot the normalized NAO index vs. the annual normalized storm frequency consisting of 39 data points. Higher NAO indeces are associated with higher

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

frequencies. This association is interesting at the very least, and is consistent with positive NAO states being characterized by a strong jet stream and in turn influencing the propagation and development of storms impacting our domain of interest (Germany, UK, France and Denmark). In Fig. 3, we plot the normalized NAO index vs. the total annual storm severity index (summing over all events in a given year). In Fig. 4 we see an association between stronger NAO states, and annual aggregated storm severity. Based on the evidence in Figs. 2–4, it appears to be the case that the state of the NAO influences and modulates to some degree the frequency and severity of windstorms impacting our region of interest. We then make the conjecture that our dataset should exhibit clustering. From the time series of storm counts derived from our 135 storms, we obtain a sample over-dispersion of 1.39, bolstering the notion that our data is clustered.

We now examine the data using a statistic that is very commonly applied in catastrophe risk management. The statistic is the exceedance probability of the annual maximum event (insured) loss on a portfolio of risk. We call this the Occurrence Exceedance Probability (OEP). For the purposes of this paper, storm severity index (SSI) is analogous to loss. We are therefore interested in looking at the exceedance probability of the annual maximum SSI. Given our time series of maximum storm severities from 1972–2010 (years with no events would be assigned zero for the storm severity), we can plot empirical exceedance probability curve which is depicted by the green line on the upper left panel of Fig. 5. Note that on the vertical axis of Fig. 5, we have return period RP, and this is defined as 1 divided by the exceedance probability. The green line on the upper panel of Fig. 5, joins up 39 data points. To determine the exceedance probability associated with the largest SSI value, we take the cumulative probability as $\frac{m-1/3}{n+1/3}$ where $m = 39$ (the order statistic, the largest SSI will have $m = 39$) and $n = 39$ equal to the number of years. This is the plotting position approximately equal to the median of the beta distribution which describes the probability distribution of the cumulative probabilities associated with the largest order statistic (David and Nagaraja, 2003). Hereafter, all other exceedance probability curves in this paper are plotted in the same manner. The exceedance probability at the maximum SSI value over the 39 years is therefore $1 - \frac{m-1/3}{n+1/3}$. The same scheme is used to compute the exceedance probability of the other 38 maximum annual SSI values. For the exceedance probability for the largest SSI value, the grey bounds depict the 5/95 percentiles derived from the beta distribution, representing the statistical sampling uncertainty associated with this exceedance probability (clearly large for the $m = 39$ order statistic).

We now examine whether or not the empirical OEP curve exhibits clustering using a simple test. The test involves building a model of the data utilizing a Poisson assumption. We then generate simulations from this Poisson model and compare the results to the data. To build the Poisson model of the data we start by taking the 135 values of historical SSI and fit a generalized pareto distribution, using the smallest SSI value as the threshold. This choice of threshold is justified in that the 135 storms is a representation of high intensity events. This generalized pareto distribution of the data represents the distribution of SSI given that there is an event. We call this the conditional SSI distribution. We then assume that the annual rate of storm occurrence is $\lambda = 135/39$ (simply the number of historical events divided by the number of years in the historical data). We then generate a timeline simulation using a Poisson assumption (as described in Sect. 2.1). We generated a 10^5 year timeline simulation of SSI under a Poisson assumption. The resulting OEP curve is depicted by blue line in the upper left panel of Fig. 5. Particularly for short return periods around 10 years, the Poisson model OEP curve is much higher than the empirical OEP. Therefore it appears as though a Poisson assumption is inadequate. The next step was to build a clustered version of the model to see whether we get a model that generates results that are more consistent with the historical data. We applied the Super-Cluster method using a Poisson mixtures formulation. In line with the notion that more intense storms are clustered, we treat all simulated events (from the generalized pareto distribution) below SSI 2.5 as Poisson, and all events greater than or equal to 2.5 in SSI as clustered with a gamma variance of 1.5. The first measure of how well this models the data is the over-dispersion we which found to be approximately 1.39 in agreement with the historical data. From this clustered version of the model, we also generated 10^5 years of timeline simulation. The resultant OEP curve from these data are depicted by the red line in the upper left panel of Fig. 5. Particularly for the return periods between 2 and 10, the red line OEP curve generated from the clustered model is more in line with the empirical OEP. The clustered version of the OEP curve is *below* the OEP curve derived from the Poisson version of the model. As well, for very large SSI thresholds, the clustered and Poisson OEP curves appear to converge towards one another. These two properties related to the clustered and Poisson OEP curves are shown informally in Appendix B. By construction, our formulation of clustering has no impact on the mean annual rate, and therefore has no impact on the exceedance frequency beyond any SSI threshold as shown in Appendix C.

Oftentimes in risk modelling applications, we are also interested in statistics associated with the distribution of the 2nd, 3rd and 4th annual maximums. In the upper right panel of Fig. 5, the green line depicts the empirical exceedance probability curve for the 2nd annual maximum (which we call the OEP2). This curve is obtained by taking the 2nd annual maximum SSI value from each year of our time series of 39 years (years with no events are assigned zero SSI). The upper right panel of Fig. 5 also depicts the OEP2 curves from the Poisson (blue) and clustered models (red). For larger SSI thresholds (greater than 3), the clustered OEP2 is larger than the Poisson OEP2. In the lower left/right panels of Fig. 5, we depict analogous results for the OEP3 (3rd annual maximum) and OEP4 (4th annual maximum). As for the OEP2, the OEP3 and OEP4 are higher for the clustered version of the model for large SSI thresholds. The implication is that the clustered model generates higher probability of getting years with larger numbers of intense SSI events, compared to the Poisson model. We would argue that across all 4 panels of Fig. 5, the clustered version of the model is much more consistent with the empirically derived exceedance probability curves. For example, the lower left and right panels demonstrate that the Poisson model is not even consistent with the considerable uncertainty band around the empirical curves.

The clustered version of the model generates timeline simulations that yield much higher probabilities of getting multiple large events. For example, in the Poisson model, the return period associated with getting 3 or more events larger than SSI equal to 4 is approximately 1000 years. The clustered version of the model has a return period of approximately 200 years. Looking at our historical data, we have for 1999 3 major historical storms Anatol, Martin and Lothar with SSI values 2.70, 6.15 and 3.77, respectively. In 1990 storms Daria, Vivian, Herta and Wiebke have SSI values of 5.53, 5.91, 4.17 and 4.14. If we take the 3 most intense storms from 1990 for example, the return period of 1990 suggested by the lower left panel of Fig. 5 is well over 1000 years for the Poisson model. This seems unreasonable given that this year exists in the historical data. The return period suggested by the clustered version of the model is order 200 years. While we have not calibrated the model specifically for the return periods of 1990 and 1999 two things are clear: (a) the Poisson frequency assumption fails to provide reasonable estimates of the return periods of key historical years and (b) the Poisson-Mixtures framework gives us the flexibility to design clustered versions of the model that provide a more accurate and reasonable description of the data. Finally, Fig. 5 demonstrates the high degree of uncertainty associated with exceedance probability statistics derived empirically from historical data. Faced with this degree of uncertainty, it is clear that there is no universal principle that we can suggest for choosing the degree of clustering to put into the model. The practitioner needs to determine the degree of clustering based on looking at a variety of statistics that are pertinent to their risk management application of interest.

4 Implications of clustering for catastrophe excess of loss contract pricing

In this Section, we develop new insights into the impact of clustering on so-called catastrophe excess of loss contracts (hereafter catXL contracts), a common type of contract in the insurance industry that is used to transfer risk from a primary insurance company to a re-insurance company. Our focus is on understanding the change in catXL contract prices when switching from a Poisson to clustered timeline simulation.

We begin by defining the type of catXL contract that we study in this paper. To limit the scope of our work, we focus on aggregate limit contracts due to their common application. We begin by re-visiting the timeline simulation concept in Fig. 1. We suppose that the timeline simulation represents simulations of loss due to European windstorms to a primary insurance company portfolio (consisting of an amalgamation of many home-owner policies). We suppose that for the level of income derived from premium payments from the home-owner policies, the level of risk is too large. To reduce the risk of having to pay out large event losses, primary insurance companies purchase catXL re-insurance from re-insurance companies. The purchase price can be modelled using output from the timeline simulation.

To develop a model for the price, we first define the attachment and exhaustion point A and E respectively. In Fig. 1, A is the loss level depicted by the red dotted line, and E is depicted by the blue dotted line. In each year of simulation, we first compute the loss of each event to the "layer" which for any event with loss $l \geq A$ is equal to $\min(E-A, l-A)$ and is otherwise 0. Suppose in year j of the timeline simulation we have N_j events. The annual loss is equal to $\sum_{i=1}^{N_j} \min(E-A, l_{i,j}-A)$ where $l_{i,j}$ is the ith event loss in year j (events with $l_{i,j} < A$ are excluded from the sum). We now define what is called the aggregate limit, $(E-A)(r+1)$, where r is the number of re-instatements. The aggregate limit is a cap on the annual aggregate loss. For zero re-instatements, the aggregate limit is simply $E-A$, and for infinite re-instatements it is unbounded. Larger numbers of re-instatements means that more losses can be potentially passed onto the re-insurer. However, this comes at the cost of a higher contract purchase price for the primary insurer as we will see.

We define the random variable AL$|r$ which is equal to the annual aggregate loss capped at the aggregate limit (indicated by the condtional notation on the number of re-instatements r). For simplicity, we define the catXL contract price using the first two moments of the distribution of AL$|r$. We take the price as the expectation

$E(\mathrm{AL}|r)$ plus some function of the standard deviation of the annual aggregate loss or $f(\sqrt{E(\mathrm{AL}|r - E(\mathrm{AL}|r))^2})$. To understand the impact of clustering on catXL contract pricing, we therefore seek to understand the impact on the expectation $E(\mathrm{AL}|r)$ and standard deviation $\sqrt{E(\mathrm{AL}|r - E(\mathrm{AL}|r))^2}$.

In what follows we discuss two extreme cases which are relatively easy to characterize. We begin by looking at the case of infinite re-instatements ($r = \infty$), and then move onto the more complicated case of zero re-instatements ($r = 0$). Understanding these two extreme cases is helpful in understanding the more difficult case of finite re-instatements, this point is discussed briefly in Sect. 5.

4.1 Infinite re-instatements

In the case of infinite re-instatements, the aggregate limit is unbounded, and the random variable we consider is the uncapped annual aggregate loss $\sum_{i=1}^{N_j} \min(E - A, l_{i,j} - A)$ (again events for which $l_{i,j} < A$ are excluded from the sum). To begin to understand the infinite re-instatement case, we first look at results from numerical simulation. We generated two different timeline simulations of 10^5 years for the Poisson and clustered models (where SSI is the analogous quantity to loss) exactly as described in Sect. 3. To determine the attachment point A and exhaustion point E thresholds, we first plotted the OEP curve from the Poisson simulation. Results shown in Figs. 6 and 7 are for two different layer defintions. In Fig. 6 the attachment point is the SSI threshold consistent with the 2 year return period on the Poisson OEP, and the exhaustion point is taken as the SSI threshold consistent with the 20 year return period. For the results shown in Fig. 6, the attachment and exhaustion points are drawn from the 20 and 50 year return periods respectively.

We now describe the results in Fig. 6 in detail. In the upper left panel of Fig. 6, we plot the OEP curve generated from our Poisson and clustered timeline simulations (again, as discussed in Sect. 3). On the upper left panel of Fig. 6, the vertical lines denote the SSI thresholds for the attachment point A which is approximately 2.8 and the exhaustion point E which is approximately 5.2. In the upper right panel of Fig. 6, we plot the mean loss (again taking SSI as our proxy for loss), for both the Poisson and clustered case, as a function of the number of re-instatements. When the number of re-instatements is 10^6 (effectively infinite in this context), the mean loss for the Poisson and clustered cases are equal to each other within numerical precision. In the lower left panel of Fig. 6, we plot the standard deviation of the loss to the layer as a function of the number of re-instatements. For 10^6 re-instatements, the standard deviation of the annual aggregate loss to layer is much higher than the Poisson. The results in Fig. 7, for 10^6 re-instatements are qualitatively identical.

So why does clustering impact the standard deviation of the annual aggregate loss, but not the mean loss? At least intuitively, these results seem to make sense. If our natural hazard peril is more volatile due to clustering, perhaps it is not surprising that clustering increases the standard deviation of the annual aggregate loss, especially in the case of infinite re-instatements where there is no cap on the annual aggregate loss, and losses due to all events are counted. Furthermore, clustering is not changing the number of occurrences of any particular event over a long timeline simulation, so in the case where there is no cap on the annual aggregate loss, the impact on the mean loss makes sense.

We now seek to understand exactly why clustering has no effect on the mean loss, while at the same time driving up the standard deviation of the uncapped annual aggregate loss to layer. We again imagine the scenario where we start with an event table and construct a timeline simulation for a Poisson and clustered model (using Poisson-Mixtures). We first note that the inclusion of clustering has no impact on the mean annual rate of occurrence of any particular event. This implies that the distribution of loss conditional on an event occurring is unchanged by the inclusion of clustering. As discussed in Appendix C, for any given loss threshold, the frequency of exceedance is by construction not impacted by the inclusion of clustering (at least for the Poisson-Mixtures formulation of the clustering model). As shown in Appendix D, the mean loss to the layer is given by the product of the event exceedance frequency evaluated at the attachment point, times the mean loss to layer for events that have losses beyond the attachment point. These quantities are unaffected by the inclusion of clustering, therefore we can draw the conclusion that the mean loss to layer for infinite re-instatements is the same for both the Poisson and clustered versions of the models.

We now address the question of why, in the infinite re-instatements case, the standard deviation of the annual aggregate loss is increased when we include clustering in our simulations. In any given year of simulation, the event losses can be sorted in descending order. We can determine which event is the 1st maximum, 2nd maximum and so on. In the infinite re-instatements case, as there is no cap on the annual aggregate loss, the loss in any particular year of simulation is simply the sum of the 1st max, 2nd max, and so on. To proceed, we treat the 1st maximum, 2nd maximum and so on as random variables. For a Poisson simulation, the 1st and 2nd maximum are correlated. For example, for our 100 000 year simulation, the linear correlation between the 1st and 2nd maximum is approximately 0.64. Although in Sect. 2 we highlighted the random nature of Poisson simulations, when we look at the 1st and 2nd maximum, we impose a sorting of the events, a source of correlation. As well, for years with large numbers of events, you essentially have more draws from the severity (SSI) distribution, increasing the likelihood of getting a 1st and 2nd maximum above their respective means. Therefore, we can think of the 1st, 2nd and so on maxima as correlated random variables for both the Poisson and clustered simulations.

Viewing the 1st, 2nd and so on maxima as correlated random variables is helpful in explaining the increased variability of the uncapped annual aggregate loss to layer in the clustered case. Consider the expression for the variance of the sum of correlated random variables. In this context, the sum is over all the event losses to the layer in any particular year. As there is no cap on the annual aggregate loss to the layer, we need to add up losses from *all* events.

Looking to the expression for the variance of the sum of correlated random variables, the variance in the clustered case will be higher if the linear correlation between the random variables is higher, in addition to the variances of the 1st, 2nd and so on maximum themselves. Our numerical simulations reveal that the correlations in the clustered case between maxima are indeed higher in the clustered case. For example, in our clustered simulations, the linear correlation between the 1st and 2nd maximum losses to the layer are 0.74 (compared to 0.64 for the Poisson case). We find qualitatively similar results for all other pairs of maxima in our simulations (we checked up to the 4th maxima). While the 1st, 2nd, 3rd and so on maxima variances are different in the Poisson and Clustered cases, the changes are small relative to the increased correlation due to the inclusion of clustering. Therefore, the source of increased standard deviation for the infinite re-instatements is the increased correlation imposed by our clustering model on the 1st, 2nd and so on maximum losses to layer.

4.2 Zero re-instatements

In the case of zero re-instatements, the cap on the annual aggregate loss to layer is $E - A$. To compute the annual aggregate loss in any given year, we first add up all the losses to layer for each event occurrence. We take the sum of event losses to layer as our annual aggregate loss, but cap it at $E - A$. We now discuss the results in Figs. 6 and 7 based on our numerical experiments. The upper right panels of Figs. 6 and 7 reveal that the mean loss is lower in the clustered case for zero re-instatements. The imposition of the cap on the annual aggregate loss has changed the situation considerably from the infinite re-instatement case where the mean losses are theoretically equivalent.

Why then is the mean loss lower for the clustered model? Suppose, for the sake of argument, that the 1st maximum in each year of simulation explains nearly all the annual aggregate loss. In this case, we can gain insight by just looking at the distribution of the maximum annual loss. In this case, the mean loss would be well approximated by the integral of the OEP curve from the attachment point A to the exhaustion point E (Klugman et al., 2012). As shown in Appendix B, clustering leads to an OEP curve that is less than or equal to the OEP curve generated from the Poisson model. When we add clustering to our simulations, the effect is to create more simulation years with large numbers of events. This has the effect of lowering the probabilities of the 1st maximum

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

exceeding a given threshold because clustering piles in events into the same years (compared to the Poisson simulation). In the case of zero re-instatements, we place a cap on the annual aggregate loss, in turn giving more importance to the 1st maximum, in turn lowering the mean loss in the clustered case. In the lower right panels of Figs. 6 and 7, we plot the percentage contribution of the annual maximum loss to the zero re-instatement mean and standard deviation. In both cases, the large majority of the mean loss is due to the contribution from the first maximum.

With regards to the standard deviation, suppose again that the 1st maximum in each year of simulation explains nearly all the annual aggregate loss (capped at $E-A$). In this case we can approximate the variance using integration of the OEP curve (Klugman et al., 2012). A lower OEP curve then implies lower annual aggregate loss variability. If we look to the numerical results in Figs. 6 and 7, we find the the following: in Fig. 6 we find a *higher* annual aggregate loss to layer standard deviation for the clustered simulations with zero re-instatements. In Fig. 7 we find a *lower* annual aggregate loss to layer standard deviation for the clustered simulations with zero re-instatements. The results in Fig. 7 are more in line with the reasoning based on the 1st maximum loss being the dominant contributor. Comparing the lower right panels of Figs. 6 and 7, we find that the maximum loss explains the vast majority of the zero re-instatement loss in the case of the catXL defined on the higher return period layer in Fig. 7.

Based on these results, it is difficult to draw a very general conclusion. However, our results do demonstrate that in the case of zero re-instatements, it is likely that the annual maximum loss explains the majority of the loss to layer. As a result, one can use integral of the OEP to infer changes to catXL contract prices when moving from a Poisson to clustered model.

5 Summary and conclusions

This paper addresses the problem of how to model clustering in the context of natural hazard risk modelling. Natural hazard risk models are used in the re/insurance industry for various functions such as pricing and efficient allocation of capital. Many natural phenomenon of interest in this context, such as large-scale European windstorms, exhibit a high degree of clustering. In the case of European windstorms, clustered behaviour arises due to the large-scale atmospheric oscillations which can dictate in any given year the frequency of very severe and damaging European windstorms.

In Sect. 2 of this paper we provide a novel framework called the Super-Clusters methodology which allows one to incorporate clustering into a natural hazard risk model in a way that enables one to model over-dispersive processes (a characteristic of clustered processes). A very special property of the Super-Clusters methodology is that it allows one to model the correlation between events. The methodology has its roots in physical arguments. The first step in applying the Super-Clusters methodology is to use a clustering algorithm (which can be defined on very general state spaces) to identify groups of historical events which can be shown to be associated with a unique set of large-scale oscillations which dictate the frequency and severity of the events. For example, in Europe, the state of the NAO is associated with variations in the frequency and severity of damaging windstorms. Once the historical data is grouped into a set of Super-Clusters, the Super-Cluster definitions can be used to group or bin the set of events which comprise the natural hazard risk model.

In Sect. 2 of the paper we also show a particular mathematical formulation of the Super-Clusters methodology. This mathematical formulation is called Poisson-Mixtures approach. The Poisson-Mixtures approach has several advantages: (1) a high degree of tunability so that one can design clustered simulations that exhibit similar behaviour seen in historical data. (2) Lends itself to analytical formulations of the probability generating function which can be advantageous for natural hazard risk model loss calculations. (3) Allows one the ability to model correlation between events within any particular Super-Cluster. Point (3) is crucial, as it has been our experience that formulations which treat events as independent but over-dispersive are oftentimes not sufficient for generating risk models which have a sufficient degree of clustering.

In Sect. 3, we provide a clear demonstration of the application of the Super-Clusters methodology to historical data. We use a set of 135 historical windfield reconstructions. Each historical windfield reconstruction is summarized by a storm severity index which is a quantity that is related to the insured loss. We look into the statistics of the historical data and find the data to be strongly clustered. We then develop a simple simulation model making use of a Poisson frequency assumption. We show that the Poisson frequency based model does not sufficiently represent the degree of clustering observed in the historical data. We then apply the Super-Clusters methodology, implementing and tuning a model using the Poisson-Mixtures formulation. We find that the clustered version of the model provides a much better representation of the historical data. In particular, the clustered model generates statistics which assign much more reasonable return periods to years with multiple intense windstorms. The simple demonstration in Sect. 3 not only demonstrates the importance of clustering in historical European windstorm data, it provides an example upon which risk modelling practitioners can build upon to apply the framework to more complex models.

In Sect. 4, we examine the impact that clustered simulations have on modelled prices of catXL contracts, used by insurance companies to transfer risk to reinsurance companies. We examine aggregate limit catXL contracts which represent one of the most important and common risk transfer contracts in the insurance industry. We focus our attention on two limits: (1) the limit of an infinite number of re-instatements and (2) zero re-instatement limit. For the infinite re-instatement limit, we find that the addition of clustering has no impact on the modelled mean loss, but leads to an appreciable increase in the contract standard deviation. Clustering has no impact on the overall frequency of event occurrences, nor does it impact the conditional event loss distributions, leading to no impact on the mean loss for infinite numbers of re-instatements as shown in the paper. Clustering does, however, lead to an increase in the standard deviation of the loss for the infinite re-instatement case. The explanation for this is that the inclusion of clustering increase drastically the correlation between 1st, 2nd and so on maximum in the timeline simulation, driving a higher standard deviation. For small numbers of re-instatements, and in particular zero re-instatements, we find that in cases where the maximum annual loss represents the vast majority of loss to the catXL contract layers, one can understand the impact that clustering has on the mean and standard deviation by understanding the impact that clustering has on the occurrence exceedance probability curve. The more difficult case to understand is the case where we have an intermediate number of re-instatements. The two extreme cases examined in this paper is a first step towards developing such an understanding.

Finally, we note that we anticipate that in the future, non-Poisson/clustered natural hazard catastrophe risk models will be more commonly used to quantify risk, and some of the understanding we have developed in this paper may be useful in that wider context.

Appendix A: Probability generating function for the Poisson-mixtures formulation

By definition, the probability generating function conditional on one particular $\Theta = \theta$ is given by,

$$P_{N_1,\ldots,N_{M_k}|\Theta}\left(z_1,\ldots,z_{M_k}\right) = E\left[z_1^{N_1},\ldots,z_{M_k}^{N_{M_k}}|\Theta = \theta\right] \tag{A1}$$

for Super-Cluster k. The expectation is over the joint frequency distribution for the random variables $N_1,\ldots,N_{M_k}$. To develop the probability generating function for the Poisson-mixtures framework frequency distribution, we need to integrate over the modulating gamma distribution ($g(\theta|1/\tau_k,\tau_k)$) as follows,

$$P_{N_1,\ldots,N_{M_k}}\left(z_1,\ldots,z_{M_k}\right) = \int_{-\infty}^{\infty} E\left[z_1^{N_1},\ldots,z_k^{N_k}|\Theta = \theta\right] g(\theta|1/\tau_k,\tau_k)\mathrm{d}\theta \tag{A2}$$

where τ_k is the variance of the gamma modulating distribution (Super-Cluster k). The above expression can be shown to be,

$$P_{N_1,\ldots,N_{M_k}}\left(z_1,\ldots,z_{M_k}\right) = \left[1-\tau_k\left(\lambda_1(z_1-1)-\ldots-\lambda_{M_k}\left(z_{M_k}-1\right)\right)\right]^{-1/\tau_k}. \tag{A3}$$

5 As noted in Wang (1998) this defines a multi-variate Negative Binomial distribution with mean annual rate of $\lambda_k = \sum_{j=1}^{M_k}\lambda_{k,j}$ and variance $\lambda_k + \tau_k\lambda_k^2$. Further properties are discussed in Sect. 2.2.

Appendix B: Impact of clustering on the OEP

Here we provide an informal demonstration of why the addition of clustering lowers the occurrence exceedance probability compared to a Poisson model. For the sake of argument, we suppose our event table consists of one event with mean annual rate of λ. Given that this event occurs, let the cumulative distribution of the loss be $F(l)$. Suppose we generate a timeline simulation off our event table consisting of this one event. We define the random variable $M_N = \max(L_1,\ldots,L_N)$, which is the maximum loss in a year with N events. The cumulative distribution of the loss up to a given loss threshold l is,

$$F_M(l) = \Pr(M \leq l) = p_0 + p_1F(l) + p_2F(l)^2 + \ldots \tag{B1}$$

where p_0 is the annual probability of getting zero events, and so on. By definition, the above expression is equal to the probability generating function of $F(l)$. For a Poisson frequency assumption $F_M(l) = e^{-\lambda(1-F(l))}$. For a Poisson-Mixtures implementation where we preserve the mean annual rate of occurrence, the probability generating function shown in Appendix A leads to the following expression for the cumulative probability of the maximum loss $F_M(l) = (1-\tau\lambda(F(l)-1))^{-1/\tau}$.

The OEP for the Poisson case is given by 1 minus the cumulative probability, or just $1-e^{-\lambda(1-F(l))}$. The OEP for the clustered case is $1-(1-\tau\lambda(F(l)-1))^{-1/\tau}$. To simplify, we define $C = 1-F(l)$. The key to understanding the relativity of the Poisson and clustered OEP is to look at the ratio of the Poisson OEP to the clustered OEP given by,

$$\frac{1-e^{-\lambda C}}{1-(1+\tau)^{-1/\tau}}. \tag{B2}$$

As the loss threshold becomes large, $C \to 0$, and the above ratio tends to 1. This implies that the Poisson and clustered OEP converge to one another for large loss thresholds. As well, the above ratio is less that one (not shown here) for all loss thresholds (for positive values of τ), which is easily demonstrated numerically or by simple proof.

Appendix C: Impact of clustering on the EEF

Assume that we have an event set comprised of M events with mean annual rates of λ_i where $i = 1,...,M$. The distribution of loss (severity) associated with event i is $p(l_i)$. Consider a simulation, using only the event i. We want to compute the average annual number of losses that exceed l^* which is called the event exceedance frequency. This is by definition the mean annual rate of event i times the probability that the loss exceeds l^* event that event i has occurred, given by,

$$\mathrm{EEF}_i(l^*) = E(N_i)p(l_i > l^*). \tag{C1}$$

For the entire event set, we take the sum over all events to get,

$$\mathrm{EEF}(l^*) = \sum_{i=1}^{M} E(N_i)p(l_i > l^*) \tag{C2}$$

which can be re-written as,

$$\mathrm{EEF}(l^*) = \sum_{i=1}^{M} \left(\sum_{n_i=0}^{\infty} p_{n_i} n_i \right) p\,(l_i > l^*) \tag{C3}$$

where p_{n_i} is the annual probability of getting n_i occurrences of event i. Note that while clustering will change the p_{n_i} values, the sum will not change, because by construction, we only consider the case where the expectation is not changed. The inclusion of clustering in our case does not change the event exceedance frequency.

Appendix D: catXL mean loss infinite re-instatements

We consider a catXL contract with an attachment point E and exhaustion point A (as defined in Sect. 4). We consider an aggregate limit $\mathrm{AL} = (E - A)(1 + r)$ where r is the number of re-instatements. We first consider the limit as $r \to \infty$. Our event set is comprised of $i = 1, \ldots, M$ events. Assume that each event has an event loss distribution $p(l_i)$. First consider the event i. Let S_{n_i} represent a random variable representing the sum of losses due to n_i occurrences of event i. Let p_{n_i} represent the annual probability of getting n_i occurrences of event i in a year. Let $E(S_{n_i})$ be the expected loss to the catXL layer due to n_i occurrences of event i. The expected loss due to n_i occurrences of event i is simply $E(S_{n_i}) = n_i \int_A^E (1 - p(l_i < l))\mathrm{d}l$, n_i times the integral of 1 minus the cumulative probability for the loss of event i. Then, by definition, the expected loss due to event i, is

$$\sum_{n_i=0}^{\infty} p_{n_i} n_i \int_A^E (1 - p(l_i < l))\mathrm{d}l = \lambda_i \int_A^E (1 - p(l_i < l))\mathrm{d}l \tag{D1}$$

Title Page
Abstract Introduction
Conclusions References
Tables Figures
|◄ ►|
◄ ►
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

where λ_i is the expected number of annual occurrences of event i. The above is completely general in that p_{n_i} need not be Poisson. Now, the mean annual loss is simply,

$$\mathrm{AAL} = \sum_{i=1}^{M} \lambda_i \int_A^E (1 - p(l_i < l))\mathrm{d}l. \tag{D2}$$

Dividing both sides above by $\lambda_{\mathrm{tot}} = \sum_{i=1}^{M} \lambda_i$ gives

$$\frac{\mathrm{AAL}}{\lambda_{\mathrm{tot}}} = \frac{\sum_{i=1}^{M} \lambda_i \int_A^E (1 - p(l_i < l))\mathrm{d}l}{\lambda_{\mathrm{tot}}} = \int_A^E \mathrm{CEP}(l)\mathrm{d}l \tag{D3}$$

where CEP is the conditional event exceedance probability by definition. Hence, we see that the mean annual loss can be written as $\lambda_{\mathrm{tot}} \int_A^E \mathrm{CEP}(l)\mathrm{d}l$. This insight is useful for a number of reasons. Suppose we have two models to compare with equal total rates. We can then attribute mean loss changes to changes in the CEP. As well, we might gain insight into the convergence of catXL contract pricing by looking specifically at the convergence of the CEP.

We can shift our point of view in the above argument and think of $E(S_{n_i})$ as the expectation of the losses to the catXL layer, *given* only the events whose loss distributions have support beyond A. We denote this $\tilde{E}(\tilde{S}_{\tilde{n}_i})$. We can think of p_{n_i} as the annual probability of getting n_i events with losses above the attachment point A, which we denote $\tilde{p}_{\tilde{n}_i}$. Finally, we can think of $\tilde{p}(\tilde{l}_i)$ as the probability distribution of the loss to the layer given that the event loss is above A. Our mean loss due to event i then becomes,

$$\sum_{\tilde{n}_i=0}^{\infty} \tilde{p}_{\tilde{n}_i} \tilde{n}_i \int_0^{\infty} \left(1 - \tilde{p}\left(\tilde{l}_i < l\right)\right) \mathrm{d}l. \tag{D4}$$

In the equation above, we integrate the loss distribution from 0 to ∞ so that we cross the exhaustion point (where there will be a delta function). Taking the sum of the above

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

equation over the M events, we get a mean loss equal to $\tilde{\lambda}\int_0^\infty$ ACEP. In this set up, $\tilde{\lambda}$ is the exceedance rate of event losses above the exhaustion point A and $\int_0^\infty ACEP$ is the mean loss to layer per event given that it causes a loss to the layer (A is used to denote above the attachment).

Finally, as we have discussed above, the exceedance frequency of event losses beyond any loss threshold is unaffected by the inclusion of clustering (as we have constructed it that way). Therefore, catXL mean losses do not change with the inclusion of clustering.

Acknowledgements. The authors would like to thank Greer Kingston and Kechi Nzerem for helpful contributions during the research and development phase of model development that was the precursor to the results presented in this paper.

References

Bonazzi, A., Cusack, S., Mitas, C., and Jewson, S.: The spatial structure of European wind storms as characterized by bivariate extreme-value Copulas, Nat. Hazards Earth Syst. Sci., 12, 1769–1782, doi:10.5194/nhess-12-1769-2012, 2012.

Cusack, S.: A 101 year record of windstorms in the Netherlands, Climatic Change, 116, 693–704, 2012.

David, H. A. and Nagaraja, H. N.: Order statistics, Wiley Series in Probability and Statistics, Wiley, Hoboken, New Jersey, USA, 2003.

Klugman, S. A., Panjer, H. H., and Willmot, G. E.: Loss models – from data to decisions, Wiley Series in Probability and Statistics, Wiley, Hoboken, New Jersey, USA , 2012.

Kossin, J. P., Camargo, S. J., and Sitkowski, M.: Climate modulation of north Atlantic hurricane tracks, J. Climate, 23, 3057–3076, 2010.

Mailier, P. J., Stephenson, D. B., Ferro, C. A. T., and Hodges, K. I.: Serial clustering of extratropical cyclones, Mon. Weather Rev., 134, 2224–2240, 2006.

Wang, S. S.: Aggregation of correlated risk portfolios: models and algorithms, available at: www.casact.com/PUBS/proceed/proceed98/980848.pdf (last access: 18 August 2014), 1998.

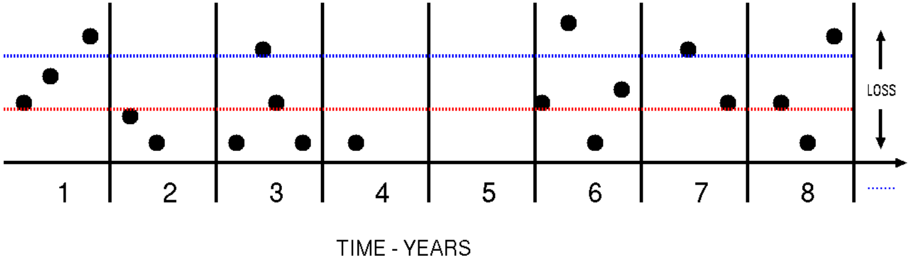

Figure 1. Above is a depiction of a simulation timeline. The dots represent event occurrences. Years are on the *x* axis. Loss is represented on the vertical *y* axis. The red bar represents an attachment point. The blue bar represents the exhaustion point.

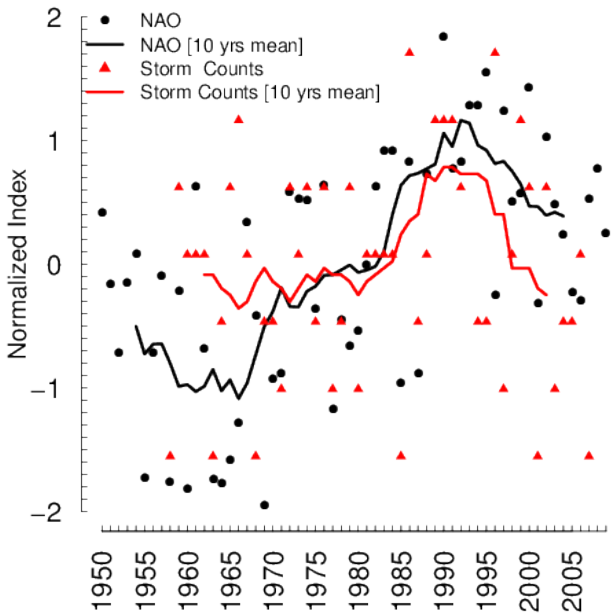

Figure 2. Data for a time series of 135 historical European windstorm events. The redline represents the decadal rolling mean of the winter time NAO indeces represented by the triangle symbols. The black line represents the decadal rolling mean of storm counts represented by the dot symbols. Data has been standardized using the sample standard deviations.

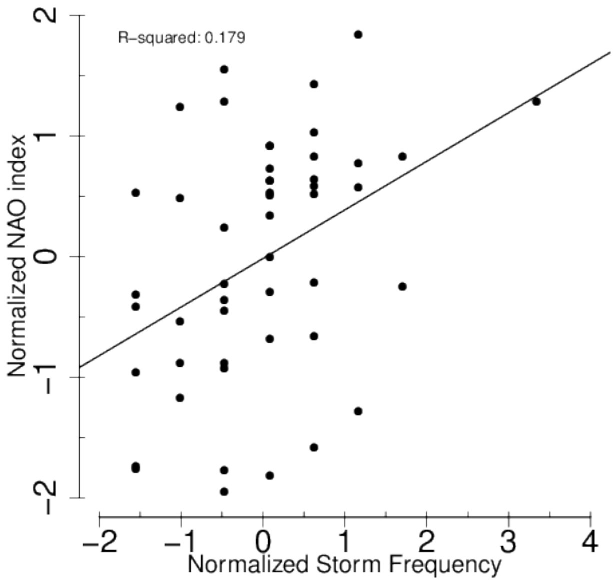

Figure 3. NAO index vs. frequency for the 135 representative European windstorm events.

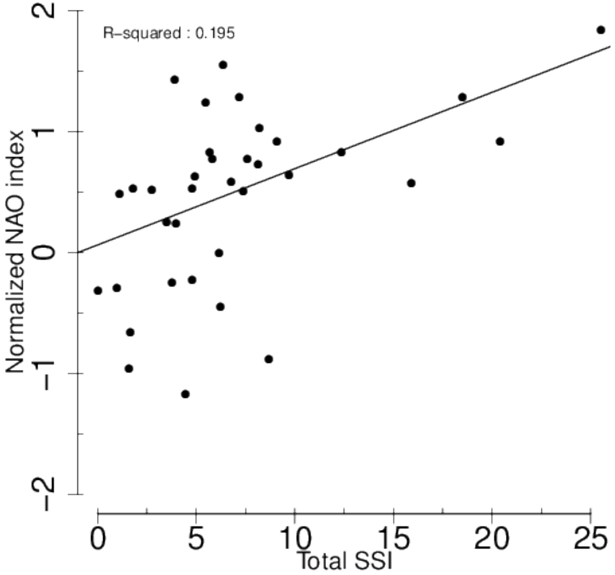

Figure 4. NAO index normalized vs. total storm severity index (SSI) for 39 years of historical European windstorm data. For each year of data, the total storm severity is defined as the sum of the SSI values for all the storms in any particular year.

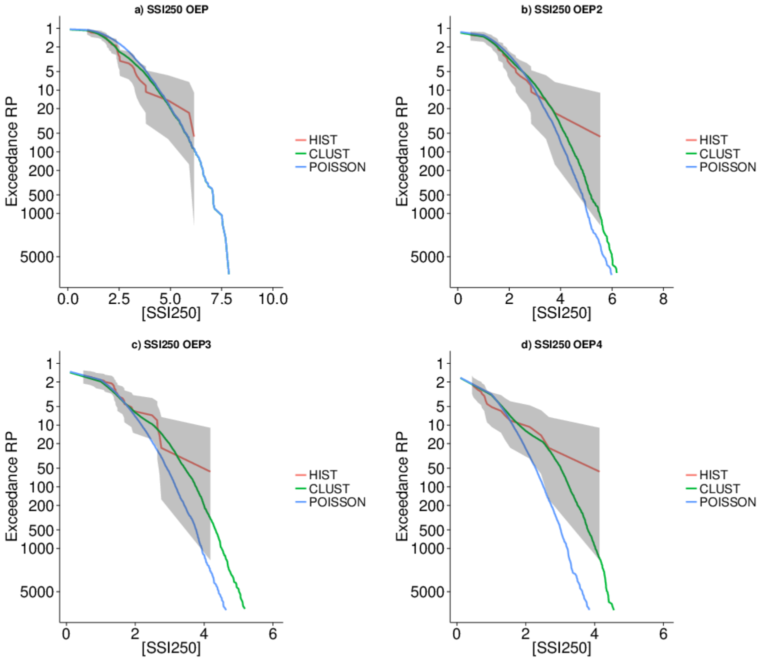

Figure 5. The upper right panel depicts occurrence exceedance probability (OEP) curves. The green line represents the OEP derived from the 135 representative European windstorm events (historical data). The blue dashed line depicts a model of the data using a Poisson frequency assumption. The red line represents a clustered model for the data using a Poisson-Mixtures formulation of the Super-Cluster methodology. The shaded grey bounds represent the 5/95 uncertainty bands on the empirical OEP curve. The upper right panel depicts the 2 event OEP derived from the distribution of the 2nd maximum loss (same plotting convention as for the upper left). The lower left and lower right panels depict analogous results for the 3rd and 4th event occurrence distributions.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀
▶
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

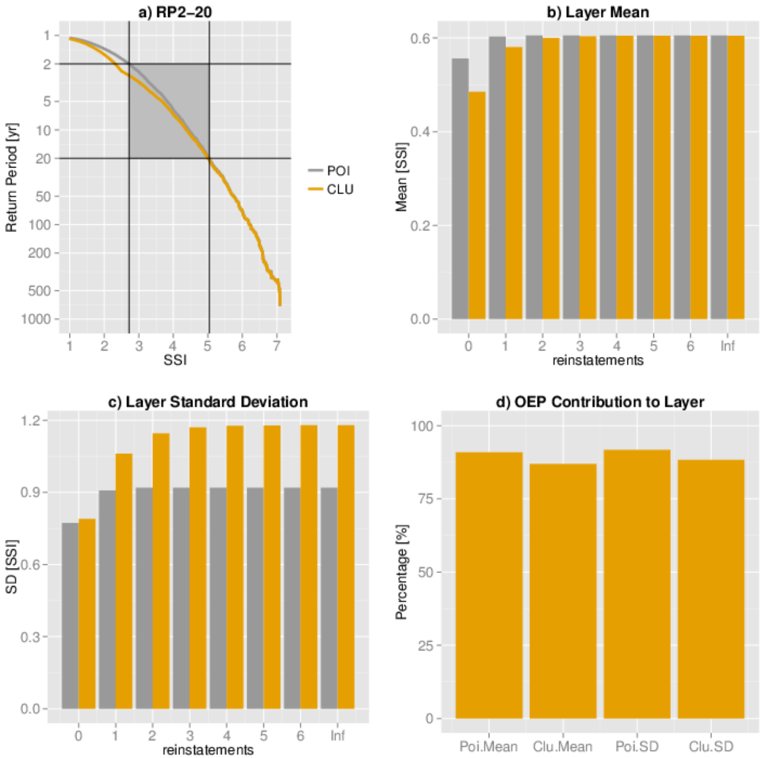

Figure 6. The upper left panel depicts the OEP curves for the Poisson (POI) and clustered (CLU) models of the data. The vertical and horizontal lines on the upper left panel depict the SSI (loss) and return period thresholds of the attachment and exhaustion point of the catXL layer under consideration. The mean loss to layer as a function of the number of re-instatements is shown in the upper right panel. The lower left panel depicts the standard deviation of the annual aggregate loss to layer as a function of the number of re-instatements. The lower right panel depicts the percentage contribution of the 1st annual maximum loss to the catXL mean and standard deviation of the loss for the case of zero re-instatements.

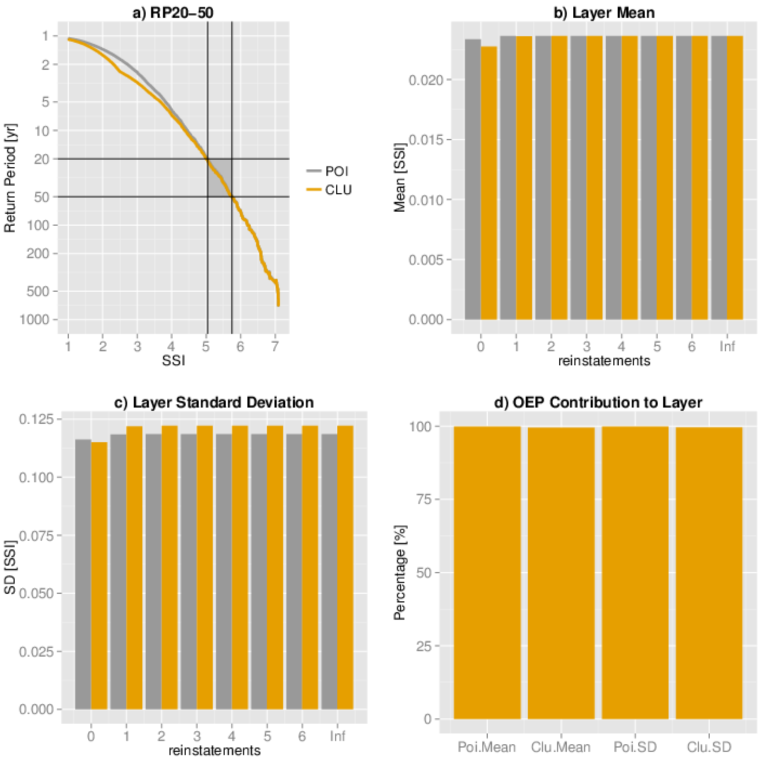

Figure 7. As in Fig. 5 except that the attachment and exhaustion point are defined at 20 and 50 year return periods respectively.