



Modelling clustering of natural hazard phenomena and the effect on re/insurance loss perspectives

S. Khare¹, A. Bonazzi², C. Mitas¹, and S. Jewson¹

¹Risk Management Solutions Ltd., 30 Monument Street, EC3R 8NB, London, UK

²Assicurazioni Generali, Piazza Duca degli Abruzzi, 2 34132 Trieste, Italy

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

Received: 19 June 2014 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 20 August 2014

Revised: 14 April 2015 – Accepted: 23 April 2015 – Published: 26 June 2015

Abstract. In this paper, we present a conceptual framework for modelling clustered natural hazards that makes use of historical event data as a starting point. We review a methodology for modelling clustered natural hazard processes called Poisson mixtures. This methodology is suited to the application we have in mind as it naturally models processes that yield cross-event correlation (unlike homogeneous Poisson models), has a high degree of tunability to the problem at hand and is analytically tractable. Using European windstorm data as an example, we provide evidence that the 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 mixture approach. We then discuss the implications of including clustering in models of prices of catXL contracts, one of the most commonly used mechanisms for transferring risk between primary insurers and reinsurers. This paper provides a number of unique insights into the impact clustering has on modelled catXL contract prices. The simple modelling example in this paper provides a clear and insightful starting point for practitioners tackling more complex natural hazard risk problems.

timeline simulation of the occurrences of a particular natural phenomenon. Timeline simulations are 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 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 hurricanes, European windstorms, Asian typhoons, global earthquakes and floods.

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 phenomena, 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/February 1990 and more recently December 1999 had in particular a high number of very intense windstorms causing considerable losses in the insurance industry (greater than EUR 10 billion). These 2 years in particular are difficult to

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

explain by European windstorm models that are based on the Poisson assumption, as we will show in this paper. Other phenomena of interest such as earthquakes, hurricanes, typhoons and severe flooding events may also exhibit some degree of clustering.

In this paper, we begin by describing a useful framework for modelling clustered processes, applied very specifically to the catastrophe modelling problem in an event-table context. We adopt what we call a Poisson mixture modelling framework, which has been previously described in Wang (1998) and for example has been explored in Eastou and Tawn (2010) in the context extreme value theory for peak flows. Having done so, we build a clustered model for historical European windstorm data, after which we explore the consequences of modelling clustering for common risk transfer contracts (catXL contracts), developing novel insights along the way.

The paper is structured as follows:

- In Sect. 2, we briefly describe a conceptual framework for modelling clustered phenomena, using historical data as a starting point. This framework identifies groups of historical events that are strongly associated with oscillations in the physical system, which are the source of clustering. We then provide a brief summary and discussion of the mathematical properties of the Poisson mixture methodology, our method of choice for modelling clustered processes.
- In Sect. 3, we apply the Poisson mixture methodology to develop a clustered model of historical European windstorm data. We demonstrate the superiority of the clustered model compared to a model based on the Poisson assumption. This section provides a simple and clear example for practitioners to build upon.
- 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 and conclusions.

2 A framework for modelling clustered natural hazard processes

2.1 Motivation and background: Poisson versus clustered timeline simulation

In Fig. 1, we provide a pictorial illustration of timeline simulation. In year 1, there are three events occurring (represented by the dots), in year 2 there are two events, and so on. On the vertical axis is “loss”, which is a representation of financial

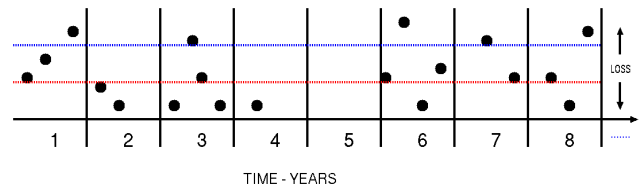


Figure 1. Above is a depiction of a simulation timeline. The dots represent event occurrences. Years are on the x axis. Loss is represented on the vertical y axis. The red bar represents an “attachment point”. The blue bar represents the “exhaustion point”.

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 the most common insurance industry practice, in which 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.

We begin with what we define as an event table. The event table is comprised of M “events” where each event has a mean annual rate λ_i , where $i = 1, \dots, M$. Assuming a Poisson frequency distribution, we can generate a timeline simulation of the number of events per year by simulating from a Poisson distribution independently for each event in our event table. This is by definition a homogeneous Poisson process where the annual rates per event are fixed (see for example Kingman 1993 for extensive discussion of the Poisson process). We emphasize that in such a Poisson simulation, there is zero correlation between the occurrences of different events in our event table. The total annual rate $\lambda = \sum \lambda_i$ is equal to the variance of the annual number of occurrences σ^2 . In the case of the Poisson simulation, the overdispersion $\sigma^2/\lambda = 1$.

In this paper, we define clustered processes as those which are overdispersive 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, historical data suggest that large-scale European windstorms exhibit overdispersion. Physically, European windstorm activity is strongly associated with a number of different teleconnection patterns, including the North Atlantic Oscillation (Mailier et al., 2006), which drive the overdispersion. If we look at years 1990 and 1999 in the recent past (Mailier et al., 2006), it appears as though European windstorms tend to happen in clusters of particu-

larly intense events, at least suggesting that event occurrences within particular years are correlated.

2.2 Conceptual framework for modelling clustering

Here our intention is to simply describe a conceptual framework for developing a model which exhibits clustering, when historical data are available. This framework for modelling 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 the North Atlantic Oscillation). 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 overdispersion. In this formulation, each of the events in the M event table belong to one unique group among the K . We call these groups “clusters”. To model the occurrence of clustering in the clusters we will need to use frequency distributions which exhibit overdispersion and are therefore not Poisson. For simplicity, we assume the behaviour of the clusters are independent of one another (although this can be relaxed).

To split our event set into the K 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 define the composition of the clusters. A typical archive of historical events for applications we have in mind consists of 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, which defines the properties of the K clusters; (2) use regression analysis to determine the physical drivers of the different clusters defined on the historical set to check for physical significance; (3) check that the K clusters are behaving independently (this can be relaxed, but in any case it is important to be able to quantify the correlation across clusters); (4) “match” the M events in the event table to the K clusters defined on the historical data using a nearest neighbour algorithm; (5) determine the target overdispersion of each of the K clusters from the historical data (with errors estimates); (6) apply a mathematical modelling framework (such as the Poisson mixture formulation described in Sect. 2.3) to generate simulations that are in the suggested range of overdispersion from the historical data and also exhibit an appropriate degree of within cluster correlation.

The idea of applying statistical clustering algorithms is now well established in the literature for tropical and extra-tropical cyclone data (see Kossin et al., 2010; Camargo et al., 2007, 2008; Gaffney et al., 2007). While the above described application of statistical clustering algorithms is convenient, we recognize there are several uncertainties associated with this scheme. The results will be sensitive to the

choice of clustering algorithm, the details of the regression analysis, the length and quality of the historical data (which leads to uncertainty in target overdispersion) and the ability to match the event set to the historical clusters. However, if such uncertainties are carefully considered, our view is that the above framework can provide a useful context for building clustered versions of models, which can be calibrated to reflect the properties of the underlying natural phenomena. In what follows, we review our chosen formulation for modelling clustered processes.

2.3 Poisson mixture methodology

As before, we assume our stochastic event set is comprised of a total of $i = 1, \dots, M$ unique events with average annual rates λ_i . Suppose we have divided this stochastic event set into $k = 1, \dots, K$ clusters. The k th cluster has M_k events where each event belongs to one of the clusters and therefore $\sum_{k=1}^K M_k = M$. Given our assumption that the clusters are mutually independent, it suffices to understand the mathematical behaviour of one of the clusters before putting all the independent clusters together in one timeline simulation.

We now focus our attention on the k th cluster comprised of $j = 1, \dots, M_k$ events. The important mathematical results underlying the Poisson mixture formulation are drawn from Wang (1998). Our intention here is to cast the Poisson mixture formulation into our context and point out the key aspects/implications of the methodology for modelling clustering.

Let the discrete random variables N_1, \dots, N_{M_k} represent the annual number of occurrences of events $j = 1, \dots, M_k$. We want to generate a timeline simulation that is overdispersive and exhibits cross-event correlation. In the Poisson mixture formulation, 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 overdispersion greater than 1 and correlation in the annual occurrences of events within the 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 k th cluster. Suppose we are generating a simulation for year 1. The total number of events n_1 to select from the 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 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 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 phys-

ical drivers that increase/decrease the annual rate of the particular types of storms in 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 $j = 1, \dots, M_k$ 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 $\text{var}[\Theta_k] = \alpha_k \tau_k^2$, this choice implies that $\text{var}[\Theta_k] = \tau_k$. So, the modulator has mean 1 and variance equal to the scale parameter. Notice that in each year of simulation, each member of the cluster is modulated by the *same* realization of Θ_k .

In what follows, we summarize a number of key points that we would like to emphasize related to using Poisson mixtures for modelling clustered phenomena:

1. Using the results in Wang (1998), Appendix A provides the expression for the probability generating function consistent with the Poisson mixture framework. Doing so allows us to compute important loss statistics (explored in Sects. 3 and 4) analytically. This is convenient from an implementation point of view.
2. As discussed in Wang (1998), the Poisson mixture 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 overdispersion is $1 + \tau \lambda_k$ for the cluster as a whole. The overdispersion is linear in both the variance of the modulation and the overall rate for the cluster. Finally, as discussed in Wang (1998), the marginal distributions are negative binomial with mean λ_j and variance $\lambda_j + \tau_k \lambda_j^2$.
3. The Poisson mixture framework implies that the annual occurrence of events within a cluster are correlated. As shown in Wang (1998), the covariance coefficient for the annual occurrences of any two events in the cluster is equal to $\frac{\text{cov}[N_a, N_b]}{E[N_a]E[N_b]} = \text{var}[\Theta_k] = \tau_k$ (where N_a and N_b are the random variables associated with any two events a and b in the set of $j = 1, \dots, M_k$ events). This is in contrast to the Poisson assumption, where the annual occurrences of any two distinct events within a cluster are uncorrelated. This Poisson mixture formulation is designed to create “big years” in which many intense events can occur in the same year, having dramatic impact on catastrophe-related exposures. As we will show in the example provided in Sect. 3, models which treat events independently do not yield model results that provide an accurate description of the historical data.
4. In Sect. 2.2, we discussed the idea that the K clusters are independent (which is convenient for implementation purposes), but we could relax this assumption by using Copula methods to rank correlate the K clusters through their modulating gamma distributions.

5. Each of the $k = 1, \dots, K$ clusters has an overdispersion of $1 + \tau_k \lambda_k$. We emphasize that the overdispersion can be ‘calibrated’ by choosing a τ_k designed to mimic results from our historical catalogue.
6. Our experience with European windstorm data suggests that the more severe/intense events tend to be the ones that exhibit clustering (consistent with the findings in Vitolo et al., 2009; Pinto et al., 2013). Hence, we can in principle divide each of our K clusters into a Poisson and overdispersive part. Physically based thresholds can be used to divide up the K clusters as desired. By doing so, the overdispersion of the k th 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 overdispersive subset (and clearly $\lambda'_k < \lambda_k$). This subsetting lowers the overall overdispersion of the k th cluster (for a fixed gamma variance). In our experience, this added flexibility is helpful in model calibration.

3 A simple clustered model of European windstorm data

We now examine 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 to 2010. We use a data set of European windstorm event 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 to 2010. Using 3 s peak gust data collected from 1972 to 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 (Catastrophe Risk Evaluation and Standardizing Target Accumulations) level for Germany, United Kingdom, France and Denmark (our chosen domain of interest for demonstration purposes). CRESTA represents a coarse spatial grid that is commonly used in the insurance industry (see www.cresta.org for example). For each storm, a single number, called the storm severity index, was produced by computing the following for each storm:

$$\text{SSI}_{250} = \frac{\sum (v_i - v_0)^3 A_i}{\sum A_i}^{1/3}, \quad (2)$$

where the sum takes place over all affected CRESTA zones for a particular storm, v_i is the wind gust at CRESTA cell i , and A_i is the area of the i th CRESTA cell. We use a threshold of v_0 set to 25.0 ms^{-1} (hence our notation SSI_{250}). The summation takes place over all CRESTA cells in the superset of Germany, United Kingdom, France and Denmark. The storm severity index correlates well to aggregated damage/loss due to the passage of the storms and is therefore par-

ticularly 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 Kyriell. 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). Using the time series of the number of storms per year over the 39-year data record, we find a sample overdispersion of 1.39, strongly suggesting the time series exhibits clustering. While it is beyond the scope of this paper to provide an in-depth study with regards to the physical drivers that cause this clustering, we note that there is a considerable literature which studies the physical drivers, and a few relevant references will be provided in Sect. 5.

In what follows, we first dig deeper into the historical data, using statistics that are commonly used in catastrophe risk management. After doing so, we build a Poisson model of the data, demonstrating the inadequacy of the model in explaining the data. We then calibrate a clustered model using the Poisson mixture formulation (using specific calibration criteria that we define), showing its superiority over the Poisson model.

In the upper left panel of Fig. 2, the red line depicts the empirical occurrence exceedance probability (OEP). On the horizontal axis we have SSI250 and on the vertical axis we have return period (RP), which is the inverse of the exceedance probability. The OEP is defined as follows: there is a probability distribution associated with the annual maximum SSI250 (loss). For any given SSI250 threshold (on the horizontal axis), the integral of this probability distribution from the threshold to infinity is the OEP (or 1 minus the cumulative probability up to the SSI250 threshold). Similarly, we can define the OEP2 as follows: there is a probability distribution associated with the second annual maximum SSI250. The OEP2 (where 2 denotes the second annual maximum) is, for any given SSI250 threshold, equal to the integral of the probability distribution associated with the second annual maximum from the threshold to infinity. We define the OEP3 and OEP4 analogously. The red curves on the upper right, lower left and lower right panels of Fig. 2 represent the empirical OEP2, OEP3 and OEP4 curves. The empirical curves are derived from our 39-year times series of SSI250 from the historical data. We provide more formal definitions of the OEP and OEP2 Appendix B (the OEP3 and OEP4 follow a similar logic to what is shown in the Appendix B). We refer readers to a book on order statistics by David and Nagaraja (2003) for a comprehensive treatment of order statistics mathematics. Note that the empirical EP curves are computed in the following way: for each case (OEP, OEP2, OEP3, OEP4) we have a time series of $n = 39$ SSI250 values (note, for example, that in the case that there is only one event in a given year, the second, third and fourth maximums are set to 0) and we order the SSI250 values in increasing order indexed by $m = 1, \dots, 39$. For the m th order

statistic, the empirical EP is set equal to $1 - \frac{m-1/3}{n+1/3}$ (using the median of the beta distribution which describes the probability distribution of the cumulative probabilities). The grey bounds on the red empirical EP curves represent the 5/95 quantiles of the beta distribution which represents the uncertainty in assigning the EP value. Note that the uncertainty increases for higher exceedance RP values, as we would expect. The reader is referred to Makkonen (2006) and references therein for an in-depth discussion related to choosing a method for assigning empirical EP values.

We now build a model of the data using a Poisson assumption and see how well the Poisson model matches empirical OEP, OEP2, OEP3 and OEP4. To build the Poisson model of the data we start by taking the 135 values of historical SSI250 and fit a generalized Pareto distribution, using the smallest SSI250 value as the threshold. This choice of threshold is justified in that the 135 storms is a representation of extreme events. This generalized Pareto distribution of the data represents the distribution of SSI250 given that there is an event. We call this the conditional SSI250 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 10^5 -year simulation of event frequency using a Poisson assumption. For each event occurrence, we sample from the generalized Pareto distribution to obtain a simulated SSI250. The resulting OEP curve is depicted by the blue line in the upper left panel of Fig. 2. Notice that at short return periods (2–5 years) the blue line is well above the empirical curve and even strays from the empirical curve uncertainty bounds. Looking at the results for the OEP2, OEP3 and OEP4, we find that the Poisson model underestimates the exceedance probability at the long return periods (or high SSI250 thresholds). In the case of the OEP4, the Poisson model is outside the empirical uncertainty bounds for return periods greater than 50. This implies that the Poisson model does not generate enough of what we call “big years” with multiple large SSI250 events. For example, in 1990, our historical data have four large events (with large insured losses): Daria, Vivian, Herta and Wiebke with SSI250 values of 5.53, 5.91, 4.17 and 4.14 respectively. If we take the 4th largest event in SSI250 (Wiebke), we find the return period to be well over 10000 years. Our opinion is that a model which assigns such a long return period to a year which has occurred in the historical record is of limited utility. Based on these results, we conclude that the Poisson model gives a poor representation of the empirical OEP, OEP2, OEP3 and OEP4. Perhaps it is not surprising given the lack of variability in the Poisson (with its overdispersion of 1.0) compared to the 1.39 in the historical data.

We now use the Poisson mixture method to develop a clustered model for the data. The intended scope of this work is to build a clustered model that satisfies what we feel are reasonable calibration benchmarks, in turn demonstrate its superiority over the Poisson assumption and then understand its

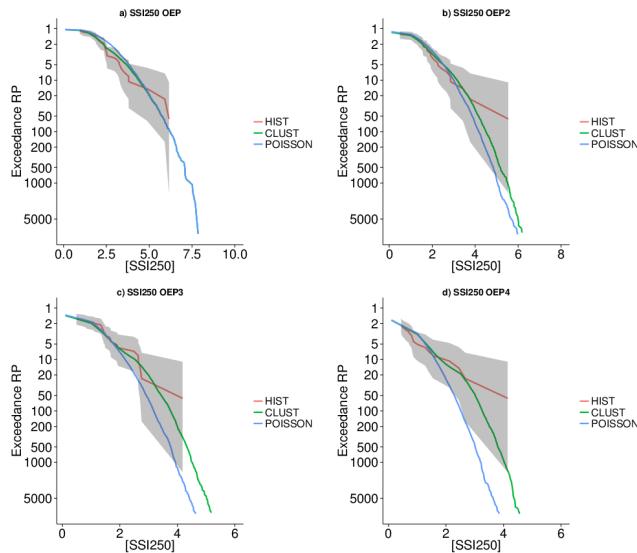


Figure 2. The upper left panel depicts a set of occurrence exceedance probability (OEP) curves. On the vertical axis we have the return period (1 divided by the exceedance probability). On the horizontal axis is the SSI250 intensity measure. The red line represents the OEP derived from the 135 representative European windstorm events (historical data). The blue line depicts a model of the data using a Poisson frequency assumption. The green line represents a clustered model for the data using a Poisson mixture methodology. The shaded grey bounds represent the 5/95 uncertainty bands on the empirical OEP curve. The upper right panel depicts the OEP2 derived from the distribution of the second maximum loss (same plotting convention as for the upper left). The lower left and lower right panels depict analogous results for the third and fourth event occurrence distributions (OEP3 and OEP4 respectively).

behaviour in the context of risk transfer contracts (Sect. 4). We now describe our calibration criteria: (1) it is now well established that a high degree of clustering is more associated with intense systems (Mailier et al., 2006; Pinto et al., 2013; Vitolo et al., 2009). As such, our intention is to build a clustering model where only the more intense storms are clustered to mimic our best understanding of the phenomena (using SSI250 as our measure of intensity). (2) We aim to build a model which yields an overdispersion that is reasonably close to what is obtained from the historical data (1.39). (3) We aim to build a model which lies within the range of uncertainty associated with the empirical OEP, OEP2, OEP3 and OEP4.

To generate the clustered simulations, we start with an event table derived from the 10^5 -year Poisson simulation. We choose an SSI250 threshold and gamma variance. Events below the SSI250 threshold are simulated using a Poisson assumption. Events with SSI250 greater than or equal to the threshold are simulated using the Poisson mixture method for the chosen gamma variance. This procedure yields a new 10^5 -year clustered simulation. To find the model configuration that matches our calibration criteria, we constructed

a large ensemble of models for various combinations of (1) SSI250 threshold and (2) gamma variance used on the clustered storms, until we found a model that met our calibration criteria. The outcome of this calibration procedure was that we chose an SSI250 threshold of 2.5 and gamma variance 1.5. In the historical data roughly 21 % of storms lie above SSI250, which shows that we have clustered the most intense events only. This model configuration happens to achieve an overall overdispersion of 1.38, very much in line with the historical data value of 1.39. Figure 3 depicts results arising from the model calibration process varying the SSSI250 threshold (labelled THRS in Fig. 3) and gamma variances (labelled ALPHA in Fig. 3). In Fig. 3, the grey uncertainty bands are associated with the empirical OEP (only up to SSI250 3.2), as in the upper left panel of Fig. 2. In the upper panel of Fig. 3 we show results for a fixed gamma variance but varying the SSI250 threshold. The lower panel of Fig. 3 depicts results for a fixed SSI250 threshold but varying the gamma variance. Figure 3 gives visual confirmation that our chosen model parameters are reasonable from the perspective of the empirical OEP and its associated uncertainty.

In Fig. 2, the green lines depict the OEP, OEP2, OEP3 and OEP4 results from the clustered model. For the OEP, we see that the clustered model lies within the uncertainty bands associated with the empirical curve. Interestingly, the clustered OEP curve is less than or equal to the Poisson OEP, which is a property of the Poisson mixture methodology as we have applied it (an informal demonstration of this is given in Appendix C) and is more in line with the empirical curve, especially at short return periods. For the OEP2, OEP3 and OEP4, the clustered model is higher than the Poisson curve (at longer return periods) and lies within the range of uncertainty associated with the empirical curves. In our view, the clustered model is a better model, because it is more consistent with the statistics associated with the historical data. The return period of the year 1990 read off the OEP4 curve is between 500 and 1000 years. Unlike the Poisson case (with a return period of well over 5000 years), our view is that the clustered model is much more reasonable, given that we have a year in the historical set where the fourth largest storm has an SSI250 of 4.14. The reason our calibrated model generates shorter return periods for years with a large number of intense SSI250 events (compared to the Poisson model) is (1) the SSI250 threshold, which emphasizes clustering on more intense events (consistent with our scientific understanding of clustering and Vitolo et al., 2009; Pinto et al., 2013); (2) the overdispersion of the clustered model, which gives a reasonable fit to historical data; and (3) the correlation across the intense SSI250 events that are clustered.

In conclusion, using the Poisson mixture methodology, we have built a clustered model that clusters intense storms, gives a reasonable match to historical overdispersion and generates OEP statistics that are within the range of uncertainty implied by the limited historical set (the Poisson model

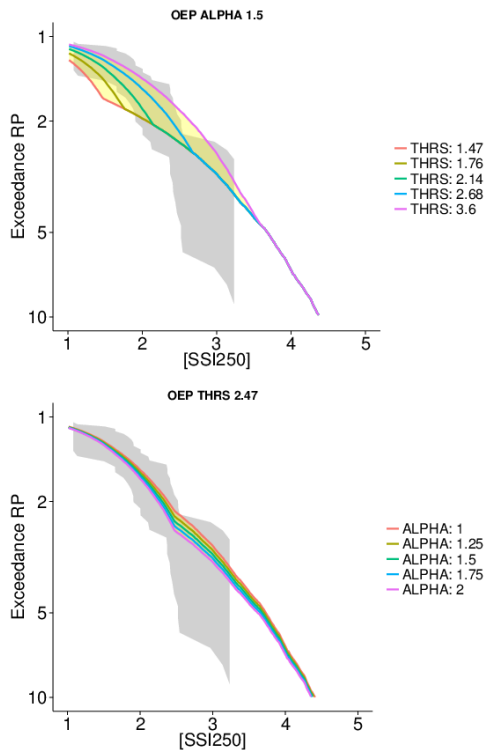


Figure 3. The upper and lower panels depict results arising from the clustered model calibration process. THRES represents the SSI250 threshold, whereas ALPHA indicates the gamma variance. The upper panel depicts OEP curves from various model combinations with fixed gamma variance 1.5 and varying SSI250 threshold. The lower panel depicts OEP curves from various model combinations for fixed SSI250 2.47 threshold but varying gamma variance. The grey uncertainty bands are associated with the empirical OEP curve derived from the historical data.

is unable to do so). The behaviour of the model is, in our opinion, vastly superior to the Poisson model, particularly due to the shorter return periods assigned to significant years with large numbers of intense storms. As such, the model we have built provides a useful demonstration of the principles, which can be built on for more complex problems and calibration methodologies/criteria.

4 Implications of clustering for catastrophe excess of loss contract pricing

Catastrophe excess of loss contracts (hereafter catXL contracts) are a common type of contract in the insurance industry. Such contracts are widely used to transfer risk from primary insurance companies to reinsurance companies. In what follows, we are interested in understanding the differences in key inputs to contract prices that arise when we switch from Poisson-based simulation models to clustered models, using the models developed in the previous section to generate numerical results.

First, we define the type of catXL contract that we study in this paper. We focus on so-called aggregate limit contracts due to their common use in the industry. Consider again the timeline simulation depicted in Fig. 1. Suppose that the timeline simulation represents losses due to European windstorms for a given insurance company portfolio. The so-called attachment point (loss level) A is depicted by the red dotted line in Fig. 1. The so-called exhaustion point (loss level) E is depicted by the blue dotted line in Fig. 1. Let the loss associated with event i in year j be $l_{i,j}$. For each event loss, we can compute the loss to the “layer”. The layer is defined by the loss values between the attachment point A and exhaustion point E . For $l_{i,j}$ the loss to the “layer” is 0 if $l_{i,j} < A$ and otherwise equal to the $\min(E - A, l_{i,j} - A)$. For year j , with N_j events, the aggregate loss to the “layer” is simply the sum of all event losses to layer. We now define the aggregate limit as $(E - A)(r + 1)$, where $r = 0, 1, 2, \dots$ is an integer which represents the number of so-called re-instatements. For any given year j , the annual loss can be at most $(E - A)(r + 1)$. In other words, for any given year j , if the sum of all the event losses to layer exceeds the aggregate limit, the annual aggregate loss is capped by the aggregate limit.

We define a random variable $AL|r$ as the annual aggregate loss for a given number of re-instatements. Again, for any given year, the annual aggregate loss is the sum of the event losses to the layer, capped by the aggregate limit. In this paper, we are interested in studying the expectation and standard deviation of the random variable $AL|r$, which we denote by $E(AL|r)$ and $\sqrt{E(AL|r - E(AL|r))^2}$ respectively. Typically, all the losses to the layer (as defined above) are covered by a reinsurance company which has a contract with an insurance company. The insurance company pays a premium for this contract. The expectation and standard deviation are key components in formulas which are used to determine the required premium (or price) that primary insurance companies pay to reinsurers.

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 ($r = 0$). The larger the number of re-instatements, the larger the potential aggregate annual loss to layer. For larger numbers of re-instatements, generally more risk is passed on from the insurance company to the reinsurer. However, the insurer will oftentimes have to pay a larger premium for contracts with larger numbers of re-instatements. Understanding these two extreme cases is helpful in understanding the more difficult case of finite re-instatements.

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, which for year j is given by

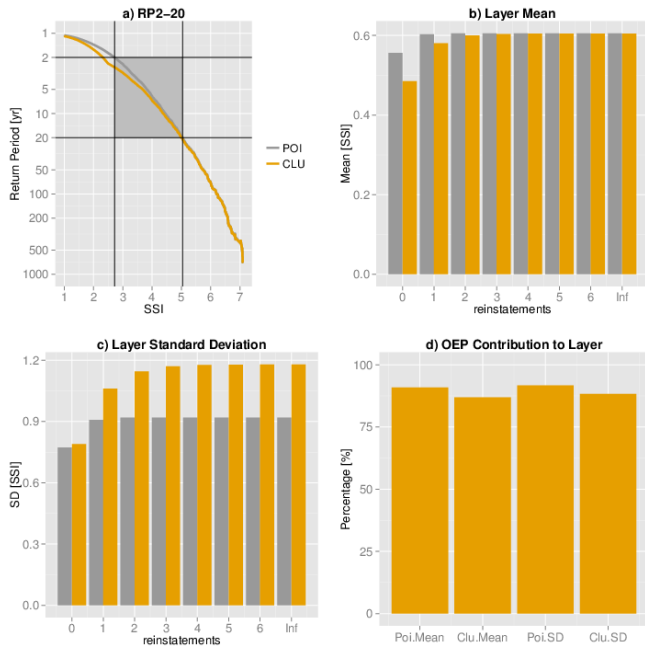


Figure 4. The upper left panel depicts the OEP curves for the Poisson (POI) and clustered (CLU) models of the data. The vertical and horizontal lines on the upper left panel depict the SSI (loss) and return period thresholds of the attachment and exhaustion point of the catXL layer under consideration (note that the SSI values are the same as in Fig. 2 which are labelled as SSI250). The attachment and exhaustion points are consistent with the 2- and 20-year return periods derived off the Poisson OEP curve. 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 first annual maximum loss to the catXL mean and standard deviation of the loss for the case of zero re-instatements.

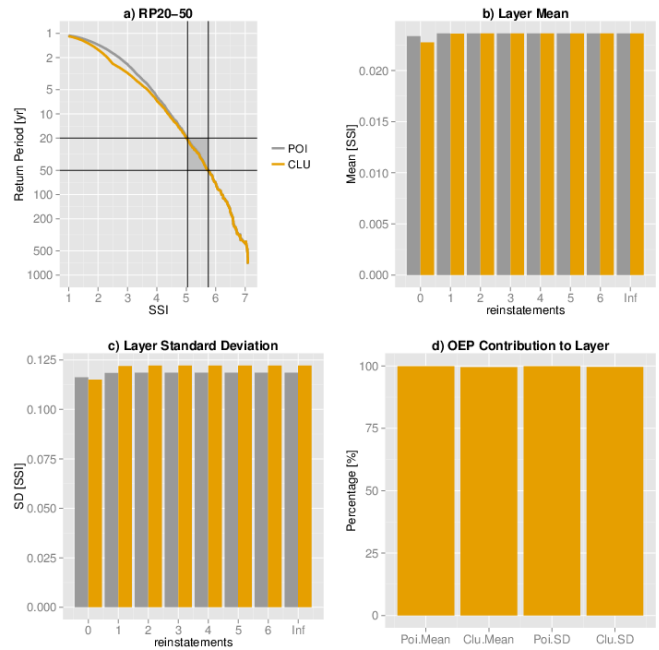


Figure 5. As in Fig. 4 except that the attachment and exhaustion point are defined at 20- and 50-year return periods respectively.

$$\sum_{i=1}^{N_j} \min(E - A, l_{i,j} - A)$$
 (note that events for which $l_{i,j} < A$ are excluded from the sum). We look at results from our 10^5 -year simulations discussed in Sect. 3 for the Poisson and clustered cases. Note again that SSI250 is analogous to loss.

Before looking at the expectation and standard deviations of the annual aggregate loss, we need to define the attachment and exhaustion points A and E . Results will be discussed for two cases which are depicted in Figs. 4 and 5. In the case of Fig. 4, the attachment and exhaustion points A and E respectively are the SSI (loss) thresholds consistent with the 2- and 20-year RP defined off the Poisson OEP. In Fig. 5, A and E are defined by the loss thresholds consistent with the 20- and 50-year RP defined by the Poisson OEP. Our experience is that similar types of definitions of A and E can be found in real contracts applied in the insurance industry. The loss thresholds which define the attachment and exhaustion points on Figs. 4 and 5 are depicted by the horizontal lines.

We now describe the results in Fig. 4 in detail. In the upper left panel of Fig. 4, we plot the OEP curve generated

from our Poisson and clustered timeline simulations (again, as discussed in Sect. 3). In the upper right panel of Fig. 4, 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 and labelled as such), the mean loss for the Poisson and clustered cases are equal to each other within numerical precision. In the lower left panel of Fig. 4, 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. 5 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 in which there is no cap on the annual aggregate loss and losses due to all events are counted. Furthermore, clustering does not change 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.

In more detail, we first note that the inclusion of clustering has no impact on the mean annual rate of occurrence of any particular event (as implemented in this study). This implies that the distribution of loss conditional on an event occurring

is unchanged by the inclusion of clustering. As discussed in Appendix D, for any given loss threshold the frequency of exceedance is, by construction not impacted by the inclusion of clustering (at least for our Poisson mixture formulation of the clustering model). As shown in Appendix E, the mean loss to the layer is given by the product of the event exceedance frequency (EEF) 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, and therefore we understand why 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 first maximum, second 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 losses to layer generated by the first, second and so on maxima. The losses incurred by the first, second and so on maxima can be thought of as correlated random variables. For example, for our 10^5 -year Poisson simulation, the linear correlation between the first and second maximum is approximately 0.64 (obtained using the losses before passing through the layer). The clustered simulations impose a higher degree of correlation between the first and second maximum, which we found to be 0.74. Recalling the expression for the variance of the sum of correlated random variables, it appears that the clustered simulations have higher variance of the annual aggregate loss due to the increased correlation. Ultimately, this higher degree of correlation arises from the property of the Poisson mixture formulation which imposes cross-event correlation of events within a cluster.

As for the price of the contract, recall that both the expectation and standard deviation are key inputs to pricing formulas. While for the infinite re-instatements case the expectation of the aggregate annual loss is unchanged, the increased standard deviation due to the correlation imposed by the Poisson mixture formulation can result in higher model-based contract prices. In this case, clustering introduces additional volatility which drives the price increase.

4.2 Zero re-instatements

In the case of zero re-instatements, the cap on the annual aggregate loss to layer is $E - A$ (since $r = 0$). 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. 4 and 5 based on our numerical experiments. The upper right panels of Figs. 4 and 5 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 annual aggregate loss can be explained nearly entirely by the first maximum 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 C, 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 first maximum 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, which gives more importance to the first maximum and in turn lowers the mean loss in the clustered case. In the lower right panels of Figs. 4 and 5, we plot the percentage contribution of the annual maximum loss to the zero re-instatement mean and standard deviation. In both cases, the first maximum explains the vast majority of the mean and standard deviation.

With regards to the standard deviation, suppose again that the first 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. 4 and 5, we find the following: in Fig. 4 we find a *higher* annual aggregate loss to layer standard deviation for the clustered simulations with zero re-instatements. In Fig. 5 we find a *lower* annual aggregate loss to layer standard deviation for the clustered simulations with zero re-instatements. The results in Fig. 5 are more in line with the reasoning based on the first maximum loss being the dominant contributor. The lower right panels of Figs. 4 and 5 confirm this reasoning: in the case of Fig. 5, the first maximum is the dominant contributor (not the case in Fig. 4). This corresponds to our intuition that the maximum loss would be the dominant contributor for layers with higher return periods for the attachment and exhaustion points (in Fig. 4 the layer is defined for return periods 2–20 and in Fig. 5 for 20–50).

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, the integral of the OEP can be used to understand, to first order, the behaviour of changes in catXL contract prices when shifting from a Poisson to clustered model.

5 Summary and conclusions

This paper addresses the problem of how to model clustered natural hazard phenomena and explores the consequences of clustering for an important class of risk transfer contracts used to manage natural hazard risk. We provide a simple conceptual framework for building a clustered model and review the properties of a Poisson mixture formulation of this framework. Using an archive of historical European windstorm data, we demonstrate the superiority of a Poisson-mixture-based clustered model over a Poisson model. We then explore the consequences of modelling clustering in the context of catastrophe excess of loss contracts, one of the most common and important risk transfer mechanisms.

Section 2 describes our conceptual framework in detail and is based on the idea of identifying cluster groups associated with unique physical drivers, an approach that is well established in the literature (Kossin et al., 2010; Camargo et al., 2007, 2008; Gaffney et al., 2007). Section 2 continues by reviewing a Poisson mixture methodology (e.g. Wang, 1998) for modelling clustered processes. The Poisson mixture methodology introduces overdispersion by modulating the annual rate of storms within a cluster group using random draws from a prescribed gamma distribution. In the context of natural hazard risk modelling, our view is that this methodology has a number of useful properties: (1) this methodology goes beyond the homogeneous Poisson process by enforcing cross-event correlation, (2) has an overdispersion that can be calibrated to reflect the properties of the underlying natural phenomena and (3) is analytically tractable which leads to useful insights and can make numerical implementation easier. Variants of the Poisson mixture approach we describe have been applied in other contexts: see for example Eastou and Tawn (2010). Beyond Poisson-mixture-type approaches, we note that many alternatives for modelling clustering exist: see for example Villarini et al. (2013) for recent work.

In Sect. 3, we begin by analyzing data derived from a set of 135 European windstorm event reconstructions, representing intense events that occurred during the period from 1972 to 2010. Each historical wind-field reconstruction is summarized by a storm severity index which is a quantity that is related to the insured loss. We find that this historical record has an overdispersion of 1.39. We then build a model of the data based on a Poisson frequency assumption. Not only is the Poisson model under-dispersive (by definition having an overdispersion 1.0), we find that the occurrence exceedance statistics (representing the statistics of the first, second, third and fourth annual storm severity maxima) given by the Poisson model do not lie within the range of uncertainty implied by the empirical data. We find that the Poisson model assigns a very long return period (> 5000 years) to a year like 1990 (with four significant events) in the historical data. We then use the Poisson mixture methodology to build a clustered model. The criteria that we chose to use in calibrating

the clustered model were as follows: (1) the model should only cluster the more intense events; this is consistent with the finding in the scientific literature that more intense events exhibit the highest degree of clustering (Mailier et al., 2006; Pinto et al., 2013; Vitolo et al., 2009). (2) The model should generate occurrence exceedance statistics that lie within the range of uncertainty implied by the historical data. (3) The model overdispersion should be reasonably close to the historical data. In the model calibration process, a large ensemble of models was tested, varying both the storm severity index threshold (beyond which clustering was applied) and the gamma variance. Our chosen model generated an overdispersion of 1.38, and the occurrence exceedance statistics are consistent with the range of uncertainty implied by the historical data. The clustered model also assigned a return period to the important year 1990 of between 500 and 1000 years. For these reasons, our view is that the clustered model is vastly superior to the Poisson model.

The historical data set used in Sect. 3 appears to exhibit strong clustering. While it is beyond the scope of this paper to explore the physical drivers in detail, we note that European windstorm clustering has been shown to be associated with a number of large-scale atmospheric patterns such as the North Atlantic Oscillation (e.g. Mailier et al., 2006; Pinto et al., 2009). We believe that these patterns are the underlying driver of the high degree of overdispersion we see in our chosen historical record. The approach we have outlined for modelling clustering is what we would call a “top-down” approach, in that we are using statistical modelling to model the behaviour we see in the data in order to generate good natural hazard risk models. One inherent limitation of this approach is the limited historical record that we have to work with. In the long run, we anticipate that clustering models will be more strongly informed by numerical modelling. For example in Pinto et al. (2013) the use of numerical modelling for understanding clustering is discussed.

Section 4 explores the impact that clustered simulations have on statistics used in models of catXL contract prices. CatXL contracts are used by insurance companies to transfer risk to reinsurance companies. In the limit of infinite re-instatements, we find that clustering has no impact on the mean loss but increases the standard deviation of the annual aggregate loss due to the increased correlation between the first, second and so on maxima in the timeline simulation. 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. These results provide unique insights into the impact modelling clustering is expected to have in catXL contract pricing. The results in this paper can be used as a starting point to understand the case where we have an intermediate number of re-instatements.

We anticipate that in the future, non-Poisson/clustering 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 mixture formulation

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

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

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

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

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

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

As noted in Wang (1998), this defines a multi-variate negative binomial distribution with a mean annual rate of $\lambda_k = \sum_{j=1}^{M_k} \lambda_{k,j}$ and variance $\lambda_k + \tau_k \lambda_k^2$.

Appendix B: Definitions of the OEP and OEP2

We start by defining the occurrence exceedance probability. Let M_1 be a random variable representing the maximum annual loss (or intensity SSI250 as in Sect. 3). Let N be a random variable representing the number of events in a year. Given an event occurrence, assume that $F(l)$ represents the cumulative probability of the loss L at loss threshold l . For a given loss threshold l the cumulative probability that the maximum annual loss is below l is given by

$$F_M(l) = \Pr(M \leq l) = p_0 + p_1 F(l) + p_2 F(l)^2 + \dots, \quad (B1)$$

where p_n is the probability of getting $n = 0, 1, 2, \dots$ events. The first term accounts for the probability of getting zero events in any given year. The second term accounts for the probability of getting one event (and if we do get one event, it needs to be below the threshold and hence the $F(l)$ in the second term on the right hand in the above equation). The third term in the above expression represents the case where we have two events in a given year, and both must be below the threshold. The summation continues up to values of N for which the annual probability is 0 (if that occurs). By definition, the OEP(l) = $1 - F_M(l)$.

To obtain the OEP2, we have the following cumulative probability for the random second annual maximum loss:

$$F_{M2}(l) = \Pr(M2 \leq l) = p_0 + p_1 + p_2(F(l))^2 + 2F(l)(1 - F(l)) + p_3 \left(F(l)^3 + 3F(l)^2(1 - F(l)) \right) + \dots \quad (B2)$$

The logic behind the third term above is as follows: if there are two events, one configuration where the second random maximum is below l is where both events are below l whose probability is given by $F(l)^2$. However, there are two other configurations that need to be accounted for: one event is below the threshold and one event is above. The total probability associated with these configurations is given by $2F(l)(1 - F(l))$ (as this can happen in two ways). By definition, the OEP2 = $1 - F_{M2}(l)$. Similar logic can be applied to the case of the OEP3 and OEP4. The basic idea is that given a fixed number of events, one needs to account for all the possible ways the maximum under consideration can be below a given threshold. Readers are referred to David and Nagaraja (2003) for a comprehensive treatment of these order statistics.

Appendix C: Impact of clustering on the OEP

In Appendix B, the expression for the cumulative probability of the maximum annual loss is shown. This, by definition, is the probability generating function of $F(l)$. For a Poisson frequency assumption $F_M(l) = e^{-\lambda(1-F(l))}$. For a Poisson mixture implementation in which 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 $1 - e^{-\lambda(1-F(l))}$. 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

$$\frac{1 - e^{-\lambda C}}{1 - (1 + \tau\lambda C)^{-1/\tau}}. \quad (C1)$$

As the loss threshold becomes large, $C \rightarrow 0$, 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 than 1 (not shown here) for all loss thresholds (for positive values of τ), which is easily demonstrated numerically or by simple proof.

Appendix D: Impact of clustering on the event exceedance frequency

Assume that we have an event table comprised of M events with mean annual rates of λ_i where $i = 1, \dots, M$. The distribution of loss 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^* (given that event i has occurred):

$$EEF_i(l^*) = E(N_i)p(l_i > l^*). \tag{D1}$$

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

$$EEF(l^*) = \sum_{i=1}^M E(N_i)p(l_i > l^*), \tag{D2}$$

which can be re-written as

$$EEF(l^*) = \sum_{i=1}^M \left(\sum_{n_i=0}^{\infty} p_{n_i} n_i \right) p(l_i > l^*), \tag{D3}$$

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 formulation of Poisson mixtures we have discussed in this paper does not, therefore, change the EEF.

Appendix E: Catastrophe excess of loss contract: mean loss for 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 $AL = (E - A)(1 + r)$ where r is the number of re-instatements. We consider the limit as $r \rightarrow \infty$. Our event table is comprised of $i = 1, \dots, M$ 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 $E(S_{n_i}) = n_i \int_A^E (1 - p(l_i < l)) dl$, which is 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

$$\sum_{n_i=0}^{\infty} p_{n_i} n_i \int_A^E (1 - p(l_i < l)) dl = \lambda_i \int_A^E (1 - p(l_i < l)) dl, \tag{E1}$$

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 (denoted ‘‘AAL’’ for average annual loss) is simply

$$AAL = \sum_{i=1}^M \lambda_i \int_A^E (1 - p(l_i < l)) dl, \tag{E2}$$

adding up the contributions from each event in our event table. Dividing both sides above by $\lambda_{tot} = \sum_{i=1}^M \lambda_i$ gives

$$\frac{AAL}{\lambda_{tot}} = \frac{\sum_{i=1}^M \lambda_i \int_A^E (1 - p(l_i < l)) dl}{\lambda_{tot}} = \int_A^E CEP(l) dl, \tag{E3}$$

where ‘‘CEP’’ is the conditional event exceedance probability by definition. Hence, we see that the mean annual loss can be written as $\lambda_{tot} \int_A^E CEP(l) dl$.

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

$$\sum_{\tilde{n}_i=0}^{\infty} \tilde{p}_{\tilde{n}_i} \tilde{n}_i \int_0^{\infty} (1 - \tilde{p}(\tilde{l}_i < l)) dl. \tag{E4}$$

In the equation above, we integrate the loss distribution from 0 to ∞ so we are guaranteed to cross the exhaustion point (where there will be a delta function). Taking the sum of the above 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. We also thank the anonymous reviewers for comments that led to significant improvements in the paper. Any views or opinions presented in this article are those of the authors and do not necessarily represent those of their employers.

Edited by: B. Merz

Reviewed by: two anonymous referees

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.
- Camargo, S. J., Robertson, A. W., Gaffney, S. J., Smyth, P., and Ghil, M.: Cluster analysis of typhoon tracks. Part I: General properties, *J. Climate*, 20, 3635–3653, 2007.
- Camargo, S. J., Robertson, A. W., Barnston, A. G., and Ghil, M.: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects, *Geochem. Geophys. Geosyst.*, 9, Q06V05, doi:10.1029/2007GC001861, 2008.
- David, H. A. and Nagaraja, H. N.: Order statistics, Wiley Series in Probability and Statistics, 2003.
- Eastou, E. F. and Tawn, J. A.: Statistical models for overdispersion in the frequency peaks over threshold data for a flow series, *Water Resour. Res.*, 46, W02510, doi:10.1029/2009WR007757, 2010.
- Gaffney, S. J., Robertson, A. W., Smyth, P., Camargo, S. J., and Ghil, M.: Probabilistic clustering of extratropical cyclones using regression mixtures models, *Clim. Dynam.*, 29, 423–440, 2007.
- Kingman, J. F. C.: Poisson Processes. Oxford Studies in Probability, 3. The Clarendon Press, Oxford University Press, New York, 1993.
- Klugman, S. A., Panjer, H. H., and Willmot, G. E.: Loss models – from data to decisions, Wiley Series in Probability and Statistics, 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.
- Makkonen, L.: Plotting positions in extreme value analysis, *J. Appl. Meteorol. Climatol.*, 45, 334–340, 2006.
- Pinto, J. G., Bellenbaum, N., Karremann, M. K., and Della-Marta, P. M.: Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions, *J. Geophys. Res.-Atmos.*, 118, 12476–12485, 2009.
- Pinto, J. G., Zacharias, S., Fink, A. H., Leckebusch, G. C., and Ulbrich, U.: Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO, *Clim. Dynam.-Atmos.*, 32, 711–737, 2013.
- Villarini, G., Smith, J. A., Vitolo, R., and Stephenson, D. B.: On the temporal clustering of US flood and its relationship to climate teleconnection patterns, *Int. J. Climatol.*, 33, 629–640, 2013.
- Vitolo, R., Stephenson, D. B., Cook, I. M., and Mitchell-Wallace, K.: Serial clustering of intense European storms, *Meteorol. Z.*, 18, 411–424, 2009.
- Wang, S. S.: Aggregation of correlated risk portfolios: Models and algorithms, available at: <https://www.casact.org/pubs/proceed/proceed98/980848.pdf>, 1998.