CHAPTER

3

Towards a More Dynamical Paradigm for Natural Catastrophe Risk Modeling

David B. Stephenson[1], *Alasdair Hunter*[2], *Ben Youngman*[1], *Ian Cook*[3]

[1]**Department of Mathematics, University of Exeter, Exeter, UK;** [2]**Barcelona Supercomputing Centre, Barcelona, Spain;** [3]**Willis Towers Watson, Willis Research Network, London, UK**

INTRODUCTION

This chapter reviews how Event Loss Tables (ELTs) produced by catastrophe models are used to simulate insurance losses due to natural hazards. A formal statistical interpretation is presented in terms of a static mixture of compound processes and implications for clustering of losses are discussed. The consistency of this approach is then critiqued on the basis of physical understanding of natural hazard processes. A more dynamic nonstationary approach is demonstrated using various idealized dynamical mixing models and results are compared for an artificial ELT typical of windstorm losses. The tail of the aggregate loss distribution is found to be highly sensitive to the modeling choices. The chapter concludes with a summary of main findings and some ideas for further developments.

LOSS SIMULATION

The most direct way to estimate insurance loss distributions would be to use historical claims data. However, such an approach is not reliable due to the rarity and quality of the loss event data, the shortness of the historical records, and the nonstationarity of the losses (e.g., due to technological and economic trends). An alternative approach known as catastrophe modeling (Friedman, 1972) relies upon developing a model to simulate a large sample of hazard events, and then for each of these synthetic hazards, vulnerability and exposure

Risk Modeling for Hazards and Disasters
http://dx.doi.org/10.1016/B978-0-12-804071-3.00003-3

information are used to calculate property damage and subsequent insurance losses. Cat(astrophe) models have become the *de facto* approach to quantify natural hazard risk after the unprecedented losses in the early 1990s; examples include the cluster of windstorms that affected Europe in 1990, Hurricane Andrew in 1992, and the Northridge earthquake of 1994 (Grossi and Kunreuther, 2005). Many cat models use a discrete event simulation approach whereby random numbers for hazard variables are drawn from a simplified parametric stochastic model of the hazard process that has been fitted to past data. Alternatively, a large sample of plausible hazard events can also be generated *ab initio* by direct numerical simulation from a complex physical model of the hazard (e.g., a regional climate model). However, such numerical simulation is often computationally expensive and requires ad hoc bias correction to adjust for model misrepresentation of extreme events. This chapter will focus on the more common stochastic simulation approach.

Event Loss Table Definition

For a given insurance portfolio (i.e., spatial vulnerability and exposure information), most probabilistic cat models produce an ELT, which can be used to create plausible simulations of losses for the forthcoming insurance period (hereafter, assumed to be the forthcoming calendar year). An ELT typically has thousands of rows each of which represents the properties of a certain possible type of hazard, namely, its rate of occurrence ρ and a set of parameters θ that describe the distribution of losses that would be incurred if such an event were to happen. The distribution is bounded from above by the maximum exposure and represents uncertainty in losses for this given hazard event due to imperfect knowledge of vulnerability and exposure in the portfolio (so called *secondary uncertainty*).

Fig. 3.1 shows a scatter plot of rate and the loss expectation (i.e., the mean of the loss distribution) for each of $I = 55{,}000$ rows from an artificial ELT created by simulation from a bivariate lognormal distribution fitted using maximum likelihood to mimic ELTs typically generated by European windstorm catastrophe models. Rates and losses vary over many orders of magnitude with rows having larger mean losses tending to have smaller and less uncertain rates.

An ELT is essentially a very large set of model parameters for a finite mixture model of compound Poisson processes that can then be used to simulate losses. Losses in a year (or other desired time period) are assumed to be a mixture of independent losses drawn from distributions specified by parameters given by each row of the ELT. The $N = \sum N_i$ losses in a year consist of $N_i \sim Poi(\rho_i)$ losses drawn independently from a known cumulative distribution function $F(x;\theta_i)$ for each row $i = 1,2,\ldots,I$ in the table. Because the total rate is usually small and there are many rows in the ELT, most of the N_i counts are 0, a few are 1, and very few are greater than 1.

In the simplest case, each type of hazard in the ELT leads to an exact loss, in other words, there is no additional so-called secondary uncertainty caused by imperfect knowledge of vulnerability and exposure. In such cases, losses from row i of the ELT are always identical fixed values equal to the expected loss $\theta_i = E(X_i)$ and the cumulative distribution function $\Pr(X \le x) = F(x;\theta_i)$ is a step function equal to 0 when $x \le \theta_i$ and 1 when $x > \theta_i$. More generally when there is secondary uncertainty, this can be represented either by specifying parameters of distributions (e.g., the 3-parameter Beta distribution) for each row in the ELT, or by

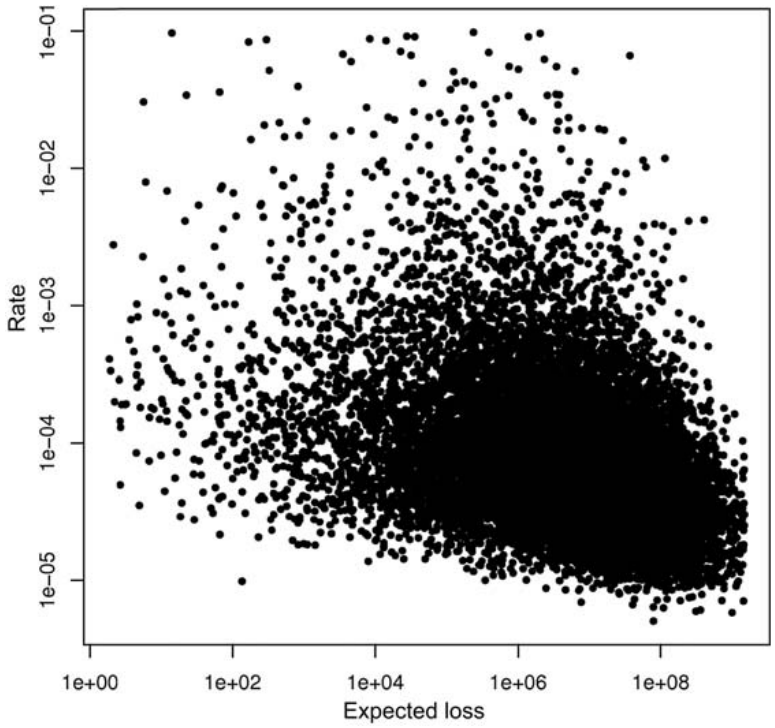

FIGURE 3.1 Scatter plot of rate against expected loss (on logarithmic axis scales) for the 55,000 rows of the illustrative ELT typical of European windstorm losses. Negative association due to smaller rates for higher losses is typical of most hazards.

replicating each row in the ELT and dividing its rate among a discrete set of fixed loss values according to their probabilities (J.C. Rougier, personal communication).

To simulate losses, one could loop over all the rows in the ELT. However, it is much faster to first find the total number of events in a year by drawing from a Poisson distribution $N \sim Poi(\rho)$, where $\rho = \sum \rho_i$, and then draw N losses from distributions having θ sampled N times by replacement from the θ_i each with probability ρ_i/ρ (for a proof see Rolski et al., 2009; Section 4.2.2). For example, in the case of no secondary uncertainty, this can be done using the following R language code:

```
rhosum <- sum(rho)
N <- rpois(rhosum)
losses <- sample(x=theta, size=N, replace=TRUE, prob=rho/rhosum)
```

where `rho` and `theta` are the column vectors of rates and expected losses, respectively, in the ELT.

Loss Distribution Functions: CEP, OEP, and AEP

Decisions about risk capital, portfolio management, and premium pricing are based on summary measures of distributions of various simulated loss variables. Following common practice, consider the time period of interest to be 1 year. Suppose that in year k there are

$N_k > 0$ single event losses $\{X_{jk}; j = 1,\ldots,N_k\}$, which have annual maximum loss $M_k = \max\{X_{jk}; j = 1,\ldots,N_k\}$ and annual aggregate loss $S_k = X_{1k} + X_{2k}\ldots + X_{N_k k}$. The following distributions are then usually estimated from simulated loss data:

- **Conditional Exceedance Probability** $\mathrm{CEP}(x) = \Pr(X_{jk} > x)$: distribution of single event loss;
- **Occurrence Exceedance Probability** $\mathrm{OEP}(x) = \Pr(M_k > x)$: distribution of maximum loss;
- **Aggregate Exceedance Probability** $\mathrm{AEP}(x) = \Pr(S_k > x)$: distribution of aggregate loss.

It should be noted that these distributions depend in different ways on the joint distribution of rate and loss parameters represented by the ELT, for example, $\mathrm{CEP}(x) = \rho^{-1} \sum_{\theta_i > x} \rho_i$ in the special case of no secondary uncertainty. Hence, caution needs to be exercised not to modify the joint distribution when merging and deleting rows in ELTs but to increase computational speed (so-called *boiling down*). Various risk measures are used to summarize exceedance probability distributions (Embrechts et al., 2005). One of the most widely used measures is the return value for a specified return period, for example, the T = 200-year return value, $\mathrm{AEP}^{-1}(1/\mathrm{T})$, is obtained by reading off the aggregate loss corresponding to the 1/T exceedance probability from the AEP curve. This is simply the $1 - 1/\mathrm{T}$'th quantile of the aggregate loss distribution sometimes also referred to as Value-at-Risk $\mathrm{VaR}_{1-1/\mathrm{T}}$ (Embrechts et al., 2005; Section 2.2). It should be noted that unless there is only one event per year, $\mathrm{CEP}^{-1}(1/\mathrm{T})$ is not the T-year return value for single event losses, i.e., the loss that is exceeded on average every T years. However, this metric has some undesirable properties: for example, it is not subadditive, i.e., the quantile of a sum of losses from two portfolios can exceed the sum of the quantiles of each individual portfolio. Subadditivity is one of the most important axioms of coherence for diversification as it guarantees that an aggregate portfolio made up of smaller portfolios is not exposed to greater risk than the sum of the constituent portfolios (Dowd and Blake, 2006). A more coherent measure is provided by *Expected Shortfall* ES_p—the expectation of losses that exceed the p'th quantile (sometimes also referred to as Tail Value at Risk TVaR and Tail Conditional Expectation). Expected shortfall can be obtained either by Monte Carlo simulation or by integration of VaR quantiles over the tail of the loss distribution:

$$ES_p = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q dq = E(X|X > \mathrm{VaR}_p)$$

Expected shortfall ES_p is never less than VaR_p, and when $P = 0$, is simply the expectation $E(X)$, i.e., the mean loss. Both VaR and ES risk measures will be used here to summarize loss.

CLUSTERING

Scientific Background

Unlike losses simulated from an ELT, natural hazard events are not generally independent; background conditions can cause hazards to arrive in close succession, which enhances the risk of large aggregate losses, e.g., the winter 2013/14 cluster of European windstorms

Christian, Xavier, Dirk, and Tini caused insured losses of 1.38, 0.96, 0.47, and 0.36 billion USD that in total exceeded 3 billion USD.

By fitting stochastic models to historical storm track data, new scientific discoveries have been made in recent years about the clustering of windstorms. Notably,

- there is more clustering than expected due to chance, i.e., there is significant overdispersion in storm counts to the south and north of the Atlantic storm track region and over West Europe (Mailier et al., 2006);
- clustering can be largely explained by dynamic modulation of Poisson rates by the North Atlantic Oscillation and other large-scale climate modes (Mailier et al., 2006);
- clustering increases for more intense windstorms especially over North Europe and Scandinavia (Vitolo et al., 2009), which is most likely related to the fact that rate and intensity of windstorms are positively correlated and increase together over large parts of North Europe (Hunter et al., 2016).

Similar findings have been reproduced in storms simulated by climate models (Kvamstø et al., 2008; Pinto et al., 2013; Economou et al., 2015). These results are stimulating new research into mechanisms for clustering of extreme storms (e.g., Rossby wave breaking; Pinto et al., 2014). Furthermore, many of these aspects of clustering also apply to other natural hazards such as hurricanes (Mumby et al., 2011) and floods (Villarini et al., 2013).

Overdispersion in Counts

A simple way to characterize clustering is to assess the variance in counts of events. If the events occur randomly in time with no clustering, then the number of events N in a fixed time period will be Poisson distributed with a variance equal to the constant mean rate. If clustering is present, this will be associated with additional variation that causes the variance to exceed the mean counts—a phenomenon known as *overdispersion.* Clustering is sometimes more narrowly defined to be additional variance in counts to that accounted for by Poisson sampling AND known variations in the rate, i.e., more variance than can be explained by a nonhomogeneous Poisson process. Overdispersion can be assessed by the statistic $\phi = Var(N)/E(N) - 1$, which indicates clustering when positive. This statistic has been extensively used in the studies mentioned in the previous section.

To understand how clustering behaves for more extreme events, it is useful to consider overdispersion in the number of events $N(u)$ that have losses/intensities exceeding threshold u. Vitolo et al. (2009) used such a quantity to show that there was increased clustering for windstorms with more extreme vorticities on the northern flank of the Atlantic storm track that extends into Scandinavia and North Europe.

For losses generated from an ELT, the Laws of Total Expectation and Variance can be used to show that the expectation and variance of $N(u)$ are given by

$$E(N) = \sum_{i=1}^{I} \rho_i (1 - F(u;\theta_i))$$

$$Var(N) = \sum_{i=1}^{I} \left(\rho_i (1 - F(u;\theta_i)) + \rho_i (1 - F(u;\theta_i))^2 \phi_i \right)$$

For generality to allow row counts to be negative binomially as well as Poisson distributed, it is assumed that the row counts N_i have expectation $E(N_i) = \rho_i$ and variance $Var(N_i) = \rho_i(1+\phi_i)$, where ϕ_i is a prescribed overdispersion for counts generated by row i. By taking the ratio of these expressions, the overdispersion in $N(u)$ is given by

$$\phi(u) = \frac{\sum_{i=1}^{I} \rho_i (1 - F(u;\theta_i))^2 \phi_i}{\sum_{i=1}^{I} \rho_i (1 - F(u;\theta_i))}$$

which leads to various interesting deductions. Firstly, $\phi(u)$ is strictly positive only if one or more of the rows has overdispersion (i.e., $\phi_i > 0$) and so if the rows all have Poisson counts with $\phi_i = 0$ then there is no overdispersion in the total counts above any threshold. Secondly, in the presence of secondary uncertainty as $u \to \infty$, $1 - F(u;\theta_i) \to 0$ and so $\phi(u) \to 0$, i.e., overdispersion and clustering vanish in total counts of extreme losses even if row counts are overdispersed (e.g., negative binomial). Finally, in the special case of no secondary uncertainty, $1 - F(u;\theta_i)$ is only either 0 or 1 and so overdispersion simplifies for independent row counts to

$$\phi(u) = \frac{\sum\limits_{\theta_i > u} \rho_i \phi_i}{\sum\limits_{\theta_i > u} \rho_i},$$

i.e., a weighted average of the overdispersion in each of the rows with losses θ_i that exceed the threshold u. When all the rows have the same dispersion $\phi_i = \phi$, this leads to overdispersion in the total counts that is independent of threshold, i.e., $\phi(u) = \phi$ but this is a singular result valid only when there is absolutely no secondary uncertainty.

Building Clustering into ELT Calculations

Scientific evidence for storm clustering has led to several recent attempts to build clustering into cat model loss simulations. The simplest approach is to assume that row counts are independently distributed as negative binomial rather than as Poisson distributions. From the arguments in Section Overdispersion in Counts above, this will result in overdispersion of losses above low thresholds that unfortunately vanishes for more extreme losses. Khare et al. (2015) describe a more sophisticated approach that splits the ELT into $K > 1$ independent clusters/groups of mutually dependent rows that have negative binomial counts with different shape parameters (overdispersion) for each cluster. Such an approach will also result in vanishing overdispersion in total counts above sufficiently high thresholds. Furthermore, there is little scientific evidence for believing that there are a finite number of independent groups of windstorm type—a continuum is perhaps more justifiable for describing vortex features in a turbulent atmosphere. An alternative approach for creating clusters of losses is to create a parent loss using Poisson sampling from the ELT and then generate offspring losses by making repeated draws from the secondary uncertainty distribution of the parent row. Such a Neymann-Scott process has been implemented on the financial platform at Willis Re but it is difficult with this approach to get overdispersion to increase for more extreme losses. Furthermore, it could

be argued that such a process is nonphysical in that it creates clusters of losses for a given hazard rather than creating losses for a cluster of hazards as occurs in the real world.

DYNAMICAL MIXTURE MODELS

Concept

Fig. 3.2 shows a simple graphical model for how ELT simulations can be made more dynamical. Random (mixing) variables Z and Z' can be introduced to represent time-varying drivers, which modulate the rate and loss distribution parameters (e.g., the North Atlantic Oscillation and Scandinavian Pattern indices for European windstorms; see Hunter et al., 2016). The drivers can be either independent or correlated with one another and can be specific to the spatial location of the portfolio.

As a first step, one can ignore modulation of losses and replace the static ρ_i rates in an ELT with dynamical rates $\lambda_i(Z)$ that depend on a time-dependent random mixing variable $Z(t)$, which modulates the hazard rate. Such an exogenous mixing approach is already a common practice in credit risk modeling (Crouhy et al., 2000; Frey and McNeil, 2003). Credit risk portfolios take a similar form to ELTs; a portfolio will consist of a bank's loans each of which is characterized by a probability of defaulting and a corresponding loss. Although loans are

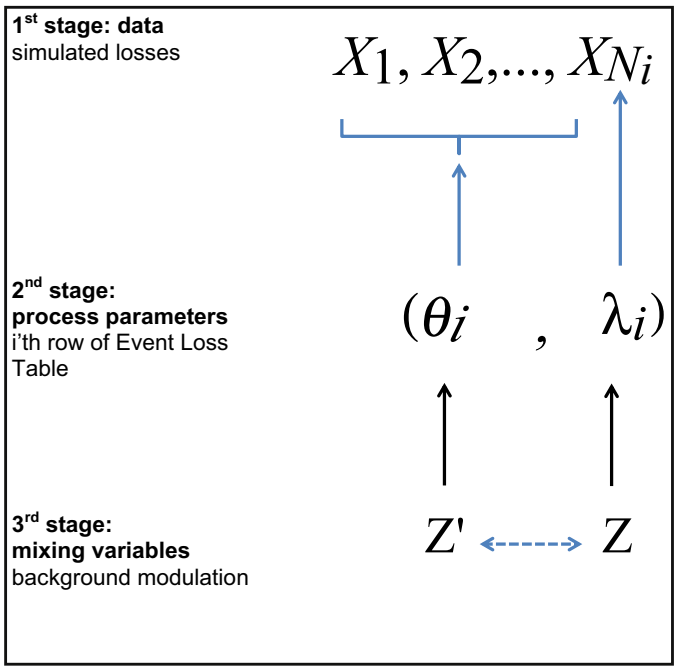

FIGURE 3.2 Schematic graph showing the hierarchical relationship between variables used to simulate losses from an ELT. *Blue arrows* (light gray in print versions) denote dependency, i.e., random variables at the tip are drawn from distributions controlled by variables at the base, for example, $X \sim F(\theta_i)$ and $N_i \sim Poi(\lambda_i)$. *Black arrows* denote deterministic functions and the *dashed arrow* represents association that may exist between variables that drive rates and those that drive hazard intensity (e.g., for windstorms over Scandinavia; see Hunter et al., 2016).

assumed to default independently of each other, they can be affected by the same set of background factors such as the rate of inflation, interest rates etc. To include these drivers in credit risk models, probabilities of loans are regressed on a common set of random variables that represent these factors.

Adopting a Generalized Linear Model approach, the counts from row i of the ELT can be modeled as $N_i \sim Poi(\lambda_i)$, with the rate depending on the mixing variable via $g(\lambda_i(Z)) = \beta_{0,i} + \beta_{1,i}Z$, where $g(.)$ is a monotonic function known as a *link function*. When $g^{-1}(\beta_{0,i} + \beta_{1,i}Z)$ is Gamma distributed, this gives a Gamma-Poisson mixture for counts, which is equivalent to assuming negative binomial counts. The simplest dynamical model is obtained by considering an identity link function and a mixing variable that can take only binary values, 0 and 1, each with probability 0.5, i.e., $Z \sim Ber(0.5)$. Binary mixing is a crude coin-toss approach for modeling years that can have either low or high states (i.e., $Z = 0$ or $Z = 1$) of some climate mode such as the North Atlantic Oscillation. A more continuous approach is to represent the climate index using a standardized Gaussian distributed variable, i.e., $Z \sim N(0,1)$. Following usual practice in Poisson regression, the link function $g(.)$ is then chosen to be the *canonical* logarithmic link, which prevents having negative rate parameters. The rates of each row are then log-normally distributed and analytic expressions can be used to find the first and second moments for each row. Since the lognormal distribution is often quite similar to the Gamma distribution, this gives counts for each row that are close to having a negative binomial distribution. However, it should be noted that the counts in all the rows (and not just groups) of the ELT are now correlated with one another since they each depend on the mutual mixing variable Z.

Estimation of the Mixing Parameters: $\beta_{0,i}$ and $\beta_{1,i}$

Because of the lack of driver information from hazard models, it is not possible to regress ELT rates on known drivers such as leading modes of climate variability. Therefore, alternative approaches are required for estimating the mixing parameters. A simple method of moments approach for finding these parameters is to impose constraints on the expectation $E(N_i) = E(g(\lambda_i))$ and the variance of row counts $Var(N_i) = (1+\phi_i)E(N_i)$, where overdispersion $\phi_i = Var(g(\lambda_i))/E(g(\lambda_i))$. To ensure that mean behavior is similar (e.g., expected loss), it is natural to assume that the expectation of the row counts should equal that of the static approach i.e., $E(N_i) = \rho_i$. Overdispersion can be set to nonnegative values $\phi_i \geq 0$ for each row to reproduce desired clustering behavior in the losses such as a believable overdispersion profile $\phi(u)$. These two constraints for each row give two equations that can be solved to find the beta parameters, e.g., $\beta_{1,i} = \sqrt{\phi_i \rho_i / VarZ}$ and $\beta_{0,i} = \rho_i - \beta_{1,i}E(Z)$ when the link function is the identity function $g(\lambda_i) = \lambda_i$. For the special case where the overdispersion of row i is chosen to be $\phi_i = \phi\rho_i/\rho$ and the rates are modulated by a binary $Z \sim Ber(0.5)$ mixing variable via an identity link function (i.e., binary mixing), beta estimates are given by

$$\beta_{1,i} = 2\rho_i\sqrt{\frac{\phi}{\rho}}$$

$$\beta_{0,i} = \rho_i - \frac{\beta_{1,i}}{2}$$

When $\phi_i = \phi\rho_i/\rho$ and the rates are modulated by a Gaussian $Z \sim N(0,1)$ mixing variable via a logarithmic link function (i.e., lognormal mixing), beta estimates are given by

$$\beta_{1,i} = \sqrt{log(1+\phi/\rho)}$$

$$\beta_{0,i} = log\rho_i - \frac{log(1+\phi/\rho)}{2}$$

Note that the $\beta_{1,i}$ slope parameters increase monotonically with the clustering parameter ϕ, i.e., increased clustering is associated with greater dependence of the rates on the mixing variable. The above expressions assume for convenience that rates increase with the mixing variable, but negative relationships (e.g., the effect of ENSO on US hurricanes) can also be easily represented, e.g., by reversing the signs of the $\beta_{1,i}$ slope parameters and for binary mixing setting $\beta_{1,i} = \rho_i + \beta_{1,i}/2$. Since rates can never be negative, $\phi > \rho$ is nonadmissible for the binary mixing model, whereas any value $\phi \geq 0$ is allowed when using the logarithmic link function.

Choice of Row Overdispersion: Scaling Properties

The special case $\phi_i = \phi\rho_i/\rho$ can be proven to give dynamical rates for each row that scale similarly with the mixing variable, i.e., $\lambda_i = \rho_i f(Z)$ where $E(f(Z)) = 1$ and $Var(f(Z)) = \phi/\rho$. Conversely, rates scaling as $\lambda_i = \rho_i f(Z)$ with $E(f(Z)) = 1$ implies $\phi_i = \phi\rho_i/\rho$ where $\phi/\rho = Var(f(Z))$. This powerful equivalence holds for all possible distributions of mixing variable and choices of link function. An important consequence of such scaling is that the CEPs (and the mean loss per event) are independent of Z and so the mixing variable only affects the rates but not the distribution of the single event losses (see Hunter, 2014; Appendix C). Finally, such scaling guarantees that $Cov(\lambda_i, \lambda_j) = \phi\rho_i\rho_j/\rho$, and hence overdispersion in the number of all losses $\phi(0) = \phi$ when the link function is the identity.

The scaling choice, $\phi_i = \phi\rho_i/\rho$, generally leads to $\phi(u)$ decreasing from ϕ to zero with increasing threshold u. However, one may wish to create dynamic mixing models that can reproduce counts that have different clustering profiles, for example, overdispersion that stays constant with threshold, e.g., $\phi(u) = 0.3$ for all u. For the simplest case with no secondary uncertainty, this can be achieved by iteratively estimating the beta parameters starting at the row in the ELT that has the largest mean loss $\theta_{(1)}$. The dispersion parameter for this row can be estimated by setting it equal to the dispersion of counts exceeding $\theta_{(1)}$, i.e., $\phi_{(1)} = \phi(\theta_{(1)})$. The dispersion parameters for rows with progressively smaller θ_i can then be found iteratively by rearranging expressions for $\phi(u)$ such as this one is valid for linear link functions when there is negligible secondary uncertainty:

$$\phi(u) = \frac{\sum_{\theta_i>u}\sum_{\theta_j>u}\sqrt{\phi_i\phi_j\rho_i\rho_j}}{\sum_{\theta_k>u}\rho_k} \tag{3.1}$$

See Section 5.3.4 of Hunter (2014) for more details.

Models and Simulations

To illustrate these approaches, simulations have been made with five different ELT simulation models:

- M1: static sampling with no mixing, i.e., constant rates giving $\phi(u) = 0$
- M2: dynamic sampling with a binary mixing and $\phi_i = \phi\rho_i/\rho$ with $\phi = 0.3$
- M3: dynamic sampling with lognormal mixing and $\phi_i = \phi\rho_i/\rho$ with $\phi = 0.3$
- M4: dynamic sampling with binary mixing designed to give $\phi(u) = 0.3$
- M5: dynamic sampling with lognormal mixing designed to give $\phi(u) = 0.3$.

For each model, losses were simulated by the following algorithm:

```
Make J=100000 random draws of Z to represent varying annual conditions;

For j in years 1,2,…,J
{
  Calculate new rates λi(zj) for each row in the ELT;
  Draw annual counts Nj~Poi(λ(zj)) with total rate λ(zj) = ∑_i λi(zj);
  Draw Nj losses from loss distribution having parameters θ sampled Nj times by
  replacement from the θi row parameters with probability λi(zj)/λ(zj).
}
```

Results

Clustering

Fig. 3.3 shows how the overdispersion of the annual counts of the simulated losses depends on the threshold for each of the models. The static model, M1, is simply a mixture of compound Poisson processes and so has no dispersion for all threshold choices. Models M2 and M3 have decreasing dispersion with threshold, which closely follows $\phi = \mathrm{CEP}(u)$, which is obtained by substituting $\phi_i = \phi\rho_i/\rho$ into Eq. (3.1). Models M4 and M5 have succeeded in generating losses that have constant dispersion of 0.3 in counts above any threshold value, although the binary model shows a drop off due to the nonadmissibility condition mentioned earlier.

Aggregate Exceedance Probability

Fig. 3.4 shows AEP curves for simulated losses from the models M1−M5. Little difference can be seen between the curves for models M1, M2, and M3. Models M4 and M5 with constant overdispersion have AEP curves that differ from those of the other models for aggregate losses that are rarer than the 0.05 exceedance probability (20 year return period).

Constant overdispersion has substantially increased the risk of extreme aggregate loss as can be seen in risk measures shown in Table 3.1. The value-at-risk $\mathrm{VaR}_{0.995}$ for the models M2 and M3 are not significantly different from that of M1, but $\mathrm{VaR}_{0.995}$ for M4 and M5 are 18% and 32% greater than that for model M1. The difference between the expected shortfalls $ES_{0.995}$ is even more pronounced, with M5 having a shortfall that is 50% greater than that of M1. All models give the same expected annual aggregate loss, with a

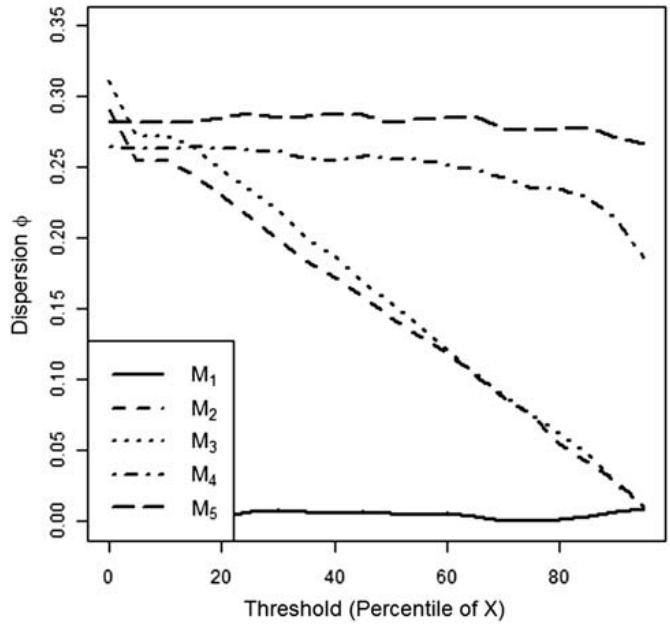

FIGURE 3.3 Overdispersion $\phi(u)$ of the number of simulated events $N(u)$ each year having losses exceeding threshold u. The threshold is expressed in terms of percentiles on the horizontal axis, i.e., $(1-\text{CEP}(u)) \times 100\%$. The static approach (model M1) has no overdispersion, models M2 and M3 have dispersion that decreases with loss intensity, and models M4 and M5 have dispersion that is constructed to stay constant with threshold.

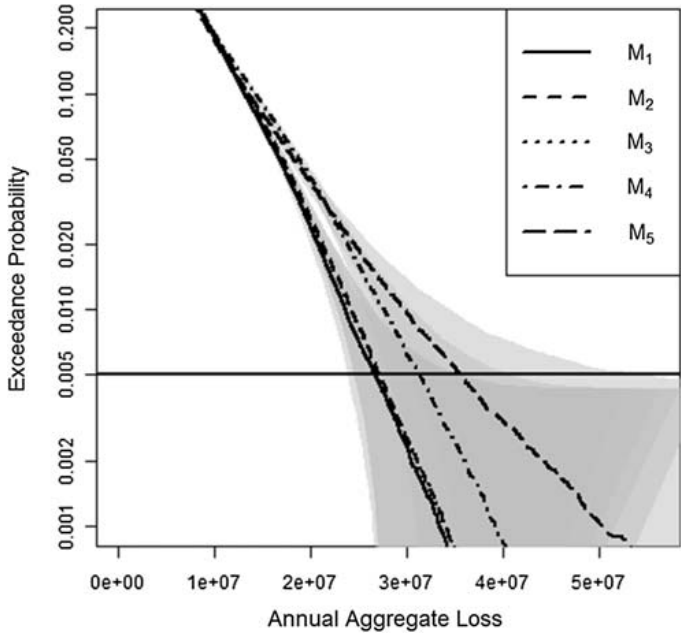

FIGURE 3.4 AEP curves for the loss models. Gray shading indicates the 95% Dvoretzky-Kiefer-Wolfowitz bounds (Hunter, 2014).

TABLE 3.1 Summary of Simulated Annual Aggregate Losses. For the VaR and *ES* Risk Measures, the Results for Models M2–M5 are Reported in Terms of Percentage Change From the Values for M1. The Overdispersion for all the Simulated Losses From Models M2–M5 are Close to the Chosen Value of 0.3—Any Differences are Solely Due to Sampling Variations

Model	Mean Loss	Overdispersion	$VaR_{0.995}$	$ES_{0.995}$
M1	6.4	0	26.8	30.5
M2	6.4	0.30	+1.4%	+3%
M3	6.4	0.32	+1.3%	+3%
M4	6.4	0.27	+17.1%	+19%
M5	6.4	0.29	+32.9%	+50%

consequence of constraining expectations of the rate parameters to be equal to the original rates for each row.

Fig. 3.5 shows AEP curves for the simulated losses conditional on specific ranges of the mixing variable Z. Not unexpectedly, the upper tail of the AEP curve depends strongly upon the value of the mixing variable for all the dynamical models M2–M5. The effect is greatest for the models with constant overdispersion.

Conditional Exceedance Probability

Fig. 3.6 shows CEP curves for the simulated losses conditional on specific ranges of the mixing variable Z. Models M2 and M3 give almost identical CEP curves to that from the M1 static approach as to be expected for rates that scale with Z (see Section Choice of Row Overdispersion: Scaling Properties). This is no longer the case for CEP curves for simulations from M4 and M5 where exceedance probability increases with Z. This positive association is clearly apparent in the scatter plot of the annual mean losses versus Z values simulated by model M5 (Fig. 3.7).

CONCLUDING REMARKS

This chapter has reviewed the assumptions used in stochastic simulations of natural catastrophe losses. For mathematical convenience, ELT simulation methods often assume that losses are identically and independently distributed in different years. Unfortunately, this static approach is inconsistent with the dynamical nonstationary behavior of the natural environment: natural hazards such as storms are known to be strongly influenced by climatic conditions that vary in time. Therefore, static simulation methods underestimate dependency and clustering especially for large loss events, which leads to large systematic underestimation of aggregate losses.

Loss calculations can be made more dynamical and realistic by introducing random variables to modulate the parameters in each row of the ELT. For hurricanes and windstorms, these mixing variables represent the behavior of large-scale indices describing modes of variability such as the El Nino/Southern Oscillation and the North Atlantic Oscillation. Such a dynamical mixture approach is similar to recent approaches used to model portfolio credit

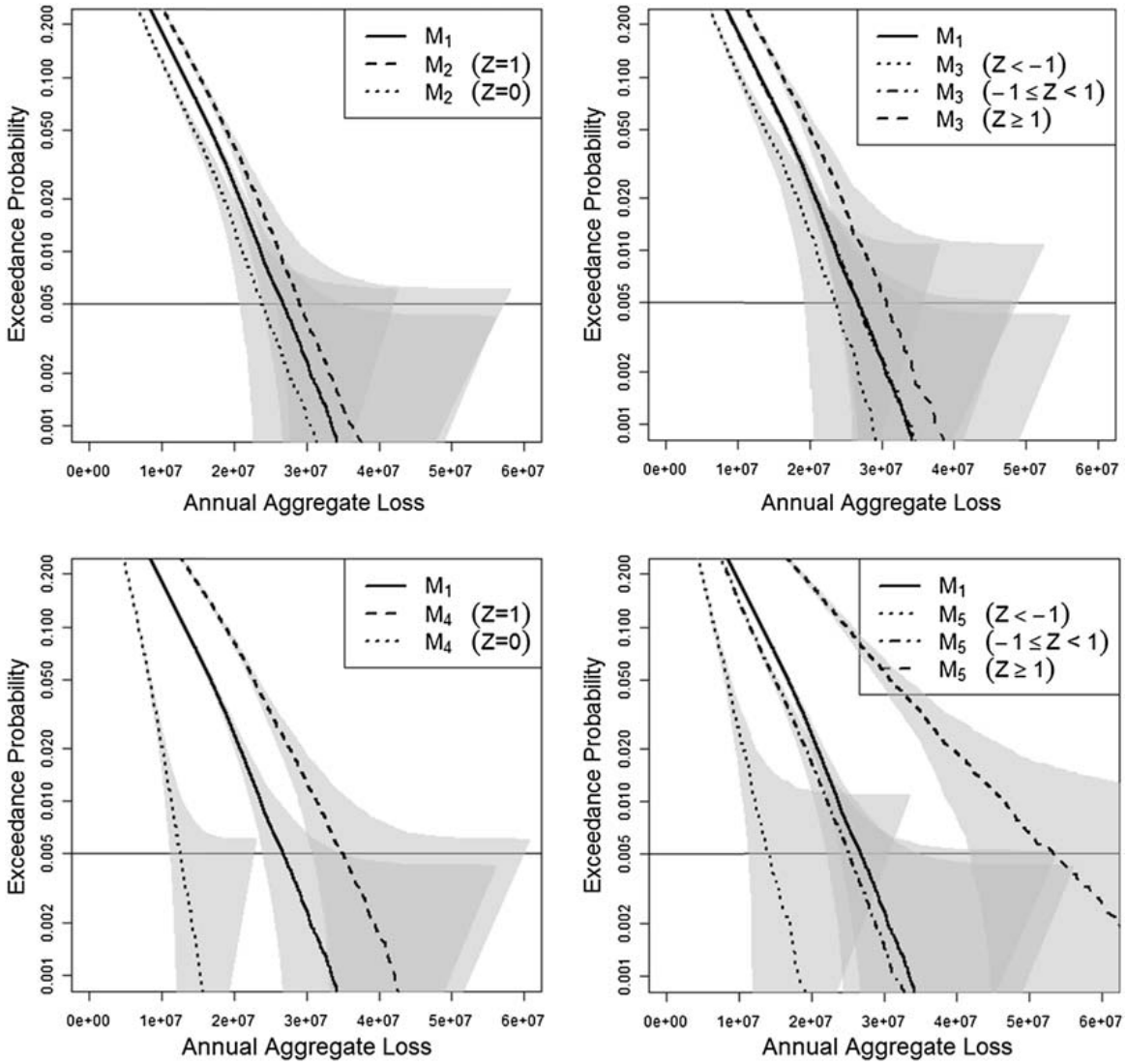

FIGURE 3.5 Conditional AEP curves simulated by models M2−M5 for different ranges of mixing variable. CEP curve for model M1 is shown as reference (*solid black line*) and gray shading indicates the 95% Dvoretzky-Kiefer-Wolfowitz bounds (Hunter, 2014).

risk, e.g., the CreditRisk + method based on binary mixing variables. This chapter has illustrated these approaches by applying four different toy mixing models to an ELT constructed to be typical of that used in practice for European windstorms. All of these models increase the probability of large aggregate losses (i.e., thicken the tail of the AEP curve), especially those models designed to maintain clustering for large loss events.

To make further progress, it is necessary to develop better more reliable methods for quantifying how clustering of hazards depends on intensity; for example, does it stay constant or sometimes even increase with threshold? To properly capture intensity-rate correlations,

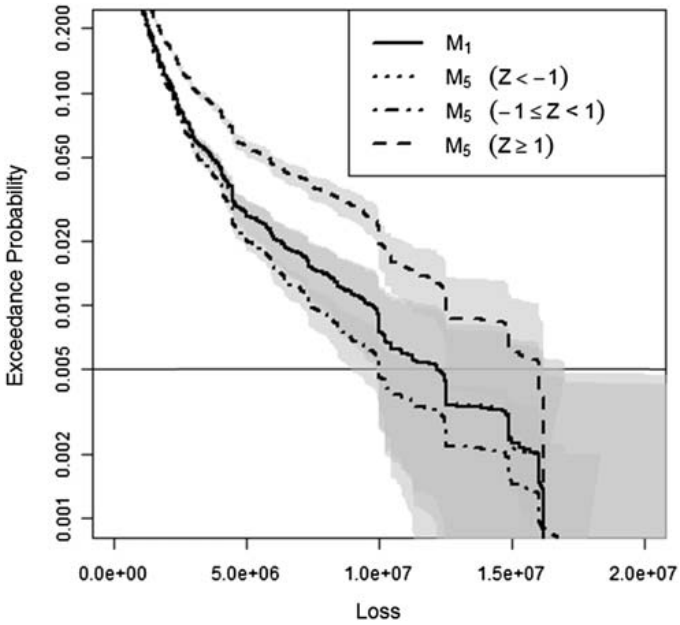

FIGURE 3.6 Conditional CEP curves simulated by model M5 for different ranges of mixing variable. CEP curve for model M1 is shown as reference (*solid black line*) and gray shading indicates the 95% Dvoretzky-Kiefer-Wolfowitz bounds (Hunter, 2014). By definition, models M1, M2, and M3 give CEP curves that do not depend on the mixing variable (see Section Estimation of the Mixing Parameters: $\beta_{0,i}$ and $\beta_{1,i}$).

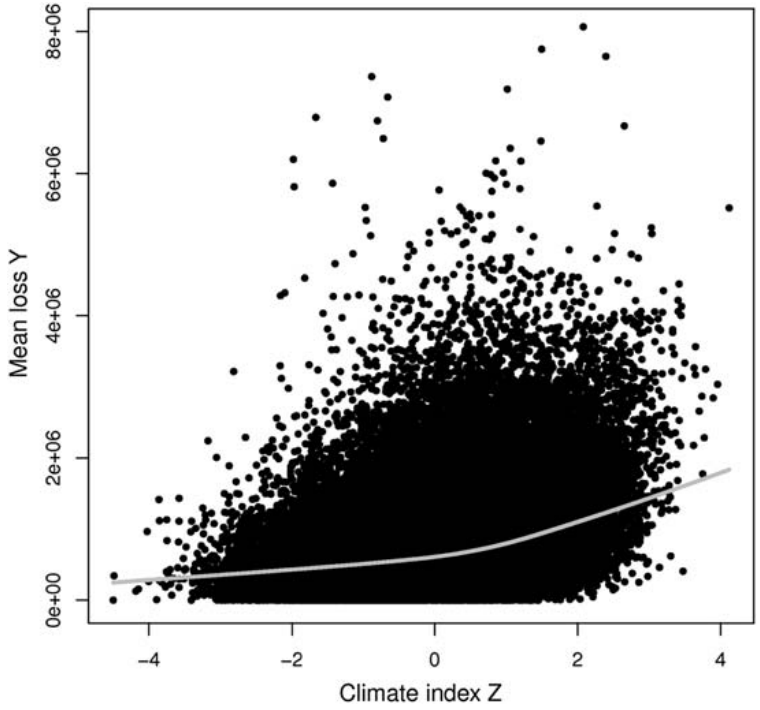

FIGURE 3.7 Scatter plot of mean loss per event (i.e., ratio of aggregate loss and number of events per year) versus mixing variable Z for 100,000 years simulated by model M5. Mean loss shows a strong nonlinear dependence on the mixing variable as indicated by the robust smoothing curve (LOWESS trend; *gray line*).

mixing models need to be developed that can allow the parameters of the loss distribution to be modulated rather than just the rate parameters. More robust estimation methods need to be developed for finding the beta parameters for each row of the ELT, e.g., for the case of increasing clustering profiles in the presence of secondary uncertainty.

Acknowledgments

We wish to thank the Willis Research Network for fellowship funding of Benjamin Youngman and for stimulating discussions especially with Rick Thomas. We are also grateful to Jonty Rougier for providing comments on an earlier draft of the chapter.

References

Crouhy, M., Galai, D., Mark, R., 2000. A comparative analysis of current credit risk models. Journal of Banking and Finance 24 (1), 59–117.

Dowd, K., Blake, D., 2006. After VaR: the theory, estimation, and insurance applications of quantile – based risk measures. Journal of Risk and Insurance 73 (2), 193–229.

Economou, T., Stephenson, D.B., Pinto, J., Shaffrey, L.C., Zappa, G., 2015. Serial clustering of extratropical cyclones in a multi-model ensemble of historical and future simulations. Quarterly Journal of the Royal Meteorological Society 141 (693), 3076–3087.

Embrechts, P., Frey, R., McNeil, A., 2005. Quantitative risk management, vol. 10. Princeton Series in Finance, Princeton.

Frey, R., McNeil, A.J., 2003. Dependent defaults in models of portfolio credit risk. Journal of Risk 6, 59–92.

Friedman, D.G., 1972. Insurance and the natural hazards. The ASTIN Bulletin: International Journal for Actuarial Studies in Non-Life Insurance and Risk Theory 7 (1), 4–58.

Grossi, P., Kunreuther, H., 2005. In: Catastrophe Modeling: A New Approach to Managing Risk, vol. 25. Springer, p. 245.

Hunter, A., 2014. Quantifying and Understanding the Aggregate Risk of Natural Hazards (Ph.D. thesis). University of Exeter, 163 pp. Available from: http://ethos.bl.uk.

Hunter, A., Stephenson, D.B., Economou, T., Holland, M., Cook, I., 2016. New perspectives on the collective risk of extratropical cyclones. Quarterly Journal the of Royal Meteorological Society 142 (694), 243–256.

Khare, S., Bonazzi, A., Mitas, C., Jewson, S., 2015. Modelling clustering of natural hazard phenomena and the effect on re/insurance loss perspectives. Natural Hazards Earth System Sciences 15, 1357–1370.

Kvamstø, N.-G., Song, Y., Seierstad, I., Sorteberg, A., Stephenson, D.B., 2008. Clustering of cyclones in the ARPEGE general circulation model. Tellus A 60 (3), 547–556.

Mailier, P., Stephenson, D., Ferro, C., Hodges, K., 2006. Serial clustering of extratropical cyclones. Monthly Weather Review 134 (8), 2224–2240.

Mumby, P.J., Vitolo, R., Stephenson, D.B., 2011. Temporal clustering of tropical cyclones and its ecosystem implications. Proceedings of the National Academy of Sciences USA 108, 17626–17630.

Pinto, J.G., Bellenbaum, N., Karremann, M.K., Della-Marta, P.M., 2013. Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions. Journal of Geophysical Research: Atmospheres 118 (22), 12476–12485.

Pinto, J.G., Gómara, I., Masato, G., Dacre, H.F., Woollings, T., Caballero, R., 2014. Large-scale dynamics associated with clustering of extratropical cyclones affecting Western Europe. Journal of Geophysical Research: Atmospheres 119 (24), 13704–13719.

Rolski, T., Schmidli, H., Schmidt, V., Teugels, J., 2009. In: Stochastic Processes for Insurance and Finance, vol. 505. John Wiley & Sons, p. 654.

Villarini, G., Smith, J.A., Vitolo, R., Stephenson, D.B., 2013. On the temporal clustering of US floods and its relationship to climate teleconnection patterns. International Journal of Climatology 33 (3), 629–640.

Vitolo, R., Stephenson, D., Cook, I., Mitchell-Wallace, K., 2009. Serial clustering of intense European storms. Meteorologische Zeitschrift 18 (4), 411–424.

This page intentionally left blank