Estimating the Volatility of Non-Life Premium Risk Under Solvency II: Discussion of Danish Fire Insurance Data

Rocco Roberto Cerchiara [1] and Francesco Acri [2],*

[1] Department of Economics, Statistics and Finance "Giovanni Anania", University of Calabria, 87036 Arcavacata di Rende (CS), Italy; rocco.cerchiara@unical.it
[2] Independent Researcher, 20135 Milan, Italy
* Correspondence: francesco_acri@yahoo.it

Abstract: We studied the volatility assumption of non-life premium risk under the Solvency II Standard Formula and developed an empirical model on real data, the Danish fire insurance data. Our empirical model accomplishes two things. Primarily, compared to the present literature, this paper innovates the fitting of Danish fire insurance data using a composite model with a random threshold. Secondly we prove, by fitting the Danish fire insurance data, that for large insurance companies the volatility of the standard formula is higher than the volatility estimated with internal models such as composite models, also taking into account the dependence between attritional and large claims.

1. Introduction

A non-life insurance company faces premium risk, among others, which is the risk of financial losses related to premiums earned. The risk in the losses relates to uncertainty in severity, frequency or even timing of claims incurring during the period of exposure. For an operating non-life insurer, premium risk is a key driver of uncertainty both from operational and solvency perspectives. In regards to the solvency perspective, there are many different methods useful to give a correct view of the capital needed to meet adverse outcomes related to premium risk. In particular, evaluation of the distribution of aggregate loss plays a fundamental role in the analysis of risk and solvency levels.

As shown in Cerchiara and Demarco (2016), the standard formula under Solvency II for premium and reserve risk defined by the Delegated Acts (DA, see European Commission 2015) proposes the following formula for the solvency capital requirement (SCR):

$$SCR = 3\sigma V \tag{1}$$

where V denotes the net reinsurance volume measure for Non-Life premium and reserve risk determined in accordance with Article 116 of DA, and σ is the volatility (coefficient of variation) for Non-Life premium and reserve risk determined in accordance with Article 117 of DA, combining the volatility σ_s according to the correlation matrix between each segment s. Then, σ_s is calculated as follows:

$$\sigma_s = \frac{\sqrt{\sigma^2_{(prem,s)} V^2_{(prem,s)} + \sigma_{(prem,s)} V_{(prem,s)} \sigma_{(res,s)} V_{(res,s)} + \sigma^2_{(res,s)} V^2_{(res,s)}}}{V_{(prem,s)} + V_{(res,s)}} \qquad (2)$$

In this paper we focus our attention on $\sigma_{(prem,s)}$, i.e., the coefficient of variation of the Fire segment Premium Risk. Under DA, the premium risk volatility of this segment is equal to 8%. As shown in Clemente and Savelli (2017), Equation (1) "implicitily assumes to measure the difference between the Value at Risk (VaR) at 99.5% confidence level and the mean of the probability distribution of aggregate claims amount by using a fixed multiplier of the volatility equal to 3 for all insurers. From a practical point of view, DA multiplier does not take into account the skewness of the distribution with a potential underestimation of capital requirement for small insurers and an overestimation for big insurers".

For Insurance Undertakings who do not believe in the standard formula assumptions, they may calculate the Solvency Capital Requirement using an Undertaking Specific Parameters approach (USP, see Cerchiara and Demarco 2016) and a Full or Partial Internal Model (PIM) after approval by Supervisory Authorities. Calculation of the volatility and VaR of independent or dependent risky positions using PIM is very difficult for large portfolios. In the literature, many different studies are based on definitions of composite models that aim to analyze loss distribution and dependence between the main factors that characterize the risk profile of insurance companies, e.g., frequency and severity, attritional and large claims and so forth. Considering more recent developments in the literature, Galeotti (2015) proves the convergence of a geometric algorithm (alternative to Monte Carlo and quasi-Monte Carlo methods) for computing the Value-at-Risk of a portfolio of any dimension, i.e., the distribution of the sum of its components, which can exhibit any dependence structure.

In order to implement PIM and investigate overestimation of the SCR (and the underlying volatility) for large insurers, we used the Danish fire insurance dataset[1] that has been often analyzed according to the parametric approach and composite models. McNeil (1997), Resnick (1997), Embrechts et al. (1997) and McNeil et al. (2005) proposed fitting this dataset using Extreme Value Theory and Copula Functions (see Klugman et al. 2010 for more details on the latter), with special reference to modeling the tail of the distribution. Cooray and Ananda (2005) and Scollnik (2007) showed that the composite lognormal-Pareto model was a better fit than standard univariate models. Following the previous two papers, Teodorescu and Vernic (2009, 2013) fit the dataset firstly with a composite Exponential and Pareto distribution, and then with a more general composite Pareto model obtained by replacing the Lognormal distribution by an arbitrary continuous distribution, while Pigeon and Denuit (2011) considered a positive random variable as the threshold value in the composite model. There have been several other approaches to model this dataset, including Burr distribution for claim severity using XploRe computing environment (Burnecki and Weron 2004), Bayesian estimation of finite time ruin probabilities (Ausin et al. 2009), hybrid Pareto models (Carreau and Bengio 2009), beta kernel quantile estimation (Charpentier and Oulidi 2010) and bivariate composite Poisson process (Esmaeili and Klüppelberg 2010). An example of non-parametric modeling is shown in Guillotte et al. (2011) with a Bayesian inference on bivariate extremes. Drees and Müller (2008) showed how to model dependence within joint tail regions. Nadarajah and Bakar (2014) improved the fittings for the Danish fire insurance data using various new composite models, including the composite Lognormal–Burr model.

Following this literature, this paper innovates fitting of the Danish fire insurance data by using the Pigeon and Denuit (2011) composite model with a random threshold that has a higher goodness-of-fit than the Nadarajah and Bakar (2014) model. Once the best model is defined, we show

---

[1] Danish insurance market data have been included in the calibration of standard parameters by EIOPA (2011) Calibration of the Premium and Reserve Risk Factors in the Standard Formula of Solvency II, Report of the Joint Working Group on Non-Life and Health NSLT Calibration.

that the Standard formula assumption is prudent, especially for large insurance companies, giving an overestimated volatility of the premium risk (and of the SCR). For illustrative purposes, we also investigate the use of other models, including the Copula function and Fast Fourier Transform (FFT; Robe-Voinea and Vernic 2016), trying to take into account the dependence between attritional and large claims and understand the effect on SCR.

The paper is organized as follows. In Section 2 we report some statistical characteristics of Danish data. In Sections 3 and 4 we posit there is no dependence between attritional and large claims. We investigate the use of composite models with fixed or random thresholds in order to fit the Danish fire insurance data, and we compare our numerical results with the fitting of Nadarajah and Bakar (2014) based on a composite Lognormal–Burr model. In Section 5 we try to appraise risk dependence through the Copula function concept and FFT, for which Robe-Voinea and Vernic (2016) provide an overview and perform a multidimensional application. Section 6 concludes the work and presents estimation of the aggregate loss volatility distribution, and results are compared under independence and dependence conditions.

2. Data

In the following, we show some statistics of the dataset used in this analysis. The losses of individual fires covered in Denmark were registered by the reinsurance company Copenhagen Re and, for our study, have been converted into euros. It is worth mentioning that the original dataset (available also in R) covers the period 1980–1990. In 2003–2004, Mette Havning (Chief Actuary of Danish Reinsurance) was on the Astin committee where she met Tine Aabye from Forsikring & Pension. Aabye asked her colleague to send the Danish million re-losses from 1985–2002 to Mette Havning. Based on the two versions from 1980–1990 and 1985–2002, Havning then made an extended version of Danish Fire Insurance Data from 1980 through 2002 with only a few ambiguities in the overlapping period. The data were communicated to us by Mette Havning and consisted of 6870 claims over a time period of 23 years. We bring to the reader's attention that, to avoid seasonal effects due to the use of the entire historical series that starts from 1980, the costs have been inflated to 2002. In addition, we referred to a wider dataset also including small losses, unlike that used by McNeil (1997), among others. In fact, we want to study the entire distribution of this dataset, while in McNeil (1997) and in other works the attention was focused especially on the right-tail distribution. We list some descriptive statistics in Table 1:

Table 1. Descriptive statistics of Danish empirical losses.

Min	Mean	Median	Q3	Max	Kurtosis	Skewness	Std Dev
27,730	613,100	327,000	532,800	55,240,000	505.592	17.635	1,412,959

The maximum observed was around €55 million and the average cost was 613,100€. The empirical distribution is definitely leptokurtic and asymmetric to the right.

To make applications of composite models and Copula functions easier, we will suppose that claim frequency k is non-random, while for the Fast Fourier Transform algorithm we consider the frequency as a random variable. The losses have been split by year, so we can report some descriptive statistics for k in Table 2:

Table 2. Empirical distribution of frequency claim statistics.

Q1	Mean	Q3	Max	Variance	Skewness
238	299	381	447	8482	−0.12

We note 50% of frequencies were included between 238 and 381 claims, and there is slight negative asymmetry. In addition, the variance is greater than mean value (299).

3. Composite Models

In the Danish fire insurance data we can find both frequent claims with low to medium severity and sporadic claims with high severity. If we want to define a joint distribution for these two types of claims we have to build a composite model.

A composite model is a combination of two different models: One having a light tail below a threshold (attritional claims) and another with a heavy tail suitable to model the value that exceeds this threshold (large claims). Composite distributions (also known as composite, spliced or piecewise distributions) have been introduced in many applications. Klugman et al. (2010) expressed the probability density function of a composite distributions as

$$f(x) = \begin{cases} r_1 f_1^*(x) & k_0 < x < k_1 \\ \vdots \\ r_n f_n^*(x) & k_{n-1} < x < k_n \end{cases} \tag{3}$$

where f_j^* is truncated probability density function of marginal distribution f_j, $j = 1, \ldots, n$; $r_j \geq 0$ are mixing weights, $\sum_{j=1}^{n} r_j = 1$; and k_j defines the range limit of the domain.

Formally, the density distribution of a composite model can be written as a special case of (3) as follows:

$$f(x) = \begin{cases} r f_1^*(x) & -\infty < x \leq u \\ (1-r) f_2^*(x) & u < x < \infty \end{cases} \tag{4}$$

where $r \in [0,1]$, and f_1^* and f_2^* are cut off density distributions of marginals f_1 and f_2, respectively. In detail, if F_i is the distribution function of f_i, $i = 1, 2$, then we have

$$\begin{cases} f_1^*(x) = \frac{f_1(x)}{F_1(u)} & -\infty < x \leq u \\ f_2^*(x) = \frac{f_2(x)}{1 - F_2(u)} & u < x < \infty \end{cases} \tag{5}$$

It is simple to note that (4) is a convex combination of f_1^* and f_2^* with weights r and $1 - r$. In addition, we want (4) to be continuous, derivable and with a continuously derivative density function, and for this scope we fix the following limitation:

$$\begin{cases} \lim_{x \to u} f(x) = f(u) \\ \lim_{x \to u^-} f'(x) = \lim_{x \to u^+} f'(x) \end{cases} \tag{6}$$

From the first we obtain

$$r = \frac{f_2(u) F_1(u)}{f_2(u) F_1(u) + f_1(u)(1 - F_2(u))} \tag{7}$$

while from the second

$$r = \frac{f_2'(u) F_1(u)}{f_2'(u) F_1(u) + f_1'(u)(1 - F_2(u))} \tag{8}$$

We can define distribution function F of (4)

$$F(x) = \begin{cases} r \frac{F_1(x)}{F_1(u)}, & -\infty < x \leq u \\ r + (1-r) \frac{F_2(x) - F_2(u)}{1 - F_2(u)}, & u < x < \infty \end{cases} \tag{9}$$

Suppose F_1 and F_2 admit inverse functions; we can define the quantile function via an inversion method. Let p be a random number from a standard Uniform distribution, then the quantile function results in

$$F^{-1}(p) = \begin{cases} F_1^{-1}\left(\frac{p}{r}F_1(u)\right), & p \leq r \\ F_2^{-1}\left(\frac{p-r+(1-p)F_2(u)}{1-r}\right), & p > r \end{cases} \tag{10}$$

To estimate the parameters of (9) we can proceed as follows: First, we estimate marginal density function parameters separately (the underlying hypothesis is that there is no relation between attritional and large claims); then, these estimates will be the start values of the density function in order to maximize the following likelihood:

$$L(x_1, \ldots, x_n; \boldsymbol{\theta}) = r^m (1-r)^{n-m} \prod_{i=1}^m f_1^*(x_i) \prod_{j=m+1}^n f_2^*(x_j) \tag{11}$$

where n is the sample dimension, $\boldsymbol{\theta}$ is a vector including composite model parameters, while m is such that $X_m \leq u \leq X_{m+1}$, otherwise m is the level of order statistics X_m immediately previous (or coincident) to u.

The methodology described in Teodorescu and Vernic (2009, 2013) has been used in order to estimate threshold u, which permits us to discriminate between attritional and large claims.

3.1. Composite Model with Random Threshold

We can define a composite model also using a random threshold (see Pigeon and Denuit 2011). In particular, given a random sample $X = (X_1, \ldots, X_n)$, we can assume that every single component X_i provides its own threshold. So, for the generic observation x_i we will have the threshold u_i, $i = 1, \ldots, n$. For this reason, $u_1, \ldots, u_n$ are realizations of a random variable U with a distribution function G. The random variable U is necessarily non-negative and with a heavy-tailed distribution.

A composite model with a random threshold shows a completely new and original aspect: Not only are we unable to choose a value for u, but its whole distribution and the parameters of the latter are implicit in the definition of the composite model. In particular, we define the density function of the Lognormal–Generalized Pareto Distribution model (GPD, see (Embrechts et al. 1997)) with a random threshold in the following way:

$$f(x) = (1-r) \int_0^x f_2(x)g(U)dU + r \int_x^\infty \frac{1}{\Phi(\xi\sigma)} f_1(x)g(U)dU \tag{12}$$

where $r \in [0,1]$, U is the random threshold with density function g, f_1 and f_2 are Lognormal and GPD density functions, respectively, Ψ is the Standard Normal distribution function, ξ is the shape parameter of GPD and σ is the Lognormal scale parameter.

3.2. Kumaraswamy Distribution and some Generalization

In this section we describe the Kumaraswamy Distribution (see Kumaraswamy 1980) and a generalization of the Gumbel distribution (see Cordeiro et al. 2012). In particular, let

$$K(x; \alpha, \beta) = 1 - (1 - x^\alpha)^\beta, x \in (0,1) \tag{13}$$

in the distribution proposed in Kumaraswamy (1980), where α and β are non-negative shape parameters. If G is the distribution function of a random variable, then we can define a new distribution by

$$F(x; a, b) = 1 - (1 - G(x)^a)^b \tag{14}$$

where $a > 0$ and $b > 0$ are shape parameters that influence kurtosis and skewness. The Kumaraswamy–Gumbel (KumGum) distribution is defined throughout (14) with the following distribution function (see Cordeiro et al. 2012):

$$F_{KG}(x; a, b) = 1 - (1 - \Lambda(x)^a)^b \tag{15}$$

where $\Lambda(x)$ is the Gumbel distribution function with density defined by (20). The quantile function of KumGum is obtained by inverting (15) and explicating Gumbel parameters (v and ϕ):

$$x_p = F^{-1}(p) = v - \varphi \log \left[-\log \left(1 - (1 - p)^{1/b} \right)^{1/a} \right] \tag{16}$$

with $p \in (0, 1)$.

The following Table 3 and Figure 1 show the Kurtosis and Skewness of the KumGum density function by varying the four parameters:

Table 3. Kurtosis and Skewness of the KumGum distribution.

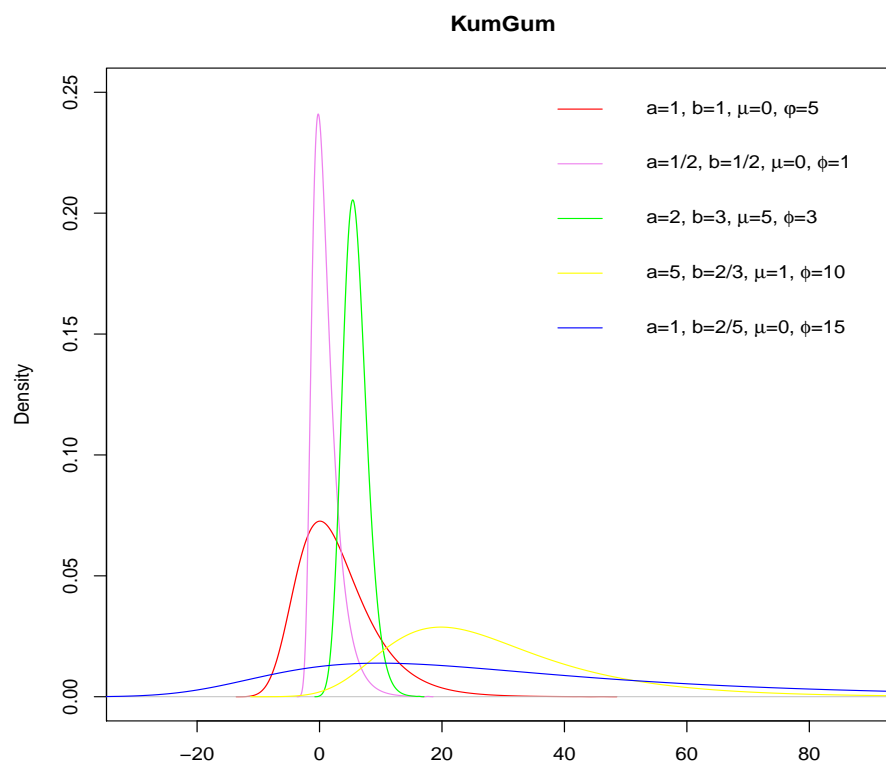| v	φ	a	b	Kurtosis	Skewness
0	5	1	1	5.4	1.1
0	1	0.5	0.5	7.1	1.6
5	3	2	3	3.6	0.5
1	10	5	0.66	6.4	1.4
0	15	1	0.4	7.6	1.7

Figure 1. KumGum density functions.

Another generalization of the Kum distribution is the Kumaraswamy–Pareto one (KumPareto). In particular, we can evaluate Equation (14) in the Pareto distribution function P which is

$$P(x; \beta, \kappa) = 1 - \left(\frac{\beta}{x} \right)^{\kappa}, x \geq \beta \tag{17}$$

where $\beta > 0$ is a scale parameter, and $\kappa \geq 0$ is a shape parameter. Thus, from (13), (14) and (17) we obtain the KumPareto distribution function:

$$F_{KP}(x; \beta, \kappa, a, b) = 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^{\kappa} \right]^{a} \right\}^{b}, x \geq 0 \tag{18}$$

The corresponding quantile function is

$$F^{-1}(p) = \beta \left\{ \left\{ 1 - \left[1 - \left(1 - p \right)^{1/b} \right]^{1/a} \right\}^{1/\kappa} \right\}^{-1} \tag{19}$$

where $p \in (0, 1)$. In the following Figure 2 we report the KumPareto density function varying the parameters:
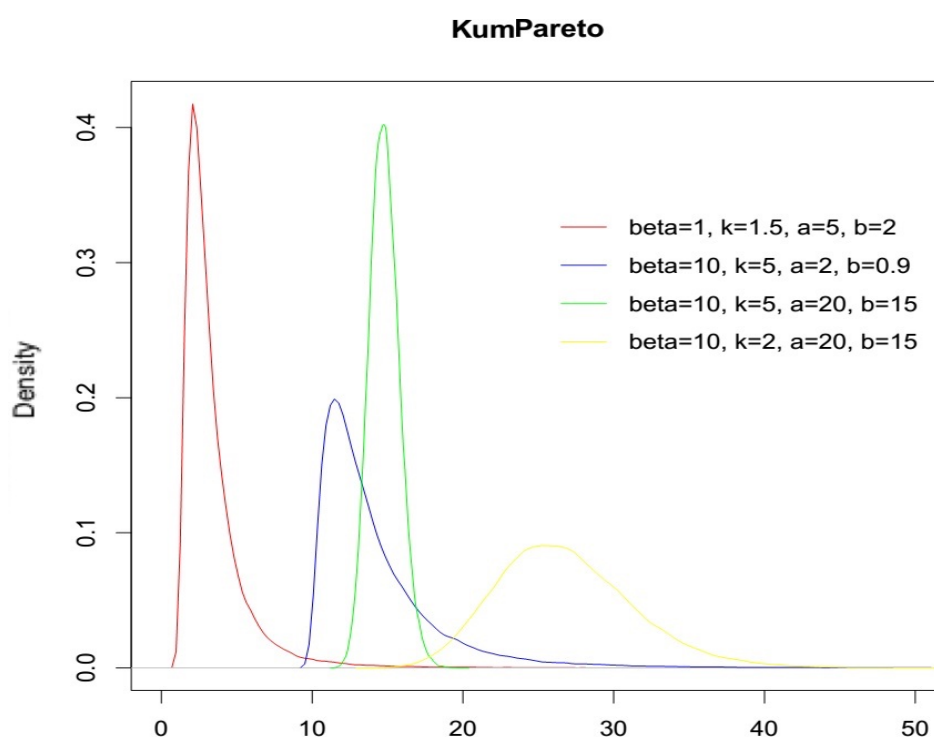

Figure 2. KumPareto density functions.

4. Numerical Example of Composite Models

In this section we present numerical results on the fitting of Danish fire insurance data by composite models with constant and random thresholds between attritional and large claims. As already mentioned, for the composite models with a constant threshold, we used the methodology described in Teodorescu and Vernic (2009, 2013), obtaining $u = 1,022,125€$.

We start with a composite Lognormal–KumPareto model, choosing $f_1 \sim$ *Lognormal* and $f_2 \sim$ *KumPareto*. From the following Table 4 we can compare some theoretical and empirical quantiles:

Table 4. Comparison between empirical and Lognormal–KumPareto quantiles.

Level	50%	75%	90%	95%	99%	99.5%
Empirical quantile	327,016	532,757	1,022,213	1,675,219	5,484,150	8,216,877
Theoretical quantile	333,477	462,852	642,196	840,161	2,616,338	4,453,476

Only the fiftieth percentile of the theoretical distribution function was very close to the same empirical quantile: From this percentile onwards the differences increased. In the following Figure 3 we show only right tails of the distribution functions (empirical and theoretical):
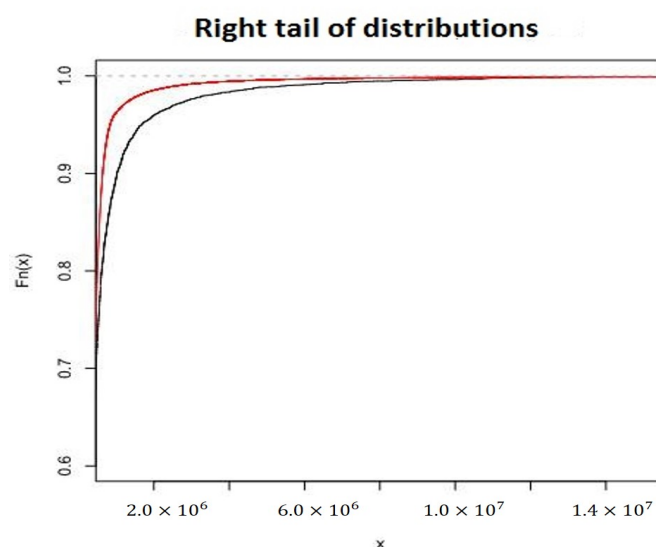

Figure 3. Right tails of Lognormal–KumPareto (red line) and empirical distribution (dark line) functions.

The red line is always above the dark line. This means Kumaraswamy-generalized families of distributions are very versatile in analyzing different types of data, but in this case the Lognormal–KumPareto model underestimated the right tail.

Therefore, we consider the composite model $f_1 \sim$ *Lognormal* and $f_2 \sim$ *Burr* as suggested in Nadarajah and Bakar (2014). The parameters are estimated using the CompLognonormal R package as shown in Nadarajah and Bakar (2014). From the following Table 5 we can compare some theoretical quantiles with empirical ones:

Table 5. Comparison between empirical quantiles and Lognormal–Burr ones.

Level	50%	75%	90%	95%	99%	99.5%
Empirical quantile	327,016	532,757	1,022,213	1,675,219	5,484,150	8,216,877
Theoretical quantile	199,681	332,341	634,531	1,029,262	3,189,937	5,181,894

This model seemed to be more feasible in catching the right tail of the empirical distribution with respect to the previous Lognormal–KumPareto, as we can see from the Figure 4 below:

Similar to the Lognormal–KumPareto model, the Lognormal–Burr distribution line is always above the empirical distribution line but not always at the same distance.

We go forward modeling a Lognormal–Generalized Pareto Distribution (GPD), that is we choose $f_1 \sim$ *Lognormal* and $f_2 \sim$ *GPD* and then we generate pseudo-random numbers from quantile function (10). In Table 6 and Figure 5 we report the estimates of parameters, 99% confidence intervals and the QQ plot (μ_1 and σ are the Lognormal parameters, while σ_μ and ξ are GPD parameters):

We observe that this composite model adapted well to the empirical distribution; in fact, except for a few points, theoretical quantiles are close to corresponding empirical quantiles. In the Figures 6 and 7 we compare the theoretical cut-off density function with the corresponding empirical one and the theoretical right tail with the empirical one.

The model exhibited a non-negligible right tail (kurtosis index is 115,656.2), which can be evaluated comparing the observed distribution function with the theoretical one.
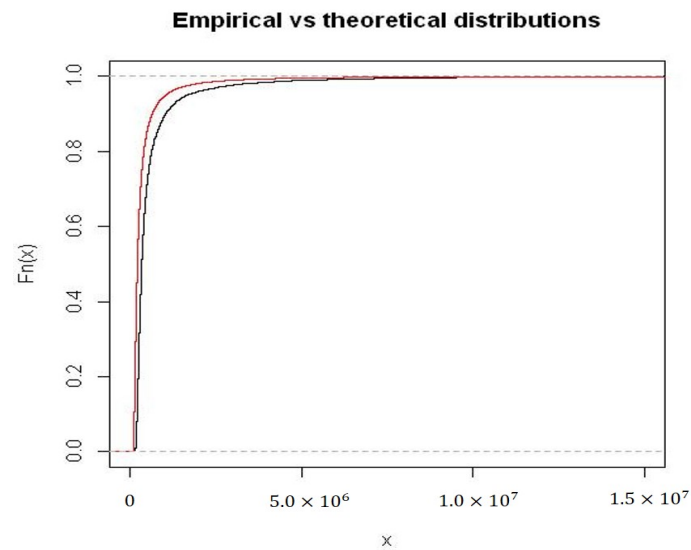
Empirical vs theoretical distributions

Figure 4. Lognormal–Burr and empirical distribution functions (red and dark lines).

Table 6. Estimated parameters and 99% confidence intervals of Lognormal–GPD.

Parameter	Low Extreme	Best Estimate	High Extreme
μ_1	12.82	12.84	12.86
σ	0.59	0.61	0.62
σ_μ	1,113,916	1,115,267	1,116,617
ξ	0.33	0.45	0.56

QQ−plot
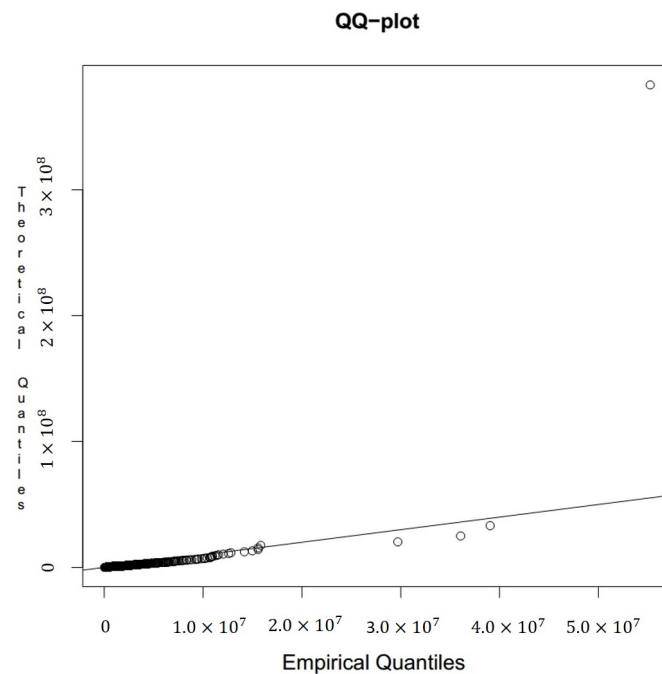
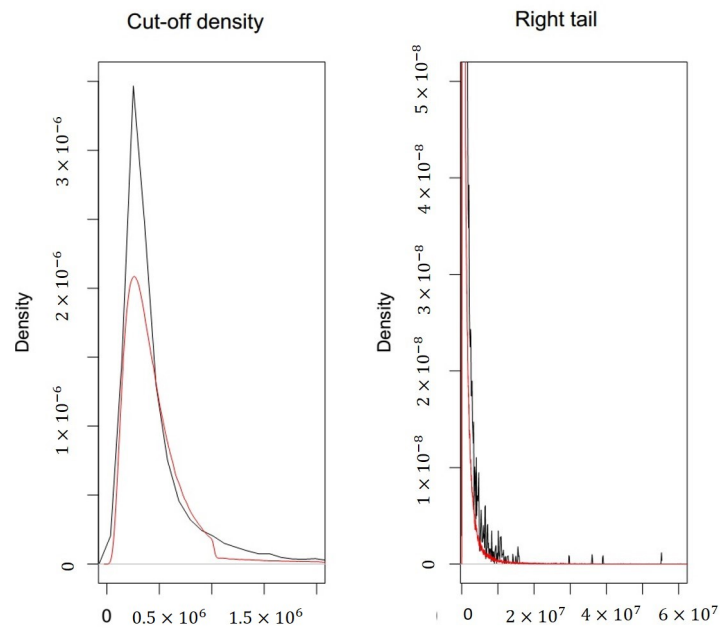**Figure 5.** Observed–theoretical quantile plot for the Lognormal–GPD model.

Figure 6. Left, comparison between cut-off density functions. Right, empirical and theoretical (red) right tail.
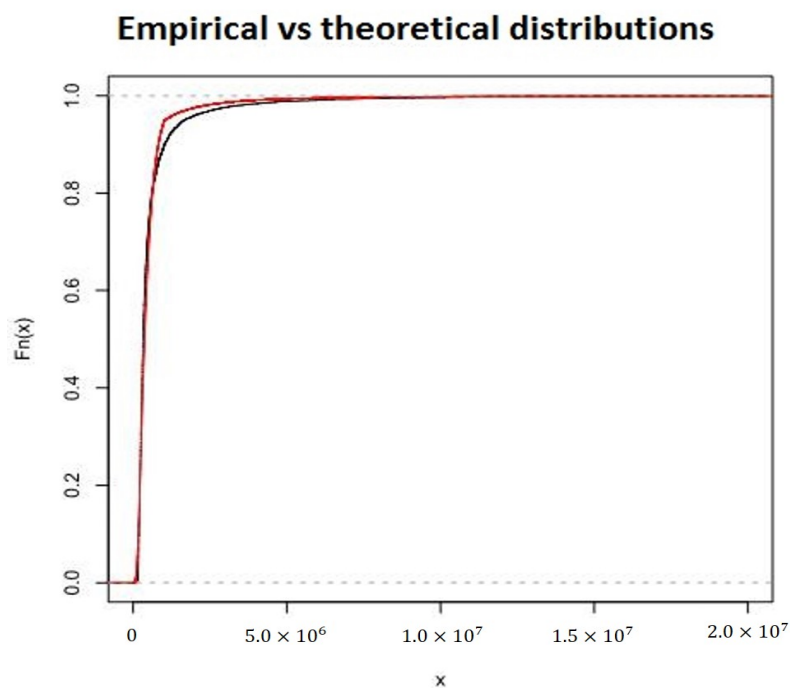

Figure 7. Lognormal–GPD (red) and empirical (dark) distribution function.

The corresponding Kolmogorov–Smirnov test returned a p-value equal to 0.8590423, using 50,000 bootstrap samples.

Finally, in Table 7 we report the best estimate and 99% confidence intervals of the composite model Lognormal–GPD with a Gamma random threshold u (see Pigeon and Denuit 2011).

Table 7. Estimated parameters and 99% confidence intervals of Lognormal–GPD–Gamma distribution.

Parameter	Low Extreme	Best Estimate	High Extreme
μ_1	12.78	12.79	12.81
σ	0.52	0.54	0.55
u	629,416	630,768	632,121
σ_μ	1,113,915	1,115,266	1,116,616
ξ	0.22	0.29	0.37

The threshold u is a parameter whose value depends on Gamma parameters. In the following Table 8 and Figure 8 we report the theoretical and empirical quantiles and the QQ plot.

Table 8. Comparison between empirical and Lognormal–GPD–Gamma quantiles.

Levels	50%	75%	90%	95%	99%	99.5%
Empirical percentile	327,016	532,757	1,022,213	1,675,219	5,484,150	8,216,877
Theoretical percentile	360,574	517,996	1,103,309	2,077,792	5,266,116	7,149,253
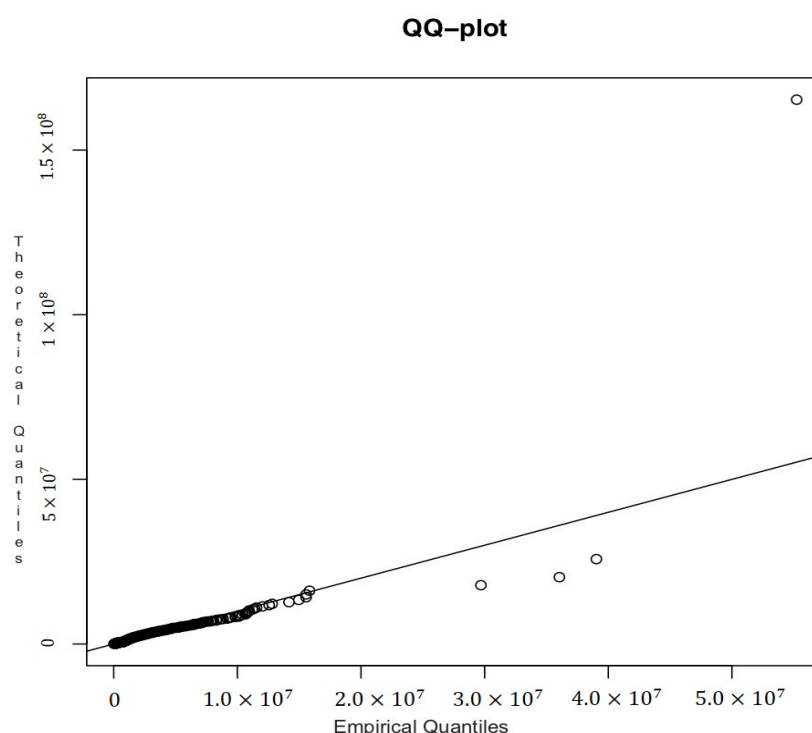
Figure 8. Observed-theoretical quantile plot for the Lognormal–GPD–Gamma model.

We can see from the Figure 9 that Lognormal–GPD–Gamma model can be considered a good fitting model.

The Kolmogorov–Smirnov adaptive test returned p-value equal to 0.1971361; therefore, we cannot reject the null hypothesis under which the investigated model is a feasible model for our data.

Finally, Lognormal–KumPareto, Lognormal–Burr, Lognormal–GPD with fixed threshold and Lognormal–GPD with a Gamma random threshold can be compared using the AIC and BIC values, Table 9.
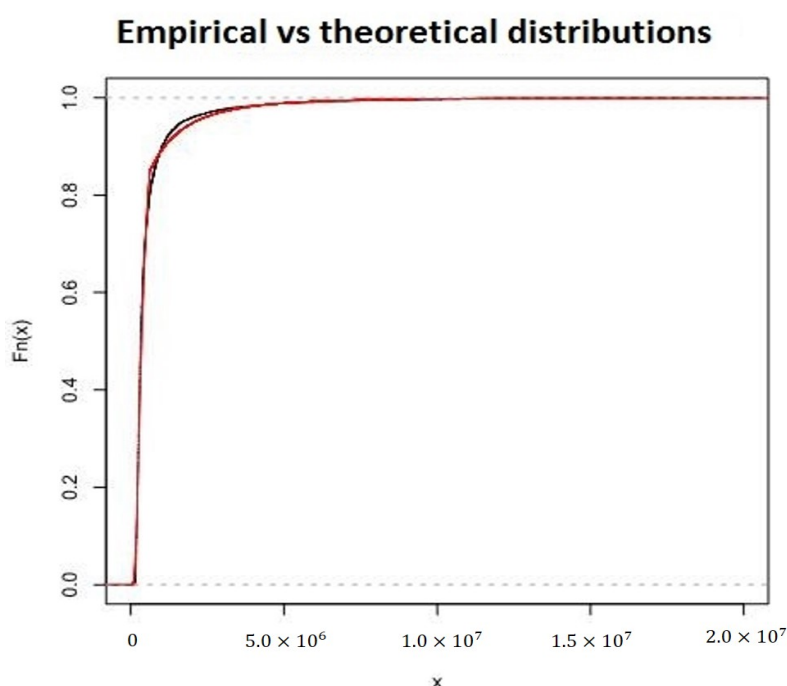
Empirical vs theoretical distributions

Figure 9. Lognormal–GPD–Gamma (red) versus empirical (dark) distribution functions.

Table 9. AIC and BIC indices for a comparison between different models.

Index	KumPareto	Burr	GPD	GPD–Gamma
AIC	193,374	191,459	191,172	190,834
BIC	193,409	191,494	191,207	190,882

The previous analysis suggests that the Lognormal–GPD–Gamma gives the best fit.

5. Introducing Dependence Structure: Copula Functions and Fast Fourier Transform

In the previous section we restricted our analysis to the case of independence between attritional and large claims. We now try to extend this work to a dependence structure. Firstly, we defined a composite model using a copula function to evaluate the possible dependence. As marginal distributions, we referenced to a Lognormal distribution for attritional claims and a GPD for large ones. The empirical correlation matrix **R**

$$\mathbf{R} = \begin{pmatrix} 1 & 0.01259155 \\ 0.01259155 & 1 \end{pmatrix}$$

and Kendall's Tau and Spearman's Rho measures of association

$$\mathbf{K} = \begin{pmatrix} 1 & 0.00252667 \\ 0.00252667 & 1 \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 1 & 0.00373077 \\ 0.00373077 & 1 \end{pmatrix}$$

suggest a weak but positive correlation between normal and large claims.

For this reason, the individuation of an appropriate copula function will not be easy, but we present an illustrative example based on a Gumbel Copula. We underline that an empirical dependence

structure is inducted by distinction between attritional and large losses. In fact, there is no unique event that causes small and large losses simultaneously, but when an insured event occurs, only an attritional or large loss is produced. For this reason, the results showed in the following should be considered as an exercise that highlights the important effects of dependence on the aggregate loss distribution.

$$C_\theta(u, v) = \exp -[\ln(u)^{-\theta} + \ln(v)^{-\theta}]^{1/\theta}, 1 \le \theta < \infty \qquad (20)$$

Table 10 reports the different methods to estimate the parameter θ:

Table 10. Different methods for estimating the dependence parameter θ of a Gumbel Copula.

Method	θ	Standard Error
Maximum pseudo-likelihood	1.11	0.008
Canonical maximum pseudo-likelihood	1.11	0.008
Simulated maximum likelihood	1.11	-
Minimum distance	1.09	-
Moments based on Kendall's tau	1.13	-
Bootstrap	1.11	0.008

We remind that Gumbel's parameter θ assumes values in $[1, \infty)$, and for $\theta \to 1$ we have independence between marginal distributions. We observe that estimates were significantly different from 1, and so our Gumbel Copula did not correspond to the Independent Copula. We can say that because we verified using bootstrap procedures, the θ parameter has a Normal distribution. In fact, the Shapiro–Wilk test gave a p-value equal to 0.08551; thus, with a fixed significance level of 5%, it is not possible reject the null hypothesis. In addition, the 99% confidence interval obtained with Maximum pseudo-likelihood method was $(1.090662; 1.131003)$, which does not include the value 1; the same confidence interval obtained with the Bootstrap procedure was $(1.090662; 1.131003)$. In the following Figure 10 we report the distribution of the Gumbel parameter obtained by the bootstrap procedure.
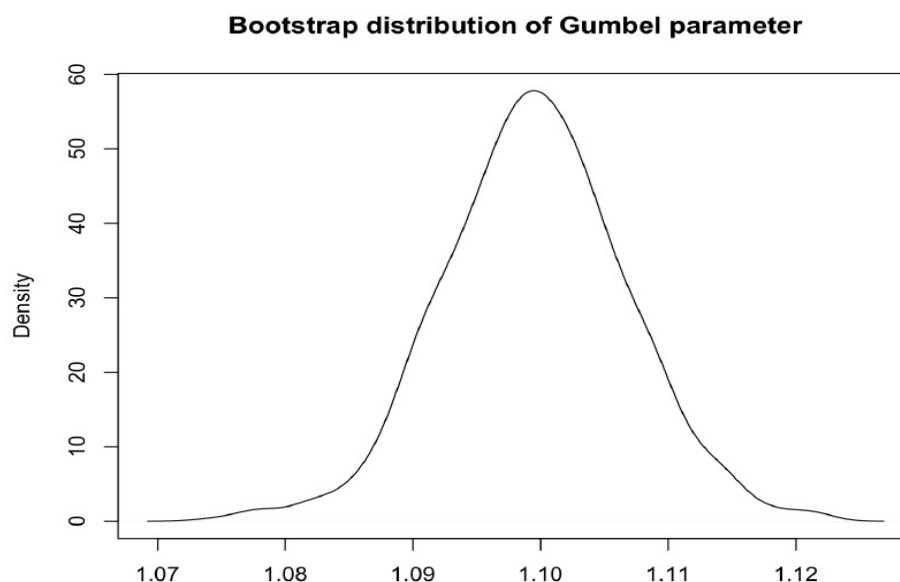

Figure 10. Normal distribution of the Bootstrap Gumbel parameter.

We report two useful graphics (Figures 11 and 12), obtained by simulation of the estimated Gumbel.
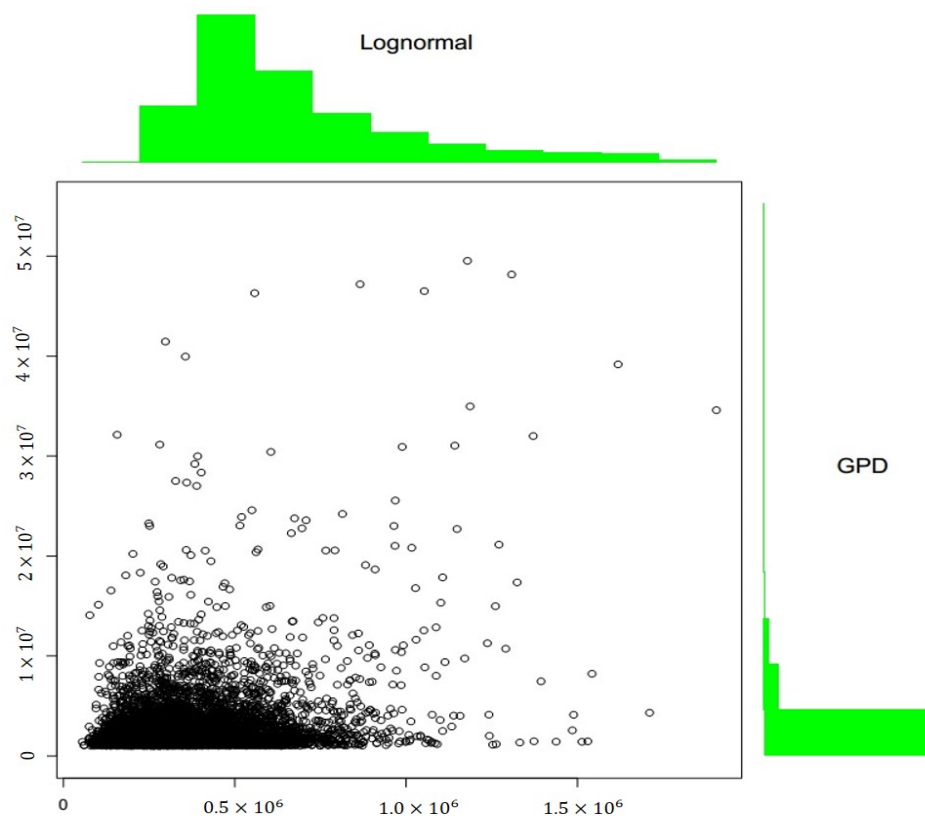
Figure 11. Lognormal (top) and GPD (right) marginal histograms and Gumbel Copula simulated values plot.
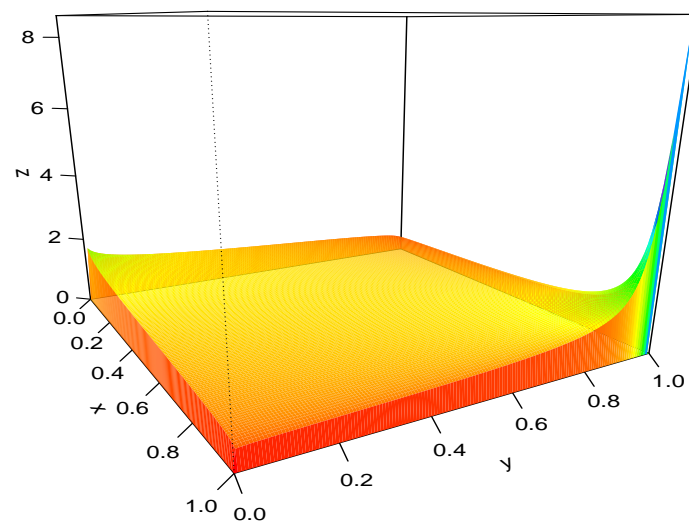

Figure 12. Density function of the estimated Gumbel Copula. Attritional claim losses on the X-axis, large claim losses on the Y-axis.

The density function (Figure 12) assumed greater values in correspondence of great values both for Lognormal and GPD marginal; in other words, using the Gumbel Copula, the probability that attritional claims produced losses near to the threshold u, and that large claims produced extreme losses, was greater than the probability of any other joined event.

We report also the result of the parametric bootstrap goodness-of-fit test performed on the estimated Gumbel Copula.

Statistic	θ	p-Value
2.9381	1.1108	0.00495

We can consider the estimated Gumbel a good approximation of dependence between data. In our numerical examples, we referred to the Gumbel Copula function despite having estimated and analyzed other copulas for which there was no significant difference for the aims of this paper. While the empirical dependence is not excessive, we will see how the introduction in the estimation model of a factor that takes it into account, such as a Copula function, will produce a non-negligible impact on the estimate of the VaR.

An Alternative to the Copula Function: The Fast Fourier Transform

Considering the fact that it is not easy to define an appropriate copula for this dataset, we next modeled the aggregate loss distribution directly with the Fast Fourier Transform (FFT) using empirical data. That approach allowed us to avoid the dependence assumption between attritional and large claims (necessary instead with the copula approach).

To build an aggregate loss distribution by FFT, it is first necessary to discretize the severity distribution Z (see Klugman et al. 2010) and obtain the vector $z = (z_0, \ldots, z_{n-1})$, of which element z_i is the probability that a single claim produces a loss equal to ic, where c is a fixed constant such that, given n length of vector z, the loss cn has a negligible probability. We considered also frequency claim distribution $\tilde{k}$ through Probability-Generating function (PGF) defined as

$$PGF_{\tilde{k}}(t) = \sum_{j=0}^{\infty} t^j Pr(\tilde{k} = j) = E[t^k] \tag{21}$$

In particular, let $FFT(z)$ and $IFFT(z)$ be the FFT and its inverse, respectively. We obtain the discretized probability distribution for the aggregate loss X as

$$(x_0, x_1, \ldots, x_{n-1}) = IFFT(PGF(FFT(\mathbf{z}))) \tag{22}$$

Both $FFT(z)$ and $IFFT(z)$ are n-dimensional vectors whose generic elements are, respectively,

$$\hat{z}_k = \sum_{j=0}^{n-1} z_j \exp(\frac{2\pi i}{n} jk) \tag{23}$$

$$z_k = \frac{1}{n} \sum_{j=0}^{n-1} \hat{z}_j \exp(-\frac{2\pi i}{n} jk) \tag{24}$$

where $i = \sqrt{-1}$.

From a theoretical point of view, this is a discretized version of Fourier Transform (DFT):

$$\phi(z) = \int_{-\infty}^{+\infty} f(x) \exp(izx) dx \tag{25}$$

The characteristic function created an association between a probability density function and continuous complex one, while the DFT made an association between an n-dimensional vector and an n-dimensional complex vector. The former one-to-one association can be done through the FFT algorithm.

For a two-dimensional case, matrix M_Z is a necessary input; this matrix contains joined probabilities of attritional and large claims such that it is possible to obtain corresponding marginal distributions by adding long rows and columns respectively. For example, let

$$\mathbf{M_z} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0.2 & 0.25 & 0 \\ 0 & 0.05 & 0 \end{pmatrix}$$

be that matrix. The vector $(0.5, 0.45, 0.05)$, obtained by adding long three rows, contains the marginal distribution of attritional claims, while the vector $(0.7, 0.3, 0)$, obtained by adding long three columns, contains the marginal distribution of large claims. The single element of the matrix, instead, is the joined probability. The aggregate loss distribution will be a matrix M_X given by

$$\mathbf{M_x} = IFFT(PGF(FFT(\mathbf{M_z})))$$ (26)

For more mathematical details, we point to Robe-Voinea and Vernic (2016) and Robe-Voinea and Vernic (2017), in which FFT is extended to a multivariate setting, and several numerical examples are illustrated.

We decided to discretize the observed distribution function, without reference to a specific theoretical distribution, using the discretize R function available in the actuar package (see Klugman et al. 2010). This discretization allows us to build the matrix M_Z to which we applied the two-dimensional FFT version. In this way, we obtained a new matrix $FFT(M_Z)$ that acted as input for the random $\tilde{k}$ probability generating function.

As reported in Section 2, in our dataset, 50% of frequencies were included between 238 and 381 claims, and there was a slightly negative asymmetry. In addition, the variance was greater than the mean value (299). Thus, it is possible to suppose a Negative Binomial distribution for frequency claims. The corresponding probability generating function is defined by

$$PGF(t) = \left(\frac{1 - p}{1 - pt} \right)^m$$ (27)

We estimated its parameters that resulted $m = 5$ and $p = 0.82$. Then, we obtained the matrix $PGF_{\tilde{k}}(FFT(\mathbf{M_z}))$. As the last stage we applied the IFFT whose output is matrix M_X. Adding long counter-diagonals of M_X we can individuate the discretized probability distribution of aggregate loss claims, having maintained the distinction between normal and large claims and, above all, preserving the dependence structure.

6. Final Results and Discussion

As shown previously, from the perspective of strict adaptation to empirical data, we can say that the best model to fit the Danish fire data is the Lognormal–GPD–Gamma one, which presented a coefficient of variation equal to $10, 2\%$, lesser than Standard Formula volatility. In fact, considering the premium risk and Fire segment only, the volatility of the Standard Formula was equal to 24% (3 times $\sigma_{(prem,fire)}$, where $\sigma_{(prem,fire)} = 8\%$; see Tripodi 2018). As written in the introduction of the present work, this result mainly was due to the fact that the DA multiplier did not take into account the skewness of the aggregate claim distribution, and it potentially overestimated the SCR for large insurers.

For illustrative purposes only, we estimated the VaR_p and the volatility of aggregate loss using the previous models, taking into account a dependence structure as well. According to the collective approach of risk theory, aggregate loss is the sum of a random number of random variables, and so it requires convolution or simulation methods. We remember that among the considered methodologies, only FFT directly returned the aggregate loss. Relating to FFT, as we mentioned above, an empirical dependence structure was inducted by discriminating between attritional and large losses, so we referred to empirical discretized severities in a bivariate mode. This is a limitation of our work that could be exceeded considering a bivariate frequency and two univariate severities, inducting such dependence by the frequency component, as it happens in practice (i.e., dependency between severities is not typical for this line of business); however, this approach would not have allowed us to apply the FFT methodology.

Considering the statistics of frequency in the Danish fire insurance data, we can assume claim frequency k distributed as a Negative Binomial, as done previously with the FFT procedure. A single simulation of aggregate loss can be achieved by adding the losses of k single claims, and by repeating the procedure n times, we obtained the aggregate loss distribution.

Contrary to the copula approach, we point out that it would be possible to obviate the need to simulate by applying FFT to generate aggregates from the fit severities.

In Table 11, we report the VaRs obtained using composite models Lognormal–GPD–Gamma, Gumbel Copula and FFT, and the corresponding coefficients of variation that give us indications of the applied models' volatilities:

Table 11. Estimates of VaR at the 99.5% level with different models and their volatilities.

Model	VaR (€)	CV
Lognormal–GPD–Gamma	216,913,143	0.102
Gumbel Copula	664,494,868	0.110
FFT	703,601,564	0.193

If we consider the independence assumption, the aggregate loss distribution will return a VaR significantly smaller (over −200%) than that calculated using the dependence hypothesis. The assumption of independence, or not, would therefore produce obvious repercussions on the definition of the risk profile and, consequently, on the calculation of the capital requirement. As seen above, for the case analyzed, the Gumbel Copula took into account the positive dependence, even if of discrete magnitude, between the tails of the marginal distributions of the severities. That is, an attritional loss close to the discriminatory threshold is, with good probability, accompanied by an extreme loss. This can only induce a decisive increase in the VaR of the aggregate distribution, as can be seen from Table 11. In the same way as Fast Fourier Transform, taking into account not only the (empirical) dependence between claims but also the randomness of frequency claims also induces a further increase in the risk estimate.

Therefore, it is fundamental to take into account the possible dependence between claims, regarding its shape and intensity, because the VaR could increase drastically with respect to the independence case, leading to an insolvent position of the insurer.This analysis highlights the inadequacy of using CV when the actual objective is to estimate VaR.

However, all previous approaches have advantages and disadvantages. With the composite models we can robustly fit each of the two underlying distributions of attritional and large claims, without a clear identification of the dependency structure. With the Copula we can model dependency, but it is not easy to determine what is the right copula to use, and this is the typical issue companies have to face for capital modeling purposes using a copula approach. FFT allows one to not simulate the claim process and to not estimate a threshold, working directly on empirical data, but includes some implicit bias due to the discretization methods; for example, since the FFT works with truncated distributions, it can generate aliasing errors. We point again to Robe-Voinea and Vernic (2016) and Robe-Voinea and Vernic (2017) for a detailed discussion and the possible solutions insurers have to consider when implementing PIM.

Finally, compared to the present literature, we remark that this paper innovates the fitting of the Danish fire insurance data, using a composite model with a random threshold. Secondly, our empirical model could have managerial implications, supporting insurance companies in understanding that the Standard Formula could lead to a volatility (and the SCR) of the premium risk that is very different from the real risk profile. It is worth mentioning CEIOPS (2009), in that "Premium risk also arises because of uncertainties prior to issue of policies during the time horizon. These uncertainties include the premium rates that will be charged, the precise terms and conditions of the policies and the precise mix and volume of business to be written. Various studies (e.g., Mildenhall (2017) Figure 10) have shown that pricing risk results in a substantial increase in loss-volatility, especially for commercial lines". Therefore, one would expect that the SCR premium charge would look high compared to a test that only considers loss (frequency and severity) uncertainty. In next developments of this research we will try to take into account these features in order to have a full picture of this comparison.

References

Ausin, M. Concepcion, Michael P. Wiper, and Rosa E. Lillo. 2009. Bayesian estimation of finite time ruin probabilities. *Applied Stochastic Models in Business and Industry* 25: 787–805. [CrossRef]

Burnecki, Krzysztof, and Rafal Weron. 2004. Modeling the risk process in the XploRe computing environment. *Lecture Notes in Computer Science* 3039: 868–75.

Carreau, Julie, and Yoshua Bengio. 2009. A hybrid pareto model for asymmetric fat-tailed distribution. *Extremes* 12: 53–76. [CrossRef]

CEIOPS. 2009. Advice for Level 2 Implementing Measures on Solvency II: SCR Standard Formula—Article 111-Non-Life Underwriting Risk. Availble onlilne: https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Final-Advice-SCR-Non-Life-Underwriting-Risk.pdf (accessed on 31 October 2009).

Cerchiara, Rocco R., and Valentina Demarco. 2016. Undertaking specific parameters under solvency II: reduction of capital requirement or not? *European Actuarial Journal* 6: 351–76. [CrossRef]

Charpentier, Arthur, and Abder Oulidi. 2010. Beta kernel quantile estimators of heavy-tailed loss distributions. *Statistics and Computing* 20: 35–55. [CrossRef]

Clemente, Gian Paolo, and Nino Savelli. 2017. Actuarial Improvements of Standard Formula for Non-life Underwriting Risk. In *Insurance Regulation in the European Union*. Edited by Pierpaolo Marano and Michele Siri. Cham: Palgrave Macmillan.

Cooray, Kahadawala, and Malwane M.A. Ananda. 2005. Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal* 5: 321–34. [CrossRef]

Cordeiro, Gauss M., Saralees Nadarajah and Edwin M. M. Ortega. 2012. The Kumaraswamy Gumbel distribution. *Statistical Methods & Applications* 21: 139–68.

Drees, Holger, and Peter Müller. 2008. Fitting and validation of a bivariate model for large claims. *Insurance: Mathematics and Economics* 42: 638–50. [CrossRef]

EIOPA. 2011. *Report 11/163. 2011. Calibration of the Premium and Reserve Risk Factors in the Standard Formula of Solvency II-Report of the Joint Working Group on Non-Life and Health NSLT Calibration*. Frankfurt: EIOPA .

Embrechts, Paul, Claudia Klüppelberg, and Thomas Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance*. Berlin/Heidelberg: Springer.

Esmaeili, Habib, and Claudia Klüppelberg. 2010. Parameter estimation of a bivariate compound Poisson process. *Insurance: Mathematics and Economics* 47: 224–33. [CrossRef]

European Commission. 2015. Commission Delegated Regulation (EU) 2015/35 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II). *Official Journal of the EU* 58: 1–797.

Galeotti, Marcello. 2015. Computing the distribution of the sum of dependent random variables via overlapping hypercubes. *Decisions in Economics and Finance* 38: 231–55. [CrossRef]

Guillotte, Simon, Francois Perron, and Johan Segers. 2011. Non-parametric Bayesian inference on bivariate extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73: 377–406. [CrossRef]

Klugman, Stuart, Harry H. Panjer, and Gordon E. Wilmot. 2010. *Loss Models: From Data to Decisions*. Hoboken: John Wiley & Sons.

Kumaraswamy, Poondi. 1980. A generalized probability density function for double-bounded random processes. *Journal of Hydrology* 46: 79–88. [CrossRef]

McNeil, Alexander J. 1997. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA* 27: 117–37. [CrossRef]

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2005. *Quantitative Risk Management. Concepts, Techniques and Tools*. Princeton: Princeton University Press.

Mildenhall, Sthepen J. 2017. Actuarial geometry. *Risks* 5: 31. [CrossRef]

Nadarajah, Saralees, and Anuar A.S. Bakar. 2014. New composite models for the Danish-fire data. *Scandinavian Actuarial Journal* 2: 180–87. [CrossRef]

Pigeon, Mathieu, and Michel Denuit. 2011. Composite Lognormal-Pareto model with random threshold. *Scandinavian Actuarial Journal* 3: 177–92. [CrossRef]

Resnick, Sidney I. 1997. Discussion of the Danish data on large fire insurance losses. *ASTIN Bulletin: The Journal of the IAA* 27: 139–51. [CrossRef]

Robe-Voinea, Elena-Gratiela, and Raluca Vernic. 2016. Fast Fourier Transform for multivariate aggregate claims. *Computational & Applied Mathematics* 37: 205–19. [CrossRef]

Robe-Voinea, Elena-Gratiela, and Raluca Vernic. 2017. On a multivariate aggregate claims model with multivariate Poisson counting distribution. *Proceedings of the Romanian Academy series A: Mathematics, Physics, Technical Sciences, Information Science* 18: 3–7.

Scollnik, David P. M. 2007. On composite lognormal-Pareto models. *Scandinavian Actuarial Journal* 1: 20–33. [CrossRef]

Teodorescu, Sandra, and Raluca Vernic. 2009. Some composite Exponential-Pareto models for actuarial prediction. *Romanian Journal of Economic Forecasting* 12: 82–100.

Teodorescu, Sandra, and Raluca Vernic. 2013. On Composite Pareto Models. *Mathematical Reports* 15: 11–29.

Tripodi, Agostino. 2018. Proceedings of the Seminar *Non-Life Premium and Reserve Risk: Standard Formula with Undertaking Specific Parameter* at Department of Economics, Statistics and Finance, University of Calabria, 5 November 2018. Available online: https://www.unical.it/portale/portaltemplates/view/view.cfm?84609 (accessed on 6 July 2019).