Sources of uncertainty in modeling operational risk losses

Giulio Mignola

Operational Risk Management, Sanpaolo IMI, Via Monte di Pietà 26, 10122 Torino, Italy

Roberto Ugoccioni

Operational Risk Management, Sanpaolo IMI, Via Monte di Pietà 26, 10122 Torino, Italy

Operational risk quantification techniques have been rapidly evolving since the first attempts in early 2000, when it appeared clear that this kind of risk would attract a specific capital requirement in the new prudential regulation. The basic component of most models developed by the industry is historical (or scenario) loss data. The modeling techniques used to obtain the risk measures are generally well developed and understood. In this work, assuming a simple but rigorous modeling framework, containing the basic features of the models which are generally observed in the industry, focus will be placed on the uncertainty of the model outputs. The sources of this uncertainty will be analyzed in a systematic way.

Introduction

Operational risk quantification techniques have been rapidly evolving since the first attempts in early 2000,[1] when it appeared clear that this kind of risk would attract a specific capital requirement in the new prudential regulation (Basel Committee 2004). The basic component of most models developed by the industry is historical (or scenario) loss data. The common paradigm used to define the risk measure is the distribution of annual aggregated losses.

To obtain the risk measures, the modeling techniques used are generally well developed and understood (see, for instance, Embrechts *et al* (2005), ch. 10 and references therein), although in practice it appears that there are as many ways to implement models as there are banks. Many models refer mostly to historical losses and others are based primarily on forecast potential losses (scenarios).

In this work, assuming a simple but rigorous modeling framework, containing the basic features of the models which are generally observed in the industry, focus will be placed on the uncertainty of the model outputs as opposed to its actual value. The sources of this uncertainty will be analyzed in a systematic way.

The views expressed in this article are those of the authors and do not necessarily represent the views of the Bank. All numerical computations described in this paper have been carried out using the R system (R Development Core Team 2006). The first author wishes to thank Ms Sarah Doring for the fruitful and challenging discussions that led to the final form of this work.

[1]Pioneering work was attempted by Bankers Trust and PricewaterhouseCoopers (New York) in the second half of the 1990s.

Section 1 describes the characteristics of the model implemented, Section 2 the structural uncertainty associated with the choice of a specific model, Section 3 the effect of the statistical errors on the results and Section 4 the controllable sources of errors connected with the actual implementation of the calculation. The problem of the aggregation of risk measures and how additional sources of uncertainty are embedded in this will be briefly touched upon.

The results which are obtained show very clearly how the requirement to express the risk measure as a very high percentile of the annual aggregated loss distribution strongly amplifies the effect of uncertainty, consequently magnifying the standard error associated with the estimates.

1 Description of the model

Models used in the industry to analyze operational risk losses, although different in the implementation, are all commonly based on the same essential methodology which will be used in its simplest form in this paper. The specific formulas used are detailed in the appendix.

The risk measure will be represented by a defined percentile p (eg, 99.9%) of the annual aggregated loss distribution $S(x)$.

The distribution $S(x)$ is obtained by studying separately the distribution of the loss frequency p_n and the loss severity $f(x)$. These two distributions are deemed to be independent and stable over time.

The aggregated annual loss distribution density can be formally obtained as the weighted average of the n-fold convolution of the severity where the weights are the frequency mass probabilities.

The n-fold convolution of the severity distribution is the probability of observing the aggregate of n individual losses.

The frequency distribution is assumed to be a Poisson distribution with the specific value of the intensity λ to be determined from data.

For the severity distributions, the families generally used in operational risk modeling will be analyzed, namely:

- the log-normal distribution;
- the Weibull distribution;
- the log-normal-gamma distribution (also known as "fat-tailed log-normal");
- the Burr distribution;
- the generalized Pareto distribution (GPD).

The specific values of the parameters of each distribution are determined by the data applying the maximum likelihood estimator. To simplify the scheme, data is assumed to be complete and all distributions to be "ground-up" (ie, threshold at zero); taking into account possible truncation in the data will add sources of uncertainty and hence strengthen the validity of the results obtained in this paper.

To choose among the various distributions the one that best describes the observed data, the following goodness-of-fit (GOF) statistics (see, for instance, D'Agostino and Stephens (1986), Chernobai *et al* (2005) and Schwarz (1978)) are used:

- the Kolmogorov–Smirnov statistics;
- the right-tailed Anderson–Darling statistics;
- the Smirnov–Cramér–Von Mises statistics.

To obtain the aggregated annual distribution, various available techniques are used:

- a Monte Carlo simulation (Klugman *et al* 2004, ch. 17);
- a direct numerical calculation, based on properties of the characteristic and generating functions, implemented with Fast Fourier Transforms (FFT) (Klugman *et al* 2004, ch. 6);
- the analytical approximation called "the single-loss approximation". This approximation, recently highlighted in Böcker and Klüppelberg (2005), gives a direct formula for the quantile of the aggregated annual distribution as

  $$S^{-1}(p) = F^{-1}\left(1 - \frac{1-p}{\mathrm{E}[n]}\right)$$

  directly linking the percentile p of S to a higher percentile $q = 1 - (1-p)/\mathrm{E}[n]$ of the single loss (severity) distribution F dependent on the expected number of yearly events $\mathrm{E}[n]$.

  This result is based on a simple and well known analytical property of distributions belonging to the class of sub-exponential distributions[2] (Asmussen 2000), which allows us to express, in the limit of high severity, the n-fold convolution as a function of the severity distribution of the single loss[3]

  $$\overline{F^{*n}}(x) \simeq n\overline{F}(x) \quad \text{as } x \to \infty \tag{1}$$

2 Uncertainty in the underlying model

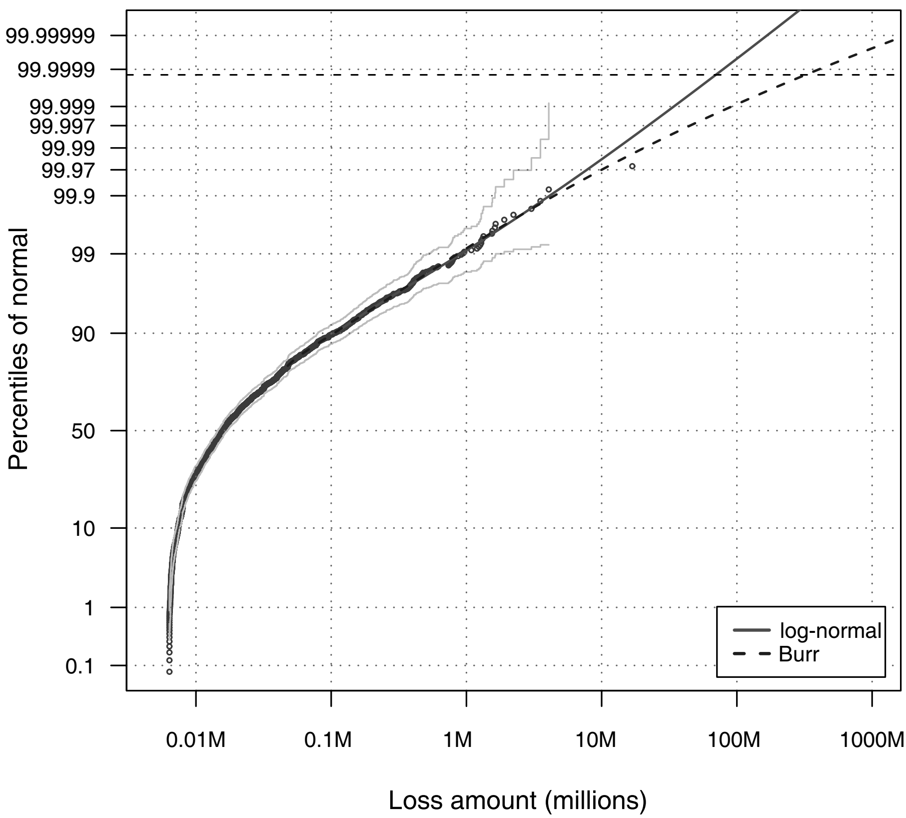
Models that describe the behavior of loss frequency and loss severity in a specific risk class should, in theory, be derived from a detailed modeling of the structure of the firm itself (eg, processes, systems, etc) and of the environment in which the firm operates. In the absence of such a model that specifies the statistical distribution to be used, the only other available option is to derive the behavior of the distribution from data samples. This means selecting the distribution that best fits the loss data actually collected. To choose this model some specific statistical tools are used, such as GOF techniques.

Unfortunately, relying on the properties of the available data alone could lead to a situation in which decisions relating to which is the best fit cannot be made.

[2]All the distributions used in this paper belong to this class.

[3]Where $\overline{F}(x) = 1 - F(x)$ is the tail distribution.

FIGURE 1 Example of divergence for high quantiles: the two distributions shown in this qq-plot fit the data equally well (p-values of GOF tests are given in the table) but differ considerably in the extrapolation to reach the high percentile called for by regulations (dashed horizontal line).



Fit	KS	ADrt	SCvM
log-normal	0.93	0.88	0.88
Burr	0.90	0.91	0.87

An example is illustrated in Figure 1 in which loss data[4] is fitted, at the same level of statistical confidence, by two different distributions. The decision here on which distribution to use has to be taken from the joint analysis of three different goodness-of-fit tests (Anderson–Darling, Smirnov–Cramér–Von Mises and Kolmogorov–Smirnov statistics). If the fits are equally good, in statistical terms, so must be the results. Indeed, the two fitted models forecast similar

[4]In this example, the data used is represented by a real case where the loss amounts have been rescaled to minimise disclosure.

TABLE 1 Some quantiles of the severity distributions in Figure 1. Notice how as long as one remains within the percentile range covered by data (ie, up to $p \simeq 0.999$), the two distributions do not differ significantly; the last row in part (b) corresponds to the single-loss approximation of the 99.9% quantile of the aggregate loss at 700 events/year, where the two distributions diverge considerably.

	p	log-normal	Burr
(a)	0.90	105,000	105,000
	0.925	144,000	145,000
	0.95	217,000	219,000
	0.98	515,000	498,000
	0.99	911,000	865,000
	0.995	1,514,000	1,450,000
	0.999	4,120,000	4,463,000
(b)	0.9999	13,084,000	20,556,000
	0.99999	34,193,000	91,635,000
	0.9999986	69,780,000	321,475,000

values for specific quantiles inside the range of available data, as can be seen in Table 1(a).

However, for the *aggregate* quantile at the 99.9% confidence level one obtains results that differ by more than a factor of 3 (120 million versus 380 million). Why is this possible? Because the forecasted percentile at the 99.9% confidence level is determined by percentiles of the single loss severity distribution far exceeding the 99.9%.

This can be easily explained using the single-loss approximation, which demonstrates directly that the annual aggregated loss at a given percentile p is linked to the loss severity distribution at a much higher percentile $q = 1 - (1 - p)/\mathrm{E}[n]$, where $\mathrm{E}[n]$ is the yearly average number of events. Extending the axes in Figure 1 to include q, shows how such large differences could be generated: indeed, in Table 1(b) it is shown that the more one extrapolates to high percentiles, the more the corresponding quantiles diverge.

This is what this paper proposes to call "the extrapolation issue" since to reach the q level required with the observed data, a number of observations of order $\mathrm{E}[n]/(1 - p)$ are needed. Since there is no practical way that such an amount of data could be made available, it appears that no practical solution to this problem exists, except reducing the p confidence level requested to a much lower value.

This source of uncertainty, that comes on the one hand from the absence of a dynamic[5] model for the operational risk losses (the true distribution) and on the other hand from the high percentile requested for the measure of risk at the annual

[5]This issue is very well described in the introduction of Hoeting *et al* (1999), although in a different technical environment.

aggregate level (the extrapolation issue), constitutes by far the major source of potential error and is structural to the approach.

A possible way out could be offered by the application of a limit theory such as the Extreme Value Theory (EVT) that asymptotically identifies a unique distributional form for the loss severity hence removing *ab initio* the problem of the choice of the best suited distribution to describe the data. In practice, however, one has to be concerned that the validity of the asymptotic regime, which is granted only in the limit of infinite severity, holds at the level at which data is available. Mignola and Ugoccioni (2005) showed that instability of results are common when the typical distributions of operational risk losses are taken into account. Furthermore, the "choice of a best fit" (the very problem dealt with in this section) could be translated in the EVT approach as a "choice of threshold", which is also a source of uncertainty (see Embrechts *et al* (2005, p. 285)). While this translation could at times ease the burden, it does not solve the problem of extrapolating to high quantiles (see Embrechts *et al* (1997, p. 364)).

3 Statistical uncertainty

Statistical errors come from the finiteness of the data sample available: clearly such errors will decrease by increasing the sample size (the rate of decrease depends on the underlying distribution). This section assumes that the underlying distribution is known but its parameters are unknown: the estimation of parameters from the data produces the statistical errors whose size is studied here.

While a thorough exploration of the likelihood profile gives the correct propagation of estimation errors to confidence bands for the quantile, this method is in practice unusable due to time constraints. In the following, approximate methods to determine the correct order of magnitude of the confidence bands will be used, namely truncated Taylor expansion (Eadie *et al* 1971, p. 26) (commonly known simply as "error propagation") and simulation.

3.1 Error propagation

3.1.1 Formal details

Let α_i $(i = 1, \dots, n)$ be the parameters estimated from the data; they include parameters of both severity and frequency, although the effect of the latter is usually negligible. Let V_{ij} be the element of the correlation matrix associated with parameters α_i and α_j, computed during the fit, eg, by likelihood profile analysis. Formally then, the aggregate loss $S(x)$ can be seen, at fixed x, as a multidimensional function of the α_i, as can the inverse $S^{-1}(p)$, which is the main concern here. Thus one can apply the well known "error propagation" theory and compute an estimate of the uncertainty Δq in the quantile estimate q due to V (see Eadie (1971, p. 26)):

$$\Delta q \approx \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial S^{-1}(p)}{\partial \alpha_i} \frac{\partial S^{-1}(p)}{\partial \alpha_j} V_{ij}} \tag{2}$$

One should remember that the above approximation can break down when $S^{-1}(p)$ is significantly nonlinear in α_i in a region of size comparable to the standard deviation of α_i.

In the simplest case of non-correlated errors, the matrix V is a diagonal matrix whose diagonal elements are the variances of the parameters, ie, what one commonly takes as the square of the standard error: $V_{ii} = (\Delta\alpha_i)^2$. Equation (2) then takes on the more familiar look:

$$\Delta q \approx \sqrt{\sum_{i=1}^{n}\left(\frac{\partial S^{-1}(p)}{\partial \alpha_i}\Delta\alpha_i\right)^2} \tag{3}$$

The partial derivatives have to be computed numerically, at least when Monte Carlo or FFT techniques are used in the calculation; only when using single-loss approximation can one compute derivatives analytically for some simple distributions.

As an example, a GPD with uncorrelated fitted parameters shape ξ and scale σ can be considered. Notice that this particular example and hence the results presented below also hold for the case in which one applies the EVT peak-over-threshold (POT) scheme to the analysis of data, being the GPD the asymptotic distribution has to use. Then

$$q = S^{-1}\left(1 - \frac{1-p}{\lambda}\right) = \frac{\sigma}{\xi}\left[\left(\frac{\lambda}{1-p}\right)^{\xi} - 1\right] \tag{4}$$

hence

$$\begin{aligned}(\Delta q)^2 = {} & \left(q\frac{\Delta\sigma}{\sigma}\right)^2 + \left[\sigma\left(\frac{\lambda}{1-p}\right)^{\xi}\frac{\Delta\lambda}{\lambda}\right]^2 \\ & + \left\{\left[-q + \sigma\left(\frac{\lambda}{1-p}\right)^{\xi}\log\left(\frac{\lambda}{1-p}\right)\right]\frac{\Delta\xi}{\xi}\right\}^2\end{aligned} \tag{5}$$

3.1.2 Results and comments

A typical case for a medium size bank for a risk category related to external frauds could be represented by a GPD fitted to 100 data points (see Table 2 for details). Applying Equation (5) one obtains a relative error $\Delta q/q \approx 110\%$; the same results are obtained using numerical derivatives with FFT techniques.

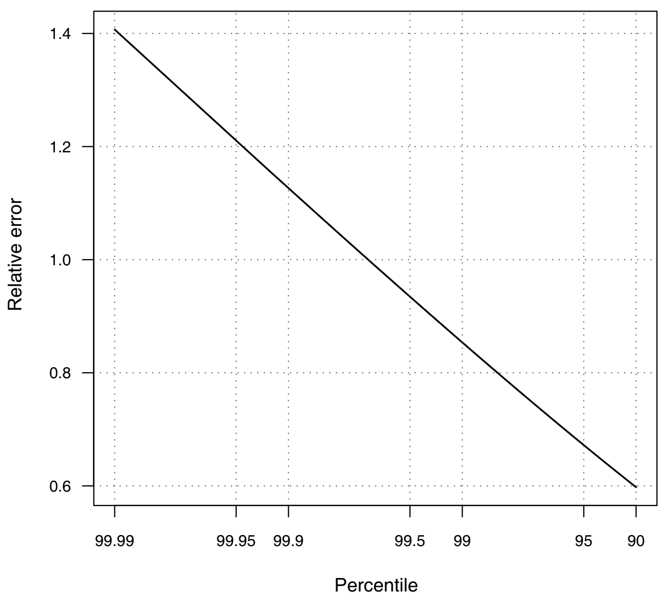
As another example, consider a risk category with again 100 data points which can be fitted by a log-normal distribution (see Table 2): this time one obtains a relative error on the quantile of "only" 60%.

Finally, consider an example where the Burr distribution is fitted to a large sample (here 4,000 data points, see Table 2): note that the Burr distribution behaves very similar to a GPD with $\xi \simeq 1/(\alpha\gamma) \approx 0.9$ here, making this tail much heavier than the GPD considered in the first example. In this case, even with small errors on the parameters, the relative error on the quantile is still at a 75% level.

TABLE 2 Three real-life examples in which $\Delta q/q$, the relative error in 99.9% quantile estimation using single-loss approximation, has been computed via Equation (3). Relative errors on fitted parameters are given in corresponding order, together with the number of fitted points.

Distribution (parameters)	# points	Errors	$\Delta q/q$
GPD ($\xi = 0.5$, $\sigma = 250{,}000$, $\lambda = 50$)	100	0.25, 0.2, 0.01	110%
log-normal ($\mu = 11$, $\sigma = 2$, $\lambda = 50$)	100	0.01, 0.07, 0.01	60%
Burr ($\alpha = 1.03$, $\gamma = 1.06$, $\theta = 50{,}000$, $\lambda = 1{,}000$)	4,000	0.06, 0.02, 0.09, 0.01	75%

FIGURE 2 Relative error in the pth quantile of the aggregate loss as a function of p, computed via Equation (5) in the case of a GPD severity with parameters as in Table 2.



Going back to the GPD example, Figure 2 shows the behavior of the relative error given by Equation (5) as a function of the percentile p of the aggregate loss that one is trying to estimate, with all parameters fixed. Notice once more the quick growth as p increases.

The error propagation formula does not provide an immediate answer to the question of dependency on the number of data points, apart from the general statement (Eadie *et al* 1971) that $\Delta\alpha_i$, and hence Δq, goes, asymptotically in the sample size, as the inverse of the square root of the sample size itself: to show how this translates for finite samples, simulation has to be used.

3.2 Simulation

3.2.1 Formal details

The following distributions are used:

(a) log-normal with $\mu = 10$, $\sigma = 2$;
(b) Burr with $\alpha = 1.1$, $\gamma = 1.1$, $\theta = 20{,}000$;
(c) generalized Pareto with shape $\xi = 0.3$, scale $\sigma = 10^6$;
(d) generalized Pareto with shape $\xi = 0.8$, scale $\sigma = 10^6$;
(e) log-normal-gamma with $\mu = 10$, $\sigma = 2$ and $\alpha = 4$;
(f) log-normal-gamma with $\mu = 10$, $\sigma = 2$ and $\alpha = 12$.

Notice that all the distributions are heavy-tailed (sub-exponential) but only (a), (e) and (f) have all moments finite; distribution (c) has finite mean, variance and skewness, while for (b) and (d) only the mean is finite. All the above distributions are representative of real-life cases for the various risk classes encountered in operational risk modeling.

For each of the above-listed distributions, and for each sample size N in this list, ie, 50, 100, 500, 1,000, 5,000, one random sample is generated, the parameters of the corresponding distribution are fitted and the 99.9% quantile of the aggregate loss distribution (assuming Poisson frequency with $\lambda = 10$) computed using FFT techniques.

The procedure is repeated $M = 1{,}000$ times ($M = 500$ times in the log-normal-gamma cases), and bias $b(\hat{q})$ and root mean square error $\mathrm{RMSE}(\hat{q})$ (see McNeil and Saladin (1997) and Embrechts *et al* (2003)) are computed for each distribution/sample size combination:

$$b(\hat{q}) \equiv \mathrm{E}[\hat{q} - q_0] = \frac{1}{M}\sum_i (\hat{q}_i - q_0) \tag{6}$$

and

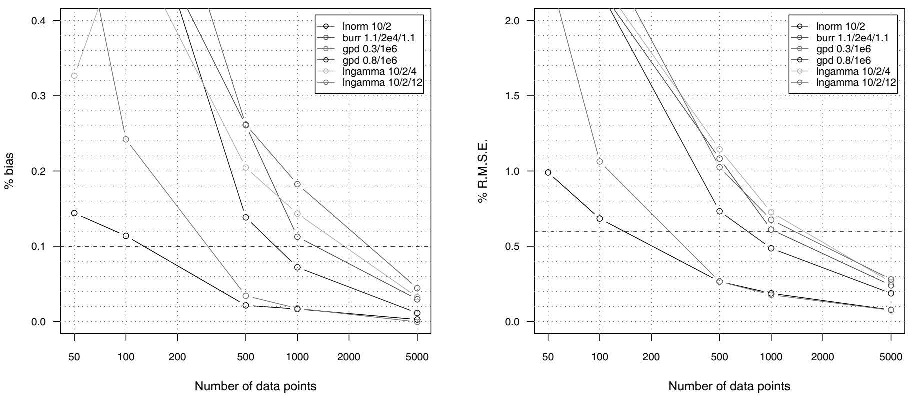
$$[\mathrm{RMSE}(\hat{q})]^2 \equiv \mathrm{E}[(\hat{q} - q_0)^2] = \frac{1}{M}\sum_i (\hat{q}_i - q_0)^2 \tag{7}$$

where q_0 is the 99.9% population quantile computed with the original distributions, with parameters as listed above, and $\hat{q}_i$ is the quantile estimate from the fit to the ith sample. Results will be given scaled by the value q_0; independence of the *scaled* results from the mean, λ, of the Poisson distribution was tested in the log-normal case, and found to hold well.

3.2.2 Results and comments

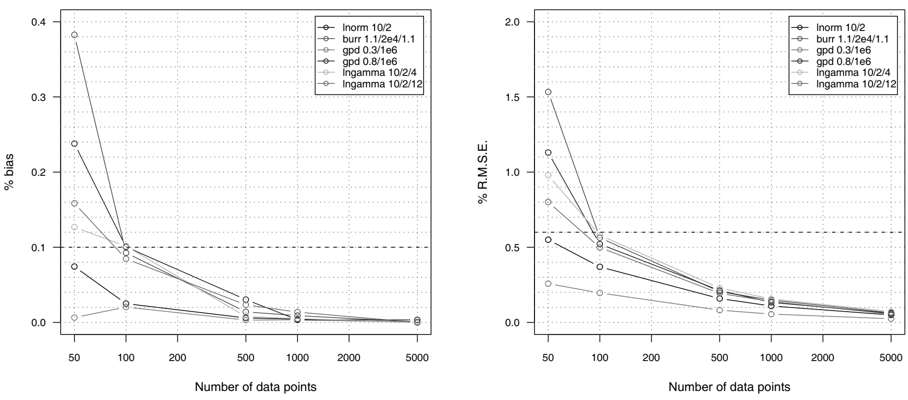
Figure 3(a) shows the percent bias of the 99.9% quantile of the aggregate loss distribution as a function of the number of data points available for the estimate, for different severity distributions. A reasonable 10% safety limit (McNeil and Saladin 1997) for lighter-tail distributions requires a few hundred data points, but for heavier-tail distributions at least a thousand points are needed.

FIGURE 3 (a) Percent bias and (b) percent root mean square error of the 99.9% quantile of the aggregate loss distribution as a function of the number of data points available for the estimate, for different severity distributions.



(a) (b)

FIGURE 4 Same as Figure 3, but for the 90% quantile.

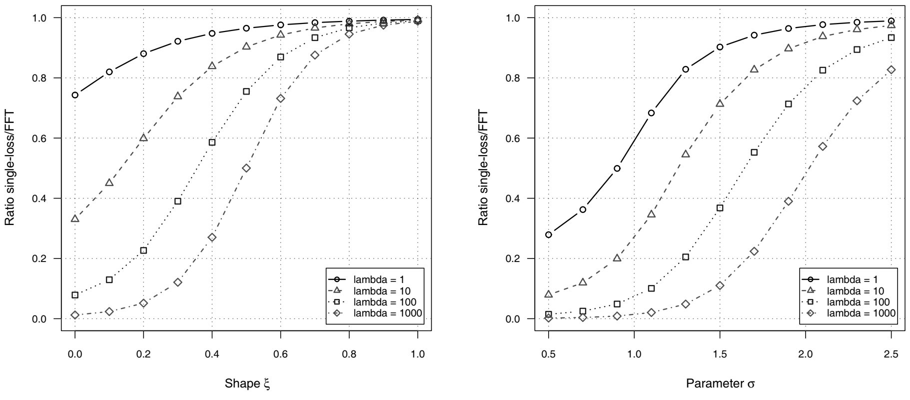


(a) (b)

Similar numerical conclusions are reached for the percent RMSE (Figure 3(b), here the safety limit was set at 60%), ie, essentially the uncertainty of the estimation process.

One can clearly see in Figure 4 how reducing the level of percentile to be estimated to 90% implies a much smaller uncertainty at a given number of data points, and hence a much smaller data sample is sufficient to reach the required accuracy.

FIGURE 5 Ratio of estimates of the 99.9% quantile of the aggregate loss using single-loss approximation and FFT for (a) GPD as a function of shape and for (b) the log-normal distribution as a function of σ, for increasing Poisson parameter λ.



(a) (b)

4 Controllable errors

As mentioned in the introduction, several approximation methods are commonly employed to compute the aggregate loss distribution and its quantiles. In particular, all of these methods introduce an error which in principle can be made arbitrarily small. In this section, comments are made on the size of controllable errors and on methods to reduce them.

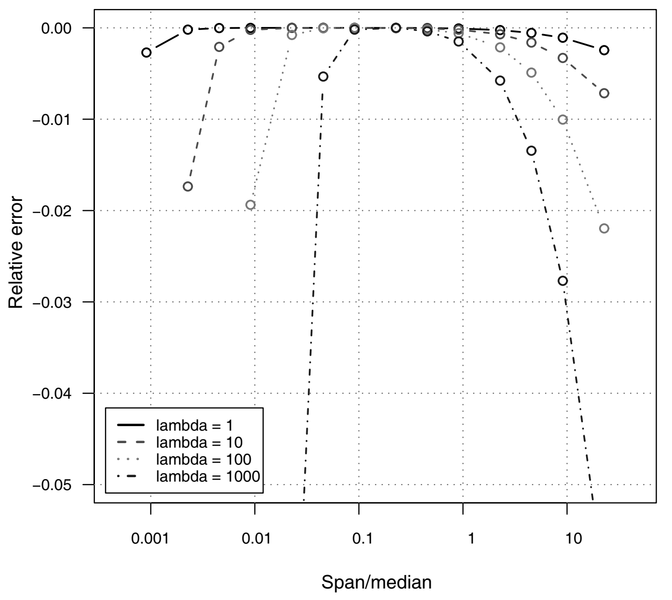
4.1 The single-loss approximation

While expressed in closed form, this method relies on an asymptotic property of the severity distribution in use: it would be very interesting and useful to compute higher-order terms in Equation (1). Numerical tests were performed (Figure 5), showing that in all cases of interest the single-loss approximation underestimates the correct aggregated quantile by an amount which grows with the mean multiplicity E[n] (λ in the figure); it appears as if on average the difference $\overline{F^{*n}}(x) - n\overline{F}(x)$ behaved proportionally to $\mathrm{E}[n]^2$, with a coefficient which is smaller, the heavier the tail of the severity F (heaviness that in the figure increases moving from left to right).

4.2 Fast Fourier Transform

A well-known method for computing the aggregate loss distribution is to use characteristic functions, as detailed, for example, in Klugman *et al* (2004). A very efficient implementation of this method relies on the application of

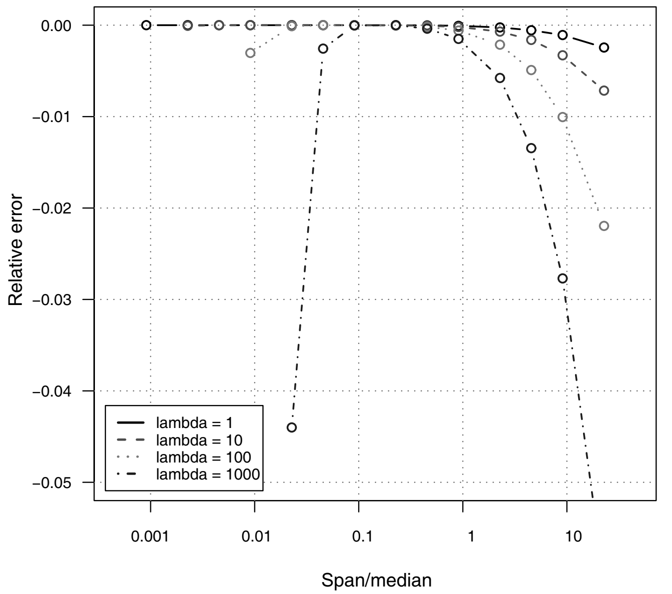
FIGURE 6 Relative error in the estimate of the 99.9% quantile of the aggregate loss for a log-normal severity with parameters ($\mu = 10$, $\sigma = 2$) and a Poisson frequency for $N_D = 2^{20}$ as a function of span for several values of the frequency parameter λ.



FFT algorithms. Although optimized for speed, care should be applied with respect to two sources of error, namely the effect of the discretization of the severity distribution (Grübel and Hermesmeier 2000) and of the truncation of the discretized distribution ("aliasing" error (Grübel and Hermesmeier 1999)). The two parameters which characterize this method are the number of points at which the severity density is discretized (N_D) and the discretization step ("span"). The natural constraint is that the product of N_D times span must be greater than the quantile to be estimated. While in principle the error can be reduced to zero by letting span go to zero and N_D to infinity, in practice time and computing power limit this possibility.

Figure 6 shows the relative error in the estimate of the 99.9% quantile of the aggregate loss for a log-normal severity with parameters ($\mu = 10$, $\sigma = 2$) and a Poisson frequency, for $N_D = 2^{20}$, as a function of span for several values of the frequency parameter λ. For values of span in the range 0.1–1 times the median the estimate is very precise (relative error much less than 1%). Lowering the span at fixed N_D first increases the aliasing error, then leads to the point at which the quantile is greater than N_D times span, so no solutions can be found. Increasing the span, on the other hand, leads eventually to a description of the density without enough detail, increasing the error. Finally, one should notice that larger values of

FIGURE 7 Same as Figure 6, but where an exponential tilting with parameter $200/q_0$, q_0 being the single-loss estimate of the quantile, has been applied.



the frequency parameter require more care in the choice of span: indeed larger λ implies larger quantile and hence a more stringent lower bound for N_D times span.

Figure 7 shows how the aliasing error can be reduced by using an "exponential tilting" technique (Grübel and Hermesmeier 1999). In this specific case a tilt parameter of $200/q_0$ was used, q_0 being the single-loss estimate of the quantile.

To show the effect of the thickness of the tail, a GPD with parameters ($\xi = 0.9$, $\sigma = 100{,}000$) is studied in Figure 8, with tilt applied. In this case the range of the ratio span/median values that give a good accuracy for all λ is shifted to larger values (1–10).

Our conclusion is that no mechanical way to fix N_D and span to optimal values is clearly indicated by the present results: case-by-case judgement should be carefully applied.

4.3 Monte Carlo method

When using the simulation method, one is estimating the quantile via order statistics: this means, essentially, finding the level q below which a given fraction p of the generated events falls. This leads to a binomial distribution in the percentile axis, which can be projected onto the quantile axis using the aggregated loss density $s(q)$ itself. Thus one arrives at the known formulae (David 1981;

FIGURE 8 Relative error in the estimate of the 99.9% quantile of the aggregate loss for a GPD severity with parameters ($\xi = 0.9$, $\sigma = 100{,}000$) and a Poisson frequency for $N_{\mathrm{D}} = 2^{20}$ as a function of span for several values of the frequency parameter λ.

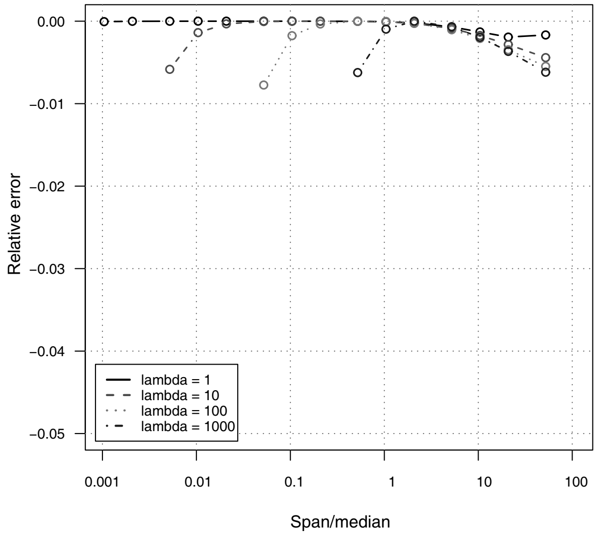
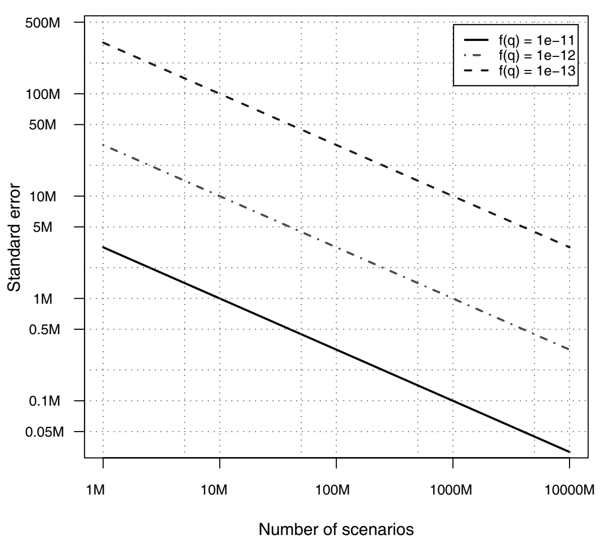


FIGURE 9 Absolute standard error (in millions, log scale) of the Monte Carlo estimation of the 99.9% quantile of the aggregate distribution versus the number of scenarios in the simulation (in millions, log scale), for three typical values of the aggregate density $f(q)$ at the quantile point.



Chen and Kelton 2001) for bias and variance in a simulation with N scenarios:

$$b(\hat{q}) = -\frac{p(1-p)}{2(N+2)}\frac{s'(q)}{s^3(q)} + O(1/N^2) \tag{8}$$

$$\operatorname{var}(\hat{q}) = \frac{p(1-p)}{(N+2)s^2(q)} + O(1/N^2) \tag{9}$$

The typical values one obtains for the value of the aggregated loss density at the quantile, $s(q)$, (eg, for the severity distributions mentioned in Table 2) vary between 10^{-11} and 10^{-13}. Taking $p = 0.999$, one obtains the curves shown in Figure 9. Although in principle it is a matter of computing power to attain arbitrary precision in the estimate, it is hardly worth the effort in view of the results of Section 3.

5 Other sources of uncertainty

To compute the whole firm risk measure, one needs both a mechanism of aggregation and the intensities of the couplings between the various risk classes. Tools that deal with the description of the joint distribution of random variables are available;[6] however, the specific model to be used and, even more critical, the intensity of the dependence between different risk classes, has to be determined from the data.

Such inference is very sensitive to the tail structure of the underlying distributions, and requires large data samples (Frahm *et al* 2005). Even though not investigated in detail here, it appears clear that this aspect, together with the high value of the percentile requested by the regulatory regime, adds another important source of uncertainty to the ones already described in the previous sections.

6 Conclusions

The sources of uncertainty embedded in the paradigm of operational risk loss modeling are examined in this paper. Far from taking into account specific details of the different models that have been developed (or are under development) in the industry, focus is put on the basic common features of the approach, in other words the loss distribution approach.

Three sources of uncertainty related to different aspects of model building are investigated in detail: the first one in order of importance is related to the absence of a causal dynamic model of loss production and hence to the ignorance of the "true" loss distribution, the second related to the statistical errors induced by the finiteness of the data sample and the third connected with the specific numerical tools used to compute the results.

The not very encouraging conclusion is that the error expected on the risk measures can be substantial and in some cases even of the same order of

[6] Such tools are generally known under the name of copulas; see, for instance, Embrechts *et al* (2005, ch. 5).

magnitude of the result itself, preventing it from having any kind of sensible meaning.

This conclusion is based on the very high percentile (driven by the new prudential regulation) at which the risk measure is calculated. This aspect, together with the scarcity of data, and with the data being heavy tailed, is responsible for most of the uncertainty that has been found in this paper; uncertainty that, in some cases (eg, choice of distribution), is not even quantifiable with precision.

A possible solution could be to considerably lower the percentile at which the risk measure is calculated (eg, to 90% or less)[7] and find some other mechanism to reach the "comfortable zone" that was meant to be achieved by the 99.9% percentile.

Appendix. Formal details of the model

The annual aggregate loss distribution $S(x)$ is obtained by studying separately the distribution of the loss frequency p_n and the loss severity $f(x)$. These two distributions are deemed to be independent and stable over time.

Based on these assumption, the aggregated annual loss distribution density can be written as

$$s(x) = \sum_{i=0}^{\infty} p_i f^{*i}(x)$$

where

$$f^{*i}(x) = \int_0^x f^{*(i-1)}(x-y) f(y)\,\mathrm{d}y$$

is called the i-fold convolution of the severity distribution and represents the probability density distribution of observing the aggregate of i individual losses that sum up at x.

The frequency distribution is assumed to be Poisson

$$p_n = \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}$$

with the value of λ determined from data.

The severity distributions used are as follows.

- The log-normal distribution

$$F(x) = \Phi(\log(x); \mu, \sigma)$$

  where $\Phi(x; 0, 1)$ is the standard normal distribution function.
- The Weibull distribution

$$F(x) = 1 - \mathrm{e}^{-(x/\theta)^{\tau}}$$

[7] A similar conclusion has been independently reached in Nešlehová *et al* (2006).

- The log-normal-gamma distribution (also known as the "fat-tailed log-normal")

$$f(x) = \int_0^{\infty} \phi\left(\log(x);\, \mu,\, \sqrt{y}\right) \gamma(y;\, \alpha,\, \sigma^2/\alpha)\, \mathrm{d}y$$

  where $\phi(x;\, \mu,\, \sqrt{y})$ is the normal density function with average μ and variance y and $\gamma(y;\, \alpha,\, \sigma^2/\alpha)$ is the gamma density function of shape α and scale σ^2/α.
- The Burr distribution

$$F(x) = 1 - \left[1 + \left(\frac{x}{\theta}\right)^{\gamma}\right]^{-\alpha}$$

- The generalized Pareto distribution (GPD)

$$F(x) = 1 - \left(1 + \xi\frac{x}{\sigma}\right)^{-1/\xi}$$

REFERENCES

Asmussen, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.

Basel Committee on Banking Supervision (2004). International convergence of capital measurement and capital standards. A revised framework (June).

Böcker, K., and Klüppelberg, C. (2005). Operational VaR: a closed-form approximation. *Risk* (December).

Chen, E. J., and Kelton, W. D. (2001). Quantile and histogram estimation. *Proceedings of the 2001 Winter Simulation Conference*, Peters, B. A. *et al* (eds). IEEE, Piscataway, New Jersey, pp. 451–459.

Chernobai, A., Rachev, S. T., and Fabozzi, F. J. (2005). Composite goodness-of-fit tests for left-truncated loss samples. Technical Report, University of California, Santa Barbara.

David, H. A. (1981). *Order Statistics*. Wiley, New York.

Frahm, G., Junker, M., and Schmidt, R. (2005). Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance: Mathematics and Economics* **37**, 80–100.

D'Agostino, R. B., and Stephens, M. A. (eds) (1986). *Goodness-of-fit Techniques*. Dekker, New York.

Eadie, W. T., Dryard, D., James, F. E., Roos, M., and Sadoulet, B. (1971). *Statistical Methods in Experimental Physics*. North-Holland, Amsterdam.

Embrechts, P., Furrer, H., and Kaufmann, R. (2003). Quantifying regulatory capital for operational risk. *Derivative Use, Trading & Regulation* **9**(3), 217–233.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modeling Extremal Events*. Springer, Berlin, Heidelberg.

Embrechts, P, McNeil, A. J., and Frey, R. (2005). *Quantitative Risk Management*. Princeton University Press, Princeton.

Grübel, R., and Hermesmeier, R. (1999). Computation of compound distributions I: aliasing errors and exponential tilting. *Astin Bulletin* **29**(2), 197–214.

Grübel, R., and Hermesmeier, R. (2000). Computation of compound distributions II: discretization errors and Richardson extrapolation. *Astin Bulletin* **30**(2), 309–331.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial (with discussion). *Statistical Science* **14**, 382–417.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2004). *Loss Models: from Data to Decisions*, 2nd edn. Wiley, New York.

McNeil, A. J., and Saladin, T. (1997). The peaks over thresholds method for estimating high quantiles of loss distributions. In *Proceedings of 28th International ASTIN Colloquium* (Cairns, Australia). Casualty Actuarial Society, Arlington, Virginia, pp. 23–43.

Mignola, G., and Ugoccioni, R. (2005). Tests of extreme value theory. *Operational Risk* **6**(10).

Nešlehová, J., Embrechts, P., and Chavez-Demoulin, V. (2005). Infinite-mean models and the LDA for operational risk. *Journal of Operational Risk* **1**(1), 3–25.

R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.