Abstract

A new methodology for financial and insurance operational risk capital estimation is proposed. It is based on using the finite time probability of (non-)ruin as an operational risk measure, within a general risk model. It allows for inhomogeneous operational loss frequency (dependent inter-arrival times) and dependent loss severities which may have any joint discrete or continuous distribution. Under the proposed methodology, operational risk capital assessment is viewed not as a one off exercise, performed at some moment of time, but as dynamic reserving, following a certain risk capital accumulation function. The latter describes the accumulation of risk capital with time and may be any non-decreasing, positive real function $h(t)$. Under these reasonably general assumptions, the probability of non-ruin is explicitly expressed using closed form expressions, derived by Ignatov and Kaishev (2000, 2004, 2007) and Ignatov, Kaishev and Krachunov (2001) and by setting it to a high enough preassigned value, say 0.99, it is possible to obtain not just a value for the capital charge but a (dynamic) risk capital accumulation strategy, $h(t)$.

In view of its generality, the proposed methodology is capable of accommodating any (heavy tailed) distributions, such as the Generalized Pareto Distribution, the Lognormal distribution the g-and-h distribution and the GB2 distribution. Applying this methodology on numerical examples, we demonstrate that dependence in the loss severities may have a dramatic effect on the estimated risk capital. In addition, we show also that one and the same high enough survival probability may be achieved by different risk capital accumulation strategies one of which may possibly be preferable to accumulating capital just linearly, as has been assumed by Embrechts et al. (2004). The proposed methodology takes into account also the effect of insurance on operational losses, in which case it is proposed to take the probability of joint survival of the financial institution and the insurance provider as a joint operational risk measure. The risk capital allocation strategy is then obtained in such a way that the probability of joint survival is equal to a preassigned high enough value, say 99.9 %.

Keywords: operational risk losses; operational risk capital assessment; dependent losses; Poisson loss arrivals; capital accumulation function; loss severity distribution; finite-time ruin probability; copulas
1. Introduction

Our aim in this paper is to propose a new methodology for modelling operational risk, based on risk and ruin theory. This is in compliance with the commitment of the Basel Committee on Banking Supervision (2001) (see its consultative report on the New Basel Capital Accord (Basel II)) to improve stability in the financial sector by reducing market risk, credit risk and operational risk. The first pillar, under the three pillar approach of Basel II, considers Minimal Capital Requirements and this is where new quantitative modelling methods, based on sound mathematical, statistical and probabilistic methodology are expected to provide a practically applicable tool for quantitative risk management. The demand for such new methods, which relate to solvency issues within the insurance industry, is also recognized within the new EU Solvency II project and the working programme of the International Association of Insurance Supervisors (IAIS) (see e.g. Linden and Ronkainen 2004). Actuarial techniques for quantifying operational risk in general insurance have recently been summarized by Tripp et al. (2004).

There are three alternative groups of methods for mitigating operational risk, outlined in the Basel Committee on Banking Supervision (2004), the basic indicator approach (BIA), the standardized approach (TSA) and the advanced measurement approach (AMA). The latter focuses on using internal and external loss data, among other techniques, and is often referred to as the Loss Distribution Approach (LDA). Under the AMA modelling framework the role of insurance in mitigating operational risk is also recognized. There are several examples of works under the LDA approach and here we will mention the common Poisson shock models of Ebnöther et al. (2001, 2002) and of Brandts (2004), and the ruin probability based models considered by Embrechts and Samorodnitsky (2003) and Embrechts et al. (2004). A more recent paper, considering the effect of insurance on setting the capital charge for operational risk is that of Bazzarello et al. (2006). The LDA approach has recently been used by Dutta and Perry (2006), who have considered fitting appropriate loss distributions to operational loss data under the 2004 Loss Data Collection Exercise (LDCE) and the Quantitative Impact Study 4 (QIS-4). Thus, it is more and more evident that LDA methods are becoming important for internal risk modelling purposes and at Basel-defined business line and
event type level modelling in order to improve the stability of the financial services industry. LDA methods are flexible and could be used within the whole financial industry sector, by central and commercial banks, insurance companies and supervisory bodies (see Cruz 2002, McNeil et al. 2005, Panjer 2006). No doubt, a great potential for developing such methods lies within the paradigm of ruin theory as has already been noted by Embrechts et al. (2004).

The classical ruin theory is over 100 years old and since the fundamental paper of Lundberg (1903), the number of publications (books, monographs and academic articles) in the probabilistic, statistical and actuarial literature is vast. Some important contributions to the field have been made by Cramér (1930), Seal (1978), Gerber (1988), Shiu (1987), Dickson (1994), Waters (1983), Grandell (1990), Picard and Lefêvre (1997), De Vylder (1999), Asmussen (2000), Willmot (2002), Gerber and Shiu (1998, 2005), Ignatov and Kaishev (2000, 2004, 2006) to mention only a few. Ruin theory may be viewed as the theoretical foundation of insolvency risk modelling. Under the classical ruin theory model, the (premium) income to an (insurance) company is modelled by a straight line \( h(t) = u + c t \), where \( u \geq 0 \) is the company's initial risk capital at time \( t = 0 \) and \( c \geq 0 \) is the premium income per unit of time, received by the company. The outgoing flow of claims paid by the company is modelled by a stochastic process,

\[
S(t) = \sum_{i=1}^{N(t)} W_i
\]

where, \( W_i, i = 1, 2, \ldots \) are assumed independent identically distributed (i.i.d.) random variables, modeling the amount of the consecutive individual losses, occurring at random moments in time. The stochastic process \( N(t) \), usually assumed a homogeneous Poisson process with parameter \( \lambda \), is counting the number of such losses up to time \( t \). The risk (surplus) process of the company is then defined as

\[
R(t) = u + c t - S(t)
\]

and the probability, \( P(T \leq \infty) \), that the aggregate amount of the loss payments, \( S(t) \), will exceed the in-flowing premium income \( h(t) = u + c t \) at some future moment, \( T \), is called the infinite-time probability of ruin of the company. In other words, this is the probability that the risk process, \( R(t) \), will become negative in some future moment, within an infinite time horizon.
The practical validity of model (1) for the aggregate operational losses under the LDA approach has been confirmed by Dutta and Perry (2006), who summarize the operational risk measuring experience of US banks under the QIS-4 submission.

Recently, Embrechts et al. (2004) proposed to take an actuarial point of view and directly apply the (classical) ruin probability model to the context of operational risk, under the LDA approach. Thus, the random variables $W_i, i = 1, 2, \ldots$ in model (1) are viewed as representing operational risk losses and the aggregate loss amount, $S(t)$, due to different types of operational risk, is expressed as a superposition of the risk processes, corresponding to each type of risk. The rate $c$ is seen "as a premium rate paid to an external insurer for taking (part of) the operational risk losses or as a rate paid to (or accounted for by) a bank internal office" (Embrechts et al., 2004). In order to reserve against operational risk, it is proposed to set the initial capital $u$ and the income rate $c$ in such a way that it satisfies the equation

$$P(T \leq x) = P\left(\inf_{0 \leq t \leq x} (u + c t - S(t)) < 0\right) = \epsilon$$

where the probability of ruin, $P(T \leq x)$, over a finite time interval, $[0, x]$, $0 < x \leq \infty$, is set to a pre-assigned appropriate (small) value $\epsilon > 0$. As noted in Embrechts et al. (2004), if the time interval is of length $x$ and $c = 0$, the risk capital $u$ is equal to the operational value at risk at significance level, $\alpha$, i.e.,

$$u = \text{OR} - \text{VaR}_{1-\alpha}^Y,$$

which is another popular risk measure considered in defining the capital charge for operational risk (see also Embrechts and Puccetti, 2006). In their paper, Embrechts et al. (2004) refer to ruin probability results, see e.g. Embrechts and Veraverbeke (1982), Asmussen (2000) and Schmidli (1999), which extend the applicability of the classical ruin probability model. However, the following major limitations may still be outlined:

- The function $h(t)$ is represented by a straight line, which is a simple but not a realistic assumption for the premium income.

- the losses, $W_i, i = 1, 2, \ldots$, are assumed independent and identically distributed which is also a restrictive assumption, not expected to hold for operational risk losses (see e.g. Panjer 2006, Chapter 8).
the ruin probability estimates quoted and discussed in Embrechts et al. (2004) are asymptotic approximations, i.e., for ruin on infinity, and as mentioned by the authors, "are not fine enough for accurate numerical approximations" and their numerical properties are "far less satisfactory", since these estimates are in an integral form.

In what follows, we propose a methodology which aims at generalizing the discussed classical ruin probability framework and making it a more practically applicable and useful approach for operational risk reserving. In particular, in our model, outlined in Section 2, we relax the above mentioned limitations and consider more general assumptions on the income function, on the distribution of the loss severities and their inter-arrival times, allowing them to be dependent. In Section 4, we consider a possible insurance coverage of the operational losses from a certain risk class (i.e., line of business or a BIS2 event type, as required by Basel Committee on Banking and Supervision 2004). Under the methodology proposed in Sections 3 and 4, it is possible to set not just a single value of the capital charge for operational risk, but to set a dynamic operational risk reserving strategy instead. This is briefly illustrated in Section 5 based on stylized numerical examples.

2. Ruin probabilities under a general model

Recently, a more general ruin probability model, relaxing the restrictive classical assumptions, has been considered by Ignatov and Kaishev (2000), where an explicit finite-time ruin probability formula was derived. Thus, the model considered by Ignatov and Kaishev (2000) assumes

- any non-decreasing (premium) income function $h(t)$ as an alternative to the classical straight line case
- any joint distribution of the losses $W_i, i = 1, 2, ..., $ allowing dependency between the loss amounts, as an alternative to the i.i.d. classical assumption
- finite time ruin probabilities, as an alternative to the asymptotic approximations of infinite ruin probabilities, suggested by Embrechts et al. (2004)
In a series of recent papers, (see Ignatov et al. 2001, 2004, Kaishev and Dimitrova 2006a, and Ignatov and Kaishev 2004, 2006) and the above mentioned ruin probability model has been explored and extended further and the following explicit non-ruin probability formulae have been derived.

Assume losses (claims) arrive at an insurance company with inter-arrival times \( \tau_1, \tau_2, \ldots \), identically, exponentially distributed r.v.s with parameter \( \lambda \), i.e., the number of the claims up to time \( t \), \( N(t) = \# \{ i : \tau_i + \ldots + \tau_i \leq t \} \), with \( \# \) denoting the number of elements in the set \{ \}. is a Poisson process with intensity \( \lambda \). In the case of discrete claim severities, the latter are modeled by the integer valued r.v.s. \( W_1, W_2, \ldots \) with joint distribution denoted by \( P_{w_1, \ldots, w_i} = P(W_1 = w_1, \ldots, W_i = w_i) \), where \( w_1 \geq 1, w_2 \geq 1, \ldots w_i \geq 1, i = 1, 2, \ldots \). The r.v.s \( W_1, W_2, \ldots \) are assumed to be independent of \( N(t) \). Then, the risk process \( R(t) \), at time \( t \) is given by

\[
R(t) = h(t) - S(t),
\]

where \( h(t) \) is a non-negative, non-decreasing, real function, defined on \( \mathbb{R}_+ \), representing the premium income of the insurance company and \( S(t) \) is the aggregate loss amount at time \( t \) defined as in (1) but assuming the losses have a joint distribution \( P_{w_1, \ldots, w_i} \).

The function \( h(t) \) is such that \( \lim_{t \to \infty} h(t) = \infty \). It may be continuous or discontinuous, in which case \( h^{-1}(y) = \inf \{ z : h(z) \geq y \} \). It will be convenient to denote the whole class of functions \( h(t) \), by \( \mathcal{H} \). We will denote also \( v_i = h^{-1}(i) \), for \( i = 0, 1, 2, \ldots \), noting that \( 0 = v_0 \leq v_1 \leq v_2 \ldots \). The time \( T \) of ruin is defined as

\[
T := \inf \{ t : t > 0, \ R(t) < 0 \}
\]

and we will be concerned with the probability of non ruin \( P(T > x) \) in a finite time interval \([0, x]\), \( x > 0 \). It has been shown by Ignatov and Kaishev (2000) that under this model the survival probability is given as

\[
P(T > x) = e^{-x \lambda} \sum_{w_1 \geq 1} P_{w_1, \ldots, w_n} \sum_{j=0}^{k-1} (-1)^j b_j(z_1, \ldots, z_j) \lambda^j \sum_{m=0}^{k-j-1} \frac{(x \lambda)^m}{m!}
\]

where \( n = \lfloor h(x) \rfloor + 1 \), \( \lfloor h(x) \rfloor \) is the integer part of \( h(x) \), \( v_{n-1} \leq x < v_n \), \( k \) is such that
\[ w_1 + ... + w_{k-1} \leq n - 1, \quad w_1 + ... + w_k \geq n, \quad (1 \leq k \leq n), \quad z_l = v_{w_1 + ... + w_l}, \quad l = 1, 2, \ldots \quad \text{and} \]
\[ b_j(z_1, \ldots, z_j) \text{ is defined recurrently as} \]
\[ b_j(z_1, \ldots, z_j) = (-1)^{j+1} \frac{z_j^j}{j!} + (-1)^{j+2} \frac{z_j^{j-1}}{(j-1)!} b_1(z_1) + ... + (-1)^{j+2} \frac{z_j^1}{1!} b_{j-1}(z_1, \ldots, z_{j-1}), \]
with \( b_0 \equiv 1, b_1(z_1) = z_1. \)

In Ignatov et al. (2001), formula (5) has been given the following exact, numerically efficient representation
\[ P(T > x) = e^{-x\lambda} \sum_{k=1}^{\infty} \sum_{w_1 \leq \ldots \leq w_k \leq 1} P(W_1 = w_1, \ldots, W_{k-1} = w_{k-1}; W_k \geq n - w_1 - \ldots - w_{k-1}) \]
\[ \cdot \left[ \sum_{j=0}^{k-1} (-1)^j b_j(z_1, \ldots, z_j) \lambda^j \sum_{m=1}^{k-1} \frac{(x\lambda)^m}{m!} \right]. \]

When claims have any continuous joint distribution, the probability of non-ruin within a finite time \( x \) has recently been shown by Ignatov and Kaishev (2004) to admit the representation
\[ P(T > x) = e^{-x\lambda} \left\{ 1 + \sum_{k=1}^{\infty} \lambda^k \int_0^{h(x)} dy_1 \int_{y_1}^{h(x)} dy_2 \ldots \right. \]
\[ \left. \int_{y_{k-1}}^{h(x)} A_k(x; h^{-1}(y_1), \ldots, h^{-1}(y_k)) f(y_1, \ldots, y_k) dy_k \right\}. \]
where \( y_0 = 0, A_k(x; v_1, \ldots, v_k), k = 1, 2, \ldots, \) are the classical Appell polynomials \( A_k(x) \) of degree \( k \) with a coefficient in front of \( x^k \) equal to \( 1/k! \), defined by
\[ A_0(x) = 1, \quad A_k(x) = A_{k-1}(x), \quad A_k(v_k) = 0, \quad k = 1, 2, \ldots. \]
and \( f(y_1, \ldots, y_k) \) is the joint density of the partial sums of consecutive claims \( Y_1 = W_1, Y_2 = W_1 + W_2, \ldots, Y_i = W_1 + \ldots + W_i, \ldots. \)

Obviously, the claim severities are related with the r.v.s \( Y_1, Y_2, \ldots \) through the equalities
\[ W_1 = Y_1, \ W_2 = Y_2 - Y_1, \ W_3 = Y_3 - Y_2, \ ... \text{i.e.,} \ Y_1, \ Y_2, \ ... \] and the joint density \( \psi(w_1, ..., w_k) \) of the r.v.s \( W_1, W_2, ..., W_k \) can be expressed as

\[ \psi(w_1, ..., w_k) = f(w_1, w_1 + w_2, ..., w_1 + ... + w_k) \]

or

\[ f(y_1, y_2, ..., y_k) = \psi(y_1, y_2 - y_1, ..., y_k - y_{k-1}). \]

It is worth mentioning that the practical use of the ruin theoretic results presented here critically depends on their numerical performance. Both formulae (6) and (7) have been implemented numerically (see Ignatov et al. 2001 and Ignatov and Kaishev 2004) and allow for the efficient computation of \( P(T > x) \) for any discrete or continuous joint distribution of the losses \( W_i, i = 1, 2, ... \). The computational properties of formulae (5) and (6) have been explored in Ignatov et al. (2001) and as has been demonstrated, formula (6) is an improved version of (5), and allows for the exact and efficient computation of the ruin probability with any prescribed accuracy (depending only on the computational resources available). This is possible because it involves only finite summation of the determinants, \( b_j(z_1, ..., z_j) \), and allows for some further recurrent enhancements. For related details, numerical results and comparisons with Monte Carlo evaluations based on Mathematica implementations, we refer to Ignatov et al. (2001). It has to be noted that, to the best of our knowledge, there are no other formulae which could produce exact ruin probability values under the general models, underlying formulae (6) and (7). Approximate values for these probabilities can be obtained by using Monte Carlo simulation. However, it is well known that Monte Carlo methods have very slow convergence, hence may require millions of time consuming simulations to achieve reasonable accuracy. Furthermore, not all distributions used to model claim amounts are analytically invertible, some multivariate dependent distributions are not so straightforward to simulate from, which additionally hinders the application of the Monte Carlo method. Attempts to improve the Monte Carlo efficiency, by applying various variance reduction techniques such as control and antithetic variates, and low discrepancy sequences have their limitations and depend on the effective dimension of the problem. For example the use of low discrepancy sequences under the Quasi Monte Carlo (QMC) approach is restricted by the dimension of the implemented existing sequences, such as
Sobol and Korobov sequences. For further details on QMC and other variance reduction techniques, we refer to Glasserman (2004).

The numerical performance of formula (7) is briefly illustrated in Ignatov and Kaishev (2004) and is explored in somewhat greater detail in a separate study, preliminary results of which have been presented at the 10th International Congress on Insurance: Mathematics and Economics (see Kaishev and Dimitrova 2006b) and a related paper is under preparation. This study indicates that although formula (7) involves infinite summation of multiple integrals with increasing dimension it still can be efficiently evaluated with any prescribed accuracy, following an appropriate algorithmic implementation. The latter substantially uses a numerically efficient recurrent representation of the classical Appell polynomials, \( A_k(x; v_1, ..., v_k) \), combined with the fact that they are multiplied by the joint probability density of the individual claim amounts. This specific structure of formula (7) allows for the accurate and efficient numerical evaluation of the multiple integrals, practically up to a dimension of several hundreds, depending on the values of the risk model parameters, such as \( x, \lambda, h(t) \) and the distribution of the claims.

Recently, a further extension of the underlying risk model, beyond Poisson claim arrivals, has been considered by Ignatov and Kaishev (2007). The authors have obtained closed form finite-time non-ruin probability expression, under the assumption that inter-arrival times, \( \tau_i, i = 1, 2, ... \) are independent, (non-identically) Erlang distributed random variables with density function,

\[
f_{\tau_i}(t) = \lambda^{g_i} t^{g_i-1} e^{-\lambda t} / \Gamma(g_i),
\]

where \( g_i, i = 1, 2, ... \) is a sequence of positive integers, (Erlang shape parameters), and \( \lambda \) is the Erlang rate parameter. This expression has been generalized further, to allow dependence in the inter-arrival times, governed by a reasonably flexible dependence structure. The latter is imposed by appropriately randomizing the Erlang shape parameters, \( g_i, i = 1, 2, ... \). This can be viewed as obtaining a dependent distribution of the inter-arrival times, \( \tau_i, i = 1, 2, ... \), by compounding their Erlang distributions with an appropriate multivariate distribution over \( g_i, i = 1, 2, ... \). Such (dependence) features of the risk model and the related non-ruin probability formulae (see Ignatov and Kaishev 2007) are especially
appropriate for modelling operational losses, since they can successfully capture variability in the loss frequency over time. The possibly substantial inhomogeneity in the loss frequency is a stylized fact in the operational loss literature (see e.g., McNeil et al. 2005, Section 10.1.4).

The flexibility of the results mentioned in this section makes them especially attractive in modelling operational risk capital allocation, which is considered in the next section.

3. Capital assessment under the general ruin probability model

The (non-) ruin probability formulae (6) and (7), are flexible and can be directly applied under the LDA approach to operational risk modelling and capital assessment, assuming ruin probability is selected as an operational risk measure. To see this, note that taking into account the general ruin probability model outlined in Section 2, equation (2) can be rewritten as

$$\begin{align*}
P(T > x) &= 1 - \mathbb{P}\left(\inf_{0 \leq t \leq x} (h(t) - S(t)) < 0 \right) = 1 - \epsilon
\end{align*}$$

where the non-ruin probability on the left-hand side can be directly expressed by formula (6) if loss severities \( W_i, i = 1, 2, \ldots \) are assumed discrete or by (7) if they are assumed continuous. Operational risk capital allocation, can now be formulated as "selecting" an appropriate "capital accumulation" function \( h(t) \in \mathcal{H} \), such that equation (8) is satisfied for a sufficiently small preassigned value \( \epsilon > 0 \). It has to be noted that there may be infinitely many solutions to the functional equation (8), since the class \( \mathcal{H} \) is rather general. In particular the functions \( h(t) \in \mathcal{H} \) need not be continuous and thus, may incorporate jump discontinuities at some points in time. Moreover, \( \mathcal{H} \) need not necessarily be strictly increasing which means that step-wise constant functions \( h(t) \) may also be considered.

Somewhat surprisingly, the flexibility of the class \( \mathcal{H} \) leads to the possibility of selecting a function \( h(t) \), which maximizes the probability of non-ruin, \( P(T > x) \), of a financial institution, say a bank, over an appropriate subclass of \( \mathcal{H} \). In other words, the bank has the flexibility of selecting different capital accumulation strategies, \( h(t) \), for reserving against operational risk, so as to maximize its chances of survival from operational losses. For example, if
the appropriate subclass is the class of all piecewise linear functions on \([0, x]\), with one jump of size \(J\), at some instant \(t_J \in [0, x]\) the bank may put aside less amount, \(u\), of initial capital at time \(t = 0\) and top up this capital by an amount \(J\) at some (optimal) later moment \(t_J\). This point is illustrated numerically in Section 5 (see Fig. 2) where it is demonstrated that one and the same high non-ruin probability \(1 - \epsilon\) can be achieved by different alternative choices of capital accumulations, \(h(t)\), whose values at the terminal time point, \(x\), coincide.

In general, to distinguish between different choices of the reserving capital accumulation function \(h(t)\), and thus to facilitate the solution of (8), these choices can be attached a different utility which may for instance be related to the cost of borrowing capital from the bank. For example, the bank may find it preferable to set less initial reserve \(u\) and top up its reserves at a later instant. In order to illustrate this point, assume that preference is measured by the Expected Present Value (EPV) of the continuous cash flow \(\dot{h}(t) = dh(t)/dt\). Then, from two different solutions of (8), which provide equal probabilities of survival, the bank will choose the solution with lower EPV. Since our purpose here is to introduce the major concepts and discuss model (8) we will restrain from going into greater details with respect to this utility modelling aspect.

A second point which deserves to be made in connection with setting operational reserves according to (8) is that the joint distribution of the operational losses \(W_1, W_2, \ldots\) can be any joint distribution, continuous or discrete. This is possible since formulae (6) and (7) are general and are valid for any i.i.d or dependent losses. Thus, they can easily accommodate any of the widely advocated one dimensional heavy tailed operational loss distributions, such as the Generalized Pareto Distribution (GPD), the Lognormal distribution or the less popular g-and-h and Generalized Beta Distribution of Second Kind (GB2) distributions, recently put forward by Dutta and Perry (2006). Properties of the g-and-h distribution in the context of operational loss modelling and extreme value theory has recently been explored further by Degen et al. (2007).

It is a common argument in the operational risk modelling literature (see e.g. Embrechts and Puccetti 2006) that operational losses do, in general, exhibit dependence in their severity. Taking account of this dependence may require significantly higher capital reserves on aggregate as illustrated in Section 5, (see Fig. 3), based on ruin probability as operational
risk measure. Thus, allowing for modelling dependence is an important feature of the methodology proposed here. Dependence can be incorporated in the loss distribution using any of the available dependence modelling techniques, for example Markovian-type dependence structures (see Albrecher and Boxma 2004), copulas or any of the existing multivariate distributions. It has to be noted that there are very few examples in the literature of dependent multivariate distributions which have been used to model dependent severities of consecutive insurance claims and operational losses. Illustrations of how this can be done are to be found in Ignatov, Kaishev and Krachunov (2001, 2004) for multivariate discrete distributions and Ignatov and Kaishev (2004) for continuous distributions.

We believe there is a great potential in exploring the applicability of appropriate classes of multivariate distributions in modelling dependence of operational losses and insurance claims. An extensive list of such candidate distributions is to be found in Johnson and Kotz (1994) Johnson, Kotz and Balakrishnan (1997) and elsewhere in the statistical literature. A few examples in this direction are the class of multivariate gamma distributions, the skew-normal distribution proposed by Azzalini and Valle (1996), the slash and skew-slash Student t distributions, recently explored by Tan and Peng (2005).

Alternatively, copulas can serve the purpose of modelling dependence in insurance losses, and for details of how this can be done we refer to Kaishev and Dimitrova (2006a). Although copulas have recently gained considerable popularity in various applications in insurance and finance (see e.g. McNeil et al. 2005, Chapter 5 and Cherubini et al. 2004) some practical difficulties, related to their multivariate versions, their appropriate parameterization and estimation, based on data, still exist. In particular there are only a few families of (multivariate) copulas which involve sufficiently many parameters, so as to be flexible enough and capture real world multivariate loss dependences. One such popular example is the family of Elliptical copulas to which Gaussian and t-copulas belong. The family of Archimedean copulas is also a popular choice in practice, although they offer less flexibility due to their symmetry (exchangeability) and scarce parameterization. Most popular Archimedean copulas, such as Gumbel, Frank, Clayton copulas involve only one parameter. More richly parameterized Archimedean copulas and their estimation have been recently considered by Lambert (2006) in the bivariate case, applying Bayesian splines and by Dimitrova et al. (2007) in the multivariate case, using so called Geometrically Designed (GeD) splines. For
extensions of the Archimedean copulas to non-exchangeability (asymmetry) see McNeil et al. (2005), where further references are to be found. For summarized information on copulas and their application, we refer to McNeil et al. (2005), Nelsen (2006) and Cherubini et al. (2004).

As has been noted, in the operational loss literature (see e.g., McNeil et al. 2005) loss frequency may exhibit strong inhomogeneity. One way of reflecting this is to depart from the assumption of Poisson loss arrivals at a constant rate $\lambda$, and assume that inter-arrival times are independent but have different distribution (as in the Sparre Andersen model), or are dependent random variables with a certain joint distribution. Under such assumptions the left hand side of (8) may be replaced by the corresponding formulae for $P(T > x)$, given in Ignatov and Kaishev (2007), for the case of (compound) Erlang distributed inter-arrival times, in order to obtain an appropriate operational risk capital accumulation function $h(t)$ for a fixed level $\epsilon$.

Further aspects of the methodology outlined in this section are discussed and illustrated numerically in Section 5.

4. Capital assessment under insurance on operational losses

Another important aspect of modelling operational risk capital assessment, recognized under the AMA approach, is the effect on it of insurance on operational losses. The latter has been considered recently by Brandts (2004) and Bazzarello et al. (2006) where it is assumed that individual operational losses are insured with an external insurer under an excess of loss (XL) contract. Under this model there is a deductible $d > 0$ and a policy limit $m > 0$, and what is covered by the insurer is

$$W_i^r = \min(\max(W_i - d, 0), m), \quad i = 1, 2, ...$$

whereas the net loss covered by the internal operational risk management (ORM) office of a financial institution is

$$W_i^c = W_i - W_i^r = \min(W_i, d) + \max(0, W_i - (d + m)), \quad i = 1, 2, ...$$

(9)
Thus, under such an arrangement, there are two parties providing the operational loss cover, the ORM office, which plays the role of an internal direct insurer and the external insurer, which could be viewed as a reinsurer. The role of the latter party is essential and the probability of it defaulting has been considered by Brandts (2004) and by Bazzarello et al. (2006).

Here, we take a different approach, motivated by the observation that both parties share the operational risk they jointly cover, and hence in defining the total risk capital, allocated overall and split by the two parties, it is meaningful to consider their joint chances of not defaulting, i.e., to consider the probability of their joint survival. To follow details of this approach we will introduce some further notation.

Denote by \( Y_1^c = W_1^c \), \( Y_2^c = W_1^c + W_2^c \), ... and by \( Y_1^r = W_1^r \), \( Y_2^r = W_1^r + W_2^r \), ... the partial sums of consecutive operational losses to the ORM office and to the external insurer, respectively. Obviously, in view of (9), we have that \( Y_i^c + Y_i^r = Y_i \), \( i = 1, 2, ... \), i.e., operational losses are shared. Under this XL reinsurance model, the total capital, \( h(t) \), accumulated by the ORM office is also divided between the two parties so that \( h(t) = h_c(t) + h_r(t) \), where \( h_c(t) \), is the ORM office's capital accumulation function and \( h_r(t) \) models premium income of the external insurer, assumed also non-negative, non-decreasing functions on \( \mathbb{R}_+ \). As a result, the risk process, \( R(t) \), can be represented as a superposition of two risk processes, that of the ORM office

\[
R_c^r = h_c(t) - Y_{N_t}^c
\]  

(10)

and of the insurer

\[
R_r^r = h_r(t) - Y_{N_t}^r
\]  

(11)

i.e., \( R(t) = R_c^r + R_r^r \).

Denote by \( P(T^c > x, T^r > x) \), the probability of joint survival of the bank ORM office and the external insurer up to time \( x \), where \( T^c \) and \( T^r \), denoting the moments of ruin of the two parties, are defined as in (4), replacing \( R(t) \) with \( R_c^r \) and \( R_r^r \) respectively. Clearly, the two events \( (T^c > x) \) and \( (T^r > x) \), of survival of the bank ORM office and the insurer are dependent since the two risk processes \( R_c^r \) and \( R_r^r \) are dependent through the common loss arrivals and the loss severities \( W_i, i = 1, 2, ... \), as seen from (10) and (11). This motivates us to con-
sider the probability of joint survival, $P(T^c > x, T^r > x)$, as a joint measure of operational risk when operational losses are insured. The following risk capital allocation problem can then be formulated, which takes into account the fact that the two parties share the risk and the total capital accumulated.

**Problem 1.** For fixed deductible $d$ and policy limit $m$, find capital accumulation function $h(t) \in \mathcal{H}$ with a representation $h(t) = h_c(t) + h_r(t), h_c(t), h_r(t) \in \mathcal{H}$ such that

$$P(T^c > x, T^r > x) = 1 - \epsilon. \quad (12)$$

Clearly, this problem may in general have more than one solution. Further conditions may be imposed to restrict the set of possible solutions.

In order to solve Problem 1, the following explicit expression for the probability of joint survival up to a finite time $x$, recently derived by Kaishev and Dimitrova (2006a), can be used in the case of continuous loss severities. We have

$$P(T^c > x, T^r > x) = \left(1 + \sum_{k=1}^{\infty} e^{-\lambda x} \frac{(\lambda x)^k}{k!} \right) \prod_{j=1}^{k} \left( \int_{0}^{\psi(x; \tilde{\nu}_1, ..., \tilde{\nu}_k)} A_k(x; \tilde{\nu}_1, ..., \tilde{\nu}_k) \psi(w_1, ..., w_k) d w_k ... d w_2 d w_1 \right) \quad (13)$$

where

$$\tilde{\nu}_j = \min(\tilde{z}_j, x), \tilde{z}_j = \max(h_c^{-1}(y_j), h_r^{-1}(y_j)), y_j = \sum_{i=1}^{j} w_i, y_j = \sum_{i=1}^{j} w_i, j = 1, ..., k,$$

$$w_i = \min(w_i, d) + \max(0, w_i - (d + m)), w_i = \min(m, \max(0, w_i - d)),$$ and

$$A_k(x; \tilde{\nu}_1, ..., \tilde{\nu}_k), k = 1, 2, ...$$ are the classical Appell polynomials $A_k(x)$ of degree $k$, defined as in (7).

Let us note that expression (13) is a generalization of formula (7) which follows from (13) in the special case of $m = 0$. Formula (13) has been implemented using the *Mathematica* system and to follow its numerical performance (also in solving optimal reinsurance problems) we refer to Kaishev and Dimitrova (2006a). Thus, formula (13) can be successfully applied to represent the left-hand side of equation (12) and solve Problem 1. In the case of discrete
claim amounts, Problem 1 can also be formulated and solved with formula (13) replaced by a
discrete analog due to Ignatov et al. (2004).

It can be argued that a typical (re)insurance company would most likely only insure a small
percentage of the bank's losses and would also insure many other banks and firms and many
other perils at the same time in order to diversify. However, often big banks and firms would
prefer to insure substantial part of their losses with one particular big (re)insurance company
and these losses would represent substantial part of the total business underwritten by the
(re)insurer. It is in such cases, where joint survival of the two parties is critical and default of
any of them with respect to the risk-sharing contract, may cause downgrading of their credit
rating or even bankruptcy, as was recently the case with the 6-th largest worldwide reinsurance
company Gerling Global Re. With the increased frequency and severity of catastrophic
events such scenarios become even more likely and this is why the simple model of joint
survival of two parties sharing the risk is relevant. Of course, in cases where there are many
parties involved in a risk sharing arrangement, the two-party model presented here may be
applied on a bilateral bases and is obviously a necessary first step towards further generaliza-
tions to the more complex multi-party multi-risk sharing reality.

Next, we provide some numerical illustrations of the methodology described in Section 3.

5. Numerical illustrations

In order to illustrate the methodology outlined in Section 3, we consider five alternative
distributions of the consecutive losses. In our first example, operational risk losses are
assumed i.i.d. with a discrete, logarithmic distribution, i.e. $W_i \sim \text{Log}(\alpha)$ with a generic p.m.f.

$$
P(W = i) = -\alpha^i / (i \ln (1 - \alpha)).
$$

We have calibrated this distribution against operational risk loss data by approximately matching its mean and variance to the Lognormal distribution fitted by Brandts (2004), (see Table 5 therein) to the aggregate losses from the 2002 LDCE data file. This is achieved for $\alpha = 0.73$. Of course, the logarithmic distribution we use, has lighter tail than the Lognormal one, but it suits our illustrative purposes here. A set of operational losses arriving in the interval $[0, 2]$ with inter-arrival times distributed as $\text{Exp}(20)$ and with severities simulated from the Log(0.73) distribution are presented in the left panel of
In the right panel of Fig. 1, for \( h(t) = u + c \, t \) with \( c = 25 \), we have presented values of the initial capital \( u \) for different choices of the probability of survival \( P(T > 2) \).

As can be seen, the capital charge \( u \) increases nonlinearly with the increase of the probability of survival, at a much higher rate as \( P(T > 2) \) approaches one. The calculations have been performed in Mathematica, solving (8) with \( P(T > 2) \) expressed by (6), applying the Newton algorithm. In particular, \( P(T > 2) = 0.99 \) is achieved for \( h(t) = 79.4 + 25 \, t \). To illustrate the fact that the same probability 0.99 can be achieved by alternative choices of the capital accumulation function \( h(t) \), we have next assumed that it belongs to the subclass of all piecewise linear functions on \([0, x]\), with one jump of size \( J \), at some instant \( t_J \in [0, x] \), i.e.,

\[
h(t) = \begin{cases} u + c_1 \, t & , \quad 0 \leq t < t_J \\ u + c_1 \, t_J + J + s_1(t - t_J) & , \quad t_J \leq t \leq x \end{cases}
\]

In the left panel of Fig. 2., two choices of \( h(t) \) are plotted, \( h_1(t) = 79.4 + 25 \, t \) and

\[
h_2(t) = \begin{cases} 59.4 + 27 \, t & , \quad 0 \leq t < 1 \\ 59.4 + 27 + 20 + 23 \, (t - 1) & , \quad 1 \leq t \leq 2 \end{cases}
\]

As illustrated in the right panel of Fig. 2, moving the location \( t_J \) of the jump \( J = 20 \) from \( t_J = 0 \) to \( t_J = 2 \), while keeping the rest of the parameters fixed, we can see that a maximum of \( P(T > 2) = 0.99 \) is achieved for \( t_J = 1 \). Indeed, both functions \( h_1(t) \) and \( h_2(t) \) provide equal chances of survival, 99% and also, accumulate equal risk capital at the end of the time interval, \( x = 2 \), i.e. \( h_1(2) = h_2(2) = 129.4 \). But obviously the choice \( h_2(t) \) is preferable since it
requires less capital, $u = 59.4$, to be put aside initially, compared to $u = 79.4$ for the choice $h_1(t)$.

![Graph showing two choices of the capital accumulation function, $h_1(t) = 79.4 + 25t$ (thick line) and $h_2(t) = (59.4 + 27t)I_{t<1} + (59.4 + 27 + 20 + 23(t-1))I_{t>1}$ (dashed line). Right panel: $P(T > 2)$ as a function of the location $t_J$ of the jump in $h_2(t)$ of size $J = 20$.](image)

In order to demonstrate how the methodology works under the assumption that losses have a continuous multivariate distribution, we consider two alternatives, a light-tail exponential distribution and a heavy-tail Pareto distribution, assuming both independent and dependent risk losses. First, the severities of the consecutive risk losses $W_i$, $i = 1, 2, \ldots$, are assumed independent, identically distributed following $\text{Exp}(0.5)$ or $\text{Pareto}(2.41, 1.17)$, so that their mean matches the mean of the 2002 LDCE data. A simulation from the joint distribution of the severities of two i.i.d. risk losses, $W_i$, $i = 1, 2$ is given in Fig. 3 (a), in the case of $W_i \sim \text{Exp}(0.5)$, and Fig. 3 (c), in the case of $W_i \sim \text{Pareto}(2.41, 1.17)$. Secondly, $W_i$, $i = 1, 2, \ldots$ are assumed dependent, with joint distribution function given by the Rotated Clayton copula, $C_{\text{RCI}}(u_1, \ldots, u_k; \theta)$ and the marginals are assumed to be $\text{Exp}(0.5)$ or Pareto$(2.41, 1.17)$ distributed. Considering these four cases allows us to study the effect of assuming a heavier tail distribution and the effect of dependence on the risk capital allocation, in particular on the size of the initial capital charge $u$.

The Rotated Clayton copula, $C_{\text{RCI}}(u_1, \ldots, u_k; \theta)$, is defined as

$$C_{\text{RCI}}(u_1, \ldots, u_k; \theta) = \frac{\sum_{i=1}^k u_i - k + 1 + \left(\sum_{i=1}^k (1 - u_i)^{-\theta} - k + 1\right)^{-1/\theta}}{1},$$

with density $c_{\text{RCI}}(u_1, \ldots, u_k; \theta) = c_{\text{CI}}(1 - u_1, \ldots, 1 - u_k; \theta)$ and parameter $\theta \in (0, \infty)$. The value $\theta = 0$ corresponds to independence. The Rotated Clayton copula has upper tail depen-
dence with coefficient \( \lambda_U = 2^{-1/\theta} \) and is suitable for modeling dependence between extreme operational losses.

Losses with dependence according to a Rotated Clayton copula with parameter \( \theta = 1 \) are illustrated through a random sample of 2000 data points in Fig. 3 (b), in the case of identical \( \text{Exp}(0.5) \) marginals, and Fig. 3 (d), in the case of identical \( \text{Pareto}(2.41,1.17) \) marginals. The presence of positive dependence, determined by \( \theta = 1 \), and of upper tail dependence, \( \lambda_U = 2^{-1} \), is clearly visible. We refer the reader to Kaishev and Dimitrova (2006a) for further applications of this copula in modelling dependence of insurance claim severities combined with other (heavy-tailed) marginal distributions.

Fig. 4 illustrates the heavy impact of dependence between loss severities on the value of the initial capital charge \( u \), given \( h(t) = u + 25t \), \( x = 2 \) and Poisson inter-arrival times \( \tau_i \sim \text{Exp}(20) \). As can be seen in the left panel of Fig. 4, in order to achieve survival probability \( P(T > 2) = 0.90 \) the capital charge should be \( u = 55.7 \) in the case of i.i.d. \( W_i \sim \text{Exp}(0.5) \) and \( u = 112 \) when assuming dependence. Furthermore, if a probability of \( P(T > 2) = 0.999 \) is to be achieved, the corresponding values are \( u = 98.3 \) for i.i.d. losses and \( u = 466 \), for dependent losses, which is 4.74 times higher. The values of the capital charge \( u \) have been calculated solving (8), with \( P(T > 2) \) given by formula (7). Similar results are presented for \( \text{Pareto}(2.41, 1.17) \) risk losses in the right panel of Fig. 4. Comparing it with the exponential case, one can see that lower level of capital is required for probabilities 0.90 and 0.95 and similar or greater \( u \) is needed for higher probabilities (between 0.99 and 0.999) both for the independent and the dependent cases, due to the effect of the heavy-tailedness of the Pareto distribution and its left truncation which rules out losses smaller than 1.17.

Of course, the choice of the Rotated Clayton copula with parameter \( \theta = 1 \), leads to Kendall's \( \tau = 0.33 \) and upper tail dependence \( \lambda_U = 0.5 \), which tailors a reasonably strong dependence, so the result could be extreme, but convincingly illustrates the importance of considering dependence when setting operational risk capital charge. However, our experience indicates that different choices of dependence structures (e.g., Markovian type dependence or different copulas) and loss distributions may have quite a different effect on the level of capital charge over the entire probability range.
Fig. 3. (a) simulated i.i.d. Exp(0.5) losses; (b) simulated dependent losses following $C^{RCLI}(u_1, u_2; 1)$ with Exp(0.5) marginals; (c) simulated i.i.d. Pareto(2.41,1.17) losses; (d) simulated dependent losses following $C^{RCLI}(u_1, u_2; 1)$ with Pareto(2.41, 1.17) marginals.

Fig. 4. Initial capital $u$, for choices of $P(T > 2)$ equal to 90%, 95%, 99%, 99.5% and 99.9%, $h(t) = u + 25 t$, $\tau_i \sim \text{Exp}(20)$, in the case of: Left panel - i.i.d. Exp(0.5) losses (thick line) and dependent losses following $C^{RCLI}(u_1, u_2; 1)$ with Exp(0.5) marginals (dashed line); Right panel - i.i.d. Pareto(2.41,1.17) losses (thick line) and dependent losses following $C^{RCLI}(u_1, u_2; 1)$ with Pareto(2.41, 1.17) marginals (dashed line).
6. Comments and conclusions

We have demonstrated that the proposed methodology which is based on solving (8) and (12) using the explicit formulae (6), (7) and (13) is a promising modelling tool for (dynamic) operational risk capital allocation. An important conclusion is that dependence of the sizes of operational losses may have a dramatic effect on the operational loss reserving strategy. A further investigation into this phenomenon would look into the effect of introducing cross-over dependence between inter-occurrence times of losses and their amounts. Further insight into the numerical methods which need to be called upon in order to implement the proposed methodology (solve equation (8) and Problem 1) is also required and is a subject of an ongoing research.

References


Tripp, M. H., Bradley, H. L., Devitt, R., Orros, G. C., Overton, G. L., Pryor, L. M. and Shaw,
