Extremes in operational risk management

E. A. Medova and M. N. Kyriacou
Centre for Financial Research
Judge Institute of Management
University of Cambridge

Abstract

Operational risk is defined as a consequence of critical contingencies most of which are quantitative in nature and many questions regarding economic capital allocation for operational risk continue to be open. Existing quantitative models that compute the value at risk for market and credit risk do not take into account operational risk. They also make various assumptions about 'normality' and so exclude extreme and rare events. In this paper we formalize the definition of operational risk and apply extreme value theory for the purpose of calculating the economic capital requirement against unexpected operational losses.

March 2001

1. Introduction

Highly publicized events such as those at LTCM, Barings and Sumitomo have all involved mismanagement leading to extraordinary losses and raising concerns about financial instability at international levels. As a result, along with the established capital charges for market and credit risks, the Basle Committee on Banking Supervision is proposing an explicit capital charge to guard the banks against operational risks. The response from the banks has been an increasing number of operational risk management initiatives with corresponding efforts to formulate a framework for capital allocation for operational risk. This paper contains a model for calculating the economic capital against extreme operational risks which is our contribution to quantification of operational risk.

One of the first definitions of operational risk (British Bankers' Association, 1997) was specified by a list of possible causes [4]:

The risks associated with human error, inadequate procedures and control, fraudulent and criminal activities;
the risks caused by technological shortcomings, system breakdowns;
all risks which are not 'banking' and arising from business decisions as competitive action, pricing, etc;
legal risk and risk to business relationships, failure to meet regulatory requirements or an adverse impact on the bank's reputation;
'external factors' include: natural disasters, terrorist attacks and fraudulent activity, etc'.

After four years of intensive debate on what constitutes an operational risk the current Basle proposal defines operational risk as [2]:

'*Operational risk is the risk of direct or indirect loss resulting from inadequate or failed internal processes, people and systems or from external events'.*

Strategic and reputational risks are not included in this new definition, but as before it focuses on the causes of operational risk and is open to endless discussion about the detailed definition of each loss category. The 'semantic Wild West' of operational risk [15] is still with us and the view of operational risk as ‘everything not covered by exposure to credit and market risk’ remains the one most often used by practitioners.

Our own operational risk study started with a search for a definition suitable for quantitative modelling. The resulting modelling approach is presented in Section 2. According to Basle: ‘A capital charge for operational risk should cover unexpected losses. Provisions should cover expected losses.’ The Committee clarifies the complex issues of risk management by adopting a “three-pillared” approach. The first pillar is concerned with capital allocation, the second pillar with supervision and controls and the third with transparency and consistency of risk management procedures. With the view that statistical analysis of loss data and consistency of modelling techniques may be considered respectively as parts of Pillars 2 and 3, we adopt the ‘practitioners’ definition of operational risk and propose a model for the capital allocation of Pillar 1. We also assume that provisions and improvements in management control (Pillars 2 and 3) will cover low value frequently occurring losses and we concentrate here on extreme and rare operational risks. A definition of operational risk suitable for quantitative modelling and our framework for economic capital allocation are presented in Section 2. This stochastic model is based on results from extreme value theory and in Section 3 we review key results on stable distributions and the classical theory of extremes. In Section 4 we detail our model and discuss related implementation issues. A Bayesian hierarchical simulation method is applied to the parameters estimation of extreme distributions from small-sized samples. The method also provides a more transparent assessment of risk by taking into account data on losses due to different risk factors or business units. We illustrate our operational risk framework on an example of anonymous European bank during the period of the Russian Crisis in Section 5 and draw conclusions and sketch future directions in Section 6.

2. Firm-wide operational risk management

Market or credit risk definitions came naturally from specific businesses, effectively market trading, lending or investment, with the corresponding consistent probabilistic definition of the *value at risk* (VaR). Operational risk definitions on the other hand are based on an identification of causes whose consequences are often not measurable. Such differences in defining types of risk result in segregated capital allocation rules for operational risk. Yet the importance of integration of all forms of risk are obvious.

Recall that VaR provides a measure of the market risk of a portfolio due to adverse market movements under *normal* market conditions and is expressed here in return terms

$$\mathrm{P}\ (\boldsymbol{R} < \mathrm{VaR}) = \int_{-\infty}^{VaR} N(R)dR = \pi, \tag{1}$$

as

where the return $\boldsymbol{R}$ is the normalised portfolio value change over a specified time horizon, N denotes a suitable normal density and π is a probability corresponding to a one-sided confidence level (typically 5% or 1%). More generally, N is replaced by an appropriate return density $f_{\boldsymbol{R}}$, for example, one which is obtained by simulation.

Similarly, credit ratings correspond to normal credit conditions, for example with default corresponding to a rating below CCC. In credit modelling the *default point* threshold is difficult to formalize as it depends on the evolution of the institution's assets (for a discussion, see M. Ong [20]). The 'value of the firm' framework as implemented by CreditMetrics defines a series of levels of the firm's assets which determine the credit rating of the firm. In Ong's interpretation: 'Assuming that asset returns denoted by the symbol $\boldsymbol{R}$, are normally distributed with mean μ and standard deviation σ, the generalization concerning the firm's credit quality immediately translates to the slicing of the asset returns distribution into distinct bands. Each band, representing the different threshold levels of asset returns, can be mapped one-to-one to the credit migration frequencies in the transition matrix.' Thus the firm's default probability expressed in terms of its asset return distribution is given by

$$\text{P}\,(\,\boldsymbol{R} < \text{CVaR}) = \int_{-\infty}^{CVaR} N(R)dR = \rho < \pi. \tag{2}$$

Again, more general or empirical (historical) distributions might be substituted for the Gaussian in (2) to represent ‘normal’ market conditions more accurately.

One might thus naturally ask how the definition of “normality” relates to operational risk and to the problem of internal bank controls and external supervision. These questions are critical when a specific loss event happens, particularly when it is related to extreme losses. As market, credit and operational risks become entangled at the time of occurrence of large losses, it is important that an operational risk analyst deals with both market and credit risk management without double-counting. While risk capital is generally understood as a way of protecting a bank against “unexpected” losses – expected losses are covered by business-level reserves – it is not clear as to what degree risk capital should be used to cover the most extreme risks. In an attempt to answer these questions we construct a framework that allows the allocation of capital against extreme operational losses while identifying the roles of credit and market risks in their occurrence.

Let us assume that a bank’s market and credit risk management is informed by quantitative models that compute the value at risk for market risk and credit risk and that allocate economic capital to these risks. It is clear that such capital allocation is not sufficient to cover unexpected losses due to natural disasters, fraudulent activities and human errors. Currently used models do not take into account operational risks. For example, VaR models allocate capital ‘under normal market conditions’ and so exclude extreme or rare events such as natural disasters and major social or political events. As a consequence, inadequate models contribute to operational losses as a part of an ‘inadequate internal process’.

The first step in operational risk management should be a careful analysis of all available data to identify the statistical patterns of losses related to identifiable risk factors. Ideally, this analysis would form part of the financial surveillance system of the bank. In the future perhaps such an analysis might also form part of the duties of bank supervisors. In other words at a conceptual level such an analysis relates to the third of the Basle Committee's three pillars. The important point is that this surveillance is concerned with the identification of the "normality" of business processes. In statistical terms it means a fundamental justification of the Gaussian or *normal* model to describe the central part of the distribution which does not allow for large fluctuations in data. The identification of suitable market and credit risk models suitable for the *tail* events forms a natural part of an operational risk assessment. It allows an analyst to classify a bank's losses into two categories:

(1) *significant* in value but *rare*, corresponding to *extreme* loss event distributions;

(2) *low* value but *frequently* occurring, corresponding to '*normal*' loss event distributions.

Thus an analysis of profit and loss data and the verification or rejection of the assumption of normality may both be considered as the part of the (usually internal) risk supervisory process. We take the view that over time control procedures will be developed by a financial institution for the reduction of the low value/frequent losses and for their illumination and disclosure -- the second pillar of the Basle approach. These control procedures, and any continuing expected level of losses, should be accounted for in the operational budget.

Any deviation from the normality assumed, or increased volatility in the markets, will tend to underestimate market value at risk. Similarly, under normal conditions for credit risk, which corresponds to credit ratings higher than BBB, credit models provide measures for credit risk. This allows us to assume that only losses of large magnitude need be considered for operational risks. With the view that control procedures verify the assumptions of internal market and credit models, and that losses within the limits of market and credit value at risk can be accommodated, we assume that only losses of larger magnitude need be considered for operational risk capital provision. Hence we adopt the accepted practice of defining operational risk as 'everything which is not market or credit risk' and *assume operational losses to be in the category of losses which are larger than those due to market or credit risks under normal market conditions.*

As all forms of risk are driven by the same fundamental market conditions, capital allocation for market, credit risks and operational risk must be derived from the same profit and loss distribution simultaneously[1]. Therefore for integrated profit and loss data at the firm- or business unit-level the following thresholds for losses are obtained from market and credit risk models as:

-- the unexpected loss level due to market risk, denoted by u_{VaR_π}, which is exceeded with probability π

-- the level of loss due to both credit and market risks, denoted u_{CVaR_ρ} which is exceeded probability $\rho \le \pi$, so that $u_{CVaR_\rho} \le u_{VaR_\pi}$.

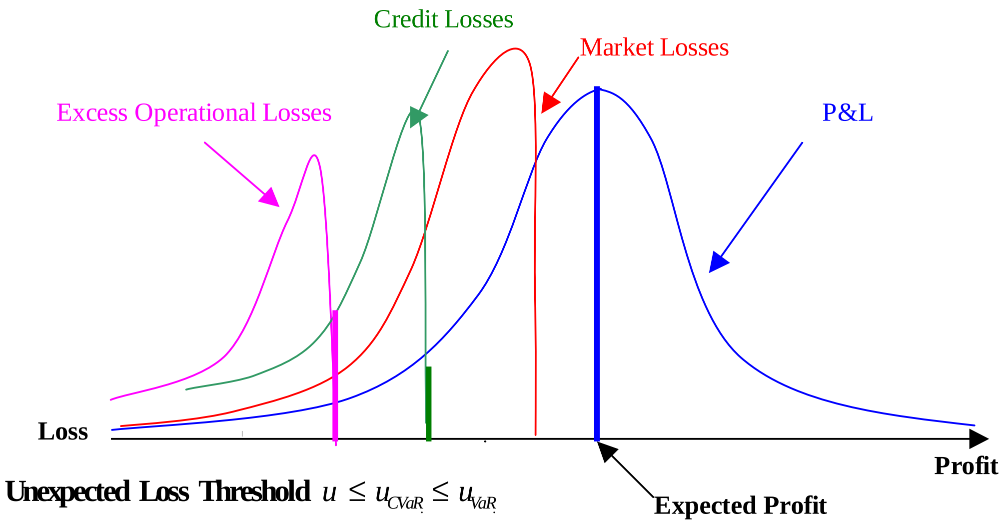

Figure 1 *Decomposition of the loss-tail of a Profit & Loss distribution into its three loss-types (market, credit and operational losses) and definition of the threshold for extreme operational losses.*

---

[1] This conceptual view of total risk modelling does not necessarily mean simultaneous implementation of market, credit and operational risk model components.

Losses beyond the u_{CVaR_ρ} level, or so called *unexpected losses,* are assumed to belong to the operational risk category. Therefore extreme operational losses are modelled as excesses over both market and credit losses on the P&L distribution as shown in Figure 1, with the risk measures corresponding to appropriate approximating distribution. The required capital allocation for operational risk will be derived from the parameters of the *asymptotic* distribution of *extremes* of profit and loss.

For the purpose of operational risk management we obtain an *unexpected loss threshold* u obtained from the operational risk model to be developed (see Section 4). We shall suppose that u_{CVaR_ρ} level approximately equals to this threshold u. Relations between the thresholds for market and credit risk may be obtained by variety of methods as implemented by internal models. These levels should be re-examined with respect to the overall implementation of risk management procedures according to the definitions of 'expected' and 'unexpected' losses.

3. Stable random variables and extreme value theory

Our formalism in defining operational risk focuses on tail events. But consistency in estimation of profit and loss distributions at different levels of a financial institution and at different time scales is difficult to achieve and any successful implementation would rely on approximation and heuristics. The asymptotic theories of sums and maxima of random variables are thus of crucial importance for risk management. Here we recall some definitions and principal results used in our proposed procedure for operational risk capital allocation.

The summary effects of daily fluctuations in price return or of a portfolio is well captured by a limiting *normal* distribution for data whose underlying distribution has finite variance, but this normal limit is often inadequate for highly variable data. *Stable* distributions approximate the distribution of sums of *independent identically distributed* (i.i.d.) random variables with infinite variance and include the Gaussian as special case.

There are many famous monographs on asymptotic theory for sums dating from the 1950s: Gnedenko and Kolmogorov (1954) [11], Feller(1966) [9], Mandelbrot (1982) [18], Samorodnitsky and Taqqu (1990) [22].

Many results of the *asymptotic theory for sums* (or *central limit theory*) have their complements in the *asymptotic theory of extreme order statistics* known as *extreme value theory* (EVT). EVT has been applied in engineering, hydrology, insurance and currently applies to financial risk management. Some of most useful references are: Galambos (1978) [10], Leadbetter *et all* (1983) [16], Du Mouchel (1983) [7], Castillo (1988) [5], Embrechts, Kluppelberg & Mikosch (1997) [8], R. Smith (1985, 1990, 1996) [24-28], Danielson and de Vries (1997) [6] and McNeil and Saladin (1997) [19].

One of the fundamental problems of risk management is identification of the functional form of a profit and loss distribution. Simulation methods will 'construct' such a distribution without requiring an analytic form, but this usually involves a complex implementation and considerable computing time.

Every random profit/loss $\boldsymbol{X}$ has associated with it a distribution function with four basic parameters that have physical or geometric meaning. These are the *location* μ, the *scale* σ, the *tail index* α, or equivalently the *shape* $\xi=1/\alpha$, and the *skewness* β.

Stable distributions have a number of equivalent definitions in terms of the 'stability' property, the domain of attraction, or as a special subclass of the infinitely divisible distributions. Most important for applications is the fact that any α-stable random variable can be expressed as a convergent sum of random variables indexed by the arrival times of a Poisson process (for definitions, see[22]).

A random variable $\boldsymbol{X}$ is said to have an α– *stable* distribution if for any n ≥ 2 there is a positive number c_n and a real number d_n such that

$$\boldsymbol{X}_1 + \boldsymbol{X}_2 + ... + \boldsymbol{X}_n \stackrel{d}{=} c_n \boldsymbol{X} + d_n \qquad (3)$$

where $X_1, X_2, ..., X_n$ are independent copies of X and $c_n = n^{1/\alpha}$ for some number α, $0 < \alpha \le 2$, called the *index of stability*.

Stable distributions are suitable to modelling a wide class of empirical distributions. In fitting such distributions to heavy-tailed samples, the parameter α measures the thickness of tails and finiteness of the moments of the distribution of X. The distribution functions of stable random variables are often not available in a closed form with the exception of a few special cases. Feller [9] describes stable distributions analytically by specifying their characteristic function given by

$$\varphi_X(t;\alpha,\beta,\mu,\sigma) := \mathrm{E}[\exp(itX) \mid \alpha,\beta,\mu,\sigma] = \begin{cases} \exp\left(i\mu t - \sigma^{\alpha}|t|^{\alpha}\left(1 - i\beta\, sign(t)\, tan\left(\frac{\pi\alpha}{2}\right)\right)\right) & ,\text{if } \alpha \neq 1 \\ \exp\left(i\mu t - \sigma|t|\left(1 + i\beta\, sign(t)\frac{2}{\pi}\log|t|\right)\right) & ,\text{if } \alpha = 1 \end{cases} \quad (4)$$

for $-\infty < t < \infty$, $0 < \alpha \le 2$, $\sigma \ge 0$, $-1 \le \beta \le 1$ and μ real, where E[·] denotes expectation. A r.v. X is has a stable distribution if, and only if, it has a *domain of attraction*, i.e. if there is a sequence of independent identically distributed (i.i.d.) random variables $Y_1, Y_2, ...$ and sequences of positive numbers $\{d_n\}$ and real numbers $\{a_n\}$ such that

$$\frac{Y_1 + Y_2 + ... + Y_n}{d_n} + a_n \overset{d}{\Rightarrow} X \,. \quad (5)$$

where $\overset{d}{\Rightarrow}$ denotes convergence in distribution as $n \to \infty$. In general $d_n := n^{1/\alpha} h(n)$, where $h(x)$, $x > 0$, is a *slowly (or regular) varying function at infinity*, i.e. for sufficiently large u>0

$$\lim_{x \to \infty} h(ux)/h(x) = 1 \,. \quad (6)$$

When X is Gaussian, i.e. α=2, and $Y_1, Y_2, ...$ are i.i.d. with finite variance, then (5) is the statement of the *Central Limit Theorem* (CLT). Generalizations of the CLT involve *infinitely divisible* random variables [11]. The family of *infinitely divisible distributions* includes the stable distributions. A random variable is *infinitely divisible* if, and only if, for every natural number n it can be represented as the sum

$$\boldsymbol{X} = \boldsymbol{X}_{n_1} + \boldsymbol{X}_{n_2} + ... + \boldsymbol{X}_{n_n} \qquad (7)$$

of n i.i.d. random variables.

.

Equivalently, for every natural number n there exists a characteristic function given by $\varphi_{\boldsymbol{X}}(t)$ whose nth power is equal to the characteristic function φ_{x_n} of $\boldsymbol{X}$, i.e.

$$\varphi_x = (\varphi_{x_n})^n . \qquad (8)$$

In terms of distribution functions, the distribution function F of $\boldsymbol{X}$ is given by a convolution of corresponding F_n's

$$F = F_n^{n*} := F_n * F_n * ... * F_n . \qquad (9)$$

Let $\boldsymbol{X}_1, ..., \boldsymbol{X}_n$ represent i.i.d. random variables with distribution function F and define their *partial sum* by $\boldsymbol{S}_n = \boldsymbol{X}_1 + \boldsymbol{X}_2 + ... + \boldsymbol{X}_n$ and their *maximum* by $\boldsymbol{M}_n = \max(\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_n)$.

It can be shown [8, 9,12] that regular variation in the tails (6) and infinite divisibility (7) together imply *subexponentiality* of a distributions, i.e. for $n \geq 2$

$$\lim_{x \to \infty} \frac{\bar{F}^{n^*}(x)}{\bar{F}(x)} = n, \qquad (10)$$

where, for example, $\bar{F} := 1 - F$ denotes the *survivor function* corresponding to F.
It follows that

$$P(\boldsymbol{S}_n > x) \sim P(\boldsymbol{M}_n > x) \qquad \text{as} \quad x \to \infty . \qquad (11)$$

Thus behaviour of the distribution for a sum in its tail may be explained by that of its maximum term, leading to many complementary results to those of central limit theory for the max-stable distributions studied in extreme value theory.

The possible limiting distributions for the maximum $\boldsymbol{M}_n$ of *n* i.i.d. random variables are identified as the class of *max-stable* distributions, the maximum domain of attraction is analogous to the domain of attraction and the Poisson representation mentioned above is the main theoretical tool for studying the process of exceedances of a specified level.

The current theoretical foundations of EVT are given in Embrecht, Kluppelberg and Mikosch's book [8]. Since [8] and R. Smith's papers [24-28] focus on applications to insurance and risk management, we will only state here results required for modelling operational risk.

The Fisher-Tippett theorem proves the convergence of the sample maxima to the non-degenerate limit distribution $H_{\xi;\mu,\sigma}$ under some linear rescaling such that for $c_n>0$ and d_n real, $c_n^{-1}(M_n - d_n) \xrightarrow{d} H_{\xi,\mu,\sigma}$, as the *sample size* n increases, i.e. for $-\infty < x < \infty$

$$P\left[\left(\frac{M_n - d_n}{c_n}\right) \leq x\right] \longrightarrow H_{\xi;\mu,\sigma}(x) \qquad \text{as } n \to \infty. \tag{12}$$

Three classical extreme value distributions of normalised sample maxima which are included in this representation are the *Gumbel, Frechet* and *Weibull* distributions. The generalised extreme value distribution (GEV) $H_{\xi;\mu,\sigma}$ provides a representation for the non-degenerate limit distribution of normalised maxima with *shape parameter* ξ

$$H_{\xi;\mu,\sigma}(x) = \begin{cases} \exp\left[-\left(1+\xi\dfrac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right] & \text{if } \xi \neq 0, \quad 1+\xi\dfrac{x-\mu}{\sigma} > 0 \\ \exp\left[-\exp\left(-\dfrac{x-\mu}{\sigma}\right)\right] & \text{if } \xi = 0 \quad . \end{cases} \tag{13}$$

For the case of α-max-stable distributions, the shape parameter ξ satisfies $1/2 \leq \xi = 1/\alpha < \infty$ and determines the existence of moments. For the Gaussian case $\alpha=1/\xi=2$, while for $\xi>1$ the distribution has no moments finite.

Modelling worst case losses will involve fitting an extreme value distribution. This can be done by grouping the data into epochs (month, years, etc) and using its maximum (minimum) over an epoch as one representative of a GEV. However the longer the epoch the larger loss of data with this approach. The central idea of a method based on *exceedances* is to avoid such a loss of information and to consider all data which lie above a given *threshold* value [16, 17, 22, 24].

Given an i.i.d. sequence of random variables $\boldsymbol{X}_1,\ldots, \boldsymbol{X}_n$ drawn from an underlying distribution F, we are interested in the distribution of *excesses* **Y**:= **X**-u over a high *threshold u*. We define an *exceedance* of the level u if in the event $\boldsymbol{X}=x$ we have $x > u$. The distribution of *excesses* is given by the conditional distribution function in terms of the tail of the underlying distribution F as

$$F_u(y) := P(\boldsymbol{X} - u \le y \mid \boldsymbol{X} > u) = \frac{F(u+y) - F(u)}{1 - F(u)} \quad \text{for } 0 \le y < \infty. \tag{14}$$

The limiting distribution $G_{\xi,\beta}(y)$ of excesses as $u \to \infty$ is known as the *generalised Pareto distribution* (GPD) with *shape parameter* ξ and *scale parameter* β given by

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \xi \dfrac{y}{\beta}\right)^{-1/\xi} & \xi \ne 0 \\ 1 - \exp\left(-\dfrac{y}{\beta}\right) & \xi = 0 \end{cases} \qquad \text{where} \quad y \in \begin{cases} [0, \infty] & \xi \ge 0 \\ [0, -\dfrac{\beta}{\xi}] & \xi < 0. \end{cases} \tag{15}$$

Pickands [21] has shown that the GPD is a good approximation of F_u in that

$$\lim_{u \to x_F} \sup_{0 \le y \le y_F} | F_u(y) - G_{\xi,\beta(u)}(y) | = 0, \tag{16}$$

where x_F (possibly infinite) is the right hand end point of the support of the distribution given by F and $y_F := x_F - u,$ for some positive measurable function of the threshold u

given by *β(u)*, provided that this distribution is in the max-domain of attraction of the generalized extreme value distribution.

For $\xi > 0$, the tail of the density corresponding to F decays slowly like a *power* function and F belongs to the family of *heavy-tailed* distributions that includes among others the Pareto, log-gamma, Cauchy and t-distributions. Such distributions may not possess moments. Indeed, for the GPD with $\xi > 0$, E[$\mathbf{Y}^k$] is infinite for $k > 1/\xi$, so that for $\xi > 1$ the GPD has no mean and for $\xi > 1/2$ it has infinite variance. For $0 \le \xi \le 1/2$, the tail of F decreases *exponentially fast* and F belongs to the class of *medium-tailed* distributions with two moments finite comprising the normal, exponential, gamma and log-normal distributions. Finally, for $\xi < 0$ the underlying distribution F is characterised by a *finite* right endpoint and such *short-tailed* distributions as the uniform and beta.

Financial losses and operational losses in particular are often such that underlying extremes tend to increase without bound over time rather than clustering towards a well-defined upper limit. This suggests that the shape parameter for the GPD estimated from such data can be expected to be non-negative.

Equation (14) may be re-written in terms of survivor functions as

$$\bar{F}(u+y) = \bar{F}(u)\ \bar{F}_u(y). \tag{17}$$

The survivor function $\bar{F}(u)$ may be estimated empirically by simply calculating the proportion of the sample exceeding the threshold u, i.e. $\bar{F}(u) = N_u/n$. The corresponding q-quantiles of the underlying distribution F are given by

$$x_q = u + \frac{\beta}{\xi}\left[\left(\frac{n}{n_u}(1-p)\right)^{-\xi} - 1\right] \qquad \xi \ge 0$$

$$x_q = u - \frac{\beta}{\xi} \qquad \xi < 0 \tag{18}$$

and the mean of the GPD or *expected excess* function equals

$$E(\boldsymbol{X}-u \mid \boldsymbol{X}>u)=\frac{\beta+\xi u}{1-\xi} \qquad \text{for } \xi<1, \qquad u>0. \tag{19}$$

These may be estimated by replacing the shape and scale parameters by their sample estimates [19, 26].

4. Stochastic model for measuring of operational risk

Occurrences of extreme losses over time may be viewed as a point process $\boldsymbol{N}_u$ of exceedances which converges weakly to a Poisson limit [7, 17, 19]. The GPD provides a model for the excesses over an appropriate threshold *u*, while the limit Poisson approximation helps to make inferences about the *intensity* of their occurrence. The resulting asymptotic model is known as the *peaks over threshold* (POT) model [8, 16, 17].

For *u* fixed the parameters of the POT model are the *shape* ξ and the *scale* β_u parameters of the GPD and the Poisson *exceedance rate* λ_u. In terms of these parameters, the alternative *location* μ and *scale* σ parameters are given respectively by

$$\mu = u + \beta\xi^{-1}\left(\lambda^{\xi} - 1\right) \tag{20}$$

$$\sigma = \beta\,\lambda^{\xi}\,. \tag{21}$$

Conversely, the location and alternative scale parameters determine the scale parameter and exceedance rate respectively as

$$\beta_u = \sigma + \xi(u-\mu) \tag{22}$$

$$\lambda_u := \left(1+\xi\,\frac{(u-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}. \tag{23}$$

The POT model captures both aspects of operational risk measures – severity and frequency of loss – in terms of excess sizes and corresponding exceedance times. The choice of threshold must satisfy the asymptotic convergence conditions in (11) and (16), i.e. be large enough for a valid approximation, but when u is too high classical parameter estimators for ξ and β_u may have too high a variance due to the small size of exceedances. In the literature [6, 7, 8, 19, 24-28] various techniques have been proposed for a statistically reliable choice of threshold. We will assume that the chosen threshold u satisfies a 'bias versus variance trade-off' optimality condition. In our operational risk framework such a u may be termed an *unexpected loss threshold*. Since in this threshold method all excess data is used for parameter estimation, the intensity is measured in the same time units as the given underlying profit and loss data.

Justified by the theoretical results presented from the asymptotic theory of extremes and based upon the point process representation of exceedances given by the POT model, we are now in a position to quantify operational risk. In summary, the operational risk measures are the expected severity and intensity of losses over a suitably chosen threshold u for this model estimated from appropriate profit and loss data.

- Severity of the losses is modelled by the GPD. The expectation of excess loss distribution, i.e. *expected severity* is our coherent risk measure [1] given by

$$E(\boldsymbol{X}-u \mid \boldsymbol{X}>u)=\frac{\beta_u+\xi u}{1-\xi} \quad \text{with } \beta{:=}\sigma+\xi(u-\mu). \tag{24}$$

- The number of exceedances $\boldsymbol{N}_u$ over the threshold u and the corresponding exceedance times are modelled by a Poisson point process with *intensity* (frequency per unit time) given by

$$\lambda_u := \left(1+\xi\,\frac{(u-\mu)}{\sigma}\right)^{-\frac{1}{\xi}} \quad . \tag{25}$$

- *Extra* capital provision for operational risk over the *unexpected loss threshold u* is estimated as the *expectation of the excess loss* distribution (expected severity) scaled by the *intensity* λ_u of the Poisson process, *viz*.

$$\lambda_u \; E(\boldsymbol{X} - u \,|\, \boldsymbol{X} > u) = \lambda_u \; \frac{\beta_u + \xi u}{1 - \xi}, \tag{26}$$

where u, β, ξ and λ are the parameters of the POT model and time is measured in the same units as data collection frequency, e.g. hours, days, weeks, etc. (Note that usually β_u and λ_u will be expresses in terms of the μ and σ as in (24) and (25).)

- The *total* amount of capital provided against extreme operational risks for the time period *T* will then be calculated by

$$u_T + \lambda_u \; T \; E(\boldsymbol{X} - u \,|\, \boldsymbol{X} > u) = u_T + \lambda_u \; T \; \frac{\beta + \xi u}{1 - \xi}, \tag{27}$$

where u_T may in the first instance be considered to be equal to *u* under the assumption of max-stability.

In general this threshold value u_T over a long horizon T should be adjusted with respect to the time horizon appropriate to integrated risk management and to the thresholds obtained from market and credit models. This is a topic of our current research. The accuracy of our economic capital allocation (26) depends of course on both the correct choice of threshold and accurate estimates of the GPD parameters.

Extreme losses are rare by definition and consequently the issue of small data sets becomes of crucial importance to the accuracy of the resulting risk measures. In addition, operational risk data sets are not homogeneous and are often classified into several subsamples, each associated with a different risk factor or business unit. The conventional *maximum likelihood* (ML) estimation method performs unstably when it is applied to small or even moderate sample sizes, i.e. less than fifty observations. *Bayesian simulation* methods for parameter estimates allow one to overcome problems associated with lack of data through intensive computation.

The *Bayesian hierarchical Markov Chain Monte Carlo* (MCMC) *simulation* model [3, 23] treats uncertainties about parameters by considering them to be random variables (Bayesian view) and generates (simulates) an empirical parameter distribution approximating the conditional posterior parameter distribution given the available loss data. A Bayesian hierarchy is used to link the posterior parameters of interest through the use of prior distribution hyperparameters – in our case estimates of the parameters are linked through the data on different risk types. Our computational procedures were built on R. Smith's statistical procedures and algorithms for GPD assumption verification and corresponding threshold choice [24] using the special library for extreme value statistics of Splus software. Stability of parameter estimation in the presence of small samples is achieved by taking as estimates the medians of the possibly disperse empirical marginal posterior parameter distributions.

Operational loss data may be organized into a matrix according to loss type and to business unit as in Table 1 (in which for simplicity only a single cell entry is shown).

Business unit **Loss factor**	1	…	j	…	N	**Firm-wide**
Technology failure	X_1^1		X_1^j		X_1^N	$X_1^1, X_1^2, ..., X_1^N$
Fraud	X_2^1		X_2^j		X_2^N	$X_2^1, X_2^2, ..., X_2^N$
…		…		…		…
External event	X_n^1		X_n^j		X_n^N	$X_n^1, X_n^2, ..., X_n^N$
Total	$X_1^1, X_2^1, ..., X_n^1$	…	$X_1^j, X_2^j, ..., X_n^j$	…	$X_1^N, X_2^N, ..., X_n^N$	$X^1, X^2, ..., X^N$

Table 1 *Firm-wide matrix of operational losses.*

The simulated values of the parameters of the POT model are used for calculation of capital provision according to formulas (25) and (26). For overall capital allocation at the top level of the bank, we hope to reduce the overall assessed capital allocation due to portfolio diversification effects and to identify the high-risk factors for specific business units of the firm.

The procedure could can be applied to one business unit across different loss types. Alternatively, it may be applied to one type of loss across all business units as will be demonstrated below in Section 6. Conceptually, both loss factor and business unit dimensions can be simultaneously accommodated at the cost of increased complexity – a topic of our current research. Essentially, the technique is to apply computational power to substitute for insufficient amounts of data, but its empirical estimation efficiency when back-tested on large data sets is surprisingly good.

5. Simulation of peaks over threshold model parameters by MCMC

Bayesian parameter estimation treats uncertainties about parameters by considering parameters to be random variables possessing probability density functions. If the *prior density* $f_{\theta|\psi}$ of the random parameter vector θ is parametric, given a vector of random *hyperparameters* ψ, and of a mathematical form such that the calculated *posterior density* $f_{\theta|X_1,\ldots,X_n,\psi} := f_{\theta|\psi+}$ is of the same form with new hyperparameters ψ^+ determined by ψ and the observations $X_1,\ldots, X_n$, then we say that $f_{\theta|\psi}$ is a parametric family of densities *conjugate prior* to the *sampling density* $f_{X|\theta}$.

The *Bayesian hierarchical model* provides a transparent risk assessment by taking into account the possible classification of the profit and loss sample according to loss data subtypes or *classes*, i.e. risk factors or business units, as well as the aggregate. In this model the prior density for the hyper-parameters ψ is *common* to all loss subtype prior densities for the parameters θ. The *hyper-hyper parameters* φ are chosen to generate a vague conjugate prior indicating a lack of information on the hyper-parameters' prior distribution before the excess loss data is seen. Thus we have a Bayesian hierarchical decomposition of the posterior parameter density $f_{\theta|\mathbf{X},\psi}$ given the observations and the *initial* hyper-hyper-parameters φ as

$$
\begin{aligned}
f_{\theta|X,\psi}(\theta \mid X,\psi) &\propto f_{X|\theta}(X \mid \theta)\ f_{\theta|\psi}(\theta \mid \psi) f_{\psi}(\psi \mid \varphi) \\
&\propto f_{X|\theta}(X \mid \theta)\ f_{\psi|\theta}(\psi \mid \theta,\varphi) \\
&\propto f_{X|\theta}(X \mid \theta)\ f_{\psi}(\psi \mid \varphi^{+})\ ,
\end{aligned} \tag{27}
$$

where $\propto$ denotes proportionality (up to a positive constant). We may thus perform the Bayesian update of the prior parameter density $f_{\theta} \propto f_{\theta|\psi}\ f_{\psi}$ in two stages -- first updating the hyper-hyper-parameters φ to φ^{+} conditional on a given value of θ and then computing the value of the corresponding posterior density for this θ given the observations X. Figure 2 depicts schematically the relationships between the 3 parameter levels and the excess loss observations for each risk class. Note that even though the prior specification of parameters for individual risk classes is as an independent sample from the *same* hyperparameter Gaussian prior distribution, their posterior multivariate Gaussian specification will *not* maintain this independence given observations which are statistically dependent.

The Bayesian *posterior density* $f_{\theta|X,\psi}$ may be computed via *Markov chain Monte Carlo* (MCMC) simulation [23, 27, 28]. The idea, which goes back to Metropolis, Teller *et al* and the hydrogen bomb project, is to simulate sample paths of a Markov chain. The states of the chain are the values of the parameter vector θ and its visited states converge to a stationary distribution which is the Bayesian joint posterior parameter distribution $f_{\theta|X,\psi}$ (termed the *target* distribution) given the loss data X and a vector ψ of hyperparameters as discussed above. In this context, a *Markov chain* is a discrete time continuous state stochastic process whose next random state depends statistically only on its current state and not on the past history of the process. Its random dynamics are specified by the corresponding state transition probability density. In this application the parameter vector state space of the chain is discretised for computation in order to create a parameter histogram approximation to the required multivariate posterior parameter distribution.

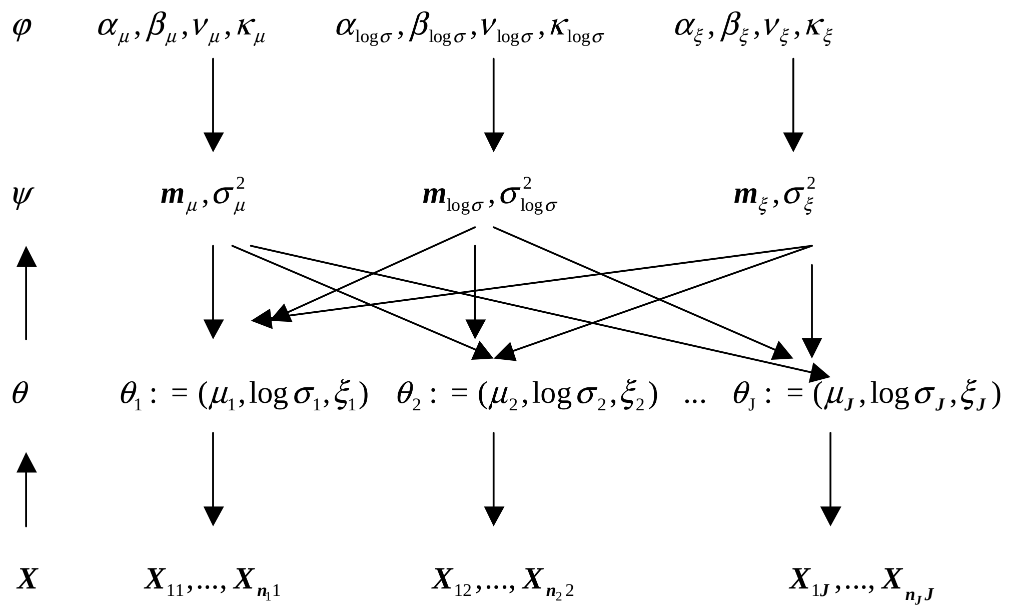

Figure 2 *Hierarchical Bayesian model parameter and observation dependencies conditional on their hyperparameters.*

For our application, the parameter vector θ represents the generalized Pareto distribution (GPD) parameters of interest $\{\mu_j, \log\, \sigma_j, \xi_j : j = 1,2\ldots J\}$ for the $j=1,\ldots,J$ data classes (business units or risk factors) and the hyperparameter vector ψ consists of $\{ m_{\mu}, s^2_{\mu}, m_{log\sigma}, s^2_{log\sigma}, m_{\xi}, s^2_{\xi} \}$ which are the parameters of a common (across all business units) multivariate Gaussian prior distribution of the GPD parameters. To implement the strategy, *Gibbs sampling* and the *Metropolis-Hastings* algorithm [3] are used to construct the Markov chain possessing our specific target posterior distribution as its stationary distribution. This target distribution is defined by standard Bayesian calculations in terms of the peaks over threshold likelihood function and appropriate prior distributions. Running the Markov chain for very many transitions (about 1M) produces an empirical parameter distribution that is used to estimate the posterior density $f_{\theta|X,\psi}$.

These MCMC dynamical methods generate the sequence $\{\theta_j^0,\ \theta_j^1,\ \theta_j^2,\ldots\}$of parameter estimates $\theta_j=\{\mu_j,\ log\sigma_j,\ \xi_j\},\ j=1,2,\ldots J\}$ for each data class with θ_j^{t+1} (for time $t\geq0$) depending solely upon θ_j^t. This process represents the traditional exchange of computational intensity for low data availability. After sufficient iterations the Markov chain will forget its initial state and converge to the stationary required posterior distribution $f_{\theta|X,\psi}$ not depending on the initial state θ_j^0 or time t. By discarding the first k (=10k) states of the chain, constituting the *burn-in period*, the remainder of the Markov chain output may be taken to be a parameter sample drawn from the high-dimensional target parameter posterior distribution.

In summary, the MCMC simulation is used to generate an empirical parameter distribution approximating the conditional posterior multivariate parameter distribution given the available loss data. A *Bayesian hierarchical* model is used to link the posterior parameters of interest through the use of common prior distribution hyperparameters. The simulation is implemented using hybrid methods and parameter estimates are taken as median values of the generated empirical marginal parameter distributions.

6. Example: Bank trading losses analysis through the Russian Crisis

We apply the framework set out above to analyse the losses of the trading activities of a major European investment bank during the period 1 October 1997 to 31 December 1998. Financial turmoil in the summer of 1998 caused by the Russian government's domestic bond default on 24 August 1998 caused losses which can be seen as *external* to the bank's normal operating conditions -- possibly in the category of unexpected large losses. In financial crises the separation of financial risks into various types (market, credit etc.) proves to be difficult and the Russian crisis is no exception. To reduce bank exposure to the consequences of such events a correct model for risk evaluation and capital provision should be identified, with the corresponding unexpected threshold level given by current historical loss data. In what follows the necessary diagnostics to test and verify the POT model assumptions for *aggregated* P&L data are first performed. Next we back-test the predictive power of the POT model in terms of the proposed capital provision estimation rule and then study its breakdown by business unit. Our data (rescaled for confidentiality reasons) contains daily P&L reports from four business unit / trading desks. Daily events are aggregated across the four desks. The *aggregated* P&L data consists of n=296 profits or losses with a net profit figure of 27,337 monetary units. They range from 2,214 loss to 1,532 profit; see Table 2 for summary statistics and Figure 3 for a time-series plot and histogram of aggregated P&L data.

Min:	-1532.394960	Mean:	-92.353455
Max:	2214.319020		
1st Qu.:	-320.839980	Median:	-119.276080
3rd Qu.:	68.261120		
Sample size:	296		
Std Dev.:	463.733057		
Excess Kurtosis:	5.047392		

Table 2 *Summary statistics for daily aggregated P&L data. Losses are positive and profits are negative.*

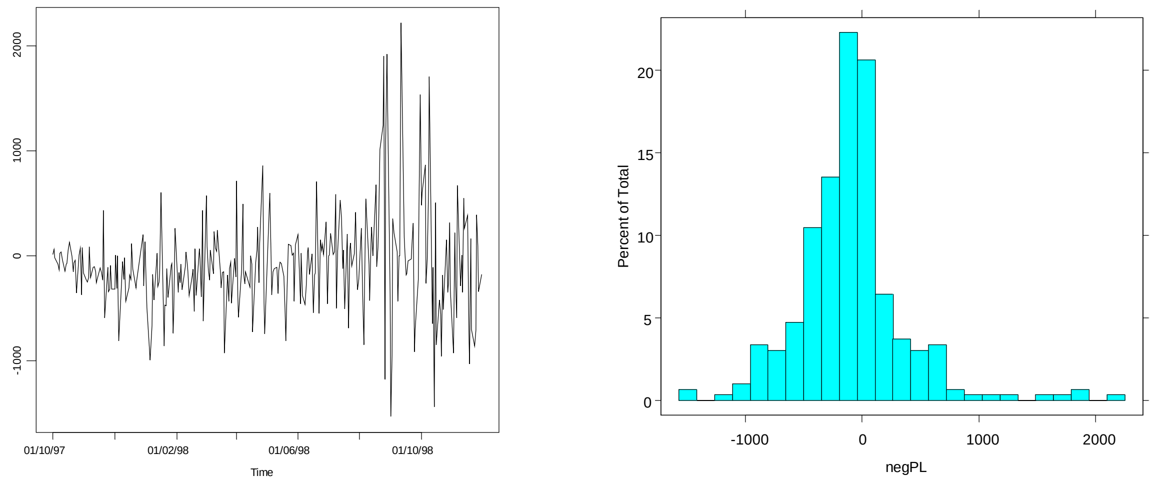

Figure 3 *Daily P&L data aggregated over the four trading desks: time-series plot (left) and histogram (right). Note that losses are positive and profits are negative.*

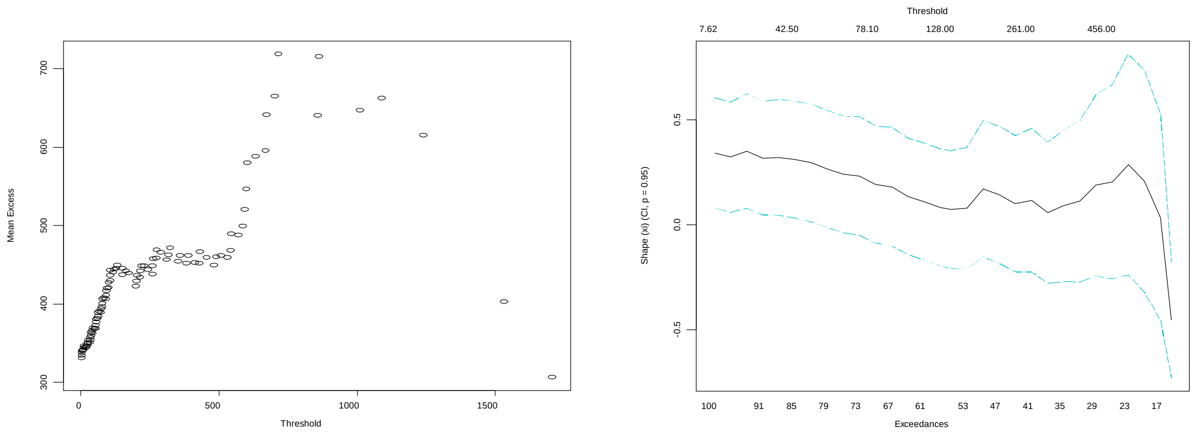

Figure 4 *Empirical mean excess plot and shape parameter ξ ML estimates for an increasing sequence of thresholds in aggregated P&L data. Dotted lines represent estimated 95% confidence intervals of ξ ML estimates.*

In Figure 4 we plot the empirical excesses of the aggregated P&L data for an increasing sequence of thresholds. The positive steep slope above a threshold of about 500 indicates a heavy loss tail. The shape parameter plot, based on maximum likelihood estimation of ξ, seems to have stable standard deviation 0.15 up to a minimum of n_u=55 exceedances. Samples of size less than n_u=55 exceedances (or equivalently thresholds higher than u=150) yield ML ξ estimates with significantly large estimated 95% confidence intervals. Hence a realistic threshold should not be set higher than u=150 when fitting the POT model with this approach.

Figure 5 shows the empirical quantiles versus a standard normal distribution and a GPD with scale parameter β =1 and shape parameter ξ = 0.25, which represents the best Q-Q plot against the GPD for various values of ξ. These Q-Q plots verify earlier observations that the loss tail is heavier than that to be expected from a normal distribution.

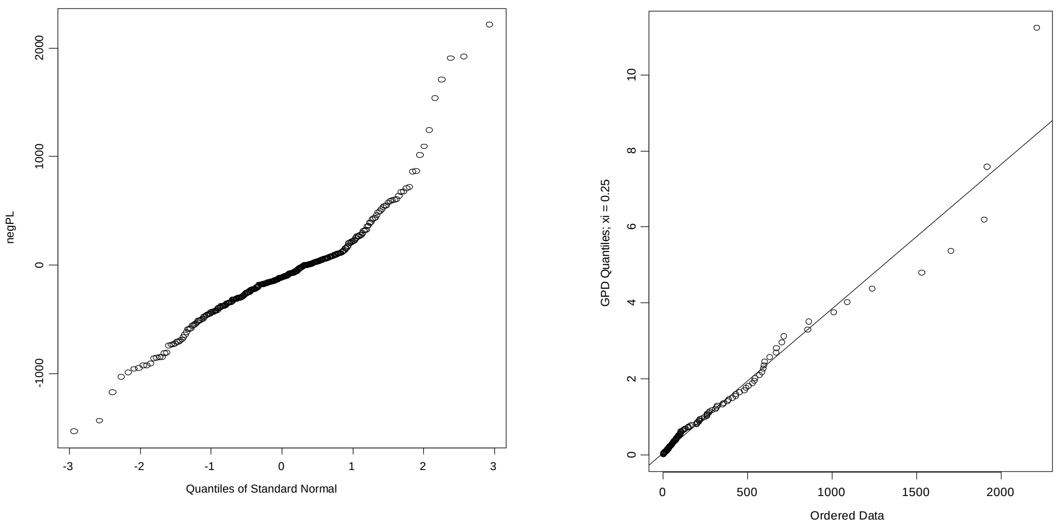

Figure 5 *Q-Q plots of aggregated P&L data. (a) Comparing the empirical quantiles (vertical axis) with the quantiles expected from a standard normal distribution (horizontal axis); losses are positive. (b) Comparing the empirical quantiles (horizontal axis) with the quantiles expected from a GPD (β=1, ξ=0.25).*

As noted above, choice of threshold should guarantee the stability of the ML estimate of the shape parameter ξ when using maximum likelihood estimation. The ML estimates of ξ together with their standard errors for an increasing sequence of thresholds, $50 \le u \le 1534$, and the corresponding MCMC Bayesian estimates of ξ based on the medians and standard deviations of the marginal posterior distributions for the same thresholds are shown in Table 3. The Bayesian estimates of the shape parameter ξ from the MCMC algorithm show relative stability in the range u=250 to 1534 (corresponding to the 15% to 1% tails of the underlying P&L empirical distribution) at a value about $\hat{\xi} = 0.53$, indicating an α-stable distribution with only a single moment finite. Moreover the Bayesian method allows estimation of the shape parameter from smaller-sized samples, less than n_u =20 exceedances, whereas the corresponding ML estimates become totally unreliable for such small samples. For example, the ML shape parameter estimate for u=600 (i.e. n_u =16 exceedances) is negative, which is not at all representative of its true value. By contrast, the Bayesian shape parameter estimates are stable even for just n_u =4 exceedances (i.e. u=1534), although, as shown in Table 3, the corresponding posterior distribution in this case is rather dispersed.

Threshold u	***Number of exceedances n_u***	***% Tail Fitted P(X>u)***	***Bayesian Shape par. $\hat{\xi}$ (posterior median estimate)***	***Maximum Likelihood Shape par. $\hat{\xi}$***
50	82	28%	**0.396 (0.195)**	0.296 (0.167)
75	72	25%	**0.311 (0.207)**	0.220 (0.163)
100	64	22%	**0.258 (0.215)**	0.154 (0.158)
150	55	18%	**0.254 (0.226)**	0.119 (0.163)
250	43	15%	**0.536 (0.268)**	0.144 (0.197)
400	30	10%	**0.520 (0.221)**	0.181 (0.261)
600	16	5%	**0.573 (0.325)**	-0.228 (0.527)
1000	8	2.7%	**0.524 (0.422)**	NA*
1534	4	1%	**0.527 (0.662)**	NA*

*NA: not available

Table 3 *Bayesian and ML estimates of the shape parameter ξ from the fitted POT model on the aggregated P&L beyond an increasing sequence of thresholds u. In parentheses are the standard errors of the ML estimates and Bayesian posterior distributions.*

The estimated standard errors of the ML estimates are merely an indication of accuracy which in fact deteriorates dangerously for higher thresholds (or, equivalently, lower tail probabilities and samples of smaller size). However, by calculating the posterior distributions of μ, σ and ξ by the Bayesian MCMC method, statistics – such as standard deviation or quantiles -- based on the entire distribution can be considered in addition to the median point estimates corresponding to absolute parameter error loss functions.
Such an EVT analysis can assist in model evaluation by more robustly identifying the heavy-tail distributions. In our example, the Bayesian estimates of the shape parameter for the aggregated data suggest that only the first moment (i.e. the mean) is finite.

Prediction of actual losses by the economic loss capital provision at firm level

To test the capital allocation rule consider five 'event' dates: 17th, 21st, 25th, 28th August 1998 and 11th September 1998. Three events are before and two after the Russian government's GKO default on 24th August 1998, *cf.* Figure 3. The fifth event-date (11th September) is selected so that the sub-sample includes the maximum historic loss as its last observation. (Note that losses are treated as positive unless stated otherwise.) For a fixed loss threshold u=150, we fit to data both the normal distribution and the POT model using both maximum likelihood and Bayesian estimation. With the threshold set at u=150 the number of exceedances for all five data sets and the full sample are equal to n_u=27, 29, 31, 33, 36 and 55 respectively. The results are illustrated in Figure 6 where the dots represent the empirical distribution function based on the full aggregated P&L data. There is a marked difference between the suggested GPD model and the normal distribution in all six experiments. The GPD approximates the excess loss distribution F_u significantly better using the Bayesian posterior median estimates of ξ, μ and σ (see Figure 7). No maximum likelihood estimates are available for the first data set (to 17th August 1998). Hosking and Wallis [13,14] show empirically that no ML estimates exist for n_u<50. Our data supports this for n_u=27. The Bayesian method yields a posterior distribution for the shape parameter with median estimate $\hat{\xi} = 0.22$. Prediction results are improved by 21st August 1998 with the Bayesian estimates still performing better than the maximum likelihood estimates. For data up to 28 August 1998 both estimation techniques start to yield comparable fits. This is so for the data up to the 11th September 1998 and indeed for the full sample. When this experiment is repeated for the threshold u=600 corresponding to the 5% tail of the empirical loss distribution only Bayesian estimates (based on 16 exceedances in the full sample) are reliable.

For the five dates selected the results of the Bayesian calculations of the operational risk capital allocation (using (25)) are given in Table 4. All estimates are based on the medians of the corresponding posterior distributions. Table 4A corresponds to the statistically fit threshold u=150, while Table 4B corresponds to the more theoretically reliable threshold u=600 at which the Bayesian estimate of the tail shape parameter

$\hat{\xi} = 0.57$ (*c.f.* Table 3) indicates that only a single moment of the underlying P&L distribution is finite. The estimated annual expected excess risk capital based on 250 trading days is also shown as a percentage of the corresponding figure estimated from the full data. Clearly for both threshold levels the more data used in the period of turmoil, the closer our model captures the estimated full-data annual excess capital requirement. Examination of Figure 3 shows visually that while daily losses have settled to early 1997 patterns by the end 1998 about 92% of in-sample annual loss capital provision for 1998 could have been predicted using the statistically determined lower threshold value by 11th September, less than half-way through the turmoil and before the Long Term Capital Management collapse added to volatility.

It is the *severity* of loss that varies between the five chosen 'event' dates, with loss *frequency* playing only a minor role. While the estimated expected excess loss using (24) increases from 232 to 587, the estimated time between exceedances decreases only moderately from about 12 to 9 days. The average number of losses per year[2] exceeding the threshold level u=150 remains approximately at 25; that is, ten trading days on average between excessive losses, which seems to be a reasonable time interval in which to liquidate some risky positions.

Data split date	**Daily Expected Excess beyond *u* (*u*=150)**	**Exponential time gap (in days) $\hat{\lambda}_u^{-1}$ between successive Loss Excesses**	**Annualised Poisson Intensity $\hat{\lambda}_u$ (Expected number of Excesses)**	**Expected Excess Annual Risk Capital *(% of the full data estimate)***
17th Aug '98	231.6	11.7	21.4	4,956 *(29.7%)*
21st Aug '98	271.0	11.1	22.5	6,098 *(36.7%)*
25th Aug '98	440.3	10.6	23.6	10,391 *(62.5%)*
28th Aug '98	513.9	10.0	24.9	12,796 *(77%)*
11th Sep '98	586.7	9.6	26.0	15,254 *(91.7%)*
Full sample	517.0	7.7	32.2	16,647 *(100%)*

Tables 4A *Expected excess annual risk capital for the five sub-samples and the full-sample based on estimated with* u=150.

---

[2] The Poisson intensity $\hat{\lambda}_u$ is calculated from equation (24) from the current posterior values for μ, σ and ξ on the MCMC simulation path. This yields an empirical distribution for λ_u from which we select the median estimate $\hat{\lambda}_u$.

However, the estimated excess provision of 16,647 based on the full sample fails to cover actual excess losses over this threshold incurred in the last 250 trading days in the sample of 23,422 -- a *deficit* of about 30%.

Data split date	Daily Expected Excess beyond *u* (*u*=600)	Exponential time gap (in days) $\hat{\lambda}_u^{-1}$ between successive Loss Excesses	Annualised Poisson Intensity $\hat{\lambda}_u$ (Expected number of Excesses)	Expected Excess Annual Risk Capital *(% of the full data estimate)*	
17th Aug '98	319.9	86.6	2.9	928	*(7.2%)*
21st Aug '98	432.0	69.9	3.6	1,555	*(12%)*
25th Aug '98	933.1	50.1	5	4,666	*(36.4%)*
28th Aug '98	1245.2	38.7	6.4	7,969	*(62.1%)*
11th Sep '98	1459.9	36.2	6.9	10,073	*(78.5%)*
Full sample	1395.4	27.2	9.2	12,838	*(100%)*

Table 4B *Expected excess annual risk capital for the five sub-samples and the full-sample based on estimated with u*=600.

On the other hand, while we see from Table 4B that only 79% of the full sample excess capital provision of 12,838 is covered by 11 September using the more theoretically justified higher threshold, the suggested annual provision at this date compares very favourably with sample excess losses of 8,737 over the last 250 trading days – a *surplus* of about 15% – which might be expected from extreme value theory appropriately applied in predictive mode.

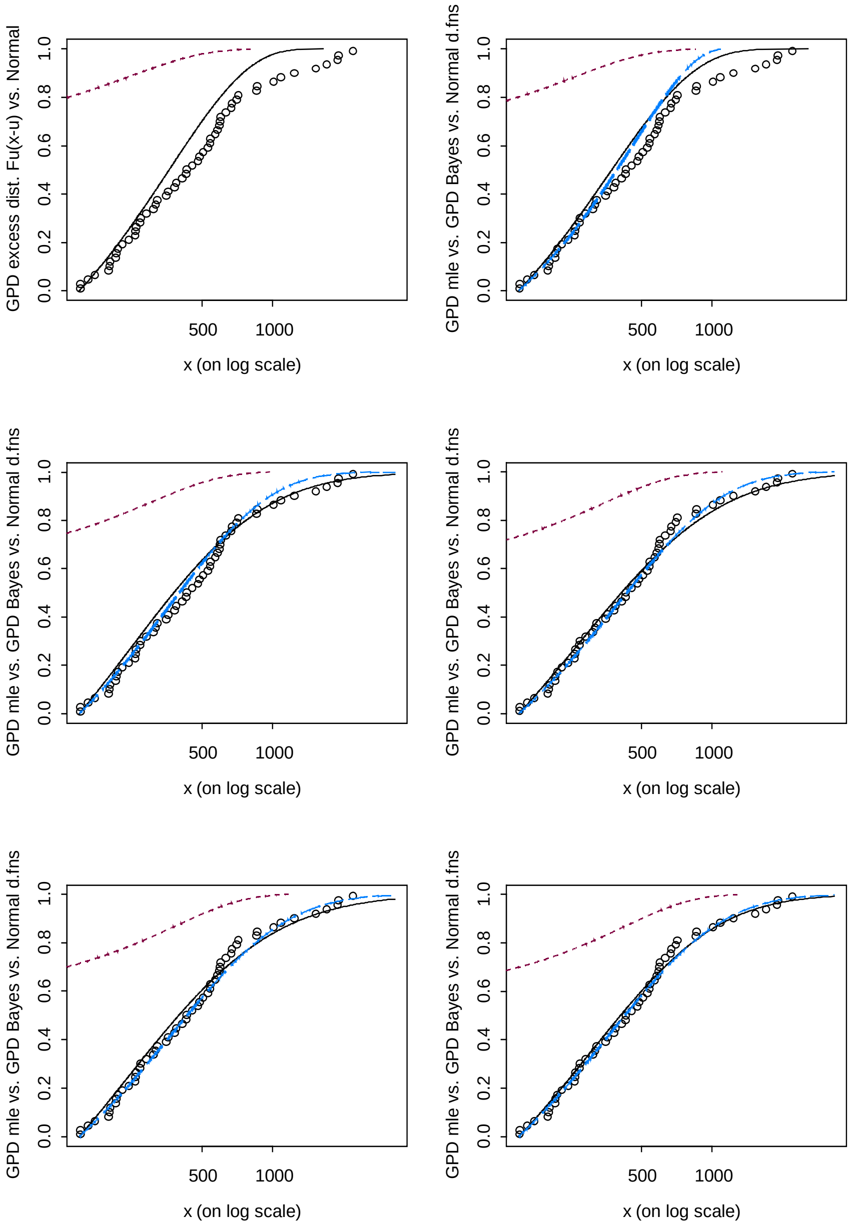

Figure 6 *Aggregated P&L with threshold u=150: Fitted GPD excess distribution functions* $G_{\xi,\beta}$ *based on ML (dashed lines) and Bayesian (solid lines) posterior median estimates of* ξ *and* β *vs. normal distribution functions (dotted lines) using data up to the 17th (top-left), 21st (top-right), 25th (middle-left), 28th August (middle-right), 11th September 1998 (bottom-left) and the full sample (bottom-right). Dots represent the empirical distribution function* F_u *for aggregated losses exceeding u.*

Economic capital for operational risk at business unit level

Having estimated the frequency and severity of the aggregated daily P&L our aim next is to use the hierarchical structure of the Bayesian model for operational risk capital allocation at the level of the four *individual* trading desks. The Bayesian hierarchical MCMC model was applied to the four desks with a fixed loss threshold u=130 for their parameter estimation to ensure a sufficient number of exceedances. The numbers of exceedances (beyond u=130) are respectively n_u= 83, 13, 22 and 8 for desks one, two, three and four, which clearly makes maximum likelihood estimation ill-suited to the task, particularly for desks two and four. The four individual desks estimates for ξ and β as well as for the aggregated P&L data and the annual risk capital (based on 250 trading days) are summarised in Table 5. The GPD-based severity quantile is specified by (18) and expected excess is calculated by (24) and (25).

Firm-wide level u=150	***Bayes posterior median estimates*** $\hat{\xi}$	$\hat{\beta}$	Daily severity q-GPD-based **95%**	**99%**	**Daily expected excess beyond *u***	**Expected number of excesses beyond *u* (per annum)**	**Expected excess annual risk capital**
	0.25	340	691.0	1,639.5	517.0	32.2	**16,646**
Business-unit level u=130							
Desk One	0.34	205.2	601.6	1,360.3	365.9	49.3	18,046
Desk Two	0.25	108.1	116.3	324.5	190.4	7.5	1,435
Desk Three	0.24	118.6	179.2	442.0	206.5	13.0	2,688
Desk Four	0.26	106.1	71.2	250.0	192.8	4.8	925
						Total:	**23,094**

Table 5 *Statistical analysis of the aggregated P&L and the four individual P&L data sets: Bayesian estimates of the GPD and Poisson parameters and their resulting risk measures, all based on the medians of the corresponding posterior distributions.*

Expected excess annual risk capital provision at the firm-wide level is less than the sum of the corresponding capital provisions across the four individual desks. Thus the *sub-additivity* -- or portfolio diversification -- property holds under the expected excess loss risk measure [1]. However, in spite of the too low threshold bias discussed above, the sum of the individual desk provisions covers actual firm-wide excess losses of 23,422 to within about 1%. In addition the hierarchical structure of Bayesian method of parameter estimation provides a more transparent risk assessment for the four business units. Based on the estimates of the severity parameters ξ, β and frequency parameter λ_u, we see that desk one is the most risky amongst the four desks. The estimated parameters are given by the respective medians of their posterior marginal distributions[3] as shown in Figure 7.

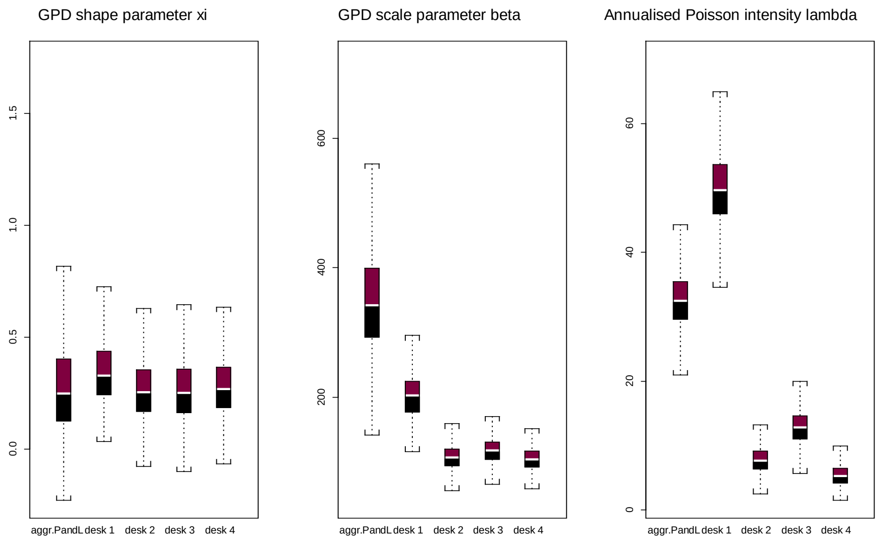

Figure 7 *Posterior distributions of the estimated shape ξ and scale β (GPD) parameters, and the annualised Poisson intensity λ_u. The posterior distributions for the aggregated P&L are estimated from losses exceeding threshold u=150 whereas the posterior distributions for the four individual desks are estimated from losses exceeding threshold u=130.*

[3] Boxplot interpretation of posterior marginal parameter distributions: White horizontal line within the whisker of the boxplot indicates the median of the posterior distribution while the whiskers' lower and upper sides represent respectively the 25% and 75% of the distribution. The lower and upper brackets represent the minimum and maximum values of the distribution.

Economic capital for operational risk at firm level

Our example consists of essentially market data with losses due to political events, i.e. operational losses. It is thus important that the unexpected loss threshold is chosen greater than or equal to the combined market and credit VaR threshold. With such a choice the capital allocation will protect against large and rare losses classified as operational. The most problematic aspect of standard VaR methods -- underestimation of capital for longer time periods -- in this case will be accounted for by exceedances. In our method we have assumed max-stability and therefore only the intensity of the Poisson process is scaled. In Table 6 we summarise the different rules for excess risk capital allocation corresponding to the 18%, 5% and 2.7% quantile thresholds of the empirical P&L distribution and compare them with actual excess losses.

Aggregated Trading P&L Loss Provision			
Threshold u	150	600	1000
Empirical P&L quantile (%)	18	5	2.7
Daily Intensity $\hat{\lambda}_u$ (days) (full sample estimate)	0.1288	0.0368	0.0180
Annual Intensity 250 $\hat{\lambda}_u$ (days)	32.2	9.2	4.5
Daily expected excess above *u* (full sample estimate)	517.0	9.2	4.5
Annual excess capital provision	16,646	12,838	6,877
Actual excess losses above *u* (last 250 trading days in sample)	23,422	8,737	4,619
Percentage safety margin (%)	-29.0	46.9	48.9

Table 6 *Firm-wide excess capital allocation rules for operational risk.*

Conclusions and future directions

Losses incurred similar to those of Barings Bank belong to the category of *extreme operational* loss and could have been mitigated through *control and capital allocation*. P&L data, volatility of returns and other factors should be constantly analysed for identification of extremes. Apparent lack of operational loss data suggests an implementation based on Bayesian hierarchical MCMC simulation, which provides us with robust parameter estimates of extreme distributions. When applied at the level of business units Bayesian procedures allow more efficient capital allocation.

In measuring operational risk we propose a framework which allows a consistent integration with market and credit risk capital allocations. Due to fuzzy boundaries between the different risk types, operational risk must be measured as an excess over levels for market and credit risk. Integrated risk management will involve different risk valuations for different business units and by different models. In our model we assume the 'ordering' of thresholds: market ≤ credit ≤ operational. For integrated risk management further careful adjustments of market and credit thresholds and time re-scaling of intensity should be performed to be comparable with market and credit risk evaluation. These are topics of our current research. Further progress in operational risk modelling depends on cooperation with industry and the wider availability of case study data.

References

1. Artzner P., F. Delbaen, J.M. Eber & D. Heath (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203-228.
2. Basle Committee on Banking Supervision, January (2001). *Operational Risk.*
3. Bernardo J.M. & A.F.M. Smith (1994). *Bayesian Theory*. Wiley, Chichester.
4. British Bankers' Association (1997). Operational Risk Management Survey.
5. Castillo E. (1988). *Extreme Value Theory in Engineering*. Academic Press, Orlando.

6. Danielson J. & C. G. de Vries ((1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance* **4**, 241-257.
7. Du Mouchel W.H. (1983). Estimating the stable index α in order to measure tail thickness: a critique. *Annals of Statistics* **11**. **4**, 1019-1031.
8. Embrechts P., C. Kluppelberg & T. Mikosch (1997). *Modelling Extremal Events*. Springer, Berlin.
9. Feller W. (1966). *An Introduction to Probability Theory and Its Applications,* **2**. Wiley, New York.
10. Galambos J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York.
11. Gnedenko B.V. & Kolmogorov A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, MA.
12. Gnedenko B.V. (1941). Limit Theorems for the Maximal Term of a Variational Series. *Comptes Rendus de l'Academie des Sciences del'URSS* **32**, 7-9.
13. Hosking, J.R., J.R. Wallis & E.F. Wood (1985). Estimation of the generalised extreme value distribution by the method of probability-weighted moments. *Technometrics* **27**, 251-261.
14. Hosking, J.R. & J.R. Wallis (1987). Parameter and quantile estimation for the generalised Pareto distribution. *Technometrics* **29**, 339-349.
15. Jameson R. (1998). Playing the name game. *Risk* **11**, 38-45.
16. Leadbetter M.R., Lindgren G. &. Rootzen H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin.
17. Leadbetter M.R. (1991). On a basis for 'Peaks over Threshold' modeling. *Statistics & Probability Letters* **12**, 357-362.
18. Mandelbrot B.B. (1982). *The Fractal Geometry of Nature,* W.H. Freeman, San Francisco.
19. McNeil A.J. & T. Saladin (1997). The peaks over thresholds method for estimating high quantiles of loss distributions. *In Proceedings of XXVII International ASTIN Colloquium*, 23-43.
20. Ong M.K. (1999). *Internal Credit Risk Models: Capital Allocation and Performance Measurement*. Risk Books, London.

21. Pickands, J. (1975). Statistical inference using extreme order statistics, *Annals of Statist*ics **3,** 119-131.

22. Samorodnitsky G. & M.S. Taqqu (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* Chapman & Hall, London.

23. Smith, A.F.M. and G.O. Roberts (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. Royal Statistical Society* **B 55**, 3-23.

24. Smith R.L. (1985). Threshold methods for sample extremes. In: J. Tiago de Oliveira, ed., *Statistical Extremes and Applications*, NATO ASI Series, 623-638.

25. Smith R.L. (1987). Estimating tails of probability distributions. *Annals of Statistics* **15**, 1174-1207.

26. Smith, R.L. (1990). Extreme value theory. In: Ledermann W. (Chief ed.). *Handbook of Applicable Mathematics Supplement.* Wiley, Chichester, 437-472.

27. Smith, R.L. (1998). Bayesian and frequentist approaches to parametric predictive insurance. In: J.M Bernado, J.O. Berger, A. Dawid, A.F.M. Smith (eds.) *Bayesian Statistics* **6.** Oxford University Press.

28. Smith R.L. (2001). Measuring risk with extreme value theory. Forthcoming in: *Risk Management: Value at Risk and Beyond*. M.A.H. Dempster & H.K. Moffat, eds. Cambridge University Press.