Measuring risk by extreme values

Elena Medova introduces extreme value theory and looks at how it relates to operational risk initiatives. The latter part of the paper describes a new EVT approach that uses Bayesian simulation techniques to measure firm-wide operational risk

Since Gnedenko's *Limit theorems for the maximal term of a variational series* (1941), and Gumbel's *Statistics of Extremes* (1958), theories concerning the calculation of extreme values have been applied to a great variety of practical problems. Extreme value theory (EVT) has found applications in structural, aerospace, ocean and hydraulic engineering, as well as in studies of pollution, meteorology and highway traffic. Actuaries also now use EVT extensively to model casualty insurance claims. So perhaps it's not surprising that researchers have begun to explore whether EVT can be used to measure operational risk in financial institutions.

The key attraction of EVT is that it offers a set of ready-made approaches to the most difficult problems in operational risk analysis: how can risks that are both extreme, and extremely rare, be modelled appropriately?

Key literature

The literature on EVT is now quite extensive. Galambos et al (1994) offers a useful exposition of EVT theory and practice, while Castillo (1998) describes its applications in engineering. The definitive exposition of EVT is given in *Extremes and Integrated Risk Management* edited by Paul Embrechts and published by *Risk Books* earlier this year. Other authoritative texts include a series of working papers and a monograph by Embrechts et al (1997). Significant theoretical and experimental results can be found in Smith (1987, 1997); McNeil and Saladin (1997); McNeil's extreme value software library written in S-plus; and Danielson and de Vries (1997).

Key components

The principal results of EVT concern the limiting distribution of sample extrema (maxima or minima). Suppose that $X = (X_1,\ldots,X_n)$ is a sequence of independent identically distributed observations with distribution function F, not necessarily known, and let the sample maximum be denoted by $M_n = \max\{X_1,\ldots,X_n\}$. Under certain assumptions – subexponential distributions – the tail of the maximum determines the tail of the sum as $n \to \infty$.

1. Standard GEV

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

More generally, the generalised extreme value distribution (GEV) given by $H_\xi(x)$ describes the limit distribution of suitably normalised maxima. The random variable X may be replaced by $(X-\mu)/\sigma$ to obtain a standard GEV with a distribution function that is specified as shown in the box above, where μ, σ and ξ are the location, scale and shape parameters respectively.

Three standard distributions correspond to different values of ξ. They are the:

- **Gumbel distribution** Λ, $\xi = 0$
- **Fréchet distribution** Φ_α, $\xi = \alpha^{-1} > 0$
- **Weibull distribution** ψ_α, $\xi = -\alpha^{-1} < 0$

The purpose of tail estimation procedures is to estimate the values of X outside the range of existing data. To do this, researchers have employed both extreme epochs (events), and exceedances of a specified level. The standard approach assumes that the tail of the population follows the selected family of distributions.

Pickands (1975) showed (with some additional assumptions) that the generalised Pareto distribution (GPD) – the limit distribution of excesses $Y := \max\{X-u, 0\}$ over sufficiently high thresholds u – offers a good approximation of the tail of F for some fixed ξ and β, which depend upon u. Similar results have been obtained for stationary sequences of observations, whose dependence extends only to a finite number of previous values, see Leadbetter et al (1983).

Thus the distribution of Y may be thought of as the conditional distribution of X given $X > u$.

The GPD with shape parameter ξ and scale parameter β is specified as:

$$G_{\xi,\beta}(y) = \begin{cases} 1-\left(1+\xi\dfrac{y}{\beta}\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1-\exp\left(-\dfrac{y}{\beta}\right) & \text{if } \xi = 0 \end{cases}$$

$$\text{where } y \in \begin{cases} [0,\infty] & \text{if } \xi \geq 0 \\ \left[0,-\dfrac{\beta}{\xi}\right] & \text{if } \xi < 0 \end{cases}$$

and the sign of the shape parameter ξ determines its tail behaviour and thus the tail behaviour of the original distribution.

For $\xi>0$, the tail of the distribution function F of X decays like a power function $x^{-1/\xi}$. In this case, F belongs to a family of heavy-tailed distributions that includes, among others, the Pareto, log-gamma, Cauchy and t-distributions.

For $\xi=0$, the tail of F decreases exponentially, and belongs to a class of medium-tailed distributions that include the normal, exponential, gamma and lognormal distributions.

Finally, for $\xi<0$, the underlying distribution F is characterised by a finite right endpoint, which class of short-tailed distributions includes the uniform and beta distributions.

It can be shown that the mean excess function (expectation) of the GDP is given by the expression:

$$e(u) = E(X-u \mid X>u) = \frac{\beta+\xi u}{1-\xi}$$

where

$$\beta = \sigma + \xi(u-\mu)$$

and $\max_n Y_n$ follows a GEV distribution with parameters ξ, μ, σ.

The POT model

The peaks-over-threshold (POT) model can be used to estimate the excess distribution with respect to a threshold level u, and to estimate the tail shape of the original distribution. It should be noted that the threshold setting of the POT model is data dependent. The model defines a two-dimensional (Y_n, N_u) space-time point process on $X_n \geq u$, $n=1,\ldots, N_u$.

Here Y_n and N_u are independent random variables, such that $Y_n \sim \text{GPD}(\xi, b)$, and the number of excesses N_u follows a Poisson process with intensity λ, representing the average number of exceedances over the time interval used for the sampling process and given by:

$$\lambda = \left(1 + \xi \frac{(x - \mu)}{\sigma}\right)^{-1/\xi} \quad \text{for } x \geq u$$

The threshold u is usually chosen using mean excess plots, and other statistical devices developed by Smith (1987, 1997), that consider the trade off between bias in estimating the excess distribution function parameters and their sampling variance.

The POT model for operational risk

The estimation of the parameters of the POT model is usually based on the maximum likelihood method, which requires a relatively large number of observations above the threshold (eg, more than 100). But in an operational risk situation, it might be more realistic to think in terms of 20 or 30 excesses. This suggests that another estimation technique will be necessary for operational risk, and we propose the use of Bayesian simulation techniques (Medova, 1999). A presentation of this estimation procedure, which has been developed with Marios Kyriacou, will be given in the second part of this paper.

Another complex issue is the question of consistency in deciding upon threshold values. In an ideal world, the threshold obtained from the POT model should correspond to the integrated market and credit risk value-at-risk (VAR) quantile. This is currently one topic of our programme of research into integrated risk management.

Some problems

To justify the modelling of operational risk using EVT, many obstacles must be overcome. Not all the obstacles are technical in nature. Many are caused by the fact that operational risk continues to be ill-defined for the purpose of calculating risk capital. For example, one might ask how any approach to operational risk using extreme value theory relates to definitions of "normality" and the problem of internal bank controls and external supervision? And how does EVT relate to the Basle Committee on Banking Supervision's present proposals for controlling operational risk?

The Committee has attempted to clarify the complex issues of risk management by adopting a "three-pillared" approach. The first pillar concerns capital allocation, the second supervision and controls and the third transparency and consistency of risk management procedures. What is the relation of EVT to these three pillars – most problematically, the second and third pillars?

Another problem is that, 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 to what degree risk capital is used to cover the most extreme risks.

Some practitioners and regulators have made it clear that they do not intend to include the risk of the most extreme losses in their calculations of either economic risk capital or regulatory risk capital.[1] So in what way is extreme value theory useful in measuring operational risk?

Lastly, how can an analyst deal with market and credit risk management without double-counting? A framework that identifies the roles of credit, market, and other risks must be constructed. Below we suggest some thoughts on these issues that help to show how they relate to the nature of EVT.

Some solutions

Let us assume that a bank's market and credit risk management is informed by quantitative models that compute the VAR for market risk and credit risk and allocate economic capital to these risks. Is such a capital allocation sufficient for unexpected losses due to human errors, fraudulent activities and other external factors? Clearly not, for two reasons.

Firstly, the models do not take into account operational risks (extreme or not). Secondly, they make various assumptions about "normality", and so exclude extreme and rare events. Such events include natural disasters, as well as major social or political events.

How can we think clearly about operational and extreme events? In our research, we termed the related risk factors primal (catastrophic). Processing all information and taking decisions at different levels of the bank may lead to further losses reflected in increased business costs. Such secondary causes include human or technological errors, lack of control to prevent unauthorised or inappropriate transactions, fraud and faulty reporting. Many of these secondary causes are used in one or other definition of operational risk. Some of them, such as the failure of a bank's internal computer system, may themselves be regarded as primal and catastrophic.

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 primal and secondary risk factors.

Ideally, this analysis would form part of the financial surveillance system for 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, it relates to the second of the Basle Committee's three pillars.

Here, the important point is that this surveillance is concerned with the identification of the "normality" of business processes. The identification of suitable market and credit risk models also forms a natural part of this operational risk assessment.

Such an analysis should allow an analyst to classify a bank's losses into two categories:

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

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

Next, we might take the view that control procedures will be developed for the reduction of the low-value/frequent losses, and for their illumination and disclosure (the third pillar of the Basle approach).

These control procedures, and any continuing expected level of loss, should be accounted for in the operational budget. This allows us to assume that only losses of a large magnitude need be considered for operational risk economic capital provision.

Again, an analysis of the profit and loss data and the verification or rejection of the assumption of normality, related to the universe of primary and secondary risks, are all part of the (usually internal) risk supervisory process.

From VAR to extreme event analysis

VAR has been adopted as the central measure of market risk by many financial institutions. Under normal market conditions, VAR provides a measure of the market risk due to adverse market movements. Any deviation from normality will tend to underestimate the VAR. Similarly, under normal conditions for credit risk, which correspond to credit ratings higher than triple B, credit models provide measures for credit risk.

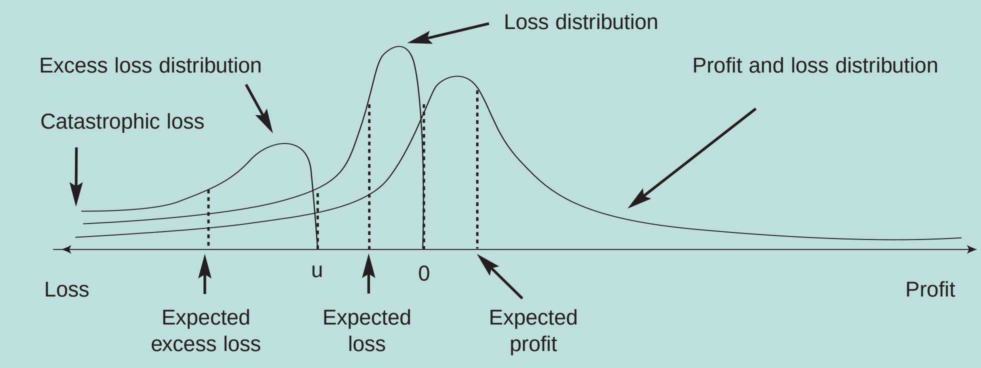
But there are theoretical alternatives to VAR that also offer a coherent risk measure (Artzner et al, 1997, 1999). One approach is to define a measure for the expected shortfall, or tail conditional expectation, with respect to the unknown maximal loss distribution.

We adopt a similar conditional measure for operational risk. But we assume that a threshold has been derived from the marginal statistical distribution of losses as a part of the operational risk supervisory process. A quantitative treatment of this approach will begin to be developed in the next section. This gives a slight twist to the usual definition of operational risk. For the purpose of calculating capital provision, operational risk is everything which is not credit and market risk under normal conditions, including catastrophic market and credit losses where appropriate.

In effect, operational risk is redefined as a tail of the profit and loss distribution of the appropriate level – business unit, or enterprise-wide – of the bank (see figure 2). In the presence of extremes, further analysis will be required for the

[1] For example, see practitioner Tony Peccia's comments on capital allocation at CIBC World Markets, one of Canada's leading banks, in April's Operational Risk newsletter, page 12, and comments by Jeremy Quick of the UK's Financial Services Authority, in February's Operational Risk newsletter, page 10

2. Profit and loss distribution and a chosen threshold for extreme operational losses



identification of a threshold, and for the evaluation of a capital requirement for unexpected operational losses.

2. Calculating operational risk by Bayesian simulation

Above, we argued that the expected severity and frequency of losses are required to evaluate a capital requirement for unexpected operational losses. A Bayesian simulation methodology will be developed in order to measure firm-wide operational risk and allocate capital risk against it. In our framework, the capital requirement is given by the expectation of the excess distribution and the frequency by the average number of exceedances over time, with respect to a suitably chosen threshold for the POT model involving GPD.

Extreme losses are rare by definition and, consequently, the issue of limited data availability becomes of crucial importance to the accuracy of the resulting risk measures against extreme losses due to all sources.

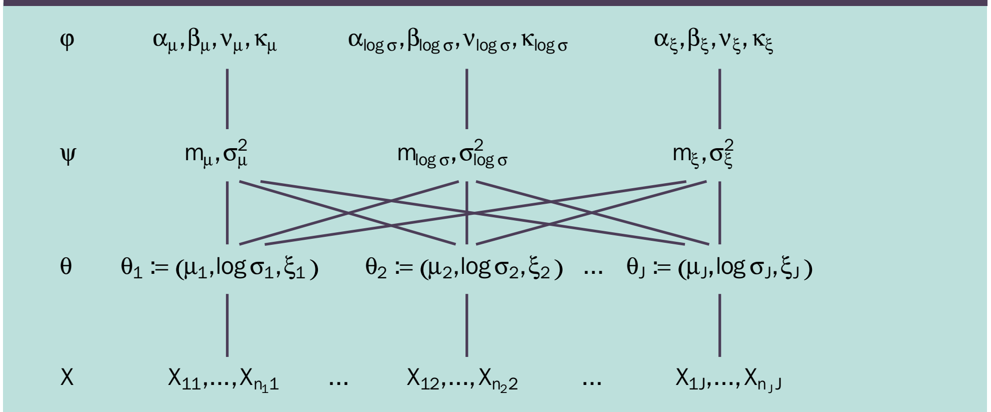
The conventional method to estimate parametric extreme loss distributions is maximum likelihood estimation (MLE), which chooses parameter estimates to make the observed sample most likely. This method is widely used because it is theoretically robust and fast to implement from a computational viewpoint. However, the sampling distributions of MLEs are based on large sample theory and they perform unstably when the MLE method is applied to small, or even moderate, sample sizes.

Our procedure for capital allocation for "unexpected" operational risks is therefore based on simulated parameter estimates of the model. Let us assume that an agreement on how to collect profit and loss data has been reached, that the data has been collected and that appropriate threshold levels have been chosen for business units and at firm-wide level (this may require a number of statistical techniques as noted above). The matrix of losses (see table A) may be classified into several sub-samples, each associated with a different risk factor and business unit.

A. Firm-wide matrix of operational losses

Business unit	**1**	**...**	**j**	**...**	**N**	**Firm-wide**
Loss type						
Technology						
Human						
:						
External events						
Total	$X_1^1,\ldots,X_n^1$		$X_1^j,\ldots,X_n^j$		$X_1^N,\ldots,X_n^N$	$X_1,\ldots,X_n$

3. Hierarchical Bayesian model parameter and observation dependencies conditional on their hyper-parameters



Since sufficient data is seldom available for accurate estimation of extreme values, we have developed a computationally intensive hierarchical Bayesian simulation technique for fitting heavy-tailed distributions to small samples. At present, the procedure is developed for 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. Essentially, the technique is trading computational power for lack of data, but its empirical estimation efficiency when tested on subsets of large data sets is surprisingly good.

A Bayesian viewpoint treats uncertainties about parameters by considering them 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 hyper-parameters 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 hyper-parameters ψ^+ determined by ψ and the observations $X_1,\ldots, X_n$, we say that $f_{\theta|\psi}$ is a parametric family of densities conjugated 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 sample according to loss data subtypes or classes (ie, risk factors or business units), as well as the aggregate (Smith, 1998a). In this model, the Bayesian prior density for the hyper-parameters ψ is common to all loss subtype Bayesian prior densities for the parameters θ. At the model initialisation stage, the hyper-hyper parameters φ, which are the Bayesian parameters of this common prior hyper-parameter, are chosen to generate a vague prior, indicating a lack of information on the hyper-parameters' prior distribution before the loss data is seen. Thus we have a Bayesian hierarchical decomposition of the posterior parameter density $f_{\theta|X,\psi}$, given the observations and the initial hyper-hyper parameters φ, as:

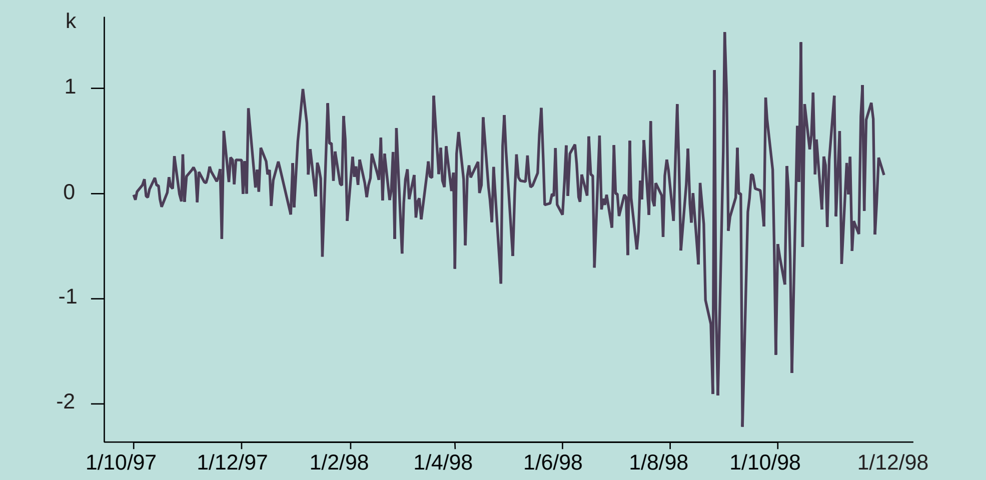
$$
\begin{aligned}
& f_{\theta|X,\psi}(\theta|X,\psi) \propto \\
& f_{X|\theta}(X|\theta) f_{\theta|\psi}(\theta|\psi) f_{\psi}(\psi|\varphi) \\
& \propto f_{X|\theta}(X|\theta) f_{\psi|\theta}(\psi|\theta,\varphi) \\
& \propto f_{X|\theta}(X|\theta) f_{\psi}(\psi|\varphi^+)
\end{aligned}
$$

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, then calculating the value of the corresponding posterior density for this q given the observations X. Figure 3 schematically depicts the relationships between the three 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 hyper-parameter Gaussian prior distribution, their posterior multivariate Gaussian specification will not maintain this independence given observations that are statistically dependent.

The Bayesian posterior density $f_{\theta|X,\psi}$ may be calculated via Markov chain Monte Carlo (MCMC) simulation (Hastings, 1970, Smith & Roberts, 1993, and Smith, 1998a). The idea is to simulate a sample path of a Markov chain whose state is the parameter vector q and whose visited states converge to a stationary distribution that is the Bayesian joint posterior distribution $f_{\theta|X,\psi}$ (termed the target distribution) given the loss data X and a vector ψ of hyper-parameters as discussed above. As used here, 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 calculation, to create a parameter histogram approximation to the required posterior parameter marginal distributions.

For our application, the parameter vector q represents the GPD parameters of interest $\{\mu_j, \log\sigma_j, \xi_j : j = 1,2,\ldots,J\}$ for the $j = 1,\ldots,J$

4. Aggregated profit and loss data from four trading desks



data classes (risks or business units) and the hyper-parameter 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 risk types, or alternatively across all business units) multivariate Gaussian prior distribution of the GPD parameters. To implement the strategy, Gibbs sampling and the Metropolis-Hastings algorithm (Smith & Roberts, 1993) 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 and appropriate prior distributions. Running the Markov chain for very many transitions (about one million) 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 $t \geq 0$) depending solely upon θ_j^t. This process represents the traditional exchange of computational intensity for low data

5. Mapping the Bayesian hierarchical framework to data

Risk factor: external event

Business unit:	1	2	.	.	.	J		
	$x_{1,1}$	$x_{1,2}$	.	.	.	$x_{1,J}$		
	$x_{2,1}$	$x_{2,2}$	.	.	.	$x_{2,J}$		
	.	.	.	.	.	.		
	.	.	.	.	.	.		
	.	.	.	.	.	.		
	$x_{n_1,1}$	$x_{n_2,2}$	.	.	.	$x_{n_j,J}$		
POT parameters							**Hyper-parameters**	**Hyper-hyper parameters**
Mean	$\mu 1,$	$\mu 2,$	.	.	.	μ_J	m_μ, s_μ^2	$\alpha_\mu, \beta_\mu, \gamma_\mu, \kappa_\mu$
Log scale	$\log\sigma_1,$	$\log\sigma_2,$	.	.	.	$\log\sigma_J$	$m_{\log\sigma}, s_{\log\sigma}^2$	$\alpha_{\log\sigma}, \beta_{\log\sigma}, \gamma_{\log\sigma}, \kappa_{\log\sigma}$
Shape	$\xi_1,$	$\xi_2,$	.	.	.	ξ_J	m_ξ, s_ξ^2	$\alpha_\xi, \beta_\xi, \xi_\gamma, \kappa_\xi$

B. Statistical analysis of the aggregated P&L and the four individual P&L data sets

Firm-level u=150	ξ	β	Daily severity distribution quantile .95	.99	Daily expected excess beyond u	Expected excesses beyond u (per year)	Excess annual risk capital
Aggregated P&L	0.25	340	691.0	1,639.5	517.0	47	24,294
Business units							
Threshold u = 130							
One	0.34	205.2	601.6	1,360.3	365.9	72	26,345
Two	0.25	108.1	116.3	324.5	190.4	11	2,095
Three	0.24	118.6	179.2	442.0	206.5	19	3,924
Four	0.26	106.1	71.2	250.0	192.8	7	1,350
Total							33,714

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 τ. By discarding the first k (= 10,000) 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 (see figure 4) 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 parameter distributions.

Example: Bank trading loss analysis through the Russian crisis

The financial turmoil in 1998 caused by the Russian government's default on August 24 was a political event that caused significant losses to major banks. In financial crises, the separation of financial risk sources into various types (market, credit, etc) can prove to be a fallacy and the Russian crisis was no exception. The credit risk in the primary market leads to market risk in the secondary markets and can lead to operational risk within the financial institutions themselves. To be more precise, operational risk in this case is associated with the losses likely to be suffered by the bank if it failed to set the appropriate controls in place to quantify a potential political event in a timely manner in order to reduce its exposure. Of course, one could argue that these losses do not belong to the category of operational risk, but we will not debate this point further here. Rather, we use trading data over this period of external market turmoil to test the procedures outlined above. Note, however, that the aggregated profit and loss data plot in figure 4 shows some extreme losses (along with a few abnormal profits) through late summer and early autumn 1998, before returning to normal volatility levels at the end of the year.

Our Russian crisis case study has two aims: first, to estimate the extreme loss severity and frequency applying our techniques described above; second, to estimate similarly the corresponding economic capital provision using the estimate of expected excess loss over a suitably chosen threshold (Medova, 2000). Here we report in detail only on the estimation of economic capital provision by business units and in total using our hierarchical Bayesian simulation techniques.

Four daily profit and loss reports on trading activities denominated in a domestic currency have been rescaled for reasons of confidentiality. The aggregated profit and loss data consists of 237 losses of a total size of 58,600 monetary units and individual losses range from 0.002 to 2,942.028 (see figure 4) in the period from October 1, 1997 to December 31, 1998.

Statistical analysis indicates heavy-tailed behaviour, which for obvious reasons cannot be explained by market factors alone. The data X consists of $J(= 4)$ classes (business units), as in figure 5. The necessary diagnostics were carried out to test the quality of fit of the POT model to the aggregated loss data. Having estimated the frequency and severity of the aggregated losses, our aim is to use the hierarchical structure of the Bayesian model to analyse jointly the four individual loss types.

The multivariate Bayesian estimation of the POT model was applied jointly to the four subsamples. The values of the simulated parameters and the resulting loss estimates are presented in table B. Decentralisation of risk management allows the identification of the risk sources and more efficient capital allocation.

The total capital estimated for the four trading desks in our case study is more than that calculated at the firm level due to partially independent tail events – capital allocation satisfies the sub-additivity property of risk measures and thus the firm's operations enjoy the effects of portfolio diversification and loss mitigation. ■

Dr Elena A Medova is part of a research team, including Professor MAH Dempster and MN Kyriacou, at the Centre for Financial Research, Judge Institute of Management, University of Cambridge, working on the topic of op risk
www.cfr.jims.cam.ac.uk

REFERENCES

Artzner P, F Delbaen, JM Eber and D Heath, 1997
Thinking coherently
Risk 10, pages 68-71

Artzner P, F Delbaen, JM Eber and D Heath, 1999
Coherent measures of risk
Mathematical Finance 9, pages 203-228

Castillo E, 1988
Extreme Value Theory in Engineering
Academic Press, Inc

Danielson J and CQ de Vries, 1997
Tail index and quantile estimation with very high frequency data
Journal of Empirical Finance 4, pages 241-257

Embrechts P, 2000
Extreme Value Theory and Integrated Risk Management
Risk Books

Embrechts P, C Kluppelberg and T Mikosch, 1997
Modelling Extremal Events
Springer Verlag, Berlin

Galambos J, J Lechner & E Simiu, ed, 1994
Extreme Value Theory and Applications
Proceedings of the Conference on Extreme Value Theory and Applications, Gaithersburg, Maryland 1993

Gnedenko BV, 1941
Limit theorems for the maximal term of a variational series
Comptes Rendus de l'Academie des Sciences del'URSS 32, pages 7-9

Gumbel EJ, 1958
Statistics of Extremes
Columbia University Press

Leadbetter MR, G Lindgren and H Rootzen, 1983
Extremes and Related Properties of Random Sequences and Processes
Springer Verlag, Berlin

McNeil AJ and T Saladin, 1997
The peaks over thresholds method for estimating high quantiles of loss distributions
Proceedings of XXVII International ASTIN Colloquium, pages 23-43

Medova E, 1999
Operational risk measures and Bayesian simulation methods for capital allocation
Working Paper 27/99, Judge Institute of Management Studies, University of Cambridge

Quick J, 1999
Other risks
Reforming the Basle Capital Accord (FSA conference), October 1999, London

Smith RL, 1987
Estimating tails of probability distributions
The Annals of Statistics, vol 15, pages1174-1207.a