Scientific Research

An Application of Bayesian Inference on the Modeling and Estimation of Operational Risk Using Banking Loss Data

Kashfia N. Rahman[1], Dennis A. Black[2], Gary C. McDonald[1]

[1]Oakland University, Rochester, USA
[2]Comerica Bank, Dallas, USA
Email: kashfia_nausheen@yahoo.com, dablack@comerica.com, mcdonald@oakland.edu

Abstract

Bayesian inference method has been presented in this paper for the modeling of operational risk. Bank internal and external data are divided into defined loss cells and then fitted into probability distributions. The distribution parameters and their uncertainties are estimated from posterior distributions derived using the Bayesian inference. Loss frequency is fitted into Poisson distributions. While the Poisson parameters, in a similar way, are defined by a posterior distribution developed using Bayesian inference. Bank operation loss typically has some low frequency but high magnitude loss data. These heavy tail low frequency loss data are divided into several buckets where the bucket frequencies are defined by the experts. A probability distribution, as defined by the internal and external data, is used for these data. A Poisson distribution is used for the bucket frequencies. However instead of using any distribution of the Poisson parameters, point estimations are used. Monte Carlo simulation is then carried out to calculate the capital charge of the internal as well as the heavy tail high profile low frequency losses. The output of the Monte Carlo simulation defines the capital requirement that has to be allocated to cover potential operational risk losses for the next year.

Keywords

1. Introduction

Operational risk modeling is receiving a lot of attention in the financial sector due to recent high profile banking

losses. These losses result from inadequate or failed internal processes, people or systems, or from external events. Traditionally, these risks were managed using insurance protection and audit. However, due to globalization, complex financial products and changes in information technology and given the rise in the number of high-profile operational losses in the banking sector, the importance of operational risk and its management have become an important topic. The Basel Committee of Banking Supervisor (BCBS) began discussion on operation risk in 1998. These discussions resulted in an international regulatory framework, Basel II (2001-2006), which requires internal risk management and risk capital requirement to manage operational risks. Now, all major banks are developing quantitative methods to satisfy the requirements defined in Basel II. Monte Carlo simulation, using point estimation of the distribution parameters, is the predominant method in the banking industry for the modeling of the operational risk. Bayesian inference method is another alternative which has been suggested for the modeling of operational risk.

Recently a text book on operational risk modeling using Bayesian Inference has been published [1]. This book presents a good reference of operational risk modeling using Bayesian Inference as well as several Bayesian model derivations. The Bayesian inference methods, in the context of operational risk, have been briefly mentioned in the earlier literature. Books such as [2]-[4], have short sections on a basic concept of a Bayesian method. The Bayesian method was implicitly used to estimate operational risk frequency in the working paper of [5]. One of the first publications to present detailed and illustrative examples of the Bayesian inference methodology for estimation of the operational risk frequency and severity, and subsequent estimation of the capital was presented in [6]. The Bayesian methodology was extended to combine three data sources (expert opinion, internal, and external data) in [7] and developed further in [8] for the multivariate case with dependence between risks. Currently, the use of Bayesian methods for modeling operational risk is an active research line.

Modeling of operational risk in financial sectors is receiving increasing attention. Frequentist methods [9] are predominantly used for the modeling of operational risk. However, these (frequentist) methods do not include the effect of parameter uncertainty while estimating the capital requirement as mandated by the Basel II accord (www.bis.org). This paper has made another attempt of Bayesian inference method to overcome this drawback. Bayesian inference is used to develop posterior distributions of the severity and frequency distribution parameters. Bank internal and external data collected over several years are used to define the prior and the posterior distribution of the severity as well as the frequency distribution parameters. To model the low frequency but high severity heavy tail loss data a somewhat ad-hoc method, which is substantially different from any of the methods presented in the earlier papers, is implemented in this paper. The loss data are divided into several high loss buckets, where the frequencies of each bucket losses are determined from the expert opinion data. Severity distribution, along with the parameter posterior distribution, as derived using the bank internal and external data is also used for the heavy tailed bucket data. However point estimation of the frequency parameters, as defined by the experts, is used.

This paper is organized as follows. Section 2 briefly describes the operational risk modeling method as defined by the Basel II accord. Section 3 presents the Bayesian model development and posterior distribution formulation. Section 4 presents the Monte Carlo simulation results and compares the results using the conventional frequentist approach with those using the Bayesian approach. Finally, Section 5 presents the summary and conclusion.

2. Modeling Operational Risk

The Basel II framework is based on a three-pillar concept (www.bis.org). *Pillar I*: *Minimum Capital Requirements*, *Pillar II*: *Supervisory Review Process*, *and Pillar III*: *Market Discipline*. This article focuses on the Pillar I and considers a probabilistic model of operational risk losses. The Basel II framework defined three approaches to quantify operational risk annual capital charge $\mathbb{C}$. The basic indicator approach, the standardized approach, and the advance measurement approach (AMA) [9] [10].

The AMA is the one adopted by most major banks. Under AMA a bank can calculate an internally developed model subject to regulatory approval. A bank intending to use AMA needs to demonstrate accuracy of the internal models within the matrix of Basel II risk cells. There are seven event types and eight business lines for each of the risk cells. The seven event types and eight business lines in combination create 56 risk cells.

Many models are suggested for modeling operational risk under Basel II AMA. These models have been discussed in more detail in [9] [11]. There are two distinct approaches, the top-down approach and the bottom up approach. Lost distribution approach (LDA) is one bottom up approach which is most widely used and is

adopted in this work. A more detailed description of the LDA modeling can be found in [2] [3] [12] [13] so a detailed description is not given here.

The LDA model is based on modeling the frequency N and severities $X_1, X_2, X_3, \cdots$ of the operational risks events. Then, the annual loss is calculated by aggregation of the severities over a one year period: $Z = X_1 + X_2 + \cdots + X_N$. LDA model for the total annual loss Z_t in a bank can be formulated as:

$$Z_t = \sum_{i=1}^{J} Z_t^{(j)}; Z_t^{(j)} = \sum_{i=1}^{N_t^{(j)}} X_i^{(j)}(t). \tag{1}$$

1) $J's$ are the risk cells at which events are divided.

2) The annual loss in risk cell j, $Z_t^{(j)}$, is modeled as a compound (aggregate) loss over one year with the frequency (annual number of events) $N_t^{(j)}$ implied by a counting process (e.g. Poisson process) and severities $X_i^{(j)}(t), i = 1, \cdots, N_t^{(j)}$.

3) Typically, the frequencies and severities are modeled by independent random variables.

Estimation of the annual loss distribution by modeling frequency and severity of losses is used extensively in actuarial science. It is also used to model solvency requirements for the insurance industry. Under model (1), the capital is defined as the 0.999 Value-at-Risk (*VaR*) which is the quantile of the distribution for the next year annual loss Z_{T+1} as shown in Equation (2):

$$VaR_q[Z_{T+1}] = \inf\{z \in \mathbb{R} : P_r[Z_{T+1} > z] \leq 1 - q\} \tag{2}$$

at the level $q = 0.999$. Here, $T + 1$ refers to the next year.

Monte Carlo simulation is a widely used method in calculating the sum defined in Equation (1). The severity data is fit into a defined parametric distribution, while the frequency is typically modeled with a Poisson distribution. Monte Carlo simulation is then run to generate very large samples from the given distribution and the above sum is calculated to define the capital charge for an event type defined earlier. This will be presented in more details in Section 4.

Operational Loss Data

Basel II specifies the data that should be collected and used for AMA. In brief a bank should have internal data, external data, and expert opinion data. Combining these different sources of information is critical for estimation of operation risk, especially for low-frequency/high severity risks.

Basel II requires that capital charge should cover unexpected loss (UL), while expected loss (EL) should be covered by the bank through internal provisions. The capital, required to cover operational losses, is the sum of EL and UL which is the 99.9% Value-at-Risk (VaR). **Figure 1** below describes this [9].

Ad-hoc procedures or general non-probabilistic methods are used, conceptually, to process different sources of data. For example under an ad-hoc procedure, one can define different Poisson frequencies for the internal and external/Expert opinion data, while the distribution of the external or expert opinion severity data may be
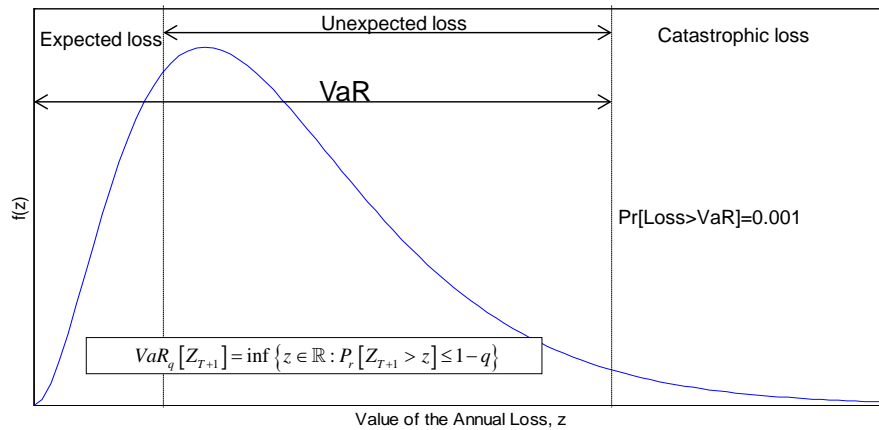

Figure 1. Illustration of the expected and the unexpected losses in the capital *Requirement at the* 99.9% *confidence level.*

modeled using the same distribution identified by the internal data. Sometimes, low frequency high severity heavy tailed data is modeled as a separate distribution and a mixed distribution model is used.

3. Bayesian Model Development

3.1. Estimation of Prior Distribution

3.1.1. Frequency Prior Distribution

A Poisson distribution is assumed for frequency and the Poisson parameter is assumed to follow a gamma distribution. These selections of the prior and the parameter distributions (conjugate prior) allow the posterior to be defined by closed form equation. Once the parameters of the prior gamma distribution are obtained the posterior frequency distribution is defined. The internal data provided by the bank has the following information:

1) Given $\Lambda_j = \lambda_j$, $N_{j,k}$ are independent random variables from Poisson (λ_j) with probability mass function $f\left(.|\lambda_j\right)$. Here $\left\{N_{j,k}, k = 1, 2, \cdots, j = 1, \cdots, J\right\}$ is the annual number of events in the *j*-th cell in the *k*-th year. Here λ_j represents the average frequency for the *j*-th cell. Unlike the frequentist approach, the Poisson parameter λ_j does not have a point estimation but a random distribution (the prior) which is defined as Gamma (α, β) with the density function as $\pi(.)$.

2) We further assume that the random variables $\Lambda_1, \Lambda_2, \cdots, \Lambda_J$ for each of the risk cells are independent but are identically distributed from Gamma (α, β).

We now define $N_j = \sum_{k=1}^{K_j} N_{j,k}$.

3) With the above information and assumptions, the joint density of all data (over all J risk cells) can be written as

$$f\left(n_{1:J}\right) = \prod_{j=1}^{J} \int \left[\prod_{k=1}^{K_j} f\left(n_{j,k}|\lambda_j\right)\right] \pi\left(\lambda_j\right) \mathrm{d}\lambda_j = \prod_{j=1}^{J} \int \left[\prod_{k=1}^{K_j} \mathrm{e}^{-\lambda_j} \frac{\lambda_j^{n_{j,k}}}{n_{j,k}!}\right] \frac{\lambda_j^{\alpha-1}\mathrm{e}^{-\lambda_j/\beta}}{\Gamma(\alpha)\beta^\alpha} \mathrm{d}\lambda_j, \tag{3}$$

where the distribution of $f\left(n_{j,k}|\lambda_j\right)$ is Poisson and the distribution of the parameter $\pi\left(\lambda_j\right)$ is gamma. The integration over the model parameters (Equation (3)) can be carried out and the joint distribution is obtained as shown in the Equation (4),

$$f\left(n_{1:J}\right) \propto \prod_{j=1}^{J} \frac{\Gamma\left(\alpha + n_j\right)}{\Gamma(\alpha)\beta^\alpha \left(K_j + \dfrac{1}{\beta}\right)^{\alpha + n_j}}. \tag{4}$$

Taking log on both sides the log likelihood function becomes

$$\log f\left(n_{1:J}\right) \propto \sum_{i=1}^{J} \left[\log\Gamma\left(\alpha + n_j\right) - \log\left(\Gamma(\alpha)\right) - \alpha\log(\beta) - \left(\alpha + n_j\right)\log\left(\frac{1}{\beta} + K_j\right)\right]. \tag{5}$$

The parameters α and β can now be estimated by maximizing the $\log f\left(n_{1:J}\right)$ of Equation (5) over α and β by the likelihood method. The SAS procedure NLP is used for this purpose. The simulation results are presented in the next section.

3.1.2. Severity Prior Distribution

Next, the prior distribution of the severity is calculated. The simulation data best fits into a log-normal distribution and is thus used to model the severity data. It is assumed that the distribution parameter μ (of the severity log-normal distribution) is normally distributed but the standard deviation σ of the log-normal distribution is known. To identify best distribution fit for the loss data the Bayesian information criterion (BIC) or the deviance information criteria (DIC) may be used. Some of the widely used distributions are Pareto, log-Gamma, mixed log-normal, Weibull, etc. Now under the log-normal/normal distribution described above, we have the following severity model for a risk cell in one bank:

Let $\Theta_\mu \sim N\left(\mu_0, \sigma_0\right)$ be a normally distributed random variable with parameters μ_0 and σ_0, which are estimated from simulated data using the maximum likelihood estimation.

Given $\Theta_\mu = \mu$ the losses $X_1, X_2, \cdots$, are conditionally independent with a common log-normal distribution $X_k \sim LN\left(\mu, \sigma\right)$ where σ is assumed known.

With above assumptions, the joint probability distribution of all data belonging to J cells over K_J years can be expressed as

$$f\left(y_1, y_2, \cdots, y_j\right) = \prod_{j=1}^{J} \frac{1}{\sqrt{2\pi}\sigma_0} \int \left\{ e^{-\frac{\left(\mu_j - \mu_0\right)^2}{2\sigma_0^2}} \prod_{k=1}^{K_j} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(y_j^k - \mu_j\right)^2}{2\sigma^2}} \right\} d\mu_j. \tag{6}$$

In Equation (6) $y_k^j = \log\left(x_k^j\right)$, where x_k^j is the severity data for the j-th cell for the k-th year. Under the log transformation of the severity data, the log-normal distribution variate transforms into a normal variate. After some steps the above expression simplifies to

$$f\left(y_1, \cdots, y_j\right) \propto \prod_{j=1}^{J} \frac{B\sigma_{0,n}}{\sigma_0} \int \left[\frac{1}{\sqrt{2\pi}\sigma_{0,n}} \exp\left\{ -\frac{\left(\mu_j - \mu_{0,n}\right)^2}{2\sigma_{0,n}^2} \right\} \right] d\mu_j, \tag{7}$$

where

$$B = \exp\left\{ \frac{\left(\mu_0^2/\sigma_0^2 + \sum_{k=1}^{K_J} y_j^k/\sigma^2 \right)^2}{2\left(\frac{1}{\sigma_0^2} + \frac{K_J}{\sigma^2} \right)} - \frac{\mu_0^2}{2\sigma_0^2} - \frac{1}{2\sigma^2}\sum_{k=1}^{K_J}\left(y_j^k\right)^2 \right\}, \tag{8}$$

$$\mu_{0,n} = \frac{\frac{\mu_0}{\sigma_0^2} + \sum_{k=1}^{K_J} \frac{y_j^k}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{K_J}{\sigma^2}} = \frac{\mu_0 + \omega \sum_{k=1}^{K_J} y_k^j}{1 + K_J \omega}, \omega = \frac{\sigma_0^2}{\sigma^2}, \text{and } \sigma_{0,n}^2 = \frac{\sigma_0^2}{1 + K_J \omega}.$$

The integration in Equation (7) is taken over a normal distribution with the distribution parameters $\mu_{0,n}$ and $\sigma_{0,n}^2$, defined by Equation (8). This integration is the total area under the probability distribution and is equated to 1. Now taking log on the remaining terms we get

$$\log f\left(y_1, \cdots, y_j\right) \propto \sum_{i=1}^{J} \left[-\log\left(\sigma_0\right) + \log\left(\frac{\sigma_0}{\sqrt{1 + K_j\omega}} \right) + \left\{ \frac{\left(\frac{\mu_0}{\sigma_0^2} + \frac{\sum_{k=1}^{K_j} y_j^k}{\sigma^2} \right)^2}{2\left(\frac{1}{\sigma_0^2} + \frac{K_j}{\sigma^2} \right)} - \frac{\mu_0^2}{2\sigma_0^2} - \frac{1}{2\sigma^2}\sum_{k=1}^{K_j}\left(y_j^k\right)^2 \right\} \right]. \tag{9}$$

Similar to the frequency likelihood the prior distribution parameters μ_0 and σ_0 are estimated by maximizing the above expression using the MLE method.

3.2. Estimation of Posterior Distribution

Once the prior distributions are obtained, the posterior distributions can be derived using the Bayesian rule [1]. If a closed form solution of the posterior distribution cannot be derived, Markov Chain Monte Carlo (MCMC) method can be used to numerically determine the posterior distribution. For the selected distributions of the frequency, severity, and the model parameters in this article, the posterior distribution for the average frequency λ_j, and average severity μ_j can be derived in closed form into some known distribution. This is illustrated in the following two sub-sections.

3.2.1. Posterior Frequency Distribution
Once the prior distribution parameters α and β are estimated, the posterior distribution of λ_j for the j-th risk cell can be expressed using Bayesian rule as [1]:

$$\pi\left(\lambda_j \mid n_j\right) \propto \frac{\left(\frac{\lambda_j}{\beta}\right)^{\alpha-1}}{\Gamma(\alpha)\beta} \mathrm{e}^{-\frac{\lambda_j}{\beta}} \prod_{k=1}^{K_j} \mathrm{e}^{-\lambda_j} \frac{\lambda_j^{n_{j,k}}}{n_{j,k}!}. \tag{10}$$

After some simplifying steps Equation (10) can be expressed as a gamma distribution as follows:

$$\pi\left(\lambda_j \mid n_j\right) \propto \lambda_j^{n_j+\alpha-1} \exp\left\{-\frac{\lambda_j\left(1+\beta K_J\right)}{\beta}\right\}, \tag{11}$$

which is proportional to a gamma distribution Gamma $\left(\hat{\alpha}, \hat{\beta}\right)$, where the parameters are defined in Equation (12)

$$\hat{\alpha} = \alpha_T = \alpha + \sum_{k=1}^{K_J} n_j^k = \alpha + n_j$$
$$\hat{\beta} = \beta_T = \frac{\beta}{\left(1+\beta K_J\right)}. \tag{12}$$

With the posterior distribution of frequency determined the expected frequency over a risk cell for next year can be determined which would be used in Monte Carlo simulation to obtain the operational risk capital.

Now the predictive distribution of N_{T+1}, that is the expected frequency of next year, can be deduced with the following steps:

$$P\left(N_{T+1} = m \mid N_T = n\right) = \int f\left(m \mid \lambda\right) \pi\left(\lambda \mid n\right) \mathrm{d}\lambda = \int \mathrm{e}^{-\lambda} \frac{\lambda^m}{m!} \frac{\lambda^{\alpha_T-1}}{\left(\beta_T\right)^{\alpha_T} \Gamma(\alpha_T)} \mathrm{e}^{-\frac{\lambda}{\beta_T}} \mathrm{d}\lambda$$

$$= \frac{\left(\beta_T\right)^{-\alpha_T}}{\Gamma(\alpha_T) m!} \int \mathrm{e}^{-\left(1+\frac{1}{\beta_T}\right)\lambda} \lambda^{\alpha_T+m-1} \mathrm{d}\lambda = \frac{\Gamma(\alpha_T+m)}{\Gamma(\alpha_T) m!} \left(\frac{1}{1+\beta_T}\right)^{\alpha_T} \left(\frac{\beta_T}{1+\beta_T}\right)^m,$$

a negative binomial distribution. It is assumed that given $\Lambda = \lambda$, N_{T+1} and N are independent. The expected number of events over the next year, given past observations, $E[N_{T+1}|N]$, *i.e.* mean of the *NegBin*(α_T,1/(1 + β_T)), which is also mean of the posterior distribution in this case, allows good interpretation as follows [1]:

$$E\left[N_{T+1} \mid N = n\right] = E\left[\lambda \mid N = n\right] = \alpha_T \beta_T = \beta \frac{\alpha + n_i}{1 + \beta \times T} = \omega_T \hat{\lambda}_T^{MLE} + \left(1 - \omega_T\right)\lambda_0.$$

Here

1) $\hat{\lambda}_T^{MLE} = \frac{1}{T}\sum_{i=1}^{T} n_i$ is the estimate of λ using the observed count only.

2) $\lambda_0 = \alpha\beta$ is the estimate of λ using the prior distribution only.

3) $\omega_T = \frac{\beta T}{1+\beta T}$ is the credibility weight in [0,1) used to combine λ_0 and $\hat{\lambda}_T^{MLE}$.

Hence, as the number of observed year T increases, the credibility weight ω_T increases and vice versa. That is, the more observations we have, the greater credibility weight we assign to the estimator based on the observed counts, while the lesser the credibility weight is attached to the prior distribution.

3.2.2. Posterior Severity Distribution
Similarly, using the Bayesian rule the severity posterior distribution can be expressed as:

$$\pi\left(\mu_j \mid y_j\right) \propto f\left(y_j \mid \mu_j\right) \pi\left(\mu_j\right) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left[-\frac{\left(\mu_j - \mu_0\right)^2}{2\sigma_0^2}\right] \prod_{k=1}^{K_J} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\left(y_j^k - \mu_j\right)^2}{2\sigma^2}\right]. \tag{13}$$

After some algebraic steps Equation (13) can be reduced to a normal distribution

$$\pi\left(\mu_j\middle|y_j\right) \propto f\left(y_j\middle|\mu_j\right)\pi\left(\mu_j\right) = \frac{1}{\sqrt{2\pi}\sigma_{0,n}}\exp\left[-\frac{\left(\mu_j-\mu_{0,n}\right)^2}{2\sigma_{0,n}^2}\right], \tag{14}$$

where $\mu_{0,n}$ and $\sigma_{0,n}$ are defined by (8). With the posterior distribution of severity mean estimated the expected value of the log severity loss for next year is

$$E\left(y_{n+1}\middle|y_n\right) = \mu_{0,n}.$$

Now the posterior severity mean defined in Equation (8) can be rewritten as

$$\mu_{0,n} = \frac{\mu_0 + \omega\sum_{i=1}^{n}y_i}{1+n\times\omega} = \omega_n\bar{y}_n + \left(1-\omega_n\right)\mu_0,$$

where

1) $\bar{y}_n = \dfrac{1}{n}\sum_{i=1}^{n}y_i$ is the estimate of μ using the observed losses only.

2) μ_0 is the estimate of μ using the prior distribution only.

3) $\omega_n = \dfrac{n}{n+\sigma^2/\sigma_0^2}$ is the credibility weight in [0,1) used to combine μ_0 and $\bar{y}_n$.

As more and more data are observed the weight increases. As a result more weight is placed on the observed data than the prior distribution.

Once the frequency $f\left(.\middle|\lambda_j\right)$ and severity $g\left(.\middle|\mu\right)$ densities in the risk cell are chosen and the posterior densities $\pi\left(\theta\middle|y\right)$, $\theta = \left(\lambda_j,\mu\right)$ are estimated, the predictive annual loss density in the risk cell can be calculated using Monte Carlo procedure. The following steps will be followed while estimating the predictive distribution.

---

1. For $k = 1,\cdots,K$

    a. For a given risk, parameters $\theta = \left(\lambda_j,\mu\right)$ are simulated from the respective posterior distributions $\pi\left(\theta\middle|y\right)$.

    b. Given $\theta = \left(\lambda_j,\mu\right)$, the annual number of events N is estimated from $f\left(.\middle|\lambda_j\right)$ and severities $X^{(1)},\cdots,X^{(N)}$ from $g\left(.\middle|\mu\right)$. The annual loss $Z^{(k)} = \sum_{n=1}^{N}X^{(n)}$ is then calculated.

2. Next k

The annual losses $Z^{(1)},\cdots,Z^{(k)}$ are samples from the predictive density. The 0.999 quantile $Q_{0.999}^{P}$ can then be calculated using the usual procedure.

Figure 2 shows the flow chart for the Bayesian approach compared to the conventional frequentist approach.

Now to add the low frequency but high severity (heavy tail) loss data a somewhat ad-hoc procedure is implemented. Frequency information for different loss buckets, e.g. \$1M - \$5M in one bucket, \$5M - \$25M in another bucket etc., is available from the simulated data. These frequency data are typically provided by the bank experts for the respective loss buckets. The Poisson distribution would be used for each of the bucket frequencies. However, point estimation of the Poisson parameters is used. For severity, the posterior distribution as derived in Equation (14) would be utilized. A special adhoc method would be utilized to make sure that the severity is estimated within the defined ranges (\$1M - \$5M, \$5M - \$25M etc.). This would be illustrated more in Section 4. The MC algorithm as described earlier would then be used to estimate the annual high severity losses. The only difference is that the Poisson parameter λ has a point estimate (provided by expert opinion) and has no uncertainty distribution. When the annual losses for the internal data and the high severity data (structured scenario data) are added to obtain the total annual loss, the parameter uncertainty is taken into consideration. The 0.999 quantile $Q_{0.999}^{P}$ for the total loss is estimated.
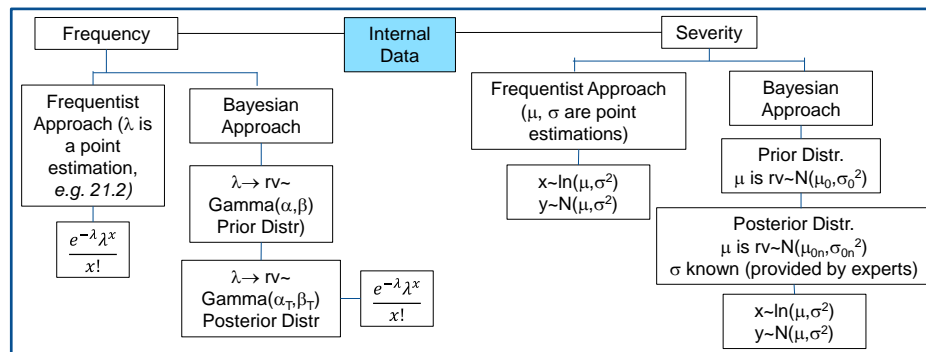
Figure 2. Bayesian flow chart as compared to the frequentist approach for the estimation of loss severity.

Figure 3 shows the flow chart of loss calculation for the expert opinion data for the Bayesian approach.

The posterior models developed in this section are used in the next section to estimate the operational risk capital using simulated bank data.

4. Estimation of Operational Risk Capital Using Bayesian Approach

In this section simulated bank data is used for the estimation of risk capital using the Bayesian inference method outlined in the earlier section. The bank data that is used belongs to the ET4, Clients, products, and business practices, event type as outlined in [10]. The simulated bank data has 152 internal and external loss data sets. These loss data have been categorized into 16 risk cells as outlined in **Table 1** and **Table 2**.

First the frequency and severity priors and posteriors are estimated.

4.1. Estimation of Frequency Prior

A maximum likelihood estimation (MLE) procedure is outlined in the previous section for the estimation of the prior distribution parameters. Equation (5) defines the log-likelihood function for the estimation of the prior gamma distribution parameters α and β.

Here J is the total number of risk cells which is 16. Data is available for the last seven years. Hence K_J, which defines the number previous years, is 7. The quantity n_j is the total number of loss events for the j-th risk cell for the past 7 (2004-2010) years. The **Table 1** summarizes the total frequency n_j for each of the risk cells. **Table 2** lists the business lines to which these risk cells belong.

A grid of α and β values (α ranging from 1 to 2.5 and β ranging from 0 to 2) is created and the log-likelihood function (5) is calculated for each of the α,β grid values. A 3D mesh and contour diagram of the log-likelihood values are plotted as shown in **Figure 4**. From the figure it is clear that the log-likelihood values would be maximized roughly for a value of $\alpha = 1.7$ and $\beta = 0.8$.

Alternatively, a non-linear optimization algorithm may be implemented for the maximum likelihood estimation (MLE) of the distribution parameters. For this purpose SAS procedure NLP is used. The NLP procedure (**N**on**L**inear **P**rogramming) offers a set of optimization techniques for minimizing or maximizing a continuous nonlinear function of decision variables, with lower and upper bound, linear and nonlinear, equality and inequality constraints.

The maximum likelihood formulation for the frequency prior is coded in SAS. The results of the NLP non-linear optimization procedure is shown in **Table 3**. The maximum optimum value of the function is −71.16. The optimal values for the distribution parameters α and β are listed in the table which are close to the values obtained in **Figure 4**.

4.2. Estimation of Severity Prior

The severity prior has been defined as a normal distribution. The likelihood function that has to be maximized to estimate the distribution parameters of μ_0 and σ_0 is defined by Equation (9).

Similar to the frequency prior the severity likelihood function is calculated for a grid of μ_0 and σ_0 values. A
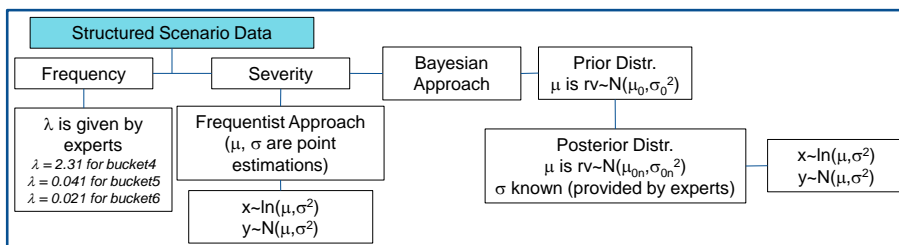
Figure 3. Bayesian flow chart for the high severity (heavy tail, expert opinion) data.

Table 1. Summary of the risk cell total frequency.

Risk Cells	Total Frequency n_j
Retail Banking, Michigan	13
Retail Banking, Ohio	8
Retail Banking, Texas	4
Commercial Banking	37
Corporate Agency	4
Corporate Trust	25
Card Service	3
Asset Management	4
Treasury	4
Market Making	2
Advisory Service	3
Proprietary Positions	13
External Clients	15
Retail Brokerage	5
Merchant Banking	4
Private Banking	8

Table 2. Business lines of the risk cells of **Table 1**.

Risk Cells	Business Line
Merchant Banking	Corporate Finance
Market Making, Proprietary Positions, Treasury	Trading and Sales
Retail Banking Michigan, Ohio, Texas, Private Banking, Card Services	Retail Banking
Commercial Banking	Commercial Banking
External Clients	Payment and Settlement
Corporate Agency, Corporate Trust	Agency Services
Asset Management	Asset Management
Retail Brokerage	Retail Brokerage

Table 3. Estimated distribution parameters.

	PROC NLP: Nonlinear Maximization Optimization Results Parameter Estimation				
N Parameter	Gradient Approximate Estimate	Std Error	Approx *t*-Value	Pr > |t|	Objective Function
1 alpha_est	1.725238	0.670045	2.574808	0.020355	1.6136283E−9
2 beta_est	0.786641	0.346161	2.272471	0.037202	2.333316E−9

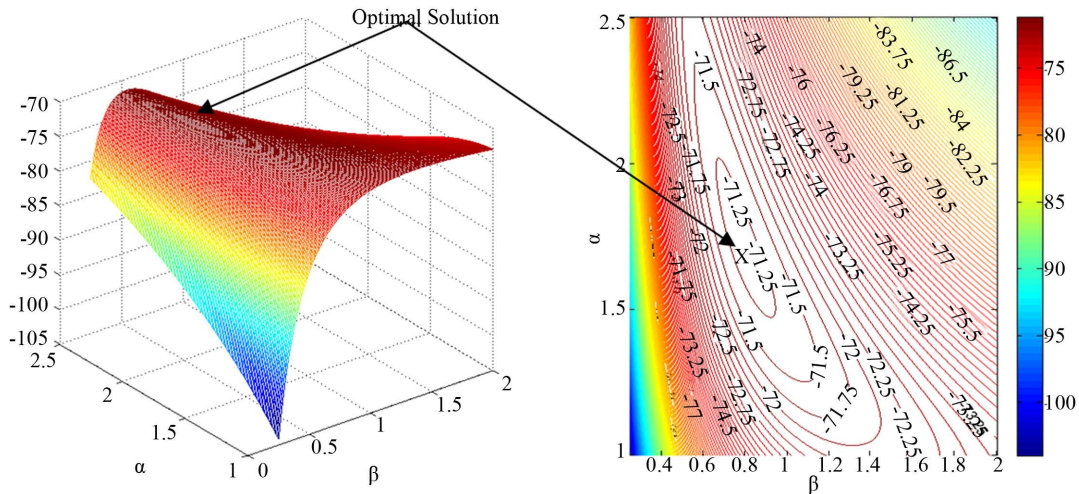Value of Objective Function = −71.16764281.

Figure 4. 3D mesh plot (a) and a contour plot (b) of the frequency prior log-likelihood values.

range of 3 to 6 is used for μ_0 and a range of 1 to 3 is used for σ_0. A 3D mesh and contour diagram of the likelihood function values is shown in **Figure 5**. It is clear from the figure that roughly for values of $\mu_0 = 4.58$ and $\sigma_0 = 1.51$ the likelihood function is maximized. The maximum value of the likelihood function is roughly −356.25.

The SAS NLP procedure is also used for the MLE of the distribution parameters. The optimal distribution parameters as computed by the SAS NLP procedure are listed in **Table 4**. The estimated optimal value of the objective function is −356.16.

4.3. Estimation of Frequency Posterior

The closed form solution of the posterior distribution is presented in Equation (11). Once the prior distribution parameters α and β are estimated the posterior distributions for each of the cells can be estimated using Equation (11). **Figure 6** shows the posterior distribution of the Poisson frequency parameter λ_j for each of the "*j*" cells outlined earlier for the event ET4. The distribution of the Poisson parameter λ_j would be used to calculate the Poisson frequency of each cell for the Monte Carlo simulation of estimating the risk capital. This is described later in this section.

4.4. Estimation of Severity Posterior Distribution

The posterior distribution for the severity mean parameter μ is derived as normal in an earlier section. Normal distribution mean and standard deviations are estimated in Equation (9). Posterior distributions of each of the severity mean for each of the cells are plotted in **Figure 7**. Once the parameter distributions are defined, we are in a position to estimate the risk capital using the Bayesian approach, which is illustrated next.

Risk Capital Estimation Using Bayesian Inference
In the Bayesian inference approach the model distribution parameter uncertainty is taken into consideration as outlined earlier. In this paper the Poisson frequency parameter λ_j and log-normal severity parameter μ_j uncer-

Table 4. Maximum likelihood estimate of the distribution parameters.

PROC NLP: Nonlinear Maximization Optimization Results Parameter Estimation					
N Parameter	Gradient Approximate Estimate	Std Error	Approx *t*-Value	Pr > |*t*|	Objective Function
1 mu_est	4.587306	0.430230	10.662455	1.1171596E−8	−6.33549E−12
2 sigma_est	1.518823	0.346161	4.406243	0.000441	2.342635E−11

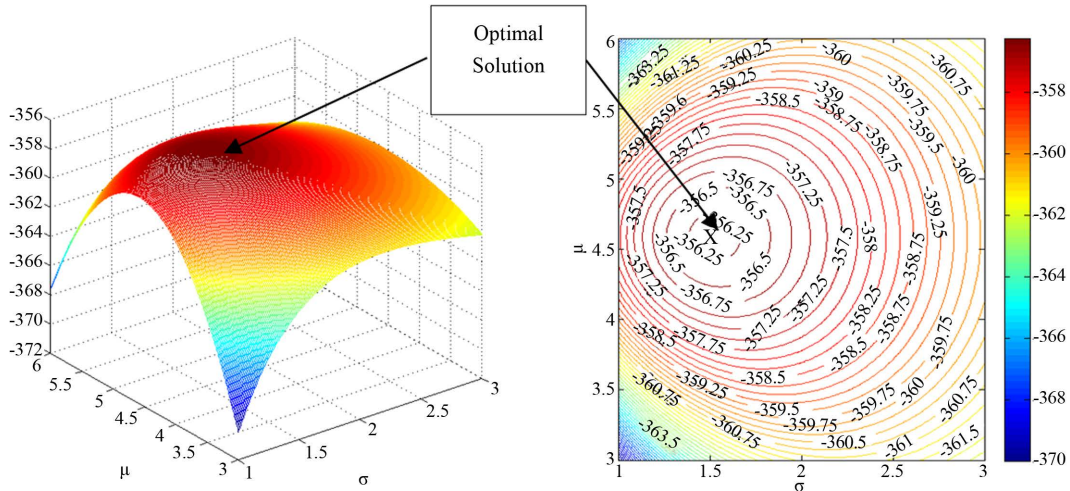Value of Objective Function = −356.1570041.

Figure 5. 3D mesh plot (a) and contour plot (b) of the severity likelihood function.

Figure 6. Probability distribution of posterior frequency mean for each of the cells.

tainties are taken into consideration by following the Bayesian inference approach. The log-normal standard deviation is assumed known. In this analysis we would fit the entire ET4 loss data into log-normal distribution to determine a point estimation of the standard deviation. SAS NLMIXED procedure is used to fit the ET4 data into a log-normal distribution. The estimated value of the standard deviation σ is shown in the **Table 5**.

This value of σ is used in determining the posterior distributions of the parameters as well as the frequency and severity distributions. Once the frequency $f\left(.|\lambda_j\right)$ and severity $g\left(.|\mu_j\right)$ densities in the risk cell are

chosen and the posterior density $\pi(\theta|y)$, $\theta = (\lambda_j, \mu_j)$ is estimated, the predictive annual loss density in the risk cell can be calculated using Monte Carlo procedures. Steps that need to be carried out for estimating the predictive distribution is presented earlier in this section.

Total number of Monte Carlo simulation K of 1,000,000 is used. The annual losses $Z^{(1)}, \cdots, Z^{(k)}$ are samples from the predictive density. This process is carried out for each of the 16 cells defined earlier.

The same Monte Carlo simulation is carried out for the structured scenario data as well. The structured scenarios are the low frequency and high severity loss data. The high severity losses are broken into six buckets, buckets 1 - 3 have the lower level of losses while bucket 4 to bucket 6 have the higher levels. In this article we would consider buckets 4 - 6 which covers the higher loss levels and would neglect buckets 1 - 3 since those would be already included by the internal and external data. Losses between ($1M - $5M) are in bucket 4, ($5M - $25M) are in bucket 5, and losses higher than $25M are in bucket 6. For high severity losses only the bucket frequencies are available which are defined by the bank experts each year. The bucket frequencies that would be used for this simulation are listed in **Table 6**. Bucket 6 frequency of 0.031 indicates that such a loss (>$25M) would occur roughly once in every 30 years.

In this simulation an adhoc procedure is adopted to estimate the high severity low frequency losses. Frequency data listed in **Table 6** would be directly used in the Poisson frequency distribution. Hence, a point estimate (the bucket frequencies) would be used and no uncertainty in the Poisson parameter distribution would be considered. However, for severity the prior distribution of the severity mean that was defined during the internal data simulation would be considered. In order to determine the posterior distribution for severity mean for the structured scenario data, the entire ET4 internal data would be used as the loss cell. Hence, for the estimation of the $\sum_{k=1}^{N} y_j^k$ in the posterior distribution parameter $\mu_{0,n}$ (Equation (8)), the entire ET4 internal/external loss data would be used as the "*j*" cell and the corresponding posterior distribution of severity parameter μ would be used in the Monte Carlo simulation. It is to be noted that the sum is taken over all data N, hence the normalizing factor N would be used instead of K_j in Equation (8). As before the standard deviation σ would be considered known and the prior established value of 2.1408 would be used, which was determined by fitting the ET4 data into log-normal distribution.

The structured scenario data follows the same Monte Carlo simulation steps outlined for the internal data. For the structured scenario data some additional steps are followed in order to make sure that the estimated loss severity falls within the specific bucket range. The steps are outlined below.

---

a. Estimate the probability for the lower range of the bucket
- *P_BEG4=CDF('LOGNORMAL', BEGPOINT4, MU ,&SIGMA);*
b. Estimate the probability for the upper range of the bucket
- *P_END4=CDF('LOGNORMAL', ENDPOINT4 ,MU ,&SIGMA);*
c. Estimate probability for any value within the upper and lower range
- *PR4 = P_BEG4 + (P_END4-P_BEG4)*RANUNI(&SEED2);*
d. Calculate the severity for the bucket
- *SEV4 = EXP(MU+&SIGMA*PROBIT(PR4)).*

Figure 8 outlines the entire procedure for the bank internal/external data as well as the structured scenario (bank expert) data.

Once both the internal and the structured scenario severity losses are estimated, these loss data are merged and added to get the aggregated total. This total is then sorted by the level of severity. Typically 99.9 percentile value is selected as the total capital charge. The 99.9, 99.93, and 99.97 percentile values of the total capital charge are listed in **Table 7**. The capital charge distribution with the 99.9% percentile value is additionally shown in **Figure 9**.

Finally, for comparison the estimation of the capital charge is carried out using the frequentist approach. In the frequentist approach the point estimation of the severity parameters μ and σ are estimated by the MLE by fitting the ET4 severity data into log-normal distribution. The NLMIXED procedure described earlier is used to estimate μ and σ for the entire ET4 internal severity data. The result is shown in **Table 8**.

For the Poisson frequency distribution, the average frequency per year (152 loss events/7 years = 21.7) is used as the mean parameter λ. The Monte Carlo simulation for the internal data and the structured scenario data is carried out. However, instead of using any parameter uncertainty the point estimation of the parameters are used. The total capital estimated using this more conventional approach is listed in **Table 9**. The capital estimated by
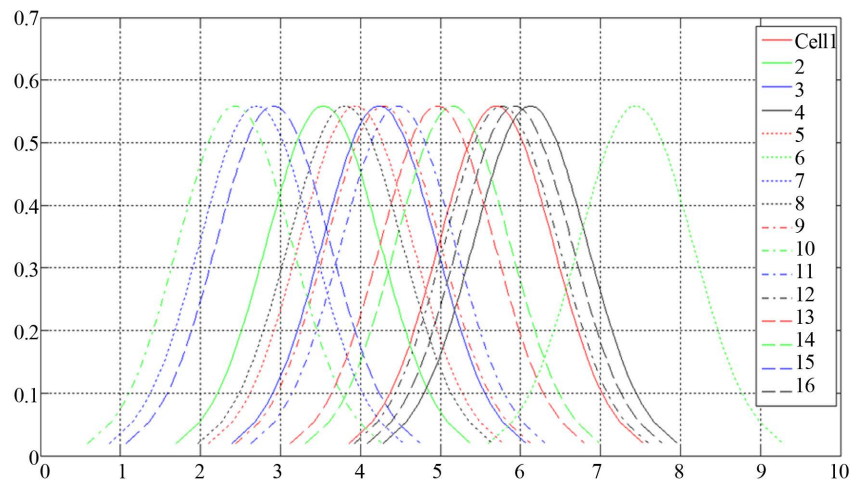
Figure 7. Posterior distribution of the severity mean.

Table 5. Estimated standard deviation of the severity log-normal distribution.

The NLMIXED Procedure Parameter Estimates								
Parameter	Estimate	Standard Error	DF	*t*-Value	Pr > |t|	Alpha	Lower	Upper	Gradient
SIGMA	2.1408	0.1226	152	17.46	<.0001	0.05	1.8986	2.3830	2.911E−7

Table 6. Structured scenario bucket frequency.

Bucket No.	4 ($1M - $5M)	5 ($5M - $25M)	6 (>$25M)
Frequency	2.22	0.262	0.031

Table 7. 99.9, 99.93, and 99.97 percentile values of the total capital charge using Bayesian inference method.

Internal and SS Data Percentile		
Obs	Severity	Percentile
1	$152,829,497	99.90%
2	$178,854,622	99.93%
3	$262,064,437	99.97%

Table 8. Estimated values of μ and σ.

The NLMIXED Procedure Parameter Estimates								
Parameter	Estimate	Standard Error	DF	*t*-Value	Pr> |t|	Alpha	Lower	Upper	Gradient
SIGMA	2.1408	0.1226	152	17.46	<.0001	0.05	1.8986	2.3830	2.911E−7
MU	9.3490	0.1734	152	53.91	<.0001	0.05	9.0063	9.6916	7.941E−7

Table 9. 99.9, 99.93, and 99.97 percentile values of the total capital charge using frequentist approach.

Obs	Severity	Percentile
1	$159,762,828	99.90%
2	$185,098,353	99.93%
3	$262,180,889	99.97%

Internal Data

	Cell 1: $n_i=13$			
Monte Carlo Simulation	Poisson Parameter from Gamma Posterior Distribution (Internal Data)	Estimated Frequency from Poisson Distribution	Parameter for Severity Posterior Distribution	Losses for Each Cells: $X_1+X_2+X_3+X_4+\ldots+X_N$
	λ_{11}	N_1	μ_{11}	X_{11}
	λ_{12}	N_2	μ_{12}	X_{12}
	λ_{13}	N_3	μ_{13}	X_{13}
	.	.	.	.
	.	.	.	.
	.	.	.	.
	.	.	.	.
	.	.	.	.
	λ_{1M}	N_M	μ_{1M}	X_{1M}

Structured Scenario Data

	Fixed (Point Estimation) Poisson Parameters for the Buckets			Estimated Frequency from Poisson Distribution	Parameter for Severity Posterior Distribution	Losses for Each Cells: $X_1+X_2+X_3+X_4+\ldots+X_N$		
Monte Carlo Simulation	Bucket 4	Bucket 5	Bucket 6			Bucket 4	Bucket 5	Bucket 6
	2.31	0.041	0.021	N_1	μ_{B41}	X_{B41}	X_{B51}	X_{B61}
	2.31	0.041	0.021	N_2	μ_{B42}	X_{B42}	X_{B52}	X_{B62}
	2.31	0.041	0.021	N_3	μ_{B43}	X_{B43}	X_{B53}	X_{B63}
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	2.31	0.041	0.021	N_M	μ_{B4M}	X_{B4M}	X_{B5M}	X_{B6M}

Loss Sum for All Cells

Losses for Each Cells: $X_1+X_2+X_3+X_4+\ldots+X_N$									Loss Sum	Sorted Loss Sum
1	2	3	4	…….	16	Bucket 4	Bucket 5	Bucket 6		
X_{11}	X_{21}	X_{31}	X_{41}	…….	X_{161}	X_{B41}	X_{B51}	X_{B61}	Z_1	Z_1
X_{12}	X_{22}	X_{32}	X_{42}	…….	X_{162}	X_{B42}	X_{B52}	X_{B62}	Z_2	Z_2
X_{13}	X_{23}	X_{33}	X_{43}	…….	X_{163}	X_{B43}	X_{B53}	X_{B63}	Z_3	Z_3
.	.	.	.	…….	.	.	.	.	.	.
.	.	.	.	…….	.	.	.	.	.	.
.	.	.	.	…….	.	.	.	.	.	.
.	.	.	.	…….	.	.	.	.	.	$Z_{99.9\%}$
.	.	.	.	…….	.	.	.	.	.	.
.	.	.	.	…….	.	.	.	.	.	$Z_{99.93\%}$
.	.	.	.	…….	.	.	.	.	.	.
.	.	.	.	…….	.	.	.	.	.	$Z_{99.97\%}$
.	.	.	.	…….	.	.	.	.	.	.
X_{1M}	X_{2M}	X_{3M}	X_{4M}	…….	X_{16M}	X_{B4M}	X_{B5M}	X_{B6M}	Z_M	Z_M

Figure 8. Monte Carlo simulation steps for the internal/external and structured scenario data.
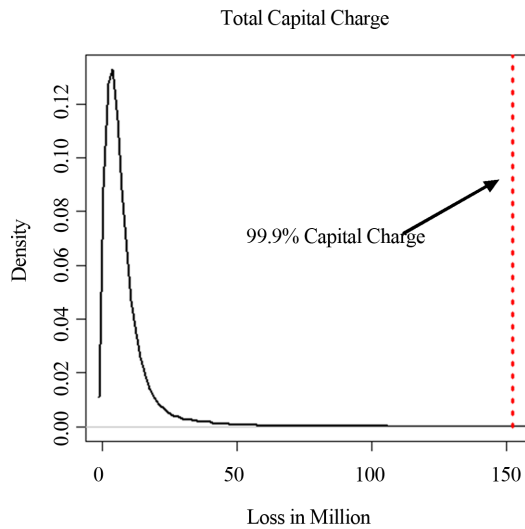

Total Capital Charge

99.9% Capital Charge

Figure 9. Capital charge distribution.

the later approach is somewhat higher than the one estimated by the Bayesian Inference method.

5. Summary and Conclusion

A Bayesian inference method is carried out for the estimation of the total capital for a bank to cover the operational risk of the ET4 category. For the Bayesian inference method the prior distribution of the frequency and

severity mean parameters λ and μ respectively are determined using the data. These prior distributions are then used to estimate the posterior distribution as additional loss data are available from the bank internal or external sources. This approach would have advantage over the more conventional frequentist approach since the Bayesian method includes the parameter uncertainties. The total capital using the Bayesian inference approach and the conventional (frequentist) approach is estimated following the Monte Carlo simulation. Total capital charges using these two approaches are somewhat similar with the frequentist approach estimating slightly higher values.

References

[1] Shevchenko, P.V. (2011) Modeling Operational Risk Using Bayesian Inference. Springer-Verlag Publishing Company. http://dx.doi.org/10.1007/978-3-642-15923-7

[2] King, J.L. (2001) Operational Risk Measurement and Modeling. Wiley.

[3] Cruz, M.G. (2002) Modeling, Measurement and Hedging Operational Risk. Wiley.

[4] Panjer, H.H. (2006) Operational Risk: Modeling Analysis. Wiley. http://dx.doi.org/10.1002/0470051310

[5] Franchot, A. and Roncalli, T. (2002) Mixing Internal and External Data for Managing Operational Risk. Working Paper, Groupe de Recherche Operationnelle.

[6] Shevchenko, P.V. and Wüthrich, M.V. (2006) The Structural Modeling of Operational Risk via Bayesian Inference: Combining Loss Data with Expert Opinion. *The Journal of Operational Risk*, **1**, 3-36.

[7] Lambrigger, D.D., Shevchenko, P.V. and Wüthrich, M.V. (2007) The Quantification of Operational Risk using Internal Data, Relevant External Data and Expert Opinions. *The Journal of Operational Risk*, **2**, 3-27.

[8] Peters, G.W., Shevchenko, P.V. and Wüthrich, M.V. (2009) Dynamic Operational Risk: Modeling Dependence and Combining Different Data Sources of Information. *The Journal of Operational Risk*, **4**, 69-104.

[9] Chernobai, A.S., Rachev, S.T. and Fabozzi, F.J. (2007) Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis. John Wiley and Sons, Inc.

[10] Basel Committee on Banking Supervision: Results from the 2008 Loss Data Collection Exercise for Operational Risk. Bank for International Settlement (2009). www.bis.org

[11] Allen, L., Boudoukh, J. and Saunders, A. (2005) Understanding Market, Credit and Operational Risk: The Value-at-Risk Approach. Blackwell Publishing, Oxford.

[12] Cruz, M.G. (Ed.) (2004) Operational Risk Modeling and Analysis: Theory and Practice. Risk Books, London.

[13] McNeil. A.J., Frey, R. and Embrechts, P. (2005) Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton, NJ.