Quantifying operational risk exposure by combining incident data and subjective risk assessments

A. B. Huseby
University of Oslo, Norway

J. Thomsen
Norges Bank Investment Management, Norway

ABSTRACT: Quantifying operational risk exposure typically involves gathering information from several sources, including historical data as well as subjective assessments. Using historical data one can estimate both an incident frequency distribution, as well as an incident consequence distribution. Based on these two distributions a simulation model can be established. However, by limiting the focus to data related to incidents which may reappear in the future, one is often left with a relatively short incident history. In order to improve the risk quantification, it is often necessary to include subjective risk assessments as well. In the present paper we propose three models for how to combine these two sources of information. In the first model we assume that the two sources are completely disjoint, while in the second model the two sources are assumed to overlap completely. The third model represents an intermediate situation where the two sources are partially overlapping. This third model contains the two first models as limiting cases. The models are illustrated and compared in an extensive numerical example.

1 INTRODUCTION

When quantifying operational risk exposure one often needs to gather information from several sources. Such sources include statistical data based on historical observations as well as assessments made by a panel of experts. For the statistical data this represents observations of actual incidents collected through a suitable number of years. By counting the number of incidents per unit of time one can estimate an incident frequency distribution. Moreover, by looking at the occurred consequences an incident consequence distribution can be estimated. Together these two distributions enable the analyst to build a well-defined simulation model. In the present paper we refer to this as the *incident database model*. While this model may represent an adequate summary of the past, it may not work quite as well for the future. Due to changes in the overall risk picture, older incidents may sometimes be less likely to occur in the future. By limiting the focus to data related to incidents which may reappear in the future, one is often left with a relatively short incident history. Moreover, there may be important risk factors missing in the incident database. Such risk factors could arise from recent threats or rare incidents with possibly severe consequences. Thus, in order to improve the risk quantification, it is often necessary to include subjective risk assessments as well. Typically, this is done by a panel of experts identifying a set of potential risk factors. For all these factors the experts assess the frequency of occurrence as well as the risk factor consequence distribution. Based on these assessments another simulation model can be established, referred to here as the *risk factor model*.

Obviously the incident database model and the risk factor model may be produce results which differ significantly. Thus, there is a need to handle these differences by developing some way of combining the two models. In this paper we shall discuss three different approaches to combining incident database model and risk factor model. Underlying all three approaches is the assumption that the incident data and the subjectively assessed risk factor model can be viewed as two different but possibly overlapping sources of information. The three approaches represent different cases with respect to the degree of overlap.

Before introducing the models we review results related to compound processes and how to incorporate parameter uncertainty into the models. The proposed models are illustrated by considering a numerical example.

2 COMPOUND PROCESSES

A common approach to modelling the accumulated consequences from a series of events is using a *compound process*. Such processes are used extensively in e.g., insurance mathematics. For an excellent introduction to this field see (Bølviken 2014). This topic of compound processes is also covered in (McNeil et al. 2005).

According to this approach the accumulated consequences in a time interval $[0, t]$, denoted $Z(t)$ can be expressed as:

$$Z(t) = \sum_{i=1}^{N(t)} X_i, \tag{1}$$

where $N(t)$ denotes the number of events in $[0, t]$, while X_i denotes the consequence of the ith event, $i = 1, 2, \ldots$. Here $N(t)$ is a non-decreasing stochastic process with non-negative integer values such that $N(0) = 0$, also referred to as a *counting process*. We refer to this as *the event generating process*. Typically one assumes that $X_1, X_2, \ldots$ are stochastically independent and identically distributed with distribution F_X, which we refer to as the *consequence distribution*.

Using standard results from the theory of compound processes it is very easy to show that:

$$\mathrm{E}[Z(t)] = \mathrm{E}[N(t)] \cdot \mathrm{E}[X_i], \tag{2}$$

$$\mathrm{Var}[Z(t)] = \mathrm{E}[N(t)] \cdot (\mathrm{Var}[X_i] + \mathrm{E}[X_i]^2). \tag{3}$$

Unless one needs to include time dependence or non-stationarity into the model, a Poisson process is a common choice for the event generating process. This type of process is also recommended in (Adachi et al. 2011), and will be used throughout the present paper. A compound process where the events are generated from a Poisson prosess, is referred to as a *compound Poisson process*. See e.g., (Adelson 1966) or (McNeil et al. 2005) for more details.

When the event generating process is a Poisson process, it follows that:

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n = 0, 1, 2, \ldots, \tag{4}$$

where $\lambda > 0$ is a parameter denoting the *event intensity* per time unit. Thus, for $t > 0$, $N(t)$ has a Poisson distribution with parameter λt, i.e., $N(t) \sim \mathrm{Po}(\lambda t)$. In this context we define a time unit to be *one year*. We also introduce:

$$N_k = N(k) - N(k-1), \quad k = 1, 2, \ldots.$$

Thus, N_k is the number of events in the kth year, $k = 1, 2, \ldots$. Since $N(t)$ is assumed to be a Poisson process, it follows that $N_1, N_2, \ldots$ is a sequence of independent variables, all with the same Poisson probability distribution:

$$P(N_k = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \; n = 0, 1, 2, \ldots, \; k = 1, 2, \ldots.$$

That is, $N_k \sim \mathrm{Po}(\lambda)$, $k = 1, 2, \ldots$.

2.1 Parameter uncertainty

Realistically, the parameter of the Poisson process is typically an uncertain quantity. In order to include this uncertainty into the model, we assess a probability distribution representing our prior knowledge about the parameter. A convenient choice is the natural conjugate distribution, i.e., a Gamma distribution with positive parameters α and β. The density of this distribution, denoted by π, is given by:

$$\pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \quad \lambda > 0.$$

In order to assess the hyperparameters α and β, one may imagine the prior distribution as a result of observing the process in β time units. (See (Huseby 1989).) When the prior knowledge is considered to be weak, the value of β should be chosen to be small, typically less than or equal to 1. The value of α is assessed indirectly using the fact that the prior expected value of λ is:

$$\mathrm{E}[\lambda] = \frac{\alpha}{\beta}.$$

Thus, by assessing a prior point estimate for λ, denoted λ_π, the parameter α is typically chosen so that:

$$\alpha = \lambda_\pi \beta.$$

If the parameter uncertainty is taken into account, the resulting distribution for $N(t)$, referred to as the *prior predictive distribution*, becomes:

$$P(N(t) = n) = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)\Gamma(n+1)} (1 - \frac{t}{\beta + t})^\alpha (\frac{t}{\beta + t})^n,$$

for $n = 0, 1, 2, \ldots$. If α is a non-negative integer, this distribution is called the *negative binomial distribution*. Moreover, by inserting $t = 1$ into this expression, we obtain:

$$P(N_k = n) = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)\Gamma(n+1)} (1 - \frac{1}{\beta + 1})^\alpha (\frac{1}{\beta + 1})^n,$$

for $n = 0, 1, 2, \ldots$ and $k = 1, 2, \ldots$.

As one gets observations from the Poisson process, the prior uncertainty about λ has to be updated using Bayes' theorem. See e.g., (Berger 2010). In particular, if we have observed the process in τ units of time, during which we have recorded ν events, then the resulting uncertainty distribution, referred to as the *posterior distribution* is:

$$\pi(\lambda|\tau,\nu) = \frac{(\beta+\tau)^{\alpha+\nu}}{\Gamma(\alpha+\nu)}\lambda^{\alpha+\nu-1}e^{-(\beta+\tau)\lambda}, \quad \lambda > 0.$$

For simplicity we denote posterior distribution by π', and we observe that this can be written as:

$$\pi'(\lambda) = \frac{\beta'^{\alpha'}}{\Gamma(\alpha')}\lambda^{\alpha'-1}e^{-\beta'\lambda}, \quad \lambda > 0.$$

Thus the posterior distribution is another Gamma distribution with parameters $\alpha' = \alpha + \nu$ and $\beta' = \beta + \tau$. Furthermore, the *posterior predictive distribution*, i.e., the distribution for $N(t)$ given τ and ν has the same form as the prior predictive distribution, except that α is replaced by α' and β is replaced by β'.

In the following subsections we will use compound Poisson processes with gamma priors for both the incident database model and the risk factor model, and we will show how these models can be fitted using the available information.

2.2 The incident database model

The incident database model is fitted by using data from incidents that have occurred. Thus, we assume that we have observed the event process for a period of τ units of time. In this period we have observed ν events. For each of these events we have also recorded their resulting consequences, denoted $X_1, \ldots, X_\nu$ respectively. We then consider the number of events in an upcoming period of length 1 year, denoted N_I, where the subscript I indicates that we are considering the incident database model. For a given λ_I, $N_I \sim \text{Po}(\lambda_I)$. Moreover, we assume that $\lambda_I \sim \text{Gamma}(\alpha_I, \beta_I)$. Using the results from Subsection 2.1, we then get that the posterior predictive distribution for N_I is given by:

$$P(N_I = n) = \frac{\Gamma(\alpha'_I + n)}{\Gamma(\alpha'_I)\Gamma(n+1)}(1 - \frac{1}{\beta'_I + 1})^{\alpha'_I}(\frac{1}{\beta'_I + 1})^n,$$

where $\alpha'_I = \alpha_I + \nu$ and $\beta'_I = \beta_I + \tau$.

Using the observed consequences, $X_1, \ldots, X_\nu$ we can fit a consequence distribution for the incident database model, denoted F_I. This can of course be done in many different ways ranging from a purely non-parametric approach using the empirical cumulative distribution to fitting various types of parametric distributions. In this context we have chosen to fit a lognormal distribution with mean value ξ_I and standard deviation σ_I, where ξ_I and σ_I are estimated using the observed sample mean and standard deviation from the given consequence data.

The distribution of the resulting accumulated consequences from the incident database model can now easily be estimated using Monte Carlo simulation. In each iteration we start out by sampling the number of events using the posterior predictive distribution for N_I. This can be done directly using the above derived distribution. Alternatively, the sampling can be done in two steps: In the first step λ_I is sampled from the posterior distribution, i.e., $\text{Gamma}(\alpha'_I, \beta'_I)$. Then in the second step N_I is sampled from the Poisson distribuion given the sampled value of λ_I.

Having sampled N_I, we then proceed by sampling N_I variables from the fitted consequence distribution, F_I, and add up the results.

2.3 The risk factor model

One of the reasons for including subjectively assessed operational risks in the calculation is the relatively short history of relevant incidents. Moreover, recent changes in both internal and external conditions may significantly change the various risk factors and even introduce new and completely unknown factors. Such changes will tyically not be represented in the incident database. As a result, the observed incidents alone may give an acceptable description of the impact distribution's body, but are in many cases not sufficient to describe the tail.

Through the use of a panel of experts a list of r potential operational risk factors is identified. A common approach to risk factor modelling is to assess the probability that the risk factor occurs within a unit of time, say a year. A weakness with this approach is that it does not allow repeated events within the same unit of time. In order to avoid this problem, and allow the risk factor to occur more than once, we prefer to use a Poisson model. That is, we assume that the number of occurrences of the risk factor within a unit of time is Poisson distributed with a suitable rate. In particular we let λ_s denote the rate parameter per year for the sth risk factor, and let N_s denote the number of events in this period from the sth process $s = 1, \ldots, r$.

Note that this implies that $P(N_s = 0) = e^{-\lambda_s}$. If λ_s is small, it follows by a Taylor expansion that $P(N_s = 0) \approx 1 - \lambda_s$, and hence, $P(N_s > 0) \approx \lambda_s$. If λ_s is small then we also have $P(N_s > 0) \approx P(N_s = 1)$. Thus, in such cases the difference between a Poisson model and a binary model is small, and the rate of the Poisson model is close to the probability that the risk factor occurs in the binary model.

As for the incident database model, we consider events in an upcoming period of length 1 year. The total number of events from all processes, denoted N_R is then:

$$N_R = \sum_{s=1}^{r} N_s.$$

Given the rates for the individual processes, $\lambda_1, \ldots, \lambda_s$, it is a well known property of Poisson processes that N_R has a Poisson distribution as well.

More specifically, $N_R \sim \text{Po}(\lambda_R)$, where:

$$\lambda_R = \sum_{s=1}^{r} \lambda_s.$$

The expert panel then assesses priors to each rate, and we assume that $\lambda_s \sim \text{Gamma}(\alpha_s, \beta_s)$, $s = 1, \ldots, r$. We recall that the β_s-parameters can be interpreted as a measure of the strength of the prior knowledge. In this context it is convenient to assume that these parameters are equal for all the r risk factors. Thus, letting β_R denote this common value, we assume that $\beta_s = \beta_R$, $s = 1, \ldots, r$. We also introduce:

$$\alpha_R = \sum_{s=1}^{r} \alpha_s. \tag{5}$$

Assuming that all the rates of the r processes are stochastically independent apriori, it is easy to show that the total rate λ_R is Gamma distributed as well. More specifically, $\lambda_R \sim \text{Gamma}(\alpha_R, \beta_R)$.

For each risk factor the expert panel assesses the mean value and standard deviation of the resulting consequences. The mean and standard deviation of the sth risk factor are denoted ξ_s and σ_s respectively, $s = 1, \ldots, r$.

Now, assume that we know that there has been an event related to one of the risk factors. However, the actual risk factor that caused this event is unknown. We denote the unknown index of this risk factor by S, and the resulting consequence by X_R. Then, it can be shown that:

$$P(S = s) = \frac{\alpha_s}{\sum_{i=1}^{r} \alpha_i} = \frac{\alpha_s}{\alpha_R}, \quad s = 1, \ldots, r.$$

By the assumptions on the consequence distributions it follows that for $s = 1, \ldots, r$:

$$\text{E}[X_R | S = s] = \xi_s,$$

$$\text{Var}[X_R | S = s] = \sigma_s^2.$$

Using elementary probability theory it is the easy to find the unconditional mean and variance of X_R, denoted respectively ξ_R and σ_R^2:

$$\xi_R = \text{E}[X_R] = \frac{\sum_{s=1}^{r} \xi_s \alpha_s}{\alpha_R}, \tag{6}$$

$$\sigma_R^2 = \text{Var}[X_R] = \frac{\sum_{s=1}^{r} (\xi_s^2 + \sigma_s^2) \alpha_s}{\alpha_R} - \xi_R^2. \tag{7}$$

As a combined consequence distribution for all the risk factors in the risk factor model we simply use a lognormal distribution with mean value ξ_R and standard deviation σ_R. We denote this distribution by F_R. The distribution of the resulting accumulated consequences from the risk factor model can then easily be estimated using Monte Carlo simulation in exactly the same way as for the incident database model.

3 THE COMBINED MODEL

In this section we turn to the problem of combining the incident database model and the risk factor model. In order to do so we will consider three different cases.

3.1 No overlap between the models

In the first case we assume that the sets of events covered by the two models do not overlap at all. Thus, when combining the two, we have to add the accumulated consequences from both models. However, the assumption that there is no overlap, also implies that during the period we have recorded events, none of these events could be related to any of the risk factors included in the risk factor model. This means that the recorded consequences does not contain any relevant information about the risk factor consequence distribution. Thus, there is no reason to change this in the simulations. The event generating process, however, needs to be updated as a result of this, which is easily done using the methodology introduced in Subsection 2.1. With zero observed events over a period of length τ, the posterior distribution for λ_R becomes another Gamma distribution with updated parameters $\alpha'_R = \alpha_R$ and $\beta'_R = \beta_R + \tau$.

The distribution of the resulting accumulated consequences from the combined model can again easily be estimated using Monte Carlo simulation. With no overlap between the models we generate N_I and N_R from their respective posterior predictive distributions. Then we generate N_I consequences from F_I and N_R consequences from F_R and add all the results.

3.2 Full overlap between the models

In the second case we assume that the set of events covered by the two models overlap completely. Thus, all the observed incidents are actually events of the kinds included in the risk factor model. Thus, a combined model should be obtained by updating the risk factor model using all the incident data. With ν observed events over a period of length τ, the posterior distribution for λ_R becomes another Gamma distribution with updated parameters $\alpha'_R = \alpha_R + \nu$ and $\beta'_R = \beta_R + \tau$.

Concerning the consequence distribution for the combined model this is derived by updating the consequence distribution for the risk factor model, F_R, according to the observed consequences. The updated consequence distribution is denoted F'_R. There are many ways of doing this ranging from stringent Bayesian methods to more ad hoc procedures. Here we take a simple weighted approach, where we replace the mean and the standard deviation of the F_R

by:

$$\xi_R' = c\xi_I + (1-c)\xi_R, \tag{8}$$

$$\sigma_R' = c\sigma_I + (1-c)\sigma_R, \tag{9}$$

where $c \in (0,1)$ is a suitable weight factor. As we have pointed out already, the α_R-parameter can be interpreted as a measure of the number of observations underlying the prior knowledge, while ν is the number of real observations. The weight factor should balance the strengths of these two sources. Hence, we propose the following weight factor:

$$c = \frac{\nu}{\alpha_R + \nu}.$$

The distribution of the resulting accumulated consequences from the combined model is once again estimated using Monte Carlo simulation. With full overlap between the models, there is just one compound process combining the two models. Thus, we generate N_R from the posterior predictive distribution with the parameters α_R' and β_R'. Then we generate N_R consequences from F_R' and add all the results.

3.3 Partial overlap between the models

In the third case we assume, perhaps more realistically, that some of the observed incidents are actually events of the kinds included in the risk factor model while the others are not. Thus, we partition the ν observations into two disjoint sets such that ν_R is the size of the set of events included in the risk factor model, while $\nu_I = \nu - \nu_R$ is the the size of the set of events not included in this model. The partitioning can be done in many different ways depending on how easy it is to identify events as being included among the risk factors. Ideally, this can be done without any uncertainty, in which case we easily obtain two clearly separated groups. In more realistic cases, however, it may be difficult to distinguish between the different types of events. Still it may be possible to assess the numbers ν_R and ν_I, and then simply create a random partition such that the two sets get their desired sizes.

Assuming that we have obtained the desired partition one way or the other, the remaining part of the procedure is a combination of the first two cases. The incident part of the model now consists of data from the ν_I events. As before N_I is sampled from its posterior predictive distribution. However, in this case the parameters of this distribution become $\alpha_I' = \alpha_I + \nu_I$ and $\beta_I' = \beta_I + \tau$. Moreover, the parameters of consequence distribution is estimated based on the consequences of the ν_I events. We denote these parameters by ξ_{II} and σ_{II}.

The risk factor part of the model is constructed similarly to the full overlap case. However, in this case only ν_R observations are used in the updating. Thus, the parameters of the Gamma distribution for the rate become $\alpha_R' = \alpha_R + \nu_R$ and $\beta_R' = \beta_R + \tau$. Moreover, the parameters of the consequence distribution are:

$$\xi_R' = c\xi_{IR} + (1-c)\xi_R, \tag{10}$$

$$\sigma_R' = c\sigma_{IR} + (1-c)\sigma_R, \tag{11}$$

where ξ_{IR} and σ_{IR} are estimated based on the consequences of the ν_R events, and the weight factor is given by:

$$c = \frac{\nu_R}{\alpha_R + \nu_R}.$$

The distribution of the resulting accumulated consequences from the combined model is estimated using Monte Carlo simulation. Here we generate N_I and N_R from their respective posterior predictive distributions, and then generate and add consequences from their respective consequence distributions.

Note that the partial overlap model has the other two models as limiting cases. If we let $\nu_R = \nu$, we get the full overlap model, while if we let $\nu_R = 0$, we get the no overlap model.

4 A NUMERICAL EXAMPLE

In this section we illustrate the proposed methods by considering a numerical example[1]. In this example the incident database consists of data from a period of $\tau = 5$ years. During this period $\nu = 460$ incidents are recorded. The prior uncertainty about the rate λ_I is modelled as a Gamma distribution with parameters α_I and β_I. To avoid that the prior affects the results significantly, we use a relatively *vague* prior. That is, we let α_I and β_I be small numbers. More specifically, we let $\alpha_I = \beta_I = 0.01$. The posterior distribution for λ_I then becomes a Gamma distribution with parameters $\alpha_I' = \alpha_I + \nu = 460.01$ and $\beta_I' = \beta_I + \tau = 5.01$. Thus, the posterior mean of λ_I is:

$$\mathrm{E}[\lambda_I|\nu,\tau] = \frac{\alpha_I'}{\beta_I'} = \frac{460.01}{5.01} = 91.818.$$

Using the consequences from the 460 incidents, we estimate the mean and standard deviation of the consequence distribution, and get that $\xi_I = 2.176$ while $\sigma_I = 8.614$.

In the risk factor model a set of 30 risk factors are included. Their respective parameters are shown in Table 1. In the table the ratio α_s/β_s, i.e., the prior mean value of λ_s, is given for each of the risk factors. In order to obtain the value of α_s, this ratio has to be multiplied with β_s. As in the previous section we assume that $\beta_s = \beta_R$, $s = 1, \ldots, 30$. In the base case we let $\beta_R = 0.2$. However, in the analysis we will vary

[1] All the incident data and risk factors used in the example are fictive. Moreover, for simplicity we have skipped the monetary unit throughout the paper.

Table 1: Risk factor model parameters

s	α_s/β_s	ξ_s	σ_s	s	α_s/β_s	ξ_s	σ_s
1	0.1	136	34	16	0.2	72	18
2	0.1	124	31	17	0.1	72	18
3	0.1	120	30	18	0.2	72	18
4	0.2	112	28	19	0.2	60	15
5	0.2	108	27	20	0.1	60	15
6	0.1	96	24	21	0.1	56	14
7	0.1	96	24	22	0.2	52	13
8	0.1	96	24	23	0.1	52	13
9	0.2	80	20	24	0.3	48	12
10	0.3	80	20	25	0.3	48	12
11	0.1	80	20	26	0.3	44	11
12	0.1	80	20	27	0.3	44	11
13	0.2	76	19	28	0.3	40	10
14	0.2	76	19	29	0.3	40	10
15	0.1	76	19	30	0.2	36	9

this in order to study how this parameter affects the results.

Using (5), (6) and (7) we then obtain $\alpha_R = 1.080$, $\xi_R = 37.926$ and $\sigma_R = 49.576$. The prior mean of λ_R is:

$$\mathrm{E}[\lambda_R] = \frac{\alpha_R}{\beta_R} = \frac{1.080}{0.2} = 5.40.$$

We start out by considering the *no overlap* model. In this case the parameters of the posterior distribution for λ_R are $\alpha'_R = \alpha_R = 1.080$ and $\beta'_R = \beta_R + \tau = 0.2 + 5 = 5.2$. Hence, the posterior mean of λ_R is:

$$\mathrm{E}[\lambda_R|\nu, \tau] = \frac{\alpha'_R}{\beta'_R} = \frac{1.080}{5.2} = 0.208.$$

We observe that the mean value of λ_R is reduced considerably (from 5.40 to 0.208) as a result of observing the process in $\tau = 5$ years without any of the risk factor events occurring.

We then turn to the *full overlap* model. In this case the parameters of the posterior distribution for λ_R are $\alpha'_R = \alpha_R + \nu = 1.080 + 460 = 461.080$ and $\beta'_R = \beta_R + \tau = 0.2 + 5 = 5.2$. Hence, the posterior mean of λ_R is:

$$\mathrm{E}[\lambda_R|\nu, \tau] = \frac{\alpha'_R}{\beta'_R} = \frac{461.080}{5.2} = 88.669.$$

Thus, in this case the mean value of λ_R is increased considerably (from 5.40 to 88.669) as a result of observing the process in $\tau = 5$ years with as many as 460 of the risk factor events occurring.

With a full overlap, however, we also need to update the consequence distribution using the 460 observed consequences. The weight factor, c, is:

$$c = \frac{\nu}{\alpha_R + \nu} = \frac{460}{1.080 + 460} = 0.998.$$

We observe that the weight factor c is relatively high. As a result when updating the mean value and the standard deviation of the consequence distribution, a lot of weight is put on the observed data and less weight is put on the assessments from the expert panel. Thus, using (8) and (9), we get:

$$\xi'_R = c\xi_I + (1 - c)\xi_R = 2.260,$$

$$\sigma'_R = c\sigma_I + (1 - c)\sigma_R = 8.710.$$

Thus, both the mean value and the standard deviation is reduced considerably as a result of the updating.

We recall that main purpose of including the risk factor model is to obtain a better estimate of the tail of the accumulated consequence distribution. If we adopt the full overlap model, this effect is reduced considerably since the risk factor consequnce distribution is strongly drawn towards the incident database consequnce distribution. Typically, the incident databse will include a lot of minor events with small consequences. Such events are not likely to contain relevant information about the risk factors. Thus, this model tends to put far to much weight on the consequences of the events in the incident database.

On the other hand, if the no overlap model is adopted, the mean value of the occurence rate λ_R is reduced since according to this model no risk factor events occurred during the observation period. This is likely not a realistic scenario either.

The partial overlap model, however, tries to balance these issues by using some but not all the incident data to update the occurrence rate uncertainty and the consequence distribution. In order to choose the set of incident data likely to be relevant to the risk factor distribution, we observe that the consequences associated with the risk factors typically are high compared to the consequences in the incident database. Thus, when selecting incident data relevant to the risk factors, we use a random criterion where the probability of being selected increases by the size of the consequence. More specifically, denoting the probability that the ith incident is selected by p_i, we let:

$$p_i = 1 - e^{-\rho X_i}, \quad i = 1, \ldots, \nu,$$

where ρ is a suitable constant. We observe that p_i is an increasing function of ρ. Hence, a high value of ρ implies that a lot of the incidents will be selected, while a small value of ρ implies that only a few incidents will be selected. Note, however, that the choice of ρ depends strongly on the scale of the consequences. Converting the monetary consequences e.g., between different currencies implies that ρ must be changed accordingly. In order to choose a value for ρ it may be easier to start out by assessing the number ν_R, i.e., the number of incidents used to update the risk factor model, and then adjust ρ so that the resulting number of selected incidents is close to ν. In this example it was assessed that about a third of the incidents should be selected. In order to accomplish this we used $\rho = 0.5$.

As a result of the random selection, we obtained that $\nu_I = 292$ and $\nu_R = 168$. The updated parameters for the rate distributions then become

$$\alpha_I' = \alpha_I + \nu_I = 0.01 + 292 = 292.01,$$

$$\beta_I' = \beta_I + \tau = 0.01 + 5 = 5.01,$$

$$\alpha_R' = \alpha_R + \nu_R = 1.080 + 168 = 169.080,$$

$$\beta_R' = \beta_R + \tau = 0.20 + 5 = 5.20.$$

The parameters of the consequence distributions are estimated for the two sets of incidents, and we get $\xi_{II} = 0.499$, $\sigma_{II} = 0.603$, $\xi_{IR} = 5.092$ and $\sigma_{IR} = 13.778$.

Finally, we calculate the weight factor c:

$$c = \frac{\nu_R}{\alpha_R + \nu_R} = \frac{168}{1.080 + 168} = 0.994.$$

Then using (10) and (11) we get the updated parameters for the risk factor consequence distribution:

$$\xi_R' = c\xi_{IR} + (1-c)\xi_R = 5.302,$$

$$\sigma_R' = c\sigma_{IR} + (1-c)\sigma_R = 14.007.$$

Before we run a Monte Carlo simulation using all these parameters, it is of interest to calculate the mean and the standard deviations of the total accumulated consequences using the three models. This is easily done analytically using (2) and (3).

Table 2: Mean and standard deviations for the three models for the case when $\beta_R = 0.2$

	No overlap	Full overlap	Partial overlap
Mean	207.674	200.369	201.472
St.dev.	89.761	84.732	85.608

We observe that the *no overlap model* has the largest mean and standard deviation, while the *full overlap model* has the smallest mean and standard deviation. Not surprisingly, the *partial overlap model* places itself between the two extremes. The same tendency can be seen when running a simulation on the three models. The results of this simulation, consisting of 100000 iterations, is shown in Figure 1.

The estimated cumulative distribution curve for the *no overlap* model lies slightly more to the right compared to the other models. However, the difference is almost neglectable. Thus, for the parameters used in this case, the potential overlap between the incident database and the risk factors does not have a significant effect on the results. One of the reasons for this is that relatively small weights are put on the prior for λ_R. That is, the parameter β_R, which we interpret as a measure of the strength of the prior knowledge about λ_R, is modest. This assessment reflects that experts often find it difficult to assess the rate parameter in a compound Poisson process. In order to study in more detail how β_R affects the results, we repeat the calculations for $\beta_R = 0.5$ and $\beta_R = 1.0$ as well. The resulting mean and standard deviations for the three models are shown in Table 3 and Table 4 respectively, while the estimated cumulative distributions for the models are shown in Figure 2 and Figure 3 respectively. We observe that the results for the full overlap model are almost the same for all three values of β_R, while both the mean and the standard deviation for the no overlap model increases as β_R increases. The results for the partial overlap model are in between the two extremes, but closer to the full overlap model.

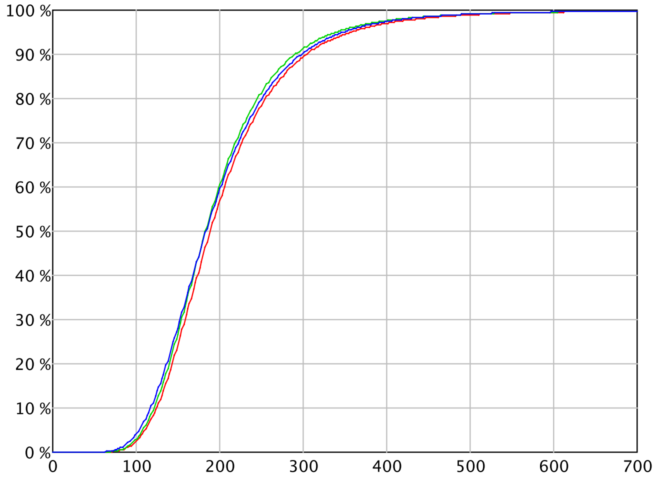

Figure 1: Estimated cumulative distributions for the accumulated consequences using the *no overlap model* (red curve), *full overlap model* (green curve) and the *partial overlap model* (blue curve), $\beta_R = 0.2$.

As already indicated, a high β_R-value implies that more weight is put on the prior estimate of the rate of the risk factor model. For the no overlap model one effect of a change of β_R from 0.2 to 1.0 is that the posterior mean rate $\mathrm{E}[\lambda_R|\nu, \tau]$ increases from 0.208 to 0.900. On the other hand, for the full overlap model a change in β_R from 0.2 to 1.0 implies that the posterior mean rate $\mathrm{E}[\lambda_R|\nu, \tau]$ decreases from 88.67 to 77.57. At the same time we recall that the α_s-parameters are not assessed directly. Thus, changes in the value of β_R also affects the values of the α_s-parameters, and hence also the α_R-parameter. More specifically, the increase in β_R also results in an increase in α_R from 1.080 to 5.400. Hence, the weight factor c is reduced from 0.998 to 0.988. From this it follows that ξ_R' and σ_R' grows. The total effect of this is that the change in the mean and standard deviation of the total accumulated consequences is virtually neglectable.

This analysis shows the effect of putting more weight on the prior estimate of the rate of the risk factor model. As β_R increases, the model becomes more sensitive to changes in the overlap conditions. Thus, it becomes more important to have more precise knowledge about the degree of overlap. In a real-life application, however, the value of β_R is more likely to be small in which case the model is robust with respect to variations in the overlap conditions. Moreover, even for slightly higher β_R-values, the results of the three models are not that different.

Table 3: Mean and standard deviations for the three models for the case when $\beta_R = 0.5$

	No overlap	Full overlap	Partial overlap
Mean	218.415	200.611	203.240
St.dev.	95.710	84.095	86.017

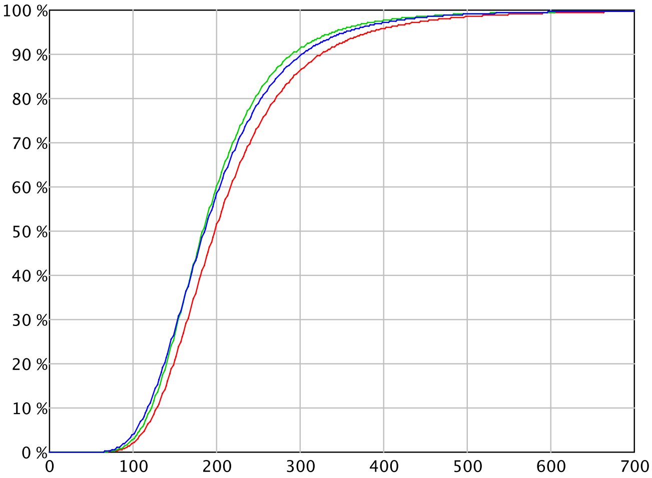

Figure 2: Estimated cumulative distributions for the accumulated consequences using the *no overlap model* (red curve), *full overlap model* (green curve) and the *partial overlap model* (blue curve), $\beta_R = 0.5$.

Table 4: Mean and standard deviations for the three models for the case when $\beta_R = 1.0$

	No overlap	Full overlap	Partial overlap
Mean	233.930	200.960	205.794
St.dev.	103.703	83.240	86.753

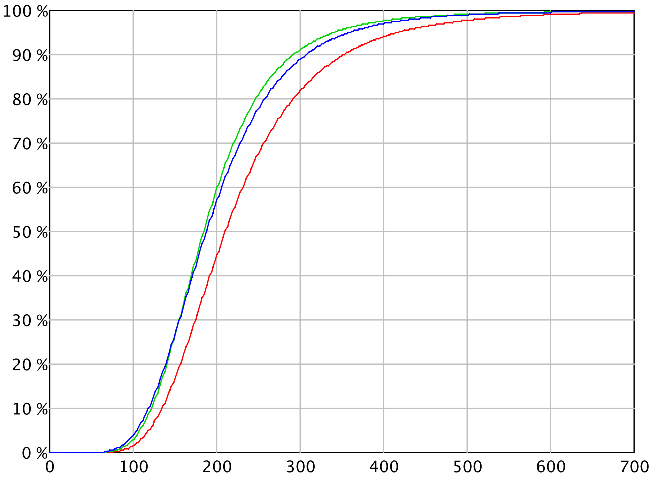

Figure 3: Estimated cumulative distributions for the accumulated consequences using the *no overlap model* (red curve), *full overlap model* (green curve) and the *partial overlap model* (blue curve), $\beta_R = 1.0$.

5 CONCLUSIONS

In the present paper we have studied how one can combine incident data with subjective assements. A very important issue in relation to this is the degree of overlap between these two sources of information. The models we have proposed cover the full range from no overlap to full overlap. In the case of no overlap the consequence data for the incidents is irrelevant with respect to the consequence distribution for the risk factors. Still having observed the processes a number of years without recording any events related to the risk factors, is vital information which need to be taken into account. In the case of full overlap the incident data is used both to update the rate distribution and the consequence distribution.

The main focus of this paper has been on the rates of occurrence of the various risk factors, and how this can be updated in a consistent way. Far less emphasis has been put on how to fit and update the consequence distributions. We have used a very simple model where lognormal distributions are used both for incidents and for the sum of the risk factors. Moreover, for the combined model we also use a lognormal distribution where the parameters are determined by simple weighted averages. In a more refined model, one may need to consider a wider variety of distribution classes, including non-parametric distributions. For the risk factors a more flexible approach could include the fitting of specific distributions for each individual risk factor. Finally, in order to combine all distributions, a full scale Bayesian updating approach should be developed. All these issues will be addressed in an upcoming paper.

Still, despite these shortcomings, the proposed model, in all its simplicity, contains the most important features, and thus enables the analyst to obtain useful results.

REFERENCES

Adachi, M. et al. *Operational Risk - Supervisory Guidelines for the Advanced Measurement Approaches* Basel Committee on Banking Supervision, 2011.

Adelson, R. M. Compound Poisson distributions. *Operational Research Quarterly* (17): 73–75, 1966.

Berger, J. O. *Statistical Decision Theory and Bayesian Analysis.* Springer Series in Statistics, 2010.

Bølviken, E. *Computation and Modelling in Insurance and Finance* Cambridge University Press, 2014.

Huseby, A. B. Combining Opinions in a Predictive Case In *Bayesian Statistics 3.*: 641–651, Oxford University Press, 1989.

McNeil, A. J., Frey, R., Embrechts, P. *Quantitive Risk Management, Concepts, Techniques and Tools* Princeton University Press, 2005.