Societal Risk Assessment

Tim Pattenden, Tessella Plc

Abstract

One measure of risk in the presence of potentially dangerous events is societal risk. Societal risk can be expressed graphically as the frequency of N or more fatalities occurring, often referred to as the 'FN curve'. The FN curve is a key input for decisions taken about the tolerability of risky activities. This paper describes how an accurate FN curve can be derived in the context of explosives storage for defence purposes, where there are a large number of potential events and a large number of populations at different locations, and the characteristics of both events and populations vary with time.

An accurate FN curve is achieved by breaking down the time period for which a risk assessment is required into a set of periods within which the relevant parameters – both the population levels and the frequency of dangerous events – do not change. Within each of these time periods the consequences of an event on each population are calculated (which requires empirical knowledge or modelling specific to the explosives domain). The overall consequences on all populations can then be calculated by convolving the probability distributions of fatality numbers for each population. Finally, the results for all events and all time periods must be added to give an overall frequency distribution of fatality numbers.

A complication with explosive events is that one event may trigger another. Accurate calculations of the FN curve must therefore account for both the single and multiple explosion events, whilst still ensuring that the correct population characteristics are selected for the time periods in which multiple explosions can occur.

The application of a number of mathematical techniques allows the general problem to be solved in cases where the number of events and population sizes previously required considerable approximations to be made. Whilst these techniques are standard in mathematical work they are newly applied in this context. This paper illustrates how the use of more sophisticated mathematics improves the fidelity of the FN curve, in some situations significantly altering the shape of this crucial indicator of risk.

1. Introduction

One measure of risk in the presence of potentially dangerous events is societal risk. Societal risk can be expressed graphically as the frequency of N or more fatalities occurring, often referred to as the 'FN curve'. The FN curve is a key input for decisions taken about the tolerability of risky activities. This paper describes how an accurate FN curve can be derived in the context of explosives storage for defence purposes. In this context there are a large number of potential accidental explosive events and a large number of population groups present at different locations. The characteristics of both explosive events and population groups vary with time.

The problem of calculating the overall frequency distribution for fatality numbers can be summarised as follows:

- Identify a set of unique (i.e. non-overlapping) time periods for which the parameters of the risk calculation remain constant within each individual period.
- For each time period identify the potential explosive events.
- For each potential explosive event calculate its impact on all population groups.
- Combine these results.

2. Identifying Time Periods

To calculate the probable impact of explosive events on a population relevant parameters include the quantity of explosive stored, the type of processing being carried out on those explosives, and the number and location of the population. All of these parameters may vary over time. For example, residential population may be largely absent during the working day whereas for on-site workers the opposite will be the case.

To define unique non-overlapping time periods within which all parameters remain constant means finding, for each new time period we consider, its intersections and relative complements with all the time periods previously defined. To take a very simple example, suppose that time-varying data is defined in terms of days of the week only and we have the following information:

- All population is present on weekdays and absent at the weekend.
- An explosive storehouse is used for processing explosives on Mondays, and is otherwise used for storage only.

We build up our set of time periods as follows:

$$\begin{aligned} T &= \varnothing \\ T &= \{\text{weekdays}, \text{weekends}\} \\ T &= \{\text{Mondays}, \text{other weekdays}, \text{weekends}\} \end{aligned} \tag{1}$$

In practice this data is defined in terms of days of the week, hours of the day and days of the year, so the above process must operate in three dimensions. Nevertheless the problem is computationally tractable so long as reasonable approximations are made (e.g. assume that all office workers leave at 5:30 rather than assuming that group one leaves at 5:00, group two at 5:10, group three at 5:20 and so on).

3. Fatality Probability Distributions

For each time period in the set T we must find the impact on all population groups for all possible explosive events. For a particular explosive event and a particular exposed population group the modelling methods used give a single number ρ for the probability of fatality for each member of the population. It is assumed that the probabilities for each member of the population are independent variables, so the overall number of fatalities follows a binomial distribution

$$P(n) = \frac{N!}{n!(N-n)!} \rho^n (1-\rho)^{N-n} \tag{2}$$

Where *n* is the number of fatalities out of a total exposed number *N*.

The total number of fatalities is the sum of the fatalities for each population group. Using $G = \{1,2,3,\ldots\}$ as a suffix for population group we have

$$n_{event} = \sum_G n_G \tag{3}$$

Thus the probability distribution of total fatality numbers for one explosive event is given by the convolution of the probability distributions for individual population groups.

$$P(n_{event}) = P(n_1) * P(n_2) * P(n_3) * \ldots \tag{4}$$

And the frequency distribution for these fatality numbers is simply the probability distribution multiplied by the frequency of the event. The fatality frequency distribution for all possible explosive events in a time period is the sum of the frequency distributions for individual events.

$$f\left(n_{period}\right) = \sum_{events} f_{event} P\left(n_{event}\right) \tag{5}$$

Finally, we add together results from all the time periods:

$$f\left(n_{total}\right) = \sum_{\text{periods in T}} \alpha_{period}\, f\left(n_{period}\right) \tag{6}$$

Where α_{period} is the fraction of a year taken up by the time period (all time periods are periodic over at most one year).

4. Multiple Explosions

A complication with explosive events is that one event may trigger another. If this happens sufficiently quickly that the inputs to the risk calculation remain substantially unchanged (i.e. people do not have time to respond to the first explosion by running away) then it is considered as a single multiple explosion event.

The fatality probability for an individual from multiple simultaneous explosions is simply calculated as

$$\rho = 1 - \prod_{i} \left(1 - \rho_i\right) \tag{7}$$

And the impact of the event can then be calculated as in section 3.

If one event may trigger another, the situation of both the initiating and responding explosive storehouses are relevant parameters in the calculation, and time periods must be found in which both of these parameter sets are constant. This can add considerably to the size of the time period set discussed in section 2 and so multiple explosions are only considered when the probability of their occurrence rises above a threshold.

5. Solving Performance Issues

For a large site with many possible explosive events, including multiple explosion events, and a large number of population groups, the set of time periods T becomes very large. Since many calculations must be performed in order to find the FN curve for each time period the speed of these calculations becomes crucial.

The convolution operation used in equation 4 is defined as

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m] g[n-m] \tag{8}$$

Where the size of a population group is large this operation this can be very slow if implemented as an explicit evaluation of the sum for each value of n . Instead the convolution theorem is used:

$$F\{f * g\} = F\{f\} \cdot F\{g\} \tag{9}$$

Where $F\{\ \}$ indicates taking the Fourier transform and the dot indicates point-wise multiplication. This reduces the complexity from $O\left(N^2\right)$ to $O\left(N \log N\right)$.

In many cases a further approximation can be made. The binomial distribution from equation 2 is well approximated by a normal distribution when n is large and ρ is not too near to zero or one. The sum of two independent normally distributed variables is itself normally distributed. The convolution can therefore be achieved by simply adding the mean and variance of the two distributions. In practice, ρ is rarely close to one and if it is close to zero the probability of large numbers of fatalities is low, meaning that the convolution is in any case fast.

Together with suitable housekeeping routines in the software to truncate distributions that have very long tails and to neglect results whose impact on overall risk is below certain thresholds, these methods allow an accurate FN calculation to be performed for sites with large populations and a large number of potential explosive events.

6. Example Results

Figure 1 below shows an FN graph produced using these methods for an illustrative site (all data is purely fictional). On this site there are:

- A population of 20 people working on the site, present during the day.
- A nearby population of 100 people, located off-site, and present largely at night time.
- One explosive storehouse, used for processing explosives during the day but for storage only at night.

For both populations, the fatality probability for each individual in the event of an explosion, ρ , is 0.1049.

The curved lines on the graph show the frequency of N or more fatalities plotted against number of fatalities for three groups: the on-site population (green), the off-site population (blue) and the combined population (yellow – the line is largely hidden underneath the other two lines). The straight red and purple lines represent criteria used in the evaluation of risk. These criteria are discussed by the Health and Safety Executive, though it should be noted that the results presented for the fictional data do not necessarily correspond to the most recent criteria issued by the HSE.

It can be seen that there is a higher frequency of fatalities amongst the on-site population, since they are present during the more dangerous time when explosives are being processed. There is the potential for higher numbers of fatalities amongst the off-site population, but this is less frequent, as an accidental explosion when a storehouse is being used for storage only is much less likely.

Figure 2 shows an FN graph generated for the same site when an approximation has been made. Rather than specifying the times at which population is present, each population is simply given a fraction of time, from zero to one, during which it is exposed to hazards. Although the fraction of time for the off-site night-time population is correct, the calculation is now unable to account for the low intersection between this population's presence and the period of most danger, and so the calculated frequencies for fatalities amongst this population are too high. Prior to the application of the techniques discussed above such approximations were necessary in order to make the problem computationally tractable and it can be seen that in some circumstances they could significantly affect this measure of risk.

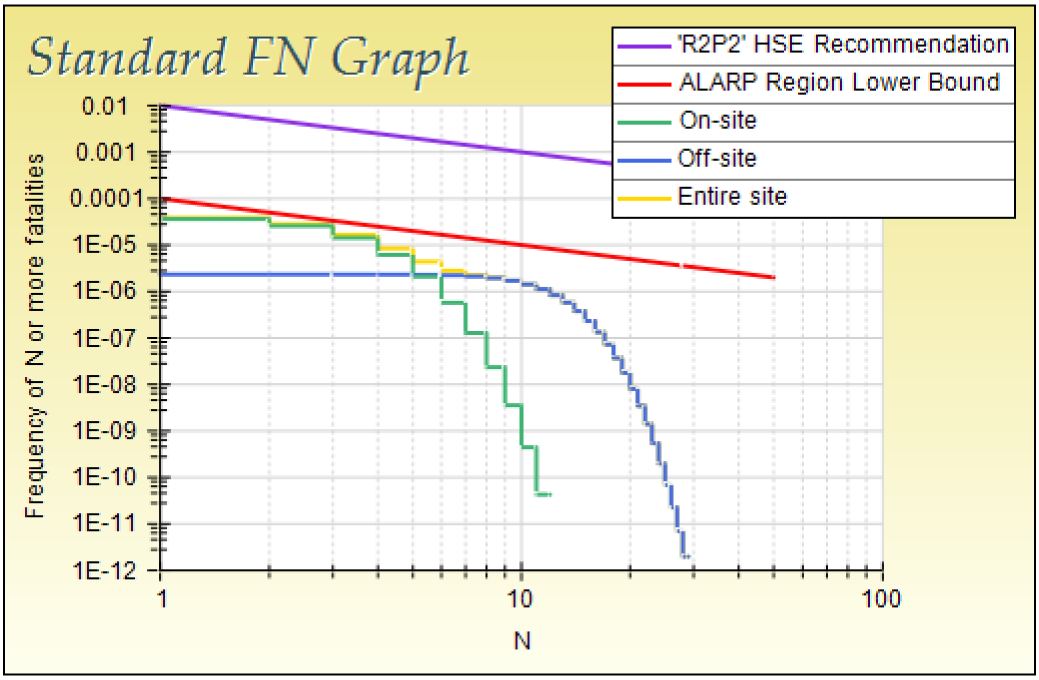

Figure 1 : FN Curve for an Example Site

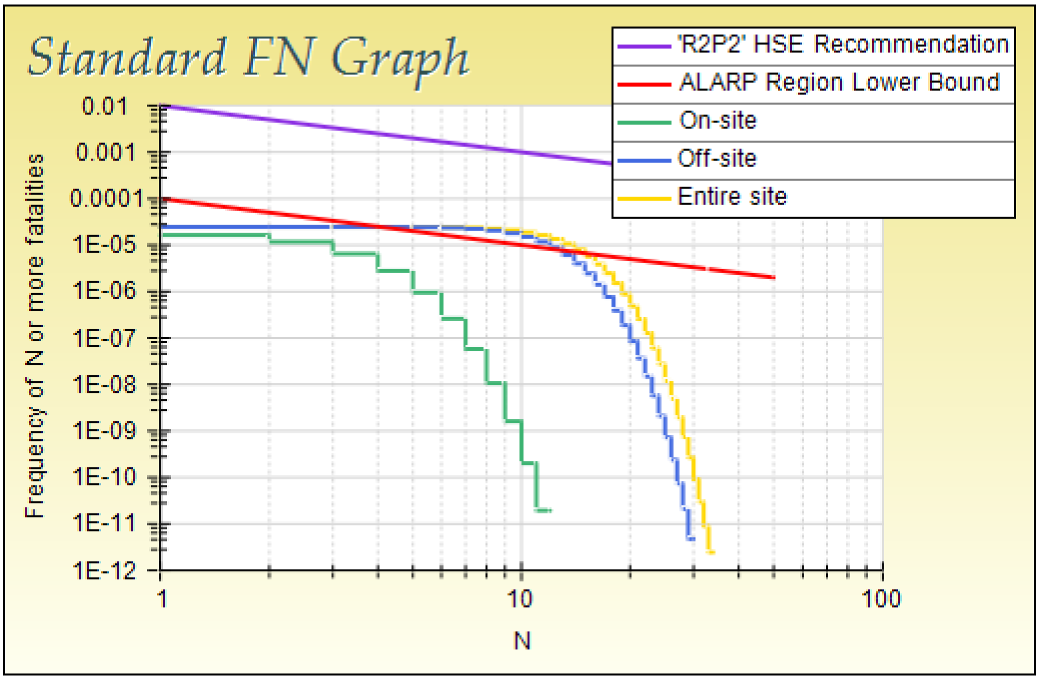

Figure 2 : FN Curve with Approximate Exposure Calculations

7. Conclusions

Calculating an accurate FN curve in the context of explosives storage for defence purposes is a computationally intensive problem. The presence of many different population groups and the time-varying exposure of these groups to different hazards mean that very many calculations must be performed in order to derive an accurate measure of risk.

Fortunately it is possible to greatly speed up the calculations by the application of some appropriate mathematical techniques. These techniques are entirely standard but are newly applied in this context. The performance benefits obtained illustrate how using the correct mathematical approach can revolutionise the performance of a software program.

The new ability to calculate FN curves without having to approximate the exposure of population groups to hazards as a simple time fraction can in some cases significantly alter the shape of this crucial indicator of risk.

REFERENCES

HEALTH & SAFETY EXECUTIVE, 2001, “Reducing Risks, Protecting People: HSE’s decision-making process”