Uncertainty in Explosion Risk Assessment

Chris Coffey, Principal Engineer, Offshore Consulting Group, London

Richard Gibson, Principal Engineer, Offshore Consulting Group, London

Marios Christou, Principal Engineer, Offshore Consulting Group, London

Probabilistic approaches are often used when considering explosion risk of offshore and onshore facilities. However, there are a number of limitations with such explosion risk assessment methodologies. These are primarily due to the uncertainties that are inherent with the technique. Statistical data is used to provide input (*e.g.* wind probability data, leak frequency data, ignition probability data). Each of these statistical distributions have their own associated uncertainties. Additional uncertainties arise in the modelling techniques used to determine the consequences of a release and subsequent explosion. Finally, there is a large element of engineering judgement to select the most appropriate scenarios to consider. This brings further uncertainties to the problem.

How these uncertainties propagate through the analysis are poorly understood and communicated. Good studies perform some level of sensitivity to try and highlight the uncertainty to the client; however, this is typically not done well or at all. Single values are often presented without any associated error bars or confidence levels, which then get fed directly into design. This presents a significant point of concern as it could lead to facilities being under designed. Here, we present a methodology that shows how uncertainties (both in the inputs and in the models used) propagate through the explosion risk analysis. These enable the results of the analysis to be presented with confidence intervals to allow informed design decisions to be made. Alternatively, the uncertainty can be included in the final estimates of design pressures.

By understanding and highlighting the uncertainties with the approach, we can better understand how improvements can be made. Efforts can be focussed on improving the data for inputs that will have the most significant impact on the accuracy of the results. In addition, we can show where there is a risk the analysis may not produce conservative results and potentially lead to dangerous design.

Introduction

Offshore installations are threatened by various hazards which if not managed may lead to consequences such as injury to or death of operators, damage to assets, environmental impact, and business disruption. The main hazards include wind and waves, fires, and explosions. The topside of an offshore platform is exposed to the risk of hydrocarbon fires and explosions that unmanaged can lead to disastrous consequences.

Loss of containment events that can lead to explosions vary significantly in their magnitude and likelihood, with larger releases generally being much less frequent than the smaller ones. The consequences of any release also varies significantly. Risk-based approaches to explosion management are therefore a key part of the design of any new installation. NORSOK Z-013 [1], for example, provides a methodology for such an approach. The application of risk-based methods is also important in supporting decision making in operations, particularly where there is a change from the original design assumptions. This can occur, for example, if a new field development is tied back to an existing facility.

However, there are a number of limitations with the explosion risk assessment methodologies. These are primarily due to the uncertainties that are inherent with the technique. Statistical data is used to provide input (*e.g.* wind probability data, leak frequency data, ignition probability data). Each of these statistical distributions have their own associated uncertainties. Additional uncertainties arise in the modelling techniques used to determine the consequences of a release and subsequent explosion. Finally, there is a large element of engineering judgement to select the most appropriate scenarios to consider. This brings further uncertainties to the problem.

Currently the understanding and communication of uncertainties associated with probabilistic explosion analysis is poor. Good studies perform some level of sensitivity to try and highlight the uncertainty to the client; however, this is typically not done well or at all. Single values are often presented without any associated error bars, which then get fed directly into design. This presents a significant point of concern as it could lead to facilities being under designed and has been raised as a concern by regulators. For example, the UK HSE reports that:

> "The methodologies used for the development of exceedance across the industry curves are not consistent. The treatment of uncertainties is not clear and the wide range of methodologies employed, ranging from use of generic curves, combinations of phenomenological and CFD modelling and solely CFD modelling, prevents comparison between approaches used."[2]

Here, we show how these uncertainties may be determined and how they impact the overall analysis. For the purposes of this paper, we focus on the how the uncertainties associated with the input data propagate though the analysis. We use a simplified "toy" model of dispersion and explosion loadings. This allows the importance of the uncertainty in any one element to be quantified and provide a framework by which a probabilistic model could be developed for real problems.

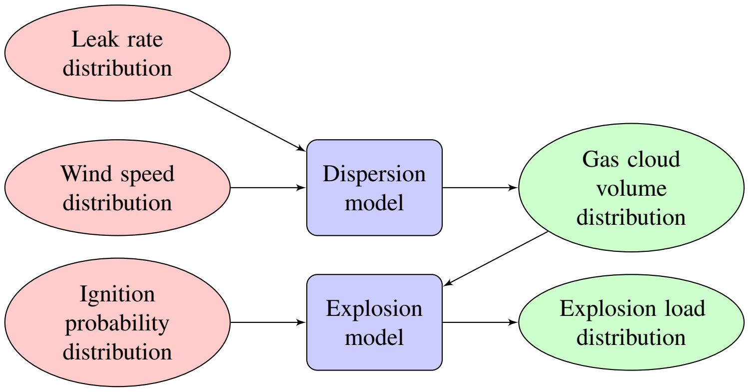


Figure 1: Simple schematic of probabilistic explosion risk assessment methodology.

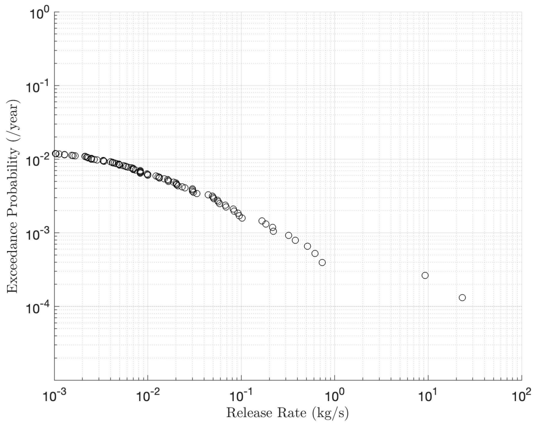


Figure 2: Empirical exceedance probability of leak by release rate for compressors.

Input distributions

Figure 1 shows a very simplified schematic of the probabilistic explosion risk assessment methodology. In its most basic form, the approach takes distribution of leak probability, wind probability and ignition probability and transforms them via deterministic models, to simulate dispersion and explosion, into the distribution of pressure on a given target of interest. For this paper we focus on the three inputs into the explosion risk assessment and the uncertainty associated with them.

Leak rate

The Hydrocarbon Release Database (HCRD) [3] has been compiled by the UK HSE over a 20-year period and contains details of leak and ignition events at oil and gas installations in the UK Continental Shelf. The HCRD has become a standard source of leak frequencies for offshore risk assessments and as such we will use it as the basis of our analysis here, though a similar approach can be taken with any source data.

For the purposes of this paper we consider a compressor leak; though it is relatively trivial to extend the analysis to all the reported equipment types (and in reality we would normally consider a combination of equipment types). Figure 2 shows the empirical exceedance probability based on the data from the HCRD. For the more significant events (release rates greater than 1 kg/s), the data are particularly sparse. As such, for this part of the parameter space we expect the uncertainty to be high.

To provide a better understanding of the distribution and the associated uncertainty, we fit the data above a given hole size (*i.e.* the tail of the distribution) using a Generalised Pareto (GP) distribution [4]. This is a technique used to model the extreme values in the tail of a distribution (where the data may be sparse) that is justified using asymptotic arguments. Below that threshold release, we assume there is sufficient data to use the empirical distribution directly. Figure 3 shows an example of the maximum likelihood fits of the distribution to the data with varying threshold values for release rate. Statistical tests (for example, consideration of the mean residual life) can be used to determine the most appropriate range of threshold release rates to consider. For the data shown here, thresholds in the range 0.05 kg/s to 0.1 kg/s are the most sensible. We include the

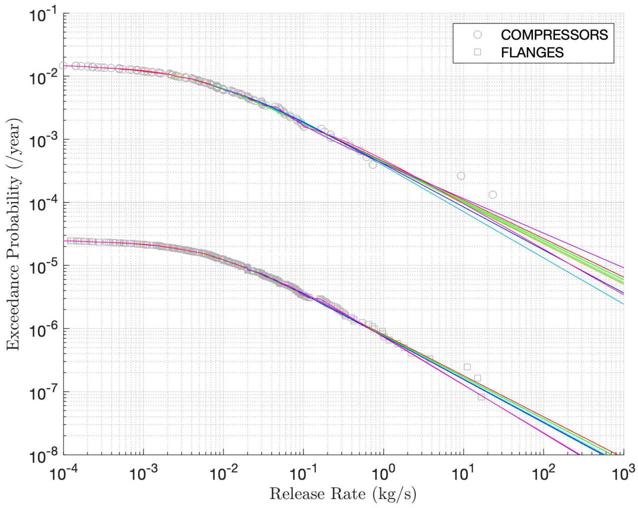
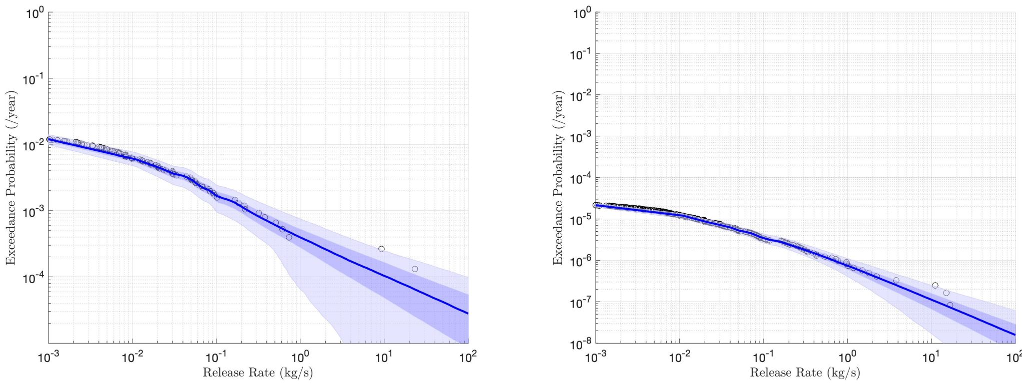


Figure 3: Fitted generalised Pareto distributions above different release rate thresholds.



(a) Compressors

(b) Flanges

Figure 4: Probability (/year) of exceeding a leak rate. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles.

uncertainty in threshold selection by averaging over the plausible range. Other techniques for including this in an analysis are discussed in [5].

To understand the uncertainty in the distribution, we use bootstrapping techniques [6] to resample the source data into 1,000 replicate datasets and fit a GP distribution to each one. Figure 4a shows a fan plot from the resulting distributions. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles. Here, we have extended the distribution out to release rates of 100 kg/s. We can see that for larger release rates the uncertainty in the distribution becomes significant. Above 1 kg/s the range between the 5% and 95% is one order of magnitude.

Figure 4b shows fan plots for leaks from flanges. We see similarly high levels of uncertainty in the leak distribution for flanges as we do for the compressors leaks. However, as the data set for flanges is larger than that for compressors (there are many more flanges than compressors across North Sea installations), the uncertainty in the distribution is comparatively smaller.

Ignition probability

There is a range of models available for the prediction of the ignition probability (see, for example, [7],[8],[9]). These models vary in their level of sophistication and the number of input parameters they consider. Some are purely statistical correlations. Others are based on models of the physics. Some, include time variance. However, they all rely on underlying statistical data to calibrate them.

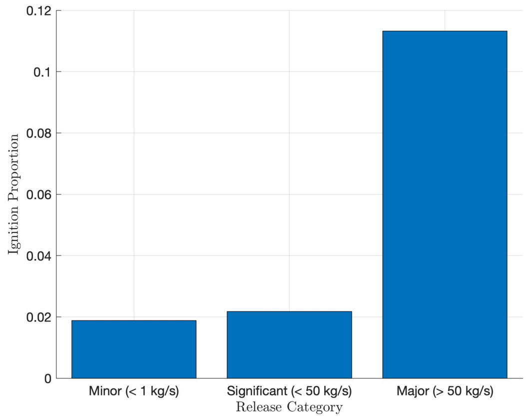


Figure 5: Proportion of release that ignited for each leak category considered.

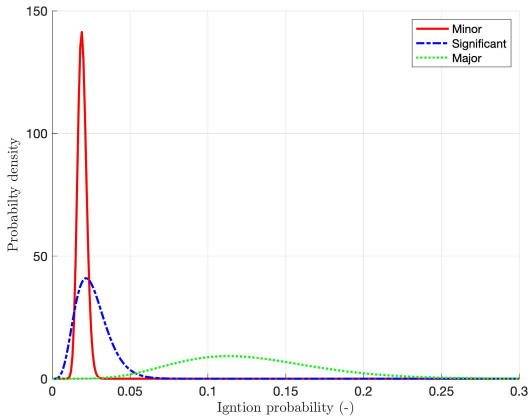


Figure 6: Ignition probability distribution for each leak category.

For the purposes of this paper, we focus on one of the simplest approaches to ignition modelling (as per [7]) where it is assumed ignition probability is a function of leak rate. Again, for simplicity's sake, we take no account of delayed versus immediate ignition and assume all ignition results in an explosion. Note, however, that similar approaches to understanding the uncertainty in the ignition probability could be applied to any of the more sophisticated models. As the models all rely on similar data sets we expect broadly similar results in the size of any uncertainty, but clearly the details will be different..

We divide the leak rates ($\dot{m}$) into three categories: Minor ($\dot{m} < 1$ kg/s), Significant (1 kg/s $\leq \dot{m} \leq 50$ kg/s), and Major ($\dot{m} >$ 50 kg/s). As with the leak probability, we use the HCRD as our source of data. Figure 5 shows the proportion of ignitions for each category (focussing on gas leaks).

Due to the sparseness of the input data, we expect some uncertainty in the value of ignition probability that needs to be quantified if we are to understand its impact on our explosion risk analysis. If we assume no prior knowledge of the ignition probability values ahead of considering the HCRD data (*i.e.* we assume the prior distribution is uniform on the interval [0,1]), then using a Bayesian analysis (see, for example, [10]) one can show that if we have y ignitions in n independent events then the probability θ of a subsequent event igniting is described by

$$p(\theta|y) = \mathrm{Beta}(y+1, n-y+1), \tag{1}$$

where Beta is the *beta* distribution. Figure 6 plots the distribution for each of the leak catogries considered. We see that the spread in ignition probability is large and increases with leak category as the data becomes more sparse.

Wind speed

Unlike leak rate and ignition probabilities, there are good data sets available for wind speed. There are numerous weather stations and numerical reanalyses providing historical wind speeds at regular intervals over large time spans. Here, we shall

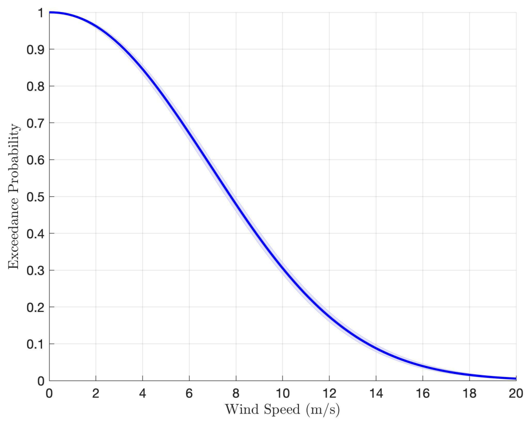


Figure 7: Probability of exceeding a wind speed. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles.

consider such a data set for a typical North Sea installation.

By using bootstrapping techniques [6] we can assess the uncertainty through resampling (with replacement) the underlying data. In this case, we sampled 1,000 data sets. Care needs to be taken in this sampling, as we expect a high degree of correlation between consecutive measurements (*i.e.* they are not independent). Analysis shows that for this data set at a 10 day interval the correlation is minimal, but it can certainly be argued that

1. we are overestimating the uncertainty by sub-sampling the data too coarsely;
2. we are underestimating the uncertainty in the very low wind speed tail.

Further work is required to explore this aspect of the uncertainty in more detail. In particular, work is needed to fully assess the uncertainty under calm conditions where we expect the explosion risk to be highest.

Figure 7 shows an exceedance fan plot of the bootstrapped wind speed data. As expected, partly due to the size of the data set and partly due to the fact that we're not interested in the storm events, the uncertainty is very small.

Consequence models

To predict the pressure load on a target due to an explosion, deterministic models are used. Dispersion models predict the extent of a flammable vapour cloud given a leak. Explosion models predict the resulting pressure loads following the ignition of a vapour cloud. For most modern analyses on offshore structures, Computational Fluid Dynamics (CFD) is used [11]. This allows a large range of parameters to be considered (leak location, ignition location, geometrical effects *etc.*). Here, as we are focussed solely on the uncertainty associated with leak, ignition and wind distributions, we use two simple 'toy' models to provide a prediction of the resulting overpressure on a target of interest.

For our dispersion model, we shall assume

$$Q \propto \left(\frac{\dot{m}}{U_w}\right)^{\frac{3}{2}}. \tag{2}$$

Here, Q is the flammable cloud size, $\dot{m}$ is the leak rate, and U_w the wind speed. This is an assumption underlying models such as [12]. For our explosion model, we shall assume

$$p \propto Q^{1/3}, \tag{3}$$

where p is the pressure on a target of interest. This is an assumption underlying models such as [13]. We have calibrated our constants of proportionality so that they give us predictions in a sensible range. While we don't expect either of our toy models to give good predictions of cloud size or pressure, they serve to help us understand how the input distributions and the uncertainty associated with them pass though the analysis.

Probabilistic analysis

To illustrate the impact of the uncertainties in the input distributions, we have performed a probabilistic explosion risk analysis. We considered a leak and subsequent ignition for a single compressor. A Monte Carlo technique was used to sample a range of

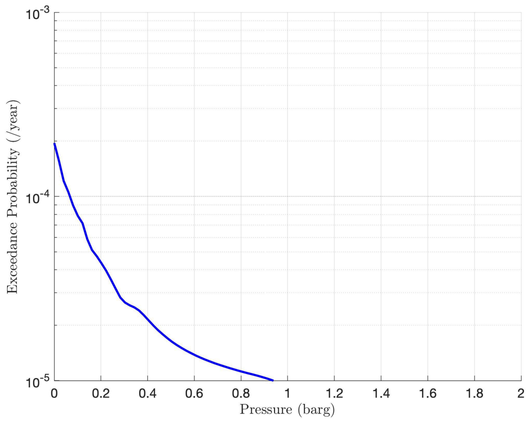


Figure 8: Probability (/year) of exceeding a pressure. Here, no uncertainty in the input distributions is considered. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles. The dashed red line shows the predictive posterior distribution.

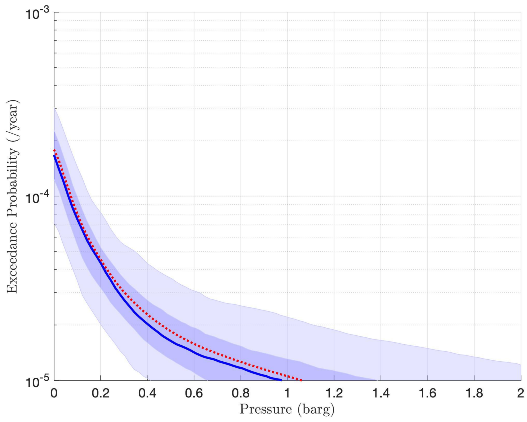


Figure 9: Probability (/year) of exceeding a pressure. Here, uncertainty in the leak, ignition, and wind input distributions is considered. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles. The dashed red line shows the predictive posterior distribution.

potential scenarios, each defined by a leak rate, a wind speed and an ignition probability. In total 10^8 scenarios were considered. For each of these scenarios, we used the simple consequence models described in equations (2) and (3) to determine the pressure loading on some target of interest. We then construct an exceedance curve that shows the probability (/year) of exceeding a given pressure on our target of interest.

Should (as is done by the majority of engineers when performing such an analysis) we neglect the uncertainty in the input distributions, we construct the single exceedance curve presenting in Figure 8. We see that, at typical probabilities of interest (*i.e.* 10^{-4} /year and 10^{-5} /year), we predict pressures on our target of approximately 0.1 barg and 1.1 barg, respectively.

However, as discussed previously, the uncertainty associated with our input distributions is not insignificant and as such should be considered in the analysis. By allowing our Monte Carlo technique to sample from the full set of possible distributions, this uncertainty is passed through the risk analysis. Figure 9 shows a fan plot of the probability (/year) of exceeding a pressure including this uncertainty. The solid blue line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles. The dashed red line shows the 'expected' value of the overpressure across the uncertainties (*i.e.* the predictive posterior distribution). The pressures exceeded at our probabilities of interest (10^{-4} /year and 10^{-5} /year) are depicted as a box plot in Figure 10. Here, the rectangle represents the 25% and 75% quantiles, the red line indicates the median value. The extent of the data are shown by the whiskers (excluding any outliers defined as 3/2 x quartile and shown by red crosses).

We see that the uncertainty is large. For this toy example, the value ranges are 0 barg to 0.3 barg and 0.1 barg to 2.1 barg

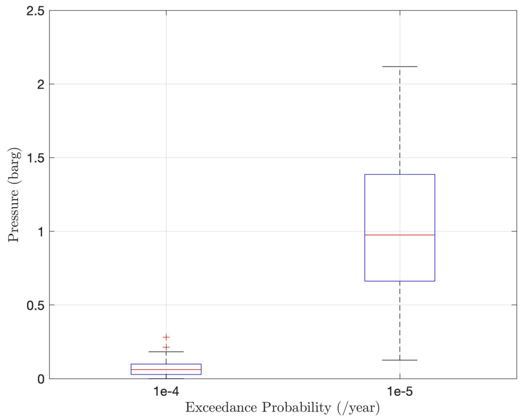


Figure 10: Box plot of pressure exceeding at 10^{-4} /year and 10^{-5} /year . Here, uncertainty in the leak, ignition, and wind input distributions is considered. The rectangle represents the 25% and 75% quantiles, the red line indicates the median value. The extent of the data are shown by the whiskers (excluding any outliers defined as 3/2 x quartile and shown by red crosses).

for 10^{-4} /year and 10^{-5} /year, respectively. Depending on what safety factor is used in any ongoing design and what, if any, uncertainty that safety factor is supposed to represent, neglecting the uncertainty may lead to significant under-conservatism in the results and significant under design.

Discussion

Impact of individual input uncertainties

Through this approach we are able to consider the uncertainties due to each individual input separately. This allows us to rank their importance and so consider where efforts might be best placed to reduce uncertainty in future analyses. For the case considered herein, Figure 11a, 11b, and 11c shows the exceedance fan plots if we only consider uncertainty in leak rate, ignition probability, or wind speed, respectively.

For our case study, the uncertainty in ignition probability is dominating the uncertainty at 10^{-4} /year. Conversely, the uncertainty in leak rate is dominating the uncertainty at 10^{-5} /year. This is because, for a single compressor, the 10^{-4} /year exceedance pressure is driven by relatively low leak rates. Figure 12 shows the cumulative probability that a pressure comes from a scenario defined by a given leak rate. We see that for pressures of 0.1 barg, the scenario leak rates are all below 0.1 barg. In this region of the parameter space, the uncertainty in our leak rate distribution is low (10^{-2} /year see Figure 4a). As we consider smaller exceedance probabilities, the overpressure is driven by larger leak rates and, hence, the uncertainty increases and eventually exceeds the uncertainty of the ignition model.

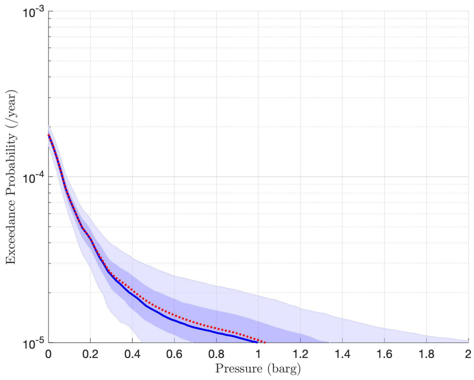
In a more typical study (such as for a full offshore module), we would expect the leak frequency to be due to multiple pieces of equipment. As such the total leak frequency would be expected to be much higher than that presented here (perhaps in the range 0.1 /year to 1 /year). As such, we would expect large releases to be responsible for the risk at the 10^{-4} /year exceedance probability and the associated uncertainty is likely to be higher than presented in this case study.

Other sources of uncertainty

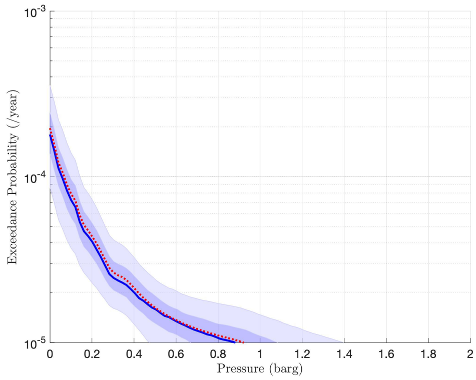
In the presented analysis we have only considered three sources of uncertainty. There are numerous other sources of uncertainty that should also be captured and considered through the whole analysis. Here, we expand on three of them.

1. **Variability between source data sets**
   For each of the input distributions, we have considered the statistical uncertainty in a single data set. This does not, however, provide an estimate of the uncertainty in the data itself.

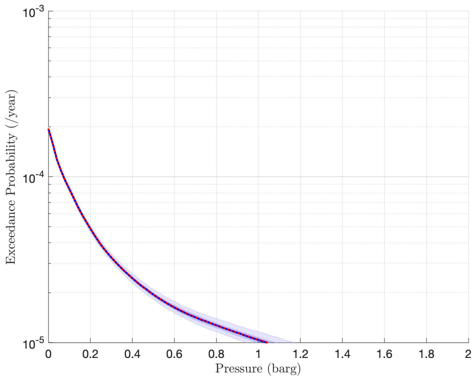
   For example, considering wind speed, we have examined a data set from a single weather station. Should the instruments on that weather station not be reading true (due to, for example, faulty equipment or the probe being in a sheltered location), the uncertainty analysis presented will not capture that. To capture that uncertainty in the data, we ideally need to compare several independent data sets. This is often possible when considering wind speeds as multiple data



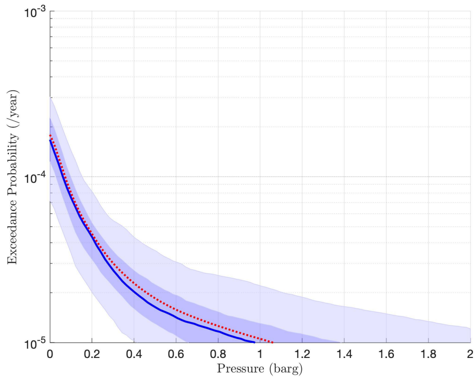
(a) Leak rate.



(b) Ignition probability



(c) Wind speed



(d) All

Figure 11: Probability (/year) of exceeding a pressure. Each plot only considers the uncertainty of the stated input distribution. The solid line shows the median value, the inner boundary the 25% and 75% quantiles and the outer boundary the 5% and 95% quantiles. The dashed red line shows the predictive posterior distribution.

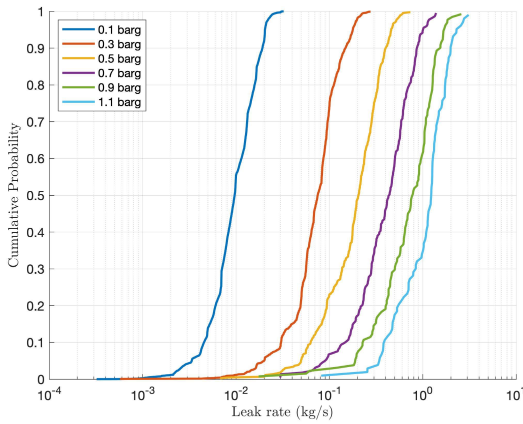


Figure 12: Cumulative probability that a given pressure comes from a scenario of a given leak rate.

sets from multiple weather stations may be available. By comparing a number from nearby locations we can estimate the measurement uncertainty in the data.

This is much more difficult when considering leak rate distributions and ignition probability distributions. Due to the generally rarity of events, we may not have multiple independent data sets to compare. One possibility (that has been consider in for example [9]) is to use data for the Norwegian Continental Shelf (NCS).

2. **Error associated with consequence models**
For our case study, we treated our consequence models as perfect — no uncertainty was accounted for in their results. In reality all consequence models will have an uncertainty associated with them. These mainly comes from two sources. The first is due to a lack of knowledge about the physical processes that go into building the model. For example, there is still a lot about turbulence modelling that is not understood. The second is due to known errors in the model due to simplifications that have been applied. These may, for example, be because of spatial or temporal discretisation or be because of deliberate simplification of the model to allow it to be effectively computed.

   CFD is generally used to perform explosion risk assessments. It is important to understand the uncertainty in these models and how it varies across the parameter space considered. Verification and validation (V&V) of CFD are the primary methods for quantifying this uncertainty (see, for example, [14]). This is challenging; but despite these challenges, it is important that a judgement on the uncertainty is included in the analysis.

3. **Error associated with sampling**
In our case study we were able to take a very large sample (10^8 scenarios). This is possible because of the very simple dispersion and explosion models used. For a real analysis, much more complicated models are used (such as CFD) that has a much more significant computational cost. As such, it is not possible to evaluate such a large number of scenarios. Most consultants will consider tens or hundreds of scenarios. With modern high performance computing, it may be possible to push that to evaluating thousands of scenarios in a sensible engineering timescale and cost. This is still far short of the number required to ensure we have converged on the solution if a naive Monte Carlo approach is used where one samples randomly from the parameter space.

   Instead, scenario selection is currently based primarily on engineering judgement. A number of simplifications are then made to expand a limited number of modelled scenarios to fill the parameter space (*e.g.* assuming linear scaling with wind speed, the frozen cloud assumption, and using equivalent stoichiometric clouds). However, it is not clear whether these simplifications are justifiable nor is it clear how many scenarios need to be considered to ensure an accurate answer. Small investigations have indicated a high degree of scatter due to the choice of scenarios by individual engineers [15].

   This reduced scenario selection, hence, introduces further uncertainty into the analysis that ideally should be properly analysed and presented as part of the final output.

Conclusions and future work

We have presented the early stages of an analysis of the uncertainty in the explosion risk assessment methodology. Focussing just on the statistical uncertainty associated with leak rate, ignition probability, and wind speed, we show that the uncertainty is significant and that it is important that it is presented in any output to prevent underdesign.

While our work highlights the uncertainty is large, we don't feel it is so large as to invalidate the technique altogether. Through further analysis we can fully quantify the uncertainty. This will allow analyses to show the full picture and remove under conservatism and under design. We can also work to bring consistency across the methodology, ensuring that different engineers arrive at the same solution when performing such analysis.

Ultimately, we aim to answer:

1. **How accurate are existing approaches?**
Our investigation highlights the weaknesses with the current techniques. Industry needs to know whether analyses performed over the last 10 – 20 years can be trusted or whether errors and uncertainties are so large as to mean we cannot be confident that facilities are designed adequately to meet the explosion hazard. As such we plan to review the full explosion risk assessment in its current forms (*i.e.* considering typically used leak and ignition models, the errors associated with the consequence modelling and the error associated with the sampling) to fully understand the true level of uncertainty. This then can be compared to safety factors used in design (and checking those safety factor aren't meant to cover other uncertainty, such as those arising in the structural analysis), we can determine whether reassessment of the risk and mitigation is required.

2. **How can we improve?**
We aim to provide a definitive answer as to whether thousands (or even tens of thousands or more) of scenarios need to be considered and modelled, or whether it is possible to select and consider a smaller number of scenarios and still produce accurate results.

By comparing possible approaches, we hope to be able to show that new methodologies for scenario selection can robustly provide the required degree of accuracy and precision.

Additionally, with a clear understanding of the uncertainties associated with the analysis, we can pinpoint areas where the largest gains can be made in the reduction of the uncertainty. For example:

- How much benefit can be made from directly modelling the full transient case rather than using the frozen cloud approximation?
- How significant are the errors associated with the equivalent stoichiometric cloud methodology to the overall uncertainty with explosion risk assessment?
- Should effort be made to improve uncertainties associated with inputs (*e.g.* leak frequencies and ignitions probabilities)?

References

[1] NORSOK. *Risk and emergency preparedness analysis*. 2001.

[2] HSE. Explosion Hazard Assessment, . URL http://www.hse.gov.uk/offshore/strategy/hazard.htm.

[3] HSE. Hydrocarbon releases system, . URL http://www.hse.gov.uk/offshore/hydrocarbon.htm.

[4] S. Coles. *An introduction to statistical modeling of extreme values*. Springer, 2001.

[5] C. Scarrott and A. MacDonald. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Statistical Journal*, 10(1):33–60, 2012.

[6] A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application*, volume 1. Cambridge University press, 1997.

[7] Andrew W Cox, Frank P Lees, and M Lm Ang. *Classification of hazardous locations*. IChemE, 1990.

[8] DNV. JIP ignition model. Technical Report DNV report 99-3193, 1999.

[9] Lloyds Register. Modelling of ignition sources on offshore oil and gas facilities - MISOF. Technical Report 106364/R1, 2016.

[10] A. Gelman, H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 2013.

[11] S. Davis and O. Hansen. Benefits of risk-based design through probabilistic consequence modeling. *Gcps 2011*, 2011.

[12] C. J. Lea and A. Kelsey. Evaluation of the workbook on gas accumulation in a confined and congested area. Technical Report HSL/2006/45, 2006.

[13] B. J. Wiekema. Vapour cloud explosion model. *Journal of Hazardous Materials*, 3(3):221–232, 1980.

[14] W. L. Oberkampf and T. G. Trucano. Verification and validation in computational fluid dynamics. *Progress in Aerospace sciences*, 38(3):209–272, 2002.

[15] C. Azzi, K. Rogstadkjernet, L .and van Wingerden, J. Choi, and Y. Ryu. Influence of scenario choices when performing CFD simulations for explosion risk analyses: Focus on dispersion. *Journal of Loss Prevention in the Process Industries*, 41:87–96, 2016.