FORMALIZING QUANTITATIVE RISK ANALYSIS

PROBABILISTIC FOUNDATIONS AND MONTE CARLO METHODS

JOHAN JÖNSSON

Bachelor's thesis
2025:K36

LUND INSTITUTE OF TECHNOLOGY
Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Formalizing Quantitative Risk Analysis: Probabilistic Foundations and Monte Carlo Methods

Johan Jönsson

December 18, 2025

Abstract

Quantitative Risk Analysis (QRA) for hazardous materials facilities remains dominated by event tree partitioning into discrete scenarios, with many calculations performed manually. This study formalizes the classical individual and societal risk metrics within the framework of probability theory. The resulting framework enables deeper theoretical inquiry and makes underlying assumptions explicit. This framework also enables quantitative risk analyses to utilize more advanced Monte Carlo methods commonly employed in more mature risk management disciplines.

A full year case study is presented for a hypothetical ammonia installation in Malmö, Sweden. The simulation combines ERA5 2024 meteorological data with a time-varying Gaussian puff model and a stochastic two-phase leak-rate generator. A total of 8 784 hourly realizations are compared against a conventional 216-scenario QRA that follows recommendations for a typical quantitative risk analysis. The Monte Carlo approach reproduces directional variability without the angular averaging required by the scenario method and eliminates the need for 100% lethality assumptions within a fixed radii.

The study demonstrates that a fully probabilistic QRA enhances clarity in the interpretation of risk measures and yields more precise spatial risk estimates with minimal additional analyst effort. Because the framework is grounded in probability theory, it imposes no strict constraints on model selection. These attributes position Monte Carlo QRA as a viable alternative to coarse scenario sets in future safety assessments.

Acknowledgements

This thesis is the culmination of what I have learned during my four years at Lund University, where I studied mathematical statistics and risk management.

I am deeply grateful to my thesis supervisor, Johan Lindström, for his guidance and support throughout the process. Because this thesis is not part of the standard program, I am especially thankful that he went out of his way to make it happen. Without him, this thesis would not have been possible.

I also extend my sincere thanks to the people at Branskyddslaget AB, who supported me by answering questions about how typical QRA is performed and by providing access to tools I otherwise would not have had.

Contents

2

List of Notation

(x, y) Spatial coordinates in the plane of the study area.

η_t Number of fatalities at time t, defined as $\eta_t = \int_{\mathbf{S}} D_t(s) P_t(s)\, ds$.

Λ_t Cumulative intensity, $\Lambda_t = \int_0^t \lambda_v\, dv$.

λ_t Intensity (rate of occurrence) of the counting process at time t.

$\mathbf{X}_t$ Stochastic process describing the effect-side state at time t.

$\mathbf{Y}_t$ Stochastic process describing the cause-side state at time t.

$\mathrm{Dose}_t(s)$ Total dose at location s due to an accident starting at time t.

Ω Sample space of all possible outcomes in the probability model.

ξ Random variable defined on the probability space $(\Omega, \mathcal{F}, P)$.

A_0, A_1 Parameters controlling the two-phase leak-rate profile (plateau height and decay amplitude).

$C(u, s, \tau)$ Dispersion model giving the concentration at location s and time τ for state u.

$D_t(s)$ Probability of death at position s given an accident starting at time t.

EN Total expected number of deaths across all scenarios.

EN_i Expected number of deaths for scenario i, i.e. N_i times the scenario frequency.

$f(n)$ Frequency that the number of deaths equals or exceeds n in the F–N curve.

f_i Annual frequency of scenario i.

F_N Annual frequency at which N or more individuals are affected.

$IR(s)$ Time-averaged individual risk at location s over a specified period.

$IR(x, y)$ Total individual risk at location (x, y), sum of all scenario contributions.

$IR_i(x, y)$ Individual risk contribution from scenario i at location (x, y).

$IR_t(s)$ Instantaneous individual risk at time t and location s.

$L(z)$ Lethality function mapping dose z to probability of fatality.

M Toxicological mapping from concentration to instantaneous dose.

n Exponent controlling the concentration–time trade-off in the toxicological model.

N_i Expected number of persons affected by scenario i.

N_t Counting process recording the cumulative number of accidents up to time t.

$P(x, y)$ Number of people at point (x, y) (population density).

$P_t(s)$ Population density at location s and time t.

$p_{f,i}(x, y)$ Conditional probability of fatality at (x, y) given scenario i.

$r(t; A_0, A_1)$ Instantaneous leak-rate function used in the release process.

t_0, τ Plateau duration (t_0) and exponential decay time constant (τ) in the leak-rate model.

t_{duration} Duration of the exposure window for an accident.

Chapter 1

Introduction

Modern societal safety analysis still leans heavily on empirically derived formulas and scenario-based event trees, with relatively limited use of formal statistical reasoning (Ingvarson et al., 2022). The aim of this thesis is to place such empirical formulas on a solid probabilistic foundation enabling the application of modern statistical methods. With the rapid growth in available computational power, far more detailed models and much larger numbers of simulations have become feasible. This makes it possible to replace coarse, discrete scenario analyses with more precise Monte Carlo simulation techniques that are widely used in other parts of reliability engineering and risk analysis (Zio, 2013).

1.1 Background on Quantitative Risk Analysis

The primary objective of this form of quantitative risk analysis (QRA) is to evaluate societal safety, a task undertaken by numerous stakeholders. QRAs are commonly applied to industrial facilities or to the transport of hazardous materials by road and rail. This thesis uses *process safety* as an example to show how re-derived risk measures can be applied, with the principal aim of assessing human consequences; in particular, fatality risk. According to Nassiri et al. (2021), a QRA seeks to estimate the risk posed to third parties (the general public) and, in some cases, to second parties. In the Swedish regulatory framework, these analyses serve as the basis for permit decisions, and the basic details are outlined in the report *process safety*[1]. These national documents are often based on the Seveso Directive (*Directive 2012/18/EU. On the control of major-accident hazards involving dangerous substances*, 2012). Although much of the referenced material pertains specifically to the Swedish context, the methods themselves are widely applicable to quantitative risk analyses conducted worldwide. A complete QRA requires many preparatory steps beyond those presented

---

[1]We use the CCPS definition: "A disciplined framework for managing the integrity of operating systems and processes handling hazardous substances by applying good design principles, engineering, and operating practices. " (Center for Chemical Process Safety (CCPS), n.d.)

here; in this work, we focus on the specific formulas for individual and societal risk. Accordingly, the analysis focuses exclusively on the derivation and notation of the formulas used to calculate these risk metrics, without surveying the full spectrum of hazard-specific methodologies (e.g., BLEVE or jet-flame analyses). For detailed guidance on calculation methods for particular hazards, see Ingvarson et al. (2022).

1.2 Previous work

1.2.1 Formalizing Time-Dependent Risk in QRA

Raveendran et al. (2022) reviews how QRA techniques are evolving beyond static analyses toward more flexible, time-adaptive frameworks. This evolution underscores the need to recast conventional risk metrics within a rigorous probabilistic framework, particularly using stochastic process theory. Such a reformulation broadly generalizes the formulas, enabling QRA to accurately capture different types of uncertainty, temporal variability, and operational dynamics. This reformulation is also compatible with the various modeling approaches discussed, including machine learning methods and Bayesian frameworks.

1.2.2 Impact of Scenario Granularity on Risk Contours

Marx & Cornwell (2009) demonstrates that the granularity of discretized meteorological input data, such as the number of wind directions considered, significantly influences the calculated spatial distribution of individual risk in quantitative risk analysis, where atmospheric dispersion is a driving factor. This spatial distribution is commonly represented using risk contours, that is, isolines on a map connecting points with the same individual risk level. By employing 64 wind directions and 21 atmospheric stability/velocity classes, the authors generated a total of 1344 distinct weather scenarios. When the number of wind directions was reduced from 64 to 16, the area enclosed by the $10^{-6}\,\mathrm{yr}^{-1}$ individual risk contour decreased by 6.3%. Further simplification to 16 directions and 6 stability/velocity classes resulted in an underestimation of 27.7% of this area. The most severe reduction to eight directions and two classes produced an underestimation of 45.9%. These findings highlight that the overall number of scenarios, rather than directional resolution alone, governs the accuracy of the risk model.

1.3   Purpose

This study extends the traditional scenario-based approach in two steps. First, in A Probabilistic Methodology for QRA, it develops a general probabilistic framework for QRA by formally deriving probabilistic formulas for classical risk measures. Second, in Case Study and Benchmarking, it applies this framework to atmospheric dispersion and examines the implications of using a scenario-based method compared to a fully time-resolved Monte Carlo simulation grounded in these probabilistic formulas.

1.3.1   Research Questions

- Can the classical QRA risk measures, such as individual risk and societal risk, be derived and defined consistently within a probability theoretic framework?

- Does a Monte Carlo simulation that directly employs time-resolved empirical weather data provide more reliable estimates of individual and societal risk than the traditional scenario-based method?

- Is the model sufficiently modular to be reused across facilities via simple data exchange, thereby reducing project lead time?

Chapter 2

A Probabilistic Methodology for QRA

2.1 Traditional definitions of individual and societal risk measures

The classical formulas for risk measures prioritize practical applicability over a formal theoretic foundation. This approach makes the expressions accessible to practitioners with limited exposure to probability theory and allow straightforward implementation in commonly used spreadsheet software. Nonetheless, this simplicity often omits critical nuances and precludes the use of more advanced statistical tools. One such nuance is the precise treatment of scenario additivity, as discussed in section 2.1.3. However, it remains important to retain these conventional formulations, as their familiarity aids in communicating the core decision principles to a broad audience. The formulas in the next sections are taken from Ingvarson et al. (2022), with only minimal modifications to clarify the spatial component.

2.1.1 Individual risk

In quantitative risk analysis, individual risk is the foundational risk measure on which other measures are built, and it represents the likelihood that a reference person suffers a fatal outcome when exposed to hazardous conditions over a defined time interval, typically one year. The reference individual is assumed to be of average sensitivity, unprotected, and remaining at a fixed location throughout the period. Formally, the individual risk at a point (x, y) is calculated by summing the contributions of all hazardous scenarios that may affect that location. Let f_i denote the annual frequency of scenario i, interpreted as the expected number of accidents per year, and let $p_{f,i}(x, y)$ denote the conditional

probability of fatality at (x, y) given scenario i.

$$IR_i(x, y) = f_i \cdot p_{f,i}(x, y) \tag{2.1}$$

$$IR(x, y) = \sum_{i=1}^{n} IR_i(x, y) \tag{2.2}$$

where $IR_i(x, y)$ denotes the individual risk contribution from event i and $IR(x, y)$ denotes the total individual risk. Individual risk is therefore interpreted as the annual frequency of fatality at (x, y), under the assumption of time-homogeneous conditions across years.

These values are often visualized using individual risk contours, which provide a spatial representation of risk levels across a geographic area (Ingvarson et al., 2022), see fig. 2.1.
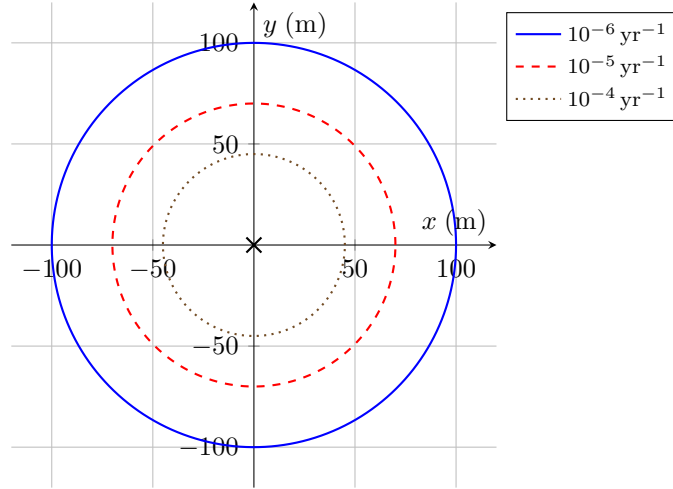

Figure 2.1: Illustrative individual-risk contours around a point source. The X marks the point-source hazard, and the red and blue contour lines bound the set of locations where the individual risk equals $10^{-6}\,\mathrm{yr}^{-1}$. In practical terms, an unprotected person of average sensitivity who remains at a fixed position anywhere within this band for one year faces a one-in-a-million chance of a fatal outcome.

2.1.2 Societal risk

Societal risk measures aim to estimate the expected number of individuals affected by the hazard per time interval and are built on individual risk measures. The two primary metrics are the expected number of persons affected and the F-N curve. The F-N(Frequency–Number of fatalities) curve represents the annual frequency at which N or more individuals are affected. The corresponding mathematical definitions are provided below (Ingvarson et al., 2022):

The expected number of affected given that scenario i has happened is given by

$$N_i = \sum_{x,y} P(x,y) \cdot p_{f,i}(x,y) \tag{2.3}$$

where $P(x,y)$ is the number of people at point (x,y), and $p_{f,i}$ is the probability of a fatal outcome at (x,y) given scenario i. The time-averaged expectation is obtained by weighting by the scenario's frequency.

$$EN_i = N_i \cdot f_i = \sum_{x,y} P(x,y) \cdot IR_i(x,y) = \sum_{x,y} P(x,y) \cdot f_i \cdot p_{f,i}(x,y) \tag{2.4}$$

which is just the expected number of deaths for scenario i. Then the total number of expected deaths is:

$$EN = \sum_{x,y} P(x,y) \cdot IR(x,y) = \sum_{x,y} \left[P(x,y) \sum_{i=1}^{n} IR_i(x,y) \right] \tag{2.5}$$

F-N Diagrams

F-N diagrams/curves can be used to present societal risk and also to determine whether, and how, risk criteria might be violated, and they provide a more nuanced explanation of the different accident death-toll outcomes that might occur. An F-N diagram is usually derived from a QRA or empirical observations, where the x-axis consists of the consequences, typically the expected number of deaths, and the y-axis represents the cumulative frequency of N or more fatalities. The formula for calculating F_N is the following:

$$F_N = \sum_{i:N_i \geq N} F_i \tag{2.6}$$

i.e., the sum runs over all scenarios with $N_i \geq N$; F_N is the cumulative frequency of scenarios in which N or more people are affected, and F_i is the frequency of scenario i.

2.1.3 Bow-tie modeling

The bow-tie analysis framework is widely used for risk assessment; it serves as a bridge between fault tree and event tree (CCPS, 2018) and is therefore a useful conceptual framework for understanding a QRA. In this approach, the Top Event denotes the point of transition, marking the endpoint of the fault tree and the starting point of the event tree. Details on how to structure a qualitative Bow-tie model are discussed in greater depth in the report (MSB, 2017).

In practice, the top-event frequency from the fault tree is apportioned across the terminal branches of the event tree. Because the terminal scenarios $\{i\}$ are,
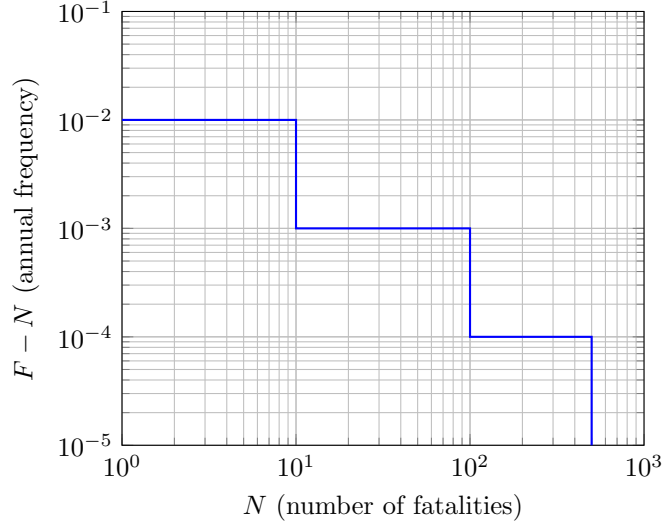
Figure 2.2: Illustrative F–N curves using scenarios.

by construction, mutually exclusive and collectively exhaustive, the additivity in (2.2) follows from the law of total probability:

$$IR(x,y) = \sum_i f_i \, p_{f,i}(x,y).$$

This partition-based calculation is a special case of the continuous-time formulation introduced later. In settings where scenarios are not disjoint, for example, when multiple failure modes may occur within the same period, the later formulation provides a time-resolved generalization that retains additivity, but only in expectation.

The bow-tie representation provides a convenient structure for both qualitative and quantitative descriptions of risk in a QRA. However, there are several ways to structure a QRA, the bow-tie framework is used here as a practical starting point because it aligns with how the standard scenario formulas in the QRA literature are typically interpreted. In this thesis, the bow-tie is therefore not presented as a unique qualitative framework for understanding risk, but rather as a conceptual starting point from which the standard formulas are developed into more advanced ones.

2.2 Mathematical Formalization of Methods Used for QRAs

The previous section presented the classical formulas employed in QRAs. However, those formulas are grounded in practical and empirical considerations
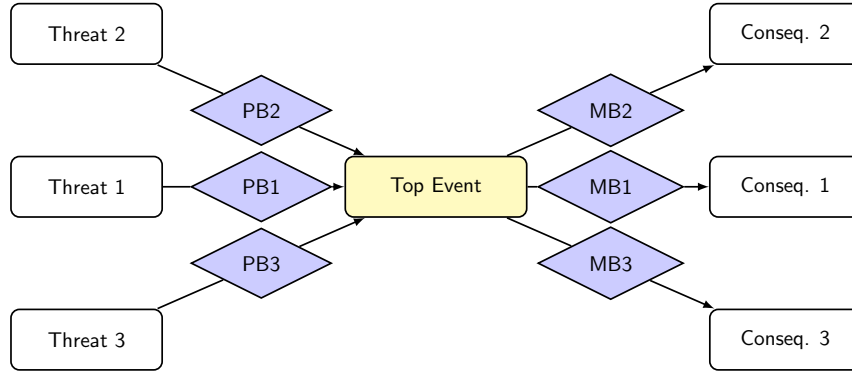
Figure 2.3: A typical bow-tie model: threats initiate the top event, and consequences are the possible outcomes once it occurs. MB denotes mitigation barriers; PB denotes preventive barriers.

rather than in a formal probability theoretic framework. The objective of this section is therefore to reformulate these expressions and to derive, from the original formulas, versions expressed in the notation of probability theory. This approach will enable the use of standard methods from modern statistics, which often form the basis of risk models in disciplines such as quantitative finance. Consequently, many advanced Monte Carlo simulation techniques become readily applicable; such as variance reduction techniques. Before developing these formalisms, we introduce several definitions and key results from probability theory.

2.2.1 Initial definitions and results

Probability theory has two fundamental objects that are necessary for working with it: the *probability space* and the *random variable*. These provide the formal foundation and notation for all subsequent results in this chapter, linking the formulas directly to probability theory.

Definition 1: *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let*

$$\xi \colon \Omega \longrightarrow \mathbb{R}$$

be a random variable.

For readers who are less familiar with probability notation, an intuitive explanation is provided. The symbol Ω, referred to as the sample space, denotes the set of all possible outcomes or scenarios that are consistent with the specified model assumptions and constraints. For each outcome $\omega \in \Omega$, the random variable $\xi(\omega)$ yields a corresponding value. The estimation and characterization of uncertainty, both within and beyond the model, lie outside the scope of this thesis. These matters can be addressed by various methods in probability theory, and we defer to the reader's preferred approach. Two common and distinct

options are a frequentist approach (estimating model parameters from data, for example using maximum likelihood) or a Bayesian approach (placing prior distributions on parameters and updating them to posteriors), and either choice is compatible.

The following theorem provides intuition as to why approximating scenarios with discrete events amounts to approximating a random variable, and it helps clarify how the classical method is a simplification of a more complex and nuanced mathematical framework. The intuition is that if we partition Ω into finitely many events and replace $\xi(\omega)$ by a representative value on each event, then we obtain a simple random variable that can approximate the original random variable ξ arbitrarily well as the partition is refined. Conceptually, this is analogous to approximating a continuous function by a piecewise-constant (finite) approximation, which can be made accurate by using sufficiently fine intervals.

Theorem 1 (Approximation by simple random variables)**:** *For every random variable $\xi : \Omega \to \mathbb{R}$ there exists an increasing sequence of simple random variables $\{\xi_n\}_{n \geq 1}$ such that*

$$|\xi_n(\omega)| \leq |\xi(\omega)|, \qquad \xi_n(\omega) \uparrow \xi(\omega) \quad \text{for all } \omega \in \Omega.$$

Proof. This proof treats the special case in which $\xi(\omega) \geq 0$. For $n \in \mathbb{N}$ define

$$\xi_{n,i}(\omega) = \frac{i-1}{2^n} \mathbf{1}_{\{(i-1)/2^n \leq \xi(\omega) < i/2^n\}}(\omega)$$

$$\xi_n(\omega) = \sum_{i=1}^{n2^n} \xi_{n,i}(\omega) + n \, \mathbf{1}_{\{\xi(\omega) \geq n\}}(\omega)$$

As the intervals $[(i-1)/2^n, \, i/2^n)$ become finer with increasing n, we have

$$\xi_n(\omega) \leq \xi_{n+1}(\omega) \leq \xi(\omega), \quad \xi_n(\omega) \uparrow \xi(\omega) \quad \text{for all } \omega \in \Omega.$$

The inequality $\xi_n(\omega) < \xi(\omega)$ is immediate from the construction. This result is standard; see, for example, Shiryaev (2016, theorem 2.4.1). $\qquad \square$

Applying this intuition more concretely to our case, suppose that the probability of death D, viewed as a random variable, has been calculated for a given hazard and that this probability varies with relevant input parameters. As the number of scenarios increases, theorem 1 implies that D can be approximated with arbitrarily fine resolution. Specifically, in a classical QRA framework, one approximates D by a simple and discrete random variable D_n that takes only finitely many values, where a representative value is assigned on each part of Ω.

Since D represents a probability, it is almost surely bounded between 0 and 1; that is,

$$\mathbb{P}(0 \leq D \leq 1) = 1.$$

Hence, the interval $[0, 1]$ may be divided into $2^n + 1$ equal-length subintervals. For example, when $n = 2$, one obtains the following event-tree representation:
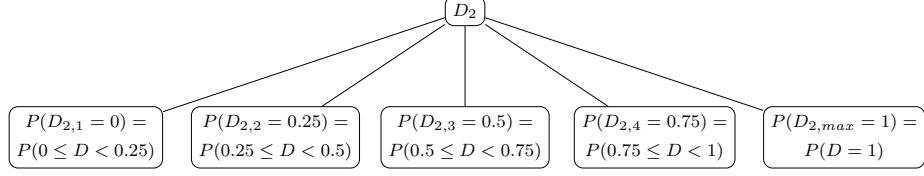
Figure 2.4: Visualization of how we split up our continuous random variable into four discrete scenarios.

By allowing n to increase without bound the theorem implies that $D_n \to D$ almost surely, thus recovering the original random variable as the scenario count grows. This convergence provides the basis for the subsequent derivation of risk formulas.

Moreover, the result is closely connected to one of the central concepts in risk science used to describe uncertainty, namely the risk triplet (s_i, p_i, c_i), which provides a direct bridge between a scenario-based QRA description and the probabilistic formulation used in this thesis. The risk triplet (s_i, p_i, c_i) naturally has three parts, where s_i denotes the ith scenario, p_i its probability, and c_i the corresponding consequence (Aven, 2020). To illustrate, consider again the discretization of D into five scenarios as in fig. 2.4. Each discrete outcome $D_{2,i}(\omega)$ then embodies the three components of the risk triplet. Specifically, the scenario element is given by

$$ S_i = \left\{ \omega \in \Omega : \frac{i-1}{2^2} \le D(\omega) < \frac{i}{2^2} \right\}, $$

the probability component is

$$ p_i = \mathbb{P}(S_i), $$

and the consequence is represented by

$$ c_i = \frac{i-1}{2^2}. $$

Thus, the process of approximating D by simple random variables yields a natural formalization of the risk triplet within a probabilistic framework.

2.2.2 Bow-tie representation of the state vectors

To structure the analysis similarly to how a traditional QRA is presented, the model is represented by *cause* and *effect* components, mirroring a bow-tie diagram. Let $\mathbf{Y}_t$ (cause) and $\mathbf{X}_t$ (effect) be stochastic processes on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0}, \mathbb{P})$ satisfying the usual conditions (right-continuous and complete) (Nikeghbali, 2006). Indexing by time, $t \mapsto (\mathbf{Y}_t, \mathbf{X}_t)$, treats $\{\mathbf{Y}_t\}_{t \ge 0}$ and $\{\mathbf{X}_t\}_{t \ge 0}$ as stochastic processes that record system evolution: for each t, $(\mathbf{Y}_t, \mathbf{X}_t)$ is a random outcome, and as t varies the collection traces the temporal dynamics.
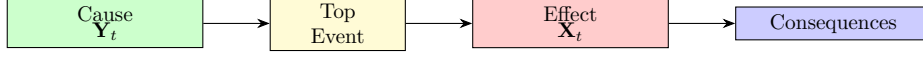
14

Figure 2.5: Bow-tie representation of the cause($\mathbf{Y}_t$) and effect($\mathbf{X}_t$) components.

2.2.3 Counting process for events

To model whether an accident has occurred at a given time, a counting process is employed. This choice is motivated by the fact that a counting process records the cumulative number of events up to time t, thereby capturing both whether an accident has occurred and when it occurs. Formally it is defined as

Definition 2: *A random process $\{N_t,\ t \geq 0\}$ is a **counting process** on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ if it is $(\mathcal{F}_t)$-adapted and satisfies the following conditions:*

1. *The trajectories of N_t are, with probability one, right-continuous and piece-wise constant.*

2. *The process starts at zero, so $N_0 = 0$.*

3. *For each t,*
$$\Delta N_t = 0 \quad or \quad \Delta N_t = 1,$$
   with probability one. Here ΔN_t denotes the jump of N at time t, more formally defined as
$$\Delta N_t = N_t - N_{t-}.$$

This is a standard definition; see Björk (2021). It encompasses well–known processes, including the homogeneous and inhomogeneous Poisson processes. Since a counting process naturally captures the instant at which an accident occurs, it serves as a good basis for temporal modeling.

To construct the top-event counter from the cause process, let h map the history of $(\mathbf{Y}_t)$ to the cumulative number of failures by time t, and defined as

$$N_t \ :=\ h\big(\mathbf{Y}_v : 0 \leq v \leq t\big).$$

Thus N_t is the top-event counting process which equals the number of accidents that have occurred up to time t. Each unit jump $\Delta N_t = 1$ marks an accident time. A representative sample path is shown in fig. 2.6.
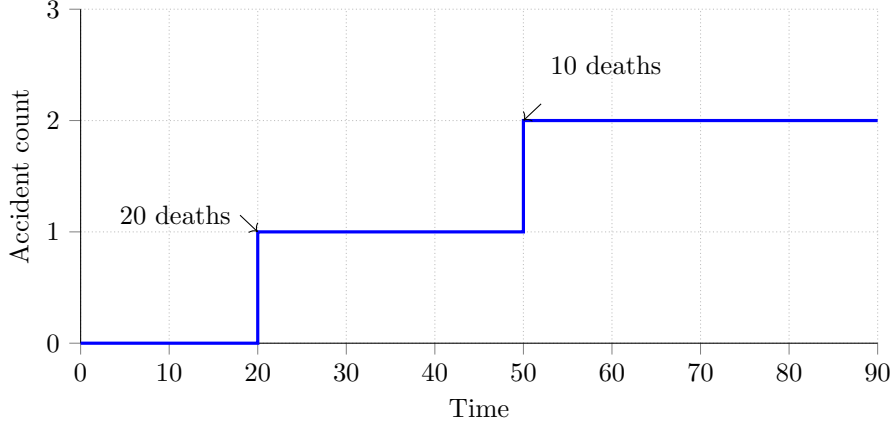
Figure 2.6: Representative sample path of the counting process N_t, in which unit jumps mark accident occurrence times; annotations indicate the associated death toll for each accident.

Definition 3 (Rate of occurrence of failure (ROCOF)): *For the purpose of keeping the theory manageable, assume that N is a* **inhomogeneous Poisson process (NHPP)**, *with properties:*

$$\lambda_t = \frac{d}{dt}\,\mathbb{E}\big[N_t\big] = \lim_{\Delta t \downarrow 0} \frac{\mathbb{P}\big(N_t - N_{t-\Delta t} = 1\big)}{\Delta t},$$

1. *$N(0) = 0$;*

2. *$\{N_t, t \geq 0\}$ has independent increments;*

3. *$\mathbb{P}\big(N_t - N_{t-\Delta t} \geq 2\big) = o(\Delta t)$;*

4. *$\mathbb{P}\big(N_t - N_{t-\Delta t} = 1\big) = \lambda_t\,\Delta t + o(\Delta t)$.*

5. *$\Lambda_t = \int_0^t \lambda_v\,dv$.*

A constant $\lambda_t = \lambda$ yields a homogeneous Poisson process (Høyland & Rausand, 1994).

To connect the stochastic formulation to the traditional Bow-tie representation used in QRAs, consider a simple fault tree with three basic events feeding a single OR gate. Let $N_t^{(i)}$ denote the counting process that records failures of component i up to time t; under the OR gate the top event occurs as soon as any $N_t^{(i)}$ jumps. The cause state is then collected in the vector process, so that

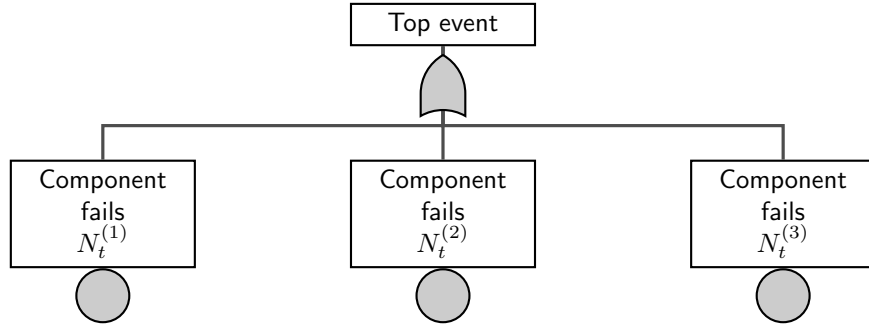$$\mathbf{Y}(t) = \big(N_t^{(1)}, N_t^{(2)}, N_t^{(3)}\big),$$

16

Figure 2.7: Fault tree with three basic events feeding a single OR gate. The top event occurs when any component process $N_t^{(i)}$ jumps.

Assuming that each component in fig. 2.7 fails independently of the others, evolves as an inhomogeneous Poisson counting process $N_t^{(i)}$ with rate $\lambda_t^{(i)} > 0$, and is restored immediately to an as-good-as-new state after failure (perfect repair), the OR gate triggers the top event whenever any component fails. The corresponding top-event counting process is therefore the sum of the component processes,

$$N_t^{\text{Top}} = h\big(\mathbf{Y}_v : 0 \leq v \leq t\big) = \sum_{i=1}^{3} N_t^{(i)},$$

and, by superposition of independent NHPPs, N_t^{Top} is itself an inhomogeneous Poisson process with rate

$$\lambda_t^{\text{Top}} = \lambda_t^{(1)} + \lambda_t^{(2)} + \lambda_t^{(3)}.$$

Hence, a fault tree with three independent, perfect-repair components feeding a single OR gate produces an NHPP for the top event. This exemplifies how the modeling choice of a general stochastic-process framework readily captures standard models.

The assumption of an NHPP, with deterministic λ_t and independent increments, is standard in reliability analysis and is sufficient for all results in Secs. 2–3. More advanced treatments that relax these simplifying assumptions would require tools from stochastic calculus, which are beyond the scope of this work. However, how one models $\mathbf{Y}_t$ and $\mathbf{X}_t$ is not restricted to bow-tie models. Rather, any model within the general theory of stochastic processes can be employed (Nikeghbali, 2006). What matters is having a well-defined scenario in which the hazards either occur or do not occur at a specified time, allowing one or more counting processes (see definition 2) to be defined. In principle, this formulation can encompass a broad range of classical dynamic QRA models (Raveendran et al., 2022).

2.2.4 Modeling the probability of effect(event-tree)

The next step is to quantify the effect of the hazard using $\mathbf{X}_t$, following the standard QRA workflow for scenario specific fatality probabilities. First, a dispersion model, for example a *Gaussian plume*, describes how the release spreads in space and time (Stellan et al., 1995). Given the concentration at a receptor location, the total dose over the exposure window is computed, and this dose is then mapped to a fatality probability through a dose–response function. A probit specification is commonly used (Stellan et al., 1995), although other choices, for example logit or Weibull type models, are equally admissible and should be selected based on the chosen exposed individuals and the available data.

For notational clarity, t denotes the accident start time, that is a jump time of N, whereas τ is reserved as the integration variable over the evolution of the accident at time t. Conditional on an accident beginning at t, the exposure window is $[t, t + t_{\text{duration}}]$, and the following integrals are taken with respect to τ over this interval. Henceforth the analysis is conditional on an accident having occurred at time t. The danger attributable to that specific accident is obtained by allowing the concentration to evolve in time and accumulating its contribution via integration, assuming a concentration profile $c_\tau(s)$ at spatial coordinate s.[1]

$$\text{Dose}_t(s) \;=\; \int_t^{t+t_{\text{duration}}} \big[c(s)_\tau\big]^n \, d\tau,$$

where $c(s)_\tau$ denotes the concentration at the receptor as the accident evolves, and $n > 0$ controls the concentration and time trade off, for example $n = 1$ recovers Haber's rule, while $n > 1$ places relatively greater weight on high concentrations (Gaylor, 2000).

In standard QRA, once the event specific dose $\text{Dose}_t(s)$ has been computed for an accident starting at t, the fatality probability is obtained by applying a dose–response, that is a lethality, function. A widely used choice is the probit model

$$L(z) \;=\; \Phi\big(\alpha + \beta_2 \log z\big),$$

where Φ is the standard normal CDF, and α, β_2 are parameters specific to the chemical and the chosen exposed individuals assessed in the QRA. Although the probit form is conventional for acute lethality in QRA, other dose–response functions can be used when supported by data. The present analysis focuses on acute effects over the exposure window, long term outcomes can be addressed by choosing an L tailored to chronic responses.

The first step in formalizing these steps is to define C, the dispersion model, over the spatial domain $S \subset \mathbb{R}^d$, and to let $T \subseteq \mathbb{R}$ denote the time domain. The purpose of this function is to determine the concentration in space and time,

---

[1]Here s denotes a generic spatial coordinate to keep the notation dimension-agnostic. This allows the same dose definition to be used whether the dispersion model is formulated on a line ($s \in \mathbb{R}$), in the plane ($s \in \mathbb{R}^2$, corresponding to (x, y)), or in three dimensions ($s \in \mathbb{R}^3$, corresponding to (x, y, z)).

and it is formally defined as

$$C \colon \mathbb{R}^m \times S \times T \to [0, \infty), \quad (u, s, \tau) \mapsto C(u, s, \tau). \tag{2.7}$$

Here, $u \in \mathbb{R}^m$ collects scenario and environment variables, for example wind speed and direction, atmospheric stability, release rate, hole diameter, and mitigation status. $s \in S$ denotes the receptor location, for example a geographic coordinate or grid cell. $\tau \in T$ denotes time, for example seconds since the start of the accident. The function $C(u, s, \tau)$ thus returns the concentration at location s at time τ under state u. $C(u, s, \tau)$ will serve as the concentration input to subsequent toxicological calculations.

Let M denote the toxicological function from concentration, with time input, to the instantaneous dose

$$M \colon [0, \infty) \times T \to [0, \infty). \tag{2.8}$$

Under the standard formulation used with probit lethality, M raises concentration to a positive exponent,

$$M(C(u, s, \tau), \tau) \ = \ \big[C(u, s, \tau)\big]^n, \qquad n > 0,$$

which captures how short high concentrations and longer lower concentrations contribute differently to dose.

M converts concentration to instantaneous dose, and combined with the dispersion model C, it yields the total dose at position s from an accident that starts at time t and lasts until $t + t_{\mathrm{duration}}$ by integrating over the exposure window

$$\mathrm{Dose}_t(s) \ = \ \int_t^{t+t_{\mathrm{duration}}} M(C(\mathbf{X}_\tau, s, \tau), \, \tau) \, d\tau.$$

The final step is to introduce the lethality function, which completes the formalization of QRA workflows by linking the accumulated dose to the probability of death following exposure to the leak.

$$L \colon [0, \infty) \to [0, 1], \tag{2.9}$$

which maps the cumulative dose to the probability of fatality. In what follows, L is treated abstractly and includes the probit specification introduced earlier as a special case. The parameterization of L should be chosen to reflect the chemical and the chosen exposed individuals assessed in the QRA.

Then the probability of death at position s following a leak at time t is thus

$$D_t(s) \ = \ L(\mathrm{Dose}_t(s)) \ = \ L\!\left(\int_t^{t+t_{\mathrm{duration}}} M(C(\mathbf{X}_\tau, s, \tau), \, \tau) \, d\tau\right). \tag{2.10}$$

Equation (2.10) completes the causal chain from event occurrence to the chosen consequence measure. Conditional on $\Delta N_t = 1$, $D_t(s) \in [0, 1]$ is the acute fatality probability for an individual located at s due to an accident that starts at time t. In the integral, the dispersion model C evaluated along the realized

state path $\mathbf{X}_\tau$ generates the concentration profile; the toxicological mapping M encodes the concentration–time trade-off via the exponent n; and the lethality function L converts the accumulated dose to a probability. Hence $D_t(s)$ may depend on factors such as weather and release characteristics. The receptor location influences the encountered concentration and its toxicological transformation through M, and the chosen chemical and exposed individuals enter through the parameterization of L. Since $\mathbf{X}_\tau$ is random, $D_t(s)$ is a random variable; event-level summaries are therefore obtained through conditional expectations given $\Delta N_t = 1$, which will be introduced more formally in the next section.

2.2.5 Individual Risk Assessment

Using the random variables N_t and $D_t(s)$ introduced in sections 2.2.3 and 2.2.4, the bow-tie representation becomes explicit: the cause side yields the top-event indicator ΔN_t, and the effect side yields the random consequence $D_t(s)$. The analysis is therefore naturally formulated in terms of conditional random variables, for example $D_t(s) \mid \Delta N_t = 1$, and their conditional expectations.
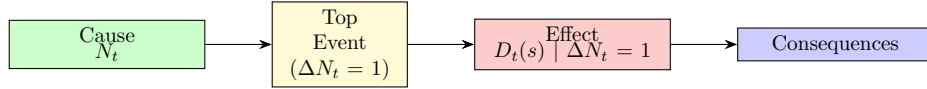

Figure 2.8: Bow-tie representation of the cause(N_t) and effect($D_t(s)$) components.

In this continuous-time formulation with an indicator of the top event, the quantities in (2.2) admit the following analogues: the scenario frequency f_i is replaced by the instantaneous event rate λ_t of the counting process N_t. The purpose of this step is to shift from a scenario-indexed annual frequency f_i to a time-dependent occurrence rate λ_t. Here the index i labels mutually exclusive scenarios in the classical formulation, whereas t indexes time in the counting-process formulation. In principle, one could introduce cause-specific intensities equivalently $\lambda_{i,t}$ to mirror the original scenario decomposition, but this refinement is omitted to keep the derivation focused and the notation manageable. The scenario-level fatality probability $p_{f,i}(x,y)$ is replaced by the conditional random variable $D_t(s) \mid \Delta N_t = 1$. Accordingly, the *instantaneous expected individual fatality probability* is defined as:

$$\lim_{\Delta t \downarrow 0} \mathbb{E}[D_{U_1} \mid N_t - N_{t-\Delta t} = 1] = \mathbb{E}[D_t \mid \Delta N_t = 1]. \tag{2.11}$$

More formal treatment can be developed within stochastic calculus, for example by working with conditional expectations and projections (Nikeghbali, 2006). For the present purposes the essential step is to condition on the occurrence of an accident at time t and to consider the conditional expectation of formula (2.10) given $\Delta N_t = 1$.

The instantaneous rate of death (instantaneous individual risk) is therefore

$$\mathrm{IR}_t(s) = \mathbb{E}\big[D_t(s) \mid \Delta N_t = 1\big]\lambda_t. \tag{2.12}$$

with a detailed derivation provided in appendix A. Since N is a counting process with unit jumps (in particular, a Poisson process), at most one accident can occur in an infinitesimal interval, i.e., $\Delta N_t \in \{0,1\}$. It follows that the instantaneous death rate equals the product of the event rate and the conditional expected fatality probability as seen in formula (2.12). If both λ_t and $\mathbb{E}\big[D_t(s) \mid \Delta N_t = 1\big]$ are constant in time, then formula (2.12) reduces to the classical QRA form $f_i\, p_{f,i}(x,y)$ under the identifications $\lambda_t = f_i$ and $\mathbb{E}\big[D_t(s) \mid \Delta N_t = 1\big] = p_{f,i}(x,y)$; see formula (2.2).

To assess the impact over a time window $[t_0, t_1]$, define the *time–averaged individual risk* at location s as the mean, across time, of the instantaneous rate $\mathrm{IR}_t(s)$. This quantity equals the expected number of fatalities for a continuously present individual at s accumulated over $[t_0, t_1]$ and divided by the window length. Equivalently, it is the expected number of fatalities per unit time (for example, per year):

$$IR(s) \;=\; \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \mathrm{IR}_t(s)\, dt. \tag{2.13}$$

When λ_t and $\mathbb{E}[D_t(s) \mid \Delta N_t = 1]$ are time-invariant, the time average equals the instantaneous value, so

$$IR(s) = \mathbb{E}[D_t(s) \mid \Delta N_t = 1]\, \lambda_t,$$

which reduces to the classical definition of individual risk; see formula (2.2).

The new definition of individual risk can be interpreted as follows: Over a short interval $(t - \Delta t,\, t]$ the expected number of accidents is $\Lambda_t - \Lambda_{t-\Delta t} \approx \lambda_t\, \Delta t$. Conditional on an accident in that interval ($\Delta N_t = 1$), $\mathbb{E}[D_t(s) \mid \Delta N_t = 1]$ is the probability that an individual located at s dies from that accident. The expected fatalities accrued in $(t - \Delta t,\, t]$ are therefore $\mathbb{E}[D_t(s) \mid \Delta N_t = 1]\big(\Lambda_t - \Lambda_{t-\Delta t}\big) \approx \mathrm{IR}_t(s)\, \Delta t$, with $\mathrm{IR}_t(s) = \mathbb{E}[D_t(s) \mid \Delta N_t = 1]\lambda_t$. Integrating this instantaneous individual risk over $[t_0, t_1]$ and dividing by the window length yields the time–averaged individual risk in formula (2.13). This formulation therefore generalizes the original definition, encompassing it as a special case.

2.2.6   Expected number of deaths

The expected number of deaths is a central risk measure in QRAs. To formalize it, let $S \subset \mathbb{R}^d$ and $T \subseteq \mathbb{R}$ denote the spatial and time domains introduced above, and define the population density field

$$P_t \colon S \to [0, \infty), \qquad s \mapsto P_t(s),$$

which gives the number of exposed individuals per unit area at location s and time t. The process $\{P_t(s),\, t \in T\}$ is taken to be $(\mathcal{F}_t)$-adapted on the same

filtered probability space as before. While temporal variation is obvious, for example, people sleep at night and go to work in the day, there is additional randomness in that pattern due to factors that are difficult to predict, for example, work-from-home days, illness, and temporary gatherings; as a result, the population density $P_t(s)$ remains stochastic even after accounting for regular schedules.

Analogous to formula (2.3), which computes the number of deaths at a location as probability of death times population present, the *realization–level* fatality density at time t and position s is thus written as

$$D_t(s)\, P_t(s).$$

It is important to distinguish this realized number of deaths from the *statistical expectation* across the randomness in exposure, toxicity, and population. Since $P_t(s)$ is a population *density* (hence may be continuous) and $D_t(s)$ is random, the expected number of fatalities accumulated over a space–time window is

$$EN \;=\; \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \int_{\mathbf{S}} \mathbb{E}[\, D_t(s)\, P_t(s) \mid \Delta N_t = 1]\, \lambda_t\, ds\, dt. \qquad (2.14)$$

Here the inner expectation gives the expected fatalities per unit area at (t, s) *given* that an accident occurs at time t, and multiplication by λ_t accounts for the event rate; integration over space and time then yields the expected total fatalities. Dividing by the time window finally yields the time-averaged number of deaths per year.

A necessary assumption is integrability of the product. The assumed requirement is that $\mathbb{E}[D_t(s)\, P_t(s) \mid \Delta N_t = 1] < \infty$, which enables the derivation from formula (2.5) to proceed by replacing the summation over (x, y) with an integral over the spatial domain. It should also be noted that formula (2.5) relies on an unstated assumption: conditional independence, given $\Delta N_t = 1$, between $D_t(s)$ and $P_t(s)$, or an equivalent factorization condition ensuring that

$$\mathbb{E}[D_t(s)\, P_t(s) \mid \Delta N_t = 1] = \mathbb{E}[D_t(s) \mid \Delta N_t = 1]\, \mathbb{E}[P_t(s) \mid \Delta N_t = 1].^2$$

This holds, for example, if $P_t(s)$ is deterministic for all t and s, or if the occupancy process is modeled as independent of the consequence $D_t(s)$ conditional on an event at time t. Utilizing this unstated assumption of conditional independence, it is possible to rewrite formula (2.14) as an equation that utilizes individual risk:

$$EN = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \int_{\mathbf{S}} IR_t(s) \mathbb{E}\left[P_t(s) \mid \Delta N_t = 1\right] ds\, dt$$

Formula (2.14) provides a long-term (time-averaged) estimate of expected fatalities and is very similar to the expression stated by Jonkman et al. (2003) on p. 7. However, the temporal dimension remains unaddressed in the formula given by Jonkman et al. (2003), even though individual risk may vary substantially over time due to factors such as weather patterns, operational states, and

time varying occupancy cycles. In contrast, Formula (2.14) makes the time aspect explicit by integrating over t and weighting by the event rate λ_t, and it retains $P_t(s)$ as a stochastic process rather than a fixed field, thereby capturing randomness in population density in addition to its regular temporal structure. This constitutes one reason why the definitions presented in this thesis differ from traditional ones: they allow a time-varying element of individual risk and explicitly incorporate the stochasticity of population density while containing the classical formulation as a special case.

2.2.7 F-N curve

Since the F-N curve is a central tool in QRA, we formalize it by first defining a random variable that counts the number of deaths at time t:

$$\eta_t \;=\; \int_{\mathbf{S}} D_t(s)\, P_t(s) ds,$$

Combining this with an indicator function that determines whether $\eta_t \geq n$, then weighting by the intensity λ_t and averaging over time, yields the frequency with which n or more fatalities are exceeded.

$$f(n) = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \mathbb{E}\left[\mathbf{1}_{\{\eta_t \geq n\}} \mid \Delta N_t = 1\right] \lambda_t\, dt. \tag{2.15}$$

This is analogous to formula (2.6) because $\mathbf{1}_{\{\eta_t \geq n\}}$ indicates whether, at time t, the number of deaths in a subset of the sample space is at least n (i.e., $\mathbf{1}_{\{\eta_t \geq n\}} = 1$ if the failure event has occurred and the death count is $\geq n$). By integrating $\mathbf{1}_{\{\eta_t \geq n\}}$ over $[t_0, t_1]$ and normalizing by the interval length, the corresponding frequency is obtained.

However, the F–N curve formula described by Jonkman et al. (2003) on page 7 addresses a different question: it considers the probability that the total loss in a year exceeds n, often under an implicit "rare-event" regime (at most one event per year). This is not the same as asking about the frequency of such exceedances within the year. In the notation used here, and emphasizing that a rigorous definition requires tools from stochastic calculus, the corresponding probability can be written informally as

$$\mathbb{P}\left(\int_{t_0}^{t_1} \int_{\mathbf{S}} D_t(s)\, P_t(s)\, ds\, dN_t \;\geq\; n\right).$$

Here, the stochastic integral with respect to dN_t should be understood (informally) as a random sum over event times $T_j \in (t_0, t_1]$: $\sum_j \int_{\mathbf{S}} D_{T_j}(s)\, P_{T_j}(s)\, ds$. This construction effectively sums, over all accident times within the window, the deaths attributable to each event and then evaluates, for the realized window, whether the total number of deaths exceeds n. A fully rigorous definition must also reflect that $D_t(s)$ is event-specific: for an accident starting at time t, $D_t(s)$ is determined by the time evolution of the accident on its own timeline, beginning at t and proceeding over the exposure window $[t,\, t + t_{\text{duration}}]$.

As noted in the paragraph above, it is easy to confuse *frequency* with *probability* in risk analysis. It is therefore important to understand that the frequency $f(n)$ is the average number per unit time of events with at least n fatalities, whereas $\mathbb{P}(\xi \geq n)$ is the probability that the total number of fatalities ξ in a specified period $[t_0, t_1]$ is at least n.

Another common assumption, which often fails to differentiate between probabilities and frequencies, is that the area under an F-N curve equals the time averaged expected number of fatalities. Because many risk measures do not explicitly account for time, it is not immediately obvious whether this property carries over when the F-N curve represents a frequency function rather than a probability distribution. Vrijling & van Gelder (1997) established this result for F-N curves interpreted probabilistically, that is, as distribution functions; invoking formula (2.15), the same conclusion follows for the frequency based definition used here, namely that the expected annual number of fatalities equals the area under the curve, as demonstrated by the short derivation that follows.

$$\int_0^\infty f(n)dn = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{\eta_t \geq n\}}\, dn \mid \Delta N_t = 1\right] \lambda_t\, dt.$$

and since

$$\int_0^\infty \mathbf{1}_{\{\eta_t \geq n\}}\, dn = \eta_t = \int_{\mathbf{S}} D_t(s)\, P_t(s)ds$$

which gives that

$$\int_0^\infty f(n)dn = \int_{t_0}^{t_1} \int_{\mathbf{S}} \mathbb{E}\left[D_t(s)\, P_t(s) \mid \Delta N_t = 1\right] \lambda_t\, ds\, dt.$$

the final result is that

$$\int_0^\infty f(n)dn = EN$$

Which means that assumption works for frequencies as well.

2.2.8 Model Overview

The bow-tie structure developed above separates *causes* from *effects* in a way that is convenient for both stochastic modeling and consequence analysis. On the *cause* side, the event process is the counting process N_t generated by the history of $\mathbf{Y}_t$, modeled as a inhomogeneous Poisson process (NHPP) with ROCOF λ_t (Def. 3). On the *effect* side, the consequence is the conditional lethality field $D_t(s)$ computed from the dispersion and toxicology mappings (Eqs. (2.7)–(2.10)). These components are coupled through the *instantaneous individual risk*

$$\mathrm{IR}_t(s) = \mathbb{E}[D_t(s) \mid \Delta N_t = 1]\, \lambda_t,$$

see (2.12), which has units of time^{-1}. Averaging over time yields the *time–averaged individual risk $IR(s)$* via (2.13). Spatial aggregation with the population density field produces the *time–averaged expected number of fatalities EN* in (2.14), and exceedance frequencies for the F–N curve are defined by (2.15).

24

In the time-constant special case, identifying $\lambda_t = f_i$ and $\mathbb{E}[D_t(s) \mid \Delta N_t = 1] = p_{f,i}(x, y)$ recovers the classical scenario formulas in (2.2). Thus, the present framework generalizes the original equations while including them as a special case.

Chapter 3

Case Study and Benchmarking

This chapter applies the probabilistic framework developed in section 2.2 to a concrete case study and benchmarks its outputs against a classical scenario-based analysis in order to assess the practical implications of the framework and its modularity. A Monte Carlo implementation is used to explicitly capture temporal variability, generating time-resolved realizations of the *same* dispersion model under identical meteorological and source inputs. By holding the physics and data fixed, the comparison isolates the effect of the modeling approach itself, enabling a like-for-like assessment of the framework versus the scenario method.

3.1   Description of Case Study

3.1.1   Ammonia (NH_3) as the Case Study Substance

Ammonia is a widely used industrial chemical, employed as a coolant and used as a raw material in plastic production. At ambient conditions, ammonia is a colorless irritant with high water solubility (MSB, n.d.). Although gaseous ammonia is lighter than air, releases from storage can be transiently cold and partially aerosolized in the near field before warming. In this study, the dispersion is modeled directly as passive from the onset, and any near source dilution is incorporated into the effective source term of the puff model.

Release-rate time histories are informed by accident reconstruction for an ammonia leak following Daly et al. (2013), which provides a physically grounded template for constructing time-varying source terms. The concentration fields produced by the passive Gaussian puff model are then mapped to dose and lethality using a probit formulation with toxicity parameters taken from Stellan et al. (1995). In this formulation the total dose depends on the concentration–time history at the receptor, and the probability of fatality is obtained by

applying the probit function calibrated for ammonia.

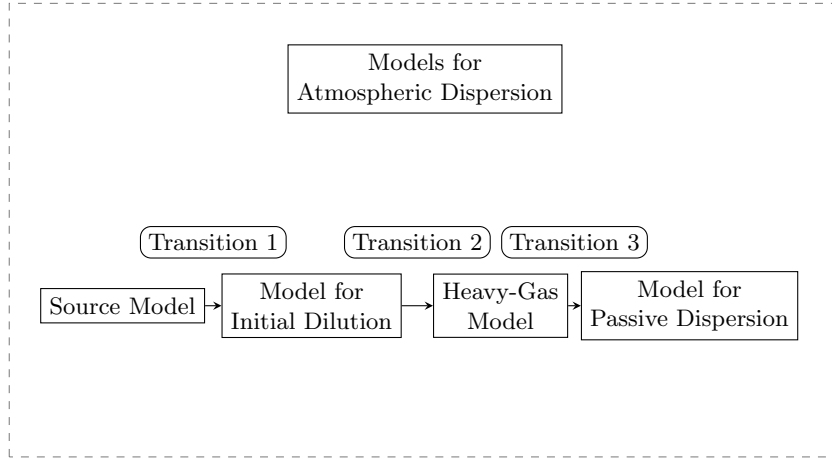### 3.1.2 Dispersion model C: passive Gaussian puff via Chama

Figure 3.1: A flat chain of model types with transitional conditions based on the chain of models in *Vådautsläpp av brandfarliga och giftiga gaser och vätskor: Metoder för bedömning av risker* (Stellan et al., 1995, p. 197).

The function $C(u, s, \tau)$ in (2.7) is implemented with the CHAMA Gaussian puff model (Klise et al., 2017). In this representation a release is described as a sequence of puffs that are transported by the local wind and spread by turbulent diffusion. The input vector u collects all driving conditions: meteorology (wind speed, wind direction, and a stability indicator) and source descriptors (time-dependent emission rate, source location, and release height). Superposition of puffs yields the concentration field, so that $C(u, s, \tau)$ returns the concentration at receptor s and time τ based on the input u.

Passive dispersion is applied directly from the source. Any very near source dilution or jet effects are absorbed into an effective source term, and the puff formulation is used without a separate heavy gas stage. This choice gives a simple yet transparent consequence model.

3.1.3 Release process for $\mathbf{X}_t$

The effect-side state $\mathbf{X}_t$ consists of the time-varying release rate together with the meteorological inputs required by C. Release rates are constructed by adapting the parameterization of the release-rate curve in Daly et al. (2013) to keep the computation reasonable and realistic. The formulation permits an elevated initial discharge that transitions to a rate that decays exponentially; the initial variability captures changes in orifice size and other factors that may produce a high early release; representative profiles are shown in fig. 3.2.

Leak rate trajectories are modeled as a two phase profile consisting of a constant plateau(A_0) followed by exponential decay of the leakage. Let $t_0 > 0$ denote the plateau duration and $\tau > 0$ the decay time constant. For a given realization the instantaneous rate $r(t; A_0, A_1)$ is

$$
r(t; A_0, A_1) = \begin{cases} A_0, & 0 \leq t \leq t_0, \\ A_1 \exp\big(-(t - t_0)/\tau\big), & t > t_0, \end{cases}
$$

with units s^{-1} interpreted as the fraction of inventory released per second. The decay amplitude A_1 is fixed by a mass balance over an infinite horizon,

$$
\int_0^\infty r(t; A_0, A_1)\, dt \;=\; A_0 t_0 + A_1 \tau \;=\; 1, \qquad \Longrightarrow \qquad A_1 \;=\; \frac{1 - A_0 t_0}{\tau} \;\geq 0.
$$

To ensure $A_1 \geq 0$ almost surely, the plateau level A_0 is drawn from a triangular distribution supported on $[\ell, m, h]$ with mode $m = \widehat{A}_0$, lower bound $\ell = \frac{1}{2}\widehat{A}_0$, and upper bound $h = (1 - \varepsilon)/t_0$ for a small numerical $\varepsilon \in (0, 1)$. Denoting $A_0 \sim \mathrm{Tri}(\ell, m, h)$, the density is

$$
f_{A_0}(a) = \begin{cases} \dfrac{2(a - \ell)}{(h - \ell)(m - \ell)}, & \frac{1}{2}\widehat{A}_0 \leq a \leq m, \\ \dfrac{2}{h - \ell}, & a = m, \\ \dfrac{2(h - a)}{(h - \ell)(h - m)}, & m < a \leq (1 - \varepsilon)/t_0, \\ 0, & \text{otherwise,} \end{cases}
$$

which implies $A_0 t_0 \leq 1 - \varepsilon$ and hence $A_1 = (1 - A_0 t_0)/\tau$ is nonnegative by construction.

Discretization follows the simulation grid used by the dispersion model. Let $\Delta t > 0$ and $T > 0$ be the time step and profile duration, and define $t_k = k\,\Delta t$ for $k = 0, 1, \ldots, N$ with $N = \lfloor T/\Delta t \rfloor$. The discrete profile associated with a realization is the vector

$$
R = \big(r(t_0; A_0, A_1),\, r(t_1; A_0, A_1),\, \ldots,\, r(t_N; A_0, A_1)\big) \in \mathbb{R}_+^{N+1}.
$$

Figure 3.2 shows representative Monte-Carlo realizations of the leak-rate profiles produced by the model.
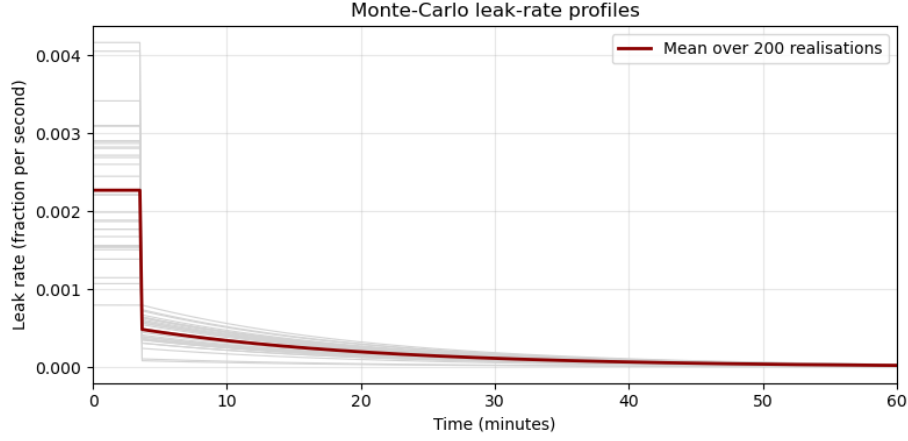
Figure 3.2: Monte-Carlo leak-rate profiles generated with the python function `generate_mc_leak_rates`; 30 individual realisations (grey) and their ensemble mean (red) are shown.

3.1.4 Meteorology for X_t

Hourly meteorological data were taken from the ERA5 (Hersbach et al., 2023) for the entire year 2024 at the grid point nearest to Malmö, Sweden. From the 10 m wind components u_{10} and v_{10} the wind speed and wind direction were derived as

$$\text{speed} = \sqrt{u_{10}^2 + v_{10}^2}, \qquad \text{direction} = \big(270 - \text{atan2}(v_{10}, u_{10}) \cdot 180/\pi\big) \bmod 360,$$

and the total cloud cover tcc with windspeed was used in the stability assignment.

Wind speed 10 m height (m/s)	Day – Solar radiation			Night – Cloudiness	
	Strong; solar elevation $> 60°$	Moderate $35°$–$60°$	Weak $< 35°$	Thin clouds or $> 4/8$ low	$\leq 3/8$
< 2	A	A–B	B	F*	F*
2–3	A–B	B	C	E	F
3–4	B	B–C	C	D	E
4–6	C	C–D	D	D	D
> 6	C	D	D	D	D

Table 3.1: Classification table as suggested by Stellan et al. (1995) for atmospheric stability classes.

The atmospheric stability class for each hour was obtained with a Pasquill Gifford scheme following table 3.1. Daylight and darkness were determined from the solar elevation angle computed at the Malmö coordinates for each timestamp. During daylight the insolation category was inferred from solar elevation, with thresholds at $60°$ and $35°$ to represent strong, moderate, and weak insolation. The stability class was then assigned from the usual function

of insolation and 10 m wind speed to classes A through D. During darkness the classification used the cloud cover to distinguish clear from cloudy conditions using a threshold of $tcc = 0.5$, together with wind speed, to assign classes D through F according to table 3.1. The resulting hourly series of stability classes is aligned with the ERA5 time index and is used by the passive puff model as a discrete proxy for atmospheric stability in the dispersion calculation.

The following figures visualize samples of the weather data from Hersbach et al. (2023) that were used in the case study.
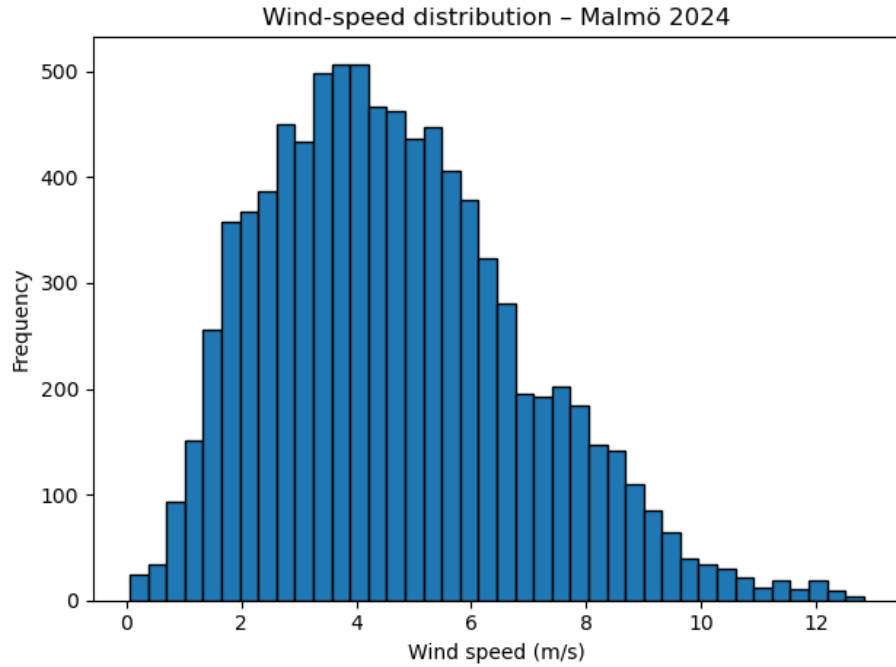

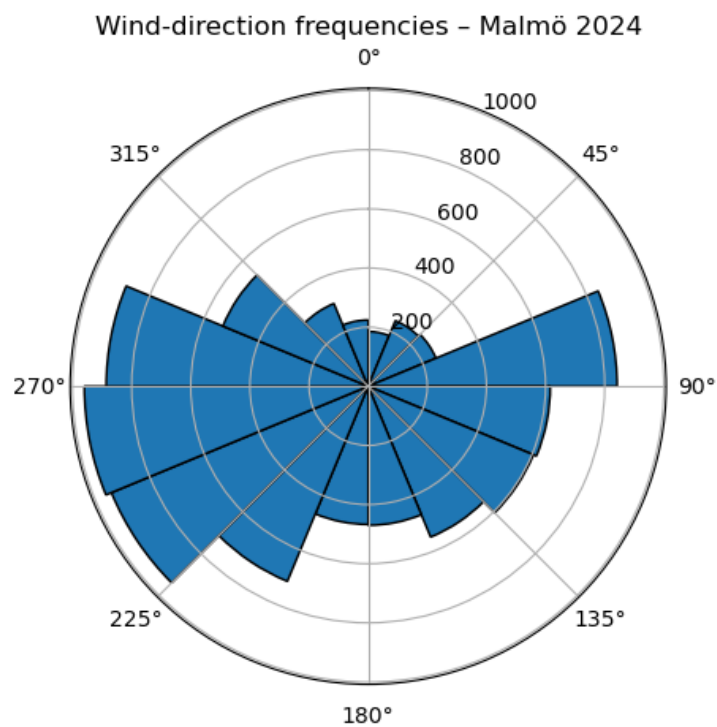Figure 3.3: Wind-speed distribution for Malmö 2024.

Figure 3.4: Wind-direction frequencies (22.5° sectors) for Malmö 2024.
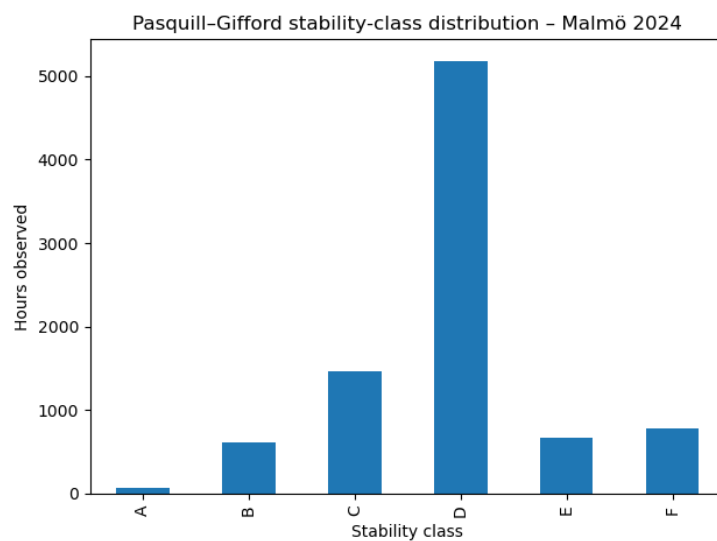

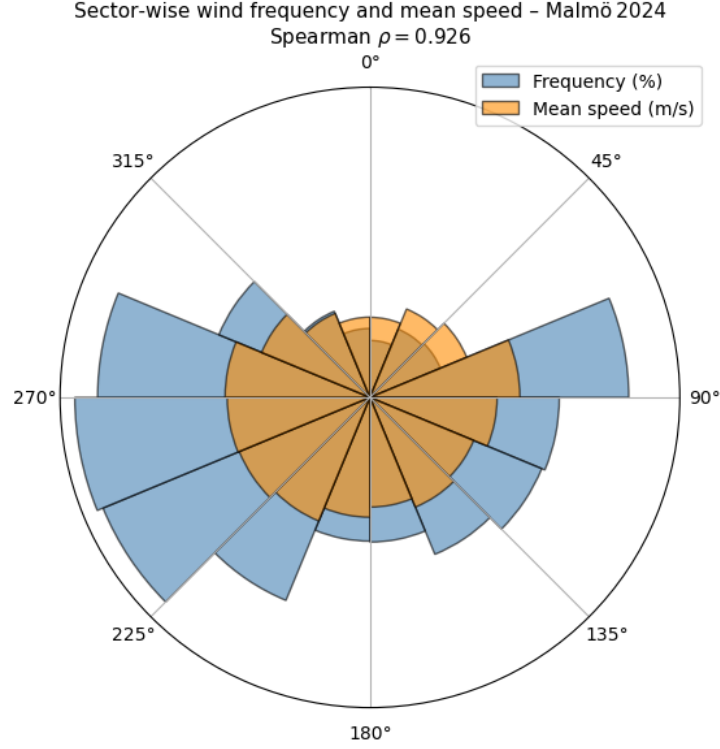Figure 3.5: Pasquill–Gifford stability-class counts for Malmö 2024.

Figure 3.6: Sector-wise comparison of wind-direction frequency (%) and mean wind speed (m/s) for Malmö 2024, with Spearman correlation.

3.1.5 Toxicological function M and Lethality L: Probit Formulation

Given a concentration history $\tau \mapsto C(\mathbf{X}_\tau, s, \tau)$, the toxicological function M in (2.8) converts instantaneous concentration to incremental dose. The dose is then described via the integral:

$$\text{Dose}_t(s) \;=\; \int_t^{t+t_{\text{duration}}} \big[\, C(\mathbf{X}_\tau, s, \tau)\,\big]^2 \, d\tau,$$

and total dose is mapped to fatality probability using a *probit* function,

$$D_t(s) \;=\; \Phi\big(\alpha + \beta \log\big(\text{Dose}_t(s)\big)\big),$$

where Φ is the standard normal CDF and $\alpha = -9.4, \beta_2 = 0.75$ are substance-specific coefficients taken from Stellan et al. (1995). This yields the theoretical calculations of the probability of death at time t and position s.

3.1.6 Exposure Window and Numerical Dose Accumulation

For a leak initiating at time t and lasting $t_{\text{duration}} = 1$ *hour*, the discretized dose at receptor s is computed on a grid $t = t_0 < t_1 < \cdots < t_N = t + t_{\text{duration}}$ by the trapezoidal rule,

$$\text{Dose}_t(s) \approx \sum_{i=0}^{N-1} \left(\frac{C(\mathbf{X}_{t_i}, s, t_i) + C(\mathbf{X}_{t_{i+1}}, s, t_{i+1})}{2} \right)^n \Delta t.$$

The corresponding lethality $D_t(s) = L(\text{Dose}_t(s))$ feeds the effect node in the bow-tie (Fig. 2.8), conditioned on a top-event occurrence ($\Delta N_t = 1$).

3.1.7 Instantaneous and Time-Averaged Individual Risk $\mathbf{Y}_t$

Coupling to the cause side proceeds via the NHPP intensity λ_t. The *instantaneous expected individual fatality probability* at s given that a top event occurs a t is $\mathbb{E}[D_t(s) \mid \Delta N_t = 1]$. The *instantaneous fatality rate* is then

$$\text{IR}_t(s) = \mathbb{E}\big[D_t(s) \mid \Delta N_t = 1\big] \lambda_t,$$

and the *time-averaged* individual risk over $[t_0, t_1]$ is

$$IR(s) = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \text{IR}_t(s) \, dt.$$

Operationally, we approximate expectations by Monte Carlo simulation over release-rate and meteorological data, while λ_t controls the temporal weighting of event occurrence. The intensity was taken constant, $\lambda_t \equiv \lambda = 10^{-6} \, \text{yr}^{-1}$ since there was little motivation to handle the more complex case of a time-varying frequency in this simplified comparison. For a one hour interval within 2024, which contains 8784 hours, the Poisson mean on that hour is

$$\mu = \lambda \frac{1 \text{ hour}}{1 \text{ year}} = \frac{10^{-6}}{8784} \approx 1.138 \times 10^{-10}.$$

The hourly top–event count is therefore Poisson(μ). The probabilities for exactly one, and at least two events in a given hour are

$$\mathbb{P}(\Delta N_t = 1) = e^{-\mu} \approx 1.138434 \times 10^{-10}$$

$$\mathbb{P}(\Delta N_t \geq 2) = 1 - e^{-\mu}(1 + \mu) \approx 6.480154 \times 10^{-21}$$

Hence, the probability of more than one event in a single hour is of order 10^{-20}. It is approximately 5.7×10^{-11} times the probability of exactly one event in that hour, which justifies treating multiple events within an hour as negligible.

3.2 Classical scenario model

For the classical scenario model, the recommended weather scenarios in Ingvarson et al. (2022) were used as the meteorological baseline. The wind rose was partitioned into twelve direction sectors of width $30°$ with centers at $0°, 30°, \dots, 330°$. For each sector, three wind speed classes were used $\leq 3 \mathrm{~m\,s^{-1}}$, 3–$7 \mathrm{~m\,s^{-1}}$, and $> 7 \mathrm{~m\,s^{-1}}$. Each speed class was split into day and night. This construction yields

$$12 \times 3 \times 2 = 72$$

distinct meteorological scenarios. For each scenario, the dispersion calculation uses the sector–center wind direction together with the scenario's wind–speed class, and the Pasquill–Gifford stability class computed for its day/night category. The following classification based on recommendations was therefore preformed. For daytime which had a probability of 0.44, three representative scenarios were used: $2 \mathrm{~m\,s^{-1}}$ with stability class B (i.e., $\leq 3 \mathrm{~m\,s^{-1}}$); $5 \mathrm{~m\,s^{-1}}$ with class C (i.e., $(3, 7] \mathrm{~m\,s^{-1}}$); and $8 \mathrm{~m\,s^{-1}}$ with class D (i.e., $> 7 \mathrm{~m\,s^{-1}}$). For nighttime, three representative scenarios were used: $2 \mathrm{~m\,s^{-1}}$ with stability class F (i.e., $\leq 3 \mathrm{~m\,s^{-1}}$); $5 \mathrm{~m\,s^{-1}}$ with class D (i.e., $(3, 7] \mathrm{~m\,s^{-1}}$); and $8 \mathrm{~m\,s^{-1}}$ with class D (i.e., $> 7 \mathrm{~m\,s^{-1}}$).

To couple meteorology with the source term, meteorological scenarios were crossed with release rate realizations. Each scenario was also evaluated for A_0, $A_0/2$, and $1/t_0$ using the same leak generating function with the prior on A_0 suppressed. For each meteorological row, three independent release rate time histories were generated, yielding three realizations per row.

$$72 \times 3 = 216$$

runs. The probability mass in table B.1 was preserved by assigning to each run a weight equal to one third of the corresponding row probability, assuming equal probability for all emission values.

The concentration and dose for each run were then computed with the passive puff model using the meteorological representatives and the selected release rate history.

Two handbook simplifications were not adopted (Ingvarson et al., 2022). The fixed geometric release angle of $30°$ was not imposed and the 50% lethality distance was not interpreted as a region of 100% lethality. The simulated concentration fields did not support those assumptions and the intention was to let the dispersion–toxicity function supply lethality directly rather than by a fixed radius rule.

3.3 Implementation details

The concentration function C is evaluated for 8784 realizations using CHAMA with the modifications, and the dose–probit pipeline is implemented in `Python` (Jönsson, 2025). All 8,784 hours of 2024 were simulated: for each hour, the

corresponding weather inputs were used and a single release-rate realization was drawn from the model in section 3.1.3.

3.3.1   Model complexity and usability

Weather data were required for both models; however, additional preprocessing was necessary for the scenario-based method, because the data had to be partitioned. In contrast, the Monte Carlo method used the samples from the processed ERA5 dataset without further modification. The simulation of Monte Carlo samples, when performed using multiprocessing on fourteen cores, required approximately 70 to 80 hours, whereas single-threaded processing for the scenario-based method required approximately 9 to 10 hours. To rerun the analysis for the Monte Carlo–based model at a different geographic location, the only required adjustment is to change the coordinates used to obtain ERA5 weather data.

3.4   Result

This section presents individual risk contours for both modeling approaches using the meteorological dataset described in section 3.1.4. In addition to the contours, two comparative diagnostics are shown to quantify pointwise model disagreement on the map. The first is a signed relative difference that indicates where the Monte Carlo estimate exceeds or falls below the classical scenario estimate. The second is a logarithmic risk ratio that reports the magnitude of the discrepancy on a multiplicative scale. Together, these figures serve as a visual tool for understanding where and by how much the modeling approaches differ.
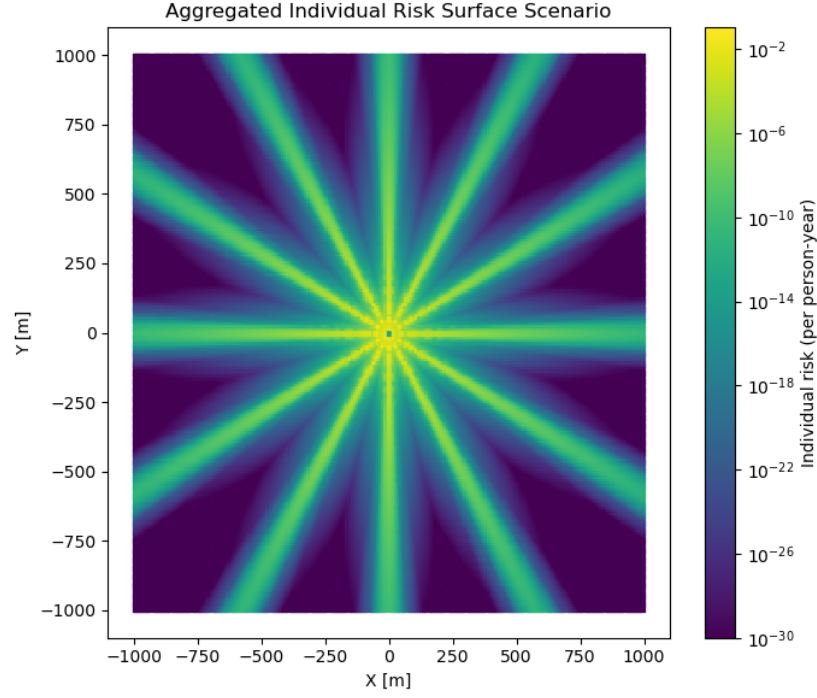
(a) $v_{\min} = 10^{-30}$
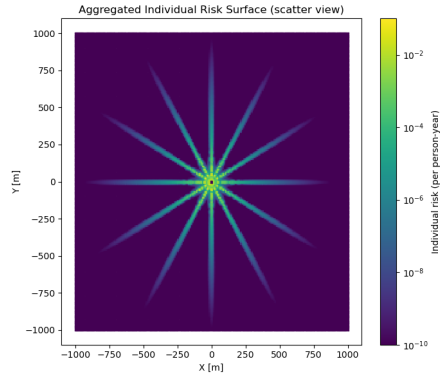
(b) $v_{\min} = 10^{-10}$
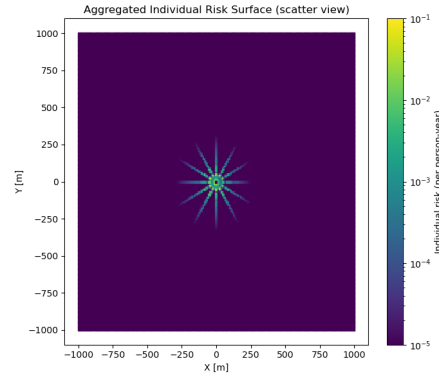
(c) $v_{\min} = 10^{-5}$

Figure 3.7: Aggregated individual-risk surfaces for the 216 deterministic scenarios under three different lower-cut-off values denoted $v_{\min}$.

Even with twelve distinct wind directions employed, comprehensive coverage of all possible orientations was not achieved. Accordingly, the scenario-based approach requires widening each directional sector to avoid omitting potential wind orientations.
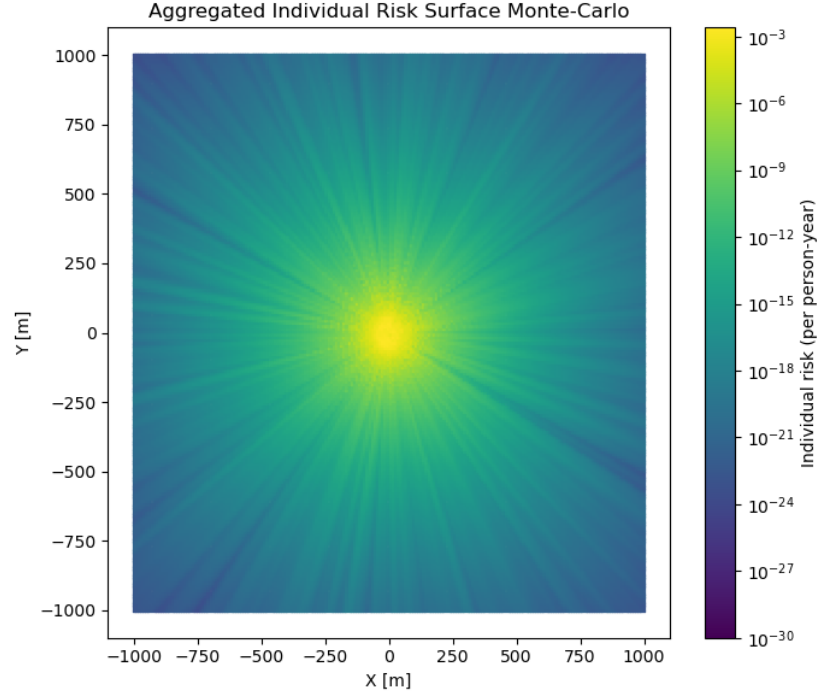
(a) $v_{\min} = 10^{-30}$
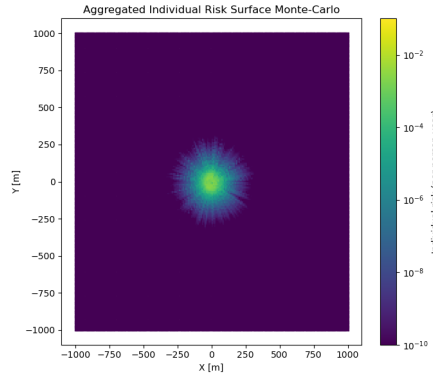
(b) $v_{\min} = 10^{-10}$
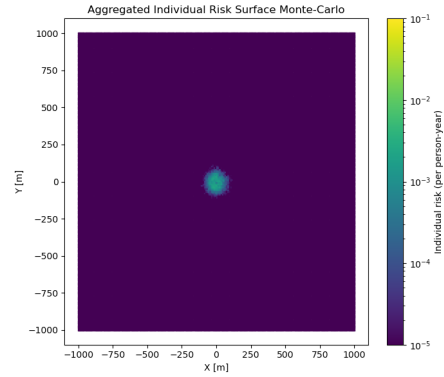
(c) $v_{\min} = 10^{-5}$

Figure 3.8: Aggregated individual-risk surfaces produced by the Monte Carlo simulation (same lower-cut-off values as Figure 3.7). The graph is produced from the simulation which is described in section 3.1.

The samples appear to be satisfactorily distributed, even without adjustments and despite the relatively small sample size. The graph's unconventional format, however, makes interpretation more challenging than in traditional rep-

resentations. It may be expected to observe a clear bias in the individual-risk estimates toward the more frequent wind directions. The lack of a clear directional bias may be partly due to the strong positive correlation between sector frequency and mean wind speed shown in fig. 3.6, leading to broader dispersion and longer transport distances, but with diminished concentrations.
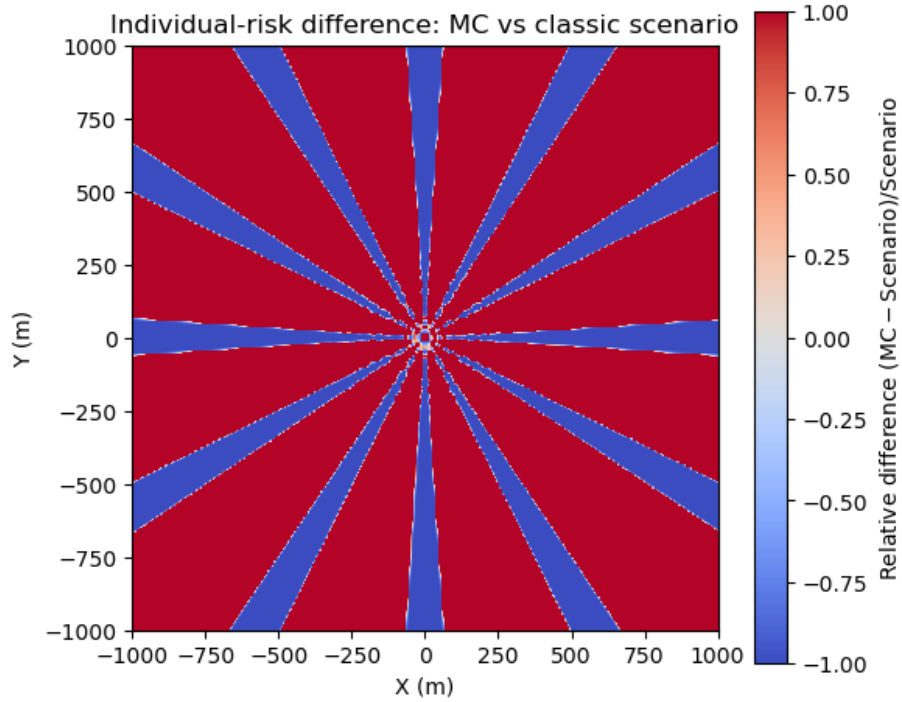
Figure 3.9: Relative individual-risk difference between the Monte-Carlo method and the classical scenario method.
Red = MC higher; blue = MC lower; white = equal.
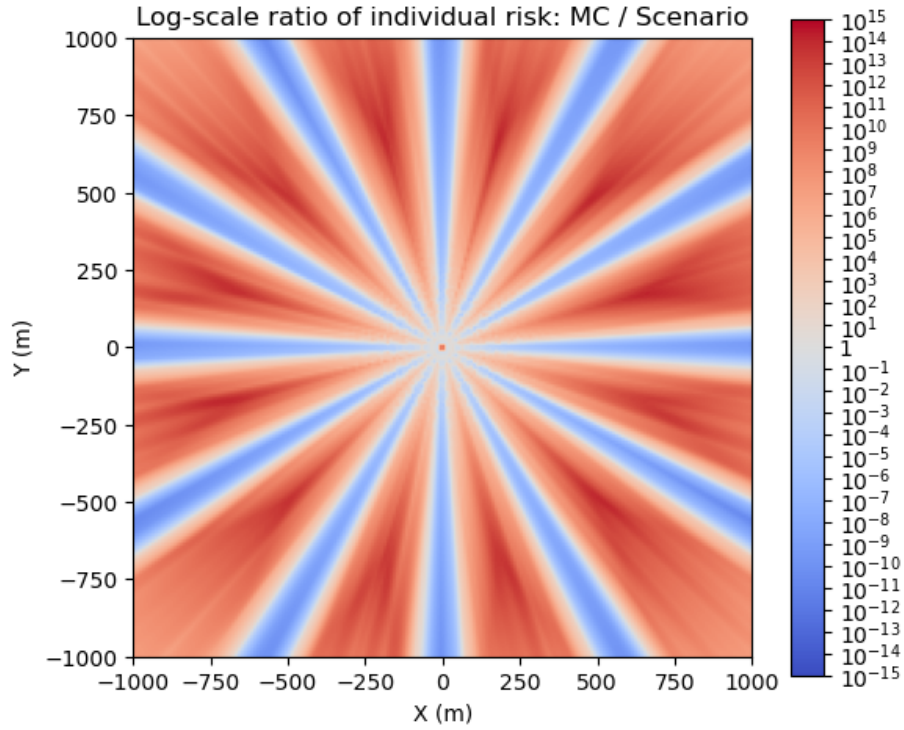
Figure 3.10: Logarithmic ratio of individual risk computed with the Monte-Carlo method to that obtained with the classical scenario method. Blues indicate sectors where the MC estimate is *lower* than the scenario result; reds where it is *higher*. Tick labels on the colour bar are in powers of ten (e.g. $10^{+2} = 100 \times$ higher, $10^{-1} = 10 \times$ lower).

A substantial discrepancy between the two methods is evident in the graphs, as the scenario-based method covers only the specified directions. In the logarithmic-scale version, the differences are particularly large, with the graphs capped at 10^{15} for both the lower and upper bounds; this cap appears to be exceeded multiple times.

Chapter 4

Discussion and Conclusion

4.1 Discussion

The results indicate that Monte Carlo simulation is a viable alternative to traditional scenario methods, particularly when the objective is to produce more realistic outcomes. Scenario methods can perform well for assessing specific event types but generally cannot match the accuracy of risk estimates obtained from Monte Carlo simulation. Given the identical dataset and the significantly reduced preprocessing requirements of Monte Carlo simulation, the scenario method retains only one clear advantage: shorter overall simulation time. However, the active effort required from the risk assessor is considerably greater for scenario methods, as these simulations cannot proceed unattended; they require smearing results to a 30° release radius and assuming a 100% fatality rate within the region defined by the 50% fatality distance. While these steps could be automated, doing so would require processes similar to those used in Monte Carlo simulation and would still yield less precise results. Consequently, although the scenario method may conserve computational time, extended background simulations may be acceptable, particularly given that computational resources are now less constrained than in the past.

The formalized framework supports more accurate error estimation by applying appropriate statistical methods. Although the present comparison was limited to a single risk measure in a fictitious scenario, extending Monte Carlo simulation to more realistic cases would enable the use of advanced techniques common in other risk management disciplines, such as variance reduction methods.

4.2   Research questions

1. Can the classical risk measures used in quantitative risk analysis be placed on a rigorous probability-theoretic foundation?

The traditional formulas for societal and individual risk were incorporated into a formal framework, which revealed new insights to how concepts use in risk science relates to probability theory. Certain components of these formulas could benefit from additional formalization, particularly the treatment of event counting, which requires more advanced mathematical methods. Nevertheless, this study provides a solid foundation for the further development of quantitative risk analysis and a deeper understanding of risk. Concretely, each reformulation of the equations using random variables introduces a better understanding of which inputs could be treated as random, thereby creating the possibility of handling uncertainty. The equations also make clear the difference between probabilities and frequencies, which are not easily distinguished without the use of counting processes.

2.   Does a Monte Carlo simulation that directly employs time-resolved empirical weather data provide more reliable estimates of individual and societal risk than the traditional scenario-based method?

Figure 3.10 clearly shows that the risk estimates differ significantly, indicating that additional scenarios or a correction method would be required to improve accuracy. Because the traditional approximation method of distributing the release over a 30° arc and assuming a 100% fatality rate within the 50% fatality radius could not be applied, it cannot be concluded that this method is invalid. However, the substantial overestimation of individual risk within the primary plume direction, and underestimation in the other directions, observed in fig. 3.10 suggests that the use of scenario methods leads to a greatly inflated estimate. This effect may occur because the leak plume covers an angle much smaller than 30°, resulting in a low probability that the leak would reach that location.

3.   Is the model sufficiently modular to be reused across facilities via simple data exchange, thereby reducing project lead time?

The Monte Carlo method demonstrates strong modularity with the chosen dispersion model, as only the coordinates for weather data retrieval and the leak size need to be adjusted. To increase its scope, the selected model could incorporate a leak characterization submodel based on orifice geometry and tank mass, followed by a heavy gas dispersion model before transition to the passive dispersion model; fig. 3.1 illustrates one possible sequence of models. Compared with the traditional scenario method, this approach offers greater modularity, since once a model for a given hazard has been developed, the only required change is to provide local input data. Traditional dispersion models can also be automated, provided they support efficient data ingestion and the concurrent

execution of multiple instances.

4.3   Further research

To enhance the utility of this framework and improve quantitative risk analysis outcomes, dispersion models should be refined to achieve greater accuracy and computational efficiency. Methods for computational speedup, such as the parallelization implemented in this work, efficient coding practices like vectorized Python, and other execution time optimizations could further accelerate simulations. This could enable a substantially larger number of runs when sufficient computational resources are available. With adequate resources, high-fidelity models such as full computational fluid dynamics (CFD) simulations may be employed, potentially by exploiting the capabilities of GPUs. Another promising approach is the use of physics-informed neural networks (PINNs), as demonstrated by Lopes et al. (2025). PINNs are neural networks that, in addition to learning from generated data, embed the governing physical principles and can produce accurate predictions within seconds. Because neural networks often execute rapidly on GPUs, they offer a compelling opportunity to combine advances in artificial intelligence with modern GPU performance.

In addition, a detailed comparison between Monte Carlo based quantitative risk analyses and scenario based methods should be undertaken to evaluate potential biases in each approach. Such an analysis is important, as underestimation of risk may lead to loss of life due to insufficient protective measures, whereas overestimation may hinder the approval of developments with substantial societal cost. The theoretical framework could also be strengthened through further formalization using stochastic calculus. Such refinement would make it possible to determine bounds, credible intervals, confidence intervals, and other essential metrics in Monte Carlo simulations.

Bibliography

Aven, T. (2020). Three influential risk foundation papers from the 80s and 90s: Are they still state-of-the-art? *Reliability Engineering System Safety*, *193*, 106680. doi: https://doi.org/10.1016/j.ress.2019.106680

Björk, T. (2021). *Point processes and jump diffusions: An introduction with finance applications.* Cambridge University Press.

CCPS, E. I. (2018). *The bow tie model.* John Wiley Sons, Ltd. doi: https://doi.org/10.1002/9781119490357

Center for Chemical Process Safety (CCPS). (n.d.). *Process safety.* Retrieved 2025-12-14, from `https://www.aiche.org/ccps/resources/glossary/process-safety-glossary/process-safety` (AIChE CCPS Process Safety Glossary)

Daly, A., Zannetti, P., & Jennings, M. (2013). Accident reconstruction and plume modeling. In *Air pollution xxi* (Vol. 174). WIT Press. doi: https://doi.org/10.2495/AIR130011

Directive 2012/18/eu. on the control of major-accident hazards involving dangerous substances. (2012, July 4). Official Journal of the European Union, L 197/1–37. Retrieved from `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012L0018`

Gaylor, D. W. (2000). The use of haber's law in standard setting and risk assessment. *Toxicology*, *149*(1), 17-19. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0300483X00002286` doi: https://doi.org/10.1016/S0300-483X(00)00228-6

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., . . . Thépaut, J.-N. (2023). *Era5 hourly data on single levels from 1940 to present.* Copernicus Climate Change Service (C3S) Climate Data Store. (Accessed: 2025-07-17) doi: https://doi.org/10.24381/cds.adbb2d47

Høyland, A., & Rausand, M. (1994). *System reliability theory: Models and statistical methods.* New York: Wiley-Interscience.

Ingvarson, J., Egeberg, T., Nystedt, F., & Nassiri, S. (2022). *Qra handledning: Del 2*. Intressentföreningen för processäkerhet (IPS).

Jonkman, S. N., van Gelder, P. H. A. J. M., & Vrijling, J. K. (2003). An overview of quantitative risk measures for loss of life and economic damage. *Journal of Hazardous Materials*, *99*(A), 1–30. doi: https://doi.org/10.1016/S0304-3894(02)00283-2

Jönsson, j. (2025). *Formalizing quantitative risk analysis: Probabilistic foundations and monte carlo methods.* GitHub. Retrieved 2025-08-25, from `https://github.com/joensson12/Formalizing-Quantitative-Risk-Analysis-Probabilistic-Foundations-and-Monte-Carlo-Methods` (GitHub repository; commit 376e02b)

Klise, K. A., Nicholson, B., & Laird, C. D. (2017). *Sensor placement optimization using chama* (Sandia Report No. SAND2017-11472). Sandia National Laboratories.

Lopes, G. M., da Silva, F. V., & Vianna, S. S. V. (2025). A physics-informed data-driven model applied for gas dispersion. *Journal of Loss Prevention in the Process Industries*, *97*, 105703. doi: https://doi.org/10.1016/j.jlp.2025.105703

Marx, J. D., & Cornwell, J. B. (2009). The importance of weather variations in a quantitative risk analysis. *Journal of Loss Prevention in the Process Industries*, *22*, 803–808. doi: https://doi.org/10.1016/j.jlp.2009.08.009

MSB, M. (n.d.). *Ammoniak, vattenfri.* Retrieved 2025-08-25, from `https://rib.msb.se/fa/Substance/Index?id=448` (MSB RIB – Farliga ämnen. In Swedish)

MSB, M. (2017, March). *Riskutredning för mindre och medelstora verksamheter: Vägledning* (Tech. Rep. No. MSB1060). Myndigheten för samhällsskydd och beredskap. (Produktion: Advant Produktionsbyrå)

Nassiri, S., Bergstrand, U., & Lindblom, V. (2021). *Qra handledning: Del 1.* Intressentföreningen för processäkerhet (IPS).

Nikeghbali, A. (2006, January). An essay on the general theory of stochastic processes. *Probability Surveys*, *3*(none). Retrieved from `http://dx.doi.org/10.1214/154957806000000104` doi: 10.1214/154957806000000104

Raveendran, A., Renjith, V., & Madhu, G. (2022). A comprehensive review on dynamic risk analysis methodologies. *Journal of Loss Prevention in the Process Industries*, *76*, 104734. doi: https://doi.org/10.1016/j.jlp.2022.104734

Shiryaev, A. N. (2016). *Probability-1* (3rd ed., Vol. 95). New York: Springer.

Stellan, F., Forsén, R., & Hertzberg, O. (1995, March). *Vådautsläpp av brandfarliga och giftiga gaser och vätskor: Metoder för bedömning av risker* (Tech. Rep.). Försvarets forskningsanstalt (FOA).

Vrijling, J. K., & van Gelder, P. H. A. J. M. (1997). Societal risk and the concept of risk aversion. In *Proceedings of esrel/psam.* Delft, The Netherlands: Department of Civil Engineering, Delft University of Technology.

Zio, E. (2013). *The monte carlo simulation method for system reliability and risk analysis* (1st ed.). London: Springer London. doi: 10.1007/978-1-4471-4588-2

Appendix A

Proof of instantaneous death rate

Assume that the intensity of λ_t is continuous and let $(U_i)_{i \in \mathbb{N}}$ denote the (a.s. strictly increasing) jump times of the counting process N inside $(t - \Delta t, t]$. For a release that begins at time u, define the *lethality mark*

$$\phi_u(s) \in \{0, 1\}, \qquad \text{with} \quad \mathbb{E}\big[\phi_{U_1}(s) \mid D_{U_1}(s), N_t - N_{t-\Delta t} = 1\big] = D_{U_1}(s) \in [0, 1],$$

so that $\phi_u(s) = 1$ indicates that the release starting at u ultimately causes a fatality at location s.

The instantaneous expected fatality rate at location s and time t is then defined by

$$\lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbb{E}\left[\sum_{i \,:\, U_i \in (t-\Delta, \, t]} \phi_{U_i}(s) \right] \tag{A.1}$$

It is then possible to rewrite the expectation via standard formulas.

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \sum_{k=0}^{\infty} \mathbb{E}\big[\sum_{i=1}^{k} \phi_{U_i}(s) \mid N_t - N_{t-\Delta t} = k \big] \mathbb{P}\big(N_t - N_{t-\Delta t} = k \big)$$

The subsequent step is to express $\mathbb{P}\big(\phi_{U_1}(s) = 1 \mid N_t - N_{t-\Delta t} = 1\big)$ in terms of $D_{U_1}(s)$, since this is the quantity of interest.

$$\mathbb{P}\big(\phi_{U_1}(s) = 1 \mid N_t - N_{t-\Delta t} = 1\big) = \mathbb{E}\big[\phi_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big]$$

The next step is to reformulate this expression by applying the law of conditioning:

$$\mathbb{E}\big[\phi_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big] =$$

$$\mathbb{E}\big[\mathbb{E}\big[\phi_{U_1}(s) \mid D_{U_1}(s), N_t - N_{t-\Delta t} = 1\big] \mid N_t - N_{t-\Delta t} = 1\big]$$

This is equivalent to expressing it as

$$\mathbb{E}\big[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big]$$

which leads to splitting up the limit into

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \mathbb{E}\big[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big] \mathbb{P}\big(N_t - N_{t-\Delta t} = 1\big).$$

and

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \sum_{k=2}^{\infty} \mathbb{E}\Big[\sum_{i=1}^{k} \phi_{U_i}(s) \mid N_t - N_{t-\Delta t} = k\Big] \mathbb{P}\big(N_t - N_{t-\Delta t} = k\big)$$

Utilizing the result from definition 3 leads to the conclusion that

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \sum_{k=2}^{\infty} \mathbb{E}\Big[\sum_{i=1}^{k} \phi_{U_i}(s) \mid N_t - N_{t-\Delta t} = k\Big] \mathbb{P}\big(N_t - N_{t-\Delta t} = k\big) = 0$$

since it is bounded by

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \sum_{k=2}^{\infty} \mathbb{P}\big(N_t - N_{t-\Delta t} = k\big)k$$

Next, since it is a NHPP with mean

$$\mu(\Delta t) = \int_{t-\Delta t}^{t} \lambda(u)\, du,$$

utilizing first that the sum can be written as

$$\sum_{k=2}^{\infty} \mathbb{P}(N_t - N_{t-\Delta t} = k)k = \mathbb{E}[N_t - N_{t-\Delta t}] - \mathbb{P}(N_t - N_{t-\Delta t} = 1)$$

which can be written as

$$\mu(\Delta t) - \mu(\Delta t)e^{-\mu(\Delta t)} = \mu(\Delta t)\big(1 - e^{-\mu(\Delta t)}\big).$$

Dividing by Δt gives

$$\frac{\mu(\Delta t)}{\Delta t}\big(1 - e^{-\mu(\Delta t)}\big).$$

Since

$$\lim_{\Delta t \downarrow 0} \frac{\mu(\Delta t)}{\Delta t} = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \int_{t-\Delta t}^{t} \lambda(u)\, du = \lambda(t) \quad \text{and} \quad \lim_{\Delta t \downarrow 0}\big(1 - e^{-\mu(\Delta t)}\big) = 0,$$

it follows that

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \sum_{k=2}^{\infty} \mathbb{P}(N_t - N_{t-\Delta t} = k)k = 0.$$

The only remaining limit to evaluate is

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \mathbb{E}\big[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big] \mathbb{P}\big(N_t - N_{t-\Delta t} = 1\big).$$

This results in the final definitional step for determining individual risk by utilizing definition 3.

$$\lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \mathbb{E}\big[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big](\lambda_t \Delta t + o(\Delta t)) = \lim_{\Delta t \downarrow 0} \mathbb{E}\big[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1\big]\lambda_t$$

Using the notation from definition 2, note that on the event $\{N_t - N_{t-\Delta t} = 1\}$ there is a unique jump time $U_1 \in (t - \Delta t, t]$, and $U_1 \to t$ as $\Delta t \downarrow 0$. If $\mathbb{E}[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1]$ is continuous at t, this yields

$$\lim_{\Delta t \downarrow 0} \mathbb{E}[D_{U_1}(s) \mid N_t - N_{t-\Delta t} = 1] = \mathbb{E}[D_t(s) \mid \Delta N_t = 1].$$

A fully rigorous interpretation can be framed using Palm calculus, which formally treats conditioning on an arrival at a given time.

Appendix B

Scenarios for classic QRA

We used the recommended weather scenario data as stated by Ingvarson et al. (2022), see section 3.2.

Table B.1: Weather scenarios for classical QRA Ingvarson et al. (2022).

Sector center (°)	Dir. range	Speed class	Day/Night	Weather code	Stab.	Probability
0	346–15°	≤3 m/s (låg)	Dag	2B	B	0.0032
0	346–15°	≤3 m/s (låg)	Natt	2F	F	0.0041
0	346–15°	3–7 m/s (medel)	Dag	5C	C	0.0076
0	346–15°	3–7 m/s (medel)	Natt	5D	D	0.0097
0	346–15°	>7 m/s (hög)	Dag	8D	D	0.0020
0	346–15°	>7 m/s (hög)	Natt	8D	D	0.0026
30	16–45°	≤3 m/s (låg)	Dag	2B	B	0.0038
30	16–45°	≤3 m/s (låg)	Natt	2F	F	0.0048
30	16–45°	3–7 m/s (medel)	Dag	5C	C	0.0090
30	16–45°	3–7 m/s (medel)	Natt	5D	D	0.0115
30	16–45°	>7 m/s (hög)	Dag	8D	D	0.0024
30	16–45°	>7 m/s (hög)	Natt	8D	D	0.0030
60	46–75°	≤3 m/s (låg)	Dag	2B	B	0.0057
60	46–75°	≤3 m/s (låg)	Natt	2F	F	0.0072
60	46–75°	3–7 m/s (medel)	Dag	5C	C	0.0135
60	46–75°	3–7 m/s (medel)	Natt	5D	D	0.0172
60	46–75°	>7 m/s (hög)	Dag	8D	D	0.0036
60	46–75°	>7 m/s (hög)	Natt	8D	D	0.0045
90	76–105°	≤3 m/s (låg)	Dag	2B	B	0.0128
90	76–105°	≤3 m/s (låg)	Natt	2F	F	0.0163
90	76–105°	3–7 m/s (medel)	Dag	5C	C	0.0305
90	76–105°	3–7 m/s (medel)	Natt	5D	D	0.0388
90	76–105°	>7 m/s (hög)	Dag	8D	D	0.0080
90	76–105°	>7 m/s (hög)	Natt	8D	D	0.0102
120	106–135°	≤3 m/s (låg)	Dag	2B	B	0.0102
120	106–135°	≤3 m/s (låg)	Natt	2F	F	0.0130
120	106–135°	3–7 m/s (medel)	Dag	5C	C	0.0243
120	106–135°	3–7 m/s (medel)	Natt	5D	D	0.0310
120	106–135°	>7 m/s (hög)	Dag	8D	D	0.0064
120	106–135°	>7 m/s (hög)	Natt	8D	D	0.0081
150	136–165°	≤3 m/s (låg)	Dag	2B	B	0.0086
150	136–165°	≤3 m/s (låg)	Natt	2F	F	0.0110
150	136–165°	3–7 m/s (medel)	Dag	5C	C	0.0205
150	136–165°	3–7 m/s (medel)	Natt	5D	D	0.0260
150	136–165°	>7 m/s (hög)	Dag	8D	D	0.0054
150	136–165°	>7 m/s (hög)	Natt	8D	D	0.0069
180	166–195°	≤3 m/s (låg)	Dag	2B	B	0.0077
180	166–195°	≤3 m/s (låg)	Natt	2F	F	0.0098
180	166–195°	3–7 m/s (medel)	Dag	5C	C	0.0184
180	166–195°	3–7 m/s (medel)	Natt	5D	D	0.0234
180	166–195°	>7 m/s (hög)	Dag	8D	D	0.0048
180	166–195°	>7 m/s (hög)	Natt	8D	D	0.0062
210	196–225°	≤3 m/s (låg)	Dag	2B	B	0.0114
210	196–225°	≤3 m/s (låg)	Natt	2F	F	0.0145
210	196–225°	3–7 m/s (medel)	Dag	5C	C	0.0270
210	196–225°	3–7 m/s (medel)	Natt	5D	D	0.0344
210	196–225°	>7 m/s (hög)	Dag	8D	D	0.0071
210	196–225°	>7 m/s (hög)	Natt	8D	D	0.0090
240	226–255°	≤3 m/s (låg)	Dag	2B	B	0.0157
240	226–255°	≤3 m/s (låg)	Natt	2F	F	0.0199
240	226–255°	3–7 m/s (medel)	Dag	5C	C	0.0372
240	226–255°	3–7 m/s (medel)	Natt	5D	D	0.0473
240	226–255°	>7 m/s (hög)	Dag	8D	D	0.0098
240	226–255°	>7 m/s (hög)	Natt	8D	D	0.0125
270	256–285°	≤3 m/s (låg)	Dag	2B	B	0.0158
270	256–285°	≤3 m/s (låg)	Natt	2F	F	0.0201
270	256–285°	3–7 m/s (medel)	Dag	5C	C	0.0376
270	256–285°	3–7 m/s (medel)	Natt	5D	D	0.0478
270	256–285°	>7 m/s (hög)	Dag	8D	D	0.0099
270	256–285°	>7 m/s (hög)	Natt	8D	D	0.0126
300	286–315°	≤3 m/s (låg)	Dag	2B	B	0.0101
300	286–315°	≤3 m/s (låg)	Natt	2F	F	0.0129
300	286–315°	3–7 m/s (medel)	Dag	5C	C	0.0241
300	286–315°	3–7 m/s (medel)	Natt	5D	D	0.0307
300	286–315°	>7 m/s (hög)	Dag	8D	D	0.0063
300	286–315°	>7 m/s (hög)	Natt	8D	D	0.0081
330	316–345°	≤3 m/s (låg)	Dag	2B	B	0.0049
330	316–345°	≤3 m/s (låg)	Natt	2F	F	0.0062
330	316–345°	3–7 m/s (medel)	Dag	5C	C	0.0116
330	316–345°	3–7 m/s (medel)	Natt	5D	D	0.0147
330	316–345°	>7 m/s (hög)	Dag	8D	D	0.0031
330	316–345°	>7 m/s (hög)	Natt	8D	D	0.0039