Contents

List of Figures

List of Tables

1 Survey of Quantitative Methods

The bread and butter of quantitative consulting is the ability to apply quantitative methods to business problems in order to obtain actionable insight. Clearly, it is impossible (and perhaps inadvisable, in a more general sense) for any given individual to have expertise in every field of mathematics, statistics, and computer science.

We believe that the best consulting framework is reached when a small team of consultants possesses expertise in 2 or 3 areas, as well as a decent understanding of related disciplines, and a passing knowledge in a variety of other domains: this includes keeping up with trends, implementing knowledge redundancies on the team, being conversant in non-expertise areas, and knowing where to find detailed information (online, in books, or through external resources).

In this section, we present an introduction for 9 “domains” of quantitative analysis:

- survey sampling and data collection;
- data processing;
- data visualisation;
- statistical methods;
- queueing models;
- data science and machine learning;
- simulations;
- optimisation, and
- trend extraction and forecasting;

Strictly speaking, the domains are not free of overlaps. Large swaths of data science and time series analysis methods are quite simply statistical in nature, and it’s not unusual to view optimisation methods and queueing models as sub-disciplines of operations research. Other topics could also have been included (such as Bayesian data analysis or signal processing, to name but two), and might find their way into a second edition of this book.

Our treatment of these topics, by design, is brief and incomplete. Each module is directed at students who have a background in other quantitative methods, but not necessarily in the topic under consideration. Our goal is to provide a quick “reference map” of the field, together with a general idea of its challenges and common traps, in order to highlight opportunities for application in a consulting context. These subsections are emphatically NOT meant as comprehensive surveys: they focus on the basics and talking points; perhaps more importantly, a copious number of references are also provided.

We will complement each section with projects projects we have tackled in our own practices. These case studies are accompanied by (partial) deliverables in the form of charts, write-ups, report extract, etc.).

As a final note, we would like to stress the following: it is **IMPERATIVE** that quantitative consultants remember that acceptable business solutions are not always optimal theoretical solutions. Rigour, while encouraged, often must take a backseat to applicability. This lesson can be difficult to accept, and has been the downfall of many a promising candidate.

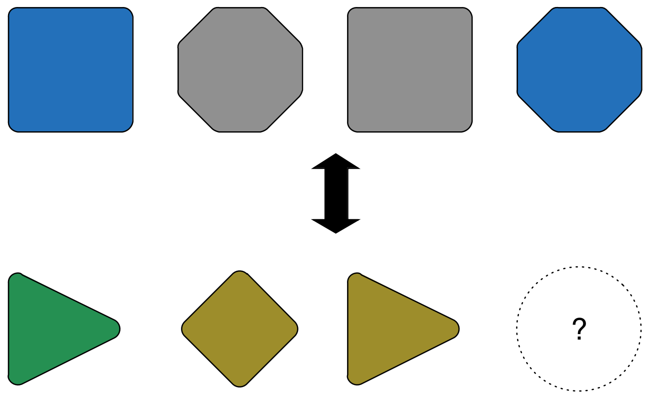


Figure 1: Can you draw an analogy between the top row of shapes and the bottom row of shapes? What should the shape and colour of the final figure in the bottom row be?

1.7 Simulations

When used in the broadest sense of the term, **modeling** is a central activity in quantitative consulting. As a result, in order to be a successful quantitative consultant, it is important to understand the different types of modeling and models, their commonalities and differences, and relevant and appropriate applications of modeling techniques. At the same time, because of its ubiquity in so many aspects of the quantitative process, the importance of modeling is often overlooked and taken for granted, since it underlies, and is incorporated into, so many other techniques. **If you are a quantitative consultant you are, by necessity, a modeler**. Consequently, having a strong general understanding of what modeling is (as distinct from particular modeling techniques) and understanding how to construct models in a more general sense, will facilitate many consulting endeavours.

Analogical reasoning is the act of reasoning from one specific occurrence to another specific occurrence, on the basis of similarity. For example,

[HAND:FINGERS, FOOT: —].

A major benefit of this type of reasoning is that it can reveal new aspects or relationships between objects that have not previously been considered. Clearly, the choice of objects used in an analogy is important:

[HAND:FINGERS, ORANGE: —]

likely yields little useful insight, but

[HAND:FINGERS, PLANT STEM: —]

might be more interesting (see Figure 1). Analogical reasoning is viewed by some as a primary **cognitive strategy**, underlying much of human cognition [1–3].

Keeping this context in mind, a model is simply an independent entity, or structure, that has useful similarities to another structure of interest, and which allows for analogical reasoning. This structure of interest is referred to as the target of the model. We can carry out inductive or deductive reasoning on the model and then, via analogical reasoning, transfer our insights about the model over to the target, and in this way learn something about the target. The target structure might be a single object or a system of objects, or a process being carried out by this system of objects.

Our ability to create a model with *useful similarities* to the target system, and then learn about our chosen target system using this model, can be extremely powerful. For example, we can make a very small model of something that is, in reality, very large or very distant – for example, a small scale model of the solar system, made out of wire and styrofoam – and use this small simple model to come up with accurate predictions about this large and distant system.

The solar system model example also showcases the importance of understanding **which parts of the model are usefully similar to the target system** in the context of our intended use of the model. If we try to use our simple solar system model to draw conclusions relating to the relative densities of planets in the solar system, we will be disappointed.

Although there are many different types of models, which we will further discussed later, in general we can say that models have two main functions: **explanation** and **prediction**.

- In some cases, we might have a system whose behaviour we do not fully understand and cannot explain. Models can help us increase our understanding of the mechanisms underlying the behaviours or properties of interest.
- In other cases, regardless of how a type of system is generating a particular behaviour, or came to have a certain property, our interest is not in understanding how this came to be, but rather in predicting the presence (or absence) of that behaviour or property in another system of the same type.

Modelers often try to create **taxonomies** or categorisations of models. These efforts have arguably not been that successful from a conceptually rigorous point of view but, pragmatically, it is still useful to consider the types of models that people commonly talk about (see [7] for a useful review and discussion of a variety model and simulation types).

IMPORTANT NOTE: it has been our experienced that clients usually take a dim view of simulations, as though they are less 'valid' or 'real' than other quantitative approaches. The reasons for this are varied, and perhaps not entirely unfounded as simulations can easily be used in the wrong way or with the wrong endgame in mind.

We discuss, in the following pages, a number of strategies to help consultants provide sound simulation solutions for their client.

1.7.1 Static Models

Central to the idea of simulations is the notion of a **model**. There are numerous modeling strategies – ultimately, however, all models are used with the goal of helping the modeler better understand a **system** (a term that we will use in an axiomatic fashion).

Conceptual Models A conceptual model is an **abstraction of a real world system** or process that defines which elements of that system or process are of interest in the current context, and how these elements and their relationships will be defined for the purposes of **drawing conclusions** about the behaviours or properties of the system. Arguably, before any other type of model can be generated, a conceptual model must first be created, either implicitly or explicitly.

Explicit conceptual models may take the form of diagrams or formalised descriptions of the system. Conceptual models may then be implemented as other types of models (e.g. mathematical, simulation).

Implicit conceptual models are often linked with gaps in the understanding of a system – assumptions that go unchallenged and unstated are often less clear and obvious than is originally believed. An engineer may, for instance, state to a consultant that the probability of a certain component failing by time t is 0 without feeling the need to specify that, in the jargon of the discipline, this really means that

$$P(\text{failure by time } t > T) > \varepsilon > 0, \quad \text{for a sufficiently large } T \text{ and a sufficiently small } \varepsilon;$$

the consultant, not knowing the conventions of the field, might understand this to mean that

$$P(\text{failure by time } t) = 0 \quad \text{for all } t,$$

and the mistake can propagate through the simulation, potentially making the simulation useless in practice.

Mathematical Models A mathematical model uses mathematical expressions to **support reasoning** about a real world system. Relationships between objects in the system, or their properties, are represented by mathematical relationships between variables. If the relationships within the mathematical model are **sufficiently similar** to relationships between objects in the system of interest, then carrying out mathematical methods (truth preserving manipulations) on the model should result in new true conclusions about the system.

Note that arguments represented by symbolic logic also fall under this category. As a result, it could readily be said that all models implemented on computers are a type of mathematical model. That being said, the expression ‘mathematical model’ typically refers to models that are not necessarily implemented on computers, and which consist of systems of mathematical equations.

Although mathematical models may represent processes and dynamic elements of systems by including time and space as variables, the models themselves are **static**, in the sense that they do not change over time in a manner that is similar to the ways in which the target system itself changes over time. Mathematical models may still be implemented on computers and methods for **solving the systems of equations** in these models (e.g. symbolic manipulations, numerical analysis) may be carried out using computer algorithms, however, the fact that this work is carried out on a computer, and the fact that simulations are also carried out on a computer, does not mean that finding solutions to equations using programmatic strategies is the same as carrying out simulations, in the way that the term ‘simulation’ is typically used. Simulations will be further discussed below.

Figure 2: To-scale architectural model of the interior of an office building [10].

Statistical Models Conceptually speaking, statistics represents the world in terms of **populations** and **processes**. These populations and processes then have certain properties, which can be represented themselves using mathematical expressions. Statistical models could thus be described as mathematical models motivated by a certain (statistical) conceptualisation of real world processes.

To-Scale Physical Models A to-scale physical model is a model that is constructed from **physical materials**, which are shaped and positioned in such a way as to accurately represent the physical layout and positions of elements of the target system, as well as their size, relative to each other (see Figure 2 for an example of an architectural model).

Data Models A data model is a conceptual model used to design the structure of data storage. Since data itself represents facts about a system, it is appropriate to first conceptually model the properties and relationships that exist within the system, and which are represented by the data, and then use this conceptual model to create a data storage structure that can be used to efficiently **hold**, **extract**, **edit** and **add** to the stored data (see Figure 3 for an example).

1.7.2 Dynamic Models

In some situations, only the static aspects of a system are interesting, or the system itself is mostly static. For example, if we build a physical model of a house, we expect both the house and the model to be **relatively unchanging** – the measurements of the rooms and the furniture in the house will not change from minute to minute (although they could change over years due to remodeling, or drastically if the house is sold to new owners with a different sense of aesthetics),

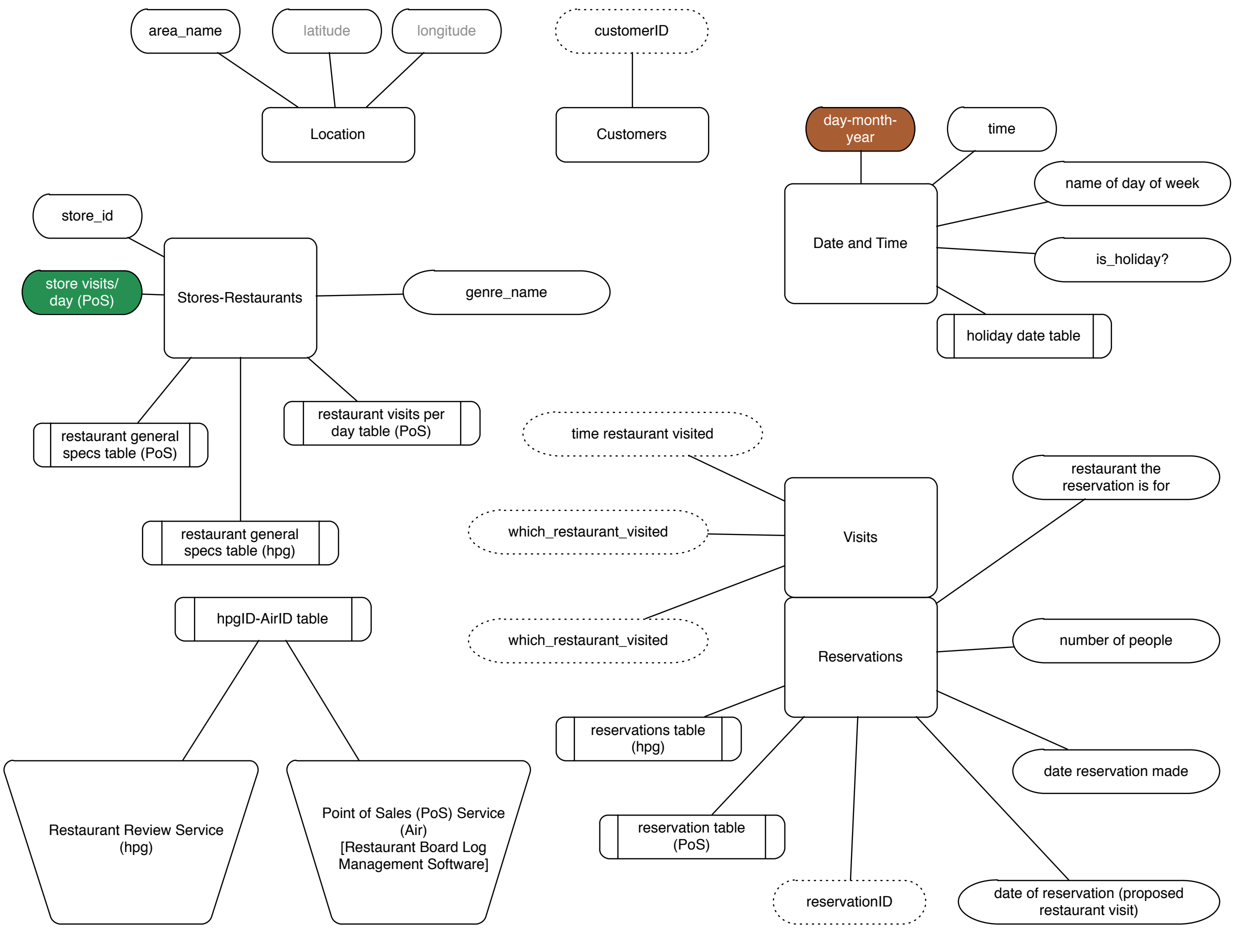


Figure 3: A preliminary data model of a restaurant reservation system, which can be used to help design an efficient data storage structure, as well as develop data analysis strategies.

and the model will not need to change either. We can then use the model to reason about the house:

- if the model couch fits against this wall in the model house, we can reason that the real couch will fit in the real house;
- if eight model chairs cannot simultaneously be placed around the model kitchen table, than eight real chairs cannot be simultaneously be placed around the real kitchen table, etc.

Other systems, however, are more active, or **dynamic**, with processes taking place within the system. When modeling these dynamic system elements, we often talk about **simulation models** or simply **simulations**. Although the term 'simulation' is not precisely defined, it typically indicates that a model is intended to **reflect the behaviour of the target system** – its processes – over time, and also that the **model itself will independently change over time**, when it is run. The goal is to construct the simulation in such a way that it will change over time in ways that are similar to the manner in which the system itself changes over time. As a result we can use the simulation to predict past, current, and future behaviours of the system. Historically speaking, simulations have often modeled individual object-level properties and behaviours, as well as the mechanisms underlying relevant behaviours, rather than group level properties or system outputs, but this does not have to be the case.

Modeling Time and Movement How do we incorporate time and movement into a model? To return to our styrofoam and wires model of the solar system, if we set up our model such that when we turn a crank the planets and moons realistically move around a light bulb in the centre of the model (representing the sun), then we have a dynamic model, or simulation of the solar system. We can simulate what will happen within the actual solar system over time.

As another example, if we wish to know how emergency responders might behave in different plane crash scenarios, we could set up a number of simulated crash scenarios, with a life-size model of a crashed plane, and actors behaving as injured people might. We can then have the emergency responders try out (i.e. simulate) different approaches and strategies to dealing with plane crashes.

The advent of computers greatly facilitated the construction and possible uses of simulations, because it made it possible to simulate dynamic systems **virtually** instead of having to create a dynamic physical model of the system, whose elements could be represented as data structures (and variables within these structures) within computer programs. The physical interactions between these system elements could then, in turn, be represented by logical rules and mathematical equations operating over these data structures.

These logical rules and mathematical equations pushed computer simulations closer to the domain of mathematical models, relative to physically constructed models. At the same time, computer simulations retained the strategy used by these physical models of determining what would happen to the system by moving the model through its expected behaviours step-by-step, over time. Rather than mechanically moving the model (or using people and other elements in this capacity) computer models rely on the computer processor to run the program that represents the system, and essentially 'move' (in an electronic sense) the model based on the behaviours the model implements. As discussed earlier, this is a different technique than the one used by mathematical models implemented on computers.

1.7.3 Uses, Data, and Contrast with Mathematical Modeling

Simulations are typically used to

- better understand actual real-world phenomena and systems, and
- explore phenomena that don't currently exist but which could exist hypothetically.

Simulations can allow us to both **predict** what our target system will do under particular circumstances, but also **explain** why a system behaves the way it does. However, given that we build simulations using only what is already known (or possibly suspected) to be either currently the case about the system, or at least plausible within the conceptual phase space in which the system resides, you may wonder how a simulation could possibly tell us anything new about the system, and thus, why we would ever bother running simulations.

Humans thinking is typically unable to capture all the possible interactions between a system's various parts, and how these parts influence each other in particular circumstances; **merely** thinking through the behaviours of a system which is even slightly complicated is likely lead us to miss implications, and, as a result, **incorrectly predict or explain** the system's behaviour. If, instead, we introduce what we do know into the simulation and allow it to behave based on these rules, behaviours that we would not easily have anticipated can emerge from the process.

Consequently, the notion of **emergence** is crucial in simulations. We can say that simulation behaviours emerge when they are not programmed in the simulation directly, but rather occur as the result of interactions between model components that are themselves programmed into the simulation directly.

The emergent behaviours may occur at different **levels of granularity** of the system. For example, if we create a simulation of people in a city, we might see emergent behaviour with respect to which people most frequently interact with which other people, and we might also see emergent behaviour at the population level, where the average number of people in a given location is equal to a particular value over time.

We can see from this example how emergence allows us both to predict and to explain elements of a system that were not previously amenable to such efforts. We can predict average numbers of people in a particular location, if this information is not available from another source; if it is, we can still use the simulation to explain the origins and underpinnings of this number, by referring to the more granular system components whose interactions lead to the value.

IMPORTANT NOTE: 'emergence' is a concept that has crossed-over into a large number of areas of human endeavour. Don't be surprised to hear clients talk about "emergent phenomena" in contexts where you would not normally expect to hear it. It is quite conceivable that they have a very thorough understanding of what emergence means and what it entails – don't make the classic quantitative consulting mistake of assuming that clients do not understand technical concepts ... you never know what their background and interests are – but, together with terms like 'synergy' or 'big data', it seems to have entered the business lexicon as a trendy but ultimately meaningless term. Be sure to clarify the situation at an early stage (by definining concepts in the proposal, say) in order to avoid the confusion and headaches that can result when deliverables are handed off.

Simulations and Data All modeling activities rely on the modeler having **accurate** and **relevant** information or data about the target system, which allows for the construction of a model with useful similarities to the target system, which is basically a data collection/information gathering problem. But even then, simulations have a particular relationship with data:

- first and foremost, data is needed in order to properly set simulation parameters – the initial simulation settings that determine how the simulation will run in a particular instance; in the absence of this type of information, although the simulation may generate outputs that could, in principle, have some relevance to the target system in some circumstances, the simulation behaviour is unlikely (or at least, should not be expected) to overlap with target system behaviours of interest within the specific context in which the simulation was generated;
- secondly, simulations have the capacity to generate large amounts of data about the behaviour of the simulation, and by extension, the target system. This data, sometimes referred to as 'synthetic data' or 'simulated data', can be uses as a stand-in for actual data about the system, just as the model is being used as a stand-in for the target system.

When very little is know about reasonable parameters values, a preliminary simulation might first be required in order to produce data which could then be used to set simulation parameters, which, in turn, could be used to produce data for analysis. It is not too difficult to conceive of multiple links being added to this chain; our advice is to keep the number of such links to a minimum (preferably zero) – in light of the point made in the first item above, it might be preferable to garner information about parameters from **first principles** (or other models).

Simulations vs. Mathematical Models The procedural element of computer models, whereby the behaviour of the target system must be, in a sense, mechanically replicated by the data structures and procedures of the computer program, distinguishes **computer simulations** from mathematical models, which, rather than modeling the temporal, dynamic components of systems by incorporating a temporal, dynamic component directly into the model, instead represent them as variables in mathematical equations that represent components and behaviours of the system.

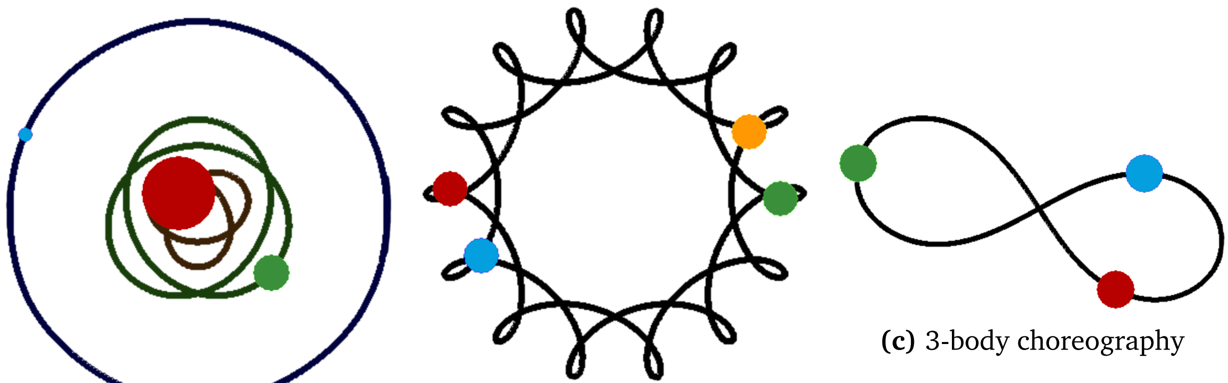
On this front, the advantage of mathematical models is that **deductive reasoning** (or first principles reasoning) can, in theory, be used to determine the target system behaviour, rather than have to resort to 'running' the model over a range of starting conditions. This is appealing, as mathematical strategies can allow for more definitive and general statements about the system (e.g. "The system will never do the following"; "The system will always do the following", etc.); these types of statements are typically outside the reach of even the most advanced mechanical or programmatic simulations. In practice, however, the underlying complexity of such models limit the usefulness of this approach in most scenarios.

Consider, for example, the n—**body problem** (nBP) of classical mechanics, which consists in predicting the individual trajectories of n celestial bodies bound by gravitational attraction.

Using Newtonian meachanics, the trajectories can be deduced to follow the paths described by the following system of differential equations:

$$
\begin{aligned}
m_1\frac{\mathrm{d}^2\mathbf{q}_1}{\mathrm{d}t^2} &= \sum_{j=2}^{n}\frac{Gm_1m_j(\mathbf{q}_j-\mathbf{q}_1)}{\|\mathbf{q}_j-\mathbf{q}_1\|^3}\\
m_2\frac{\mathrm{d}^2\mathbf{q}_2}{\mathrm{d}t^2} &= \sum_{j\neq 2}^{n}\frac{Gm_2m_j(\mathbf{q}_j-\mathbf{q}_2)}{\|\mathbf{q}_j-\mathbf{q}_2\|^3}\\
&\vdots\\
m_n\frac{\mathrm{d}^2\mathbf{q}_n}{\mathrm{d}t^2} &= \sum_{j=1}^{n-1}\frac{Gm_nm_j(\mathbf{q}_j-\mathbf{q}_n)}{\|\mathbf{q}_j-\mathbf{q}_n\|^3},
\end{aligned}
$$

where m_i and $\mathbf{q}_i(t)$ are, respectively, the mass and the trajectory of the i^{th} celestial body in 3-space, and G is Newton's constant. These equations describe, in principle, the behaviour of stars in a globular cluster, say, or of the Earth-Sun or the Earth-Moon system. They cannot provide a **complete** description as the range of gravitational attraction is infinite – every 'object' in the Universe influences every 'other' object to some extent, no matter how distant (see the precession of Mercury, for instance), and other forces may also act on the bodies (but at large distances, the force due to gravity overwhelms the other 3 forces), but, for most practical applications (if one can consider astronomy a practical discipline), they are more than sufficient, as long as we are willing to ignore relativistic effects.



(a) Planetary 3BP

(b) 4BP pairs of bodies orbiting each other

(c) 3-body choreography

Figure 4: Possible solutions of the n—body problem, based on [14].

What do the solutions look like? A typical mathematical approach would be to try to solve the 2BP, and to see if the solution can be generalised to more complex cases.

The **two-body problem** has an exact solution. The centre of mass of the two bodies is the vector

$$\mathbf{x}(t) = \frac{m_1\mathbf{q}_1(t) + m_2\mathbf{q}_2(t)}{m_1 + m_2}.$$

In the centre-of-mass frame (that is, in the frame that moves along with the centre of mass), it can be shown, using physical conservation laws, that the trajectories of the two bodies are co-planar and 'orbit' the system's barycentre, with an angle $\theta_i(t)$ which depends on the reduced mass of the system $m_* = \frac{m_1 m_2}{m_1+m_2}$ and on the effective potential $U(r(t), \ell, m_*)$, where $r(t) = \|\mathbf{q}_2 - \mathbf{q}_1\|$ and ℓ is the system's angular momentum.

Various combinations of parameters lead to various orbits; if the effective potential admits a local minimum, for instance, the orbits will oscillate around the barycentre (elliptic or precessing elliptic paths, in each Sun-planet system); if the effective potential does not admit a minimum, then the orbits may escape to infinity (hyperbolic or parabolic paths, such as in some Sun-comet systems).

Under some restrictions on the masses and momenta of the bodies, the nBP can be shown to have closed-form solutions or theoretically understood approximate solutions (see [11] for a list, and Figure 4 for some illustrations), including:

- **Euler's Problem of Two-Fixed Centres** allows for colinear motion in systems where two of the three masses are comparable and fixed;
- the **restricted** 3BP shows the existence of 5 fixed configurations (involving the **Lagrangian points**) which rotate around the system's barycentre in cases where one of the masses is negligible, such as is the case in the Sun-Jupiter-Trojans systems (there are two);
- the **planetary** nBP admits quasi-periodic solutions in systems where one of the masses is significantly larger than the other $n-1$ masses, which shows that planets in stable, planar, and nearly circular orbits around a star *can* transition to chaotic orbits, but that these orbits would be bounded by quasiperiodic tori and so would preserve some regularity, and

- n—body choreography in which all the masses move on the same manifold, without collisions.

The **general** n—body problem can be solved analytically using Taylor Series (known as Sundman's series), but the series converge so slowly as to be of no practical use for astronomical results (which would require at least $10^{8000000}$ terms in the 3BP case, well beyond even what modern computers can produce [12]. A whimsical take on the effects of such unpredictable behaviour is offered in Liu Cixin's *The Three-Body Problem* [13].

By contrast, to draw conclusions from a simulation you must set certain initial conditions and then run the simulation and examine the resulting output. Each simulation run represents only one specific instance in the model space. As a result, it can be difficult, if not downright impossible, to draw general conclusions from the results of one or even multiple simulation runs (to say nothing of exploring the outcome of using different parameter values).

This has lead to criticism over the use of simulations in some milieus, on the basis that simulations should **never** be used if mathematical models can be used instead.

However, the nBP illustrates why taking this hardline position may be inadvisable; clearly, there are circumstances in which it is difficult to create solvable (actionable) mathematical models of a system that represents the target system in ways sufficiently similar to the system in relevant respects in order to for salient and accurate conclusions to be drawn about that system, in which case a simulation might provide greater insight. It is also possible to create hybrids of mathematical and simulation models to allow for increased insight into system behaviours.

If n is relatively small, the nBP trajectories can be approximated to a high-level of accuracy by using numerical methods to solve the corresponding system of differential equations (see [16] for an example of planetary system formation). For astronomical bodies that avoid collisions (or near encounters), there are two main technical issues:

- the first one is that the nBP problem is **chaotic** for $n > 2$, so that small errors (such as are generated by truncating initial conditions or intermediate calculations) may lead to simulated solutions that are wildly divergent from the true paths;
- astronomical simulations typically run over million of years, leading to an accumulation of integration errors; this is problematic as the approximate solutions are only mathematical objects, whereas the actual bodies they represent have to satisfy physical laws (including the various conservation laws); this can be tackled by using analytical methods such as the **variational principle** and **perturbation theory** to produce trajectory manifolds on which to 'project' the integrated approximations.

For many bodies, the time complexity is related to the square of the number of bodies (more on this later), which can make the direct simulation unpractical. In that case, useful simulations must approximate the essential character of the actual trajectories while reducing the computational complexity. There are many dedicated methods to achieve this goal (including so-called **tree code** and **particle mesh** methods) [11].

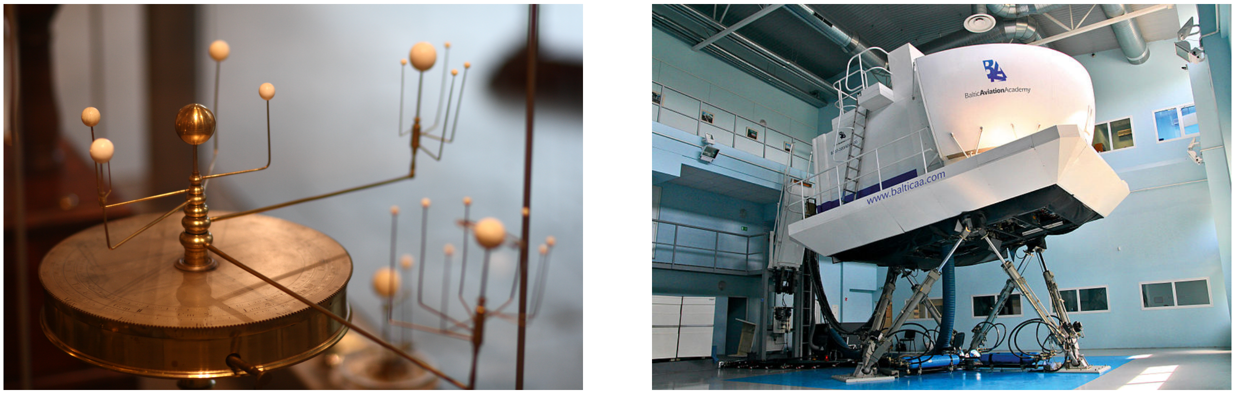


Figure 5: Harvard orrery [17], and Baltic Aviation Academy Airbus B737 Full Flight Simulator (FFS) in Vilnius (public domain).

While these particular issues may not apply to general simulations, the interplay of valid approximation and computational feasibility lies at the core of successful simulations.

1.7.4 Types of Simulations

We have already alluded to some types of simulations; in this section we provide more concrete descriptions of the avenues available to modelers.

Full-Scale Physical Simulations Full-scale physical simulations are **life-sized**, **physically realistic** simulations, which make use of structures that already exist to replicate or reproduce target system behaviours.

For example, to simulate boat rescue situations (and then practice responding under various scenarios), the Coast Guard might make use of existing vessels and emergency personnel, and introduce actors playing the part of accident victims, a wave machine to simulate possible environmental conditions, etc.

Mechanical Simulations A mechanical simulation is one that is physically implemented but which is not necessarily full-scale, to-scale or physically realistic in various other respects. It simulates dynamic behaviours using electro-mechanical components. Mechanical simulations were popular prior to the advent of computers. The 'orrery', a classic type of clockwork model of the solar system, is a typical example of a mechanical simulation (see Figure 5, left). Another example would be a CPR dummy that can be used to practice proper CPR technique, and which may have sensors to simulate certain heart behaviours and then provide feedback regarding the effectiveness of the applied CPR.

Computer (Programmatic) Simulations Programmatic simulations represent the target system or process using **data structures** and **algorithms**. The data structures are sets of variables that represent the properties of system objects, and the algorithms determine how these properties change over time. When quantitative consultants produce simulations, they are usually programmatic.

- **Event-Centric Computer Simulations** – This type of computer simulation models **activity** (and is dynamic in this sense), but the focus is not accurate modeling of time. The goal, rather, is to represent an event or sequence of events. For example, we might simulate the selection, and result, of sampling a population, or simulate possible outcomes of a series of events that themselves occur with particular probabilities.
- **Discrete Time Computer Simulations** – As suggested by the name, discrete time simulations treat time as a **discrete series of consecutive steps**, rather than continuously. A common example of this is the **agent-based** model (or multi-agent simulation); in this type of simulation, the time step may range from seconds to years, and the goal of the simulation is to explore how individual agents interact with each other over this time span.
- **Continuous Time Computer Simulations** – In contrast to discrete time simulations, continuous time simulations treat time as a continuous property. However, there is a challenge here, as continuous time simulations are generally implemented on a computer, and computers are necessarily discrete. Thus, in practice, a continuous time simulation is one where the discrete time steps are simply **very small**. Note that this is not equivalent to implementing a continuous-time mathematical model on a computer and solving it using mathematical methods implemented as algorithms.

Hybrid Models It is also possible to create a model of a system where one part of the model is of one type and another part is of another type. A realistic flight simulator, for instance, might consist of a few full-scale physical components such as the cockpit, seats, etc. (possibly using part of an actual plane), while the experience of actually flying the plane is simulated via computer, and perhaps integrated with the physical part of the simulation by projecting a computer controlled image onto the cockpit window (see Figure 5, right). The computer simulation might also controls the physical behaviour of the motion of the cockpit – its pitch, yaw, and roll, for example.

1.7.5 Strategies for Creating Models and Simulations

Among practitioners, it sometimes said that modeling is as much an art as it is a science. While there are no tested and true approaches that will work no matter the situation under consideration, the following steps, illustrated in Figures 6 to 11 with the simulation of a school of fish, often end up having practical importance in the process:

- gather information about the target system;
- create a conceptual model;
- build the model;
- verify and validate, and
- run and analyse.

Gathering Information About the Target System As domain experts or modeling specialists, it can be tempting to believe that the understanding of the target system is so strong that that we can forgo collecting and validating information about that system and jump right into implementing a model of the system. However, modelers tend to be experts in specific techniques rather than in the behaviour of the target system, and *vice-versa* for the domain experts – **teamwork** is usually required to properly construct the model.

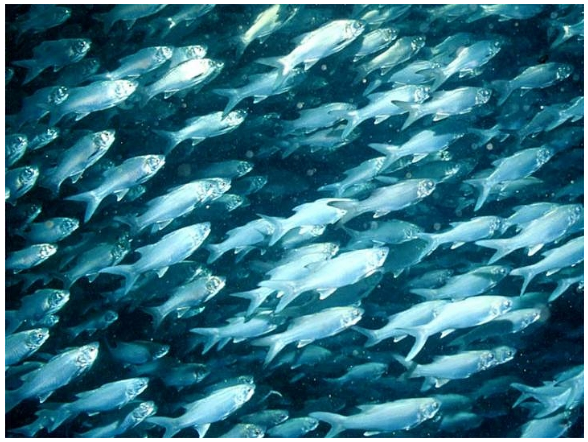


Figure 6: A school of fish – an example of a target system to be simulated [18].

In such a case, the modeler and domain expert must work together closely to **gather the information** about the system that the domain expert believes will be required to understand or predict the relevant behaviours of the target system. The modeler must also keep in mind the types of information required to create a comprehensive and consistent model of the system, given the proposed model type. Creating a **conceptual model** (see below) will greatly assist with the process of determining what information is necessary to properly **represent** the target system.

There is also an opportunity to **validate** the structure of the model at this stage. Even when a domain expert is involved, ensuring that the information being incorporated into the model comes from **rigorous and reliable sources**, and **documenting these sources** early on, will enhance the likelihood that the model will be valid, as well as **increasing the credibility** of the model in the eyes of those using the model.

Creating a Conceptual Model A **conceptual model** is a clearly defined description of those **components**, **properties**, and **relationships** of the system that are believed to be important, relative to the system behaviours or properties of interest (i.e. the **modeling context**). A conceptual model may be:

- a verbal description of the system, structured or organised in some way;
- a collection of diagrams depicting elements of the system and their relationships, or
- a combination of both.

The conceptual model can be thought of as the **blueprint** that will be followed during construction of the model.

At this stage, the modeler will also often discover that it is necessary to define concretely the more abstract or less well defined elements of the target system, in preparation for implementing the model. As well, it may be determined during construction of the conceptual model that there are gaps in the understanding of the system itself, which prevent the construction of a complete model of the system. If this occurs, it may be necessary to return to gathering information about the target system itself. If the required information is not readily available it is important at this step to indicate which parts of the model are based on reliable knowledge about the system and which parts are speculative.

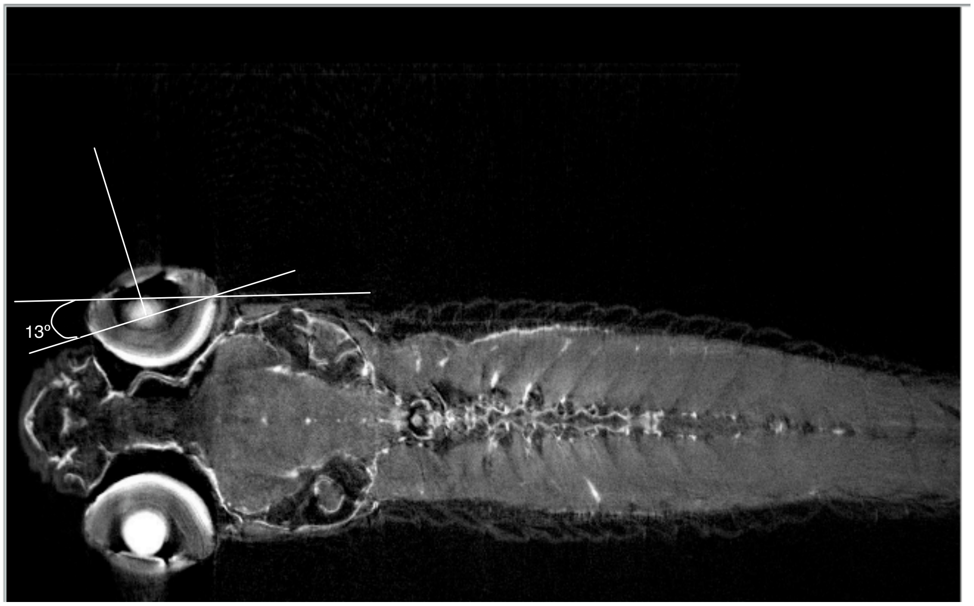


Figure 7: Gathering information: relevant perceptual mechanics information about a single fish, to be incorporated into the model [6].

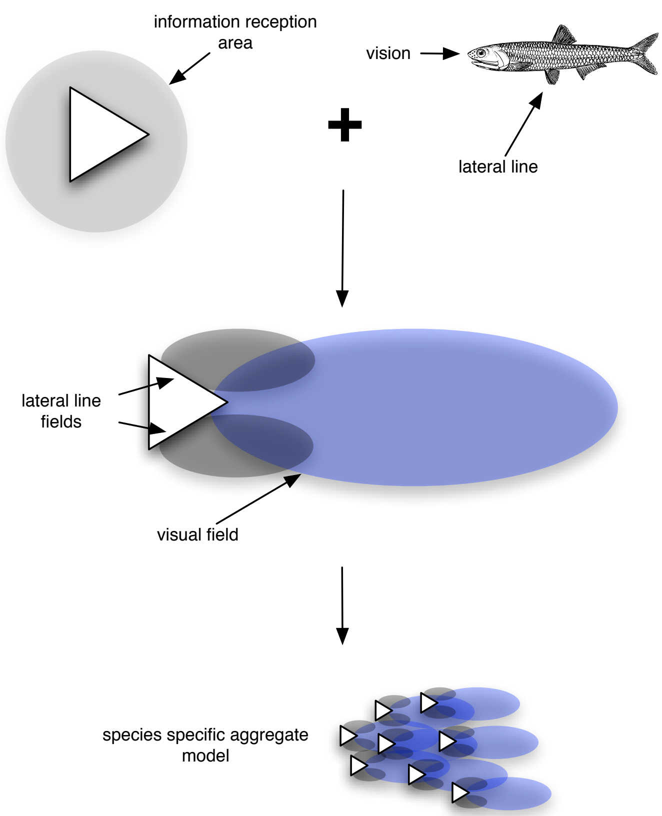


Figure 8: Creating a conceptual model: a conceptual model showing how elements of the target system – the fish in a fish school – will be represented in the model of the fish school [6].

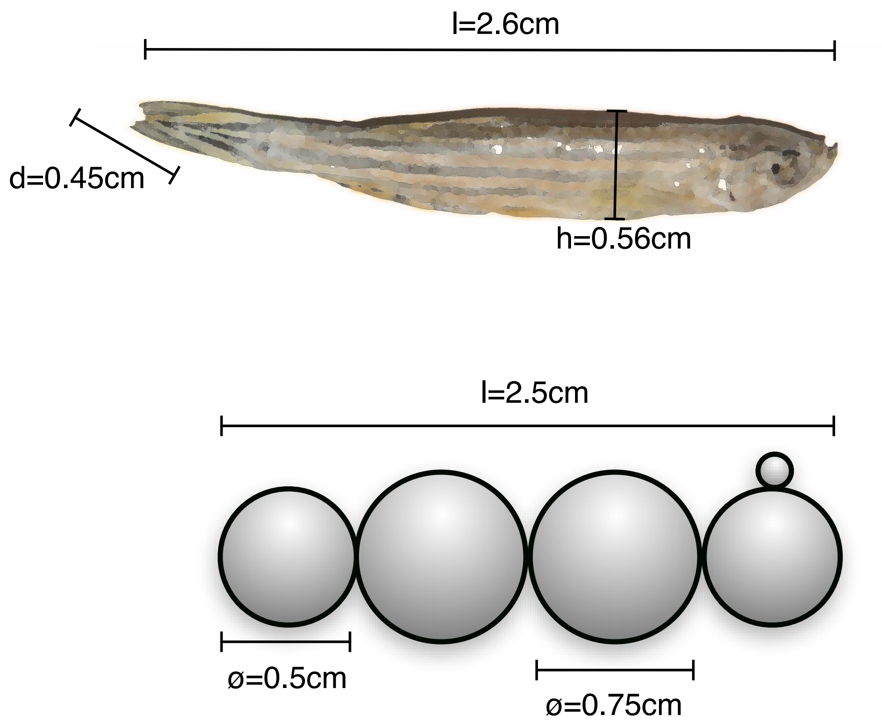


Figure 9: Creating a conceptual model: determining how specific relevant physical characteristics of individual fish will be represented and incorporated into the model [6].

This step can be challenging from an interdisciplinary perspective because, as we have already mentioned, it requires the modeler and the domain expert to work together to create the conceptual model. For this step to proceed as it should, the domain expert must, in a sense, enter into the world of the modeler, just as the modeler must enter into the world of the domain expert. This can be difficult to achieve, for a variety of reasons, and as a result it can be tempting to to skip this step outright – to leave the conceptual model in an implicit stage rather than an explicit stage – and to jump straight into building the model. However, unless the modeler is a domain expert and the system itself is relatively simple, this can lead to models that do not perform satisfactorily when all is said and done.

Building the Model Once the conceptual model is in place, a model type (e.g. mathematical, simulation) can be selected in order to **build the model** itself, using the conceptual model as a blueprint. Target system objects, properties and relationships are translated into model structures.

Verifying and Validating Verifying the model means going over the model in order to confirm that it has been **constructed as intended**, given the conceptual blueprint that has been developed. Validation refers to a process of confirming that the constructed model is in fact a **good match** for the target system. Thus, a model could be verified as having been constructed as intended, but the model might still be invalid if, for example, the modeler was misinformed about the actual workings of the target system. A thoughtful discussion of model validation, in the context of building population-based disease simulation models, can be found in [4].

```
TIMESTEP(O, λ, Φ, N)
for each agent in λ
    do L′ ← ATTENTION(O)
       I ← COGNITIVE-PROCESSING(L′)
       ACTION(I)
for each agent in Φ ▷ perception deprived agents
    do L′ ← ()
       I ← COGNITIVE-PROCESSING(L′)
       ACTION(I)

ATTENTION(O)
L′ ← MERGE-LISTS(O)
L′ ← PICK-NEIGHBOURS(N, L′)
▷ the appropriate PICK-NEIGHBOURS procedure (below) is called for each scenario
return L′

PICK-NEIGHBOURS-RANDOM(N, L′)
return RANDOM(N, L′)

PICK-NEIGHBOURS-NEAREST(N, L′)
return NEAREST(N, L′)
```

Figure 10: Building the model: pseudo-code describing how the simulation of the fish school is created [6].

Running and Analysing Once the model has been verified and validated, it may then be analysed in order to **draw conclusions** about the target system. In the case of simulations, model parameters have to be selected, and 'runs' of the model carried out for each set of parameters. A 'run' here means that the model is given certain initial starting conditions and then the behaviour of the simulation allowed to proceed and produce various outputs of interest. If the model has stochastic components, it may be necessary to carry out multiple runs using the same parameter settings in order to produce posterior distributions for the outputs. Once the simulation has been run with all of the relevant parameter settings, the resulting output of the simulation can be analysed. At this point, the analysis may follow any of a vast number of methods: trend extraction and forecasting, classification, data visualisation, etc.

1.7.6 Computational Complexity of Simulations

Because simulations are computer programs, it remains crucial to be aware of the broader issue of **computational complexity** when constructing simulations. The computational complexity of an algorithm is based on the number of possible steps in the algorithm and how they interact with different types of data to lead to different run times.

Although a detailed discussion of computational complexity is beyond the scope of this section, understanding that the manner in which the simulation is programmed will influence its run time is very important, as this might limit the options for the exploration of parameter space.

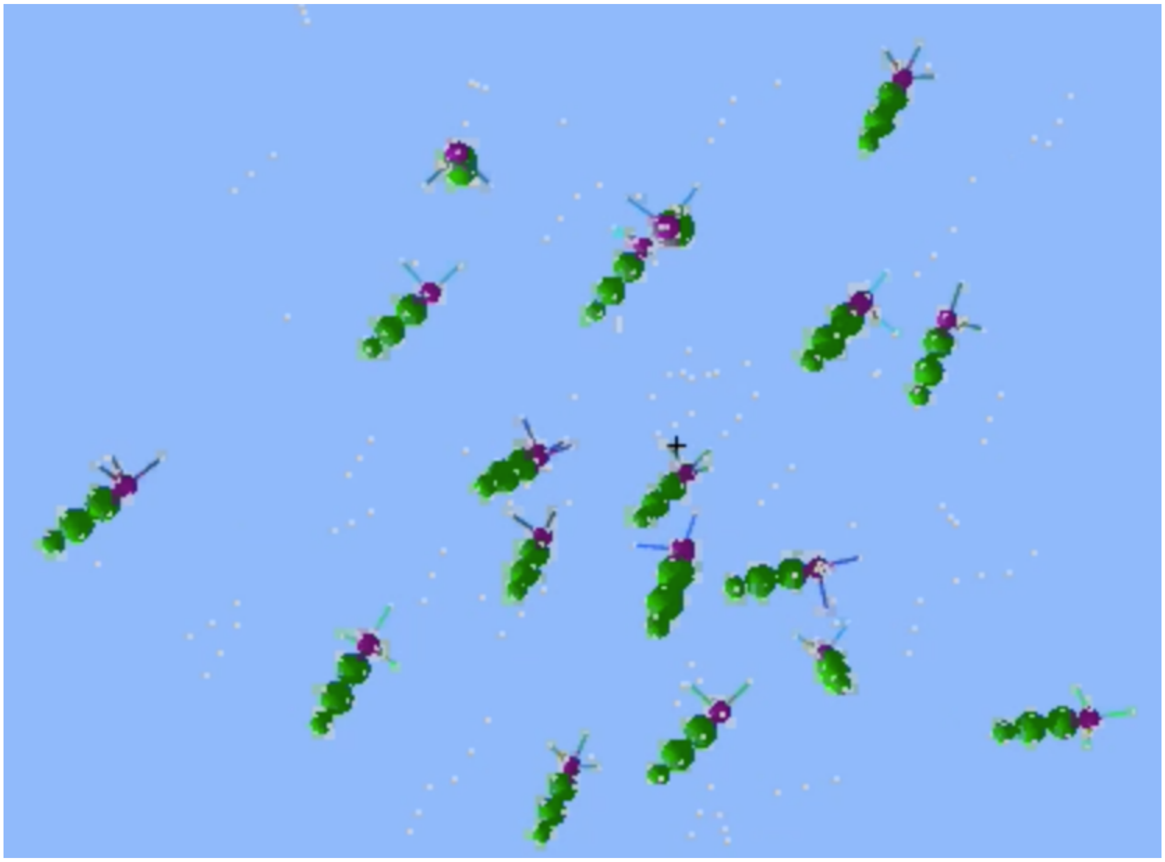


Figure 11: Building the model: the resulting simulation of the fish school. The schooling behaviour is an emergent property of the simulation, coming out of programmed individual simulated-fish behaviours [6].

As previously discussed, when a simulation is created, a set of parameters to vary has to be explicitly selected in order to explore the behaviour of the simulation. However, because specific parameter values have to be chosen for each run of the simulation, and because multiple simulations have to be run in order to get a general sense of the behaviour of the simulation (i.e. building a posterior distribution for the behaviour), and by extension the system, the problem of **combinatorial explosion** is encountered very quickly. The problem cannot always be bypassed, and it might be that the best that can be hoped for is to maximise the number of simulation runs that the computer can support in the available time.

1.7.7 Model and Simulation Applications

Science The appropriate role of models and simulations within science is a topic for debate within scientific circles. Statistical models are well accepted and used extensively. Mathematical models are generally accepted if used in a theoretical context. In our experience, however, the use of simulations is currently not well tolerated. In situations where carrying out actual experiments would be difficult (e.g. for ethical reasons), simulations may be viewed as a type of virtual experiment. In such situations the results of the virtual experiment, although not viewed in the same light as actual experimental results, may, at the very least, usefully fuel the discovery of hypotheses, which may then be tested using other methods.

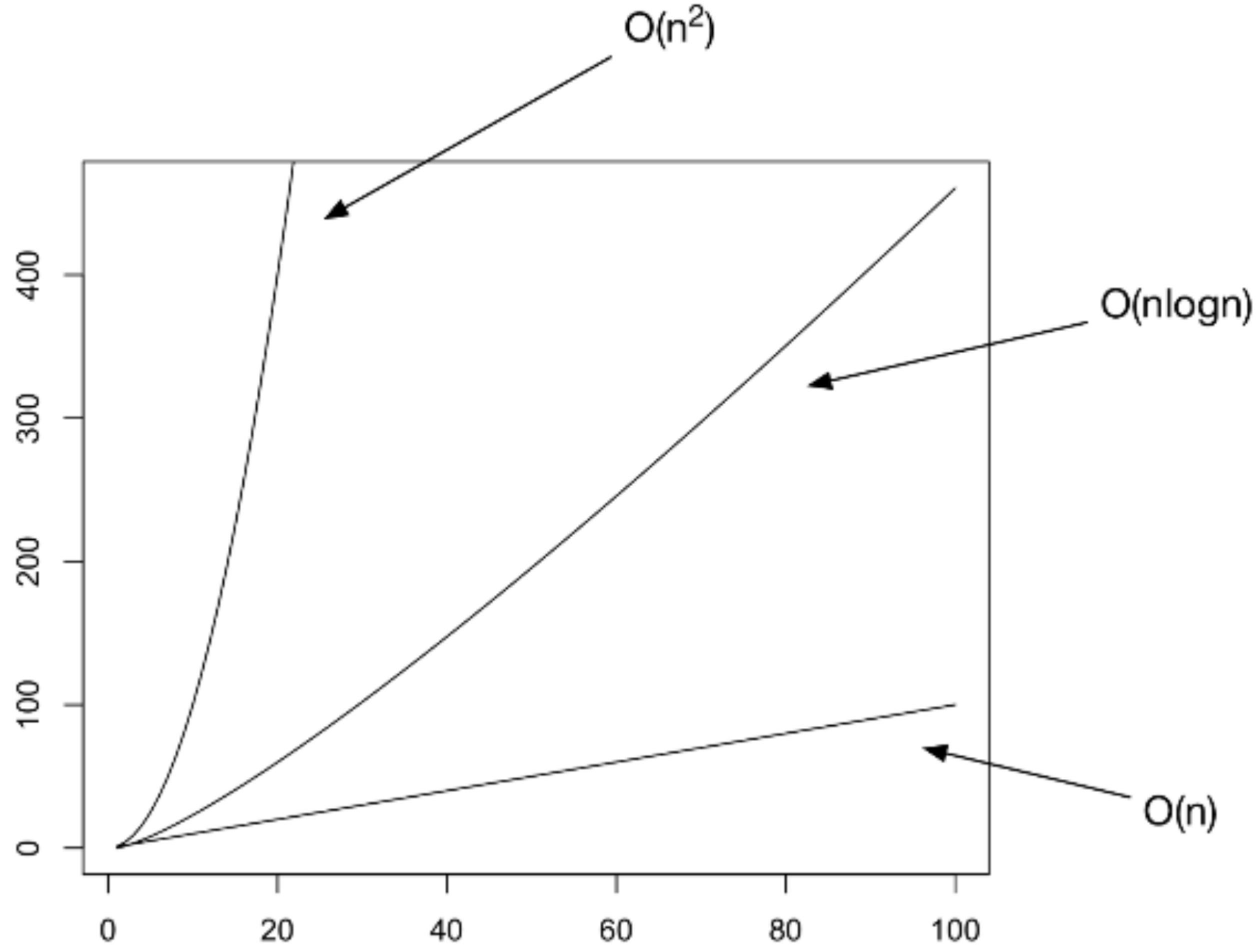


Figure 12: A sketch of some different possible computational complexities of a computer program, as represented in Big-O notation.

Business Accurate prediction of events is highly valued in a business context. As a result the emphasis for models in this domain is on predictive accuracy, rather than on being able to use the model for explanatory purposes. Businesses use models to, for example, predict customer behaviour, how their business will be affected in certain market situations, and how they might reorganise their business structure to reduce overhead and increase profitability.

Government A major activity within the government is setting policy. Within this context, it is often important to explore different possible policy scenarios, and gain a better understanding of which policies will be more or less effective under a variety of circumstances. Models that provide explanatory power can be particularly helpful in this type of work, because it allows for an understanding of why one approach might work better than another. This can then be taken into account in order to ensure good policy.

In addition, as with businesses, the government is interested in making its own operations more efficient and effective. From an organisational perspective models can be useful in determining the best strategies for internal structures and processes, as well as the conditions under which such structures may function more or less optimally.

Education Simulations play an important role in education. They allow students to explore and experience scenarios in a virtual manner, which both decreases the potential consequences of learning through doing, and increases the possibility to learn from experiences in controlled and monitored conditions. For a very thorough discussion of the role of simulations in education, see [5].

Entertainment It might be argued that most forms of entertainment are simply reflections or representations of real world experiences, and thus are, in some sense, models of life, broadly speaking. More specifically, simulations and models frequently play an important role in theatre, television and film – allowing the creators of such media to convincingly mimic real life situations without needing to entirely re-create or enact them. Use may be made of physical small-scale models (e.g. a small-scale model of a cityscape), life-size models of particular environments (e.g a life-size model of a submarine) or computer simulations (e.g. simulated flocks of birds and artificially generated clouds, added to give more realism and detail to the backdrop of a scene).

1.7.8 Modeling and Simulation Software

It is quite possible to create models by hand, without the use of computers, and it is also possible to create computer models or simulations without using a particular programming environment. Some programming environments have been specifically designed for creating simulations. Some of these currently available (as of 2018) include:

- Matlab Simulink (commercial simulation software)
- Simio (commercial simulation software)
- Netlogo (free software, mainly for teaching and prototyping)
- SymPy (a python library for discrete time simulations)

1.7.9 Case Study: NWMO

Canada has a long history with nuclear power: the first self-sustained Canadian nuclear reaction was achieved at Chalk River's ZEEP reactor in 1945. Over the years, numerous research reactors and power reactors have been built and decommissioned – as of 2014, electricity is currently being produced by 19 CANDU reactors in Ontario and New Brunswick. Given that the existence of high energy nuclear waste in Canada is a *fait accompli* – we have already chosen, as a society, to use nuclear power and create nuclear waste – it is paramount that we find ways to safely dispose of this waste.

In 2002, the *Nuclear Fuel Waste Act* (NFWA) was enacted to study possible strategies for the management of Canada's used nuclear fuel. As a result, the *Nuclear Waste Management Organization* (NWMO) was formed by the Canadian nuclear power companies, with the mandate to provide recommendations to the Canadian Government for the long-term management of used nuclear fuel. One such recommendation, which was accepted in 2007, was the establishment of Adaptive Phased Management (APM) as both a social and technical approach to permanently manage Canada's used nuclear fuel. Canadian citizens determined that the optimal strategy, given the current state of technology in Canada, is the construction of a deep geological repository to contain and isolate the fuel.

This decision puts the NWMO in a unique and demanding position, as it is the first group in Canada to design and build a unique but extremely performance-critical engineering structure: a long term Canadian repository for high energy nuclear waste. By its very nature, this structure as a whole cannot be tested in advance of use and essentially cannot be maintained once it is built. Furthermore, the environment and materials involved are themselves volatile and their long term behaviour is difficult to predict.

Under such challenging circumstances, engineers must do their best to use all of the expertise at their disposal to create as perfect a design as possible for the required structure. Despite the uniqueness of the structure, they need to produce a design that will meet the requirements that have been set out, and then, once built, function exactly as predicted on the first try. Such a design process is necessarily a lengthy one, involving many designers with high levels of expertise. Many designs would be proposed and rejected before a final design is selected, based on all the evidence and expertise the design team have at their disposal.

At the end of the process the engineering team will have high confidence in the final design that is put forward. The success of the structure in question is critical, and, as responsible, professional engineers, they would not put forward a design for such a structure without being entirely certain, to the best of their collective ability, that this structure will not fail.

Despite this confidence, due diligence requires more than the simple assurance (and belief) from the design team that the structure will not fail. It is not enough, from a societal perspective, for the team to simply provide a "vote of confidence:" it also requires the provision of more quantitative information about the failure aspects of the structure. Those responsible for the structure need to be able to determine (and to help the stakeholders understand) what are the structure's necessary and sufficient conditions for failure (and by extension, the conditions for non-failure). To produce these answers they need to be able to quantitatively examine what circumstances the structure might encounter, and under these circumstances, what the probability of failure is.

From an ideal testing point of view, the entire proposed structure would be built many times over to run trials relating to each of the foreseen circumstances. Data would then be gathered and analyzed to determine the failure tolerance of the structure. Failure probabilities would be calculated based on this data, along with an understanding of possible failure circumstances – the structure might even be redesigned to take into account the results of the testing.

However, as we have already noted, this idealistic testing scenario is simply not an option in this case. The structure as a whole cannot be directly tested even once, let alone multiple times. And on top of this, even were many replications of the structure itself available for testing, not all failure circumstances (in particular those involving major geological forces and long time spans) would be possible to re-create in a test environment.

An alternative strategy is centered around a combination of physical testing and modeling of the behaviour of the structure and environment. More specifically, a larger structure is built up of many component parts, which themselves may be built up of many components. The failure parameters of these component parts may be tested, even if the structure as a whole cannot.

Similarly, while the structure itself, and perhaps even in some cases the components themselves, cannot be tested repeatedly, there remains the option of creating models of the structure and components in question, and then using the behaviour of these models to predict the behaviour of the components and, in turn, of the structure at large.

In the absence of the ideal testing scenario, understanding and quantifying the failure of the system as a whole can be carried out by understanding and quantifying the failure circumstances of the components of the system, understanding the causal relationships between these components, creating models of the system as a whole based on these relationships, determining the failure circumstances and probabilities of the constructed structure level models and then transferring these findings over to the structure itself. This results in an estimate of the failure circumstances and probabilities of the actual engineered structure as a whole.

The end result of this exercise will thus be, rather than a simple yes/no statement (such as “No, the structure will not fail”, for instance), a list of the possible failure circumstances and an estimate of the failure probabilities for both the structure components and the structure itself, along with a confidence measure indicating a level of confidence in the failure probabilities calculated for each failure circumstance.

Such a table of failure circumstances, probabilities, and confidence measures will allow those building the structure to open a legitimate dialogue with those responsible for, and those being affected by, the resulting structure. In essence, this deliverable will allow the designers of the structure to provide their stakeholders with a clearer and more detailed picture of the risks they are likely to encounter when undertaking the construction of such a structure.

General Objectives The general objective of this Failure Analysis project as a whole is to estimate the failure probability of the Mark II canister and engineered barrier system immediately surrounding the canister. In order to achieve that larger objective, we anticipate that we will be using a combination of statistical analysis, mathematical modeling, and simulations, much as in this prototype. More specifically, we will take the approach that our model is meant to answer a specific question, as well as to provide outputs that can be fed into other models, as may be required by already-developed NWMO models.

In this prototype phase, however, the objective is to develop a methodology and implementation framework to confirm that interactions (both planned and emergent) can in principle be captured by the modeling process, both at the repository and the manufacturing level. For both the manufacturing process and the interactions models, a specific selection of a small number of sub-components of the entire system will be considered in this phase, in order to maintain focus on the development and testability of the methodology itself.

In the following extract from the report Failure Analysis Simulation Model for the APMRD-II, we discuss some of the strategies that could be used to extract information and knowledge about the engineered barrier system, which could then be incorporated in any interaction model of its components. A discussion of system complexity and the effect it had on our choice of modeling approach is also provided. We also provide a prototype UFC manufacturing process model: potential states, actions and variables are introduced, as well as the underlying modeling assumptions and families of parameters. The model is illustrated *via* a specific parameter set; a series of 8 scenarios showcase the effect of various parameter combinations.

It should be noted that due to the uncertainty relating the manufacturing process parameters, the numbers presented in this section mostly play the roles of placeholders: rea-

sonable estimates for a large number of these parameters will be required before the model can output meaningful failure estimates.

References

[1] Clement, J.J. [2008], Creative Model Construction in Scientists and Students: The Role of Imagery, Analogy, and Mental Simulation, Springer Netherlands, 10.1007/978-1-4020-6712-9

[2] Holyoak, K., Gentner, D., and Kokinov, B. [1998], Introduction: The place of analogy in cognition, In Holyoak, K., Gentner, D., and Kokinov, B. (eds.), Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences, NBU Press, Sofia.

[3] Hofstadter, D.R. [2001], Analogy as the core of cognition, In Gentner, D., Holyoak, K., and Kokinov, B., (eds.), The Analogical Mind: Perspectives from Cognitive Science, The MIT Press Bradford Book, Cambridge MA.

[4] Kopec, J.A., Finès, P., Manuel, D., Buckeridge, D., Flanagan, W.M., Oderkirk, J., Abrahamowicz, M., Harper, S., Sharif, B., Okhmatovskaia, A., Sayre, E.C., Rahman, M.M., and Wolfson, M. [2010], Validation of population-based disease simulation models: A review of concepts and methods, BMC Public Health, 10. 10.1186/1471-2458-10-710.

[5] Landriscina, F. [2013], Simulation and Learning: A Model-Centered Approach, New York, NY: Springer, 10.1007/978-1-4614-1954-9.

[6] Schellinck, J. [2009]. A general perception based framework for modelling animal aggregation, Carleton University, Ottawa.

[7] Yilmaz, L. (ed.) [2015], Concepts and Methodologies for Modeling and Simulation: A Tribute to Tuncer Ören, Springer International Publishing, Switzerland, 10.1007/978-3-319-15096-3.

[8] Pearl, J. [2009], Causality: Models, Reasoning, and Inference (2nd ed.), Cambridge University Press.

[9] Templ, M. [2016], Simulation for Data Science with R, Packt Publishing.

[10] Attribution: By Fourdee (talk) (Uploads) - Own work, Public Domain, https://en.wikipedia.org/w/index.php?curid=8650757

[11] https://en.wikipedia.org/wiki/N-body_problem

[12] Beloriszky, D. [1930], Application pratique des méthodes de M. Sundman à un cas particulier du problème des trois corps, Bulletin Astronomique 6 (series 2), 417âĂŞ434.

[13] Cixin, L. [2008], The Three Body Problem, Chongqing Press.

[14] http://weusemath.org/?didyouknow=new-discoveries/n-body-problem

[15] https://en.wikipedia.org/wiki/N-body_simulation

[16] Momkov, T. [2013], Planetary System Formation 2, $N-$body simulation, http://trekto.info/n-body-simulation

[17] Ross, S. [2009], A 1766 Benjamin Martin Orrery, used at Harvard (photo), https://upload.wikimedia.org/wikipedia/commons/b/bf/Planetarium_in_Putnam_Gallery_2%2C_ 11-24.jpg

[18] https://en.wikipedia.org/wiki/Shoaling_and_schooling

Failure Analysis Simulation Model for the APMRD-II

Jennifer Schellinck[1], Patrick Boily[2], Shintaro Hagiwara[2], Korey MacDougall[2]
[1]Sysabee
[2]Centre for Quantitative Analysis and Decision Support, Carleton University

Original Issue R00: June 30, 2016
Revised Version R01: July 10, 2016

Contents

2 Methodology Considerations

The NWMO has a number of highly detailed models of specific components or aspects of the barrier system (e.g. models of the stability of the rock surrounding the placement rooms, models of the rates of corrosion). However, interactions between components and processes can have significant causal effects on the behaviour of the system over all. If existing models are largely single component focused, these interaction effects may not be fully take into account.

Although, in principle, the creation of a highly detailed full system model, possibly via the amalgamation and expansion of existing component models, might allow for the most comprehensive exploration of unanticipated interactions between system components, the level of detail in a such model must be properly managed to avoid the model become both unwieldy and inaccurate. Rather than resulting in a highly detailed, highly accurate model, acceptable inaccuracies in individual model components may be magnified in an amalgamated system model, resulting in a model that has behaviours which do not realistically reflect the behaviour of the modeled system as a whole. As well, pragmatically speaking, the creation, validation and management of such a highly detailed, large scope model can quickly become infeasible. In these ways model inaccuracy, invalidity and poor-functionality become introduced into the model.

Failure Mode and Effects and Criticality Analysis (FMECA) seeks to take into account interactions between system components by eliciting possible sources and pathways towards failures of the system from subject matter experts (SMEs) via a structured interview process. Experts are also asked to estimate the possibility or likelihood of each of these pathways or failure modes. These estimates may then be used to generate a semi-quantitative failure model for the system.

Although this approach avoids both the issues that can arise from narrowly-scoped models, as well as those that may occur when attempting to create highly detailed full-system models, it also lacks one of the recognized advantages of such models- their ability to produce previously unanticipated results, emergent from the structure of the model. Given this, it would seem a hybrid approach combining the advantages of both of these strategies would be a useful step forward. A number of hybrid approaches have already been proposed (e.g. Baldwin et al. 1995, Eusgeld et al. 2011).

In order to develop a system-wide model of the engineered barrier system that can take into account relevant but possibly unanticipated interactions into account, we have also taken a hybrid approach. The resulting modeling framework incorporates process behaviours of system components (which may be modeled by outputs from existing detailed system component models) within a higher level causal framework that incorporates system event probability. The framework itself is created based on a combination of input from system experts and an analysis of documented system information.

2.1 Data Collection and System Analysis

One goal of the prototype project was to explore strategies for methodically and, when possible, automatically, extracting information about system components and component interactions, from documentation, and then systematically incorporating this extracted knowledge into the model of the system in such a way that previously implicit interactions would be captured and

Key NWMO Documents Provided for Model Creation
APM Repository Design –Mark II Proof Test Plan
Mark II Design Concept for Crystalline and Sedimentary Host Rock
Used Fuel Deep Geological Repository Facility Requirements
Multiple-Barrier System
Thermal Modelling of a Mark II Container
Used Fuel Container Retrieval from a Deep Geological Repository in Crystalline Rock Vertical Borehole Configuration
Adaptive Phased Management Postclosure Safety Assessment of a Used Fuel Repository in Sedimentary Rock Pre-Project Report
Adaptive Phased Management Used Fuel Repository Conceptual Design and Postclosure Safety Assessment in Crystalline Rock Pre-Project Report
Choosing a Way Forward The Future Management of Canada's Used Nuclear Fuel
Implementing Adaptive Phased Management 2015 to 2019
Technical Program for Long-Term Management of Canada's Used Nuclear Fuel – Annual Report 2014
Preliminary Design Report for Used Fuel Container: CV-HH-4L-12
Overview Of The NWMO And The Mark II Used Fuel Container
Assessing Reliability and Useful Life of Containers for Disposal of Irradiated Fuel Waste
Radiation Sources in Nuclear Waste and Their Properties
Overview of the NWMO and the Mark II Used Fuel Container
Used Fuel Deep Geological Repository and Transportation System - Used Fuel Container Preliminary Design Requirements for the APM Update
Adaptive Phased Management Facility Used Fuel Container Preliminary Design Requirements
Long-Term Stability Analysis of APM Mark II Conceptual Design in Sedimentary and Crystalline Rock Settings

Table 1: Barrier system documentation provided by NWMO and used as a basis for the prototype phase models

incorporated into the model behaviour. This structured knowledge would then be further supplemented and verified by system expert knowledge.

Knowledge extraction began by a preliminary pass through the system overview documents provided by NWMO (see Table 1). A high level system component schematic was created based on this review (see Figure 1), with a preliminary review by system experts to confirm that no significant system components had been neglected. The goal behind generating this system component model was to develop an understanding of parts of the system which might possibly interact with each other during system operation.

This system component model was then used to methodically extract and structure information contained in the provided documents. Specifically, facts about the system were tagged with system component labels if they provided information about these system components (see Figure 2). Facts tagged with multiple component labels could then provide support for hypotheses that these system components would potentially interact during operation of the system (see Figure 3). Information was extracted from the documents both manually and automatically, and the results of these two extractions compared. The goal of the automatic extraction was both to validate the manual extraction process and also to provide a test of feasibility of data extraction on a larger scale. The results of the data extraction were then used to generate a conceptual model of the engineered barrier system. Construction of model schematics and implemented models was supported by the structured system data.

Two stand alone models were constructed – the **barrier system component interaction** (causal chain) **model** (selected components) and a detailed model of the **manufacturing process** (see Sections 4 and 3, respectively). The manufacturing process model can also be viewed as a process that provides inputs into the system level model relating to the properties of the UFC.

Simulations provided a relatively straightforward approach to programmatically represent the system, including complex interactions between system components. This came at the expense, however, of requiring exploration of a relatively large parameter space in order to determine system behaviour. The resulting simulation was a discrete time simulation, with the behaviour of

Figure 1: High level schematic of system components and their relationships (see accompanying file for larger figure)

each model element determined at each time step based on the states of elements in the preceding time step. Basic output of the models consisted of the system states and properties at each time-step of the model.

2.2 System Complexity

During the model development phase of this project, research was undertaken to determine which approach would be most suitable for the modeling of the engineered barrier system, generated from the conceptual model of the system that had been developed during the conceptual modeling phase. Broadly speaking, the research involved determining which combination of the two major branches of modeling – simulation models (which for greater clarity we will also refer to as programmatic models) and mathematical models – should be used. It is worth noting here that this distinction can be confusing and imprecise, as both types of models frequently employ computer code and mathematical equations. As well, it is possible for a given model to combine both methods. Nonetheless, there are some key differences in these methods.

Internal residual gases

Hierarchy

- Deep geological waste management system
  - Deep geological repository facility
    - Underground structures and facilities
      - Repository
        - Placement rooms
          - Buffer box
            - UFC
              - **Internal residual gases**

Summary

	Total	Qualitative	Quantitative
Number of sentences:	4	4	0

Data

DOC1:0195:J [Internal Radiation Effects]
Within an extensive study of atmospheric conditions, both containers filled with inert gases and atmospheric gases will be assessed within and external to radiation environments to assess the susceptibility of these containers to SCC, crevice corrosion, localized corrosion (pitting) and general corrosion processes.

DOC1:0196:J [Internal Radiation Effects]
Radiation may inhibit these processes (i.e. by enhancing general corrosion), or it may exasperate the situation by producing SCC agents or iron oxides within the crevice.

DOC1:0197:J [Internal Radiation Effects]
As the internal volume is relatively small, mass balance scenarios will be used to bound damage functions for these scenarios, and a combination of experimental analysis and steady-state radiolysis kinetic models will be used to define actual internal conditions.

DOC1:0198:J [Internal Radiation Effects]
This task will contribute to SCC and crevice corrosion programs ___(above); the terminal goal is to define internal corrosion allowances for each of these processes for the container. Issues within this topic have been identified as minor.

Figure 2: Screenshot of structured database of system component facts (see accompanying files for complete database).

In the case of mathematical modeling, sets of equations are developed to describe the system, and then these sets are solved to find, effectively, all (or any) possible solutions, which then define all possible behaviours of the system. In the case of programmatic models, the behaviour of the system cannot be deduced by 'solving' the system. Rather system behaviour must be simulated, piecemeal, by setting up specific initial conditions in the system, simulating the behaviour of the system at each moment in time over a defined period of time and then reading off the results of the simulation at the end point to determine the state of the system at that end point, given those starting conditions and that passage of time.

Both approaches have strengths and weaknesses. Representing a system mathematically allows for highly generalizable conclusions to be drawn about the system. However as the mathematical representation of the system becomes more complicated, as perhaps is necessitated by the underlying complexity of the system itself, solving the mathematical system may become impractical or no more efficient than the simulation approach.

Simulations provide a relatively straightforward approach to programmatically representing the system, including complex interactions between system components, but analysis of the behaviour of the simulation is often challenging and only inductive conclusions about the system as a whole can be drawn based on the behaviour of system under specified circumstances. As well, running the system over a lengthy period of time in order to generate outcomes may be prohibitive from a computational cost point of view.

The overall conclusion from research carried out during the prototype project was that the simulation approach was most suitable for modeling the engineered barrier system, due to the fact that a major emphasis of this model was incorporating interaction effects between system components. The resulting simulation was a discrete time simulation, with the behaviour of each model element determined at each time step based on the states of elements in the preceding time step. Outputs of the model consisted of the model element states at each time step.

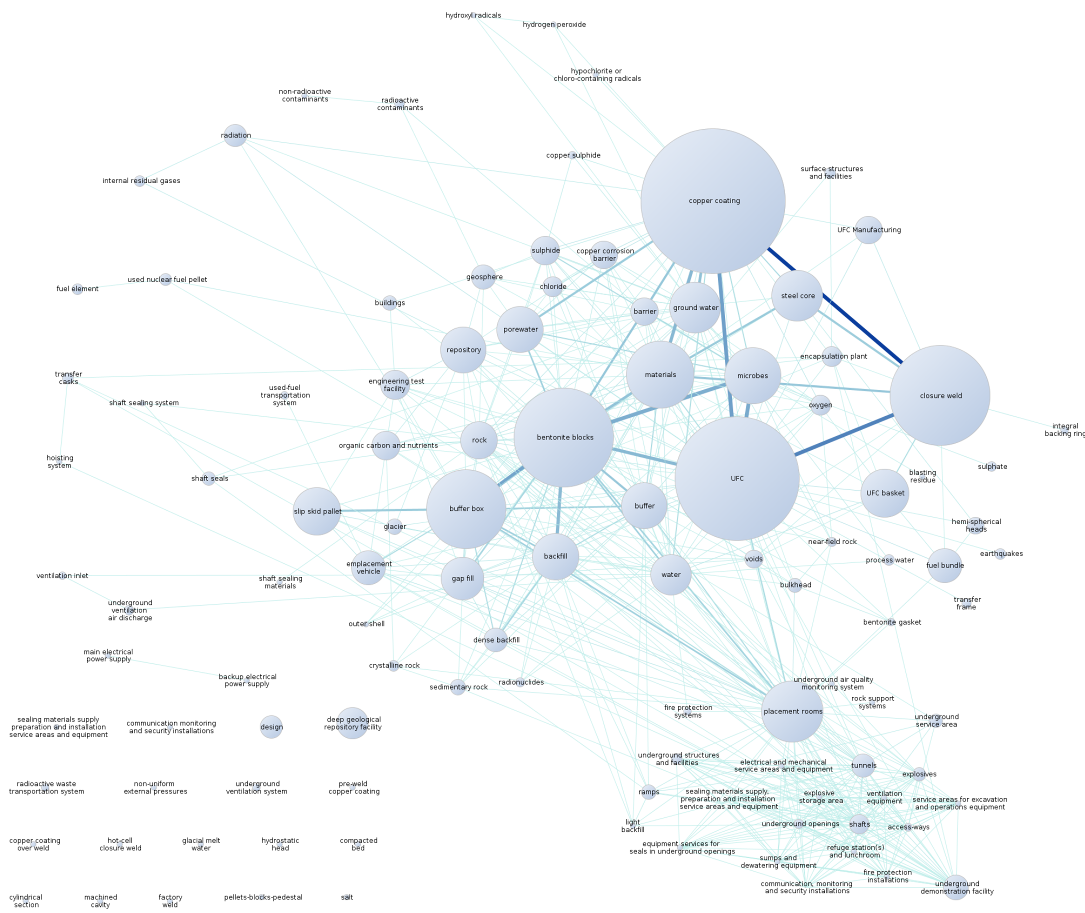


Figure 3: Bubble graph showing objects that were connected based on structured information extraction (see accompanying file for larger figure)

To deal appropriately with the complexity of the system while maintaining a manageable model, the model framework allowed for the ability to vary the level of detail of implementation, allowing for a very high abstraction representation of system, moving all the way to possible incorporation of very highly detailed models, all within the same implemented structure. Detail could be added or subtracted if it were determined that representing interactions required either a greater level of detail, or if components of the model could realistically be further simplified. Some structural elements were also incorporated in order to allow for the exploration of counterfactual scenarios- (i.e. in the absence of knowledge about what could cause a particular event to happen, it is still possible to consider the effects if it were to happen at a chosen point in time, and determine the potential consequences to the system).

References

- Baldwin, James F., Martin, Trevor P. and Athena Tocatlidou (1995). Uncertainty management in radioactive waste repository site assessment. Fuzzy Sets and Systems Volume 74(1), pp. 81-91.

- Eusgeld, I., Nann, C. and S. Dietz (2011). 'System-of-systems' approach for interdependent critical infrastructures. Reliability Engineering and System Safety 96, pp. 679-686.

UFC State	Component	Description
A1	Steel Canister	poor mechanical properties (tensile and fracture toughness)
A2	Steel Canister	physical dimensions outside of acceptable range
A3	Steel Canister	poor surface finish of steel
A4	Steel Canister	defective weld (LW)
A5	Copper Coating	poor ductility and adhesion of copper coating (LH,SH,UH,LW)
A6	Copper Coating	depth of copper coating outside of acceptable range (LH,SH,UH,LW)
A7	Steel Canister	defective weld (CW)
A8	Copper Coating	poor adhesion of copper coating (CW)
A9	Copper Coating	depth of copper coating outside of acceptable range (CW)
A10	Copper Coating	thin wall and through wall defects (minimum coating thickness: LH,SH,UH,LW,CW)

Table 2: Manufacturing process model – States

3 Prototype Modeling of the UFC Manufacturing Process

Throughout, we assume that N containers are taken through the manufacturing process, independently of one another.

The UFC Manufacturing Process model requires inputs in four categories:

- **Model structure**: states, actions, essential variables, parent/child links
- V **parameters**: parameter values and ranges to determine the essential variable output values for each state
- A **parameters**: threshold values which determine the state value for each state
- R **parameters**: parameter values and ranges which determine whether unacceptable state values are accepted or rejected by the various tests.

There are 2 main applications for the prototype's use: exploring the parameter space, and running multiple simulations with a single set of parameters to determine intrinsic variability (which could be used to indicate the absence of crucial variables). In this report, we will mostly focus on the second application.

3.1 Manufacturing States, Actions, and Variables

The prototype model consists of 10 states (see Table 2), characterized by 50 variables (see Table 4), and linked according to the schematics shown in Figure 4. The UFC is assembled state-by-state (and step-by-step), following a series of manufacturing actions (both external and internal); after certain specific actions, some non-destructive examinations of the UFC components are conducted and non-compliant components are re-sent up the chain for repairs or, presumably, to be removed from the process if they have suffered irreparable damage at a prior stage (see Table 3 for a list and Figure 5 for a visual representation).

Actual state values (denoted by A_i throughout) are given **as of the last time a given state is tested during the manufacturing process**. Consequently, the value of A_i may be recorded after any number of times the UFC has been re-sent for repairs; this number includes the possibility that the UFC is never sent for repair at a given stage.

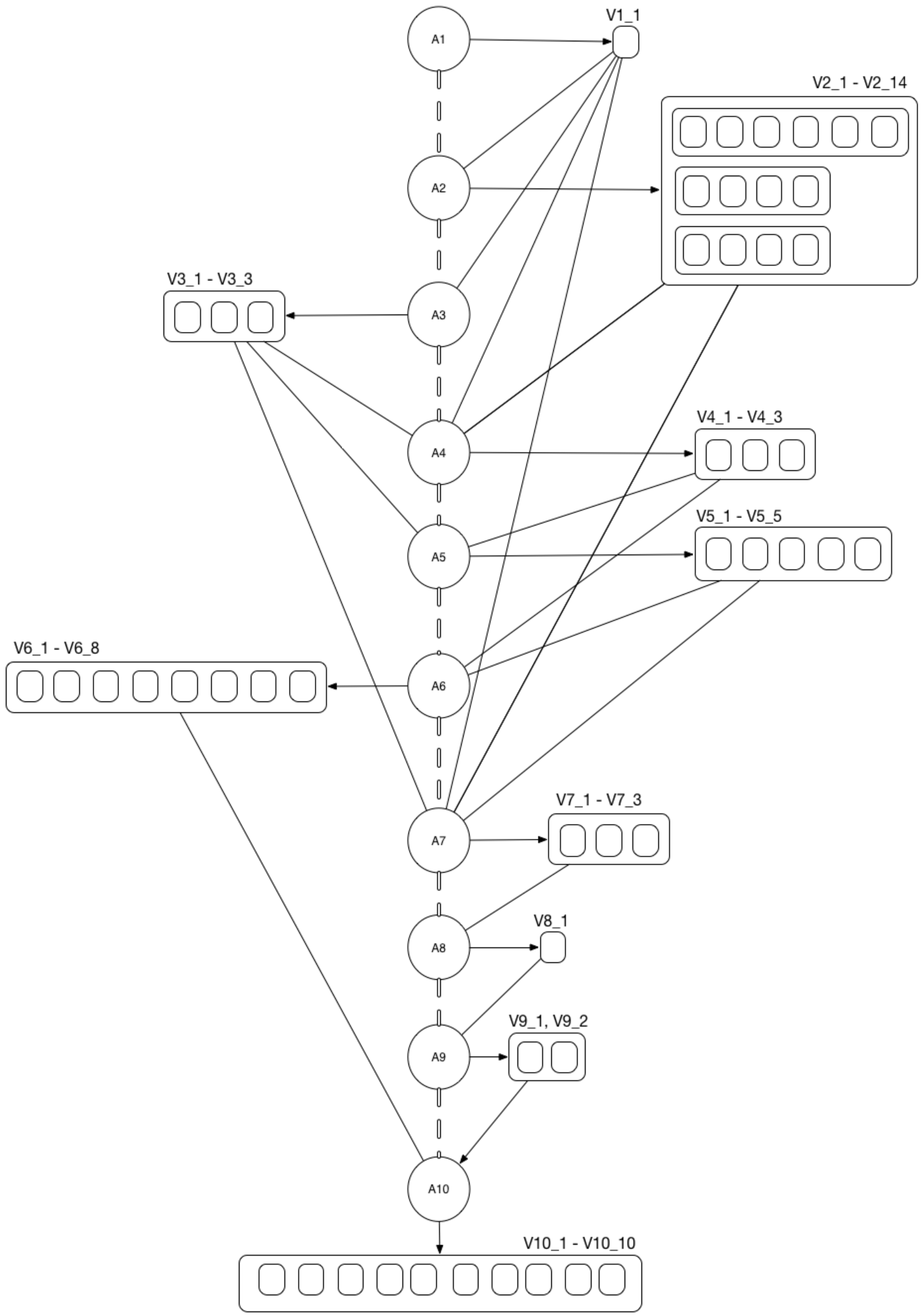


Figure 4: Manufacturing process model – Parent/child schematic links

Order	Description	Type
1	Test of UFC steel tensile and fracture toughness	Test
2	Machining of shell/head components for assembly	Action
3	Welding of Lower Assembly	Action
4	Machining of Lower Assembly Weld Cap	Action
5	NDE Test and repair of Lower Assembly Weld Zone	Test
6	Copper coating of Lower Assembly and Upper Head via Electrodeposition	Action
7	Machining of Copper Coated Surfaces	Action
8	NDE Test and repair of Copper Coated Surfaces	Test
9	Closure Welding after Fuel Loaded	Action
10	Machining of Closure Weld Cap	Action
11	NDE Test and repair of Closure Weld Zone	Test
12	Copper Coating of Closure Weld Zone via Cold Spray	Action
13	Annealing (heat treating) of Copper Coating at Weld Zone	Action
14	Machining of Copper Coating at Weld Zone	Action
15	NDE Test and repair of Copper Coating at Closure Weld Zone	Test

Table 3: Manufacturing process model – Actions and tests

A value of 1 corresponds to a “bad” state outcome (that is, at least one of the state’s essential variables falls outside its acceptable range), a value of 0 corresponds to a “good” state outcome (all of the state’s essential variable falls within their acceptable ranges). Recorded state values (R_i) follow the same valuation scheme.

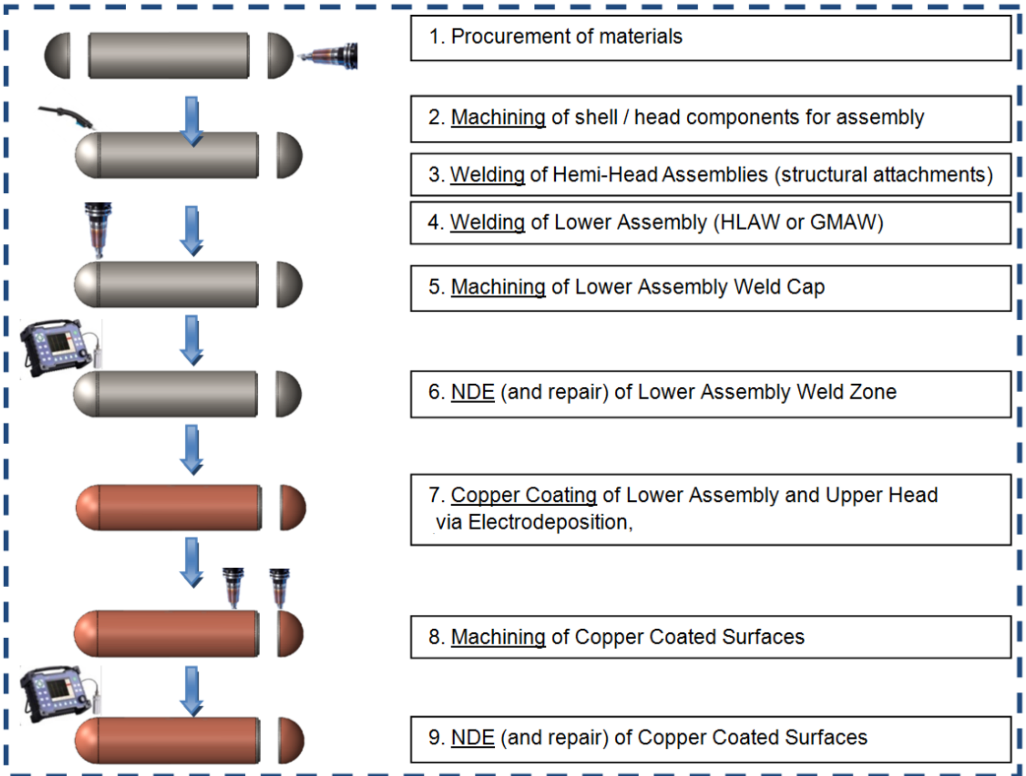
Each container state, then, belongs to one of 4 categories:

- $A = 0$ and $R = 0$ represent UFCs which were correctly retained
- $A = 0$ and $R = 1$ represent UFCs which were incorrectly removed from the process
- $A = 1$ and $R = 0$ represent UFCs which were incorrectly retained
- $A = 1$ and $R = 1$ represent UFCs which were correctly removed from the process

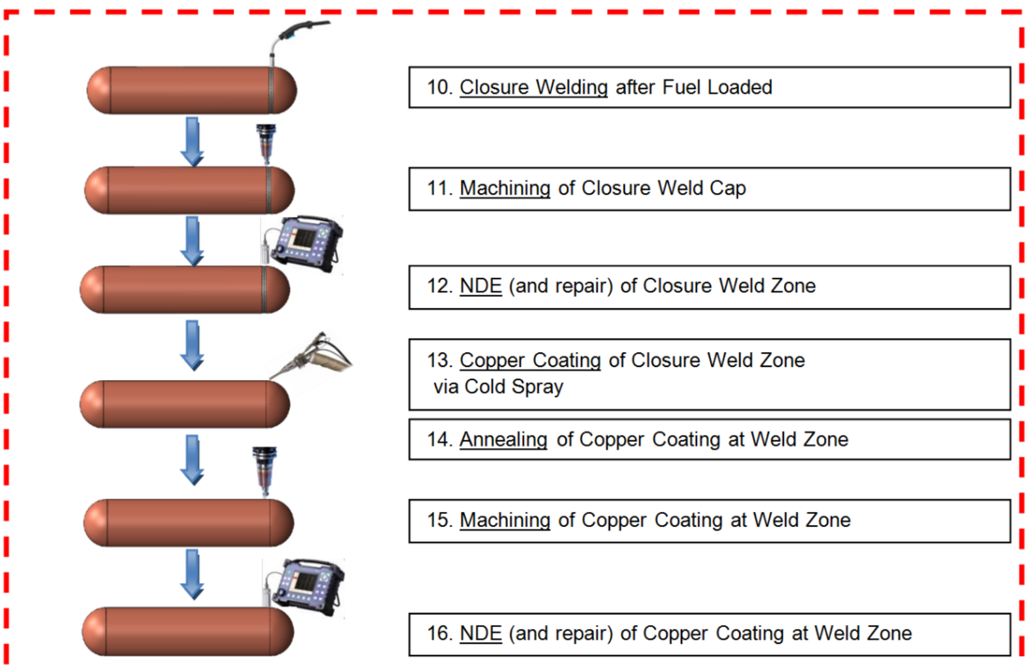
The number of UFCs which are removed from the process (whether correctly or incorrectly) may be important to the NWMO (after all, manufacturing a UFC is both financially and temporally costly), but from the point of view of the failure analysis, that figure is not relevant; the emphasis is devoted to estimating conditional probabilities

$$P(A = 1|R = 0) = \frac{P(A = 1, R = 0)}{P(R = 0)},$$

where A and R correspond to any pair of actual and recorded states in the manufacturing process. Of special interest will be the state 10 probabilities of sending a UFC to the repository with a **through-wall** or a **thinwall** defect.



(a) Off-site operations



(b) Radiological operations

Figure 5: Manufacturing process model – Visual representation (from *Overview of the NWMO and the Mark II Used Fuel Container*, C. Boyle)

.

Label	Description	Variable Type	Target Values	Distribution
V1	tensile and fracture toughness	Strength	Unknown, but presumably exists	Normal
V2_1	minimal length of shell (SH) component	Length	Given by specs	Normal
V2_2	maximal length of SH component	Length	Given by specs	Normal
V2_3	minimal thickness of SH component	Length	Given by specs	Normal
V2_4	maximal thickness of SH component	Length	Given by specs	Normal
V2_5	minimal radius of SH component	Length	Given by specs	Normal
V2_6	maximal radius of SH component	Length	Given by specs	Normal
V2_7	minimal thickness of lower head (LH) component	Length	Given by specs	Normal
V2_8	maximal thickness of LH component	Length	Given by specs	Normal
V2_9	minimal radius of LH component (at the SH joint)	Length	Given by specs	Normal
V2_10	maximal radius of LH component (at the SH joint)	Length	Given by specs	Normal
V2_11	minimal thickness of upper head (UH) component	Length	Given by specs	Normal
V2_12	maximal thickness of UH component	Length	Given by specs	Normal
V2_13	minimal radius of UH component (at the SH joint)	Length	Given by specs	Normal
V2_14	maximal radius of UH component (at the SH joint)	Length	Given by specs	Normal
V3_1	surface finish variable for LH	Roughness	Unknown	Normal
V3_2	surface finish variable for SH	Roughness	Unknown	Normal
V3_3	surface finish variable for UH	Roughness	Unknown	Normal
V4_1	minimal thickness of lower assembly weld (LW)	Length	Given by specs	Normal
V4_2	minimal thickness of LW	Length	Given by specs	Normal
V4_3	brittleness of LW material	Brittleness	Unknown, but presumably exists	Normal
V5_1	ductility of copper used in coating	Ductility	Unknown, but presumably exists	Normal
V5_2	adhesion strength of LH copper coating	Unknown	Unknown	Normal
V5_3	adhesion strength of SH copper coating	Unknown	Unknown	Normal
V5_4	adhesion strength of UH copper coating	Unknown	Unknown	Normal
V5_5	adhesion strength of LW copper coating	Unknown	Unknown	Normal
V6_1	minimal depth of LH copper coating	Length	Given by specs	Folded Normal?
V6_2	maximal depth of LH copper coating	Length	Given by specs	Folded Normal?
V6_3	minimal depth of SH copper coating	Length	Given by specs	Folded Normal?
V6_4	maximal depth of SH copper coating	Length	Given by specs	Folded Normal?
V6_5	minimal depth of UH copper coating	Length	Given by specs	Folded Normal?
V6_6	maximal depth of UH copper coating	Length	Given by specs	Folded Normal?
V6_7	minimal depth of LW copper coating	Length	Given by specs	Folded Normal?
V6_8	maximal depth of LW copper coating	Length	Given by specs	Folded Normal?
V7_1	minimal thickness of closure weld (CW)	Length	Given by specs	Normal
V7_2	minimal thickness of CW	Length	Given by specs	Normal
V7_3	brittleness of CW material	Brittleness	Unknown, but presumably exists	Normal
V8	adhesion strength of CW copper coating	Unknown	Unknown	Normal
V9_1	minimal depth of CW copper coating	Length	Given by specs	Folded Normal
V9_2	maximal depth of CW copper coating	Length	Given by specs	Folded Normal
V10_1	number of through wall spots on the LH	Number	Unknown, but presumably exists	Poisson?
V10_2	number of thin wall spots on the LH	Number	Unknown, but presumably exists	Poisson?
V10_3	number of through wall spots on the SH	Number	Unknown, but presumably exists	Poisson?
V10_4	number of thin wall spots on the SH	Number	Unknown, but presumably exists	Poisson?
V10_5	number of through wall spots on the UH	Number	Unknown, but presumably exists	Poisson?
V10_6	number of thin wall spots on the UH	Number	Unknown, but presumably exists	Poisson?
V10_7	number of throughwall spots on the LW	Number	Unknown, but presumably exists	Poisson?
V10_8	number of thinwall spots on the LW	Number	Unknown, but presumably exists	Poisson?
V10_9	number of throughwall spots on the CW	Number	Unknown, but presumably exists	Poisson?
V10_10	number of thinwall spots on the CW	Number	Unknown, but presumably exists	Poisson?

Table 4: Manufacturing process model – Variables

A number of simplifying assumptions have been made regarding the distributions from which the various values are drawn; these may need to be changed once better information becomes available to the modelers or to the NWMO.

3.2 Modeling Assumptions and Parameters

Our approach is to model the statess stochastically rather than modeling them physically after each of the actions. We also assume that states are only affected by a subset of the actions and tests, and so the various state parameters have to reflect those. The dependencies are listed below:

UFC STATE 1 – poor mechanical properties of UFC steel (tensile and fracture toughness)

- Action: None
- Test: UFC steel tensile and fracture toughness

UFC STATE 2 – dimensions of UFC components outside of acceptable ranges

- Action: Machining of shell/head components for assembly
- Test: NDE and repair of Lower Assembly Weld Zone

UFC STATE 3 – poor surface finish of UFC steel

- Action: Machining of shell/head components for assembly
- Test: NDE and repair of Lower Assembly Weld Zone

UFC STATE 4 – defective weld of UFC Lower Assembly

- Actions: Welding of Lower Assembly; Machining of Lower Assembly Weld Cap
- Test: NDE and repair of Lower Assembly Weld Zone

UFC STATE 5 – poor ductility and adhesion of UFC copper coating (LH,SH,UH,LW)

- Action: Copper coating of Lower Assembly and Upper Head via Electrodeposition
- Test: NDE and repair of Copper Coated Surfaces

UFC STATE 6 – depth of copper coating outside of acceptable range (LH,SH,UH,LW)

- Action: Machining of Copper Coated Surfaces
- Test: NDE and repair of Copper Coated Surfaces

UFC STATE 7 – defective weld of UFC closure zone (CW)

- Actions: Closure Welding after Fuel Loaded; Machining of Closure Weld Cap
- Test: NDE and repair of Closure Weld Zone

UFC STATE 8 – poor adhesion of UFC copper coating (CW)

- Actions: Copper Coating of Closure Weld Zone via Cold Spray; Annealing (heat treating) of Copper Coating at Weld Zone
- Test: NDE and repair of Copper Coating at Closure Weld Zone

UFC STATE 9 – depth of copper coating outside of acceptable range (CW)

- Action: Machining of Copper Coating at Weld Zone
- Test: NDE and repair of Copper Coating at Closure Weld Zone

UFC STATE 10 – thin-wall and through-wall defect (minimum coating thickness)

- Action: None
- Test: None

3.2.1 Modeling Procedure

The modeling procedure for each state follows the same steps:

1. the state variables V_i are generated according to the appropriate parameters, the parent states A_j and essential variables V_j;
2. the actual state A_i (and its sub-states, corresponding to each of the variables) are updated according to the appropriate parameters, and
3. the recorded state R_i (and its sub-states) are generated according to the appropriate parameters.

The specifics of the parent/child relationships are listed in the last 2 columns of Table 5.

3.2.2 Effect of Undetected Flaws in Parent States

In the prototype, an undetected flaw at a parent stage **can only affect the process used to generate state variable values** – it does not affect the process by which the state is recorded.

For UFCs which were deemed acceptable at a given state ($R_i = 0$), we distinguish between 2 cases:

- UFCs which were correctly recorded as acceptable ($A_i = 0$ and $R_i = 0$), and
- UFCs which were incorrectly recorded as acceptable ($A_i = 1$ and $R_i = 0$).

Recall that the UFCs for which the recorded state were not deemed acceptable ($R_i = 1$) have been removed from the manufacturing process; what follows does not apply to them.

Assume that, when the parent states are such that $A = 0$ and $R = 0$, the variable V follows a distribution $\mathscr{D}(\mu, \sigma)$ (not necessarily normal) with mean μ and standard deviation σ (see leftmost target illustration in Figure 6).

How should the distribution of values be affected, however, if the parent states incorrectly identified the UFC as acceptable? For $k, \ell > 0$, we answer the question with the help of the distribution $\mathscr{D}(k\mu, \ell\sigma)$. If the flaw at the parent state is:

- unlikely to change the accuracy while changing the precision, we use $k = 1, \ell \neq 1$ (see third target diagram);
- likely to change the accuracy while preserving the precision, we use $k \neq 1, \ell = 1$ (see second target diagram);
- likely to change both the accuracy and the precision, use $k, \ell \neq 1$(see fourth target diagram).

State	Affected	Child Mean	Child SD	Child V			Parent A, R											
1				V1														
2	mean	mu2_i	sigma2	V2_1	...	V2_14	A1	R1										
3	mean	mu3	sigma3	V3_1	V3_2	V3_3	A1	R2										
4	sd	mu4_12	sigma4_12	V4_1	V4_2		A1	A2_1	...	A2_6	A2_7	...	A2_10	A3_1	A3_2	R3		
		mu4_3	sigma4_3	V4_3														
5	sd	mu5_1	sigma5_1	V5_1			A3_1	A3_2	A3_3	A4_1	A4_2	A4_3	R4					
		mu5_2345	sigma5_2345	V5_2	...	V5_5												
6	mean	mu6	sigma6_135	V6_1			A5_1	A5_2	R5									
				V6_3			A5_1	A5_3	R5									
				V6_5			A5_1	A5_4	R5									
			sigma6_7	V6_7			A4_1	A4_2	A4_3	A5_1	A5_5	R5						
7	sd	mu4_12	sigma4_12	V7_1	V7_2		A1	A2_1	...	A2_6	A2_11	...	A2_14	A3_1	A3_2	A5_3	A5_4	R6
		mu4_3	sigma4_3	V7_3														
8	sd	mu5_2345	sigma5_2345	V8			A7_1	A7_2	A7_3	R7								
9	mean	mu6	sigma6_7	V9			A5_1	A7_1	A7_2	A7_3	A8	R8						
10				A10_1			R9	V6_1										
				A10_2			R9	A6_1	A10_1									
				A10_3			R9	V6_3										
				A10_4			R9	A6_3	A10_3									
				A10_5			R9	V6_5										
				A10_6			R9	A6_5	A10_5									
				A10_7			R9	V6_7										
				A10_8			R9	A6_7	A10_7									
				A10_9			R9	V9_1										
				A10_10			R9	A9	A10_9									

Table 5: Manufacturing process model – Specifics of parent/child relationship.

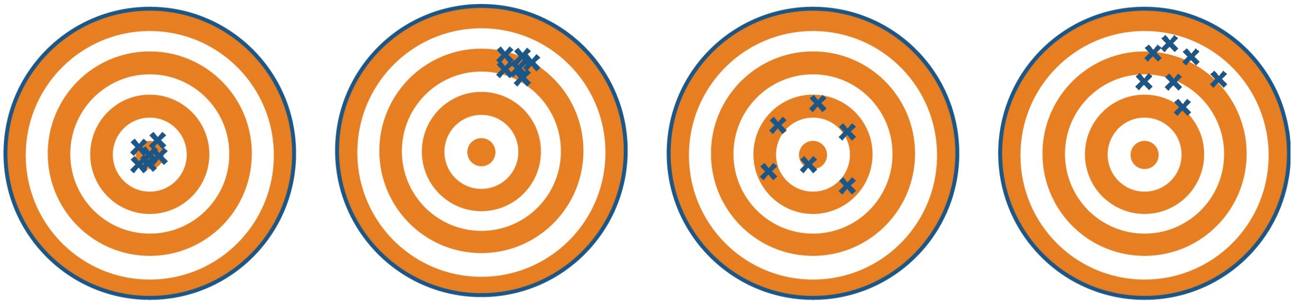


Figure 6: Illustration of the effect of flaw in parent state on a generic variable in a children state for the prototype. From left to right: no effect; mean only; variance only; both mean and variance.

Of course, the parent flaw could affect the variable V in different ways, such as changing the distribution type altogether, instead of simply modifying the distribution parameters), or being dependent on the "strength" with which a parent state fails. While other options can be implemented, we only consider situations given by the first, second or third diagram in this prototype model.

While this has the added benefit of not increasing the number of required input parameters, the main reason we are not considering cases described by the fourth diagram is that we have as yet very little information as to how exactly k, ℓ would be derived in each case. This state of affairs is surely temporary, however, and we have left some place holder parameters in the prototype to be able to make modifications simply in future iterations of the model.

The states for which only the mean is affected by defective parent states are labeled 'mean' in the second column of Table 5, while those for which it is the standard deviation that is affected are labeled 'sd'.

3.3 Illustration of the Method

We illustrate the model with the help of a step-by-step interpretation of the input parameters. In the tables of this section, yellow rows correspond to parameters that are used to generate variable values, orange rows to parameters that are used to determine actual states, and green rows to parameters that are used to determine whether a UFC is removed from the chain, or accepted and passed along to the next stage.

As discussed above, the process is fairly similar from one step to the next; there are, however, various technical details which vary with each state. We shall only present these details when they are unique.

Furthermore, the level of certainty is not the same for each parameter and variable combination. For instance, the specs for the UFC dimensions are known, but the standard deviation of the lengths of the various component are unknown; the mean of the steel tensile strength is not currently known but could be easily estimated; same goes for the tolerance of the various tests, and so forth. **The parameter values listed in this section are thus not to be taken as realistic values in general; their sole purpose is to illustrate how the model works.**

State 1

The parameter values (with descriptions) for state 1 are shown in Table 6.

V Parameters

- Inputs: μ_1, σ_1
- Output: $V_1 \sim N(\mu_1, \sigma_1^2)$

A Parameters

- Input: ρ_1
- Output: $A_1 = \begin{cases} 0 & \text{if } V_1 > \rho_1 \\ 1 & \text{else} \end{cases}$

R Parameters

- Inputs: tol_1, prot_1, $p_{1,1}$, $p_{1,2}$
- Temporary:
  - acceptable region AcReg_1: (ρ_1, ∞)
  - tolerance region TolReg_1: $(\rho_1 - \text{tol}_1, \rho_1)$
  - probabilities: $P(R_1 = 1|V_1) = \begin{cases} 1 - \text{prot}_1 & \text{if } V_1 \in \text{AcReg}_1 \\ (1 - p_{1,2}) \times \text{prot}_1 & \text{if } V_1 \in \text{TolReg}_1 \\ (1 - p_{1,1}) \times \text{prot}_1 & \text{else} \end{cases}$
- Output: R_1

Notes

- ρ_1 is known with certainty. The other parmater values are assumed to be reasonable, but experiments will have to be conducted to determine valid estimates.

State 2

The parameter values (with descriptions) for state 2 are shown in Table 7.

V Parameters

- Inputs: outer radius, head thickness, shell length, shell thickness, σ_2, f_2
- Transition:
  - $\mu_1 =$ shell length
  - $\mu_3 =$ shell thickness
  - $\mu_5 =$ head outer radius
  - $\mu_7 =$ head thickness

- $\mu_9 =$ outer radius
- $\mu_{11} =$ head thickness
- $\mu_{13} =$ outer radius

- Outputs: $V_{2,2j-1:2j} \sim (1-A_1)N(\mu_{2j-1}, \sigma_2^2) + A_1 N(\mu_{2j-1}/f_2, \sigma_2^2)$

A Parameters

- Input: threshold$_2$
- Outputs:
  - $A_{2,2j-1} = \begin{cases} 0 & \text{if } V_{2,2j-1} < \mu_{2,2j-1} - \text{threshold}_2 \\ 1 & \text{else} \end{cases}$
  - $A_{2,2j} = \begin{cases} 0 & \text{if } V_{2,2j} > \mu_{2,2j-1} + \text{threshold}_2 \\ 1 & \text{else} \end{cases}$
  - $A_2 = 0$ unless at least one of $A_{2,i} = 1$

R Parameters

- Inputs: tol$_2$, prot$_2$, $p_{2,1} = p_{2,\text{small}}$, $p_{2,2} = p_{2,\text{big}}$
- Transition:
  - acceptable region AcReg$_2$: $(\mu_{2j-1} - \text{threshold}_2, \mu_{2j-1} + \text{threshold}_2)$
  - tolerance region TolReg$_2$: $(\mu_{2j-1} - \text{threshold}_2 - \text{tol}_2, \mu_{2j-1} - \text{threshold}_2) \cup (\mu_{2j-1} + \text{threshold}_2, \mu_{2j-1} + \text{threshold}_2 + \text{tol}_2)$
  - probabilities: $P(R_{2,i} = 1 | V_{2,i}) = \begin{cases} 1 - \text{prot}_2 & \text{if } V_{2,i} \in \text{AcReg}_2 \\ (1 - p_{2,2}) \times \text{prot}_2 & \text{if } V_{2,i} \in \text{TolReg}_2 \\ (1 - p_{2,1}) \times \text{prot}_2 & \text{else} \end{cases}$
- Output: $R_2 = 0$ unless of one of $R_{2,i} = 1$

Notes

- The mean of the variables is divided by f_2 in cases where the parent state was defective: f_2 values smaller than 1 increase the mean, values greater than 1 decrease it. If defective parent states have no effect on the mean, set $f_2 = 1$.

State 3

The parameter values (with descriptions) for state 3 are shown in Table 8. The parameters are as in States 1 and 2, with the distinction that the acceptable and tolerance regions for surface roughness point in the other direction (by comparison with State 1):

- acceptable region AcReg$_3$: $(0, \mu_3 + \text{threshold}_3)$

- tolerance region TolReg$_3$: $(\mu_3 + \text{threshold}_3, \mu_3 + \text{threshold}_3 + \text{tol}_3)$

State 4

The parameter values (with descriptions) for state 4 are shown in Table 9. The parameters are as in States 1, 2, and 3, with the distinction that a defective parent state affects only the standard deviation in a child, and that the various sub-states have different test tolerances and region types:

- $V_{4,1:2} \sim (1-A_1)(1-A_2)(1-A_3)N(\mu_{4,1:2}, \sigma^2_{4,1:2}) + A_1(1-A_2)(1-A_3)N(\mu_{4,1:2}, (f_{4,1,1:2}\sigma_{4,1:2})^2) + \cdots + A_1A_2A_3N(\mu_{4,1:2}, (f_{4,1,1:2}f_{4,2,1:2}f_{4,3,1:2}\sigma_{4,1:2})^2)$
- $V_{4,3} \sim (1-A_1)(1-A_2)(1-A_3)N(\mu_{4,3}, \sigma^2_{4,3}) + A_1(1-A_2)(1-A_3)N(\mu_{4,3}, (f_{4,1,3}\sigma_{4,3})^2) + \cdots + A_1A_2A_3N(\mu_{4,3}, (f_{4,1,3}f_{4,2,3}f_{4,3,3}\sigma_{4,3})^2)$
- $\text{AcReg}_{4,1:2}$ and $\text{TolReg}_{4,1:2}$ are as in State 2, $\text{AcReg}_{4,3}$ and $\text{TolReg}_{4,3}$ as in State 1.

Notes

- The standard deviation of the variables is multiplied by products of f in cases where the parent state were defective: f_4 values smaller than 1 decrease the standard deviation, values greater than 1 increase it. If a defective parent state have no effect on the mean, set its factor to 1.
- The factors affect the standard deviations by combinatorial multiplication: the more parent states are defective, the more terms enter the final factor.
- The values used for $\mu_{4,1:2}$ are close to the expected depth of the welding substrate; but the values for brittleness $V_{4,3}$ have been arbitrarily selected – reasonable values will need to be provided in future iterations of the model.

State 5

The parameter values (with descriptions) for state 5 are shown in Table 10. The parameters are as in States 1 – 4, with no major difference.

Notes

- $A_{5,1}$ is independent of previous states, whereas States $A_{5,2}$ to $A_{5,4}$ depend on the previous states.

State 6

The parameter values (with descriptions) for state 6 are shown in Table 12. The parameters are as in previous states, with a number of major differences:

- copper coating thicknesses are drawn from a folded normal with mean 3 and standard deviations depending on the UFC component of interest (weld vs. non-weld)
- the parameter strength_6 is used to increase the likelihood of pinhole through-wall defects in the copper coating (large values increase the likelihood)

- the parameter $p_{6,0}$ is the probability of detecting a pinhole through-wall defect (as opposed to detecting a thin-wall defect).

- the parameter threshold_6 can be used to determine what constitutes a thin-wall defect

- the tolerance region recognizes 0 as a copper coating thickness of special importance:

$$P(R_{6,i}=1|V_{6,i}) = \begin{cases} 1-\text{prot}_6 & \text{if } V_{6,i} \in \text{AcReg}_6 \\ (1-p_{6,2})\times\text{prot}_6 & \text{if } V_{6,i} \in \text{TolReg}_6 \\ (1-p_{6,0})\times\text{prot}_6 & \text{if } V_{6,i} = 0 \\ (1-p_{6,1})\times\text{prot}_6 & \text{else} \end{cases}$$

State 7

The parameter values (with descriptions) for state 7 are shown in Tables 9 and 11. The parameters are exactly those of State 4, but extra parameters have been provided in case it turns out that the copper coating on the shell and on the upper head have an effect on the closure weld.

State 8

The parameter values (with descriptions) for state 8 are shown in Table 10. The parameters are exactly those of State 5.

State 9

The parameter values (with descriptions) for state 9 are shown in Table 12. The parameters are exactly those of State 6.

State 10

A Parameters

- Odd csub-states represent the presence of a through-wall defect on each of the 5 UFC components; even sub-states represent the presence of thin-walls on each of the 5 UFC components.
- The thin-wall level is controlled by the value of threshold_6.

R Parameters

- All containers getting to this point are assumed to be without defect, thus $R_{10,j}=0$, independently of the actual state of the container.

Parameter	Value	Description
mu1	300	mean of steel tensile and fracture strengths
sigma1	10	standard deviation of steel tensile and fracture strength
tol1	5	tolerance of test to detect tensile and fracture strength
prot1	0.995	probability that Test will correctly accept tensile strength within the acceptable region
p1_1	0.001	probability that Test will incorrectly accept length outside of tolerance region
p1_2	0.002	probability that Test will incorrectly accept tensile strength within the tolerance region

Target	
rho1	275
temp1_1	0.04
temp1_2	0.02

Table 6: Parameters for State 1 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
threshold2	2	absolute threshold for lengths
sigma2	0.5	standard deviation of lengths around target
factor2	1.05	correcting factor for UFC state 2 with defective A1
tol2	0.7	tolerance of test to detect absolute lengths
prot2	0.995	probability that NDE Test will correctly accept length within the acceptable region
p2_small	0.001	probability that NDE Test will incorrectly accept length outside of tolerance regions
p2_big	0.002	probability that NDE Test will incorrectly accept length within the tolerance regions

Targets	
head outer radius	279.4
head thickness	36.116
shell length	1881
shell thickness	41.275
min_target	36.116
temp2_1	0.1
strength2	1

Table 7: Parameters for State 2 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
threshold3	0.5	absolute threshold for surface finish roughness
sigma3	0.2	standard deviation of surface finish roughness around target
factor3	0.95	correcting factor for state 3 with defective A1
tol3	0.05	tolerance of test to detect surface finish roughness
prot3	0.995	probability that NDE Test will correctly accept surface finish roughness within the acceptable region
p3_small	0.001	probability that NDE Test will incorrectly accept surface finish roughness outside the tolerance region
p3_big	0.002	probability that NDE Test will incorrectly accept surface finish variable within the tolerance region

Targets	
mu3	2.7
temp3_1	0.1
temp3_2	0.02

Table 8: Parameters for State 3 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
threshold4_12	1	absolute threshold for weld depths
threshold4_3	5	absolute threshold for weld brittleness
sigma4_12	0.3	standard deviation around target
sigma4_3	1	standard deviation around target
factor4_1_12	1.05	correcting factor for weld depths with defective A1
factor4_1_3	1.05	correcting factor for weld brittleness with defective A1
factor4_2_12	1.05	correcting factor for weld depths with defective A2
factor4_2_3	1.05	correcting factor for weld brittleness with defective A2
factor4_3_12	1.05	correcting factor for weld depths with defective A3
factor4_3_3	1.05	correcting factor for weld brittleness with defective A3
tol4_12	0.1	tolerance of test to detect weld depths
tol4_3	0.2	tolerance of test to detect weld brittleness
prot4	0.995	probability that NDE Test will correctly accept state 4 variables within the acceptable region
p4_12_small	0.001	probability that NDE Test will incorrectly accept weld depths outside the tolerance regions
p4_12_big	0.002	probability that NDE Test will incorrectly accept weld depths within the tolerance regions
p4_3_small	0.001	probability that NDE Test will incorrectly accept weld brittleness 3 outside the tolerance region
p4_3_big	0.002	probability that NDE Test will incorrectly accept weld brittleness within the tolerance region

Targets	
mu4_12	10
mu4_3	30
temp4_1	0.1
temp4_2	0.02

Table 9: Parameters for State 4 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
mu5_1	90	mean of copper ductility
threshold5_2345	5	absolute threshold for adhesion measure of copper coating
sigma5_1	10	standard deviation of copper ductility
sigma5_2345	2	standard deviation of copper adhesion measure
factor5_3_2345	1.05	correcting factor for copper adhesion measure with defective A3 for variables 2,3,4,5
factor5_4_2345	1.05	correcting factor for copper adhesion measure with defective A4 for variables 2,3,4,5
tol5_1	5	tolerance of test to detect ductility
tol5_2345	0.2	tolerance of test to detect adhesion measures
prot5	0.995	probability that NDE Test will correctly accept ductility and adhesion measure within the acceptable region
p5_1_1	0.001	probability that NDE Test will incorrectly accept ductility and adhesion measure outside of tolerance region
p5_1_2	0.002	probability that NDE Test will incorrectly accept ductility and adhesion measure within tolerance region
p5_2345_small	0.001	probability that NDE Test will incorrectly accept adhesion measure outside of tolerance region
p5_2345_big	0.002	probability that NDE Test will incorrectly accept adhesion measure within tolerance region

Targets	
rho5_1	75
mu5_2345	20
temp5_1	0.04
temp5_2	0.02

Table 10: Parameters for State 5 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
factor7_5_12	1.05	correcting factor for weld depths with defective A5
factor7_5_3	1.05	correcting factor for weld brittleness with defective A5
prot7	0.995	probability that NDE test will correctly accept state 7 variables within the acceptable region

Table 11: Parameters for State 7 (yellow – V; orange – A; green – R; white – unused)

Parameter	Value	Description
threshold6	1	absolute threshold for copper coating thickness
sigma6_135	0.15	standard deviation of coating depth on LH, SH, UH
sigma6_7	0.2	standard deviation of coating depth on LW
factor6_4	2	correcting factor for copper coating thickness with defective A4
factor6_5_1	5.001	correcting factor for copper coating thickness with defective A5_1
factor6_5_2345	2	correcting factor for copper coating thickness with defective A5_2345
tol6	0.3	tolerance of test to detect absolute coating depths
prot6	0.995	probability that NDE Test will correctly accept coating depths within acceptable region
p6_small	0.001	probability that NDE Test will incorrectly accept coating depth outside of tolerance region
p6_big	0.002	probability that NDE Test will incorrectly accept coating depths within the tolerance region

Targets	
copper coating	3
p6_0	1
strength6	1.5

Table 12: Parameters for State 6 (yellow – V; orange – A; green – R; white – unused)

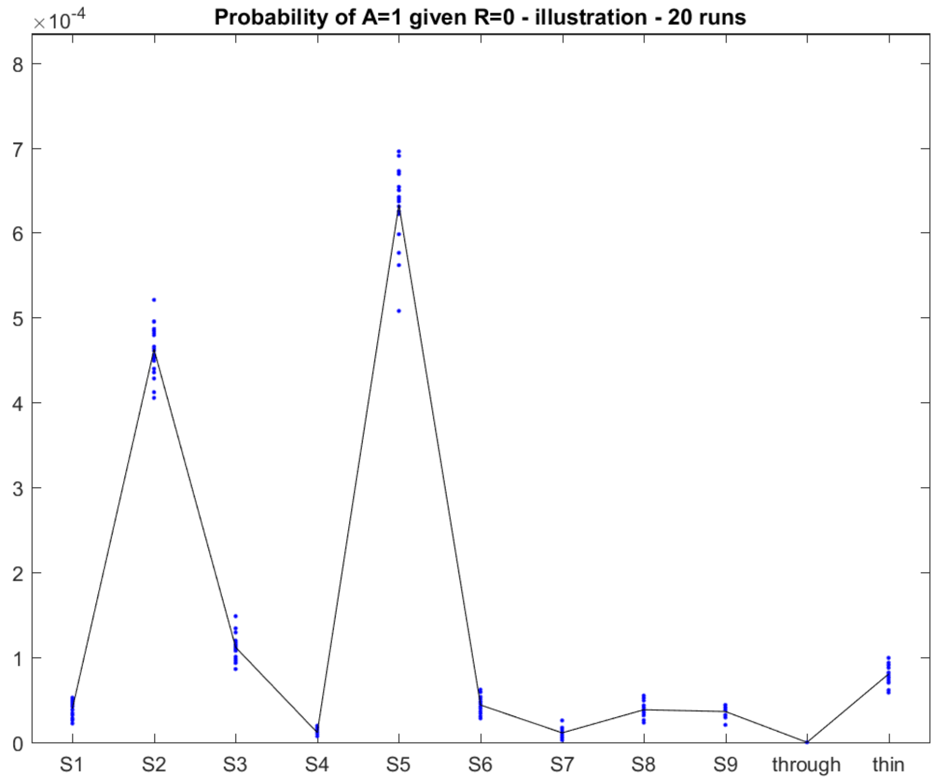
Simulation Results

20 replicates of 500,000 containers have been simulated. The results are shown in Figures 7 and 8. Interesting features include:

- At each of the states 1 to 9, there is a non-zero probability that a defective UFC will have been passed along to the next state.
- The probabilities differ from simulation run to simulation run, but they tend to cluster around specific values, which supports the likelihood that the model is stable for a given parameter set.
- States S_2 and S_5, are substantially more likely to be erroneously accepted by the process, for the given parameter set.
- In none of the simulation were through-wall defects present, although a number of thin-wall defects went undetected.
- At each stage, the probability of a damaged UFC container being sent to the next stage is never more than 7.5×10^{-4}, which seems encouraging.

In hindsight, these results are not entirely surprising, since the probability of capturing a defective UFC are highly correlated with the magnitudes of the various tests' tolerance and the accompanying probabilities in the tolerance regions, as well as with the number of sub-states where something could go wrong (from a detection standpoint).

The graphs in Figure 8 also highlight an important property of the model: the cumulative number of rejected UFCs ($R = 1$) naturally increases at every stage. Since we only care about those containers that eventually will find their way to the repository ($R = 0$), we are looking for high ratios between the first column and second column of each histogram. However, if the combined heights of the last two columns becomes too important, this could be a sign that the probabilities that are used to reject UFCs (defective or not) may be too stringent.



State	Mean	SD	Min	Q1	Median	Q3	Max
1	4.0E-05	9.1E-06	2.2E-05	3.4E-05	4.2E-05	4.7E-05	5.3E-05
2	4.6E-04	2.9E-05	4.1E-04	4.5E-04	4.6E-04	4.8E-04	5.2E-04
3	1.1E-04	1.5E-05	8.6E-05	1.0E-04	1.1E-04	1.2E-04	1.5E-04
4	1.1E-05	3.5E-06	7.3E-06	9.1E-06	1.1E-05	1.3E-05	2.0E-05
5	6.3E-04	4.4E-05	5.1E-04	6.2E-04	6.4E-04	6.6E-04	7.0E-04
6	4.4E-05	9.6E-06	2.8E-05	3.9E-05	4.2E-05	5.1E-05	6.2E-05
7	1.1E-05	5.3E-06	2.9E-06	7.9E-06	1.0E-05	1.4E-05	2.6E-05
8	3.8E-05	8.0E-06	2.3E-05	3.2E-05	3.8E-05	4.4E-05	5.5E-05
9	3.6E-05	6.8E-06	2.0E-05	3.2E-05	3.8E-05	4.1E-05	4.4E-05
Through	0	0	0	0	0	0	0
Thin	8.1E-05	1.1E-05	5.9E-05	7.3E-05	7.9E-05	9.1E-05	9.9E-05

Figure 7: Simulation results of the illustration example, for 20 replicates: the conditional probabilities $P(A_j = 1|R_j = 0)$ and their descriptive statistics are given for each state.

Of course, this is the result of a single run of 20 simulations, **with a single (and arbitrarily selected) parameter set**. Are these results robust? How likely are they to survive a switch to a different parameter set? We attempt to answer some of these questions in the following section. In the meantime, let us urge caution: without a set of reasonable parameter values, a significant amount of parameter space exploration is required before general conclusions can be reached.

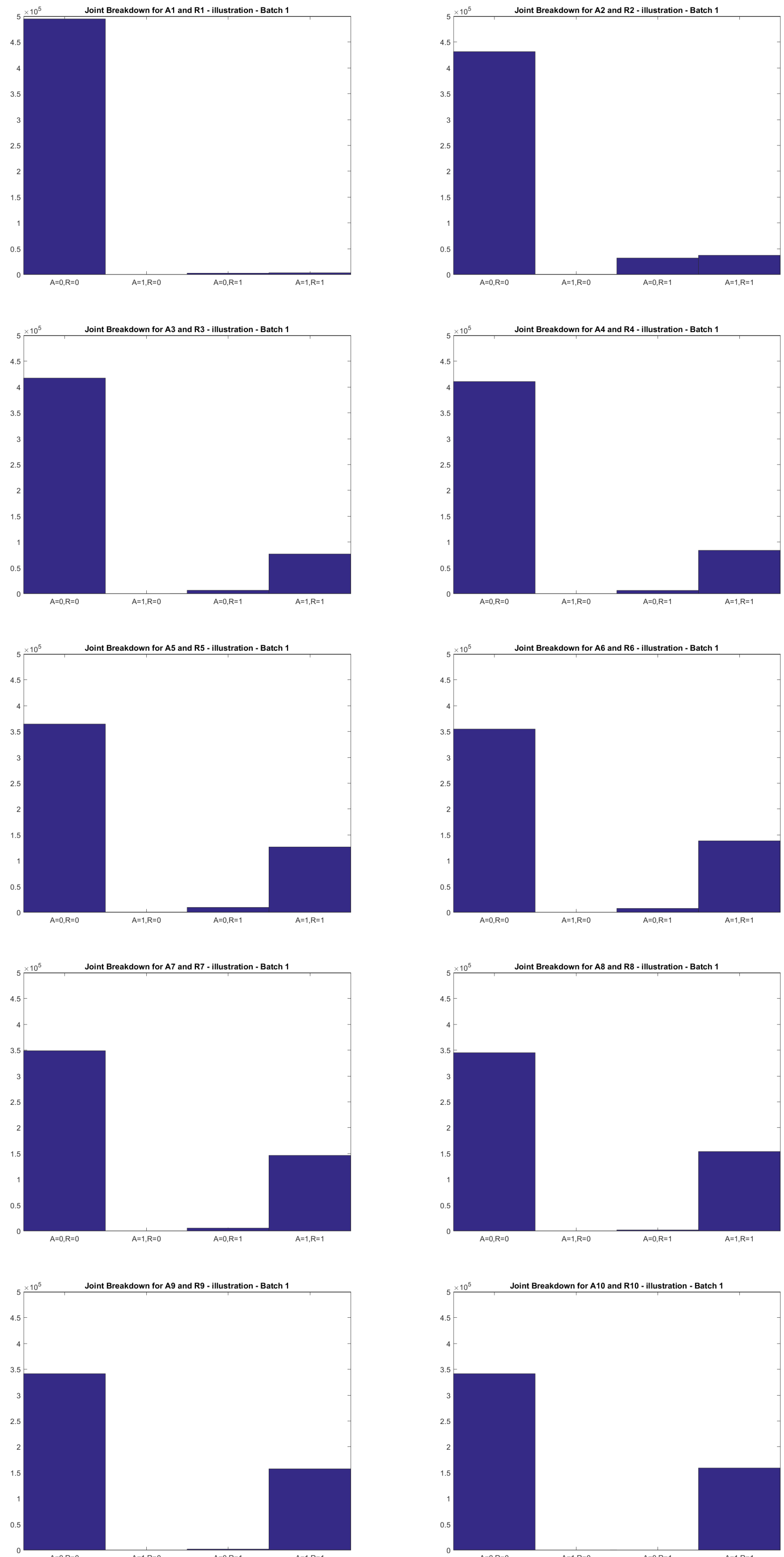


Figure 8: Simulation results of the illustration example, for the first of the 20 replicates: from left to right (within a graph), relative frequencies $P(A_j = 0, R_j = 0)$, $P(A_j = 1, R_j = 0)$, $P(A_j = 0, R_j = 1)$, $P(A_j = 1, R_j = 1)$ for all states.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
2	14	V2_1 to V2_14	0.8	0.7	2	0.875	2.5	2.857142857	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2	0.15	5	0.075	2.5	33.33333333	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
6	3	V6_1, V6_3, V6_5	0.4	0.3	1	0.75	2.5	3.333333333	0.001	0.002	0.995
	1	V6_7	0.4	0.3	1	0.75	2.5	3.333333333	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
9	1	V9	0.4	0.3	1	0.75	2.5	3.333333333	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 13: Parameter set for Scenario 1

3.4 Simulation Results for Eight Scenarios

We now present the results of 8 different simulation scenarios. As before, the emphasis should be taken away from the final probabilities as absolute numbers: rather, we aim to show how varying the parameters affects the final probabilities.

Scenario 1 (Baseline)

The parameter set for Scenario 1 is shown in Table 13; simulation results can be found in Figure 9.

Parameter Set Description:

- σ-to-threshold ratio is set to 2:5 for all states.
- For states 2, 6, and 9, the tolerance-to-threshold ratios vary from 2.9 to 3.3; for the rest of the states, these ratios vary between 5 and 25.
- $p_{i,1}$, $p_{i,2}$, and prot are set to 0.001, 0.002, and 0.995 for all states.
- The factors influencing mean or standard deviation when the parent states are defective are all set to 1.05.
- The tolerance varies from state to state.

This parameter set should give, on average, a value of $P(A_1 = 1|R_1 = 0) = 4.25 \times 10^{-5}$.

Expected Results:

- Since the ratio of σ-to-threshold is held constant at 2:5 throughout the process, with the average $P(A_1 = 1|R_1 = 0)$ of 4.25×10^{-5}, we expect that probability of undetected through-wall defect to be small for this scenario.
- Since the ratio of σ-to-threshold is held constant, we also expect the conditional probability $P(A_i = 1|R_i = 0)$ at each state to be heavily dependant of the number of variables introduced at each state.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
2	14	V2_1 to V2_14	0.8	0.1	2	0.125	2.5	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2	0.15	5	0.075	2.5	33.33333333	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
6	3	V6_1, V6_3, V6_5	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V6_7	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
9	1	V9	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 14: Parameter set for Scenario 2

Results:

- Given the parameter sets specified above, the shape of $P(A_i = 1|R_i = 0)$ in Figure 9 seems adequate at all states. States 1, 3, 4, 5, 7, and 8 have relatively small probabilities of being mistakenly accepted, and this can be explained by the fact that each of these states has between one to five components, and tolerance-to-threshold ratios are very high.
- On the other hand, States 2, 6, and 9 have large spikes. State 2 has the largest spike, which is explained by it consisting of 14 sub-states. State 6 has a lesser, but still noticeable, spike compared to the rest, which may be due to the low tolerance-to-threshold ratio. It should also be noted that as State 9 uses the same parameters as State 6 with fewer sub-states, it should also show a spike, but with a lesser intensity.
- A direct consequence of having large spikes in states 6 and 9 is that we observe large values for $P(A_{\text{thin}} = 1|R_{\text{thin}} = 0)$ value for the thin-wall. Finally, we also observe small but not negligible values of $P(A_{\text{through}} = 1|R_{\text{through}} = 0)$ for through-wall defects.

Scenario 2

The parameter set for Scenario 2 is shown in Table 14; simulation results can be found in Figure 10 and a comparison with Scenario 1 can be seen in Figure 17.

Parameter Set Description:

- The new parameter set is based on the baseline setting.
- Tolerance-to-threshold ratio for States 2 and 6 are increased to similar levels given in other states. (i.e., At States 2 and 6, the measurement error become smaller compared to baseline scenario)
- $p_{i,1}$, $p_{i,2}$, and prot are set to 0.001, 0.002, and 0.995 for all states.
- The factors influencing mean or standard deviation in the case of defective batch, are all set to 1.05.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
2	14	V2_1 to V2_14	0.8	0.1	2	0.125	2.5	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2	0.15	5	0.075	2.5	33.33333333	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
6	3	V6_1, V6_3, V6_5	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V6_7	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
9	1	V9	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 15: Parameter set for Scenario 3 (not shown, smaller values for strength_2 and strength_6.

Expected Results:

- Since we are only changing the tolerance levels at States 2 and 6, States 1, 3, 4, 5, 7, and 8 should remain at the similar level compared to Scenario 1.
- Since the tolerance levels are increased in States 2 and 6, $P(A=1|R=0)$ for States 2, 6, 9, thin-wall and through-wall should be reduced.

Results:

- Compared to the Scenario 1, the change in tolerance-to-threshold ratio shows a minor reduction in $P(A=1|R=0)$ at States 2 and 6. However, as States 2, 6, and 9 still show much higher spikes compared to States 1, 3, 4, 5, 7, and 8, we believe that there are parameters that are unique to States 2 and 6 that are causing such high spikes.

Scenario 3

The parameter set for Scenario 3 is shown in Table 15; simulation results can be found in Figure 11 and a comparison with Scenario 2 can be seen in Figure 18.

Parameter Set Description:

- In this scenario, we will investigate the effect of parameters strength_2 and strength_6, which represent the magnitude of the effect of the folded normal distribution in States 2 and 6.
- In Scenarios 1 and 2, strength_2 and strength_6 were set to 1, and 1.5, respectively. Both of these values are reduced to 0.2. These will become the default values for the next 6 scenarios.

Expected Results:

- Since strength_2 and strength_6 affect the magnitude of shift in means, we expect that the reduction in these terms will affect $P(A_1=1|R_1=0)$ at States 2 and 6.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	15	10	25	0.666666667	1.666666667	2.5	0.001	0.002	0.995
2	14	V2_1 to V2_14	1.1	0.1	2	0.090909091	1.818181818	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2.5	0.8	5	0.32	2	6.25	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.001	0.002	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
6	3	V6_1, V6_3, V6_5	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V6_7	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.001	0.002	0.995
9	1	V9	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 16: Parameter set for Scenario 4

Results:

- As shown in Figure 18, there are clear reductions in $P(A_1 = 1|R_1 = 0)$ at States 2 and 6.
- Also, with given parameter sets, the probability of through-wall is zero for all runs.
- At the same time, the probability of thin-wall is greatly decreased.

Scenario 4

The parameter set for Scenario 4 is shown in Table 16; simulation results can be found in Figure 12 and a comparison with Scenario 3 can be seen in Figure 19.

Parameter Set Description:

- The following scenario is a modification to the parameter sets provided in Scenario 3.
- σ-to-threshold ratios are reduced in States 1, 2, and 3. The ratios are now between 1.67 to 2.
- We are interested in determining whether failure at earlier stages affect the probability of through-wall.

Expected Results:

- We expect that $P(A = 1|R = 0)$ will be inflated at States 1, 2, and 3.

Results:

- Figure 19, clearly indicate inflation of $P(A = 1|R = 0)$ at Stages 1, 2, and 3.
- However, there is no visible impact on latter stages.
- Furthermore, with given parameter sets, the probability of through-wall is zero for all runs.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	15	10	25	0.666666667	1.666666667	2.5	0.001	0.002	0.995
2	14	V2_1 to V2_14	1.1	0.1	2	0.090909091	1.818181818	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2.5	0.8	5	0.32	2	6.25	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.01	0.02	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
6	3	V6_1, V6_3, V6_5	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
	1	V6_7	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
9	1	V9	0.4	0.1	1	0.25	2.5	10	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 17: Parameter set for Scenario 5

Scenario 5

The parameter set for Scenario 5 is shown in Table 17; simulation results can be found in Figure 13 and a comparison with Scenario 4 can be seen in Figure 20.

Parameter Set Description:

- The following scenario is a modification to the parameter sets provided in Scenario 4.
- In previous scenarios, $p_{i,1}$, and $p_{i,2}$ were set to 0.001, and 0.002, respectively at all states. They are inflated to 0.01 and 0.02.
- In this scenario, two main goals are to determine the effect of inflated $p_{i,1}$, and $p_{i,2}$, as well as their impact on subsequent states.

Expected Results:

- $P(A_1 = 1|R_1 = 0)$ at States 4 and 5are expected to inflate.
- Since States 7 and 8 share the parameters with States 4 and 5, these states are likely affected as well.

Results:

- Due to inflation in $p_{i,1}$, and $p_{i,2}$, States 4 and 5 show greater $P(A = 1|R = 0)$.
- While $P(A = 1|R = 0)$ are at a similar level in States 4 and 5, these probabilities are impacted at different magnitudes between States 7 and 8. This is likely due to the fact that State 7 has three components, while State 8 has only one component.
- Even though State 9 is dependent on State 7, no significant impact is observed at State 9.

Scenario 6

The parameter set for Scenario 6 is shown in Table 18; simulation results can be found in Figure 14 and a comparison with Scenario 5 can be seen in Figure 21.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	15	10	25	0.666666667	1.666666667	2.5	0.001	0.002	0.995
2	14	V2_1 to V2_14	1.1	0.1	2	0.090909091	1.818181818	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2.5	0.8	5	0.32	2	6.25	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.01	0.02	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
6	3	V6_1, V6_3, V6_5	0.5	0.1	1	0.2	2	10	0.001	0.002	0.995
	1	V6_7	0.5	0.1	1	0.2	2	10	0.001	0.002	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
9	1	V9	0.5	0.1	1	0.2	2	10	0.001	0.002	0.995
10	1	Through wall									
	1	Thin wall									

Table 18: Parameter set for Scenario 6

Parameter Set Description:

- Here, parameters in States 6 and 9 are introduced. Parameters at other states are held constant.
- σ-to-threshold ratio is reduced from 2.5 to 2 in States 6 and 9.
- As State 10 is built on States 6 and 9, our main interest lies in determining the effect on thin-wall, as well as through-wall.

Expected Results:

- $P(A = 1|R = 0)$ at States 6 and 9 should be inflated.
- The likelihood of thin-wall and through-wall should also be inflated.

Results:

- As expected, $P(A = 1|R = 0)$ at States 6 and 9 are increased significantly.
- While the likelihood of thin-wall has increased, no case of through-wall is reported in all runs.

Scenario 7

The parameter set for Scenario 7 is shown in Table 19; simulation results can be found in Figure 15 and a comparison with Scenario 6 can be seen in Figure 22.

Parameter Set Description:

- In addition to the changes implemented in Scenario 6, $p_{i,1}$, and $p_{i,2}$ are changed to 0.01, and 0.02 in States 6 and 9.

Expected Results:

- In a similar way to Scenario 6, $P(A_1 = 1|R_1 = 0)$ are expected to increase in States 6 and 9.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	15	10	25	0.666666667	1.666666667	2.5	0.001	0.002	0.995
2	14	V2_1 to V2_14	1.1	0.1	2	0.090909091	1.818181818	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2.5	0.8	5	0.32	2	6.25	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.01	0.02	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
6	3	V6_1, V6_3, V6_5	0.5	0.1	1	0.2	2	10	0.01	0.02	0.995
	1	V6_7	0.5	0.1	1	0.2	2	10	0.01	0.02	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
9	1	V9	0.5	0.1	1	0.2	2	10	0.01	0.02	0.995
10	1	Through wall									
	1	Thin wall									

Table 19: Parameter set for Scenario 7

- The probabilities of thin-wall and through-wall are positively affected.

Results:

- States 6 and 9 have further increased $P(A_1 = 1|R_1 = 0)$.
- While the likelihood of thin-wall become extremely large, the event of through-wall is not reported in any runs.

Scenario 8

The parameter set for Scenario 8 is shown in Table 20; simulation results can be found in Figure 16 and a comparison with Scenario 7 can be seen in Figure 23.

Parameter Set Description:

- In the last scenario, σ-to-threshold ratio in States 6 and 9 are further reduced to 1.42.

Expected Results:

- Inflation in $P(A_1 = 1|R_1 = 0)$ at States 6 and 9 is expected.
- Possible increase in the likelihood of thin-wall and through-wall.

Results:

- Again, $P(A_1 = 1|R_1 = 0)$ at States 6 and 9 are increased.
- While it is small, the probability of through-wall is positive in all twenty runs. The average likelihood is now 1.14×10^{-4}.

State	n.variable	Variable	sigma	tol	threshold	sig:tol	sig:threshold	tol:threshold	p1_1 (p1_5)	p1_2 (p1_4)	prot
1	1	V1	15	10	25	0.666666667	1.666666667	2.5	0.001	0.002	0.995
2	14	V2_1 to V2_14	1.1	0.1	2	0.090909091	1.818181818	20	0.001	0.002	0.995
3	3	V3_1, V3_2, V3_3	2.5	0.8	5	0.32	2	6.25	0.001	0.002	0.995
4	2	V4_1, V4_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V4_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
5	1	V5_1	10	5	25	0.5	2.5	5	0.01	0.02	0.995
	4	V5_2 to V5_5	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
6	3	V6_1, V6_3, V6_5	0.7	0.1	1	0.142857143	1.428571429	10	0.01	0.02	0.995
	1	V6_7	0.7	0.1	1	0.142857143	1.428571429	10	0.01	0.02	0.995
7	2	V7_1, V7_2	0.4	0.1	1	0.25	2.5	10	0.01	0.02	0.995
	1	V7_3	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
8	1	V8	2	0.2	5	0.1	2.5	25	0.01	0.02	0.995
9	1	V9	0.7	0.1	1	0.142857143	1.428571429	10	0.01	0.02	0.995
10	1	Through wall									
	1	Thin wall									

Table 20: Parameter set for Scenario 8

Lookup Tables

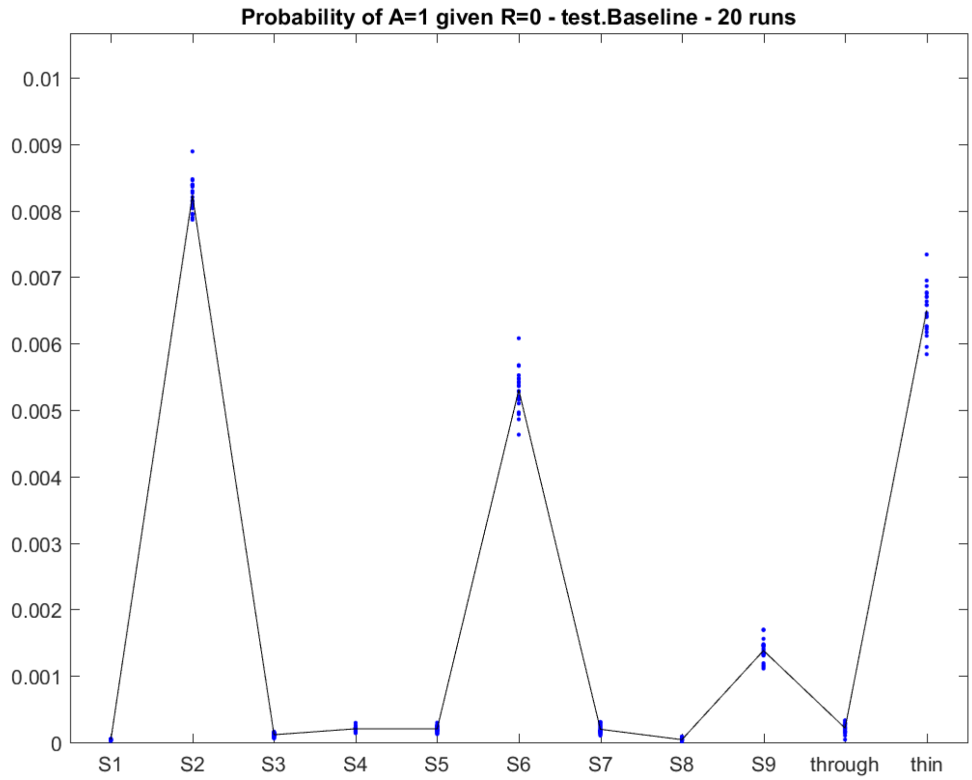
The given lookup tables serve as guidelines to determine the expected conditional probability of a damaged UFC misreported as undamaged (i.e., $P(A = 1|R = 0)$). The values in the table are generated based on normal distributions with one-sided thresholds (such as is the case with V_1 in State 1). Other variables are also generated based on normal distributions or on order statistics based on normal distributions and folded normal distributions.

Therefore, while these tables do not provide the exact expected probability $P(A = 1|R = 0)$ for each state, they provide a rough estimate of expected probability for each component at each state.

At the same time, comparisons within these tables provide an insight as to the effect that the parameters σ, tolerance, threshold, $p_1 = p_{\text{small}}$, and $p_2 = p_{\text{big}}$ have on $P(A = 1|R = 0)$. Consider, for instance, the third table of Table 21, where it is assumed that $p_1 = 0.001$, $p_2 = 0.002$, prot $= 0.99$, and tol $= 5$. The first row of that table shows that when σ is held constant at 1 (and the tolerance is held constant at 5), an increase in threshold values results in rapid reduction in $P(A = 1|R = 0)$. Similarly, a focus on the first column shows that an increase in sigma leads to inflated values of $P(A = 1|R = 0)$. The other tables can be used in a similar fashion.

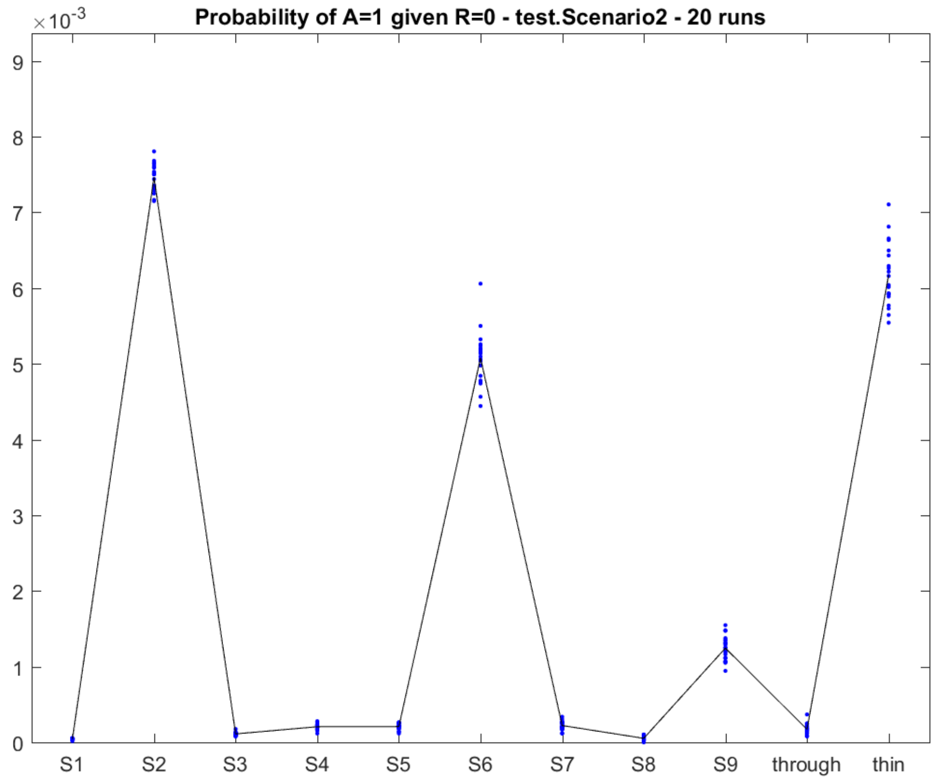
References

- Boyle, C., Overview of the NWMO and the Mark II Used Fuel Container, presentation deck from the NWMO.



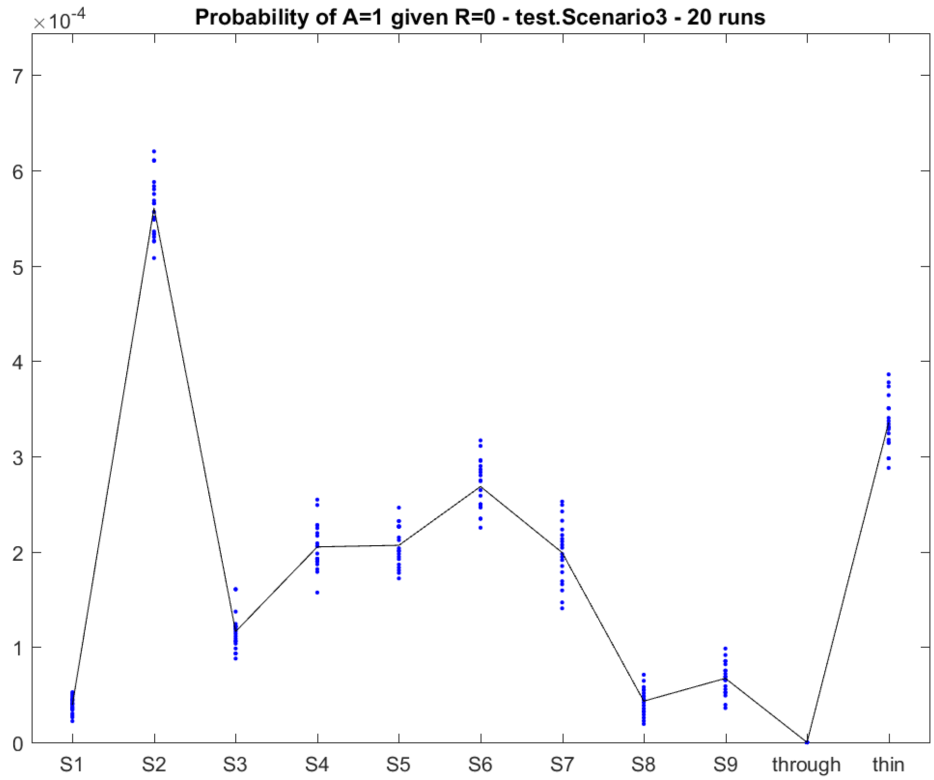
State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.9E-05	8.1E-06	2.2E-05	3.4E-05	4.0E-05	4.5E-05	5.3E-05
2	8.2E-03	2.5E-04	7.9E-03	8.1E-03	8.2E-03	8.4E-03	8.9E-03
3	1.2E-04	2.8E-05	6.3E-05	9.8E-05	1.1E-04	1.4E-04	1.6E-04
4	2.0E-04	3.7E-05	1.4E-04	1.8E-04	2.1E-04	2.2E-04	2.9E-04
5	2.0E-04	4.5E-05	1.2E-04	1.7E-04	2.0E-04	2.4E-04	3.0E-04
6	5.3E-03	3.2E-04	4.6E-03	5.1E-03	5.3E-03	5.5E-03	6.1E-03
7	2.0E-04	6.2E-05	1.0E-04	1.5E-04	1.8E-04	2.4E-04	3.1E-04
8	4.0E-05	1.9E-05	1.7E-05	3.4E-05	3.4E-05	5.1E-05	8.6E-05
9	1.4E-03	1.6E-04	1.1E-03	1.3E-03	1.4E-03	1.5E-03	1.7E-03
Through	2.1E-04	7.6E-05	4.1E-05	1.6E-04	2.2E-04	2.7E-04	3.3E-04
Thin	6.5E-03	3.6E-04	5.8E-03	6.2E-03	6.5E-03	6.7E-03	7.3E-03

Figure 9: Simulation results for Scenario 1: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



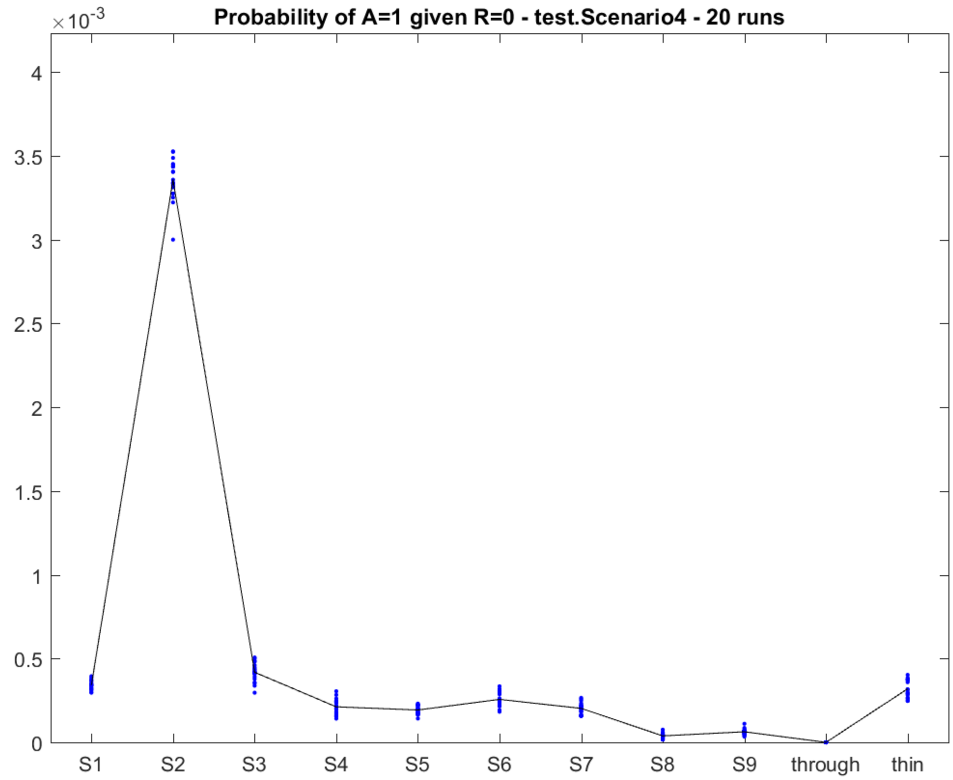
State	Mean	SD	Min	Q1	Median	Q3	Max
1	4.4E-05	9.8E-06	1.6E-05	4.0E-05	4.6E-05	5.1E-05	5.7E-05
2	7.5E-03	1.8E-04	7.2E-03	7.3E-03	7.5E-03	7.6E-03	7.8E-03
3	1.1E-04	2.1E-05	7.8E-05	9.8E-05	1.1E-04	1.2E-04	1.8E-04
4	2.1E-04	4.1E-05	1.2E-04	1.8E-04	2.1E-04	2.3E-04	2.8E-04
5	2.1E-04	4.2E-05	1.2E-04	1.9E-04	2.2E-04	2.4E-04	2.6E-04
6	5.1E-03	3.5E-04	4.4E-03	4.8E-03	5.1E-03	5.2E-03	6.1E-03
7	2.2E-04	5.7E-05	1.2E-04	1.9E-04	2.1E-04	2.6E-04	3.4E-04
8	5.3E-05	2.8E-05	0.0E+00	3.4E-05	5.1E-05	6.9E-05	1.0E-04
9	1.2E-03	1.6E-04	9.4E-04	1.1E-03	1.3E-03	1.3E-03	1.5E-03
Through	1.7E-04	6.9E-05	8.2E-05	1.2E-04	1.7E-04	2.0E-04	3.7E-04
Thin	6.2E-03	4.1E-04	5.5E-03	5.9E-03	6.1E-03	6.4E-03	7.1E-03

Figure 10: Simulation results for Scenario 2: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



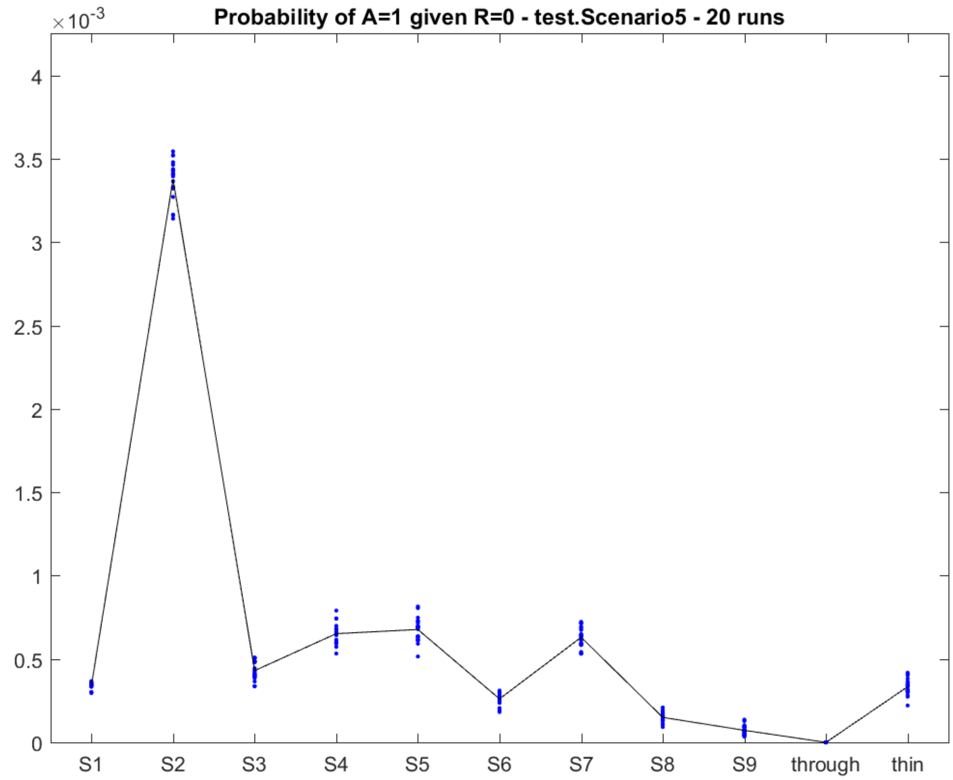
State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.9E-05	8.1E-06	2.2E-05	3.4E-05	4.0E-05	4.5E-05	5.3E-05
2	5.6E-04	3.2E-05	5.1E-04	5.3E-04	5.6E-04	5.8E-04	6.2E-04
3	1.2E-04	1.9E-05	8.8E-05	1.1E-04	1.2E-04	1.2E-04	1.6E-04
4	2.1E-04	2.4E-05	1.6E-04	1.9E-04	2.0E-04	2.2E-04	2.5E-04
5	2.1E-04	2.1E-05	1.7E-04	1.9E-04	2.0E-04	2.3E-04	2.5E-04
6	2.7E-04	2.6E-05	2.3E-04	2.5E-04	2.7E-04	2.9E-04	3.2E-04
7	2.0E-04	3.3E-05	1.4E-04	1.8E-04	2.0E-04	2.2E-04	2.5E-04
8	4.3E-05	1.4E-05	1.9E-05	3.2E-05	4.5E-05	5.2E-05	7.1E-05
9	6.7E-05	1.7E-05	3.6E-05	5.5E-05	6.7E-05	7.7E-05	9.8E-05
Through	0	0	0	0	0	0	0
Thin	3.4E-04	2.7E-05	2.9E-04	3.2E-04	3.3E-04	3.5E-04	3.9E-04

Figure 11: Simulation results for Scenario 3: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



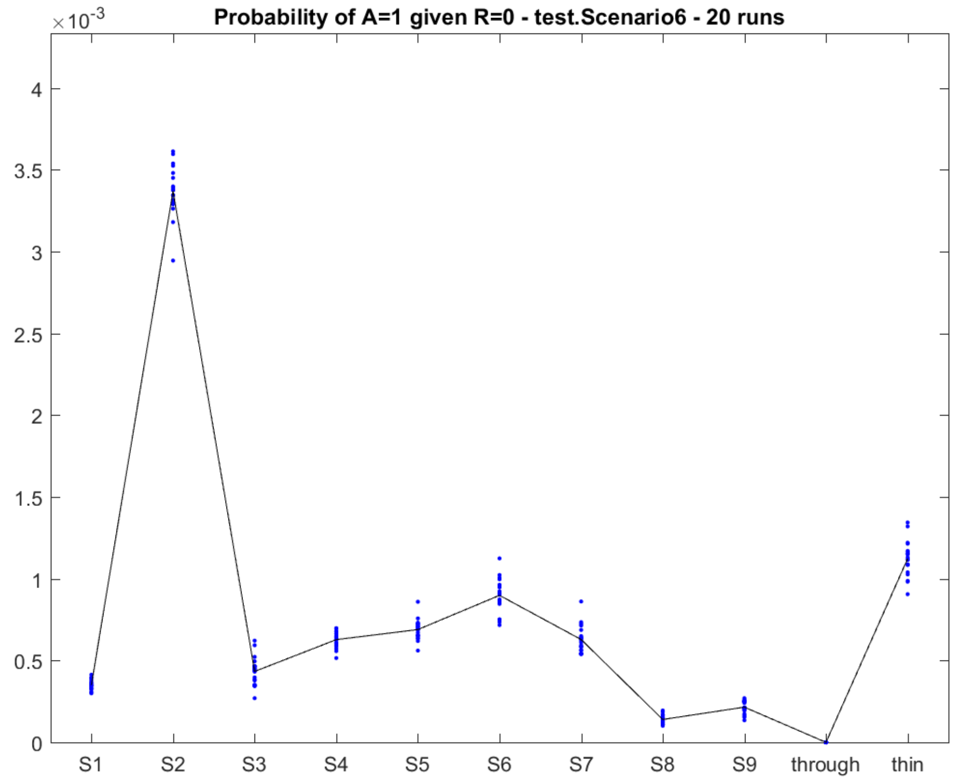
State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.4E-04	2.7E-05	3.0E-04	3.2E-04	3.4E-04	3.6E-04	3.9E-04
2	3.4E-03	1.2E-04	3.0E-03	3.3E-03	3.3E-03	3.4E-03	3.5E-03
3	4.2E-04	5.5E-05	3.0E-04	3.9E-04	4.2E-04	4.5E-04	5.1E-04
4	2.1E-04	4.5E-05	1.4E-04	1.8E-04	2.1E-04	2.4E-04	3.0E-04
5	1.9E-04	2.5E-05	1.4E-04	1.7E-04	1.9E-04	2.2E-04	2.3E-04
6	2.6E-04	4.7E-05	1.8E-04	2.2E-04	2.5E-04	3.0E-04	3.3E-04
7	2.0E-04	3.4E-05	1.6E-04	1.8E-04	2.0E-04	2.2E-04	2.7E-04
8	3.9E-05	1.6E-05	1.7E-05	2.5E-05	3.8E-05	5.0E-05	7.6E-05
9	6.3E-05	1.7E-05	3.4E-05	5.1E-05	6.4E-05	6.9E-05	1.1E-04
Through	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
Thin	3.2E-04	5.2E-05	2.5E-04	2.8E-04	3.1E-04	3.8E-04	4.0E-04

Figure 12: Simulation results for Scenario 4: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



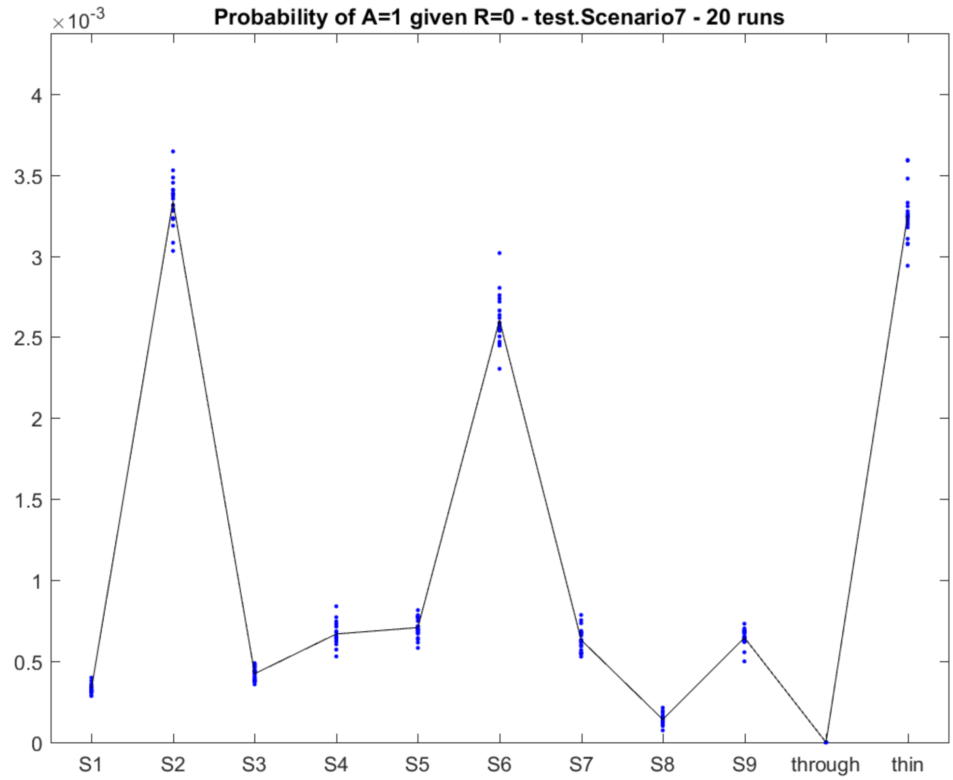
State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.4E-04	2.0E-05	3.0E-04	3.4E-04	3.4E-04	3.5E-04	3.7E-04
2	3.4E-03	1.2E-04	3.1E-03	3.3E-03	3.4E-03	3.4E-03	3.5E-03
3	4.3E-04	5.5E-05	3.4E-04	4.0E-04	4.2E-04	4.9E-04	5.1E-04
4	6.5E-04	6.2E-05	5.3E-04	6.0E-04	6.5E-04	6.8E-04	7.9E-04
5	6.8E-04	7.2E-05	5.2E-04	6.3E-04	6.9E-04	7.2E-04	8.1E-04
6	2.6E-04	3.6E-05	1.8E-04	2.4E-04	2.7E-04	2.9E-04	3.1E-04
7	6.3E-04	5.9E-05	5.3E-04	5.9E-04	6.3E-04	6.8E-04	7.2E-04
8	1.5E-04	2.9E-05	9.2E-05	1.3E-04	1.5E-04	1.7E-04	2.1E-04
9	7.2E-05	2.9E-05	3.4E-05	4.9E-05	6.8E-05	9.4E-05	1.4E-04
Through	0	0	0	0	0	0	0
Thin	3.3E-04	4.9E-05	2.2E-04	3.1E-04	3.4E-04	3.6E-04	4.2E-04

Figure 13: Simulation results for Scenario 5: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



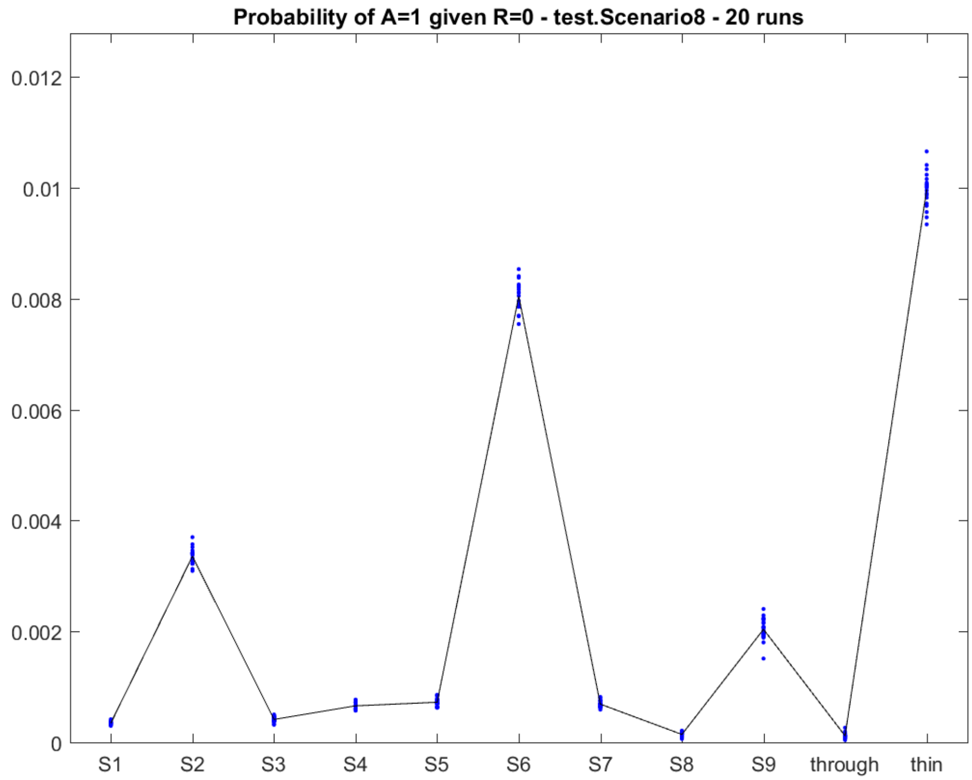
State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.5E-04	3.1E-05	3.0E-04	3.3E-04	3.5E-04	3.7E-04	4.1E-04
2	3.4E-03	1.5E-04	2.9E-03	3.3E-03	3.4E-03	3.5E-03	3.6E-03
3	4.3E-04	8.6E-05	2.7E-04	3.7E-04	4.4E-04	4.7E-04	6.2E-04
4	6.3E-04	4.8E-05	5.2E-04	6.0E-04	6.3E-04	6.6E-04	7.0E-04
5	6.9E-04	6.1E-05	5.6E-04	6.5E-04	6.9E-04	7.2E-04	8.6E-04
6	9.0E-04	1.0E-04	7.2E-04	8.5E-04	9.1E-04	9.6E-04	1.1E-03
7	6.3E-04	8.3E-05	5.4E-04	5.6E-04	6.2E-04	6.6E-04	8.6E-04
8	1.4E-04	2.7E-05	1.0E-04	1.2E-04	1.3E-04	1.6E-04	1.9E-04
9	2.2E-04	3.8E-05	1.3E-04	2.0E-04	2.1E-04	2.5E-04	2.7E-04
Through	0	0	0	0	0	0	0
Thin	1.1E-03	1.1E-04	9.1E-04	1.1E-03	1.1E-03	1.2E-03	1.3E-03

Figure 14: Simulation results for Scenario 6: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.3E-04	2.8E-05	2.9E-04	3.1E-04	3.2E-04	3.4E-04	4.0E-04
2	3.3E-03	1.4E-04	3.0E-03	3.3E-03	3.4E-03	3.4E-03	3.6E-03
3	4.2E-04	3.9E-05	3.6E-04	3.9E-04	4.3E-04	4.5E-04	4.9E-04
4	6.7E-04	7.1E-05	5.3E-04	6.3E-04	6.5E-04	7.1E-04	8.4E-04
5	7.1E-04	6.2E-05	5.8E-04	6.7E-04	7.0E-04	7.7E-04	8.2E-04
6	2.6E-03	1.5E-04	2.3E-03	2.5E-03	2.6E-03	2.7E-03	3.0E-03
7	6.3E-04	7.4E-05	5.3E-04	5.6E-04	6.3E-04	6.8E-04	7.9E-04
8	1.4E-04	3.3E-05	7.4E-05	1.2E-04	1.4E-04	1.6E-04	2.1E-04
9	6.5E-04	5.2E-05	5.0E-04	6.2E-04	6.5E-04	6.8E-04	7.3E-04
Through	0	0	0	0	0	0	0
Thin	3.3E-03	1.6E-04	2.9E-03	3.2E-03	3.2E-03	3.3E-03	3.6E-03

Figure 15: Simulation results for Scenario 7: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.



State	Mean	SD	Min	Q1	Median	Q3	Max
1	3.5E-04	2.7E-05	3.0E-04	3.3E-04	3.5E-04	3.6E-04	4.2E-04
2	3.4E-03	1.5E-04	3.1E-03	3.2E-03	3.4E-03	3.4E-03	3.7E-03
3	4.1E-04	5.0E-05	3.2E-04	3.8E-04	4.1E-04	4.5E-04	5.0E-04
4	6.6E-04	6.2E-05	5.7E-04	6.1E-04	6.6E-04	7.2E-04	7.7E-04
5	7.2E-04	7.3E-05	6.2E-04	6.8E-04	7.1E-04	7.7E-04	8.6E-04
6	8.0E-03	2.6E-04	7.5E-03	7.9E-03	8.1E-03	8.2E-03	8.5E-03
7	6.9E-04	6.0E-05	5.9E-04	6.5E-04	6.9E-04	7.2E-04	8.2E-04
8	1.4E-04	3.9E-05	6.2E-05	1.2E-04	1.5E-04	1.7E-04	2.1E-04
9	2.0E-03	2.0E-04	1.5E-03	1.9E-03	2.0E-03	2.2E-03	2.4E-03
Through	1.2E-04	5.0E-05	4.2E-05	8.3E-05	1.1E-04	1.3E-04	2.6E-04
Thin	1.0E-02	3.2E-04	9.3E-03	9.8E-03	1.0E-02	1.0E-02	1.1E-02

Figure 16: Simulation results for Scenario 8: conditional probabilities $P(A_j = 1|R_j = 0)$ for each state and descriptive statistics.

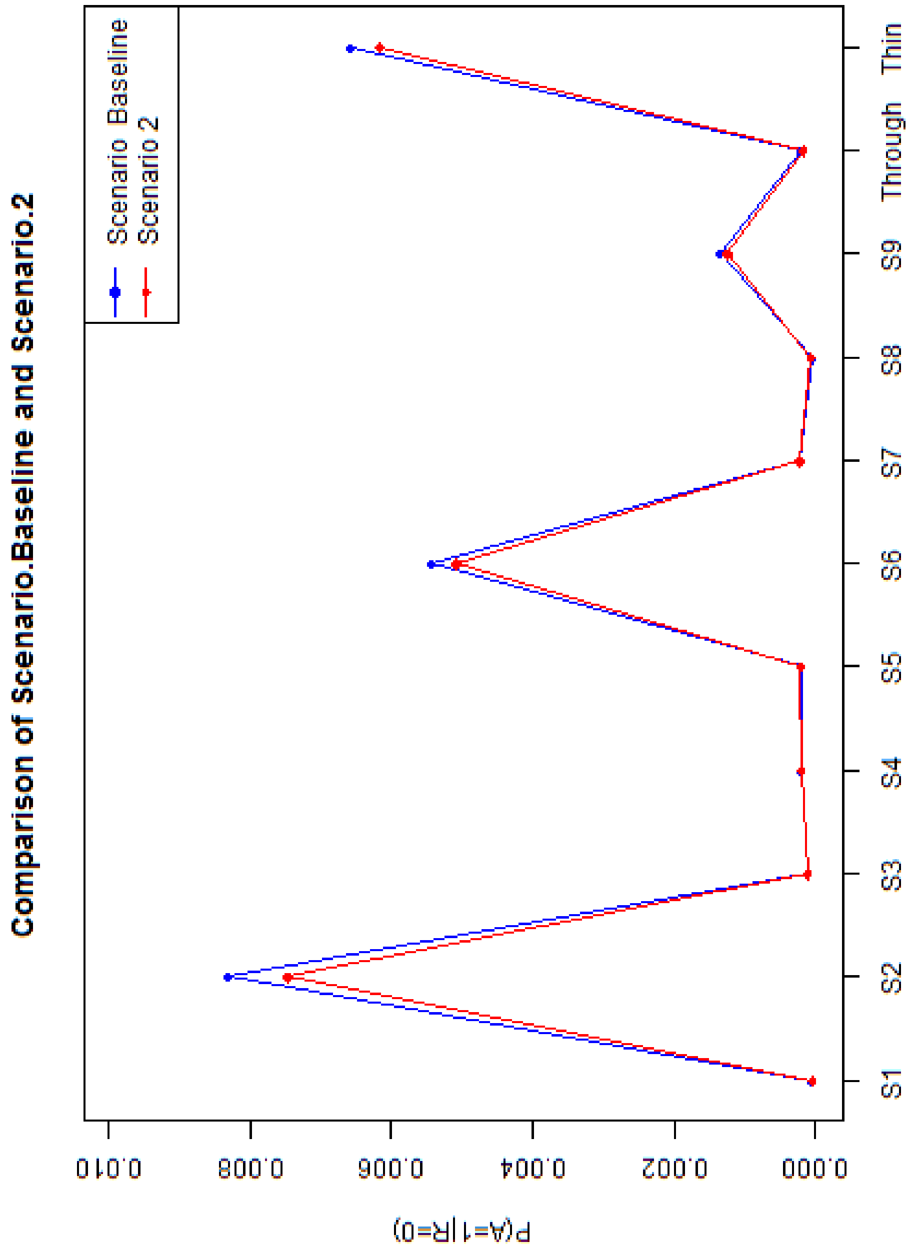


Figure 17: Comparison of simulation results between Scenarios 1 and 2.

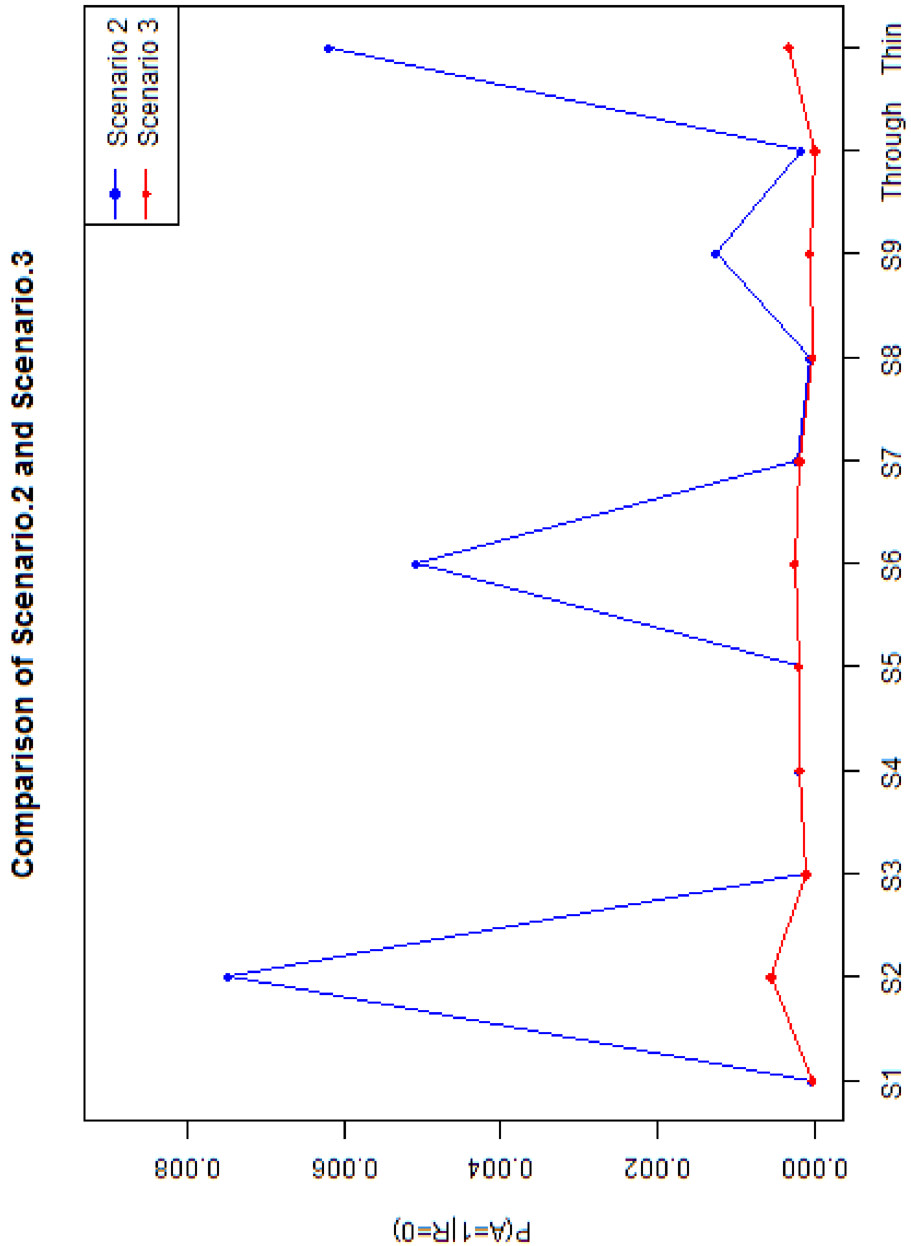


Figure 18: Comparison of simulation results between Scenarios 2 and 3.

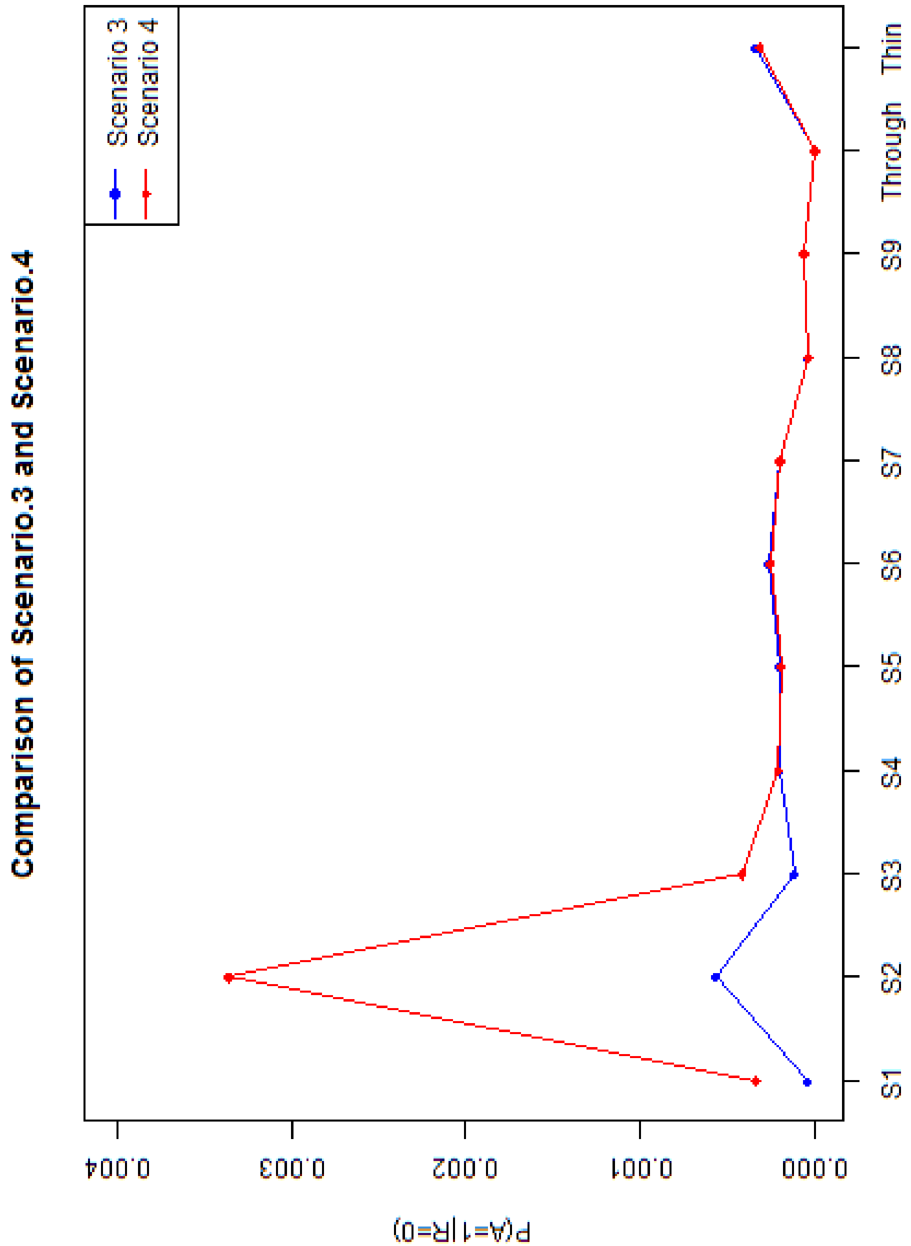


Figure 19: Comparison of simulation results between Scenarios 3 and 4.

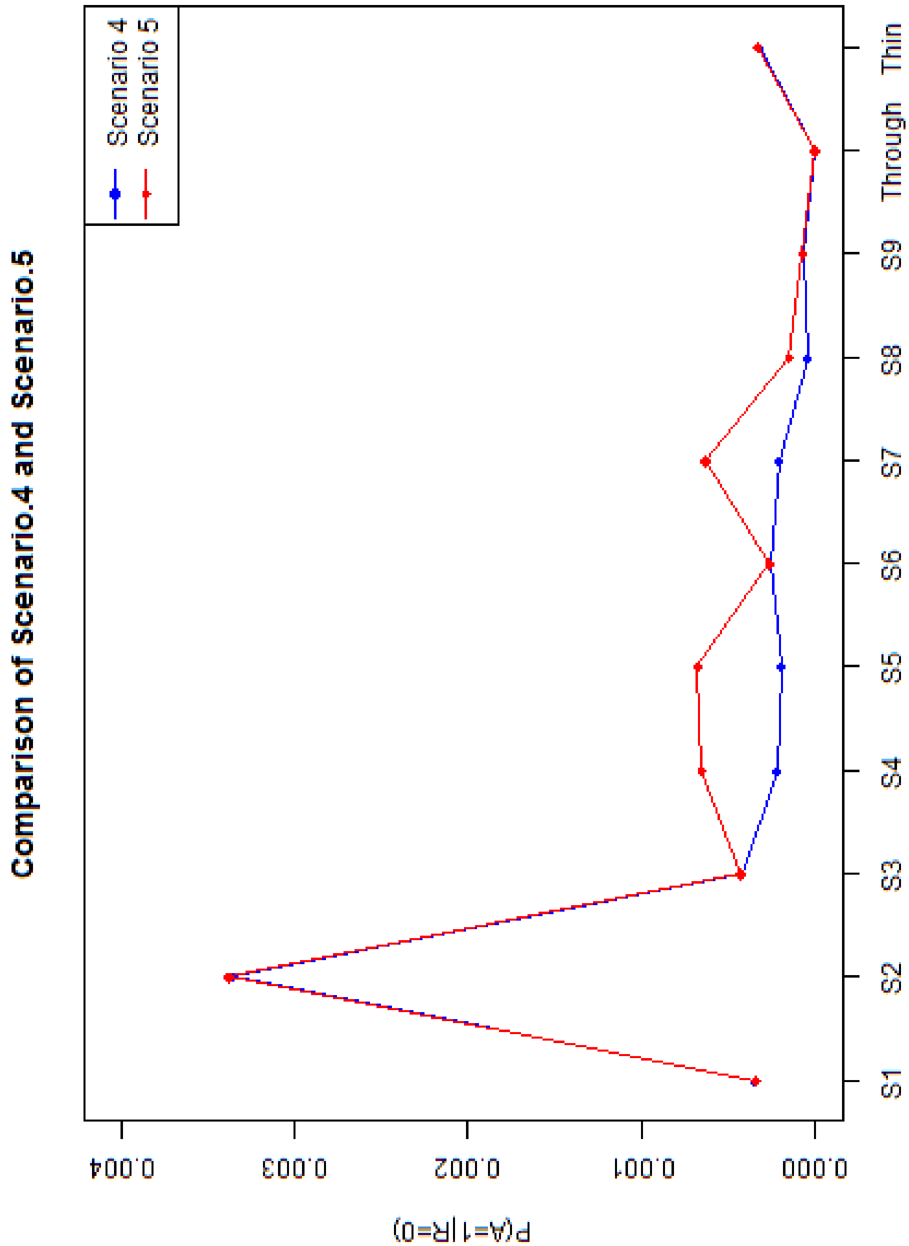


Figure 20: Comparison of simulation results between Scenarios 4 and 5.

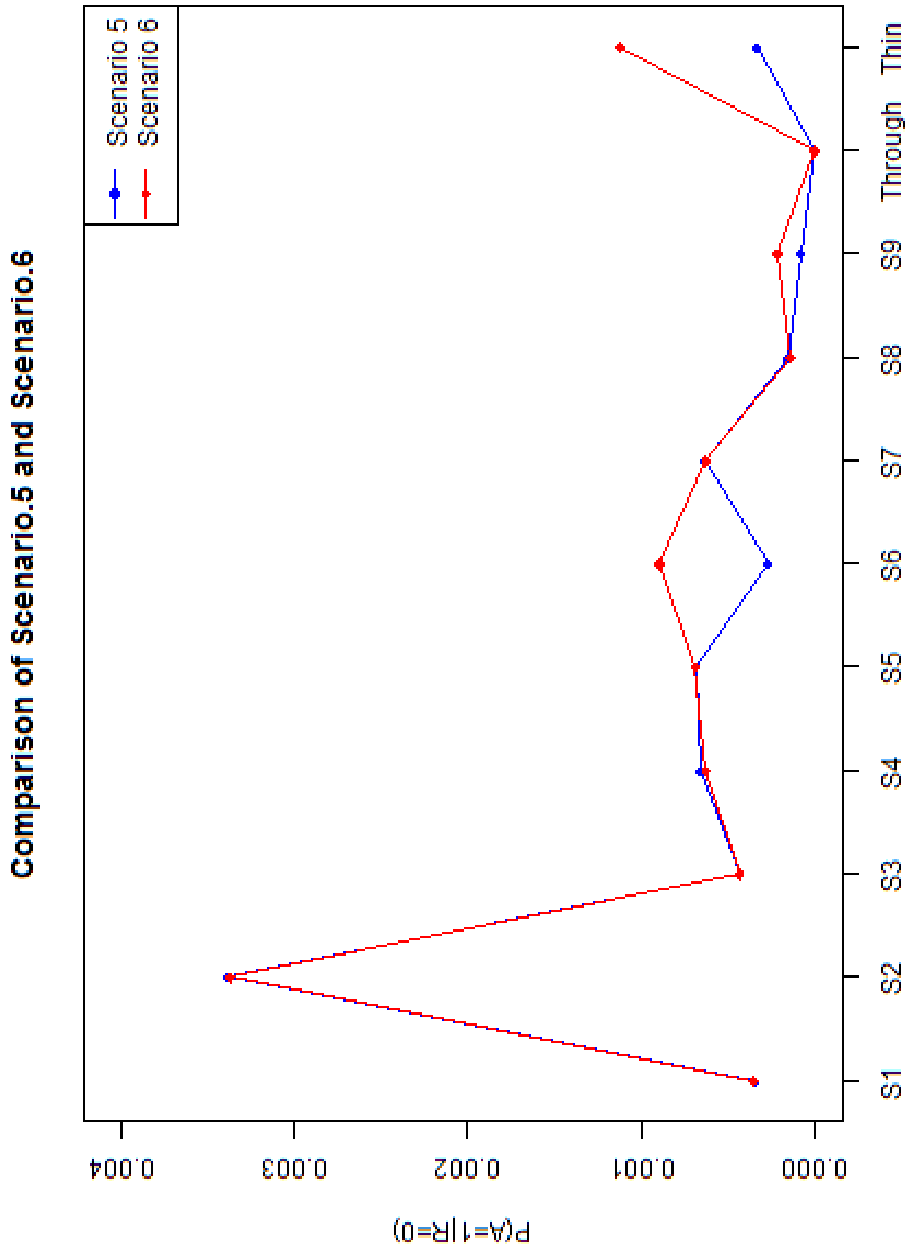


Figure 21: Comparison of simulation results between Scenarios 5 and 6.

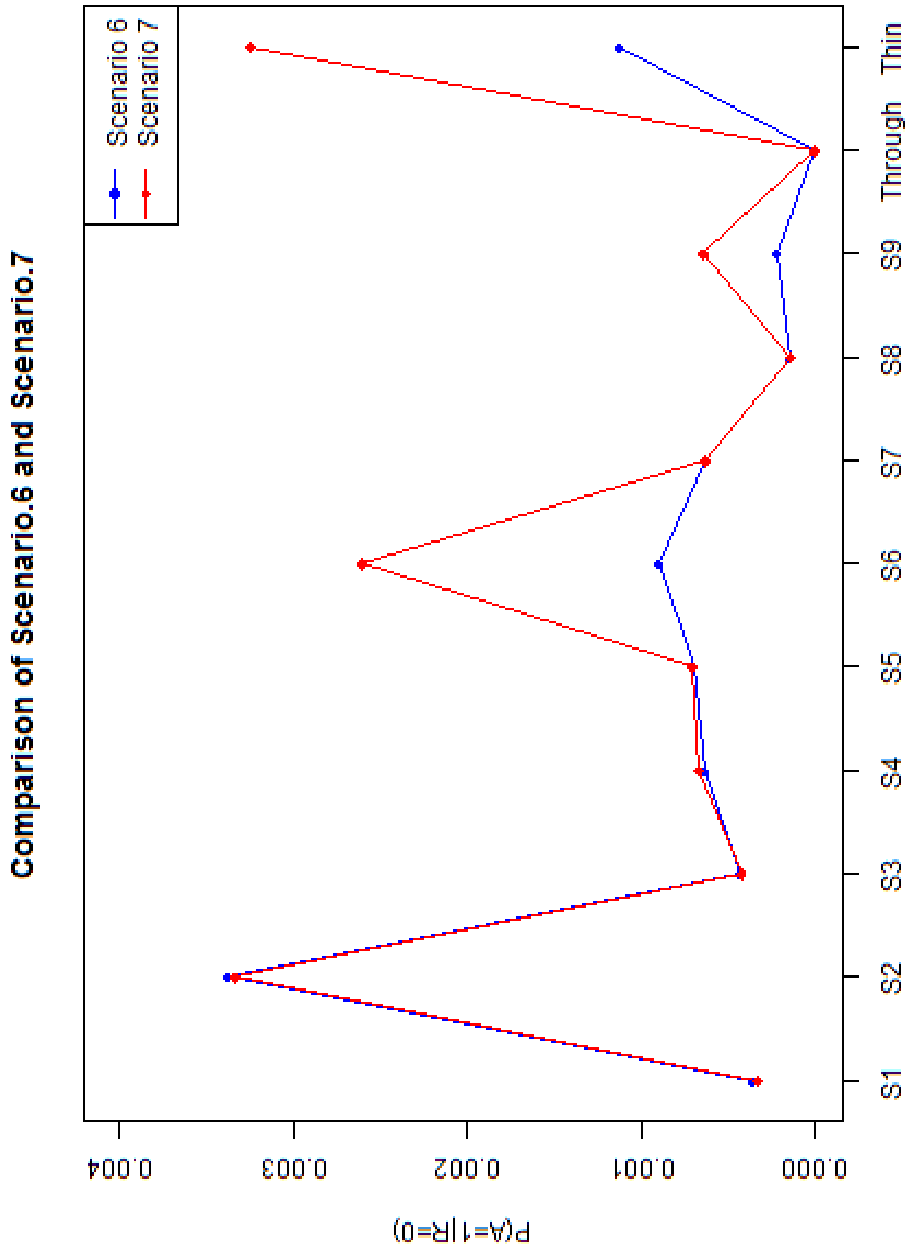


Figure 22: Comparison of simulation results between Scenarios 6 and 7.

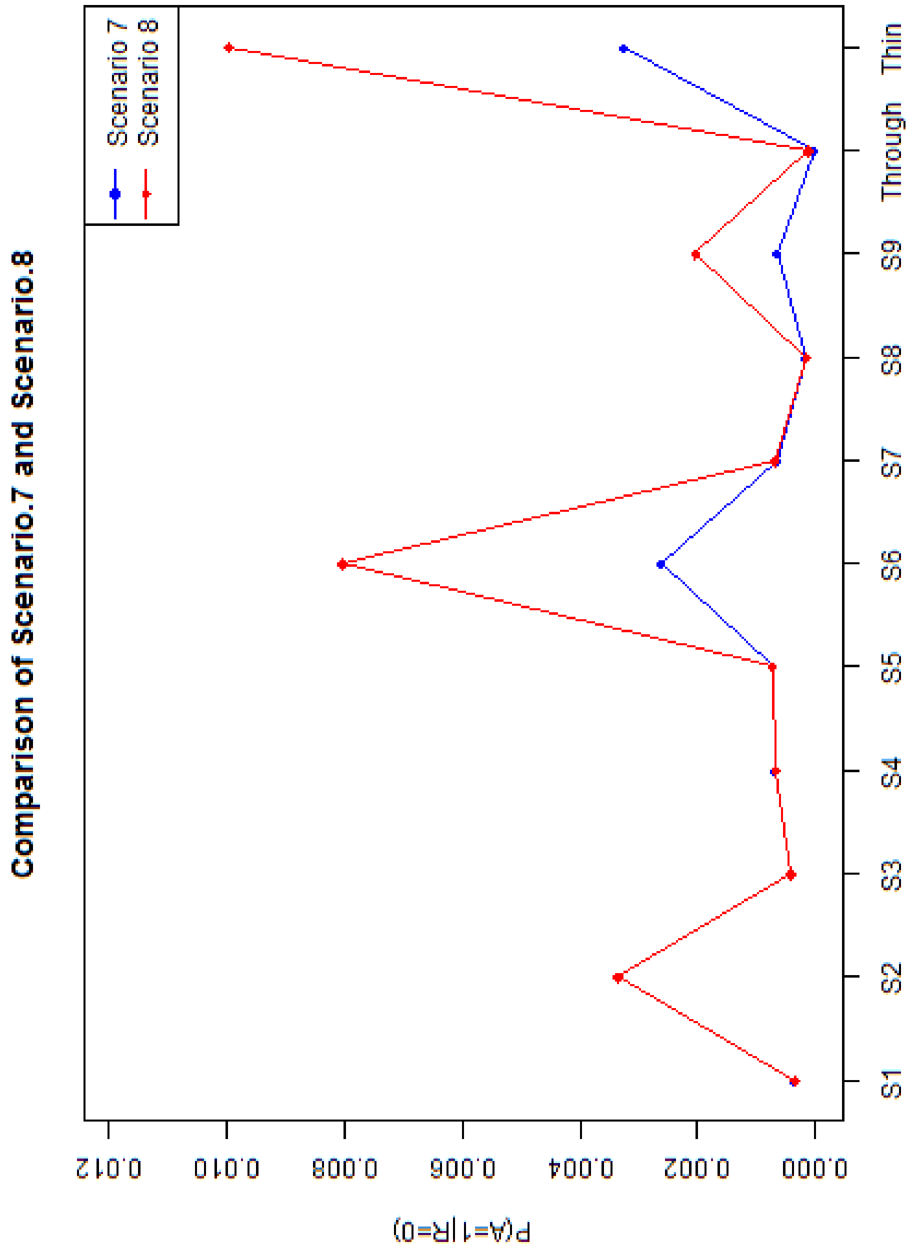


Figure 23: Comparison of simulation results between Scenarios 7 and 8.

sigma=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
tol=1	0.008128	0.00590837	0.0042627	0.0030422	0.00214057	0.00148007	0.00100247	0.00066312	0.00042722	0.00026741
tol=2	0.00824742	0.00600244	0.00433537	0.00309716	0.00218118	0.00150932	0.00102295	0.00067704	0.00043639	0.000273245
tol=3	0.00835336	0.00608264	0.00439491	0.00314043	0.00221189	0.00153058	0.00103726	0.00068639	0.0004423	0.000276862
tol=4	0.00844367	0.00614832	0.00444177	0.00317315	0.00223422	0.00154542	0.00104687	0.00069242	0.00044596	0.000279016
tol=5	0.00851764	0.00620003	0.00447722	0.00319694	0.00224982	0.00155539	0.00105306	0.00069616	0.00044815	0.000280249
tol=6	0.00857587	0.00623915	0.00450298	0.00321356	0.00226028	0.00156182	0.0010569	0.00069838	0.0004494	0.000280926
tol=7	0.00861991	0.00626758	0.00452098	0.00322471	0.00226703	0.0015658	0.00105919	0.00069965	0.00045008	0.000281285
tol=8	0.00865193	0.00628743	0.00453306	0.0032319	0.00227122	0.00156817	0.00106049	0.00070035	0.00045045	0.000281466
tol=9	0.00867429	0.00630076	0.00454085	0.00323636	0.00227371	0.00156953	0.00106121	0.00070072	0.00045063	0.000281555
tol=10	0.0086893	0.00630936	0.00454568	0.00323902	0.00227514	0.00157028	0.00106159	0.00070091	0.00045072	0.000281597

threshold=5	tol=1	tol=2	tol=3	tol=4	tol=5	tol=6	tol=7	tol=8	tol=9	tol=10
sigma=1	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09
sigma=2	7.42E-05	7.54E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05
sigma=3	0.0005831	0.00059667	0.00060295	0.00060555	0.00060652	0.00060684	0.00060693	0.00060696	0.00060697	0.000606967
sigma=4	0.00135296	0.00138279	0.00140209	0.00141382	0.00142053	0.00142413	0.00142595	0.00142681	0.00142719	0.001427356
sigma=5	0.00214057	0.00218118	0.00221189	0.00223422	0.00224982	0.00226028	0.00226703	0.00227122	0.00227371	0.002275136
sigma=6	0.00286229	0.00290839	0.00294635	0.00297677	0.00300046	0.00301842	0.00303167	0.00304116	0.00304778	0.003052276
sigma=7	0.00350075	0.00354897	0.00359078	0.0036263	0.00365586	0.00367998	0.00369925	0.00371434	0.00372592	0.003734627
sigma=8	0.00405976	0.00410819	0.0041516	0.00418992	0.00422321	0.00425169	0.00427568	0.00429557	0.00431181	0.004324858
sigma=9	0.00454882	0.00459642	0.00464008	0.00467963	0.00471503	0.00474631	0.00477363	0.00479719	0.00481725	0.004834135
sigma=10	0.00497806	0.00502429	0.0050674	0.00510719	0.00514355	0.00517646	0.00520593	0.00523208	0.00525504	0.005275003

tol=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
sigma=1	0.00227673	0.00028163	1.64E-05	3.83E-07	3.47E-09	1.19E-11	1.55E-14	7.53E-18	1.37E-21	9.22E-26
sigma=2	0.00536866	0.00227646	0.00086553	0.00028163	7.56E-05	1.64E-05	2.82E-06	3.83E-07	4.11E-08	3.47E-09
sigma=3	0.00700444	0.00405781	0.0022722	0.00121157	0.00060652	0.0002815	0.00011991	4.65E-05	1.64E-05	5.19E-06
sigma=4	0.00793585	0.00531328	0.00350431	0.00226227	0.00142053	0.00086237	0.00050333	0.00028104	0.0001495	7.55E-05
sigma=5	0.00851764	0.00620003	0.00447722	0.00319694	0.00224982	0.00155539	0.00105306	0.00069616	0.00044815	0.000280249
sigma=6	0.00891158	0.00684972	0.00524008	0.00398218	0.00300046	0.00223724	0.0016477	0.00119644	0.00085506	0.000600454
sigma=7	0.0091954	0.00734392	0.00584744	0.0046362	0.00365586	0.00286368	0.00222557	0.00171407	0.00130671	0.000984951
sigma=8	0.00940967	0.007732	0.00634029	0.00518414	0.00422321	0.00342494	0.00276282	0.00221511	0.00176373	0.001393563
sigma=9	0.0095773	0.00804477	0.00674747	0.00564781	0.00471503	0.00392377	0.00325305	0.00268532	0.00220585	0.001802125
sigma=10	0.00971213	0.00830225	0.00708925	0.00604434	0.00514355	0.00436678	0.0036971	0.0031202	0.00262388	0.002197683

Table 21: Lookup tables – $p_1 = 0.001$, $p_2 = 0.002$, prot $= 0.99$

sigma=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
tol=1	0.01566569	0.01150683	0.00838486	0.00604149	0.00429008	0.00299255	0.00204409	0.00136315	0.00088508	0.000558161
tol=2	0.0168406	0.0124362	0.00910514	0.00658754	0.00469426	0.00328404	0.00224848	0.0015022	0.00097668	0.00061648
tol=3	0.01788072	0.0132272	0.00969446	0.007017	0.0049998	0.00349582	0.00239119	0.0015955	0.00103574	0.000652621
tol=4	0.01876579	0.0138743	0.01015792	0.00734164	0.00522178	0.0036437	0.00248696	0.00165567	0.00107235	0.000674143
tol=5	0.01948971	0.01438313	0.01050823	0.00757749	0.00537678	0.00374293	0.00254872	0.00169296	0.00109415	0.000686459
tol=6	0.02005888	0.01476771	0.01076272	0.00774217	0.00548079	0.00380693	0.00258699	0.00171517	0.00110662	0.000693232
tol=7	0.02048901	0.01504708	0.01094041	0.00785267	0.00554787	0.00384658	0.00260979	0.00172787	0.00111348	0.00069681
tol=8	0.02080144	0.01524213	0.01105964	0.00792393	0.00558943	0.0038702	0.00262283	0.00173486	0.0011171	0.000698627
tol=9	0.02101956	0.01537301	0.01113652	0.00796809	0.00561418	0.00388371	0.00263	0.00173855	0.00111894	0.000699514
tol=10	0.02116591	0.0154574	0.01118417	0.00799439	0.00562835	0.00389114	0.00263379	0.00174043	0.00111984	0.00069993

threshold=5	tol=1	tol=2	tol=3	tol=4	tol=5	tol=6	tol=7	tol=8	tol=9	tol=10
sigma=1	8.62E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09	8.63E-09
sigma=2	0.00017447	0.00018571	0.00018773	0.00018802	0.00018805	0.00018805	0.00018805	0.00018805	0.00018805	0.00018805
sigma=3	0.0012702	0.00140568	0.00146835	0.00149432	0.00150397	0.00150717	0.00150813	0.00150838	0.00150844	0.001508455
sigma=4	0.00280098	0.0030983	0.0032906	0.0034075	0.00347429	0.00351016	0.00352825	0.00353683	0.00354066	0.003542258
sigma=5	0.00429008	0.00469426	0.0049998	0.00522178	0.00537678	0.00548079	0.00554787	0.00558943	0.00561418	0.005628348
sigma=6	0.0056144	0.00607263	0.00644974	0.00675166	0.0069868	0.00716495	0.00729625	0.00739038	0.00745602	0.007500552
sigma=7	0.00676454	0.00724339	0.00765822	0.00801038	0.00830336	0.00854221	0.00873302	0.00888239	0.00899698	0.009083113
sigma=8	0.00775955	0.00824004	0.00867043	0.00904999	0.00937958	0.00966138	0.00989861	0.01009524	0.01025571	0.010384648
sigma=9	0.00862296	0.00909487	0.00952734	0.00991886	0.01026899	0.01057829	0.01084818	0.01108084	0.01127894	0.011445562
sigma=10	0.00937635	0.0098344	0.01026112	0.01065473	0.01101422	0.0113393	0.01163037	0.0118884	0.01211489	0.012311734

tol=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
sigma=1	0.00564421	0.00070025	4.07E-05	9.53E-07	8.63E-09	2.97E-11	3.85E-14	1.87E-17	3.40E-21	2.29E-25
sigma=2	0.01323436	0.00564148	0.00214996	0.00070022	0.00018805	4.07E-05	7.00E-06	9.53E-07	1.02E-07	8.63E-09
sigma=3	0.01698188	0.00993647	0.0055992	0.00299724	0.00150397	0.000699	0.00029797	0.00011565	4.07E-05	1.29E-05
sigma=4	0.0187034	0.01268896	0.00845448	0.00550053	0.00347429	0.0021185	0.00124054	0.00069436	0.00037003	0.00018716
sigma=5	0.01948971	0.01438313	0.01050823	0.00757749	0.00537678	0.00374293	0.00254872	0.00169296	0.00109415	0.000686459
sigma=6	0.01985916	0.01546356	0.01196732	0.00918855	0.0069868	0.00525182	0.00389557	0.0028465	0.00204556	0.001443437
sigma=7	0.02003579	0.01619045	0.01303116	0.01043475	0.00830336	0.00655838	0.00513582	0.00398292	0.00305556	0.002316409
sigma=8	0.02011863	0.01670528	0.01383337	0.01141523	0.00937958	0.00766801	0.0062322	0.00503178	0.00403263	0.003205658
sigma=9	0.02015382	0.01708642	0.01445782	0.01220324	0.01026899	0.0086103	0.0071896	0.0059751	0.00493969	0.004060065
sigma=10	0.02016404	0.01737911	0.01495741	0.01284963	0.01101422	0.00941606	0.00802526	0.00681625	0.00576701	0.004858478

Table 22: Lookup tables – $p_1 = 0.01$, $p_2 = 0.02$, prot $= 0.99$

sigma=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
tol=1	0.008128	0.00590837	0.0042627	0.0030422	0.00214057	0.00148007	0.00100247	0.00066312	0.00042722	0.00026741
tol=2	0.00824742	0.00600244	0.00433537	0.00309716	0.00218118	0.00150932	0.00102295	0.00067704	0.00043639	0.000273245
tol=3	0.00835336	0.00608264	0.00439491	0.00314043	0.00221189	0.00153058	0.00103726	0.00068639	0.0004423	0.000276862
tol=4	0.00844367	0.00614832	0.00444177	0.00317315	0.00223422	0.00154542	0.00104687	0.00069242	0.00044596	0.000279016
tol=5	0.00851764	0.00620003	0.00447722	0.00319694	0.00224982	0.00155539	0.00105306	0.00069616	0.00044815	0.000280249
tol=6	0.00857587	0.00623915	0.00450298	0.00321356	0.00226028	0.00156182	0.0010569	0.00069838	0.0004494	0.000280926
tol=7	0.00861991	0.00626758	0.00452098	0.00322471	0.00226703	0.0015658	0.00105919	0.00069965	0.00045008	0.000281285
tol=8	0.00865193	0.00628743	0.00453306	0.0032319	0.00227122	0.00156817	0.00106049	0.00070035	0.00045045	0.000281466
tol=9	0.00867429	0.00630076	0.00454085	0.00323636	0.00227371	0.00156953	0.00106121	0.00070072	0.00045063	0.000281555
tol=10	0.0086893	0.00630936	0.00454568	0.00323902	0.00227514	0.00157028	0.00106159	0.00070091	0.00045072	0.000281597

threshold=5	tol=1	tol=2	tol=3	tol=4	tol=5	tol=6	tol=7	tol=8	tol=9	tol=10
sigma=1	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09	3.47E-09
sigma=2	7.42E-05	7.54E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05	7.56E-05
sigma=3	0.0005831	0.00059667	0.00060295	0.00060555	0.00060652	0.00060684	0.00060693	0.00060696	0.00060697	0.000606967
sigma=4	0.00135296	0.00138279	0.00140209	0.00141382	0.00142053	0.00142413	0.00142595	0.00142681	0.00142719	0.001427356
sigma=5	0.00214057	0.00218118	0.00221189	0.00223422	0.00224982	0.00226028	0.00226703	0.00227122	0.00227371	0.002275136
sigma=6	0.00286229	0.00290839	0.00294635	0.00297677	0.00300046	0.00301842	0.00303167	0.00304116	0.00304778	0.003052276
sigma=7	0.00350075	0.00354897	0.00359078	0.0036263	0.00365586	0.00367998	0.00369925	0.00371434	0.00372592	0.003734627
sigma=8	0.00405976	0.00410819	0.0041516	0.00418992	0.00422321	0.00425169	0.00427568	0.00429557	0.00431181	0.004324858
sigma=9	0.00454882	0.00459642	0.00464008	0.00467963	0.00471503	0.00474631	0.00477363	0.00479719	0.00481725	0.004834135
sigma=10	0.00497806	0.00502429	0.0050674	0.00510719	0.00514355	0.00517646	0.00520593	0.00523208	0.00525504	0.005275003

tol=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
sigma=1	0.00227673	0.00028163	1.64E-05	3.83E-07	3.47E-09	1.19E-11	1.55E-14	7.53E-18	1.37E-21	9.22E-26
sigma=2	0.00536866	0.00227646	0.00086553	0.00028163	7.56E-05	1.64E-05	2.82E-06	3.83E-07	4.11E-08	3.47E-09
sigma=3	0.00700444	0.00405781	0.0022722	0.00121157	0.00060652	0.0002815	0.00011991	4.65E-05	1.64E-05	5.19E-06
sigma=4	0.00793585	0.00531328	0.00350431	0.00226227	0.00142053	0.00086237	0.00050333	0.00028104	0.0001495	7.55E-05
sigma=5	0.00851764	0.00620003	0.00447722	0.00319694	0.00224982	0.00155539	0.00105306	0.00069616	0.00044815	0.000280249
sigma=6	0.00891158	0.00684972	0.00524008	0.00398218	0.00300046	0.00223724	0.0016477	0.00119644	0.00085506	0.000600454
sigma=7	0.0091954	0.00734392	0.00584744	0.0046362	0.00365586	0.00286368	0.00222557	0.00171407	0.00130671	0.000984951
sigma=8	0.00940967	0.007732	0.00634029	0.00518414	0.00422321	0.00342494	0.00276282	0.00221511	0.00176373	0.001393563
sigma=9	0.0095773	0.00804477	0.00674747	0.00564781	0.00471503	0.00392377	0.00325305	0.00268532	0.00220585	0.001802125
sigma=10	0.00971213	0.00830225	0.00708925	0.00604434	0.00514355	0.00436678	0.0036971	0.0031202	0.00262388	0.002197683

Table 23: Lookup tables – $p_1 = 0.1$, $p_2 = 0.2$, prot $= 0.99$

sigma=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
tol=1	0.07962138	0.05999977	0.04467581	0.0327879	0.02365353	0.01672636	0.0115616	0.00779079	0.00510519	0.003245971
tol=2	0.08979193	0.06833725	0.05131934	0.03793383	0.02752581	0.01955417	0.01356302	0.00916168	0.00601266	0.003825723
tol=3	0.09862963	0.07532983	0.05669374	0.04194663	0.03043479	0.02159957	0.01495622	0.01007972	0.00659708	0.004184698
tol=4	0.10603021	0.08098094	0.06088213	0.04496009	0.0325385	0.02302325	0.01588916	0.01067093	0.00695893	0.004398361
tol=5	0.11200336	0.08538144	0.06402596	0.04713877	0.03400256	0.02397652	0.01648994	0.01103699	0.00717431	0.004520592
tol=6	0.1166499	0.08868247	0.0662981	0.04865472	0.03498276	0.02459038	0.01686192	0.01125486	0.00729752	0.004587795
tol=7	0.12013277	0.09106715	0.06787868	0.04966953	0.03561392	0.02497043	0.01708331	0.01137951	0.00736527	0.004623301
tol=8	0.12264736	0.09272548	0.06893659	0.05032292	0.03600468	0.02519664	0.01720997	0.01144803	0.00740106	0.004641329
tol=9	0.12439531	0.09383518	0.06961764	0.05072742	0.03623724	0.02532604	0.0172796	0.01148424	0.00741923	0.004650125
tol=10	0.12556466	0.09454947	0.07003923	0.05096815	0.03637028	0.02539718	0.01731639	0.01150262	0.0074281	0.004654249

threshold=5	tol=1	tol=2	tol=3	tol=4	tol=5	tol=6	tol=7	tol=8	tol=9	tol=10
sigma=1	5.75E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08	5.76E-08
sigma=2	1.12E-03	1.23E-03	1.25E-03	1.25E-03	1.25E-03	1.25E-03	1.25E-03	1.25E-03	1.25E-03	1.25E-03
sigma=3	0.00764001	0.00897593	0.00959284	0.0098483	0.0099431	0.00997462	0.00998401	0.00998651	0.00998711	0.009987237
sigma=4	0.01601376	0.01890102	0.0207604	0.02188764	0.02253064	0.02287557	0.02304954	0.02313201	0.02316876	0.023184141
sigma=5	0.02365353	0.02752581	0.03043479	0.0325385	0.03400256	0.03498276	0.03561392	0.03600468	0.03623724	0.036370277
sigma=6	0.03015618	0.03449758	0.03804435	0.04086709	0.04305525	0.04470714	0.04592131	0.04679008	0.04739509	0.047805108
sigma=7	0.03563466	0.04012993	0.04399394	0.04725244	0.04994815	0.0521357	0.05387683	0.05523587	0.05627603	0.057056568
sigma=8	0.04027199	0.04474844	0.04872633	0.05220994	0.05521646	0.0577735	0.05991651	0.06168618	0.06312599	0.064280044
sigma=9	0.0442311	0.04859927	0.05257119	0.0561414	0.0593138	0.0621004	0.06452002	0.06659679	0.06835864	0.069835965
sigma=10	0.04764255	0.05185918	0.05575746	0.05932792	0.06256778	0.0654804	0.06807453	0.07036353	0.07236445	0.074097181

tol=5	threshold=1	threshold=2	threshold=3	threshold=4	threshold=5	threshold=6	threshold=7	threshold=8	threshold=9	threshold=10
sigma=1	0.03651924	0.00465746	2.72E-04	6.37E-06	5.76E-08	1.98E-10	2.57E-13	1.25E-16	2.27E-20	1.53E-24
sigma=2	0.08214226	0.03649357	0.01418222	0.00465712	1.25E-03	2.72E-04	4.68E-05	6.37E-06	6.83E-07	5.76E-08
sigma=3	0.10246091	0.06242461	0.03609643	0.01963194	0.0099431	0.004645	0.0019853	7.72E-04	2.71E-04	8.62E-05
sigma=4	0.10997658	0.07740165	0.05299371	0.03516857	0.02253064	0.01387511	0.00817986	0.00459896	0.00245793	1.25E-03
sigma=5	0.11200336	0.08538144	0.06402596	0.04713877	0.03400256	0.02397652	0.01648994	0.01103699	0.00717431	0.004520592
sigma=6	0.1118905	0.0896425	0.07106158	0.05567428	0.04305525	0.03282252	0.02463194	0.01817239	0.01316229	0.009347683
sigma=7	0.11099166	0.09196977	0.0756582	0.06174448	0.04994815	0.04001731	0.03172483	0.02486475	0.01924924	0.014706435
sigma=8	0.10984891	0.09326495	0.07877499	0.06615834	0.05521646	0.04577076	0.0376601	0.03073839	0.02487239	0.019939891
sigma=9	0.10868041	0.09398919	0.08097047	0.06946114	0.0593138	0.05039533	0.04258538	0.03577473	0.02986377	0.024761189
sigma=10	0.10757152	0.09438677	0.08257237	0.07200378	0.06256778	0.05416167	0.0466922	0.04007453	0.03423122	0.02909126

Table 24: Lookup tables – $p_1 = 0.1$, $p_2 = 0.2$, prot $= 0.999$