WAGENINGEN
UNIVERSITY & RESEARCH

Assessing the impact of variability in microbial inactivation kinetics on the efficacy of thermal, high pressure, and pulsed electric fields processing

An evaluation of kinetic variability and product-specificity effects on process efficacy within a quantitative microbial risk assessment framework

Author:

Rick Mannes | 1298917
Wageningen University & Research - Laboratory of Food Microbiology
MSc. Food Technology

Supervisors:

Georgios Pampoukis
Prof. dr. ir. Heidy den Besten

Examiner:

Prof. dr. ir. Marcel Zwietering

Abstract

Novel food processing technologies such as high-pressure processing (HPP) and pulsed electric field (PEF) are increasingly explored as alternatives to conventional thermal processing (TP) due to growing consumer demand for minimally processed foods and enhanced quality retention. However, microbial inactivation kinetics for these technologies vary widely across the literature, limiting their reliable integration into quantitative microbial risk assessments (QMRA). This study focuses on the effect of variability in published kinetic data on the predicted inactivation of Shiga-toxin producing *Escherichia coli* (STEC). The aim was to evaluate how the modelling approach and product specificity influence final risk estimates, thereby improving efficacy predictions for novel, non-thermal technologies. For each technology, a standardised database has been built using strict inclusion criteria. Three modelling approaches, deterministic, stochastic, and Bayesian, were applied to describe inactivation kinetics in two scenarios: a generic case (all liquid products) and a specific case (high-acid fruit juices). These models were integrated into a Monte Carlo-based QMRA with one million iterations, simulating kinetic parameters, STEC contamination, processing inactivation, consumer exposure, and the probability of illness. Incorporating kinetic variability into the QMRA significantly altered the illness probability across all technologies. The Bayesian models, which also included parameter uncertainty, resulted in the broadest risk distributions. The impact of product specificity was technology-dependent: generic models led to lower illness estimates for TP and PEF, while for HPP, the specific model for high acidic fruit juice was more effective. Among all inputs, variability in kinetic parameters had a greater influence on the illness probability than process variability, variability in initial contamination or consumption differences. Additionally, the choice of dose-response model was the largest source of uncertainty in absolute illness probabilities, though it had less effect on the relative impact of interventions between technologies. These findings highlight the limitations of using point estimates from literature and demonstrate that failure to account for variability may lead to an underestimation of food safety risks. This study provides a framework for incorporating kinetic variability and uncertainty into QMRA, showing that the estimated efficacy of novel food processing technologies is significantly reduced when realistic variability is considered.

Key words: QMRA, *Escherichia.coli*, Food safety, Fruit juice, Bayesian regression, High pressure processing, pulsed electric fields

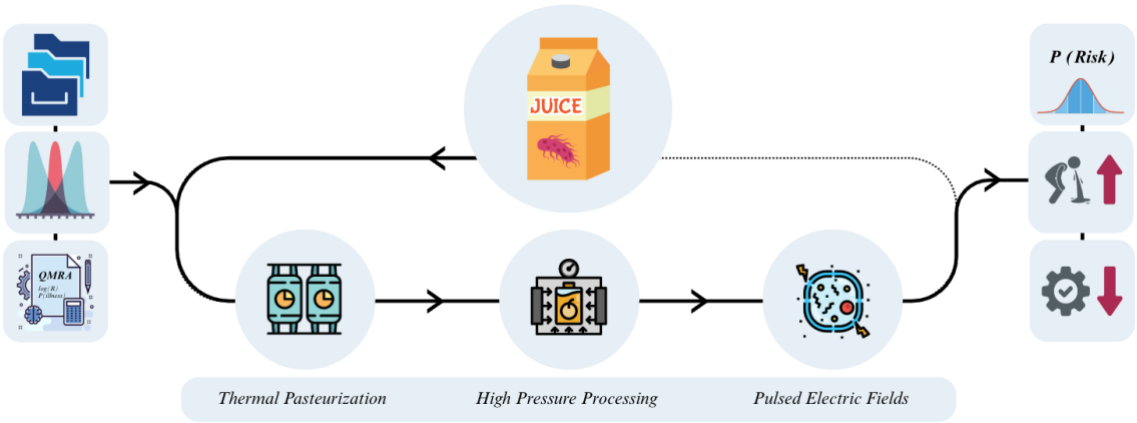


Figure 1, Graphical abstract: the impact of variability in reported microbial inactivation kinetics on the efficacy of processing technologies, using a QMRA framework and a case study focussed on STEC in fruit juice. In general, incorporating variability decreased efficacy significantly, connected to a meaningful increase in predicted STEC illness cases amongst all processing technologies.

Table of content

Chapter 1

Introduction

1.1 The role of non-thermal inactivation in risk mitigation

Microbial inactivation plays an important role in microbial risk mitigation within the food industry by controlling the safety and quality of food products, with the primary goal being the reduction of pathogenic microorganisms (Barba et al., 2017). Effective microbial inactivation reduces the likelihood of foodborne disease outbreaks and helps to control that food products remain safe for consumption during their shelf-life. Effective processing methods are vital for achieving microbial inactivation while maintaining the nutritional and sensory properties of food. While traditional methods such as thermal processing (TP) and the addition of chemical preservatives remain widely used to control microbial safety, they often compromise food quality by discolouration or nutrient deterioration, particularly for fresh produce such as fruit juices (Aganovic et al., 2017). Furthermore, the increasing global population drives a growing demand for food, which in turn contributes to rising energy consumption during production and processing. Within the European food supply chain, food processing is responsible for nearly a quarter of the total energy use, second only to agriculture (Monforti et al., 2015). To address these energy challenges while meeting consumer preferences for natural and minimally processed products, non-thermal processing (NTP) technologies such as high pressure processing (HPP) and pulsed electric fields (PEF) have gained attention as promising alternatives (Rodriguez-Gonzalez et al., 2015). These technologies hold potential for maintaining food quality while controlling microbial safety while also serving as an additional step in the supply chain to mitigate foodborne disease risk, extend shelf-life, and reduce food waste.

1.2 HPP and PEF and their key process characteristics

High Pressure Processing (HPP), also referred to as High Pressure (HP), High Hydrostatic Pressure (HHP) or Ultra-High Pressure (UHP) is a non-thermal pasteurisation technology whose industrial application has been growing steadily in recent decades (Chiozzi et al., 2022). Meanwhile, research interest in high pressure technologies has increased exponentially, as reflected by the surge in scientific publications since 2000 (Morata et al., 2021). The technology uses pressure ranges between 100 and 700 MPa and temperature ranges between -20 and 60 °C (Ozkan et al., 2019). At present, HPP is applied on fruits and vegetables, dairy products, juices and beverages, meats and ready to eat products (Aganovic et al., 2021). Unfortunately, it is insufficient for the inactivation of bacterial spores, since extremely high pressures (800 – 1700 MPa) or the combination with thermal processing is then required (Katsaros et al., 2015). Due to the high pressure, which is applied uniform and simultaneous in all directions (Chiozzi et al., 2022), the technology disrupts cell walls and membranes, inactivates enzymes and denatures proteins (Katsaros et al., 2015). The most important process parameters next to the applied pressure are temperature and holding time, which are besides the type of microorganism, strain, or food matrix effect, the most important source of variability as reported in literature (Škegro et al., 2021).

Compared to HPP and other non-thermal processing technologies, PEF offers advantages such as continuous processing, and shorter treatment times, making it particularly suitable for fruit juice

processing (Ranjha et al., 2021). The PEF technology applies short pulses of electric fields, with a duration of microseconds to milliseconds, while the product is placed between a set of two electrodes (Dziadek et al., 2019). The number of pulses conveyed to the product generates a low amount of heat. The procedure combines electroporation and electro-permeabilization of the cell-membrane (Lindgren et al., 2002). This electroporation can be permanent, depending on the strength of the electric field. For microbial inactivation, only liquid products can be used, since the current flows more efficiently through the liquid product. However, the technology knows other uses for solid products, such as physicochemical properties modification, through the modification of the membranes. The PEF system consists of a treatment chamber, a high voltage power source, a pulse generator, a cooling system for temperature rise compensation and an energy discharging switch (Aghajanzadeh & Ziaiifar, 2017). The efficacy of the technology is mainly determined by the conductivity of the product, the chamber geometry, the circuit parameters, the intensity of the electric field, the frequency, the pulse shape and the specific energy input (Salehi, 2020). The total energy input is describing the most critical and variable parameter and is mostly estimated by the parameters described above.

Variability in the microbial reduction kinetics in HPP and PEF has been found as even or more pronounced than those obtained in conventional thermal processing (Chacha et al., 2021). NTP technologies operate based on different physical principles, such as pressure or electric fields, which can lead to various effects on the food components. But most importantly, the design and operation of the NTP equipment can vary widely, affecting the uniformity and efficacy of the technology. Moreover, non-thermal methods may not always achieve the same level of microbial inactivation as thermal methods, leading to variability in the safety estimates (Chacha et al., 2021). Variability in process efficacy leads to substantial differences in microbial inactivation kinetics and reported process conditions (EFSA, 2019). Especially for PEF technology, there are other critical aspects that contribute to variability, due to the complexity of the process. Despite the fact that PEF technology follows a general principle, the broad experimental conditions, the diversity of available equipment and unstandardised experimental procedures results into considerable variability in reported microbial inactivation kinetics (Raso et al., 2016).

1.3 Generic microbial inactivation kinetics

To quantify microbial inactivation, mathematical models are used to describe inactivation kinetics (McMeekin et al., 2002). The foundation of most quantitative models is based on microbial inactivation kinetics, extracted from lab-scale experiments or literature. Primary and secondary inactivation models can quantitatively describe the microbial kinetics and are often based on the *D*-/z- approach (McMeekin et al., 2002). The *D*-value (decimal reduction time), in the primary model, is defined as the time required to inactivate 90% of the microbial population (1 log CFU/mL). The z-value provides an extra dimension and is defined as the change in temperature (or pressure) to reduce the *D*-value with 90% (1 log unit of time). These inactivation models are defined as the log-linear primary and the secondary model respectively, and are for thermal processing described by the Bigelow model:

$$\log_{10}(N_t) = \log_{10}(N_0) - \frac{t}{D} \qquad [1]$$

where $\log_{10} N_0$ is the initial microbial $\log_{10}$ counts, $\log_{10}(N_t)$ is describing the microbial $\log_{10}$ counts after treatment, and t is the treatment time (min) (Bigelow, 1921). To describe the processing conditions and assess the effect of a numerical parameter of interest on the *D*-value, a log-linear secondary model can also be used. For conventional thermal processing, the equation becomes:

$$\log_{10} D = \log_{10}(D_{Tref}) + \frac{T_{ref} - T}{z} \quad [2]$$

where D_{Tref} is the D-value at a reference temperature T_{ref} (ºC), T is the treatment temperature (ºC). For HPP the equivalent approach was implemented using the D_{Pref} (defined at a specific pressure) and z_P (the pressure change needed for 1 $\log_{10}$ reduction of the D) parameters. For pulsed electric field processing, it is not necessary to define a secondary model. Instead, a primary model can be used with the volumetric energy input (E_{in} in kJ/L) as the equivalent of time t and the D_{Ein} (kJ/L) parameter i.e., the energy input needed for 1 $\log_{10}$ reduction of the microbial population, as the equivalent of the D-value (Huang et al., 2012):

$$\log_{10}(N_t) = \log_{10}(N_0) - \frac{E_{in}}{D_{Ein}} \quad [3]$$

1.4 Quantitative Microbial Risk Assessment (QMRA)

In food safety management systems, quantitative approaches are often used and are based on the development and implementation of quantitative microbial risk assessments (QMRA) (Zwietering, 2009). A QMRA is a structured method used to evaluate the risk of human exposure to pathogenic microorganisms (Haas, 2020). Besides risk assessment, a QMRA can also provide support in policy development and decision-making processes, but most importantly, it provides a perspective on the magnitude of different intervention strategies. When combined with economic analysis, QMRAs can be used to evaluate cost-effectiveness of various processes (Zwietering & Havelaar, 2006). The backbone of a detailed QMRA consists of four main aspects: 1) Hazard identification: identifying pathogenic microbes that could pose a risk to human health; 2) Exposure assessment: determining the relationship between the dose of ingesting a pathogen and the probability of a negative health effect; 3) Hazard characterisation: estimating the extent of human exposure to a hazardous dose; 4) Risk characterisation: combining the data from the previous steps to estimate the risk of infection/illness (EC, 1997). During exposure assessment, the microbial kinetics of the identified pathogenic microorganism are explained quantitatively, using the inactivation models as defined before. Considering variability in microbial inactivation kinetic parameters is crucial in creating a realistic risk analysis (Zwietering, 2015). The final risk depends not only on the dose-response relation, which is shaped by factors such as strain variability, and variability in food matrices, but also on the dose itself. This variability in this dose is influenced by variable serving sizes and microbial concentrations at the time of consumption. These variable microbial concentrations are dependent on a variable process efficacy and initial concentration levels. Quantifying all these variabilities and identifying the sources is of high importance for realistic risk prediction (Zwietering, 2015). Besides variability, uncertainty plays an important role in QMRAs (Garre et al., 2020). Where variability in QMRAs is often based on natural differences in all parameters, uncertainty is due to the lack of knowledge or data limitations. Both variability and uncertainty can have an equal or even higher relevance than kinetic and process parameters themselves in a QMRA, pointing out the significance of including them both in the analysis (den Besten et al., 2017).

Meta-analysis in food microbiology is a statistical approach used to synthesize and analyse microbial inactivation kinetics across multiple independent studies, addressing variability and uncertainty in predictive microbiology (den Besten & Zwietering, 2012). Unlike single-study analyses, which rely on isolated datasets, meta-regression aggregates data from different sources to quantify how inactivation parameters (the D-/z- approach) change under varying conditions. The complexity of microbial responses to these intrinsic and extrinsic conditions requires different modelling approaches, with each having its own advantages and limitations.

Deterministic models assume that microbial inactivation follows fixed parameter values, often derived from regression models that fit empirical data to the predefined equations (McMeekin et al., 2002). These models provide a straightforward overview of predicting microbial reduction under controlled conditions, making them useful for defining baseline inactivation. However, they fail to capture the natural variability between different microbial strains, experimental setups and food matrices. To address these limitations, stochastic models, often based on Monte Carlo simulations, are used to incorporate random variability into parameter estimates, allowing microbial inactivation to be described by probability density distributions rather than fixed values (Zwietering et al., 2021). This allows for an estimation of the microbial reduction while accounting for natural variability by sampling from a distribution of $(\log_{10})D$-values. This approach has been found to be mathematically equivalent to the probability integral method based on Equation 1 (Equation 4), which evaluates the expected reduction by integrating over the probability distribution of *D*-values (Zwietering et al., 2021). Since *D*-values are not fixed, but are assumed to follow a log-normal distribution, the expected reduction can be calculated as:

$$\log_{10}(\frac{N_t}{N_0}) = \log_{10}\int_{-\infty}^{\infty} f(u)10^{-\frac{t}{10^{logD+\sigma\mu}}}du \quad [4]$$

With f representing the probability density function of a normal distribution with estimated value μ and standard deviation σ. This approach can be interpreted either a stochastic simulation, where Monte Carlo sampling of *D*-values represents the variability, or as a kinetic process where the D-value is treated as a normally distributed parameter. Both interpretations approximate the same biological reality that D-values are not fixed but vary naturally. By assigning similar probability functions to key parameters, stochastic models enable more realistic predictions of microbial survival and inactivation. The challenge by using stochastic approaches lies in defining appropriate distributions for the different sources of variability. Therefore, the accuracy of these models is highly dependent on the availability of experimental- or metadata to properly defined these distributions. Probabilistic models, mostly based on Bayesian approaches, extend these stochastic approaches by openly quantifying uncertainty in addition to variability (McElreath, 2015). These models leverage prior knowledge, often derived from the metadata to estimate posterior distributions. Unlike traditional frequentist regression, which produces single best-fit estimates, Bayesian regression generates full probability distributions for model parameters, integrating both prior knowledge and experimental data. Once the priors are defined, Bayesian models create posterior distributions using Bayes theorem, which updates prior beliefs based on the observed data. This distribution follows:

$$P(0|data) \propto P(data|0) * P(0) \quad [5]$$

where $P(0|data)$ represents the posterior probability of the model parameters given the data, and $P(data|0)$ the likelihood function describing the probability of observing the data specific model parameters. $P(0)$ is the prior distribution reflecting the prior knowledge before the data is observed. By integrating this approach, Bayesian models offer a more flexible and transparent framework for modelling (microbial) kinetics (van Boekel, 2020).

1.5 Study scope and case study

This study considers a general approach to assess and compare the impact of kinetic- and process variability on the efficacy of HPP, PEF and TP pasteurisation in fruit juices. In order to make a fair comparison while considering a real-world application, the study focusses on the inactivation of Shiga toxin-producing *Escherichia coli* (STEC) in fruit juices. Outbreaks of STEC have been associated with the consumption of unpasteurised apple cider and apple juice through outbreaks

that occurred in the previous decades (Topalcengiz & Danyluk, 2017). STEC are strains of *Escherichia coli* that produce Shiga toxins, which are proteins that have a cytotoxic effect (Scheutz et al., 2012). The optimum temperature for STEC growth is between 36 and 40 degrees Celsius, but the cells can survive between 6 to 45 degrees Celsius. Another important aspect is that cells can survive relatively low pH levels between 4.4 and 9, but grow optimally around neutral pH (EFSA, 2019). Moreover, prior studies have shown that STEC are capable of adapting and surviving lower pH environments, which on in turn may result in cross-protection against other stressors (Rivera-Reyes et al., 2019). The main process of contamination of STEC is in the gut of ruminants, which can spread the bacteria with their droppings. Infected animal faeces can contaminate the farm environment (mostly through water). Therefore, STEC can be found in soil, feed, water and plants (Bach et al., 2002). In the scope of this project, the main and primary source of contamination in vegetables, fruits and sprouts is the use of contaminated manure and water for fertilisation and irrigation (ECDC, 2020). In 2023, STEC infections ranked as the third most frequently reported foodborne illness in the European Union. Surveillance data from 21 EFSA member states revealed that 3.1% of 9254 non-ready-to-eat food samples tested positive for STEC (EFSA, 2024). Therefore, it is crucial to implement prevention strategies for fruit juices at the manufacturing level, by applying microbial inactivation strategies, such as pasteurisation by HPP, PEF or TP (EFSA, 2024). The addressed variability in microbial inactivation for these technologies directly influences food safety risk predictions. If a processing method results in insufficient microbial reduction the likelihood of viable STEC cells remaining in fruit juice is high, substantially increasing the probability of illness. However, even when the required reduction level is achieved, a non-zero probability of illness persists due to natural variability in initial contamination levels, microbial resistance and processing conditions. Understanding the sources and magnitude of this variability is essential for designing robust NTP processes that ensure food safety under all probable scenarios. Although previous studies have examined microbial inactivation kinetics for HPP and PEF, they primarily relied on deterministic models, which do not fully account for the range of variability in real-world processing scenarios. This study addresses these gaps by adding a stochastic and probabilistic QMRA to assess the impact of variability on food safety risks. By systematically quantifying variability in reported microbial inactivation kinetics, this study provides critical insights into process optimisation for HPP and PEF, contributing to improved food safety management strategies and evidence-based regulatory decision-making. This study investigates how variability in reported kinetic parameters for microbial inactivation based on biological variability, and variability in processing factors affect the efficacy of HPP, PEF and conventional TP in inactivating STEC in fruit juice.

Chapter 2

Methodology

2.1 Metadata collection & analysis

This study used three comprehensive datasets that were previously collected to evaluate the inactivation kinetics of *E.coli* in liquid food products, subjected to TP, HPP and PEF. These data form the basis of the modelling process and the comparison of assessing the impact of variability within each technology. Although the data was pre-collected, the collection methods and criteria used during the collection process are critical for ensuring consistency and transparency. Therefore, the data collection processes for TP, HPP and PEF are explained in the following sections. The databases for the three technologies including explanations can be found here.

2.1.1 Thermal processing data collection

The data collection for thermal processing was based on two corresponding sources: the inactivation dataset by van Asselt and Zwietering (2006), and an updated literature search performed in early 2025, specifically targeting the inactivation of *Escherichia coli* in fruit juices and ciders (van Asselt & Zwietering, 2006). From the original dataset of van Asselt and Zwietering, a subset of 142 rows was selected by applying filters on microbial species and the type of food matrix. Only entries referring to *E.coli* were retained, and datapoints were included if the reported matrix was categorised under 'liquids', 'dairy', 'juices', and 'media'. These criteria were chosen to ensure that the reference dataset focussed exclusively on liquid matrices for juice and cider processing. To expand this base dataset, a structured literature search was carried out by Pampoukis and Van der Sluis (2025), following a protocol tailored to capture recent inactivation data for *E.coli* in beverages. This search targeted five scientific databases (Scopus, Web of Science, FSTA, CAB Abstracts and SciFinder, using combinations of the keywords '*Escherichia coli'*, 'thermal inactivation', pasteurisation', 'D-value', 'fruit juice', and 'cider'. Only original, peer-reviewed English language research articles were included. Articles were retrieved in January 2025, yielding a total of 2479 initial hits. After deduplication, 1350 unique entries remained and were subjected to a two-phase screening process. The first screening step involved title and abstract review to exclude irrelevant studies. In the second phase, full-text articles were evaluated against the following exclusion criteria: 1) reviews, editorials or conference proceedings, 2) studies involving microorganisms other than *E.coli,* 3) studies involving hurdle technologies, not only thermal processing, 4) unreported reductions, D-values or beginning/end concentrations or absence of an inactivation curve, 5) processing temperatures < 50°C or > 100 °C. After this screening process, 82 new datapoints were extracted from 20 articles. If a paper overlapped with those already included in the dataset made by van Asselt and Zwietering, only the most recent or most complete version was retained to prevent duplication. Each included article was manually reviewed, and all relevant data were extracted following a redefined extraction protocol. This included the thermal conditions (e.g. temperature, duration, cooling), microbial inactivation (e.g. species, strain, growth phase), and matrix characteristics (e.g. product type, pH, Aw). When only *D*-values were not explicitly reported in the original article, they were calculated using the Bigelow model by fitting datapoints manually extracted from inactivation curves using WebPlotDigitizer V4.5 (Garre et al., 2025). The resulting log-survivor curved were fitted using the open-source software Bioinactivation4 (Garre et al., 2018),

which allowed for parameter estimation under first order kinetics. In cases where only a start and endpoint of microbial reduction were provided, the *D*-value was estimated by dividing the treatment time by the log reduction. However, to ensure the reliability of these estimates, a minimum log reduction of 0.5 log CFU/mL was required. Reduction below this threshold were excluded from the original dataset, as such minor decrease in microbial reduction fall within the uncertainty range of most enumeration techniques and could reflect background noise or method detection limits (Pampoukis et al., 2024). In total, 224 $\log_{10} D$ values for *E.coli* in thermal inactivation were retained, covering a range of fruit juices, ciders, dairy products and liquid media.

2.1.2 High Pressure Processing data collection

The data collection for HPP in this study followed a comparable approach as done for thermal processing, adapted from an existing large-scale HPP inactivation database. This database was compiled using a standardised screening protocol developed in 2023 and included articles retrieved from the five databases as defined in previous section. The search strategy included combinations of keywords such as 'high pressure processing', HPP, *E.coli,* inactivation and D-value, and was designed to broadly capture research articles related to microbial inactivation under HPP conditions. The initial search yielded a total of 17182 articles. After removal of duplicates, 4567 unique articles remained and were subjected to title and abstract screening using predefined criteria. Exclusion criteria included combined technologies (e.g. with heat or irradiation), and papers lacking sufficient data to estimate inactivation parameters. This initial screening resulted in 467 selected articles, of which 248 were found to contain quantitative microbial inactivation data and were included in the master HPP database. For this study, a filtered subset of this master database was created, specifically targeting the inactivation of *Escherichia coli* in liquid food matrices. The first filtering step was focussed on food matrix, selecting only studies that involved fruit juices, ciders, dairy-based liquids or liquid culture media, aligning the data scope with the thermal processing dataset to ensure comparability. Lastly the dataset was restricted to experiments conducted at temperatures <45°C to avoid inclusion of thermally assisted HPP conditions, ensuring that the observed reductions were based to pressure alone. After applying these filters, 25 studies were retained, contributing to a total of 326 $\log_{10} D$ values to the final dataset. For each row, relevant parameters were extracted manually using a structured template. This included the pressure level (MPa), come-up times, holding times, temperature, food matrix, pH, strain information (when available) and any available D- or z-values. When D-values were not reported explicitly, they were calculated from reduction data using similar digital tools as done for thermal processing.

2.1.3 Pulsed electric fields data collection

The data collection for PEF also structured to ensure high reproducibility and consistency. The approach followed a two-stage systematic literature review supported by ASreview. In the first stage, queries combining microbial inactivation and PEF related terms were used to retrieve articles involving *E.coli, Listeria monocytogenes, L.innocua, Saccharomyces cerevisiae, Bacillus subtills and lactic acid bacteria, Pseudomonas, Salmonella* and *Staphylococcus aureus*. From 1232 screened articles, 89 were marked relevant contributing 1036 inactivation experiments. After correcting for duplication, the full database included 175 experiments across 78 publications. A critical inclusion criterion was the ability to calculate or reliably estimate the volumetric energy input (kJ/L), which is the primary driver of microbial inactivation under PEF. Energy inputs per treatment were estimated using:

$$\Delta E = E^2 * f * \tau * \sigma * Rt * n_s \quad [6]$$

Where E equals the electric field strength (kV/cm), f equals the pulse frequency (pps), τ is the pulse width (µs), σ is the conductivity (s/cm), Rt is the estimated residence time (s) and n_s equals the number of treatment chambers. These parameters were either extracted directly from the articles or estimated under specific conditions. The electric field strength was calculated by dividing reported voltage by the electrode gap for parallel electrode configurations, assuming uniform field distribution. The pulse width and frequency were taken from reported values or estimated based on the effective treatment duration when not explicitly available. Residence time was either reported or calculated using the chamber volume and flow rate. Flow rates were estimated based on the chamber dimensions, pulse characteristics and product throughput. It is important to note that not all studies report all parameters required to calculate the estimated energy input. Therefore, articles were only included it at minimum the following parameters were reported: electric field strength, pulse frequency, pulse width, conductivity (at initial temperature), and number of treatment chambers. For this study, a subset of this larger database was used. The selected data included only continuous PEF treatment studies reporting inactivation data for *E.coli* in liquid matrices compared to both TP and HPP databases, including fruit juices, ciders, liquid dairy products and liquid media. Furthermore, only experiments resulting in a microbial reduction of at least 0.5 log CFU/mL were included. To avoid the influence of a heat effect, and the change of false-estimated energy inputs, which occur at higher values, this study only focused on datapoints where the estimated energy input was equal to or lower than 300 kJ/L. Moreover, only square-wave pulses were included, because the applied processing time and the corresponding energy input could be estimated more precisely. For modelling the reduction of *E. coli* in this study, 153 datapoints were used, extracted from 15 distinct articles.

2.2 Model fitting

For each of the three processing technologies, three types of models were developed: a deterministic, a stochastic and a Bayesian probabilistic model. All models were constructed to describe the kinetic response (the $\log_{10} D$ for TP and HPP, and $\log_{10}\left(\frac{N_0}{N_t}\right)$ for PEF) to given process inputs (temperature, pressure or energy input). To assess the influence of product specificity on the model fit and its effect in the QMRA, two scenarios were created by filtering the metadata regarding product specificity. First, a generic approach was developed by using all data for liquid products. Then, a subset was created by filtering the dataset to include only fruit juice matrices with a pH of 4.5 or lower, focussing on high acidic fruit juices explicitly. This two-level approach allowed for the evaluation of generalizability and specificity in model outcomes. The structure or the approach is visualised in Figure 2.

2.2.1 Linear models

The deterministic models represent conventional linear regressions, providing a single best-fit estimate based on the available metadata. For TP and HPP, the dependent variable was $\log_{10} D$, estimated as a linear function of temperature (TP) or pressure (HPP). This reflects the classical secondary inactivation model explained in equation 2, where the slope equals $\frac{-1}{z_x}$, and the intercept equals $\log_{10} D_{ref} + \frac{x_{ref}}{z_x}$. The z_T for thermal processing could be calculated by taking the negative inverse of the slope. For HPP the equivalent approach was implemented using the D_{ref} and z_P parameters. For PEF, the log reduction was modelled directly as a linear function of the estimated energy input (kJ/L). Since inactivation at zero energy input is not realistic, the intercept of the regression was constrained to zero. These deterministic models served as baseline estimators of microbial inactivation, based on central tendency.

To incorporate the variability present in the relation between the kinetic parameters and the process inputs for the technologies, stochastic models were created by extending the deterministic regression with a normally distributed error term. For TP and HPP, the single point predicted $\log_{10} D$ value was treated as the mean of a normal distribution with a standard deviation equal to the residual standard error (RSE) of the regression. Instead of a single point estimate, a distribution of possible $\log_{10} D$ values could now be estimated for a given processing input, enabling the prediction across the full prediction interval, and be implemented as shown in Equation 4. This approach is expressed as:

$$\log_{10} D = \sim N\left(\mu = -\frac{1}{z} * x + \log_{10} D_{ref} + \frac{x_{ref}}{z_x}, \sigma = RSE\right) \quad [7]$$

For PEF, the same approach was applied to the predicted log reduction, explained as:

$$\log_{10}\left(\frac{N_0}{N_t}\right) = \sim N\left(\mu = \frac{1}{D_{Ei}} * x, \sigma = RSE\right) \quad [8]$$

In the case of PEF, the lower tail of the reduction distribution was truncated at zero to avoid unrealistic negative inactivation values. This stochastic modelling approach allowed for a more realistic representation of the variability observed in the inactivation data collected.

2.2.2 Bayesian models

To account for parameter uncertainty in the models, a Bayesian fixed-effects regression approach was applied for each of the three technologies. Unlike the deterministic and the stochastic models, which use single best-fit estimates and fixed residual variation to describe the data variability, the Bayesian models explicitly quantify uncertainty in the regression coefficients themselves. This uncertainty arises from several factors, including the data availability, variation in study conditions and the indirect nature of the data extracted from literature. As such, Bayesian modelling provides a framework for integrating prior knowledge and observed data to produce posterior distributions that reflect both the known variability in the response and the uncertainty in the model structure. So instead of using single point estimates for the model's regression coefficients, this approach allows for a distribution of possible coefficients later used in the QMRA. A fixed effects approach was chosen to contain generalizability beyond the present data and amongst the different technologies, allowing for the reuse of the same approach in order to make a fair comparison. Other options such as random effects modelling were considered, where study specific effects as random effect to account for variability and uncertainty in the metadata might show potential. However, due to the lack of sufficient data and generalised grouping, posterior estimates based on study or strain variability would be highly uncertain, reducing the

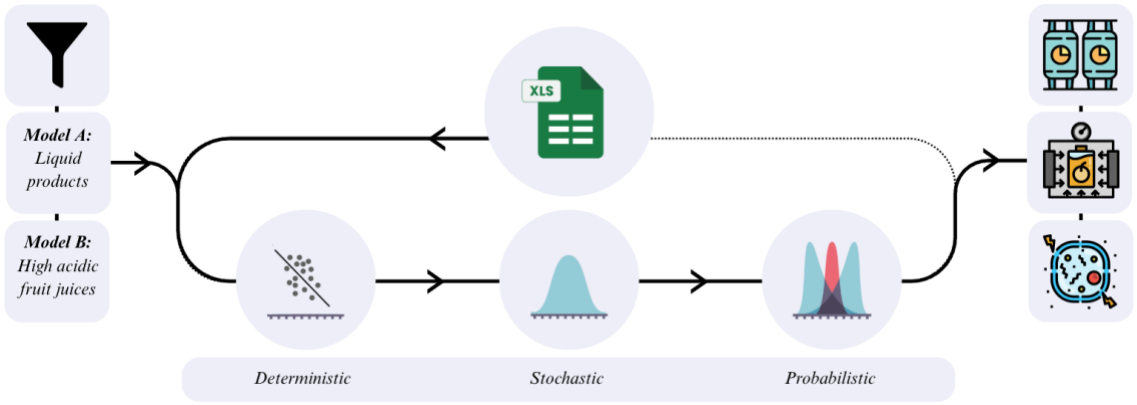


Figure 2, Model fitting approach

based on the frequentist estimates. Specifically, the intercept and slope priors were specified as normal distributions centred on their frequentist estimates, with standard deviations equals to their respective standard errors.

For intercept:

$$\beta_0 \sim N(\beta_0, SE_{\beta_0}) \quad [9]$$

For slope:

$$\beta_1 \sim N(\beta_1, SE_{\beta_1}) \quad [10]$$

For residual variability:

$$\sigma \sim N(0, \sigma_f\ ) \quad [11]$$

This setup ensures that the priors are informative enough to reflect the structure observed in the data, while still allowing the model to update these beliefs when confronted with the actual observations. The posterior distribution for the model parameters is then derived from the combination of the data likelihood and the specified priors:

$$P(\beta_0, \beta_1, \sigma|data) \propto P(data|\beta_0, \beta_1, \sigma) * P(\beta_0) * P(\beta_1) * P(\sigma) \quad [12]$$

The Bayesian regressions were implemented using the BRMS (Bayesian regression modelling using Stan) package in R. Monte Carlo sampling was performed using four independent chains, each with 4000 iterations, of which the first 2000 iterations were used as warm-up, following standard guidelines for convergence and posterior estimation (McElreath, 2015). Posterior predictive checks were performed to evaluate whether the simulated data generated by the model matched the observed distributions. In the context of this study, the Bayesian regression approach captures uncertainty in model coefficients, next to the data-level variability. While the stochastic model simulate the spread in observed D-values at a given condition, the Bayesian models reflect how uncertain one is about the true values of the slope and intercept that describe the process due to limitations in the data. This distinction is critical when later translating model results into risk estimates, as uncertainty in model structure may propagate differently than natural variability alone. The difference between this Bayesian approach and the stochastic model defined earlier thereby quantified the effect of model uncertainty in the QMRA.

2.2.3 Variability assessment

To evaluate potential sources of unexplained variability in the linear regression models, a residual variance analysis was conducted. After fitting the deterministic model for each scenario, the model residuals were analysed using a linear model with metadata covariates as explanatory variables, including *Strain*, *Food matrix*, *pH of the food*, and *Study*. This approach allowed for quantifying the proportion of residual variance explained by each metadata factor using ANOVA. The aim was to distinguish the contribution of general experimental factors (e.g. strain differences or matrix effects) from technology-specific sources of variability. Only the most influential parameters were included per technology, and variables that explained no meaningful variability were omitted from the final results. This ensured a focused interpretation of the key drivers of kinetic variation, without overfitting the model with non-informative covariates.

2.3 Quantitative microbial risk assessment

To evaluate the magnitude of the effect of variability on the efficacy of TP, HPP and PEF in reducing the risk of STEC in fruit juices, a QMRA was developed. The framework was specifically designed to integrate the outputs of the deterministic, stochastic and Bayesian models into a structured simulation, allowing for direct comparison of model predictions across the technologies and product specificities. The framework was focussed on European fruit juice consumption and was implemented in R. To cover the full range of possibilities, the framework consisted of a full Monte Carlo simulation with 1 million iterations per scenario. Random sampling from defined input distributions was used to transmit variability in key components, including initial contamination, serving sizes, and process characteristics. The QMRA followed five sequential modules: 1) prevalence and initial contamination - describing the probability of contamination and the initial STEC load in untreated fruit juice; 2) processing and inactivation – modelling the microbial inactivation under specific processing conditions using the three defined modelling approaches; 3) exposure assessment – simulating consumer level exposure by combining demographic data with per-serving consumption volumes; 4) hazard characterisation – applying a dose-response model to estimate the probability of illness from an ingested dose of STEC; 5) risk characterisation – estimating the number of cases across the European population based on the cumulative outputs of the previous findings. To account for variability between individual simulations, 5 replicates were conducted after which the results could be presented as means and standard deviations between the 5 replicates. All probability distributions, assumptions and data sources used across the QMRA framework are summarised in Table 1.

2.3.1 Prevalence and initial concentration

The probability of STEC contamination in fruit juice was estimated using data from the EFSA zoonoses monitoring dashboard (EFSA, 2019). Between 2019 and 2023, 778 samples from the category ‘ready-to-eat fruits, vegetables, and juices’ were collected across eight European countries. Among these, one sample tested positive for STEC, resulting in an observed prevalence of 0.13% (1 out of 778). To model the initial contamination of STEC in contaminated samples, it was assumed that the presence of a single cell in 25 mL of product corresponds to a positive detection outcome. This assumption aligns with current detection limits for STEC in food safety monitoring. Since the EFSA sampling data are presence-absence based, the actual concentration in positive samples was further refined using available outbreak data and expert based assessments. A PERT distribution was used to reflect the possible range of initial concentration in contaminated servings, expressed in log CFU/mL. This distribution was informed by dose-estimates from outbreak reconstructions (Strachan et al., 2005) and the EFSA pathogenicity assessment for STEC (Koutsoumanis et al., 2020a). This distribution had a minimum of -1.4 log CFU/mL (5 CFU per serving of 125 mL), corresponding to the regulated detection limit of 1 cell in 25 mL of product, a mode of -1.2 log CFU/mL (8 CFU per serving of 125 mL) and a maximum of -0.5 log CFU/mL (39 CFU per serving of 125 mL). These values correspond to total doses between 0 and 39 CFU per serving of 125 mL, which is consistent with outbreak-derived exposure estimates. To account for natural variation in serving sizes, the consumption module (section 2.3.3) incorporates a normally distributed serving size with a mean of 125 mL and a standard deviation of 22 mL.

2.3.2 Fruit juice processing parameters

For each of the three technologies, industrially realistic process conditions were defined based on literature and expert consultation, and variability in operational parameters was incorporated

using probability functions. The goal of this step was to simulate industrially relevant conditions for fruit juice processing in a manner that reflects both typical practices and process variability.

For fruit juices, thermal processing is often applied as high temperature, short processing time (HTST). According to the EFSA Biohazards panel, typical pasteurisation parameters that achieve sufficient reduction of vegetative pathogens range between 70 and 72 °C for 15 to 30 seconds (Koutsoumanis et al., 2020a). For this study, the central processing values were chosen as 72°C and 0.5 minutes (30 seconds), consistent with the mid-range of common pasteurisation conditions. To simulate natural variability in processing, normal distributions were applied to both the temperature and treatment time. The standard deviations were based on assumptions regarding temperature fluctuations in thermal processing lines, as described by Den Besten et al. (2017),

Table 1, QMRA inputs, distributions and equations

ID	Input variable	Distribution or Formula	Unit	Source/reference
Initial concentration & prevalence				
$\log_{10} N_0$	Initial concentration	$\sim PERT\,(-1.4, -1.2, -0.5)$	log CFU/mL	EFSA, 2020
Prev	Prevalence	0.13	%	EFSA, 2023
Inactivation by HPP				
t	Treatment time	$\sim N(\mu = 8, \sigma = 0.1)$	Minutes	Mode, metadata
P	Pressure	$\sim N(\mu = 550, \sigma = 0.1)$	Mpa	Anonymous HPP expert (2024)
Inactivation by TP				
t	Treatment time	$\sim N(\mu = 0.5, \sigma = 0.016)$	Minutes	Duan et al., 2011
T	Temperature	$\sim N(\mu = 72, \sigma = 0.1)$	Degrees celsius	Den Besten et al., 2017
Inactivation by PEF				
E_{in}	Energy input	$\sim N(\mu = 120, \sigma = 0.1)$	kJ/L	Aganovic et al., 2022
Microbial inactivation				
$\log_{10} D_T$	Log Decimal reduction	$log_{10}D_T = model\ slope * T + model\ intercept + \sigma$(optional)		Calculated
$\log_{10} D_P$	Log Decimal reduction	$log_{10}D_P = model\ slope * P + model\ intercept + \sigma$(optional)		Calculated
$\log_{10} N_t$	Concentration after treatment	$logN_t = logN_0 - \frac{t}{D}$ $OR\ logN_t = logN_0 - \frac{E_{in}}{D}$	log CFU/mL	Calculated
R	Reduction	$\log(\frac{10^{logN0}}{10^{logNt}})$ $OR\ Model\ slope * E_{in}$	log CFU/mL	Calculated
Dose-Respone model				
r (STEC)	Dose-Respone coefficient	0.00208	-	Strachan et al., 2005
Dose	Exposure dose	$10^{Nt} * Serving\ size$	CFU/serving	Calculated
$P_{illness}$	Probability of illness	$P_{illness} = 1 - e^{-r*Dose}$	Exponential model	Calculated
$R_{serving}$	Risk per serving	$P_{illness} * P_{rev}$		Calculated
Demographic data				
Pop	European population 2023	446000000	People as of 2023	Eurostat, 2023
Pop_{adults}	EU population: Adults	58.6	%	Eurostat, 2023
Pop_{teens}	EU population: Teens	5.2	%	Eurostat, 2023
$Pop_{children}$	EU population: Children	14.9	%	Eurostat, 2023
Consumption data				
Cons	Annual fruit juice consumption per capita	15.4	Liters fruit juice	Statista, 2023
$Serv_{adults}$	Average serving size adults	150	mL	Walton & Kehoe, 2024
$Serv_{teens}$	Average serving size teens	117	mL	Walton & Kehoe, 2024
$Serv_{children}$	Average serving size children	108	mL	Walton & Kehoe, 2024
Ss	Serving size	$\sim N(\mu = 125, \sigma = 22)$	mL	Assumption
$EU_{servings}$	Annual servings in Europe	$EU_{servings} = \sum_{i=1}^{n} \frac{Cons}{S_i} * P_i$	Servings	Pi = Population in target group i, Si = Serving size in target group i
Illness occurance				
C_{mean}	Cases calculated based on arithmetic mean	$\mu(Risk_{serving}) * EU_{servings}$	Cases per year	Based on mean Risk
$C_{bernoulli}$	Estimated cases based on bernoulli trials	$Bernoulli\ (<P_{illness}), if\ illness = 1, not = 0$	Cases per year	Calculated by summation with n = 1.000.000 iterations and multiplied by Prev and EU_servings

where temperature fluctuations of ±0.1°C were assumed. The distributions for thermal processing in this study were thereby:

$$Temperature \sim N(72\ ℃, 0.1\ ℃) \quad [13]$$

$$Treatment\ time \sim N(0.5\ min,\ 0.016\ min) \quad [14]$$

For HPP, the absence of a universally standardizes process made parameter selection mode dependent on expert knowledge. Input values were derived from practical experience in HPP operations, where pressure and treatment time are typically tightly controlled. The expert indicated that most in commercial juice applications, processes are set at 600 MPA for durations ranging between 3 and 8 minutes. However, both the literature and data suggest that for high acidic fruit juice, a pressure of 550 MPa applied for 8 minutes was the most common combination. These values were therefore selected as the central process conditions in the simulation. Since modern HPP equipment achieves and maintains setpoints with minimal deviation, a standard deviation of 0.1 MPa for pressure and 0,1 minute for treatment time was considered appropriate (to correct for pressure release).

$$Pressure \sim N(550\ MPa,\ 0.1\ MPa) \quad [15]$$

$$Treatment\ time \sim N(8\ min,\ 0.1\ min) \quad [16]$$

For PEF, the microbial inactivation is directly modelled as function of volumetric energy input in kJ/L, which integrates all relevant electrical and treatment parameters. Based on Aganovic and Smetana (2021), an energy input of 120 kJ/L was selected as a realistic value for the treatment of heat-sensitive fruit juices, aiming for sufficient inactivation of vegetative pathogen (Aganovic et al., 2021). Given that the pulse generation and delivery in industrial PEF systems is well regulated and stable, variability in energy input was assumed to be minimal and modelled with a standard deviation of 0.1 kJ/L.

$$E_{in} \sim N(120\ KJ.L^{-1}, 0.1\ KJ.L^{-1}) \quad [17]$$

2.3.3 Monte Carlo simulations for accounting variability in kinetic parameters

To accurately assess the impact of variability on the risk estimates, the three modelling approaches (deterministic, stochastic and probabilistic) were implemented in the QMRA framework. These approaches correspond to increasing levels of complexity in how variability and uncertainty are incorporated into the inactivation estimates. While each model used the same QMRA framework and the same input distributions for prevalence, initial contamination, consumption and dose-response modelling, they differed in how the inactivation step was parameterised. This allowed to isolate the influence of the inactivation model itself on the final risk predictions. In all three approaches, the Monte Carlo simulation accounted for the variability in industrial processing conditions as defined in previous sections. However, the representation of the inactivation kinetics differed.

In the deterministic approach, inactivation was calculated using point estimates from the best-fit regression model for each technology. The predicted $\log_{10} D$ or $\log_{10} Reduction$ was fixed for each iteration based on the specific process condition (e.g. temperature or pressure), without incorporating variability in the kinetic model itself. However, all other inputs such as treatment conditions, serving size, and initial concentration remained variable. This setup ensured that the impact of a fixed inactivation estimate could be assessed under realistic, variable conditions. In the stochastic model, the regression prediction was supplemented with a residual error term representing unexplained variability in the metadata. For each iteration, $\log_{10} D$ or

$\log_{10} Reduction$ values were sampled from the distributions defined before. This allowed the simulation to reflect the spread observed in the literature due to unmodelled factors such as strain or matrix differences. The probabilistic approach used Bayesian regression to account for both variability and uncertainty in the model parameters. For each iteration, regression coefficients were sampled from their posterior distributions, and a normally distributed residual error term was added. This allowed the simulation to explore a full range of plausible model predictions, where broader posteriors reflected more uncertainty due to limited or scattered data.

In all models, microbial inactivation was translated into the STEC concentration after treatment (N_t) and final dose. To prevent over estimation of the microbial reduction, the overall reduction was calculated based on the arithmetic mean of survivors:

$$Reduction_{overall} = -\log_{10}\left(\frac{mean(N_t)}{mean(N_0)}\right) \quad [18]$$

which was found to be consistent with the probability integral method explained in section 1.4 (Zwietering et al., 2021). By using consistent inputs and structure, differences in QMRA outcomes could be directly attributed to the treatment of variability and uncertainty in inactivation kinetics. This allowed a fair assessment of each modelling approach across the three technologies and product specificity levels.

2.3.4 Demographic data and fruit juice consumption

To estimate the annual exposure, demographic and consumption data were sources from Eurostat (2024) and Statista (2023). As of 2023, the European Union counts 446 million residents, of which 58.6% are adults, 5.2% are teens and 16.9% are children (Eurostat, 2024). In 2023, the average annual fruit consumption of fruit juice per capita was estimated on 15.4 L in Europe (Statista, 2023). The estimated annual servings were estimated based on the average portion sizes reported by Walton and Kehoe in 2024, where they found that the average portions sizes for adults, teens and children are 150, 117 and 108 mL respectively (Walton & Kehoe, 2024). The total annual servings in Europe was estimated by dividing the average fruit juice consumption by the average portion size per category. Then these number of servings was multiplied by the number of people in that category. Summing the servings of the three respective categories leads to the estimated annual number of fruit juice servings (~46 billion). In the QMRA, a variable serving size was used, with a mean of 125 mL and standard deviation of 22 mL, based on the average portion sizes reported by Walton and Kehoe (Walton & Kehoe, 2024).

2.3.5 Hazard characterisation

To evaluate the dose-response models and identify the most appropriate fit for low-dose scenarios relevant to the (QMRA), a comparative approach was applied, based on the outbreak data reported by Strachan et al. (2005). Table 2 presents these estimated doses and their corresponding estimated illness probabilities (Strachan et al., 2005). Nine different dose-response models were considered and fitted against the outbreak data. These models included both exponential:

$$P_{ill} = 1 - e^{-r*Dose} \quad [19]$$

as Beta-Poisson models:

$$P_{ill} = 1 - \left(1 + \frac{Dose}{\beta}\right)^{-\alpha} \quad [20]$$

The fit of the models was estimated by calculating the residual sum of squares (RSS) and the Akaike Information Criterion (AIC). The models were fitted on both the arithmetic illness probability and the log-transformed illness probability, as this approach might be more

representative for low-dose (<50 CFU) cases. The dose-response model with the lowest AIC based on the log transformed illness probability data, will be used as default in the QMRA.

2.3.6 Risk characterisation

For each combination of model type and specificity level within each technology, Monte Carlo simulations with 1 million iterations were performed. Each iteration represented an individual serving scenario, incorporating variability in contamination, processing conditions, dose-response, and consumption. The probability of illness for each serving was calculated and combined with the prevalence of contamination to determine the individual risk. The annual number of illness cases was then estimated using three independent approaches, each providing a different interpretation of the risk outputs and ensuring robustness in the final conclusions. The first approach used Bernoulli trials to simulate illness outcomes based on the probability of illness per serving. For each iteration, a binary outcome was generated by drawing a random number and comparing it to the predicted illness probability. If the random draw was lower than the probability, the iteration was marked as an illness case (1), otherwise as no case (0). The total number of cases was then scaled by multiplying the proportion of illness events with the prevalence of contamination and the total number of annual servings in the European population. This method captures the full stochastic character of low-probability events and is expressed as:

$$Cases_{Bernoulli} = \frac{\left(\sum_{i=1}^{n_{iterations}} 1(Bernoulli_{,i} > 0)\right) * prevalence * EU_{servings}}{n_{iterations}} \quad [21]$$

The second approach was based on the arithmetic mean risk per serving. For each iteration, the probability of illness was multiplied by the contamination prevalence to yield the risk per serving. The mean risk across all iterations was then scaled up using the annual number of servings. This method directly integrates the full distribution of serving-level risk:

$$Cases_{Average} = \overline{Risk_{serving}} * EU_{servings} \quad [22]$$

The third approach was based on the probability integral method, which calculates the overall reduction by using the arithmetic mean of the surviving pathogens after processing. From each iteration, the surviving concentration (N_t) was recorded, and the overall mean (N_t) was used to estimate a mean dose across the population. This average dose was then entered into the exponential dose-response model to calculate an overall probability of illness, which was scaled by the contamination prevalence and number of annual servings:

$$Cases_{Overall} = \left(1 - e^{-r*(\overline{Nt}*\overline{Serving_{size}})}\right) * prevalence * EU_{servings} \quad [23]$$

These three approaches reflect different perspectives on the use of variability and probability in risk estimation. The Bernoulli method models illness occurrence as a random binary outcome per serving, the average-risk method integrates probabilities across all servings, and the probability

Table 2, Outbreak data for STEC, as reported in Strachan et al., 2005

Outbreak	Outbreak site	Vehicle	Estimated dose	Total subjects	Infected subjects	Estimated probability
Strachan et al., 2001	UK, New Deer	Sheep fasces/soil	14	228	20	0.0877
Nauta et al., 2001; Shinagawa et al., 1997	Japan, Moroika	Salad/seafood sauce	31	871	215	0.2468
Keene and Sazie, 1997	USA, Oregon	Deer jerky	10000	12	10	0.8333
Uchimura et al., 1997	Japan, Kashiwa	Melon	1100	71	32	0.4507
Bell et al., 1994; Tuttle et al., 1999	USA, Wasington	Hamburger	23	5634	398	0.0706
Tilden et al., 1996	USA, Wasington/ California	Salami	23	2778	17	0.0061
Warner et al., 1995	USA, illinois	Water	75	2350	12	0.0051
Anon., 1997	UK, Wyre	Cheese	380	360	2	0.0056

integral method focuses on the overall effect of microbial inactivation. By using three methods across these methods, the study ensures the robustness and consistency of its risk predictions across a wide range of model structures and assumptions.

2.4 Sensitivity analysis

To further assess the impact of variability on the risk outcomes, a sensitivity analysis and probability threshold evaluation were performed. These analyses aimed to identify the most influential parameters affecting the probability of illness and to determine the probability of achieving a safe microbial reduction, given the variability in the predictive models.

2.4.1 Spearman's rank correlation

A Spearman's rank correlation analysis was conducted to assess how variation in input parameters affects variation in model outputs, specifically the probability of illness. This approach was selected because it provides a universal, rank-based measure that does not assume linearity or normally distributed data, making it applicable across all model types and input variables. The analysis was applied across all technologies and modelling approaches, including the deterministic models. While the kinetic predictions ($\log_{10} D$ or reduction) in the deterministic models are fixed, other input parameters (e.g. initial concentration, serving size, treatment conditions) remained variable. This enabled a meaningful sensitivity analysis for deterministic scenarios as well. The analysis framework was custom made in R using the purrr package.

2.4.2 Safe threshold evaluation

In addition to the correlation-based sensitivity analysis, a probability threshold evaluation was performed to assess the likelihood of achieving different log-reduction targets under realistic processing conditions. This analysis helps to evaluate how reliably each model achieves microbial safety benchmarks. A well-known rule of thumb to control general food safety in designing processes is a 5-log reduction, as recommended by the FDA (FDA, 2024). The evaluation was conducted using the stochastic linear regression models for TP, HPP, and PEF, which include variability in their predictions through a normally distributed error term. For each model, a Monte Carlo simulation with 100,000 iterations was performed, simulating $\log_{10} D$ -values at the respective processing input (temperature, pressure, or energy). For each processing time (TP, HPP) or energy input (PEF) in a predefined range, the corresponding achieved log-reduction were calculated. Next, for each point, the probability of achieving a minimum target reduction (e.g., 2-, 3-, 4-, or 5-log reduction) was calculated as the proportion of iterations where the achieved log-reduction was reached or exceeded that threshold. This resulted in a probability curve that captures the likelihood of sufficient inactivation at each treatment duration:

$$P(R \geq r) = \frac{1}{n}\sum_{i=1}^{n} 1\left(\frac{t}{D} \geq r\right) \quad [24]$$

where r is the log-reduction target and t the processing time (energy input for PEF). To provide variability bounds for each probability, the standard deviation of the binary outcomes (reduction met or not) was calculated and used to estimate lower and upper 95% confidence intervals. This approach was applied to both Model A (all liquid products) and Model C (high-acid fruit juices), using their respective regression coefficients.

Chapter 3

Results & Discussion

3.1 Thermal Processing

3.1.1 Modelling kinetic parameters: $log_{10} D$ estimation

As thermal processing serves as the baseline for pasteurisation and is the most used method for microbial inactivation, it provides a reference point for evaluating variability and model performance before comparing non-thermal technologies. Table 3 summarises the regression coefficients for the deterministic models. Despite the filtering of the data towards product specificity, a steeper slope was found in the generic model where data for all liquid products was used, meaning a stronger temperature effect was obtained. The variance displayed is defined as the average squared distance between the frequentist regression line and the actual datapoints (σ^2), explaining how much unexplained variability in $\log_{10} D$ prediction remains after accounting for the effect of temperature. One should note that these values are in squared log-units, making them less interpretable. Therefore, the residual standard error (RSE) explains the typical deviation of the observed $\log_{10} D$ values from the frequentist regression line. In both cases, a lower value was found for the model for high acidic fruit juice, explaining a lower overall variability in the model fit. The $\log_{10} D$ prediction models for all liquid products and high acidic fruit juices are displayed in Figure 3. As explained, three approaches for $\log_{10} D$ prediction were applied. The deterministic approach (solid black line) provides a baseline estimate of $\log_{10} D$ as function of temperature, where no variability or uncertainty was included. The stochastic approach, extended by the 95% prediction intervals (dashed black lines) adds variability in outcomes around the deterministic line, meaning each temperature input gets a distribution of possible $\log_{10} D$ values. The Bayesian probabilistic regressions (thin red lines) incorporate parameter uncertainty by drawing from posterior distributions for the model slope, intercept and residual variance. As there might be less natural variability present in the model for high acidic fruit juice, fewer datapoints have been used, meaning there is more uncertainty in the slope and intercept of the model. This explains the wider spread in the Bayesian lines relative to the 95% prediction interval. The breakdown of the posterior outputs can be found in appendix 1. The contrast between variability and uncertainty becomes clearer when interpreting the graphs in Figure 3 with the regression summary in Table 3. In the model for all liquid products (Figure 3A), the broader scatter of individual datapoints and the higher RSE (0.441) reflect a greater degree of natural variability around the $\log_{10} D \sim temperature$ relationship. This results in wider prediction intervals. However, in the Bayesian approach, the larger dataset provides stronger validity about the model coefficients, reducing parameter uncertainty. In contrast, the model for high acidic fruit juices (Figure 3B) includes fewer datapoints. Despite showing a lower RSE (0.307), the limited sample size increases the uncertainty in the regression parameters, resulting

Table 3, Regression coefficients and parameters for the deterministic models by thermal processing

Model & Specificity	Intercept	Slope	Datapoints	RSE	Variance	z-value (°C)
Thermal processing						
All liquid products	8.600	-0.14028	224	0.441	0.194	7.1
High acidic fruit juices	6.588	-0.10832	58	0.307	0.092	9.2

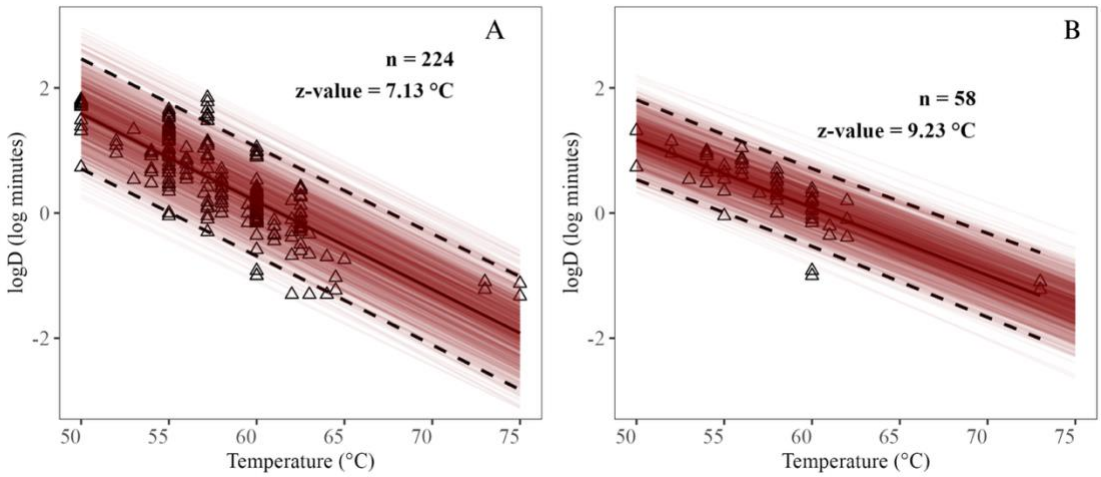


Figure 3, Thermal inactivation models based on collected metadata showing the relationship between temperature and log_{10} D values for all liquid products (left) and high acidic fruit juices (right). Each panel displays three layers of model output: a frequentist point estimate (solid black line), the 95% prediction interval (dashed black lines) and the posterior draws from the Bayesian regression (thin red lines).

in a notable wider distribution of potential $\log_{10} D$ outcomes for a given temperature. This is explained further in the QMRA section of thermal processing.

3.1.2 The overall impact on the estimated risk: a deterministic approach

The deterministic approach provides a baseline risk estimation by applying fixed model parameters without incorporating variability and uncertainty in kinetic predictions. By using the models and their respective parameters in the QMRA framework, the predicted overall reduction, risk of illness, and estimated number of cases on the European level were determined. Table 4B summarises the overall reduction in STEC levels across the two specificity levels based on 5 repeated Monte Carlo simulations, each with 1 million iterations. The estimated overall reductions were 15.2, and 8.0 log CFU/mL in the models for all liquid products and high acidic fruit juices respectively. These results align with the mean reduction over all iterations, although minor variability was found between the five simulations. These results also indicate that the more specific model for high acidic fruit juices predicts a significantly lower microbial reduction compared to the generalised scope of all liquid products. The percentile distributions of the overall reduction in Table 4A show minimal increases, which is expected since the temperature, treatment time, and serving size were treated as variable inputs reflecting realistic variability. Due to the relatively high overall reduction in the two levels of specificity, the final average STEC dose was still found to be reaching 0 ($5.85 \cdot 10^{-15}$ and $9.84 \cdot 10^{-8}$) CFU per serving. The most crucial outcome

Table 4a, Microbial reduction based on the arithmetic mean of survivors (TP). Results were based on the mean of 5 repeated simulations with each 1,000,000 iterations and displayed for the 2.5, 50, 97.5, and 99th percentile.

Model & Specificity	Mean reduction	Overall reduction (R)	R(2.5%)	R(50%)	R(97.5%)	R(99%)
Level 1: All Liquid Products						
Linear model (Deterministic)	15.8 ± 3.2	15.2 ± 0.0	14.4 ± 0.0	15.8 ± 0.0	17.3 ± 0.0	17.5 ± 0.0
Linear model (Stochastic)	25.7 ± 31.2	2.9 ± 0.0	2.2 ± 0.0	15.8 ± 0.0	115.7 ± 0.4	167.9 ± 0.6
Bayesian model (Probabilistic)	32.0 ± 47.5	1.9 ± 0.0	0.7 ± 0.0	14.5 ± 0.0	286.2 ± 1.0	Inf ± 0.0
Level 2: High Acidic Fruit Juice						
Linear model (Deterministic)	8.1 ± 1.6	8.0 ± 0.0	7.5 ± 0.0	8.1 ± 0.0	8.8 ± 0.0	8.9 ± 0.0
Linear model (Stochastic)	10.5 ± 8.9	3.0 ± 0.0	2.0 ± 0.0	8.1 ± 0.0	32.5 ± 0.1	42.2 ± 0.1
Bayesian model (Probabilistic)	21.0 ± 37.5	1.4 ± 0.0	0.3 ± 0.0	7.5 ± 0.0	179.2 ± 1.1	Inf ± 0.0

Table 4b, QMRA results for the thermal processing models. Results were based on 5 repeated simulations with each 1,000,000 iterations. Mean serving risk defined as the probability of illness corrected for the prevalence. The sum of Bernoulli trials present positive illness cases before prevalence per million iterations.

Model & Specificity	Mean serving risk	Sum bernoulli trials	Cases / million servings	Annual cases EU (bernoulli)	Annual cases EU (mean risk)
Level 1: All Liquid Products					
Linear model (Deterministic)	$2.9 \cdot 10^{-19} \pm 0.0$	0.0 ± 0.0	0.0 ± 0.0	0 ± 0	0 ± 0
Linear model (Stochastic)	$3.8 \cdot 10^{-8} \pm 4.0 \cdot 10^{-10}$	27.6 ± 4.7	0.0 ± 0.0	1412 ± 242	1477 ± 19
Bayesian model (Probabilistic)	$3.9 \cdot 10^{-7} \pm 2.1 \cdot 10^{-9}$	296.8 ± 13.5	0.4 ± 0.0	15187 ± 692	15274 ± 82
Level 2: High Acidic Fruit Juice					
Linear model (Deterministic)	$2.7 \cdot 10^{-13} \pm 0.0$	0.0 ± 0.0	0.000 ± 0.0	0 ± 0	0 ± 0
Linear model (Stochastic)	$3.1 \cdot 10^{-8} \pm 3.4 \cdot 10^{-10}$	21.8 ± 5.4	0.028 ± 0.0	1116 ± 279	1203 ± 13
Bayesian model (Probabilistic)	$1.1 \cdot 10^{-6} \pm 2.9 \cdot 10^{-9}$	854.2 ± 19.8	1.110 ± 0.0	43709 ± 1015	43948 ± 114

of a QMRA is the estimated number of foodborne illness cases, relative to the targeted population. The results of the QMRA for thermal processing are summarised in Table 4B, displaying estimate risk of illness, illness cases per million servings and the total estimated cases at European level. In both models, the risk of illness was found to be extremely low, effectively approaching zero for practical purposes. For both deterministic scenarios, zero cases were estimated due to their relatively higher overall reduction, suggesting a sufficient pasteurisation process across all levels of specificity. Another important finding is that when the variance might be reduced in the specific model, the temperature effect due to the model slope is the main factor determining the final model outcome. To further investigate the influence of the individual model parameters on the final risk estimation, a Spearman's correlations analysis was conducted (Figure 4), based on all the simulated data. For the deterministic influences (dark grey), the strongest correlation was observed between the processing time and temperature on the probability of illness, reflecting the impact of process variability on the estimated probability of illness. Serving sizes and initial contamination also were a major driver of the estimated probability, as larger contamination levels increase estimated exposure levels. The deterministic QMRA results demonstrate that no illness cases were found across both levels of specificity, indicating that using this simple and frequentist approach might indicate sufficient efficacy for thermal processing. The regulatory threshold of reaching a 5-log reduction was achieved for all deterministic scenarios (FDA, 2024). This aligns with the findings from the meta-regression models in Figure 3, where the increased specificity resulted in narrower prediction intervals and lower estimated thermal resistance as well. However, the deterministic approach has the limitation of not accounting for the full extent of variability in

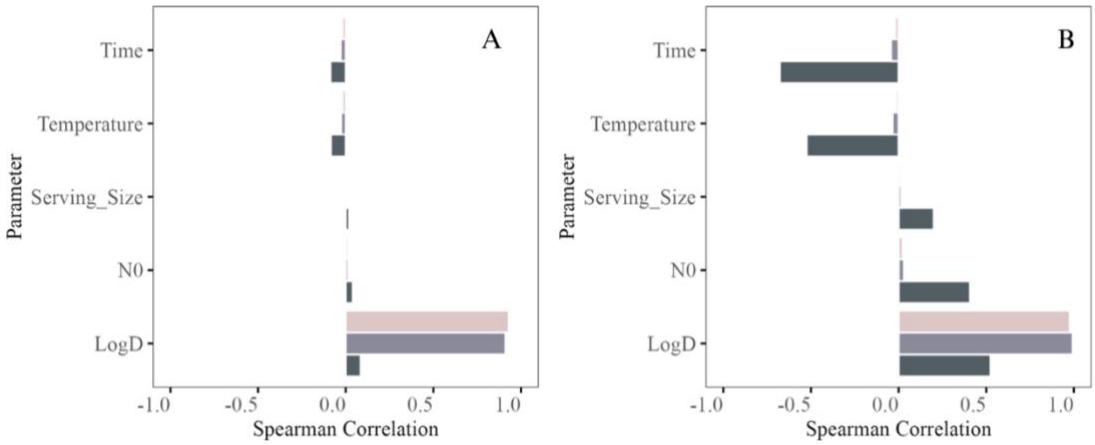


Figure 4, Spearman's correlation coefficients for all liquid products (A), and High acidic fruit juices (B). Correlations present the impact of variability in input parameters on the probability of illness. Analysis performed on deterministic models ,Stochastic models , and probabilistic models . Results displayed for TP.

its predictions. While both the specificity levels suggest virtually no illness risk, this might be overly optimistic. This because these models rely on single point estimates for $\log_{10} D$ and z-values, ignoring the influence of variability in kinetic parameters and environmental factors. As a result, this approach might truly underestimate the final risk in real-world applications. This is also evident in the sensitivity analysis in Figure 3, where the deterministic approach highlights process parameters and initial contamination as key drivers, but fail to capture the influence of the variability in the $\log_{10} D$ prediction models. This issue will be further explored in the following sections where stochastic and probabilistic models do incorporate variability into risk estimations.

3.1.3 The impact of variability on the estimated risk: a stochastic and probabilistic approach

Building on the deterministic findings, this section continuous with evaluating the stochastic and Bayesian probabilistic approaches, which incorporate variability and uncertainty in kinetic parameters in their risk predictions. These models account for the variability in the $\log_{10} D$ prediction, process deviations and uncertainties in regression parameters, providing a more realistic risk estimation. Next to the deterministic results, Table 4A and 4B are also showing the results for the stochastic and probabilistic approaches. An important result is the difference between the mean reduction, defined as the average reduction over 1 million iterations, and the overall reduction, based on the arithmetic mean of survivors amongst those 1 million iterations. Where the mean reduction increases when introducing variability and uncertainty into the model, does the overall reduction decrease significantly. The increase in mean reduction is explained by the small but relevant heavy tails in the reduction distribution, whereas the overall reduction accounts for the surviving cells. In the model made for all liquid products, the estimated overall log reduction dropped from 15.2 log CFU/mL to 2.9 log CFU/mL (stochastic) and 1.6 log CFU/mL (Bayesian). This large decrease in reduction directly points out the effect of considering variability and uncertainty in the prediction model, taking the process-resistant cells into account. As visible in the percentiles, the overall reduction follows a distribution, where extremely high and low values are found in the tails. These extremely low values are often found for the most process-resistant strains (Zwietering et al., 2021). Therefore, one should be careful by simply taking the mean of kinetic parameters when modelling based on reported data. The lower overall reductions are directly connected to the estimated annual number of STEC cases on the European scale. In the models based on data for all liquid products, the estimated number of cases was increasing from zero to ± 1400 for the stochastic approach, to even ± 15000 in the Bayesian approach, pointing out a large effect of variability on the efficacy of thermal processing. In the models based on the data for high acidic fruit juice, the estimated number of cases increased to ± 1200 for the stochastic model and ± 43000 for the probabilistic model. It is important to realise that the difference in overall z-values (the temperature effect) plays a crucial role in the final inactivation and risk estimates. In the stochastic approach, fewer illness cases were observed for the model based on all liquid products compared to the high acidic fruit juice model. This can be attributed to the higher residual variance in the former, which leads to a broader distribution of $\log_{10} D$ values and consequently a higher proportion of iterations achieving sufficient inactivation. In contrast, the high acidic fruit juice model, despite having a lower residual variance, results in a narrower distribution around the mean $\log_{10} D$ estimate, increasing the likelihood of partial inactivation and illness cases. An opposite trend is present in the Bayesian probabilistic models, where the number of illness cases is higher for high acidic fruit juices than for all liquid products. This can be explained by the additional layer of uncertainty introduced through the Bayesian framework, which accounts not only for residual variability but also for uncertainty in the regression parameters. Fewer datapoints were available for the high acidic fruit

juice model, making the posterior distributions for the slope and intercept wider, leading to greater uncertainty in model predictions and thereby a higher probability of insufficient inactivation across the iterations. This effect is further explored in Figure 5, where these distributions are visualised.

Spearman's correlations matrices (Figure 4) provide further insights into the impact of variability on the probability of illness across these modelling approaches. Unlike the deterministic approach, where the process parameters were the dominant predictor of illness probability, the stochastic and probabilistic models highlight the variability in kinetic predictors ($\log_{10} D$) as strongest factor. Despite the temperature effect on the $\log_{10} D$ prediction, the variability in this reported relation can play a larger role. This is found in the risk estimation as well, where the stochastic model for all liquid products have led to a comparable estimation as found in the model for high acidic fruit juice, indicating that the variability can overcome the $\log_{10} D$ precision. This aligns with earlier research towards the effect of both variability and uncertainty, where was found that these can have an equal or even higher relevance than kinetic and process parameters themselves in a QMRA, pointing out the significance of including them both in the analysis (den Besten et al., 2017). To further assess where illness cases occur in the simulations of the overall risk, probability density distributions were generated for the most contributing factors and are visualised in Figure 5. Barcode plots illustrate the distribution of Bernoulli trial outcomes (illness cases) for both the stochastic (top) and probabilistic (bottom) approaches, following the conceptualisation proposed by Abe et al. (2023) (Abe et al., 2023). One should realise that these results are directly extracted from the simulations, where prevalence was not considered yet. The occurrence of the positive Bernoulli trials is therefore based on the assumption that every serving was contaminated. The probability density distributions of the risk of illness per serving are visualised for all liquid products (Figure 5A) and high acidic fruit juices (Figure 5B). When using data for all liquids in the modelling approach (Figure 5A), a smaller peak at higher risks is observed compared to the more specific approach. In line with earlier findings, where Bayesian probabilistic models resulted in a higher number of estimated illness cases for high acidic fruit juice, this pattern is also reflected in the distribution of the risk of illness. Despite both distributions being heavily skewed towards zero, a clear difference is present where the model for high acidic fruit juices shows a broader and higher density at elevated risk levels. This peak is also where most of the illness cases occur. For all liquid products, the first illness cases were detected at a risk of $8.3 \bullet 10^{-8}$ (stochastic) , and $1.3 \bullet 10^{-7}$ (probabilistic). This risk distribution substantiates the idea that simply taking the average risks for these models, being $3.7 \bullet 10^{-8}$ and $6.2 \bullet 10^{-7}$ is underestimating true risk based on the true variability. However, this is not necessarily true, since the estimation of cases based on the arithmetic mean of the risk per serving leads to comparable results as the approach based on Bernoulli trials. The effects of these approaches are explained in more detail in section 3.5.3. For all specificity levels, a difference is visible between the stochastic and probabilistic approaches, where a broader risk distribution is detected for the probabilistic approach.

Zooming in towards the underlying levels of the prediction model, the distributions of the microbial reductions on iteration level are presented in Figure 5 C-D, for all liquid products and high acidic fruit juices respectfully. When these distributions are interpreted in the context of the reductions and QMRA results in Table 3B, the differences in model behaviour become clear. The heavy tails at high reductions are increasing the mean reduction significantly, while the large density of insufficient reductions shows an accumulation of illness cases in both modelling approaches. Although the difference in overall reduction is not extreme, a higher probability of achieving both extremely low and high reductions is again present in the Bayesian probabilistic

approaches, leading to the larger number of estimated illness cases. As the model implies, considering the variability in kinetic inactivation data has a significant effect on the model outcome, thus the efficacy of the technology. As identified in the sensitivity analysis (Figure 4), $\log_{10} D$ variability is the most significant contributor to the variability in the probability of illness. To further visualize this relationship, Figure 5E-F presents the $\log_{10} D$ density distributions for each level of specificity, explaining the origin of the distribution of the microbial reduction. These graphs are comparing the stochastic and Bayesian probabilistic models, giving a deeper insight

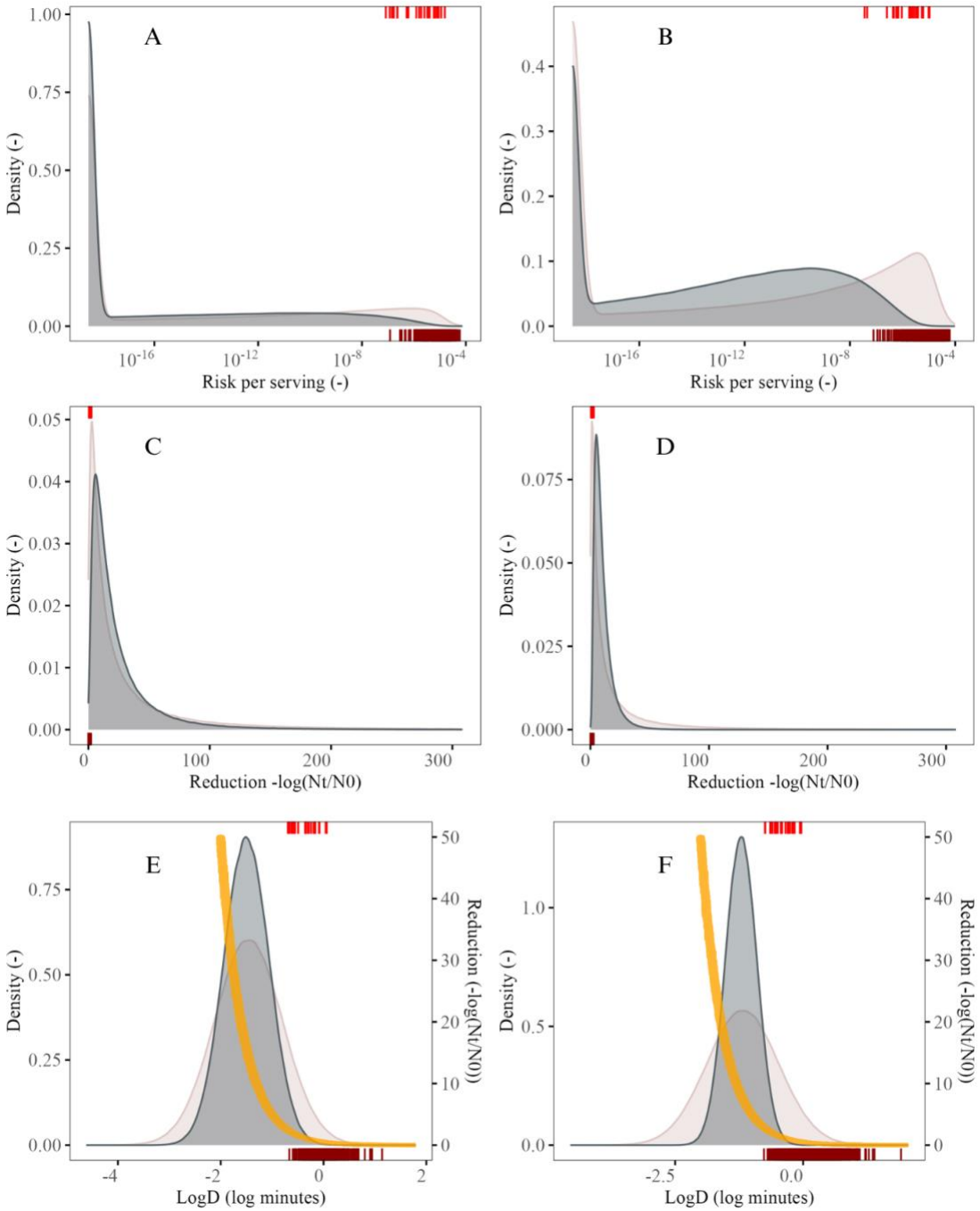


Figure 5, Probability density distributions for the Risk of illness per serving using all liquid products (A), and high acidic fruit juices (B). Probability density distributions for overall reduction using all liquids (C) and high acidic fruit juices (D). Probability density distributions for $\log_{10} D$ using all liquids (E) and high acidic fruit juices (F). Layers present the modelling approach (■ = stochastic, ■ = probabilistic). Barcode plots show the occurrence of illness cases with respect to the x-variable (top = stochastic, bottom = probabilistic). Orange lines represent the reduction of STEC with respect to the $\log_{10} D$ distribution. Results displayed for TP.

in how the variability and uncertainty in the $\log_{10} D$ prediction influences the food safety risk. The x-axis represents the distribution of $\log_{10} D$ estimates, which were obtained by sampling estimated $\log_{10} D$ values from the normal distribution based on the RSE (stochastic) and by sampling regression coefficients from the Bayesian posterior distributions while also incorporating residual variability (probabilistic). The barcode overlays in the density distributions represent the individual Bernoulli trials that resulted in an illness case, pointing out where the illness cases occur. In addition, the orange line (connected to the right y-axis) presents the estimated reduction at the iteration level, as a function of the $\log_{10} D$ distribution. This line effectively demonstrates how changes in $\log_{10} D$ values impact the reduction across the Monte Carlo simulations. Across the two levels of specificity, the Bayesian probabilistic model showed a consistently broader distribution of estimated $\log_{10} D$ values compared the stochastic approach, indicating that this approach captures a wider range of possible $\log_{10} D$ values. This phenomenon is mostly present in the model for high acidic fruit juice, where the use of less datapoints has led to an increased uncertainty in model fit. The barcode plots for the illness cases show that the cases are not evenly distributed across the $\log_{10} D$ range but are more concentrated at the upper tails of the distribution. This highlights that the most process-resistant cells play a dominant role in driving overall risk. Across the two levels of specificity, the first illness case occurred at -0.69, and -0.75 for the stochastic approach, and -0.66, and -0.76 for the Bayesian approach respectively. However, it is important to emphasise that the distributions are not centred around similar mean $\log_{10} D$ estimates. When considering the percentile position, the first illness cases occur at the 96.6%, and 93.5% percentiles for the stochastic approach, and the 88.6%, and 72.2% percentiles for the Bayesian approach, in all liquid products (Figure 5E) and high acidic fruit juices (Figure 5F) respectively. This suggests that even small variability within the dataset is sufficient to produce illness cases, disproving the assumption that only extreme values contribute to risk. Furthermore, these results are evident that filtering the data with respect to food matrix specificity is not always creating more specific and reliable prediction model. Moreover, most illness cases occur from the 95% percentile all the way up to the 99.99% percentile of the $\log_{10} D$ distribution, meaning that simply taking 95% confidence, which is commonly used in risk assessments, is still underestimating the full risk of STEC illness cases.

As recommended by multiple organisations and regulatory agencies, achieving a 5-log (100.000-fold) reduction is commonly used as guideline for evaluating the efficacy of inactivation processes (FDA, 2024). To translate the findings into an industrial relevant context, the probability of achieving different levels of STEC inactivation (2-5 log reductions) were assessed over a range of processing times, based on the stochastic model outputs. Contrary, the last "safe" reduction found in the stochastic model before an illness case occurred were found to be 3.6, and 4.3 log CFU/mL for both specificity levels respectively. These probability functions are visualised in Figure 6A (all liquid products) and Figure 6B (high acidic fruit juices). Each curve represents the estimated probability of reaching a specific log reduction at a given treatment duration, based on processing at 72 degrees Celsius. The shaded regions around each curve reflect the variability resulting from the variability in $\log_{10} D$ prediction, as captured in the stochastic modelling approach. For both models, the probability of achieving lower log reductions increases more rapidly and reaches near full certainty within a relatively short treatment window. However, the probability does not reach full certainty unless process conditions exceed industry standard processing parameters of 72 °C for 20-30 seconds (Koutsoumanis et al., 2020b). In the model for high acidic fruit juices, the variability in $\log_{10} D$ predictions results in broader variability bands, especially at higher reduction thresholds. The increasing width of the variability bands at higher log reduction levels reflect the accumulation of variability over time. In the stochastic models,

the predicted $\log_{10} D$ vary due to residual variability in the collected metadata. At longer treatment times, the absolute differences in reduction caused by small variations in $\log_{10} D$ become more pronounced due to the non-linear relationship between the treatment time and the denominator ($10^{\log_{10} D}$), meaning that a small change in $\log_{10} D$ leads to exponential changes in the D-value, directly affecting the slope of the reduction curve. As stated earlier in this section, it is of high importance to consider the whole range of reduction probability, as most risk estimates originate from extreme scenarios where insufficient reduction was present.

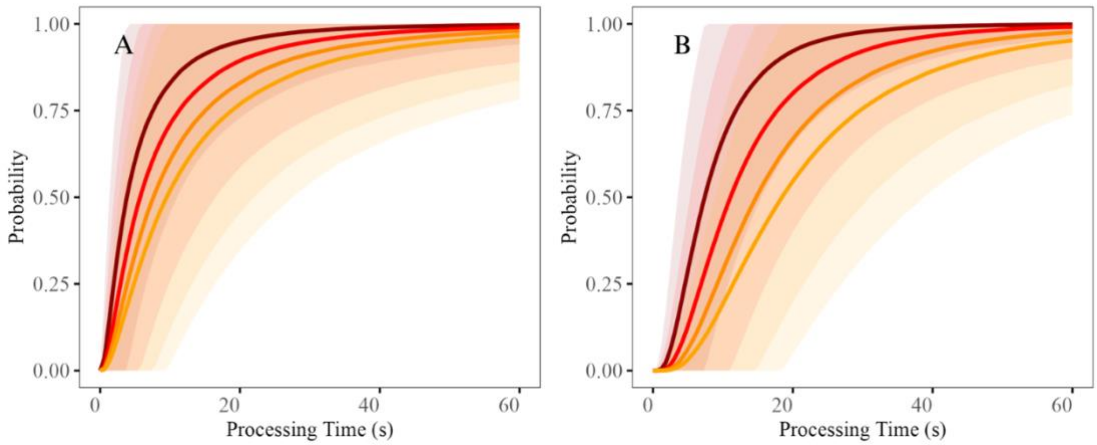


Figure 6, Probability functions showing the probability to achieve a certain log reduction based on the process time and temperature. Data shown for all liquids (A), and high acidic fruit juices (B). Probabilities calculated based on the Stochastic $\log_{10} D$ distributions. Shades show the variability range for the probability function. Colours indicate 2-log, 3-log, 4-log, 5-log reductions.

3.2 High Pressure Processing

3.2.1 Modelling kinetic parameters: $log_{10} D$ estimation

As with thermal processing, HPP was evaluated across two levels of specificity: all liquid products, and high acidic fruit juices. This section addresses how variability in $\log_{10} D$ values was influenced by pressure levels across these specificity levels and compare the three modelling approaches of deterministic, stochastic and Bayesian probabilistic regression models. Table 5 summarises the regression coefficients for the deterministic models. Similar as in the approach for thermal processing was the variance defined as the average squared distance between the frequentist regression line and the actual datapoints (σ^2), explaining how much unexplained variability in $\log_{10} D$ prediction remains after accounting for the effect of pressure in Megapascal (MPa). For both the overall variance as the residual standard error, a smaller value (0.175, 0.424) was found in the model for high acidic fruit juices. The relation between pressure inputs and $\log_{10} D$ values is displayed in Figure 7, where the deterministic model (solid black line) provides the baseline estimate of the function, while the Bayesian regressions (thin purple lines) incorporate full variability and parameter uncertainty. These parameters were based on the posterior distributions for the model slope, intercept and overall variance (σ^2), as displayed in appendix 2. The negative slopes across these models indicate that increasing pressure reduces the $\log_{10} D$ estimation as expected. However, the rate of this reduction varies extremely across the specificity levels, with the steepest decline observed for the model using data for high acidic fruit juices, visible in Figure 7B. This was supported by the z-values, where the all-liquid products approach has an extremely high z-value of 1591.7 MPa, indicting that a great increase in pressure is required to achieve 1 log reduction in $\log_{10} D$. This unrealistically high z-value suggests a negligible pressure effect on the inactivation of STEC within this diverse dataset. This is likely due to the heterogeneity of the food matrices included, which include a wide range of physicochemical properties such as pH, and molecular composition. These factors can influence the microbial resistance to pressure significantly (Syed et al., 2016). In contrast, the model for high acidic fruit juice shows a more pronounced and significant pressure effect, as displayed by a steeper slope and thereby lower z-value (269.7 MPa). This indicates that within this more homogenous subgroup, pressure has a more consistent and predictable impact on STEC inactivation. This improved model is not necessarily a function of reduced variability in the metadata only, but also reflects the specific characteristics of acidic fruit juice that enhance pressure sensitivity. The underlying cause of this is primarily food-matrix dependent, as the efficacy of high pressure processing is known to vary with the composition of food products (Sehrawat et al., 2020). Acidic environments can enhance the susceptibility of microorganisms to pressure-induced inactivation. In a comparative study done by Gouvea et al. in 2020, it was confirmed that strains were more sensitive to HPP treatment when they were present in high acidic fruit juices (Gouvea et al., 2020). These findings have several implications for process design and predictive modelling, where the limitations of univariate modelling in systems where microbial inactivation is influenced by multiple interacting factors was pointed out. While this study was focussed on pressure as the main predictor, the influence of matrix properties or additional control

Table 5, Regression coefficients and parameters for the deterministic models by high pressure processing.

Model & Specificity	Intercept	Slope	Datapoints	RSE	Variance	z-value (MPa)
High pressure processing						
All liquid products	1.073	-0.00063	326	0.539	0.289	1591.7
High acidic fruit juices	1.538	-0.00371	41	0.424	0.175	269.7

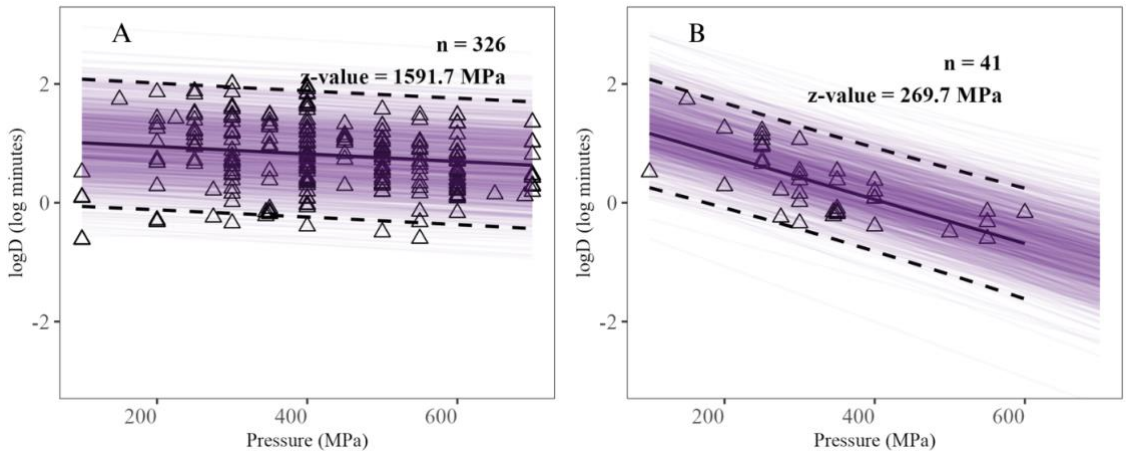


Figure 7, Meta-regression models for HPP for the two levels of specificity: All liquid products (A), and high acidic fruit juices (B). Solid lines represent the deterministic estimate and dashed lines its 95% prediction intervals. Thin purple lines represent the Bayesian regression estimates.

steps, such as temperature effects, may modify the pressure sensitivity of microorganisms. Although the model for high acidic fruit juices shows higher accuracy, it might not be generalisable. A model with higher predictive power within a narrow matrix category may not perform well when applied to a broader product group or unknown formulations.

3.2.2 The overall impact on the estimated risk: a deterministic approach

To establish a baseline risk estimation under high pressure processing conditions, the deterministic approach was applied, providing fixed parameter estimates without incorporating variability or uncertainty. The overall microbial reduction was again based on the mean survivor count across the Monte Carlo simulations, following the probability integral method. This allowed for the calculation of the resulting STEC doses after HPP treatment, which directly influenced the estimated number of cases in the QMRA framework. Table 6A presents the mean $\log_{10}$ reduction values, percentiles for overall reduction, and the overall reduction across the two levels of specificity. The results proved that modelling specific data in food matrix composition led to greater reduction. The overall reductions for all liquid products and high acidic fruit juices, were found to be 1.5 log CFU/mL, and 25.4 log CFU/mL respectively. These results align directly with the previously obtained z-values and regression coefficients, where high acidic fruit juices required lower pressure to achieve equivalent microbial inactivation. As expected, using the mean values for reduction (on iteration level) creates comparable results as the overall reduction, calculated based on the arithmetic mean of survivors. The small difference visible is due to the use of variable process and demographic model inputs. The implications of these reduction levels

Table 6a, Microbial reduction based on the arithmetic mean of survivors (HPP). Results were based on the mean of 5 repeated simulations with each 1,000,000 iterations and displayed for the 2.5, 50, 97.5, and 99th percentile.

Model & Specificity	Mean reduction	Overall reduction (R)	R(2.5%)	R(50%)	R(97.5%)	R(99%)
Level 1: All Liquid Products						
Linear model (Deterministic)	1.5 ± 0.0	1.5 ± 0.0	1.5 ± 0.0	1.5 ± 0.0	1.5 ± 0.0	1.5 ± 0.0
Linear model (Stochastic)	3.2 ± 6.0	0.8 ± 0.0	0.1 ± 0.0	1.5 ± 0.0	17.1 ± 0.1	26.9 ± 0.1
Bayesian model (Probabilistic)	3.2 ± 6.5	0.8 ± 0.0	0.1 ± 0.0	1.4 ± 0.0	17.7 ± 0.0	28.6 ± 0.2
Level 2: High Acidic Fruit Juice						
Linear model (Deterministic)	25.4 ± 0.4	25.2 ± 0.0	24.6 ± 0.0	25.4 ± 0.0	26.1 ± 0.0	26.3 ± 0.0
Linear model (Stochastic)	38.7 ± 40.8	3.6 ± 0.0	3.7 ± 0.0	25.4 ± 0.0	171.6 ± 0.6	245.4 ± 0.9
Bayesian model (Probabilistic)	37.3 ± 49.0	2.3 ± 0.0	1.3 ± 0.0	19.5 ± 0.0	282.0 ± 0.9	Inf ± 0.0

Table 6b, QMRA results for the HPP models. Results were based on 5 repeated simulations with each 1,000,000 iterations. Mean serving risk defined as the probability of illness corrected for the prevalence. The sum of Bernoulli trials present positive illness cases before prevalence per million iterations.

Model & Specificity	Mean serving risk	Sum bernoulli trials	Cases / million servings	Annual cases EU (bernoulli)	Annual cases EU (mean risk)
Level 1: All Liquid Products					
Linear model (Deterministic)	$8.8 \cdot 10^{-7} \pm 0.0$	689.2 ± 33.5	0.9 ± 0.0	35266 ± 1716	34503 ± 0
Linear model (Stochastic)	$4.0 \cdot 10^{-6} \pm 2.6 \cdot 10^{-9}$	3153.8 ± 40.0	4.1 ± 0.1	161378 ± 2045	159332 ± 101
Bayesian model (Probabilistic)	$4.4 \cdot 10^{-6} \pm 4.5 \cdot 10^{-9}$	3394.4 ± 46.9	4.4 ± 0.1	173689 ± 2398	173506 ± 178
Level 2: High Acidic Fruit Juice					
Linear model (Deterministic)	$2.9 \cdot 10^{-19} \pm 0.0$	0.0 ± 0.0	0.000 ± 0.0	0 ± 0	0 ± 0
Linear model (Stochastic)	$6.8 \cdot 10^{-9} \pm 5.3 \cdot 10^{-11}$	2.6 ± 0.5	0.003 ± 0.0	133 ± 28	266 ± 2
Bayesian model (Probabilistic)	$1.4 \cdot 10^{-7} \pm 2.3 \cdot 10^{-9}$	105.6 ± 12.4	0.137 ± 0.0	5403 ± 636	5694 ± 89

are clear in the resulting risk estimates (Table 6B). For all liquid products, the limited inactivation leads to a mean serving risk of $8.8 \bullet 10^{-7}$ and approximately 0.9 STEC illness cases per million servings. When scaled to the European fruit juice consumption, this results in an estimated 35266 cases. In contrast, the high acidic fruit juice model results in no predicted illness cases and a per-serving-risk of $2.0 \bullet 10^{-19}$. This difference supports the importance of food matrix composition on modelling HPP efficacy, as highlighted before.

A sensitivity analysis was conducted to assess the influence of (variability in) key input parameters on the probability of illness. The input distributions for pressure, treatment time, serving size and initial STEC concentration ($\log_{10} N_0$), were variable based on realistic assumptions as defined in section 2.3, while $\log_{10} D$ was treated as fixed deterministic estimate derived from the regression models. The resulting Spearman's correlation coefficients (Figure 8) show the strength and direction of the relationship between each input and the probability of illness. The dark grey bars represent the deterministic approach, only present in panel A. In the model for all liquid products (Figure 8A), a clear pattern is observed, where the variability in the initial contamination and serving size show the strongest positive correlation on the probability of illness. This reflects the native relationship where higher contamination levels and larger consumption volumes increased exposure and thereby risk. Treatment time and Pressure had only marginal influence on the illness probability. For the high acidic fruit juice model (Figure 8B), the picture is slightly different. Due to the extremely low overall illness probability observed in this model, none of the input parameters had a significant effect on the outcome. The correlation values remain close to zero for all parameters, indicating that nearly all simulations resulted in sufficient inactivation. From a risk assessments perspective, the deterministic model suggest that

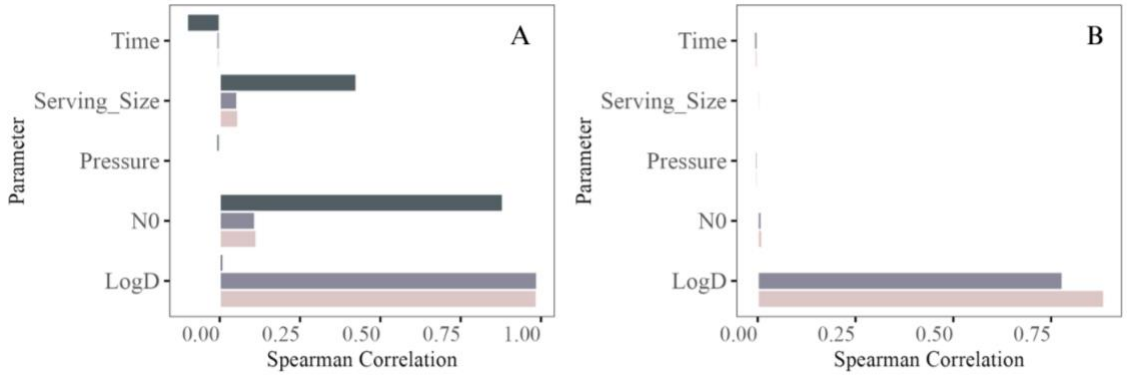


Figure 8, Spearman's correlation coefficients for all liquid products (A), and High acidic fruit juices (B). Correlations present the impact of variability in input parameters on the probability of illness. Analysis performed on deterministic models ,Stochastic models , and probabilistic models . Results displayed for HPP.

HPP provides effective control over STEC contamination under evaluated conditions. These findings establish a clear foundation for understanding the effectiveness of a deterministic baseline. However, to fully assess the risk associated with STEC inactivation by HPP, it is essential to account for variability and uncertainty in inactivation kinetics and process conditions. It has been found that the food product composition, including fats, proteins, sugars, salts and acids might affect the HPP resistance of different strains significantly (Georget et al., 2015). In addition, it is evident that some strains have larger general genetic resistance to HPP (Hauben et al., 1997). The next section will explore the impact of incorporating stochastic and probabilistic models into risk estimation.

3.2.3 The impact of variability on the estimated risk: a stochastic and probabilistic approach

Building further on the deterministic findings, stochastic and Bayesian probabilistic models were applied to evaluate how the incorporation of variability and uncertainty influences risk estimations under HPP conditions. This allowed for variation in inactivation kinetics across Monte Carlo iterations, reflecting the natural spread observed in the underlying data. In the Bayesian framework, uncertainty was incorporated by drawing the slope, intercept and residual variance from their posterior distributions, adding an additional layer of uncertainty to the inactivation predictions. These models did not aim to predict single point outcomes but provided a more robust foundation for risk-based decision making. Considerable variability has been reported in resistance to HPP amongst different strains of *E. coli* O157:H7, varying from 0.6 to 9.7 log reductions (Hauben et al., 1997; Malone et al., z.d.; Patterson et al., 1995). While commercial HPP treatments (e.g. 600 MPa for 3 minutes) are often effective in reducing pathogen levels, does the observed variability in inactivation responses of STEC across different datasets underly the need for probabilistic evaluation (Gouvea et al., 2020). This approach is consistent with regulatory guidance, such as that of the FDA which emphasizes the importance of validating microbial inactivation efficacy for each product under specific conditions (FDA, 2017).

Similar to thermal processing, this study developed modes for both all liquid products and high acidic fruit juices, and their outputs are summarised in Table 6A. As expected, the inclusion of variability in $\log_{10} D$ for a given pressure input, and in the Bayesian case, additional parameter uncertainty, led to substantial changes in the predicted reductions when compared to the deterministic baseline. A key observation is that both the stochastic and Bayesian approaches predict lower overall reductions, which directly influence the probability of illness in the QMRA framework. In the generic all liquid products model, the predicted overall reduction decreased from 1.5 log CFU/mL under the deterministic assumption to 0.8 log CFU/mL for both the stochastic and the Bayesian simulations. The effect was more pronounced for the model for high acidic fruit juice, where the deterministic reduction of 25.4 log CFU/mL dropped to 3.6 log CFU/mL in the stochastic model and 2.5 log CFU/mL for the Bayesian model. These findings emphasize the importance of considering real-world variability in STEC inactivation simulations. Such variability is clearly present in the underlying metadata. In a study done by Jordan et al. in 2001, where the inactivation of *E. coli* O157:H7 was assessed in high acidic fruit juices, a 5 log CFU/mL reduction was found at a pressure input of 550 MPa for 5 minutes (T = 20°C) (Jordan et al., 2001). However, in similar settings this study found log reduction values ranging from 1 – 6 log CFU/mL in different high acidic food products (Jordan et al., 2001). Another study based on a similar purpose showed a reduction of 7.4 log CFU/mL at processing conditions of 500 MPa for 1 minute (T = 5°C) (Petrus et al., 2020). Another important point is that mean reductions, as often reported in literature, can be misleading when used directly in risk models. Table 6A demonstrates this clearly. In the stochastic and Bayesian simulations for all liquid products, the

mean log reduction was found to be 3.2 ± 6.5 log CFU/mL. In high acidic fruit juices, the mean reduction rose to 38.7 ± 40.8 (stochastic) and 38.3 ± 49.0 log CFU/mL (Bayesian). These extremely high standard deviations are a clear indication of a heavily skewed reduction distribution and highlight the statistical misrepresentation that can occur when reductions are not calculated through proper survivor-based approaches. In contrast, the overall reduction, derived from the arithmetic mean of survivors counts over the same 1 million Monte Carlo simulations, reflects a much more realistic process outcome, where reductions of 0.8 log CFU/mL for both models in all liquid products were found. In high acidic fruit juices, the overall reduction decreased to 3.6 and 2.5 log CFU/mL for the stochastic and Bayesian models respectively. This difference highlights the risk of overestimating the efficacy of the processing technology when relying solely on mean reductions, especially when reduction distributions are highly variable or long tailed. It is also important to consider that this analysis focusses on one of the most favourable matrix-pathogen combinations for microbial inactivation. Both the pathogen and the matrix promote sensitivity to pressure and yet even in this optimal case, predicted reductions under variable conditions shows a high degree of variability. If this same modelling framework were applied to more pressure resistant organisms, such as *Listeria monocytogenes,* in more protective matrices such as milk or protein-rich foods, significantly more processing would be required to achieve similar levels of inactivation. Under such conditions, the inclusion of broad variability and parameter uncertainty as in the Bayesian models, might lead to conservative estimates that overstate the actual residual risk.

The lower overall reductions observed under these models translate directly into higher surviving STEC concentrations, thereby increasing the estimated probability of illness per serving and consequently, the number of illness cases at European level. These outcomes are summarised in Table 6B, which presents the mean risk per serving, the number of illness cases per million servings and the number of cases on European scale. In the models for all liquid products, the number of illness cases increased significantly across the modelling approaches, from approximately 35,000 cases under the deterministic approach to 162,000 under the stochastic models and 174,000 under the Bayesian model. The high acidic fruit juice model shows similar results, only at a lower magnitude. While the deterministic model predicts no illness cases, the inclusion of variability in the stochastic model yields an estimated 266 cases annually, and the Bayesian model increases this to nearly 6000 cases. All results are based on the arithmetic means of the serving level risk outputs from 5 independent Monte Carlo simulations with 1 million iterations. These findings demonstrate that even when deterministic models predict no risk, the incorporation of variability can uncover low probability but plausible illness outcomes, underscoring the importance of considering full outcome distributions. Further insight into the drivers of the variability in illness probability under the stochastic and Bayesian approach is given by the Spearman's rank coefficients shown in Figure 8. Across both specificity levels, the variability in $\log_{10} D$ estimates emerges as the dominant factor influencing risk, exceeding the influence of the variability in processing and demographic inputs. In deterministic models, where the $\log_{10} D$ was fixed, input variability in initial concentration and serving size had greater impact on these outcomes. However, when the true variability is introduced, these factors become relatively less influential. In the high acidic fruit juice models, where deterministic modelling predicted no illness cases, the spearman's correlations revealed a clear association between illness probability and $\log_{10} D$ related variability under both stochastic and Bayesian conditions. This shows that even in matrices considered highly effective for pathogen inactivation, variability in kinetic parameters can produce meaningful illness estimations. To further examine the underlying dynamics of the Monte Carlo simulations, probability density functions were generated for key

outputs and parameters on iteration level. The risk per serving (Figure 9A-B), log reductions (Figure C-D), and $\log_{10} D$ estimates at pressure P (Figure 9E-F). These plots separate between the stochastic (grey) and Bayesian (beige) models. Overlaid barcodes indicate the iterations in which Bernoulli trials resulted in illness, aligned to the corresponding values on the x-axis. It is important to note that these cases reflect conditional illness probabilities, assuming all servings are contaminated and do not yet include prevalence adjustment. In the model for all liquid products (Figure 9A), the distribution of the risk per serving is skewed towards higher risk levels,

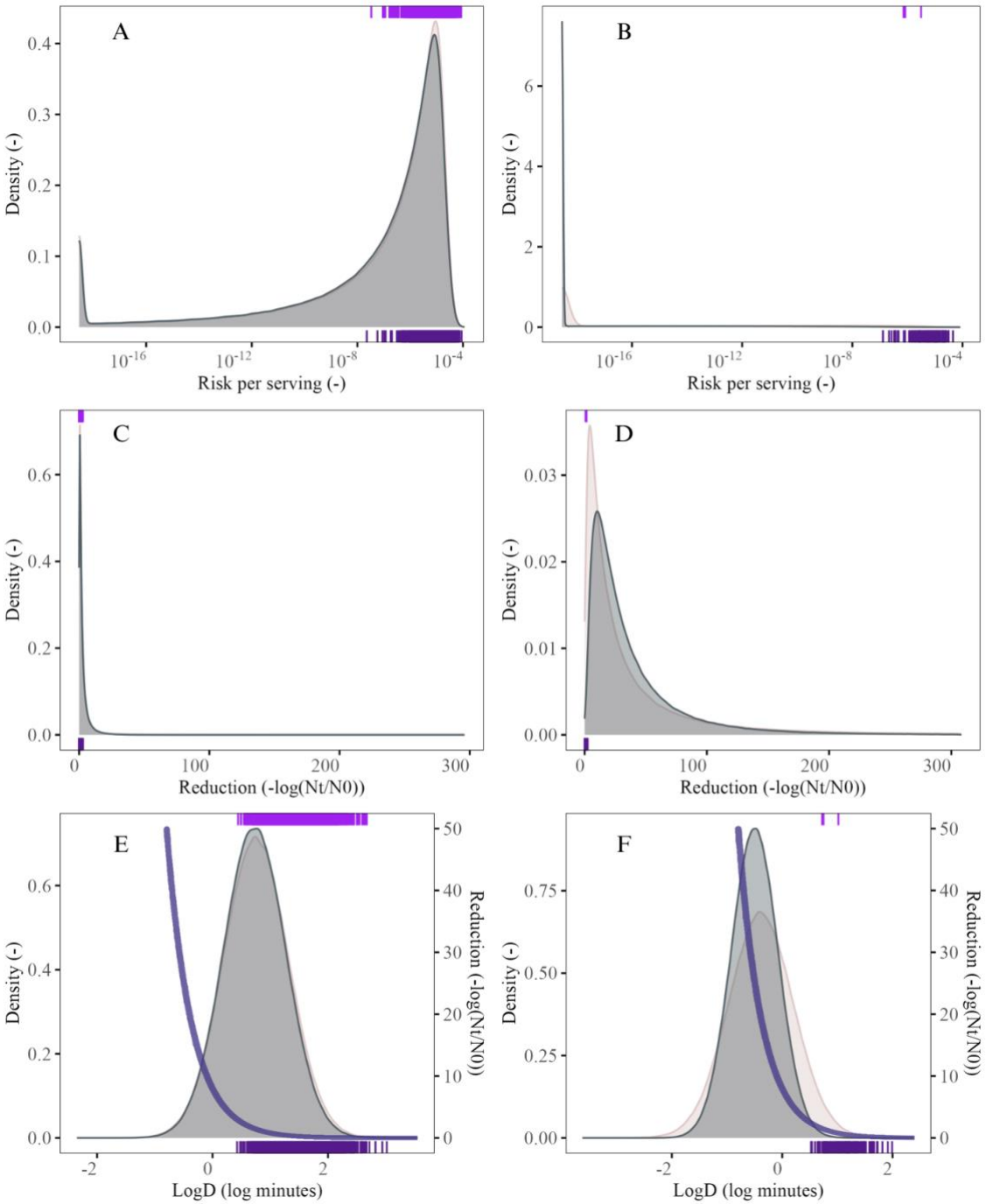


Figure 9, Probability density distributions for the Risk of illness per serving using data for all liquid products (A), and high acidic fruit juices (B). Probability density distributions for overall reduction using all liquids (C) and high acidic fruit juices (D). Probability density distributions for $\log_{10} D$ using all liquids (E) and high acidic fruit juices (F). Layers present the modelling approach (■ = stochastic, ■ = probabilistic). Barcode plots show the occurrence of illness cases with respect to the x-variable (top = stochastic, bottom = probabilistic). Purple lines represent the reduction of STEC with respect to the $\log_{10} D$ distribution. Results displayed for HPP.

with a noticeable peak where the majority of illness cases occur. Both the stochastic and Bayesian models show similar distributions in this scenario, and the positive Bernoulli outcomes cluster together in the higher end of the range. This suggests that illness risks are largely driven by conditions where reduction are lower and doses remain relatively high. In contrast, the high acidic fruit juice models (Figure 9B), show a much sharper density concentrated at low risks. For all liquid products, the first illness cases were detected at a risk of $1.5 \bullet 10^{-8}$ (stochastic) , and $4.8 \bullet 10^{-8}$ (probabilistic). For all high acidic fruit juices, the first illness cases were detected at a risk of $9.7 \bullet 10^{-8}$ (stochastic) , and $3.2 \bullet 10^{-6}$ (probabilistic). No major peak is visible at higher risk values, but a small number of positive illness cases occurs at the upper end of the distribution, showing that even rare lower efficacies can influence the overall illness case estimates when scaled to a European level exposure.

The variability in risk estimates originates from the variability in log reductions, simulated across all iterations (Figure 9C-D). For all liquid products (Figure 9C), both models produce a small distribution of reductions, where most reduction are reaching poor efficacy. The Bayesian model is slightly more dispersed due to the additional parameter uncertainty, but both show that the majority of reductions fall in the lower range. In high acidic fruit juices (Figure 9D), the distributions are more skewed and show greater separation between models. The stochastic model produces a broader distribution, while the Bayesian approach introduces more uncertainty, resulting in a higher proportion of simulations with insufficient reduction. This explains the elevated number of illness cases observed under the Bayesian model for this product group. For all liquid products, the first illness case occurred at a log reduction of 3.19 log CFU/mL for the stochastic model, and 2.69 log CFU/mL for the Bayesian model. In contrast, for high acidic fruit juices illness cases were already observed at much lower reductions, at 1.12 and 2.91 log CFU/mL respectively. To understand where these insufficient reductions originate from, the true variability in $\log_{10} D$ for pressure input ~N(550, 0.1) was visualised in Figure 9E-F. These distributions present the spread in $\log_{10} D$ across the iterations, with overlaid purple lines indicating the log reductions for a given $\log_{10} D$. For all liquid products (Figure 9E), the illness cases appear within the lower range, below the distribution peak where the inactivation is the slowest and is representative for the most process-resistant cells. The minimum safe $\log_{10} D$ at which an illness case was first observed were 0.40 and 0.47 $\log_{10} minutes$. for the stochastic Bayesian models respectively. In high acidic fruit juices, illness cases occurred at higher $\log_{10} D$ values of 0.86 (stochastic) and 0.44 (Bayesian). This substantiates the fact that stochastic model shows fewer low values for $\log_{10} D$, resulting in fewer illness cases while the Bayesian model, due to its parameter uncertainty, generates more lower values for $\log_{10} D$. A critical finding is that illness case occur from $\log_{10} D$ values that are present within the tails of the distribution where probability remains non-negligible. This supports the idea that the risk is not only driven by average behaviour but by the upper tails of the dose and survival distributions, regions that deterministic models do not consider. This becomes clear when comparing the two modelling approaches. Although the Bayesian model introduced only slightly more spread, this results in a notable increase in predicted illness cases. This difference is driven by a small subset of iterations where the combination of a high $\log_{10} D$ and lower pressure effectiveness leads to more survival and illness cases.

To translate the findings into an industrial relevant context, the probability of achieving different levels of STEC inactivation (2-5 log reductions) were assessed over a range of processing times, based on the stochastic model outputs. Figure 10 visualizes this for both all liquid products as high acidic fruit juices. In the model for all liquid products (Figure 10A), the probability curves increase gradually with treatment time and remain well below certainty even at extended

treatment times. The probability of achieving a 5-log reduction remains below 50% for most of the simulated time range, indicating both the limited predictability and broad variability of inactivation outcomes when using generalised metadata. In contrast, the model based on high acidic fruit juices (Figure 10B), demonstrates a notable different pattern. All probability curves rise sharply and reach near certainty within a few minutes of processing, even at higher reduction levels. The 5-log reduction target, commonly recommended by regulatory agencies such as the FDA (2024), is achieved with over 90% certainty in less than 5 minutes. These results reflect the strong pressure sensitivity of STEC in high acidic environments and the more precise prediction capability of models based on product-specific data. Together, these findings emphasize the importance of tailoring HPP treatments to the characteristics of specific food matrices and suggest that a one-size-fits-all approach are not suited for reliably predicting microbial safety standards across diverse liquid products.

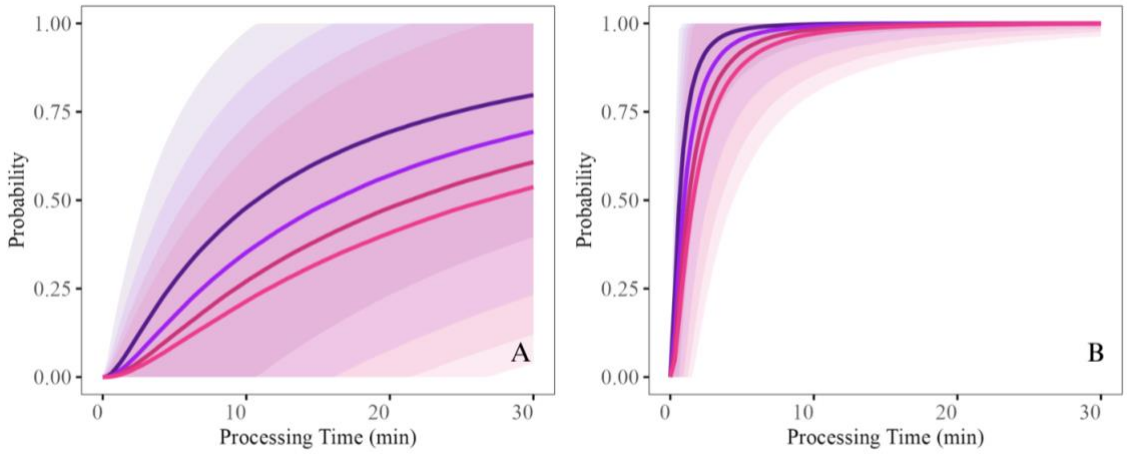


Figure 10, Probability functions showing the probability to achieve a certain log reduction based on the process time and temperature. Data shown for all liquids (A), and high acidic fruit juices (B). Curves present the probability of achieving a certain log reduction, while shaded areas reflect the variability derived from underlying log_{10} D distributions and process metadata. Colours indicate 2-log, 3-log, 4-log, 5-log reductions.

3.3 Pulsed Electric fields processing

3.3.1 Modelling kinetic parameters: log Reduction estimation

Unlike thermal and high pressure processing, microbial inactivation under pulsed electric fields (PEF) is primarily driven by the energy input rather than the treatment time, which typically occurs on the millisecond to microsecond scale (Walkling-Ribeiro et al., 2011). Therefore, the modelling approach in this section focusses on creating a relationship between the reduction of STEC and the total specific energy input (kJ/L). Table 7 summarises the regression coefficients for the deterministic models at the two specificity levels 1) all liquid products, and 2) high acidic fruit juices, providing an estimate for slope, residual standard error, the overall variance and the *D*-values. To ensure more realistic behaviour at the origin, both regression models were forced through zero, thereby removing the intercept. This approach corrects for the physically impossible prediction of microbial reduction at zero energy input, which is sometimes encountered when the intercept is freely estimated. However, this model specification comes with some trade-offs, especially in the Bayesian setting as discussed below. Notably, the slopes of the regression lines are highly consistent across the specificity levels. This close agreement suggests that matrix specificity has a limited impact on PEF inactivation kinetics, at least when modelled via energy input alone. This finding is further supported by the comparable *D*-values of 31.16 and 31.45 kJ/L for the general and specific model respectively. Unlike thermal and high-pressure processing, where acidity has a pronounces effect on inactivation, PEF appears less sensitive to this matrix acidity. Instead, other physicochemical parameters, such as electrical conductivity, charge distribution and pulse waveform can play a more dominant role in determining microbial susceptibility to PEF (Roobab et al., 2022). Interestingly, the model variance remains high and nearly unchanged between the all liquid products model (1.187), and the high acidic fruit juice model (1.086), despite the decrease in datapoints. This is in contrast with the TP and HPP models in this study, where specificity reduced variance and improved model performance. Here, the comparable variance suggests that even in refined datasets, microbial inactivation by PEF remains highly variable, particularly at higher energy inputs where data inputs were less frequent. This indicates a greater uncertainty in slope estimation due to limited observations in that range. The corresponding regression plots (Figure 11) illustrate these trends. Both the models show a relatively consistent trend between energy input and log reduction, with Bayesian regression estimates (blue thin lines) overlaid in the deterministic estimate (solid black line). As explained, the spread in Bayesian regression lines is comparable in the model for high acidic fruit juice. A key limitation in the implementation of both the stochastic and Bayesian models for PEF lies in how variability was incorporated int the simulation framework. In contrast to the $\log_{10} D$-based modelling used for TP and HPP, the PEF models treat the predicted log reduction at a given energy input as the central value of a distribution. Specifically, reduction were sampled from a normal distribution centred on the regression prediction, using the residual standard error (stochastic) or the overall variance (Bayesian) as standard deviation to define the spread. While this mirrors the approach used for prediction intervals in the other technologies, it introduced a certain issue in the context of PEF. At lower energy inputs, the lower tail of the normal distribution extend below zero, resulting in biologically impossible negative reductions. To address this, both

Table 7, Regression coefficients and parameters for the deterministic models by High Pressure Processing.

Model & Specificity	Intercept	Slope	Datapoints	RSE	Variance	D-value (kJ/L)
Pulsed electric fields processing						
All liquid products	0.000	0.03209	153	1.114	1.187	31.16
High acidic fruit juices	0.000	0.03178	34	1.053	1.086	31.46

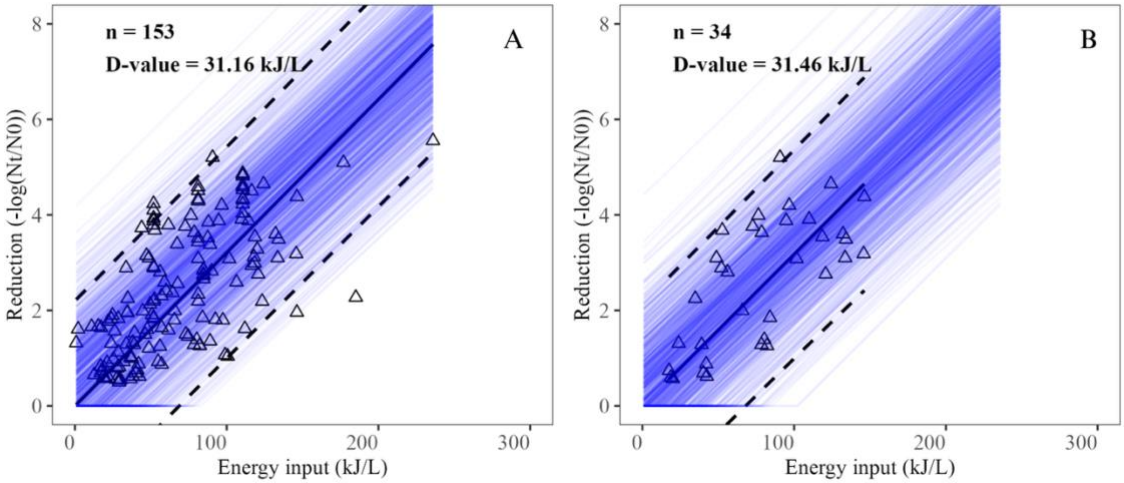


Figure 11, Meta-regression models for PEF for the two levels of specificity: All liquid products (A), and high acidic fruit juices (B). Solid lines represent the deterministic estimate and dashed lines its 95% prediction intervals. Thin blue lines represent the Bayesian regression estimates.

models were truncated at zero, ensuring that all sampled reduction values remained non-negative. Although this approach slightly alters the distribution shape, its impact on the final outcome is limited since most of the interest is focussed on higher energy inputs.

3.3.2 The overall impact on the estimated risk: a deterministic approach

To establish a baseline risk estimation for PEF processing, a deterministic modelling approach was applied using fixed inputs for predicted reduction. Although microbial inactivation was modelled as function of the energy input directly, the overall reduction per serving was derived using the probability integral method, based on the mean survivor count across the iterations. This allowed for consequent calculation of remaining STEC doses and the estimation of illness probabilities at European level. The deterministic outputs are summarised in Table 8A-B. The overall reductions were found to be consistent across the specificity levels, with 3.9 log CFU/mL for all liquid products and 3.8 log CFU/mL for high acidic fruit juices. This reflects the minimal observed differences in the regression coefficients and D-values for the models. Moreover, both models followed a relatively fixed energy input of 120 ± 0.1 kJ/L, which reflects industrially relevant and tightly controlled processing settings (Table 2). Under these assumptions, the process showed sufficient inactivation and reductions. The deterministic QMRA outputs (Table 8B) show that these reductions translate into relatively low estimates of illness risk. When scaled to the European fruit juice consumption, the number of the predicted annual cases based on the Bernoulli framework is approximately 112 for all liquid products and 194 for high acidic fruit juices. These results appear to be relative to the baseline incidence of STEC illness cases in the

Table 8a, Microbial reduction based on the arithmetic mean of survivors (PEF). Results were based on the mean of 5 repeated simulations with each 1,000,000 iterations and displayed for the 2.5, 50, 97.5, and 99th percentile.

Model & Specificity	Mean reduction	Overall reduction (R)	R(2.5%)	R(50%)	R(97.5%)	R(99%)
Level 1: All Liquid Products						
Linear model (Deterministic)	3.9 ± 0.0	3.9 ± 0.0	3.8 ± 0.0	3.9 ± 0.0	3.9 ± 0.0	3.9 ± 0.0
Linear model (Stochastic)	3.9 ± 1.1	2.5 ± 0.0	1.7 ± 0.0	3.9 ± 0.0	6.0 ± 0.0	6.4 ± 0.0
Bayesian model (Probabilistic)	3.6 ± 1.1	2.3 ± 0.0	1.5 ± 0.0	3.6 ± 0.0	5.7 ± 0.0	6.1 ± 0.0
Level 2: High Acidic Fruit Juice						
Linear model (Deterministic)	3.8 ± 0.0	3.8 ± 0.0	3.8 ± 0.0	3.8 ± 0.0	3.8 ± 0.0	3.8 ± 0.0
Linear model (Stochastic)	3.8 ± 1.1	2.6 ± 0.0	1.7 ± 0.0	3.8 ± 0.0	5.9 ± 0.0	6.3 ± 0.0
Bayesian model (Probabilistic)	3.7 ± 1.1	2.3 ± 0.0	1.5 ± 0.0	3.7 ± 0.0	5.8 ± 0.0	6.2 ± 0.0

Table 8b, QMRA results for the PEF processing models. Results were based on 5 repeated simulations with each 1,000,000 iterations. Mean serving risk defined as the probability of illness corrected for the prevalence. The sum of Bernoulli trials present positive illness cases before prevalence per million iterations.

Model & Specificity	Mean serving risk	Sum bernoulli trials	Cases / million servings	Annual cases EU (bernoulli)	Annual cases EU (mean risk)
Level 1: All Liquid Products					
Linear model (Deterministic)	$3.9\square10^{-9} \pm 0.0$	2.2 ± 1.1	0.0 ± 0.0	112.4 ± 56.3	154.0 ± 0.0
Linear model (Stochastic)	$9.3\square10^{-8} \pm 1.4\square10^{-9}$	70.6 ± 5.9	0.1 ± 0.0	3612.8 ± 299.5	3658.4 ± 54.0
Bayesian model (Probabilistic)	$1.4\square10^{-7} \pm 1.2\square10^{-9}$	100.2 ± 14.2	0.1 ± 0.0	5127.2 ± 724.9	5445.4 ± 48.3
Level 2: High Acidic Fruit Juice					
Linear model (Deterministic)	$4.3\square10^{-9} \pm 0.0$	3.8 ± 3.0	0.0 ± 0.0	194.4 ± 155.2	167.0 ± 0.0
Linear model (Stochastic)	$7.5\square10^{-8} \pm 8.5\square10^{-10}$	60.6 ± 8.3	0.1 ± 0.0	3100.8 ± 423.1	2946.2 ± 33.6
Bayesian model (Probabilistic)	$1.4\square10^{-7} \pm 1.2\square10^{-9}$	102.0 ± 7.0	0.1 ± 0.0	5219.4 ± 360.0	5543.4 ± 48.8

EU, where 321 confirmed annual cases were found between 2019 and 2023 (EFSA, 2019). To better understand the drivers of illness risk under these fixed conditions, a Spearman's rank correlation was performed for the deterministic models (Figure 12, dark grey bars). The primary finding was that the variability in initial contamination and the serving size had the strongest positive correlation with the variability in the probability of illness across both specificity levels. These results align with the percentiles of the overall reduction, where despite the fixed model parameters, variability in initial contamination and serving size had a large effect on the final reduction. Unlike the models for TP and HPP, no log-linear secondary model was used here since the reduction could be predicted directly, making the influence of variability in these parameters more significant. In contrast, the energy input shows minimal correlation with illness probability. However, one should note that this reflects the modelling assumption that the energy input is tightly controlled around 120 kJ/L and thereby introduces almost no variability into the simulations.

3.3.3 The impact of variability on the estimated risk: a stochastic and probabilistic approach

Building on the deterministic models, stochastic and Bayesian probabilistic approaches were applied to incorporate variability and uncertainty into the prediction of STEC reduction. The distinction in model structure between PEF and TP and HPP strongly influences how variability propagates through the risk assessment framework. The resulting log reduction and risk outcomes are presented in Tables 8A and 8B. The most notable observation is that, although the mean reductions remain around 3.6-3.9 log CFU/mL across all models and specificities, the overall reduction drops to approximately 2.3-2.6 log CFU/mL. This difference between the mean and

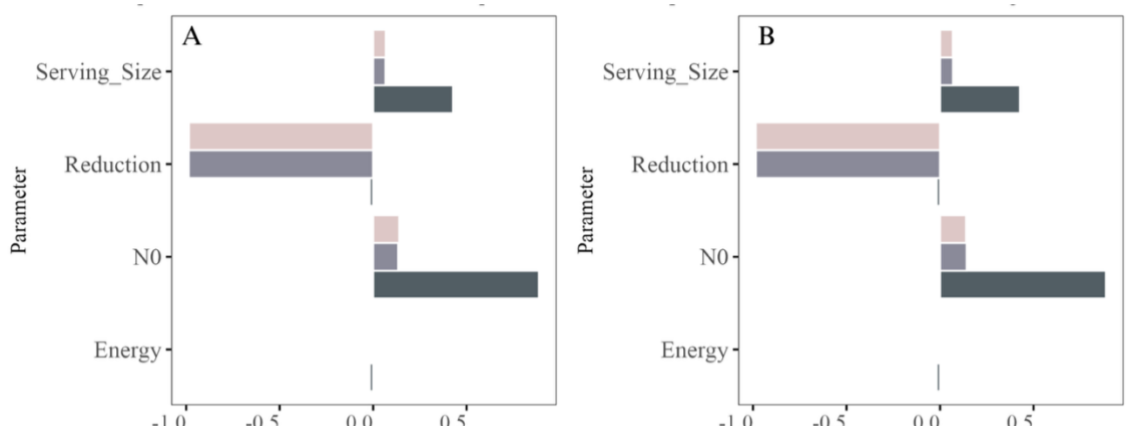


Figure 12, Spearman's correlation coefficients (PEF) for all liquid products (A), and High acidic fruit juices (B). Correlations present the impact of variability in input parameters on the probability of illness. Analysis performed on deterministic models ,Stochastic models , and probabilistic models .

overall reduction again highlights the difference in reduction distributions. The effect of variability is the clearest in the QMRA estimates. Compared to the deterministic models, which predict 112 and 194 annual illness cases for all liquid product and high acidic fruit juices respectively, the stochastic models predict 3612 and 3100 cases. These increases point out that even a small shift in the lower tail of the reduction distribution can result in significantly more estimated illness events once scaled to European level. For all liquid products, the stochastic model predicts more cases based on all liquid products than in high acidic fruit juice, while the opposite is true for the Bayesian approach. This does not necessarily reflect a failure in the modelling system but rather originates from slight differences in Bernoulli trial outcomes. Since illness cases are rare events driven by low-reduction iterations, small shifts in the distribution can have oversized effects on the total risk estimate. These observations are visualised in Figure 13, where probability density distributions are shown for both the risk per serving (Figure 13A-B) and reduction (Figure 13 C-D). For both specificity levels, the stochastic (grey) and Bayesian (beige) curves closely overlap, indicating that the central trends of the models are consistent. However, the Bayesian models show slightly wider distributions, especially for high acidic fruit juice, reflecting the added uncertainty from parameter estimation. The barcode overlays indicate the specific Monte Carlo iterations in which illness cases occurred. In both product categories, all illness cases cluster together in the far-right tails of the risk distribution and the far-left tails of the reduction distributions. The barcode overlays indicate the specific Monte Carlo iterations in which illness cases occurred. In both product categories, all illness cases cluster together in the

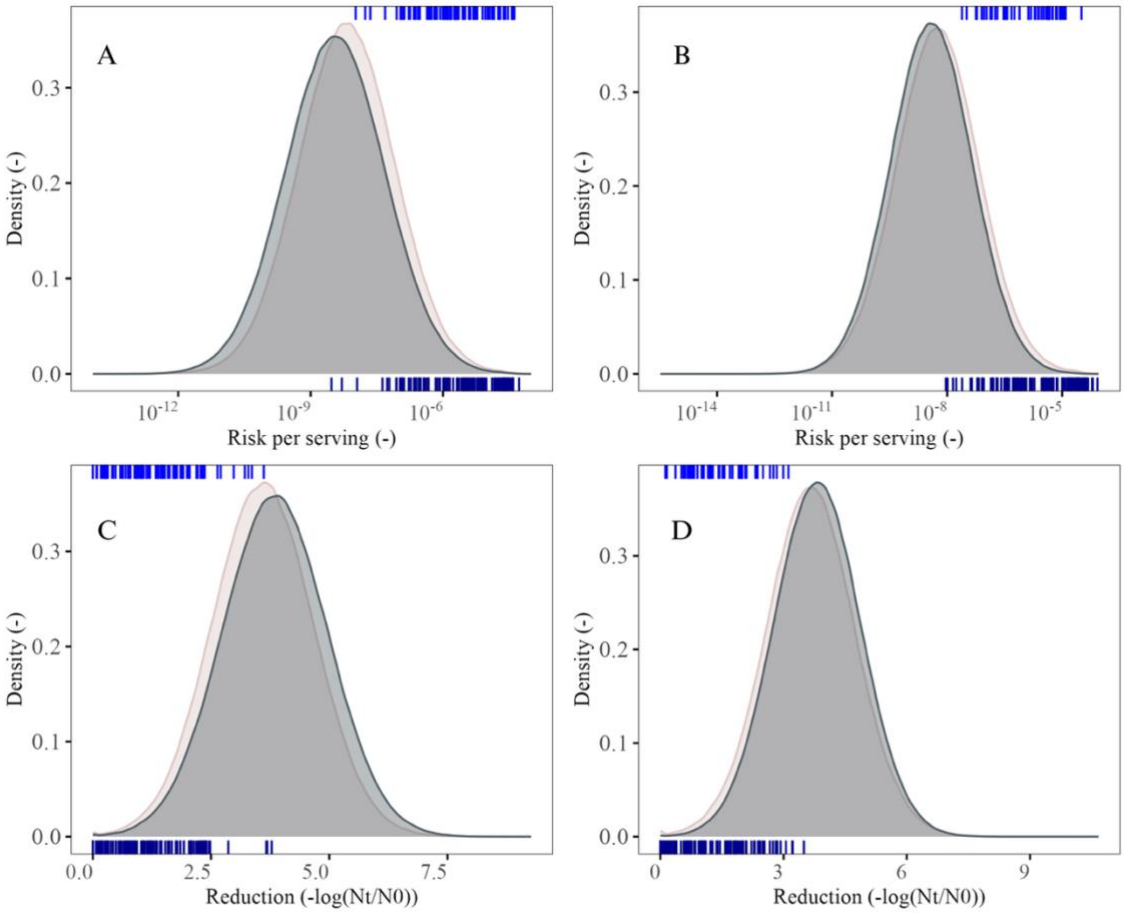


Figure 13, Probability density distributions for PEF for the Risk of illness per serving using data for all liquid products (A), and high acidic fruit juices (B). Probability density distributions for overall reduction using all liquids (C) and high acidic fruit juices (D). Layers present the modelling approach (■ = stochastic, ■ = probabilistic). Barcode plots show the occurrence of illness cases with respect to the x-variable (top = stochastic, bottom = probabilistic).

far-right tails of the risk distribution and the far-left tails of the reduction distributions. This pattern confirms that inadequate microbial reductions are the primary drivers of illness events. Based on the simulation outputs, the minimum reductions at which the first illness cases occurred were 3.84 log CFU/mL and 3.43 log CFU/mL for the stochastic model in liquid products and high acidic fruit juices, respectively. For the Bayesian model, these thresholds were 3.05 and 3.94 log CFU/mL. This confirms that rare, low-reduction outcomes are driving the estimated risks. The inclusion of variability also affects the parameter sensitivity profile. As shown in Figure 12, Spearman's rank correlation coefficients for key variables reveal a shift in influence when moving from deterministic to stochastic and Bayesian models. In the deterministic case, initial contamination and serving size were the dominant factors affecting variability in the probability of illness. However, as variability in reduction is introduced, their relative influence reduces. The correlation between reduction and illness probability remains strong but weakens slightly in the Bayesian case. As explained, the influence of energy input remains minimal across all models, since its variability was intentionally kept small to reflect well-controlled industrial settings. Consequently, other factors such as field uniformity, temperature spikes or conductivity grains may have a greater influence on microbial inactivation than energy input alone (Roobab et al., 2022). The findings from the risk assessment have already demonstrated that microbial inactivation by PEF is highly variable, and that a notable proportion of illness cases can still occur when the process is applied at industrially relevant conditions. To better understand the reliability of achieving sufficient reductions, the probability of achieving a certain reduction based on the total energy input (kJ/L) was created for both models based on the stochastic approach. These curves are shown in Figure 14, where each line presents a different target reduction level and the shades are representing the variability in the metadata. Comparing the two models, the specificity level has limited influence on the shape or steepness of the probability curves. Based on the models made for current data and including its full variability, higher total energy inputs than industrially desired are required to obtain sufficient reduction. These probability functions provide critical insight into the practical limitations of modelling PEF as microbial inactivation technology. While previous models suggested that PEF can achieve sufficient log reductions under optimised conditions, these findings show that ensuring a 5-log reduction with high certainty remains difficult, even with consistent energy inputs. A major contributor to this variability is the quality of the available data. As previously highlighted in literature, many studies on PEF do not report sufficient methodological details, such as pulse form, waveform,

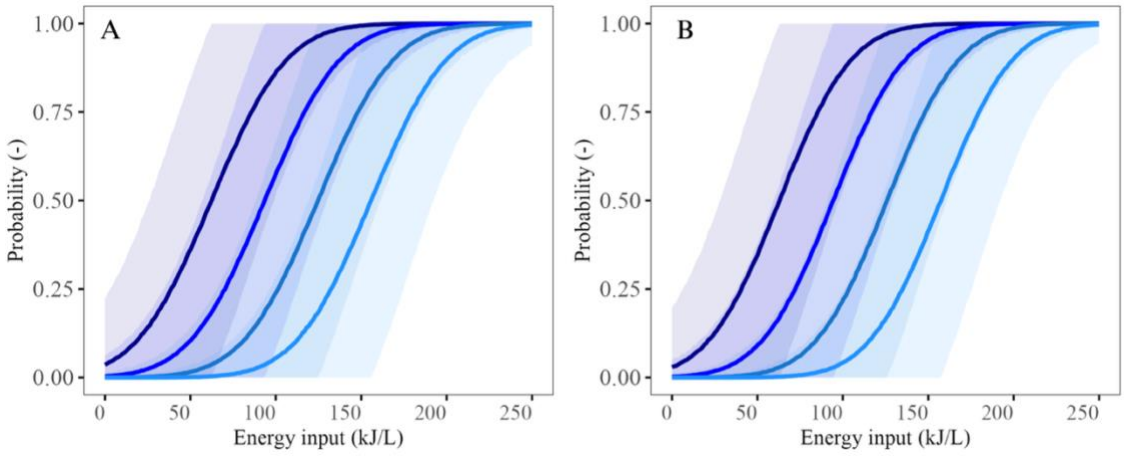


Figure 14, Probability functions showing the probability to achieve a certain log reduction based on the total energy input (kJ/L) for PEF. Data shown for all liquids (A), and high acidic fruit juices (B). Curves present the probability of achieving a certain log reduction, while shaded areas reflect the variability derived from underlying log_{10} D distributions and process metadata. Colours indicate 2-log, 3-log, 4-log, 5-log reductions.

conductivity or local heating effects, to allow for reproduceable modelling (Raso et al., 2016). Unlike TP or HPP, where standardisation has progressed further, the metadata for PEF remains more fragmented and less harmonised, limiting the accuracy of prediction models.

3.4 Technology comparison

3.4.1 The effect variability on the efficacy of the processing technologies

To understand the drivers of variability in STEC inactivation, the variability in the datasets was analysed to identify which factors in the collected data contributed most to the observed spread in $\log_{10} D$ or $\log_{10} Reduction$ values. These results are summarised in Table 9, providing insight into how strain, food matrix, pH, and study level effects influence variability across the technologies, and how the explanatory power of each factor shifts when moving from generic to product-specific datasets. Across all three technologies, the strain of *E.coli* consistently emerged as a dominant contributor to the explained variance. This effect was the most pronounced in HPP, where the strain accounted for over 50% of the total variability at both levels. Thermal processing showed a similar trend, although the strain effect slightly deviated in the specific model for high acidic fruit juice, likely due to the reduced heterogeneity in strains used within these studies. The effectiveness of microbial inactivation in HPP not only depends on pressure and holding time, but also on these parameters, which are not always consistently reported across studies (Chacha et al., 2021). For PEF, the strain effect was dropped significantly in the specific model, suggesting that here the strain variation was limited in this subgroup as well. The food matrix composition also contributed to variability in all models, explaining between 5% and 15% of the total variability. The effect of pH was found limited, although this study already took this into account by creating the models for high acidic fruit juice. Unlike TP and HP, where temperature and pressure are well-defined parameters, the microbial inactivation under PEF is influenced by multiple factors including electric field strength, pulse width, frequency and conductivity, which are often reported inconsistently, making the modelling often based on estimations (Raso et al., 2016) (Salehi, 2020). Another important note is the unexplained variability in all models, where up to 50% can be found in thermal processing. This indicates that the variability is at least partly explainable by the metadata structure, but that the extent of explanation depends on how well strain and matrix are represented. One should also be critical towards these percentages, as they may shift as more explanatory variables are added to the mode. This since the total variance

Table 9, Explained variance in the relation between $\log_{10} D$ ~ Temperature/Pressure for TP and HPP and log Reduction ~ Energy input for PEF in the collected metadata.

Thermal processing	**Sum of squares**	**Explained variance (%)**	**Sum of squares**	**Explained variance (%)**
Thermal Processing	*Level 1: All Liquid Products*		*Level 2: High acidic fruit juices*	
Strain	3.15	27.2%	1.00	19.0%
Food matrix	1.06	9.2%	0.27	5.1%
pH	0.12	1.0%	0.01	0.2%
Study	1.21	10.5%	1.41	26.8%
Residuals	6.02	52.1%	2.58	49.0%
High Pressure Processing	*Level 1: All Liquid Products*		*Level 2: High acidic fruit juices*	
Strain	44.21	52.3%	3.58	53.4%
Food matrix	9.27	11.0%	0.37	5.5%
pH	4.72	5.6%		0.0%
Study	0.34	0.4%		0.0%
Temperature	1.54	1.8%	0.30	4.5%
Residuals	24.40	28.9%	2.45	36.6%
Pulsed Electric Fields	*Level 1: All Liquid Products*		*Level 2: High acidic fruit juices*	
Strain	68.45	43.3%	0.80	4.9%
Food matrix	23.51	14.9%	1.32	8.0%
pH	15.22	9.6%		0.0%
Residuals	50.94	32.2%	14.34	87.1%

explained is distributed among all included variables. The total unexplained residual variance can only decrease as more explanatory variables are added.

Across the three technologies evaluated in this study, the incorporation of variability through stochastic and Bayesian had markedly different effects on the final risk estimates, as captured within the QMRA framework. Interestingly, when moving from deterministic to stochastic and Bayesian models, thermal processing showed the most substantial relative increase in estimated illness cases. While deterministic predictions resulted in negligible risk under all conditions, introducing variability led to a significant amount of projected illness cases annually, especially in the Bayesian model for high acidic fruit juice. This large shift reflects the effect of parameter uncertainty within the thermal inactivation used in this study, where minor reductions in treatment efficacy can shift predicted risk from virtually zero to a significant public health concern. High pressure processing showed a different pattern. Unlike thermal processing, HPP already presented risk in its deterministic models, especially for the generic all liquid products approach. However, the effect of incorporating variability was mode modest in relative terms, though still substantial in absolute impact. This suggests that while HPP models already capture some of the risk in STEC inactivation, additional stochastic or Bayesian approaches still revealed meaningful tails in risk distributions that elevated the projected number of cases. Pulsed Electric Field processing behaved differently from TP and HPP. Although risk estimates increased when variability was introduced, the magnitude of the shift was smaller relative to the other technologies. This more controlled effect is partly due to the way variability was handled in the PEF models, where reductions were directly modelled and not log-transformed (based on $\log_{10} D$ values) as in the other two technologies. This difference compresses the influence of extreme values in the lower tails of the reduction distribution, which are often the primary drivers of illness events in probabilistic outcomes. Consequently, while the number of predicted illness cases increased with variability, the rise was less pronounced. These results highlight that the impact of variability is not uniform across technologies and is shaped by both the structure of the underlying model and by how inactivation is biologically and mechanistically expressed. Additionally, metadata gaps, such as missing strain identifiers, incomplete process parameter reporting or inconsistent measurement conditions further constrain the ability to fully capture the variability present in inactivation kinetics. Together, these findings show that while the framework developed in this study allows for consistent incorporation of variability across technologies, its expression and impact on risk estimates remains technology specific.

3.5 General discussion

As in most QMRAs, input parameters are often based on previous quantitative findings, expert knowledge or assumptions, making the whole risk assessment highly uncertain (Nauta et al., 2002). Uncertainty cannot always be quantified or considered in QMRA frameworks. Possibilities are present where 2-dimensional Monte Carlo simulations are created, where uncertainty is treated as an outer loop. For every uncertain scenario, variability can be added, allowing to distinguish between variability and uncertainty. However, this approach does not make the overall model more certain or decisive. Therefore, this study only incorporated uncertainty in the regression approach by applying Bayesian regression and will elaborate on the effect of uncertainty separately in this section. The largest uncertain parts in a QMRA are often the biological parameters, such as the initial contamination and the prevalence, which describe the probability of an initial contamination being present. Also, the dose-response model is highly uncertain, due to a lack of reported data, the presence of many assumptions or limited outbreak data.

3.5.1 Dose-Response models

For the quantification of a food-safety risk, it is of high meaning to relate the probable exposures during consumption (the dose) to a connected health impact (the response). The main problem in dose-response quantification is that the effect of exposing humans to pathogens of varying virulence cannot be measured by direct experimentation (Zwietering & Havelaar, 2006). Therefore, indirect approaches are often applied to estimate this dose-response relation, generally resulting in present uncertainties. To address for these uncertainties, nine different dose-response models were found for *E. coli*, where most models were based on similar outbreak data. A previous study by Strachan et al. in 2005, summarised these STEC outbreak data (Table 2) and made new up-to-date dose-response models (Strachan et al., 2005). This study uses that summary of outbreak data, to fit the nine different dose-response models for *E. coli* found in literature. These models are summarised in Table 10. As explained in section 2.3.5, the models were fitted against the outbreak data in Table 2. To test the best fit, the AIC and RSS were evaluated, where the lowest AIC suggests the best fit with respect to the data. When evaluating the model fit, the lowest AIC was found for the original Beta-Poisson (BP) model found by Crockett et al. in 1996, where the model was used to describe the generalised dose-response relation for *Shigella* and *Escherichia coli* (Crockett et al., 1996). The next best fit includes the BP models by Strachan et al. (2005) and Teunis et al. (2004). However, while AIC and RSS provide useful statistical rankings, they do not reflect biological realism especially at low doses. One of the critical limitations of dose-response modelling is the high uncertainty in the low-dose range, where most real-world foodborne exposures occur (Teunis & Havelaar, 2000). Since many outbreak datasets are based on high-dose experimental infections, the extrapolation of dose-response models to low doses remains challenged (Pratt et al., 2021). To evaluate this uncertainty, the probability of illness in the same models was fitted on a logarithmic scale, and the AIC and RSS was calculated separately.

The fittings of the models with respect to the outbreak data is also visible in Figure 15, where the initial fittings are displayed in Figure 15A. When looking to the $\log_{10} P$ specifically, the same graph was made, displaying the probability of illness in logarithmic scale (Figure 15B). This allows to concentrate more on the lower doses respectively. At high doses, there is a high variability in the alignment of most of the outbreak data where a clear distinction is visible between the exponential and Beta-Poisson models. Based on the data distribution, there is no clear trend visible between the ingested doses and their connected probabilities, making it hard to find the most suitable model. At lower doses, the substantial spread in predicted illness probabilities does not reduce. The BP model from Cassin et al. (1998) has a lower AIC (-22.72) but performs worse when using the logarithmic probability data (17.59), suggesting that its parameterisation

Table 10, Dose-response models for E. coli extracted from literature, with corresponding colour marks (Figure 15), dose-response parameters and their connected model fit with respect to the outbreak data displayed in Table 2.

Reference	Model	D/R parameters	RSS	AIC	RSS (logP)	AIC (logP)
Strachan et al., 2005 (EXP)(used in QMRA)	EXP (used)	r = 0.00208	0.58	-18.99	40.17	14.91
Strachan et al., 2005 (BP, mean)	BP (mean)	a = 0.057, b = 2.2183	0.35	-21.01	38.28	16.52
Strachan et al., 2005 (BP, median)	BP (median)	a = 0.2241, b = 4.880	0.87	-13.78	62.24	20.41
Cassin et al., 1998 (BP)	BP	a = 0.267, b = 5.435	1.08	-12.05	65.43	20.81
Cassin et al., 1998 (EXP, calculated)	EXP	r = 0.049	3.53	-4.54	86.81	21.07
Teunis et al., 2004 (teachers)	BP	a = 0.0844, b = 1.442	0.38	-20.34	48.46	18.41
Teunis et al., 2004 (children)	BP	a = 0.0496, b = 1.001	0.37	-20.49	39.62	16.80
Nauta et al., 2002	EXP	r = 0.0051	1.56	-11.08	61.39	18.30
Crockett et al., 1996 (*Shigella*)	BP	a = 0.162, b = 15.86	0.28	-22.72	43.71	17.59

BP = Beta-Poisson model, EXP = Exponential model, AIC = Akaike Information Criteria, RSS = Residual sum of Squares

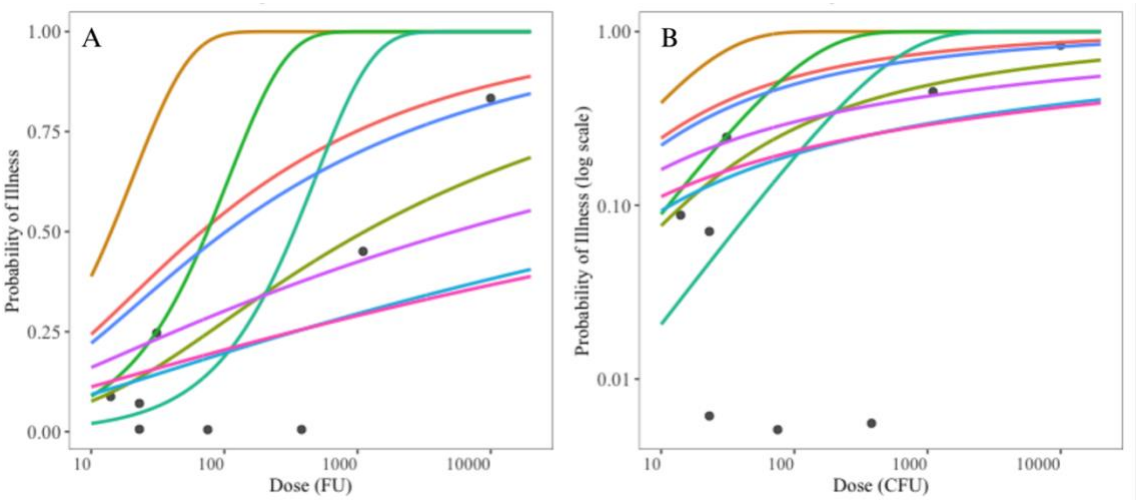


Figure 15, Model fitting of the different dose-response models displayed in table 8, with respect to the outbreak data displayed in Table 2. Data shown for all datapoints (A), and on a log(10) scale (B).

may not be well-suited for low-dose exposure scenarios (Cassin et al., 1998). Based on the logarithmic probability, the exponential model found in Strachan et al. (2005) was found to fit the best (14.91) and was therefore used as baseline in the QMRAs in this study. To assess the practical impact of dose-response model uncertainty, the effect on estimated illness cases was evaluated using the mean risk per serving for HPP models applied to high acidic fruit juice. The results are summarised in Table 11, where the estimated number of cases for each dose-response model are displayed for both the stochastic and probabilistic approach, alongside the fold increase relative to the used model. These case estimates vary substantially depending on the model applied, showing that the choice of dose-response model can introduce large uncertainty effects in a QMRA. These extreme variations in estimated cases demonstrate the high sensitivity of risk estimated. In general, the Beta-Poisson models predict higher risk than the exponential models. This align with earlier findings by Teunis and Havelaar in 2000, stating that these are not "single hit" models, especially at lower doses (Teunis & Havelaar, 2000). For example, using Cassin et al. (1998) BP instead of the exponential model from Strachan et al. (2005), the number of estimated cases increased by 20-fold. However, where most models substantially increased risk estimates, the exponential model from Nauta et al. (2002) "only" shows a 2-fold increase, indicating relatively small deviations from the used model (Nauta et al., 2002). Generally, these large variations across dose-response models highlight the critical uncertainty in risk estimation

Table 11, The impact of using different dose-response models on the estimated number of annual cases in Europe. Results are based on the Stochastic and Probabilistic models for high acidic fruit juice, using HPP technology. The fraction increase was calculated with respect to the model from Strachan et al. (2005) which was used as baseline in this study.

Dose response model	**Model**	**Stochastic**		**Probabilistic**	
		STEC cases*	Fraction increase	STEC cases*	Fraction increase
Strachan et al., 2005 (EXP)(used in QMRA)	EXP (used)	266		3173	
Strachan et al., 2005 (BP, mean)	BP (mean)	2807	11	27655	9
Strachan et al., 2005 (BP, median)	BP (median)	5346	20	56993	18
Cassin et al., 1998 (BP)	BP	5523	21	59953	19
Cassin et al., 1998 (EXP, calculated)	EXP	6197	23	71886	23
Teunis et al., 2004 (teachers)	BP	5842	22	55721	18
Teunis et al., 2004 (children)	BP	4797	18	42685	13
Nauta et al., 2002	EXP	630	2	7870	2
Crockett et al., 1996 (*Shigella*)	BP	1282	5	14405	5

* Estimated number of annual STEC cases on European scale, based on prediction models used in this study. Cases calculated by mean risk per serving

and can display a more dominant influence on the absolute outcomes of risk estimates, more than variability in inactivation kinetics. Although, it had less effect on the relative impact of interventions between technologies. Given this, it is not only justified but necessary to emphasize that dose-response model selection is one of the most critical sources of uncertainty in a QMRA.

3.5.2 Prevalence and initial contamination

Besides dose-response model selection, another critical source of uncertainty in QMRAs is the prevalence of contamination and the initial concentration of pathogens in a food product (Duarte et al., 2015). Both factors directly influence the estimated risk outcomes, as they determine the probability and severity of exposure scenarios in the model simulations. In this study, the prevalence of STEC contamination was determined based on sampling data from 2019 – 2023. The recorded prevalence for non-ready-to-eat fruits and vegetables was 1 out of 778 samples (0.13%) (EFSA, 2019). However, for fruit juices specifically, no positive samples were detected. This is a critical consideration when interpreting risk estimates, as the lack of reported contamination in fruit juices suggests a potential overestimation. Despite the low prevalence, it is important to realise that prevalence data are inherently limited by the detection limits of testing procedures (Byrne et al., 2021). If a higher prevalence was assumed, the risk estimation would increase proportionally, as the probability of a contaminated product reaching consumers would be greater. This highlights a large limitation in microbial risk assessment. While prevalence data provide a baseline probability of contamination, they do not account for sporadic contamination events or outbreaks, which can lead to significantly higher risks.

In addition to prevalence, the initial concentration of STEC in contaminated samples is another large driver of the final risk estimation. Microbial contaminants are not uniformly distributed in a food matrix but rather follow heterogeneous distributions (Byrne et al., 2021). Typically, a Poisson distribution is used to model the random distribution of bacteria in a sample, reflecting the fact that not all samples of a contaminated product will contain the same concentration of cells (Koyama et al., 2017). For this study, the initial concentration was based on the prevalence data, where the positive sample (1 in 778) was assumed to contain STEC (presence) in a 25 mL test sample. To account for uncertainty in the actual initial concentration, a PERT distribution was created, allowing for a more realistic representation of initial concentration in fruit juices. The implications of initial concentration on risk estimation are significant. Higher initial concentrations mean that greater microbial reductions are required to achieve food safety targets. In the simulations, overall reductions of 2-4 log CFU were found in the best performing models, often sufficient to reduce the risk to acceptable levels. However, if the initial contamination levels were underestimated, these reductions may no longer be adequate, leading to higher risks. This supports the importance of regulatory guidelines, recommending to achieve at least a 5-log reduction (FDA, 2024).

3.5.3 Estimation of illness cases

Given the critical uncertainty in risk estimation, three different approaches were used to calculate the estimated number of cases in this study. Each method incorporates the Monte Carlo simulations but differ in the way that illness probability and microbial reduction are interpreted. The use of multiple approaches ensures the robustness of the estimated and allows for a comparative comparison of their impact on the final risk predictions.

The first method relies on the Bernoulli trials, introducing an additional probabilistic dimension to the risk estimations. As explained in section 2.3.5, the Bernoulli trials function as a binary simulation, where each iteration has one attempt to generate a random number between 0 and 1.

If the number is lower than the probability of illness of that serving, the outcome is recorded as 1 (an illness case). The total number of illness cases (sum of 1s) across all iterations was then corrected for prevalence, ensuring it aligned with real-world contaminations. Then this number was scaled to the total number of annual European servings and divided by the number of iterations to normalize the results across the simulations. This approach follows the method used in earlier QMRAs (van der Vossen-Wijmenga et al., 2024). Mathematically, this approach was expressed as:

$$Cases_{Bernoulli} = \frac{\left(\sum_{i=1}^{n_{iterations}} 1(Bernoulli_{,i} > 0)\right) * prevalence * EU_{servings}}{n_{iterations}} \quad [28]$$

This method captures stochastic variability in individual exposure scenarios, making it sensitive to low-probability illness events. The second method was based on the average risk of illness per serving, integrating all simulated exposure scenarios. As explained, each iteration corresponded to a unique scenario, with its illness probability. These probabilities were multiplied with the prevalence to calculate the risk for each serving. The mean risk is then taken across all iterations and scaled up by multiplying it with the annual European servings. This method was also used in earlier QMRAs (van der Vossen-Wijmenga et al., 2024). Mathematically, this approach was given by:

$$Cases_{Average} = \overline{Risk_{serving}} * EU_{servings} \quad [29]$$

The last method applies the probability integral method for overall reduction based on $\log_{10} D$ variability, estimating the arithmetic mean survivor count in the Monte Carlo simulations (Zwietering et al., 2021). This approach differs from the previous two as it directly incorporates the microbial reduction data rather than focussing on iteration-level illness probabilities. The mean concentration of surviving pathogens per serving is taken across all iterations and multiplied with the mean serving size to calculate the mean dose. This dose is used to calculate the overall illness probability, after which it is multiplied with the prevalence and annual European servings. This approach was defined as:

$$Cases_{Overall} = \left(1 - e^{-r*(\overline{Nt}*\overline{Serving_{size}})}\right) * prevalence * EU_{servings} \quad [30]$$

Where $\overline{Nt}$ is the arithmetic mean survivor count and $\overline{Serving_{size}}$ is the average serving size. To assess the impact of using these three methods on the QMRA, a similar approach as in the previous section was applied, where the estimated annual illness cases were calculated for the high acidic fruit juice models in HPP, tabulated in Table 12. The results from Table 12 confirm that all methods produced very similar estimates amongst the different approaches, indicating that despite the different calculations, the overall risk prediction remains stable. The differences are minor, especially for the mean risk and overall reduction approaches. The Bernoulli trials tend to result in slightly lower estimates, most likely due to their binary nature. The application of three independent approaches to estimate illness cases confirms that the overall risk prediction is consistent, ensuring the robustness of the QMRA approach.

Table 12, Estimated number of cases based on the three calculation methods. Results displayed for the high acidic fruit juice models using HPP technology.

Model & Specificity	Estimated cases by Bernoulli trials	Estimated cases by mean Risk	Estimated cases by overall reduction
Level 2: High Acidic Fruit Juice			
Linear model (Deterministic)	± 0	± 0	± 0
Linear model (Stochastic)	± 154	± 266	± 267
Bayesian model (Probabilistic)	± 3019	± 3173	± 3177

3.5.4 Efficacy versus Efficiency

The assessment of efficacy and efficiency is critical for determining the feasibility of the NTP technologies. While all technologies, TP, HPP and PEF have demonstrated the ability to reduce microbial contamination in liquid products, their energy consumption varies largely. The findings presented in this study, based on the meta-regression models and industry relevant conditions, highlight the trade-off between efficacy and energy efficiency. The efficiency of these technologies were researched by Aganovic et al. in 2017, where common process parameters for microbial inactivation for TP, HPP and PEF were used, and their respective energy consumptions were calculated (Aganovic et al., 2017). Key findings of these study were contractionary, where thermal processing showed the lowest overall energy consumption of approximately 0.04 kWh/L, compared to 0.2 kWh/L for high pressure processing and 0.12 kWh/L for pulsed electric fields. The primary reason for this is the efficient re-use of thermal energy by heat exchanging during thermal pasteurizing processes, in which the energy in HPP and PEF cannot be efficiently recycled (Aganovic et al., 2017). Importantly, the energy input in HPP is nearly independent of the holding time, where most energy is used to build up pressure and not maintain it. As such, a 3 minute or 8-minute treatment may consume nearly the same amount of energy, making treatment duration less influential on energy use. In contrast, the energy efficiency in PEF is strongly affected by product conductivity and thermal losses. While some of the heat generated in PEF can theoretically be recovered, Aganovic et al. noted that this is potential not typically realised at scale. Another important distinction raised by this study is that HPP and PEF operate on electricity and can therefore be decoupled from fossil fuel use if renewable energy is used. In terms of process configuration, PEF offers the advantage of continuous operation, unlike the batch-based nature of HPP.

To address the efficacy of the technologies against the energy consumption, the process conditions from Aganovic et al. were used on the models created in this study, visualised in Table 13. In this table, where the efficacy of the technology, the impact of variability and the energy efficiency come together. The table displays the overall reduction, based on the mean survivor count, for TP, HPP and PEF, based on these process characteristics and their respective estimation of annual illness cases. When using these process parameters in the similar QMRA framework, based on the modes defined before, one can observe significant differences in microbial inactivation amongst the technologies and specificity levels. Unlike many expectations, thermal processing remains the most efficient technology to use in terms of energy use and corresponding reduction. Even when considering all variability, the lower energy efficiency still leads to a positive result. The most specific models for high acidic fruit juice show comparable reductions in the stochastic and probabilistic approaches. However, the higher initial energy consumption

Table 13, Overall reductions based on process parameters described in Aganovic et al. (2017) and their estimated energy consumptions. Overall reductions based on the arithmetic mean of survivors in 1000000 monte carlo iterations. Normalised energy inputs presented as kWh/L/log reduction.

Model & Specificity	-log(Nt/No) TP	STEC cases*	-log(Nt/No) HPP	STEC cases*	-log(Nt/No) PEF	STEC cases*
Processing conditions (Aganociv et al., 2017)	74°C, 30 seconds		600 Mpa, 5 minutes		188 kJ/L	
Level 1: Liquid Products						
Linear model (Deterministic)	28.2	0	1.0	107106	6.0	1
Linear model (Stochastic)	3.7	224	0.7	230444	4.6	26
Bayesian model (Probabilistic)	2.0	11152	0.7	238117	4.5	34
Level 2: High Acidic Fruit Juice						
Linear model (Deterministic)	13.7	0	24.3	0	6.0	1
Linear model (Stochastic)	3.7	181	3.6	307	4.7	22
Bayesian model (Probabilistic)	1.3	45336	2.4	4155	4.2	62
Estimated energy usage (Aganociv et al., 2017) (kWh/L)	0.04		0.2		0.12	

* Estimated number of annual STEC cases on European scale, based on prediction models used in this study. Cases calculated by mean risk per serving

for HPP and PEF based on these process characteristics is significant. The PEF technology models did not improve much with increased matrix specificity and remains much less efficient than TP with respect to energy use. When considering food safety, PEF processing shows the most promising results, where the least estimated illness cases occur. Unfortunately, the relation between the estimated energy input and the overall reduction is not straightforward and one cannot simply assume a linear relationship for estimated energy input and log reductions in microbial kinetic modelling. Overall, these findings highlight the complex trade-off between efficacy, energy inputs and industrial feasibility. Although HPP and PEF offer quality preservation advantages, their implementation must consider energy intensity and system design limitations. Evaluating these technologies under realistic operating conditions and within a full life cycle perspective, is highly recommended for guiding future innovation in sustainable food processing.

Chapter 4

Conclusions

This study evaluated the impact of variability on the efficacy of microbial inactivation for thermal-, high pressure-, and pulsed electric field processing, using a consistent quantitative microbial risk assessment framework. The models incorporated reported inactivation data for *Escherichia coli* O157:H7 across different levels or product specificity, allowing a comparative assessment of both technological performance and the influence of variability in reported inactivation kinetics. A key finding is that the effect of variability on microbial reductions is highly significant, as process efficacy decreases amongst all technologies. Thermal processing models, build on standard process parameters and widely reported inactivation data, displayed the lowest variability but yielded relatively inconsistent predictions when variability was considered. In contrast, high pressure processing showed more pronounced sensitivity to product specificity. When using product-specific data for high acidic fruit juices, the model's predictive performance improved meaningfully, capturing greater pressure sensitivity and leading to more effective reductions. For pulsed electric field processing, the predictive reliability remained constant. While the estimated energy input was standardised in the modelling approach, the associated reduction estimates remained highly variable. This is attributed to broader process diversity in the literature and to variability in the underlying food matrix, strain type and experimental conditions. The incorporation of variability in the QMRA, using both stochastic and Bayesian uncertainty quantification, had a pronounced effect on predicted risk outcomes. Across all technologies, deterministic models consistently underestimated the probability of illness, often predicting zero or near zero case estimates. In contrast, the stochastic and Bayesian models captured the full variability of reduction outcomes, revealing that rare but possible low-inactivation events can considerably affect the risk profile. These events, represented in the upper tails of the $\log_{10} D$ and log reduction distributions, were shown to be critical drivers for predicted illness cases. As a result, the estimated risk increased considerably when variability was accounted for, underscoring the limitations of using single point estimates in representing real world process performance.

The study also highlighted dose-response modelling as a major source of uncertainty within a QMRA. Different published models, though based on similar outbreak data, produced different risk predictions, particularly in the low-dose exposure range where most foodborne risks are expected to occur. Using a set of nine literature models, differences in annual illness estimates reached up to 25-fold, showing the sensitivity of QMRA outcomes to the chosen function. Lastly, this study drew attention to the distinction between efficacy and efficiency in processing technologies. While thermal processing achieved high reduction levels with relatively low energy input, HPP and PEF required substantially more energy to reach comparable reduction levels. Although HPP and PEF offer quality preservation advantages, their implementation must consider energy intensity and system design limitations. Together, these findings emphasize that variability is not merely a source of uncertainty, but a defining factor in the actual efficacy of food processing technologies. By systematically incorporating variability into a unified QMRA framework, this thesis provides a realistic and transferable approach for modelling microbial inactivation. This work reinforces the need for transparent, data-driven modelling in microbiology and offers a foundation for more robust validation of novel processing technologies.

Chapter 5

Recommendations

The findings of this thesis support several recommendations to improve the quality and robustness of predictive microbiology and microbial risk assessments for novel processing technologies. First, variability should be incorporated as a standard component in both model development and quantitative microbial risk assessments (QMRA). The demonstrated impact of variability on predicted illness outcomes underlines the inadequacy of deterministic models as the default approach, particularly when dealing with emerging or non-standardised processing methods. Instead, the use of stochastic or probabilistic modelling frameworks should be encouraged in order to capture the inherent variability in microbial inactivation and better reflect real-world process behaviour. Second, the quality and consistency of metadata in published inactivation studies should be improved. In this thesis, especially for high-pressure processing (HPP) and pulsed electric fields (PEF), the usability of literature data was frequently hindered by the incomplete reporting of critical parameters such as microbial strain, food matrix composition, pH, and process conditions. These limitations reduce the reliability of meta-regression models and restrict their generalizability. To address this, authors and journals should adopt and enforce standardised reporting guidelines aligned with the FAIR data principles (Findable, Accessible, Interoperable, Reusable), ensuring that microbial inactivation data can be effectively integrated into future predictive modelling efforts. Finally, this thesis highlights the need to define fit-for-purpose specificity levels in model development. The comparative modelling approach revealed that the optimal level of specificity is not universal, but technology dependent. For HPP, models based on highly specific subsets of data (e.g., high-acid fruit juices) yielded more precise and reliable predictions. In contrast, for thermal processing (TP) and PEF, the added specificity provided limited improvement and may not justify the associated data constraints. Future risk assessments should therefore approach model specificity as a flexible, context-dependent decision, rather than adhering to a fixed standard across technologies or applications.

Chapter 6

References

Abe, H., Garre, A., Koseki, S., den Besten, H. M. W., & Zwietering, M. H. (2023). Analysis of a quantitative risk assessment of listeriosis from pasteurised milk: The combinations of which factors cause listeriosis in this low-risk food? *Food Control*, *152*, 109831. https://doi.org/10.1016/j.foodcont.2023.109831

Aganovic, K., Hertel, C., Vogel, Rudi. F., Johne, R., Schlüter, O., Schwarzenbolz, U., Jäger, H., Holzhauser, T., Bergmair, J., Roth, A., Sevenich, R., Bandick, N., Kulling, S. E., Knorr, D., Engel, K., & Heinz, V. (2021). Aspects of high hydrostatic pressure food processing: Perspectives on technology and food safety. *Comprehensive Reviews in Food Science and Food Safety*, *20*(4), 3225-3266. https://doi.org/10.1111/1541-4337.12763

Aganovic, K., Smetana, S., Grauwet, T., Toepfl, S., Mathys, A., Van Loey, A., & Heinz, V. (2017). Pilot scale thermal and alternative pasteurisation of tomato and watermelon juice: An energy comparison and life cycle assessment. *Journal of Cleaner Production*, *141*, 514-525. https://doi.org/10.1016/j.jclepro.2016.09.015

Aghajanzadeh, S., & Ziaiifar, A. (2017). A review of pectin methylesterase inactivation in citrus juice during pasteurisation. *Trends in Food Science & Technology*, *71*. https://doi.org/10.1016/j.tifs.2017.10.013

Bach, S. J., McAllister, T. A., Veira, D. M., Gannon, V. P. J., & Holley, R. A. (2002). Transmission and control of *Escherichia coli* O157:H7—A review. *Canadian Journal of Animal Science*, *82*(4), 475-490. https://doi.org/10.4141/A02-021

Barba, F. J., Koubaa, M., Do Prado-Silva, L., Orlien, V., & Sant'Ana, A. D. S. (2017). Mild processing applied to the inactivation of the main foodborne bacterial pathogens: A review. *Trends in Food Science & Technology*, *66*, 20-35. https://doi.org/10.1016/j.tifs.2017.05.011

Bigelow, W. D. (1921). The logarithmic nature of thermal death time curves. *The Journal of Infectious Diseases*, *29*(5), 528-536. https://doi.org/10.1093/infdis/29.5.528

Byrne, D. M., Hamilton, K. A., Houser, S. A., Mubasira, M., Katende, D., Lohman, H. A. C., Trimmer, J. T., Banadda, N., Zerai, A., & Guest, J. S. (2021). Navigating data uncertainty and modeling assumptions in quantitative microbial risk assessment in an informal settlement in Kampala, Uganda. *Environmental Science & Technology*, *55*(8), 5463-5474. https://doi.org/10.1021/acs.est.0c05693

Cassin, M. H., Lammerding, A. M., Todd, E. C. D., Ross, W., & McColl, R. S. (1998). Quantitative risk assessment for *Escherichia coli* O157:H7 in ground beef hamburgers. *International Journal of Food Microbiology*, *41*(1), 21-44. https://doi.org/10.1016/S0168-1605(98)00028-2

Chacha, J. S., Zhang, L., Ofoedu, C. E., Suleiman, R. A., Dotto, J. M., Roobab, U., Agunbiade, A. O., Duguma, H. T., Mkojera, B. T., Hossaini, S. M., Rasaq, W. A., Shorstkii, I., Okpala, C. O. R., Korzeniowska, M., & Guiné, R. P. F. (2021). Revisiting non-thermal food processing and preservation methods: Action mechanisms, pros and cons: A technological update (2016–2021). *Foods*, *10*(6), 1430. https://doi.org/10.3390/foods10061430

Chiozzi, V., Agriopoulou, S., & Varzakas, T. (2022). Advances, applications, and comparison of thermal (pasteurisation, sterilisation, and aseptic packaging) against non-thermal (ultrasounds, UV radiation, ozonation, high hydrostatic pressure) technologies in food processing. *Applied Sciences*, *12*(4), Article 4. https://doi.org/10.3390/app12042202

Crockett, C. S., Haas, C. N., Fazil, A., Rose, J. B., & Gerba, C. P. (1996). Prevalence of shigellosis in the U.S.: Consistency with dose-response information. *International Journal of Food Microbiology*, *30*(1-2), 87-99. https://doi.org/10.1016/0168-1605(96)00993-2

den Besten, H. M. W., Aryani, D. C., Metselaar, K. I., & Zwietering, M. H. (2017). Microbial variability in growth and heat resistance of a pathogen and a spoiler: All variabilities are equal, but some are more equal than others. *International Journal of Food Microbiology*, *240*, 24-31. https://doi.org/10.1016/j.ijfoodmicro.2016.04.025

den Besten, H. M. W., & Zwietering, M. H. (2012). Meta-analysis for quantitative microbiological risk assessments and benchmarking data. *Trends in Food Science & Technology*, *25*(1), 34-39. https://doi.org/10.1016/j.tifs.2011.12.004

Duarte, A. S. R., Stockmarr, A., & Nauta, M. J. (2015). Fitting a distribution to microbial counts: Making sense of zeroes. *International Journal of Food Microbiology*, *196*, 40-50. https://doi.org/10.1016/j.ijfoodmicro.2014.11.023

Dziadek, K., Kopeć, A., Dróżdż, T., Kiełbasa, P., Ostafin, M., Bulski, K., & Oziembłowski, M. (2019). Effect of pulsed electric field treatment on shelf life and nutritional value of apple juice. *Journal of Food Science and Technology*, *56*(3), 1184-1191. https://doi.org/10.1007/s13197-019-03581-4

EC. (1997). *Health and consumer protection—Scientific Committee on Food—Previous outcome of discussions—07*. https://ec.europa.eu/food/fs/sc/oldcomm7/out07_en.html

ECDC. (2020). *STEC infection Annual Epidemiological Report 2020*.

EFSA. (2019, 2023). *STEC dashboard | EFSA*. https://www.efsa.europa.eu/en/microstrategy/stec-dashboard

EFSA, L. updated. (2024, december 10). *Shiga toxin-producing Escherichia coli*. ArcGIS StoryMaps. https://storymaps.arcgis.com/stories/ff3eb57b8d40474485d533b3cd2cfa63

Eurostat. (2024). *Demography in Europe 2024: Take a guess!* https://ec.europa.eu/eurostat/cache/interactive-publications/demography/2024/00/index.html

FDA. (2017). *Pressure Safe, LLC - 520442—06/07/2017*. FDA. https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/pressure-safe-llc-520442-06072017

FDA. (2024, januari 10). *Guidance for industry: Juice hazard analysis critical control point hazards and controls guidance, First Edition*. FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-juice-hazard-analysis-critical-control-point-hazards-and-controls-guidance-first

Garre, A., Clemente-Carazo, M., Fernández, P. S., Lindqvist, R., & Egea, J. A. (2018). Bioinactivation FE: A free web application for modelling isothermal and dynamic microbial inactivation. *Food Research International*, *112*, 353-360. https://doi.org/10.1016/j.foodres.2018.06.057

Garre, A., Georgalis, L., Lindqvist, R., & Fernandez, P. S. (2025). Development and validation of microbial inactivation models using Bioinactivation4. In *Basic Protocols in Predictive Microbiology Softwares* (pp. 13-58). Humana, New York, NY. https://doi.org/10.1007/978-1-0716-4112-5_2

Garre, A., Zwietering, M. H., & den Besten, H. M. W. (2020). Multilevel modelling as a tool to include variability and uncertainty in quantitative microbiology and risk assessment. Thermal inactivation of *Listeria monocytogenes* as proof of concept. *Food Research International*, *137*, 109374. https://doi.org/10.1016/j.foodres.2020.109374

Georget, E., Sevenich, R., Reineke, K., Mathys, A., Heinz, V., Callanan, M., Rauh, C., & Knorr, D. (2015). Inactivation of microorganisms by high isostatic pressure processing in complex matrices: A review. *Innovative Food Science & Emerging Technologies*, *27*, 1-14. https://doi.org/10.1016/j.ifset.2014.10.015

Gouvea, F. S., Padilla-Zakour, O. I., Worobo, R. W., Xavier, B. M., Walter, E. H. M., & Rosenthal, A. (2020). Effect of high-pressure processing on bacterial inactivation in açaí juices

with varying pH and soluble solids content. *Innovative Food Science & Emerging Technologies*, *66*, 102490. https://doi.org/10.1016/j.ifset.2020.102490

Haas, C. N. (2020). Quantitative microbial risk assessment and molecular biology: Paths to integration. *Environmental Science & Technology*, *54*(14), 8539-8546. https://doi.org/10.1021/acs.est.0c00664

Hauben, K. J., Bartlett, D. H., Soontjens, C. C., Cornelis, K., Wuytack, E. Y., & Michiels, C. W. (1997). *Escherichia coli* mutants resistant to inactivation by high hydrostatic pressure. *Applied and Environmental Microbiology*. https://doi.org/10.1128/aem.63.3.945-950.1997

Huang, K., Tian, H., Gai, L., & Wang, J. (2012). A review of kinetic models for inactivating microorganisms and enzymes by pulsed electric field processing. *Journal of Food Engineering*, *111*(2), 191-207. https://doi.org/10.1016/j.jfoodeng.2012.02.007

Jordan, S. L., Pascual, C., Bracey, E., & Mackey, B. M. (2001). Inactivation and injury of pressure-resistant strains of *Escherichia coli* O157 and *Listeria monocytogenes* in fruit juices. *Journal of Applied Microbiology*, *91*(3), 463-469. https://doi.org/10.1046/j.1365-2672.2001.01402.x

Katsaros, G., Alexandrakis, Z., & Taoukis, P. (2015). *High Pressure Processing of Foods: Technology and Applications*. https://doi.org/10.1201/b19397-13

Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bolton, D., Bover-Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Jacxsens, L., Skjerdal, T., Da Silva Felicio, M. T., Hempen, M., … Lindqvist, R. (2020a). Guidance on date marking and related food information: Part 1 (date marking). *EFSA Journal*, *18*(12), e06306. https://doi.org/10.2903/j.efsa.2020.6306

Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bolton, D., Bover-Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Jacxsens, L., Skjerdal, T., Da Silva Felicio, M. T., Hempen, M., … Lindqvist, R. (2020b). Guidance on date marking and related food information: Part 1 (date marking). *EFSA Journal*, *18*(12), e06306. https://doi.org/10.2903/j.efsa.2020.6306

Koyama, K., Hokunan, H., Hasegawa, M., Kawamura, S., & Koseki, S. (2017). Modeling stochastic variability in the numbers of surviving *Salmonella enterica*, *Enterohemorrhagic* , *Escherichia coli*, and *Listeria monocytogenes* cells at the single-Cell Level in a desiccated environment. *Applied and Environmental Microbiology*, *83*(4), e02974-16. https://doi.org/10.1128/AEM.02974-16

Lindgren, M., Aronsson, K., Galt, S., & Ohlsson, T. (2002). Simulation of the temperature increase in pulsed electric field (PEF) continuous flow treatment chambers. *Innovative Food Science & Emerging Technologies*, *3*, 233-245. https://doi.org/10.1016/S1466-8564(02)00044-9

Malone, A. H., Chung, Y., & Yousef, A. E. (z.d.). *Genes of* Escherichia coli *O157:H7 That are involved in high-pressure resistance*. Geraadpleegd 25 februari 2025, van https://journals-asm-org.ezproxy.library.wur.nl/doi/full/10.1128/aem.72.4.2661-2671.2006

McElreath, R. (2015). *Statistical Rethinking*.

McMeekin, T. A., Olley, J., Ratkowsky, D. A., & Ross, T. (2002). Predictive microbiology: Towards the interface and beyond. *International Journal of Food Microbiology*, *73*(2), 395-407. https://doi.org/10.1016/S0168-1605(01)00663-8

Monforti, F., Dallemand, J., Pinedo Pascua, I., Motola, V., Banja, M., Scarlat, N., Medarac, H., Castellazzi, L., Labanca, N., Bertoldi, P., Pennington, D., Goralczyk, M., Schau, E., Saouter, E., Sala, S., Notarnicola, B., Tassielli, G., & Renzulli, P. (2015). *Energy use in the EU food sector: State of play and opportunities for improvement*. https://doi.org/10.2790/158316

Morata, A., Escott, C., Loira, I., López, C., Palomero, F., & González, C. (2021). Emerging Non-Thermal Technologies for the Extraction of Grape Anthocyanins. *Antioxidants*, *10*(12), Article 12. https://doi.org/10.3390/antiox10121863

Nauta, M. J., Evers, E. G., & Takumi, K. (z.d.). *Risk assessment of Shiga-toxin producing Escherichia coli O157 in steak tartare in the Netherlands*.

Ozkan, G., Guldiken, B., & Capanoglu, E. (2019). 12—Effect of novel food processing technologies on beverage antioxidants. In A. M. Grumezescu & A. M. Holban (Red.), *Processing and Sustainability of Beverages* (pp. 413-449). Woodhead Publishing. https://doi.org/10.1016/B978-0-12-815259-1.00012-4

Pampoukis, G., Zwietering, M. H., & den Besten, H. M. W. (2024). Ranking factors affecting the decontamination efficacy of non-thermal plasma: The approach of dissipated power per plasma volume through machine learning modeling. *Innovative Food Science & Emerging Technologies*, *96*, 103773. https://doi.org/10.1016/j.ifset.2024.103773

Patterson, M. F., Quinn, M., Simpson, R., & Gilmour, A. (1995). Sensitivity of vegetative pathogens to High Hydrostatic Pressure treatment in phosphate-buffered saline and foods. *Journal of Food Protection*, *58*(5), 524-529. https://doi.org/10.4315/0362-028X-58.5.524

Petrus, R. R., Churey, J. J., & Worobo, R. W. (2020). High pressure processing of apple juice: The most effective parameters to inactivate pathogens of reference. *British Food Journal*, *122*(12), 3969-3979. https://doi.org/10.1108/BFJ-03-2020-0178

Pratt, A., Bennett, E., Gillard, J., Leach, S., & Hall, I. (2021). Dose–response modeling: Extrapolating from experimental data to real-world populations. *Risk Analysis*, *41*(1), 67-78. https://doi.org/10.1111/risa.13597

Ranjha, M. M. A. N., Kanwal, R., Shafique, B., Arshad, R. N., Irfan, S., Kieliszek, M., Kowalczewski, P. Ł., Irfan, M., Khalid, M. Z., Roobab, U., & Aadil, R. M. (2021). A critical review on pulsed electric field: A novel technology for the extraction of phytoconstituents. *Molecules*, *26*(16), Article 16. https://doi.org/10.3390/molecules26164893

Raso, J., Frey, W., Ferrari, G., Pataro, G., Knorr, D., Teissie, J., & Miklavčič, D. (2016). Recommendations guidelines on the key information to be reported in studies of application of PEF technology in food and biotechnological processes. *Innovative Food Science & Emerging Technologies*, *37*, 312-321. https://doi.org/10.1016/j.ifset.2016.08.003

Rivera-Reyes, M., Campbell, J. A., & Cutter, C. N. (2019). Survival of acid-adapted and non-adapted Shiga toxin-producing *Escherichia coli* using an *in vitro* model. *Food Control*, *104*, 28-33. https://doi.org/10.1016/j.foodcont.2019.04.009

Rodriguez-Gonzalez, O., Buckow, R., Koutchma, T., & Balasubramaniam, V. M. (2015). Energy Requirements for Alternative Food Processing Technologies—Principles, Assumptions, and Evaluation of Efficiency. *Comprehensive Reviews in Food Science and Food Safety*, *14*(5), 536-554. https://doi.org/10.1111/1541-4337.12142

Roobab, U., Abida, A., Chacha, J. S., Athar, A., Madni, G. M., Ranjha, M. M. A. N., Rusu, A. V., Zeng, X.-A., Aadil, R. M., & Trif, M. (2022). Applications of innovative non-thermal pulsed electric field technology in developing safer and healthier fruit juices. *Molecules*, *27*(13), Article 13. https://doi.org/10.3390/molecules27134031

Salehi, F. (2020). Physico-chemical properties of fruit and vegetable juices as affected by pulsed electric field: A review. *International Journal of Food Properties*, *23*, 1036-1050. https://doi.org/10.1080/10942912.2020.1775250

Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R., Morabito, S., Strockbine, N. A., Melton-Celsa, A. R., Sanchez, M., Persson, S., & O'Brien, A. D. (2012). Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *Journal of Clinical Microbiology*, *50*(9), 2951-2963. https://doi.org/10.1128/JCM.00860-12

Sehrawat, R., Kaur, B. P., Nema, P. K., Tewari, S., & Kumar, L. (2020). Microbial inactivation by high pressure processing: Principle, mechanism and factors responsible. *Food Science and Biotechnology*, *30*(1), 19-35. https://doi.org/10.1007/s10068-020-00831-6

Škegro, M., Putnik, P., Bursać Kovačević, D., Kovač, A. P., Salkić, L., Čanak, I., Frece, J., Zavadlav, S., & Ježek, D. (2021). Chemometric comparison of high-pressure processing and thermal pasteurisation: The nutritive, sensory, and microbial quality of smoothies. *Foods (Basel, Switzerland)*, *10*(6), 1167. https://doi.org/10.3390/foods10061167

Strachan, N. J. C., Doyle, M. P., Kasuga, F., Rotariu, O., & Ogden, I. D. (2005). Dose response modelling of *Escherichia coli* O157 incorporating data from foodborne and environmental outbreaks. *International Journal of Food Microbiology*, *103*(1), 35-47. https://doi.org/10.1016/j.ijfoodmicro.2004.11.023

Syed, Q.-A., Buffa ,Martin, Guamis ,Buenaventura, & and Saldo, J. (2016). Factors Affecting Bacterial Inactivation during High Hydrostatic Pressure Processing of Foods: A Review. *Critical Reviews in Food Science and Nutrition*, *56*(3), 474-483. https://doi.org/10.1080/10408398.2013.779570

Teunis, P. F. M., & Havelaar, A. H. (2000). The Beta Poisson dose-response model is not a single-hit model. *Risk Analysis*, *20*(4), 513-520. https://doi.org/10.1111/0272-4332.204048

Topalcengiz, Z., & Danyluk, M. D. (2017). Thermal inactivation responses of acid adapted and non-adapted stationary phase Shiga toxin-producing Escherichia coli (STEC), Salmonella spp. And Listeria monocytogenes in orange juice. *Food Control*, *72*, 73-82. https://doi.org/10.1016/j.foodcont.2016.07.014

van Asselt, E. D., & Zwietering, M. H. (2006). A systematic approach to determine global thermal inactivation parameters for various food pathogens. *International Journal of Food Microbiology*, *107*(1), 73-82. https://doi.org/10.1016/j.ijfoodmicro.2005.08.014

van Boekel, M. A. J. S. (2020). On the pros and cons of Bayesian kinetic modeling in food science. *Trends in Food Science & Technology*, *99*, 181-193. https://doi.org/10.1016/j.tifs.2020.02.027

van der Vossen-Wijmenga, W. P., den Besten, H. M. W., & Zwietering, M. H. (2024). Temperature status of domestic refrigerators and its effect on the risk of listeriosis from ready-to-eat (RTE) cooked meat products. *International Journal of Food Microbiology*, *413*, 110516. https://doi.org/10.1016/j.ijfoodmicro.2023.110516

Walkling-Ribeiro, M., Rodríguez-González, O., Jayaram, S., & Griffiths, M. W. (2011). Microbial inactivation and shelf life comparison of 'cold' hurdle processing with pulsed electric fields and microfiltration, and conventional thermal pasteurisation in skim milk. *International Journal of Food Microbiology*, *144*(3), 379-386. https://doi.org/10.1016/j.ijfoodmicro.2010.10.023

Walton, J., & Kehoe, L. (2024). Current perspectives and challenges in the estimation of fruit juice consumption across the lifecycle in Europe. *Nutrition Research Reviews*, 1-12. https://doi.org/10.1017/S095442242400009X

Zwietering, M. H. (2009). Quantitative risk assessment: Is more complex always better?: Simple is not stupid and complex is not always more correct. *International Journal of Food Microbiology*, *134*(1), 57-62. https://doi.org/10.1016/j.ijfoodmicro.2008.12.025

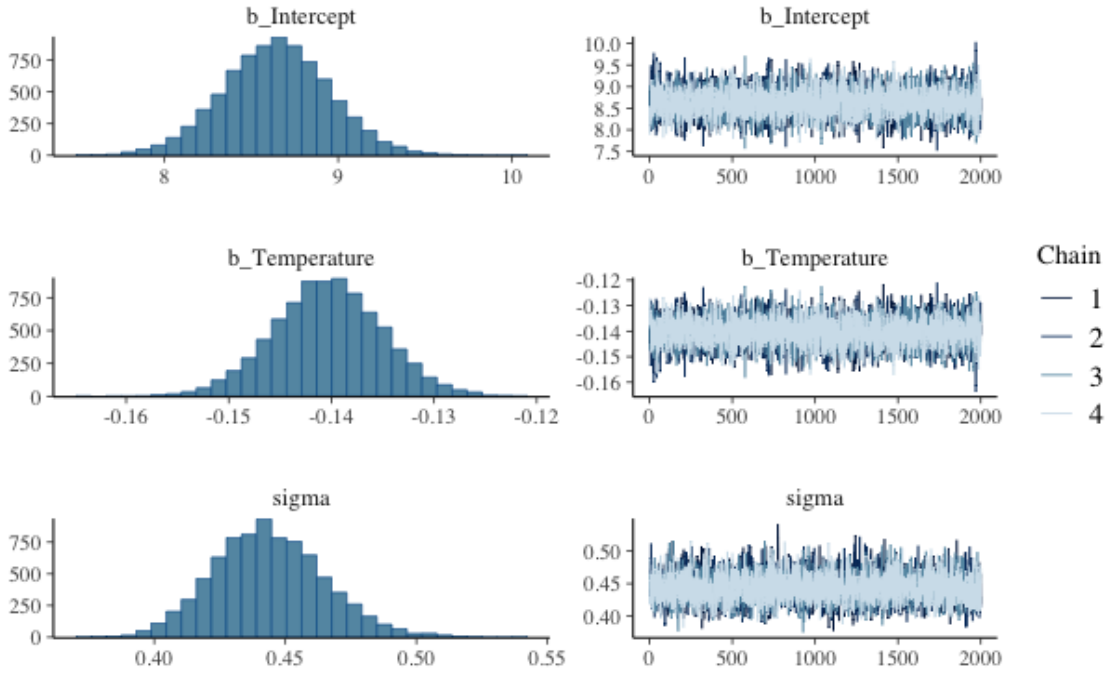
Zwietering, M. H. (2015). Risk assessment and risk management for safe foods: Assessment needs inclusion of variability and uncertainty, management needs discrete decisions. *International Journal of Food Microbiology*, *213*, 118-123. https://doi.org/10.1016/j.ijfoodmicro.2015.03.032

Zwietering, M. H., Garre, A., & den Besten, H. M. W. (2021). Incorporating strain variability in the design of heat treatments: A stochastic approach and a kinetic approach. *Food Research International*, *139*, 109973. https://doi.org/10.1016/j.foodres.2020.109973

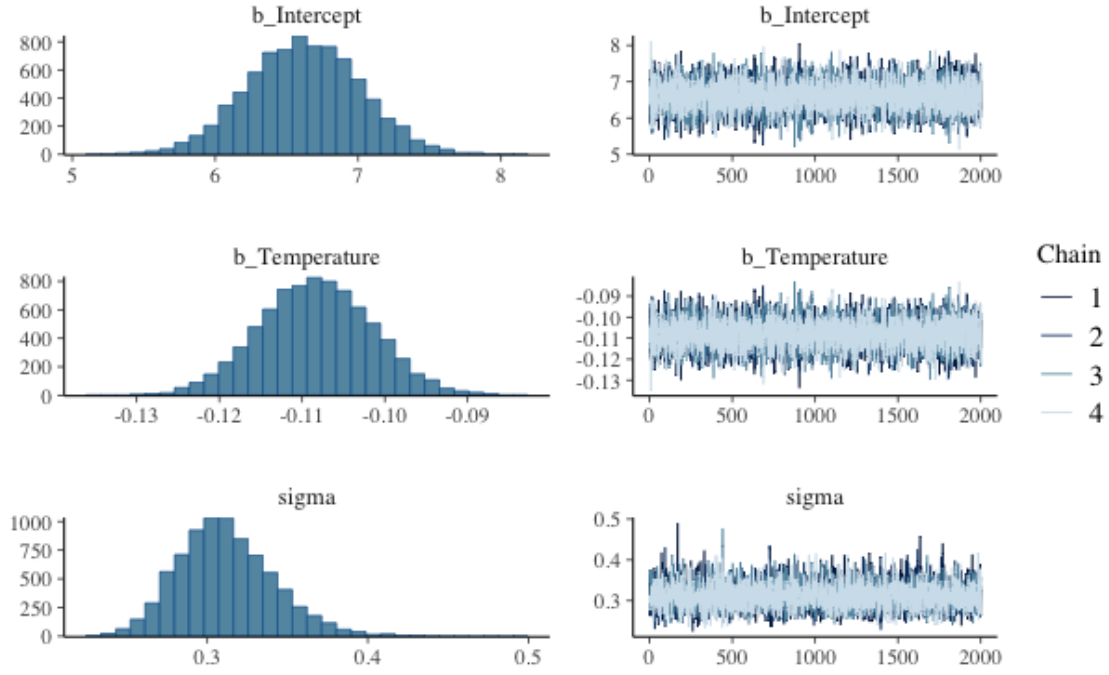
Zwietering, M. H., & Havelaar, A. H. (2006). Dose-response relationships and foodborne disease. In *Food consumption and disease risk* (pp. 422-439). Woodhead Publishing Limited.

Appendices

Appendix 1, Posterior distributions for thermal processing

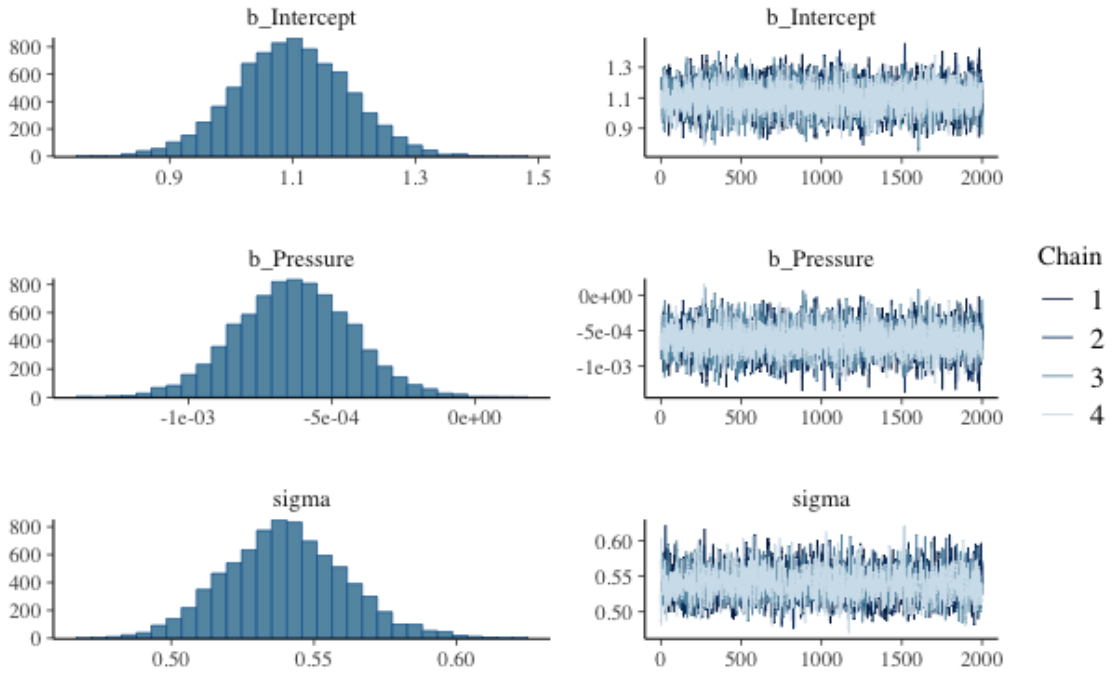


Appendix 1, Posteriors distributions (log_{10} D ~Temperature): All liquid products

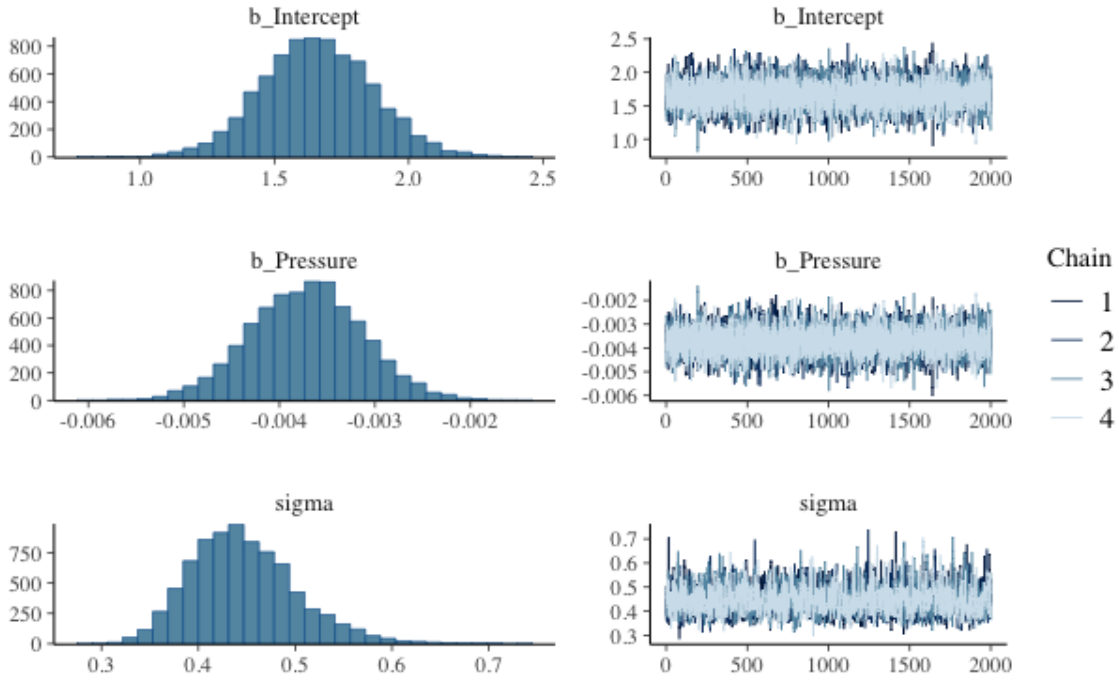


Appendix 2, Posteriors distributions (log_{10} D ~Temperature): High acidic fruit juices

Appendix 2, Posterior distributions for High Pressure Processing

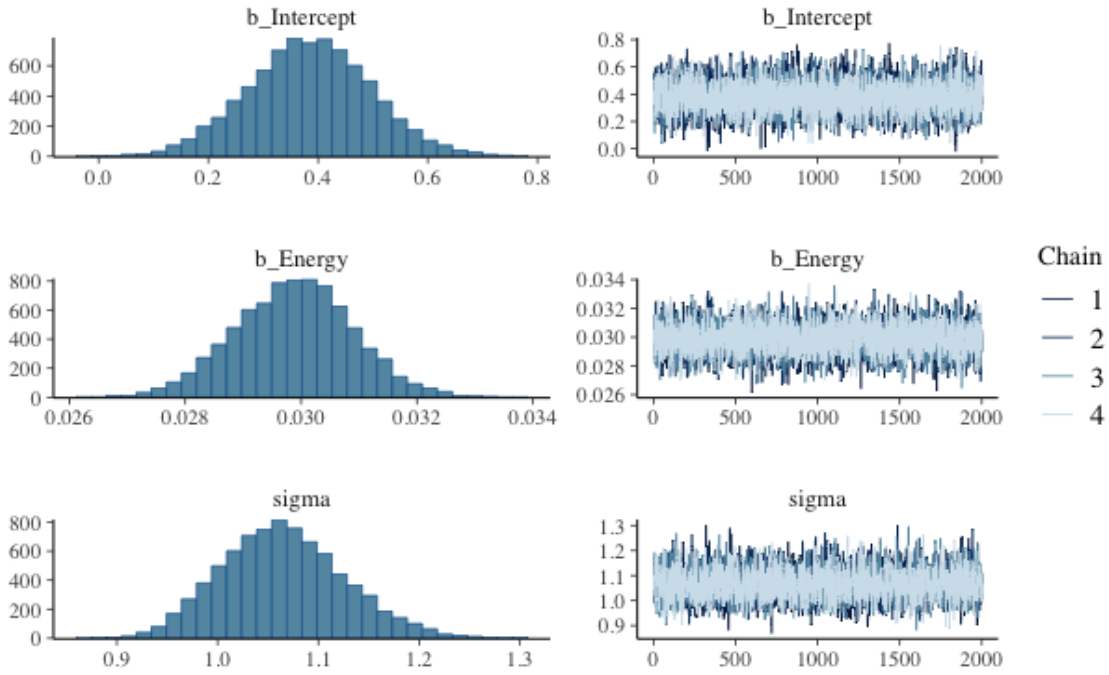


Appendix 3, Posteriors distributions (log_{10} D ~Pressure): All liquid products

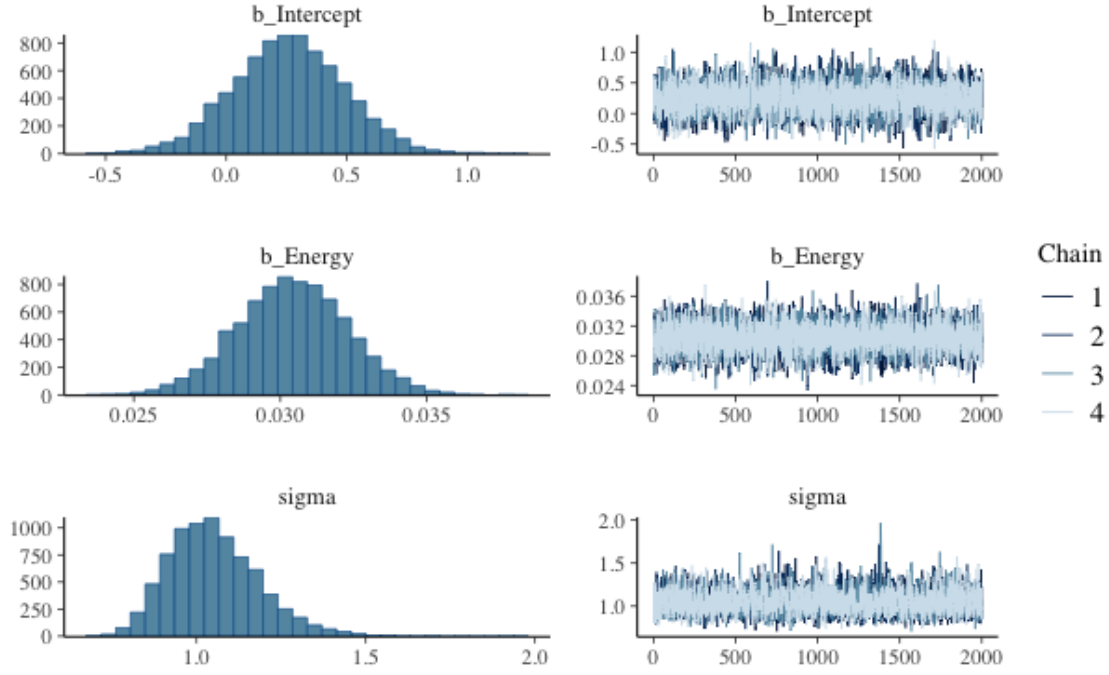


Appendix 4, Posteriors distributions (log_{10} D ~Pressure): High acidic fruit juices

Appendix 3, Posterior distributions for Pulsed Electric field Technology



Appendix 7, Posteriors distributions (log reduction~energy input): All liquid products



Appendix 9, Posteriors distributions (log reduction ~ energy input): High acidic fruit juices)

Author:

Rick Mannes | 1298917
Wageningen University & Research -
Laboratory of Food Microbiology
MSc. Food Technology

Supervisors:

Georgios Pampoukis
Prof. dr. ir. Heidy den Besten

Examiner:

Prof. dr. ir. Marcel Zwietering

WAGENINGEN
UNIVERSITY & RESEARCH