# *Bayesian Inference and Failure Analysis for Risk Assessment in Quality Engineering*

Doctoral Thesis

**Yacine Koucha**

Supervisors: Dr QingPing Yang & Prof. Alistair Forbes

*In Partial Fulfillment of the Requirements for the Doctor of Philosophy in Mechanical Engineering*

College of Engineering, Design and Physical Sciences

Brunel University London

United Kingdom

# Abstract

Failure is the state of not achieving a desired or intended goal. Failure analysis planning in the context of risk assessment is an approach that helps to reduce total cost, increase production capacity, and produce higher-quality products. One of the most common issues that businesses confront are defective products. This issue not only results in monetary loss, but also in a loss of status. Companies must improve their production quality and reduce the quantity of faulty products in order to continue operating in a healthy and profitable manner in today's very competitive environment. On the other hand, there is the ongoing COVID-19 pandemic, which has thrown the world's natural order into disarray, and has been designated a Public Health Emergency of International Concern by the World Health Organization. The demand for quality control is rapidly increasing. Failure analysis is thus an useful tool for identifying common failures, their likely causes, and their impact on the health system, as well as plotting strategies to limit COVID-19 transmission. It is now more vital than ever to enhance failure analysis methods.

The traditional FMEA (Failure mode and effects analysis) is one of the most widely used approaches for identifying and classifying failure modes (FMs) and failure causes (FCs). It is a risk analysis tool for coping with possible failures and is widely used in the reliability engineering, safety engineering and quality engineering. To prioritize risks of different failure modes, FMEA uses the risk priority number (RPN), which is the product of three risk measures: severity ($S$), occurrence ($O$) and detection ($D$). Traditional FMEA, on the other hand, has drawbacks, such as the inability to cope with uncertain failure data, such as expert subjective evaluations, the failure events' conditionality, RPN has a high degree of subjectivity, comparing various RPNs is challenging, potential errors may be ignored in the conventional FMEA process, etc. To overcome these limitations, I present an integrated Bayesian

approach to FMEA in this thesis.

In this proposed approach, I worked with experts in quality engineering and used Bayesian inference to estimate the FMEA risk parameters: $S$, $O$ and $D$. The proposed approach is intended to become more practical and less subjective as more data is added. Bayesian statistics is a statistical theory that is based on the Bayesian interpretation of probability, which states that probability expresses a degree of belief or information (knowledge) about an event. Bayesian statistics addresses the issues with uncertainties found in frequentist statistics, such as the distribution of contributing factors, the implications of using specific distributions and specifies that there is some prior probability. A prior can be derived from previous information, such as previous experiments, but it can also be derived from a trained subject-matter expert's purely subjective assessment. Frequentist statistics (also known as classical statistics) has several limitations, including a lack of uncertainty information in predictions, no built-in regularisation, and no consideration of prior knowledge. Due to the availability of powerful computers and new algorithms, Bayesian methods have seen increased use within statistics in the twenty-first century, and this thesis highlights the effective use of Bayesian analyses to address the shortcomings of the current FMEA with the revamped Bayesian FMEA. As a demonstration of the approach, three case studies are presented.

The first case study is a Bayesian risk assessment approach of the modified SEIR (susceptible-exposed-infectious-recovered) model for the transmission dynamics of COVID-19 with an exponentially distributed. The effective reproduction number is estimated based on laboratory-confirmed cases and death data using Bayesian inference and analyse the impact of the community spread of COVID-19 across the United Kingdom. The value of effective reproduction number models the average number of infections caused by a case of an infectious disease in a population that includes not only susceptible people. The FMEA is then applied to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic. In the FMEA, the focus was on COVID-19 infections and therefore the failure mode is taken as positive cases. The model is applied to COVID-19 data showing the effectiveness of interventions adopted to control the epidemic by reducing the effective reproduction number of COVID-19. The risk measures were estimated from the case

fatality rate ($S$), the posterior median of the effective reproduction number ($O$) and the current corrective measures used by government policies ($D$).

The second case study is a Bayesian risk assessment of a coordinate measuring machine (CMM) process using failure mode, effects and criticality analysis (FMECA) and an augmented form error model. The form error is defined as the deviation of a manufactured part from its design or ideal shape, and it is a key characteristic to evaluate in quality engineering and manufacturing. The form error is presented as a probabilistic model using symmetric unimodal distributions. Bayesian inference is then used to identify influence factors associated with the measurement process due to form error, environmental, human and random effects. A risk assessment is then performed by combining Bayesian inference, FMECA and conformity testing, to quantify and minimise the risk of wrong decisions. In the FMECA, the focus was on CMM measurement process and I identified four major FMs that can occur: probe, mechanical, environmental and measurement performance failure. Eleven FCs were also observed, each of which was linked to one of the four FMs. The risk measures were estimated from the posterior probability of failure causes associated with the CMM measurement process ($O$), the severity of a specific consumer's risk ($S$) and the detectability of failures from the posterior standard deviation of the form error model ($D$).

The third case study is a Bayesian risk assessment of a CMM measurement process using an autoregressive (AR) form error model and a combined Fault tree analysis (FTA) and FMEA approach to predict significant failure modes and causes. The main idea is to estimate and predict the form error based on CMM data using Gibbs sampling and analyse the impact of the CMM measurement process on product conformity testing. The FTA is used to compare the actual and predicted form error data from the Bayesian AR plot to determine the likelihood of the CMM measurement process failing using binary data. The acquired binary data is then classified into four states (true positive, true negative, false positive, and false negative) using a confusion matrix, which is subsequently utilized to calculate key classification measures (i.e., error rate, prediction rate, prevalence rate, sensitivity rate, etc). The classification measures were then used to assess the FMEA risk measures $S$, $O$, and $D$, which were critical for determining the RPN and making decisions.

Analytical and numerical methods are used in all case studies to highlight the practical implications of our findings and are meant to be practical without complex computing. The proposed methodologies can find applications in numerous disciplines and wide quality engineering.

***Keywords***— Bayesian inference, failure analysis, quality engineering, failure mode and effects analysis

# Acknowledgements

I would like to express my deepest appreciation to all those who provided me the possibility to complete this thesis. A big thank you goes to my friends and family, including my wife, who helped me to put together the parts and suggested many elements of this project. I would also like to thank all members involved in the joint collaboration project between Brunel University London and National Physical Laboratory for the support and guidance they gave me. I am also especially grateful to Brunel University London for funding my PhD studies for the entire three-year period. A special thank you to my thesis supervisors, Dr. QingPing Yang and Professor Alistair Forbes, whose contributions to stimulating suggestions and encouragement helped me to coordinate my project, especially in forming and writing this thesis. They have both constantly helped me throughout my PhD studies and have greatly motivated me into research work. I am grateful to have been guided by some of the world's best supervisors! Last but not least, thank you very much to the Almighty, Master of the Universe, for bringing me this far into my life.

# Contents

# List of Figures

# List of Tables

# Acronyms

$\mathcal{R}_t$ effective reproduction number

$D$ detection

$FN$ false negative

$FP$ false positive

$O$ occurrence

$S$ severity

$TN$ true negative

$TP$ true positive

**ACF** autocorrelation function

**AR** autoregressive

**ARIMA** autoregressive integrated moving average

**BN** Bayesian network

**BT** Bowtie

**CFR** case fatality rate

**CMM** coordinate measuring machine

**COVID-19** coronavirus disease

**DRBFM** Design Review by Failure Mode

**ETA** Event Tree Analysis

**EU** European Union

**FC** failure cause

**FM** failure mode

**FMEA** Failure mode and effects analysis

**FMECA** Failure mode, effects, and criticality analysis

**FTA** Fault tree analysis

**GUM** Guide to the Expression of Uncertainty in Measurement

**HA** Hazard Analysis

**ICU** intensive care unit

**IG** inverse gamma

**JCGM** Joint Committee for Guides in Metrology

**LS** least squares

**MA** moving average

**MAP** maximum a posteriori probability

**MAR** median autoregressive

**MCMC** Markov chain Monte Carlo

**ML** maximum likelihood

**MZ** minimum zone

**NGM** next-generation matrix

**NHS** National Health Service

**PACF** partial autocorrelation function

**QAR** quantile autoregressive

**RPN** risk priority number

**SARS-CoV-2** severe acute respiratory syndrome coronavirus 2

**SEIQR** susceptible-exposed-infectious-quarantined-recovered

**SEIR** susceptible-exposed-infectious-recovered

**SIR** susceptible-infectious-recovered

**SPC** statistical process control

**UK** United Kingdom

**WHO** World Health Organization

# Chapter 1

# Introduction

## 1.1 Overview and research objectives

Failure is the state of not achieving a desired or intended goal. This term is defined in manufacturing environments as a component that damages engineering equipment, manufactured products, affecting operation, production, and performance, as well as the company's reputation [1, 2]. Failure analysis planning in the context of risk assessment is an approach that helps to reduce total cost, increase production capacity, and produce higher-quality products. One of the most common issues that businesses confront are defective products. This issue not only results in monetary loss, but also in a loss of status. Companies must improve their production quality and reduce the quantity of faulty products in order to continue operating in a healthy and profitable manner in today's very competitive environment [3–5]. On the other hand, there is the ongoing COVID-19 pandemic, which has thrown the world's natural order into disarray, and has been designated a Public Health Emergency of International Concern by the World Health Organization (WHO) [6]. The demand for quality control is rapidly increasing. Failure analysis is thus an useful tool for identifying common failures, their likely causes, and their impact on the health system, as well as plotting strategies to limit COVID-19 transmission. It is now more vital than ever to enhance failure analysis methods.

Some systematic procedures are required to evaluate failures and comment on their likely implications, such as: the Bowtie (BT), Design Review by Failure Mode (DRBFM), Hazard Analysis (HA), Fault tree analysis (FTA) and its extension Event Tree Analysis (ETA), What-if/Checklist, and Failure mode and effects analysis (FMEA) [7]. FMEA is one of the most widely used approaches for identifying and classifying failure modes (FMs)

and failure causes (FCs). FMEA was first used by reliability engineers in the late 1950s to analyse problems that might arise from malfunctioning of military systems [8]. As a quantitative analysis, a systematic approach to the use of FMEA consists of identifying ways in which FMs can occur, then identifying FCs, the effects of each FM, evaluating the corresponding risks, and then taking the necessary steps to minimise the risks and consequences. By including a criticality analysis, the FMEA can be extended to an FMECA (Failure mode, effects, and criticality analysis). The strictly qualitative FMEA can be made more quantitative in this way. The effects of FMs on system/process performance can be assessed via an FMEA analysis, allowing essential corrective measures to be taken [9]. In the FMEA/FMECA study, the risk priority number (RPN) is employed as the prioritizing index. Product failure severity ($S$), likelihood of failure occurrence ($O$), and probability of failure detection ($D$) are the three risk factors utilized in the calculation of RPN. $S$, $O$, and $D$ parameters are multiplied to get the RPN [10]. Traditional FMEA/FMECA has drawbacks, such as the inability to cope with uncertainty, subjective expert assessments, failure event conditionality, and statistical inference of each risk factor, to name a few. As a result, I have extended this to include Bayesian inference of risk parameters, Bayesian network to conditionally seek out primary FCs, and prior/posterior updating (e.g., to reflect actions the government could have taken to better manage the COVID-19 pandemic).

Bayesian inference is a valuable technique that addresses the uncertainties found in frequentist statistics, such as the distribution of contributing factors, the implications of using specific distributions, and the specification of some prior probability. A prior can be derived from previous information, such as previous experiments, but it can also be derived from the purely subjective assessment of a trained subject-matter expert. Bayesian inference is widely used in a variety of fields, including: machine learning, quality engineering, smart systems and more [11–13]. Bayesian inference begins from two indicators: the *prior distribution* and the *likelihood function* derived from the observed data. The prior probability is a central aspect of Bayesian inference and reflects knowledge on unknown parameters, which is combined with the likelihood of new data. The *posterior distribution* is the result of combining these two indicators. In recent years, Bayesian methods have generated a lot of attention because they can combine expert knowledge with experimental data easily while taking into account uncertainty [14–16]. Bayesian inference has become increasingly essential for the development of more realistic models based on the subjective properties of posterior distribution for complex phenomena and multi-parameter systems or processes, due to the rapid advances in computer technology and demand in many areas of science

and engineering. A deterministic model for the measurement process is not viable due to the uncertainty of model parameters (such as temperature changes in measured items) and the complexity of measurement systems in quality engineering applications. Because of this, a probabilistic model must be derived and it is essential to use a Bayesian approach to calculate the posterior distribution of the unknown parameters and relate them to the FMEA/FMECA risk factors $S$, $O$ and $D$ to make decisions.

However, critical problems arise in the research of advanced computing methods and tools for failure analysis, for example:

1. What are the best ways to accurately determine FMs/FCs in complex manufacturing, measurement systems, or epidemics (among other things)?

2. Are there any ways to improve risk assessment methods by using Bayesian inference together with failure analysis methods?

3. What are the options for assessing the effectiveness of the corrective measures taken to minimize the risk and to improve product quality or process efficiency?

To answer these questions, a Bayesian risk assessment was proposed using FMEA/FMECA that monitors the significant failure modes/causes, the relevant failure effects using a rigorous risk assessment, which accurately identifies steps leading to corrective actions. In this proposed approach, posterior estimates from Bayesian inference were used to characterize the unknown parameters from a given model to estimate the FMEA/FMECA risk factors. To demonstrate the new approach, three case studies were considered in this thesis:

- a Bayesian risk assessment of the COVID-19 pandemic using a modified Susceptible-Exposed-Infectious-Removed (SEIR) epidemic model and FMEA (subsection 1.4.1),

- a Bayesian risk assessment of the coordinate measuring machine (CMM) measurement process using an augmented form error model and FMECA (subsection 1.4.2),

- a Bayesian approach to failure prediction using an autoregressive form error model, Gibbs sampling and an integrated FTA/FMEA (subsection 1.4.3).

These case studies are intended to show the effectiveness of combining Bayesian statistics with improved failure analysis methods.

## 1.2 Contributions

The contributions may be condensed as follows in light of the above-mentioned comments on the thesis's goals:

- Bayesian inference, FMEA and a modified SEIR model has been utilised to study the risks of COVID-19 infections and to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic by combining together. The risk measures, $S$ and $O$ parameters, were evaluated using the CFR rate and posterior median of the effective reproduction number, respectively. Parameter $D$ was estimated using the current detection measures in place.

- Bayesian inference has been used to summarize the process phase, FCs, current measures, and risk measures, and then gave the subsequent government actions and re-evaluated risk measures using 8 months' worth of daily COVID-19 data. Thus, I have assessed the impact of English government protective measures to COVID-19 by comparing the empirical and posterior statistics of the effective reproduction number at different time periods with and without the measures being implemented. The findings are in line with empirical results.

- By combining Bayesian inference, FMEA, and a compartmental model, the strategy presented here can be applied to future pandemics which best simulates how individuals in different compartments in a population interact.

- Using an integrated Bayesian FMECA approach and an augmented form error model, the risks were investigated and significant observable FMs/FCs of the CMM measurement process has been identified in order to effectively assess the corrective/preventive measures in place. Additional sources of uncertainty (e.g., random, environmental, human, etc.) were implemented as additive factors, resulting in more reliable product conformity rates that matched a Pareto-like distribution.

- The FMECA risk parameters, $S$ from the severity of the specific consumer's risk, $O$ from the posterior probability of failures associated with the CMM measurement process under a Bayesian network, and $D$ from the detectability of the posterior standard deviation of the proposed form error model, were estimated using the integrated Bayesian approach.

- A Bayesian risk assessment that incorporates a Markov chain Monte Carlo (MCMC) algorithm (i.e. Gibbs sampling), an autoregressive form error model, and an integrated

FTA/FMEA has been shown to be particularly effective in predicting failure and making decisions.

- The risk measures $S$, $O$, and $D$, which were critical for determining the RPN and making decisions, were assessed using classification measures derived from a combination of the integrated FTA/FMEA and a confusion matrix.

## 1.3 Publications and presentations

### 1.3.1 Publications

I published two substantial articles as a main author and delivered a virtual conference at the IMEKO World Congress in Japan, both of which received great feedback. A large proportion of chapter 2 was submitted as a manuscript to the *International Journal of Metrology and Quality Engineering (IJMQE)* in December 2020. It was accepted by the journal to strong positive feedback and very minor revisions. Similarly, parts of chapter 3 was submitted as a conference paper to *IMEKO XXIII World congress* and eventually updated as a manuscript to the *Measurement: Sensor* journal. It was accepted in June 2021 and published in December 2021.

I also collaborated on another colleague's article, "UML knowledge model for measuring process including measurement uncertainty," in which I concentrated on the third section of her paper: "Case study: Gauge block length measurement using a CMM." In this study, I was involved in investigating and calculating distinct uncertainty components into two types of CMM evaluations: I Type A (based on a series of observations using a t-distribution); (ii) Type B (based on a series of observations using a t-distribution) (evaluated using available information). The standard, combined, and enlarged uncertainty were all evaluated. I also devised the uncertainty budget and provided the result to the expression of uncertainty.

A full list of publications is found below:

- Koucha, Y., & Yang, Q. (2021). A Bayesian risk assessment of the COVID-19 pandemic using FMEA and a modified SEIR epidemic model. International Journal of Metrology and Quality Engineering, 12, 14. doi:10.1051/ijmqe/2021012

- Koucha, Y., Forbes, A., & Yang, Q. (2021). A Bayesian conformity and risk assessment adapted to a form error model. Measurement: Sensors, 18, 100330. doi:10.1016/j.measen.2021.100330

- Bharti, P., Yang, Q., Forbes, A. B., & Koucha, Y. (2021). UML knowledge model for measurement process including uncertainty of measurement. International Journal of Metrology and Quality Engineering, 12, 26. doi:10.1051/ijmqe/2021024

### 1.3.2 Presentations

We had a number of weekly meetings during the PhD, and several of them required me to present a portion of my "work to date." I also gave a number of talks at the NPL and presented my findings at a number of university-sponsored events. Since the COVID-19 pandemic broke out, I've been doing all of my presentations remotely, and I've stopped keeping track of the dates because I've been doing them every week. Table 1.1 is a list of the presentations I gave prior to the COVID-19 pandemic's outbreak.

| Date | Presentations delivered | Location |
|---|---|---|
| 05/12/2018 | Bayesian Inference on Quality Engineering | Brunel University London |
| 08/01/2019 | Review of Bayesian computational methods | Brunel University London |
| 28/01/2019 | Application of Bayes Theorem on time-series models | Brunel University London |
| 18/02/2019 | Bayesian Inference on Quality Engineering | Brunel University London |
| 18/02/2019 | BDA & MCMC implementation ideas | Brunel University London |
| 28/02/2019 | Conformance assessment methods | National Physical Laboratory |
| 13/03/2019 | Bayesian decision theory to minimise posterior loss | Brunel University London |
| 07/05/2019 | A Bayesian SPC for decision-making | Brunel University London |
| 22/05/2019 | Bayesian conformance assessent: Circularity | National Physical Laboratory |
| 12/10/2019 | Bayesian computational methods | Brunel University London |
| 14/11/2019 | Bayesian smart building and metrology examples | Brunel University London |
| 02/12/2019 | Task 6 Project Update | Brunel University London |
| 23/01/2020 | Bayesian computing methods and tools for quality engineering and smart systems | Brunel University London |
| 17/02/2020 | Bayesian methods in manufacturing metrology for quality engineering | Brunel University London |
| 12/03/2020 | Copula in conformance assessment | National Physical Laboratory |

**Table 1.1:** Presentations delivered throughout the course of this PhD.

## 1.4 The three cases studies

### 1.4.1 A Bayesian risk assessment of the COVID-19 pandemic using a modified SEIR epidemic model

In the first case studying, a Bayesian risk assessment approach of the modified SEIR model for the transmission dynamics of COVID-19 with an exponentially distributed infectious period was considered.

The main idea was to estimate the effective reproduction number based on laboratory-confirmed cases and death data using Bayesian inference and analyse the impact of the community spread of COVID-19 across the United Kingdom. To achieve this objective, an improved FMEA tool was used to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic. In the FMEA, the focus was on COVID-19 infections and therefore the FMs were taken to be positive cases.

The model was applied to COVID-19 data showing the effectiveness of interventions adopted to control the epidemic by reducing the effective reproduction number of COVID-19. A summary of the first case study is presented in Figure 2.1 (see chapter 2).

### 1.4.2 A Bayesian risk assessment of the CMM measurement process using an augmented form error model

In the second case study, a Bayesian risk assessment of a coordinate measuring machine (CMM) measurement process using an augmented form error model and an integrated FMECA approach was considered. The form error is defined as the deviation of a manufactured part from its design or ideal shape, and it is a key characteristic to evaluate in quality engineering and manufacturing [17–19]. A CMM was used to obtain measurement data and the form error was then computed. The proposed model was expressed as an additive model dependent on a number of factors which allows for a measurement system analysis approach based on analysis of variance [20–23]. The factors include: (I) uniformly distributed form error estimates; (II) random errors (associated with the measurement system); (III) human errors (operational, etc); (IV) environmental errors (temperature, humidity, dust, etc). These factors were then combined to generate a posterior distribution using Bayesian inference.

The main idea was to estimate the form error based on CMM data for product conformity testing using Bayesian inference and analyse the impact of the CMM measurement process. To achieve this goal, the FMECA was implemented to assess observable failures that may occur during the CMM measurement process. I've identified four significant FMs that can arise throughout the CMM measurement process based on my observations: probe failure (FM1), measurement performance failure (FM2), mechanical failure (FM3) and environmental failure (FM4). Eleven FCs were also observed, each of which was linked to one of the four FMs. The FCs were item displaced slightly or greatly (FC1/2), dust on time (FC3), mechanical stress (FC4), wrong probe type used (FC5), valve not set to required pressure level (FC6), temperature set too high or too low (FC7/8), CMM making noise

(FC9), probe not fitted properly (FC10) or probe not calibrated (FC11). I chose the names of these FMs in order to classify the FCs into the appropriate group and build a hypothesis.

The model was applied to CMM form error data (using a touch-trigger probe) to demonstrate the impact of corrective/preventive measures taken to monitor and minimize measurement failures. A summary of the second case study is presented in Figure 3.3 (see chapter 3).

### 1.4.3 A Bayesian risk assessment of the CMM measurement process using an autoregressive form error model and Gibbs sampling

In the third case study, a Bayesian risk assessment of a CMM measurement process using an autoregressive (AR) form error model and a combined FTA/FMEA approach to predict significant failures was considered. The AR form error model, which is the time-series equivalent of the classic form error model, allowed for the forecasting of form errors. An important parameter used is the AR lag parameter which specifies how many measured points should be used at an instant.

The main idea was to estimate and predict the form error based on CMM data using Gibbs sampling (i.e., an example of MCMC that allows one to estimate the joint marginal distribution using draws from the conditional distribution) and analyse the impact of the CMM measurement process on product conformity testing. A combined FTA/FMEA approach to the CMM to predict significant observable failures was considered to achieve this goal. The FTA was used to compare the actual and predicted form error data from the Bayesian AR plot to determine the likelihood of the CMM measurement process failing using binary data. The acquired binary data was then classified into four states (true positive, true negative, false positive, and false negative) using a confusion matrix, which is subsequently utilized to calculate key classification measures (i.e., error rate, prediction rate, prevalence rate, sensitivity rate, etc). The classification measures were then used to assess the FMEA risk measures $S$, $O$, and $D$, which were critical for determining the RPN and making decisions.

The model was applied to CMM form error data (using an automatic scanning probe) to demonstrate the effectiveness of failure prediction and the impact of corrective/preventive measures taken to monitor and minimize measurement failures. This case study supplemented and extended the ideas and methods described in the first and second case studies, respectively, by: (I) extending the time-series approach of updating daily data from the

first case study; (II) transforming the augmented form error model from the second case study into a time-series equivalent (i.e., an autoregressive form error model). As a result, the FMs/FCs were identical to those in the second case study. The algorithms and coding was completed on *RStudio*.

## 1.5 Background and literature review

In this section, a background and literature review of Bayesian methods and failure analysis techniques is presented. In the first part of this chapter, I review relevant Bayesian methods used in this thesis (subsection 1.5.1—1.5.7). In the second part of this chapter, I review current failure analysis methods for quality engineering (subsection 1.5.8—1.5.10). Both parts will review the relevant literature and methods necessary for all three case studies.

### 1.5.1 Bayesian history and analysis

In numerous journals published across the late 18th century to the early 19th century, Pierre-Simon Laplace (1749–1827), a French scholar pivotal to the development and fundamentals of engineering, mathematics, statistics, physics and astronomy, worked ambitiously on the Bayesian interpretation of probability for a brief timeline of events spanning the advent of Bayesian statistics to modern applications of Bayesian inference. Interestingly, the methodology used by Laplace to solve statistical problems is said to be analogous to the Bayesian approach of solving problems, but was not coined as such, that is, until the 1950s. Philosophical, practical and computational considerations were some of the biggest drawbacks surrounding the use of Bayesian methods among statisticians during the latter half of twentieth century. Bayesian methods simply required a lot of calculations to actually solve problems, which paled against the frequentist rendition.

In recent years, Bayesian methods have generated a lot of attention because they can combine expert knowledge with experimental data easily while taking into account uncertainty [15, 16]. Bayesian inference has become increasingly essential for the development of more realistic models based on the subjective properties of posterior distribution for complex phenomena and multi-parameter systems or processes, due to the rapid advances in computer technology and demand in many areas of science and engineering. A deterministic model for the measurement process is not viable due to the uncertainty of model parameters (such as temperature changes, measurement errors, sampling errors, etc) and the complexity of measurement systems in quality engineering applications. Because of this, a probabilistic model must be derived and it is essential to use a Bayesian approach to calculate the

posterior distribution of unknown parameters.

However, with the rise of supercomputers in the contemporary era of Big Data [24] and new Bayesian techniques[1] as well as newer practices such as objective Bayesian analysis for high-dimensional models and deep learning neural networks in recent papers [25–27]; Bayesian statistics have seen far more use coming into the early-mid 21st century.

I conducted some data analysis on the number of times Bayesian terms appear in a search query (using Brunel Library's Summon portal with results that extend beyond Brunel Library's collection). In Table 1.2a, the number of published content refers to (but not limited to): journal articles, publications, conference proceedings, book reviews (usually a short review of the original work), books, eBooks and more (e.g., video recordings, posters, student thesis, etc). The number of published items where "Bayesian" content was searched continues to increase over each period (excluding the cumulative period up to 1985); it has increased steadily over each period since 1985. From 1985-89 (i.e., 1st January 1985, to 31st December 1989), there were 14,192 (0.024%) "Bayesian" related publications; this increased to 377,322 (0.080%) in 2010-14. Most recently, in 2015-18; this percentage reached at an all time high of 0.094%. In Table 1.2b, the number and percentage of search queries were compared with additional "Bayesian" related terms as a fraction of the number of content published with the term "Bayesian" only. The results indicates that Bayesian data analysis (BDA) is more prevalent in searches than all other items in each period. Interestingly, BDA has remained constant at two-thirds of all Bayesian related searches throughout each period. The percentage of published content with Bayesian inference (BI), Bayesian computation (BC) and Bayesian decision theory (BDT) has each decreased significantly since 1985-89, from 42.7/35.6/60.4 percentage points to 31.1/17.5/15.3 in 2015-18. The most notable decrease came from BDT searches, which decreased by almost four times. However, the percentage of published content for Bayesian Markov chain Monte Carlo (BMCMC) has generally increased, both in terms of raw numbers and percentages, from 8 publications (0.1%) in 1985-89 to 34,787 (9.2%) in 2010-14; then fell to 29,971 (8.4%). The percentage of published content with Bayesian statistical process control (BSPC) has remained low, with minuscule increases from 1.0% before 1985 to 1.7% in 1990-94; then dropping steadily to 0.3% in 2015-18. Nonetheless, the number of publications for all search queries has increased over time, indicating a sustained interest in Bayesian research.

[1]For instance: Bayesian linear regression, Bayesian estimator and MCMC.

| | Before 1985 | 1985-89 | 1990-94 | 1995-99 | 2000-04 | 2005-09 | 2010-14 | 2015-18 |
|---|---|---|---|---|---|---|---|---|
| **No. of published content in all fields / disciplines** | 418,465,841 | 59,559,312 | 79,178,723 | 120,016,512 | 168,694,592 | 255,958,958 | 470,260,799 | 377,742,130 |
| **No. of published content with the word / term: "Bayesian" searched** | 24,326 | 14,192 | 21,718 | 32,971 | 63,323 | 167,816 | 377,322 | 355,389 |
| **% of published content with the word / term: "Bayesian" searched** | 0.006 | 0.024 | 0.027 | 0.027 | 0.038 | 0.066 | 0.080 | 0.094 |

**(a)** Number (and %) of published content available for specified terms.

| | "Bayesian data analysis" | | "Bayesian inference" | | "Bayesian computation" | | "Bayesian decision theory" | | "Bayesian MCMC" | | "Bayesian SPC" | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Number* | *%* | *Number* | *%* | *Number* | *%* | *Number* | *%* | *Number* | *%* | *Number* | *%* |
| **Pre 1985** | 16,177 | 66.5 | 10,379 | 42.7 | 8,670 | 35.6 | 14,696 | 60.4 | 108 | 0.4 | 254 | 1.0 |
| **1985-89** | 9,704 | 68.4 | 6,075 | 42.8 | 4,904 | 34.6 | 7,736 | 54.5 | 8 | 0.1 | 210 | 1.5 |
| **1990-94** | 13,931 | 64.1 | 8,639 | 39.8 | 6,911 | 31.8 | 9,984 | 46.0 | 60 | 0.3 | 369 | 1.7 |
| **1995-99** | 21,508 | 65.2 | 12,112 | 36.7 | 9,005 | 27.3 | 12,741 | 38.6 | 896 | 2.7 | 436 | 1.3 |
| **2000-04** | 42,177 | 66.6 | 22,269 | 35.2 | 14,557 | 23.0 | 18,383 | 29.0 | 4,192 | 6.6 | 429 | 0.7 |
| **2005-09** | 114,675 | 68.3 | 58,152 | 34.7 | 34,953 | 20.8 | 37,424 | 22.3 | 13,639 | 8.1 | 666 | 0.4 |
| **2010-14** | 258,636 | 68.5 | 125,400 | 33.2 | 69,778 | 18.5 | 66,737 | 17.7 | 34,787 | 9.2 | 1,017 | 0.3 |
| **2015-18** | 239,041 | 67.3 | 110,646 | 31.1 | 62,228 | 17.5 | 54,488 | 15.3 | 29,971 | 8.4 | 897 | 0.3 |

**(b)** Number and % of "Bayesian X" terms expressed over the number of published Bayesian content in sub-table (A).

**Table 1.2:** The total number and percentage of published content in all fields and items containing selected "Bayesian" terms, distributed by years. Coverage: Updated as at 14th February 2019. Source: Brunel Library (Summon) with results beyond Brunel Library's collection.

### 1.5.2 Bayesian data analysis

Bayesian data analysis is based on the Bayesian approach to statistical inference. In summary, there is some data to explain and some candidate explanations. Before obtaining new information, each candidate explanation has prior credibility of providing the best explanation. When new data is provided, the superior candidate explanation that best accounts for the data is chosen, and all other candidate explanations that do not account for the data are shifted away [28]. Naturally, all Bayesian problems begin with Bayes' theorem:

$$\pi_*(\theta|\mathcal{D}) = \frac{\mathcal{L}(\mathcal{D}|\theta)\pi_0(\theta)}{\mathbb{P}(\mathcal{D})}, \tag{1.1}$$

where $\theta \in \Theta$ is an unknown parameter of interest, $\mathcal{D}$ is the data observed, $\pi_0(\theta)$ is the prior probability distribution for $\theta$ which represents the opinion and state of knowledge before observing the current data, $\mathcal{L}(\mathcal{D}|\theta)$ is the likelihood distribution of $\theta$ associated with the model for the data as collected, and $\pi_*(\theta|\mathcal{D})$ is the posterior distribution for $\theta$ after computing everything on the right, taking into account the data observed. $\mathbb{P}(\mathcal{D})$ is the marginal likelihood and should be the probability of generating the data, but it is difficult

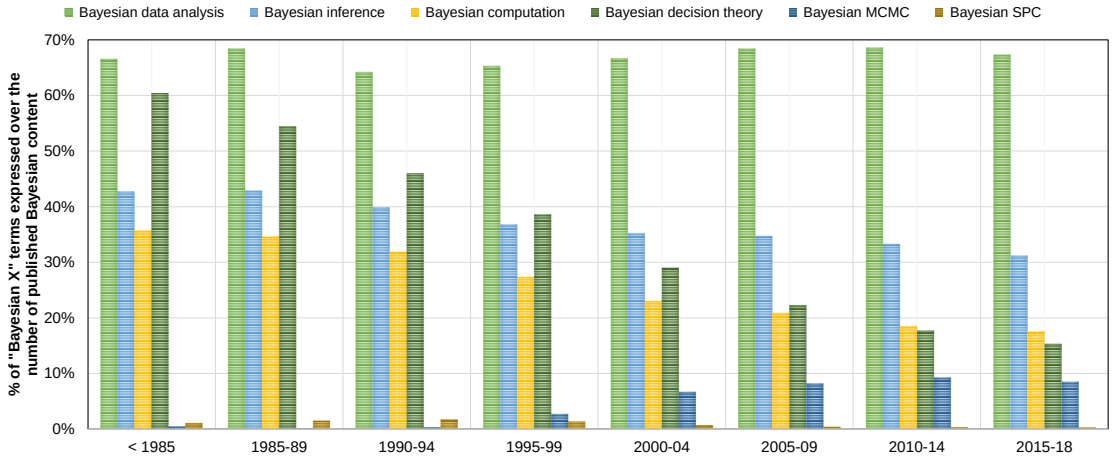



**Figure 1.1:** The % of published content with "Bayesian" terms searched as a fraction of the number of published content with the term "Bayesian" only, distributed by years. Coverage: Updated as at 14th February 2019.

to calculate on its own, so the alternative way of describing this relationship is to be one of proportionality:

$$\pi_*(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta)\pi_0(\theta). \tag{1.2}$$

Since the posterior function is a probability, the sum or integral over all possible $\theta$ should be 1, i.e.

$$\mathbb{P}(\mathcal{D}) = \int_{\Theta} \mathcal{L}(\mathcal{D}|\theta)\pi_0(\theta)\, d\theta.$$

In this case, the value of $\mathbb{P}(\mathcal{D})$ is the normalising constant which ensures that the posterior PDF integrates to unity. If $\mathbb{P}(\mathcal{D})$ included additional parameters, then it would compose of multiple integrals/sums.

Bayes' theorem is used to update the probability for a hypothesis as prior information $I$ becomes available. The value of $\theta$ is distinguished by a prior density and is independent of the measuring system but dependent by prior information $I$, i.e. $\pi_0(\theta) \equiv \pi_0(\theta|I)$ . The likelihood function is determined based on the system design, calibration information and knowledge of relevant quantities of influence, for example: environmental settings and material types. A Gaussian distribution may describe the likelihood function in many practical cases. The measuring system is used in several cases to replace precise measurement information with comparatively minuscule prior measurement knowledge. In this case, only the likelihood (processing the measurement information) defines the posterior distribution, that is, $\pi_*(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta)$.

The key difference between the classical frequentist approach and the Bayesian approach is that the parameters of the model are solely dependent on the data, while the Bayesian approach allows one to integrate other information through the use of a prior. These differences are summarized in Table 1.3.

| | Frequentist | Bayesian |
|---|---|---|
| Parameters (unknown) | Fixed | Random |
| Data (known) | Random | Fixed |
| Probability model | $\mathcal{L}(\mathcal{D}\|\theta)$ | $\mathcal{L}(\mathcal{D}\|\theta)\pi_0(\theta)$ |

**Table 1.3:** Differences between frequentist and Bayesian approaches.

The essence of conditional probability in beliefs and hypotheses, is based on a philosophical rendition of probability where probability is understood to represent a certain "degree of belief" in a given event, i.e. prior knowledge about the event, which updates as more information is presented rather than depend on a numerical value based on frequency. This "degree of belief" differs from the frequentist rendering that sees probability as the mathematical limit of the relative frequency as the sample size increases towards infinity [25].

#### *1.5.2.1 Estimating a quantity*

An estimate of a quantity is frequently used to illustrate a measurement result. A best estimate $\mu_*$ is commonly used to represent that estimate and is assumed to be the posterior distribution's statistical expectation (mean). Similarly, the corresponding dispersion parameter $\sigma_*^2$ is the square of the standard deviation (uncertainty). If both quantities are from a continuous distribution, then $\mu_*$ and $\sigma_*^2$ can be calculated as follows:

$$\mu_* = \int_\Theta \theta \; \pi_*(\theta|\mathcal{D}) \; d\theta, \quad \sigma_*^2 = \int_\Theta (\theta - \mu_*)^2 \; \pi_*(\theta|\mathcal{D}) \; d\theta.$$

#### *1.5.2.2 Credibility intervals*

The Bayesian credibility interval (CI) is another useful posterior indicator that shows where an unobserved parameter value falls in the domain of a posterior probability distribution given the available information or a predictive posterior distribution (PPD). For two reasons, CIs differ from classical confidence intervals:

- CIs include problem-specific contextual information from the prior distribution, whereas confidence intervals are solely based on data;
- CIs and confidence intervals treat nuisance parameters very differently.

The definition of a $(1-\alpha)$ CI for the unknown true value of $\theta$, given $\mathcal{D}$, is

$$\mathbb{P}(\theta \in [L(\mathcal{D}), U(\mathcal{D})]|\mathcal{D}) = \alpha,$$

where $\alpha \in [0, 1]$ is the credibility value and $[L(\mathcal{D}), U(\mathcal{D})]$ are the lower and upper limits of the data.

### 1.5.3 Selecting suitable prior distributions

Selecting a suitable prior distribution for $\pi_0(\theta)$ for the prior beliefs of $\theta$ has always been a major obstacle to widespread acceptance of Bayesian inference. The inclusion of the prior distribution causes the majority of the disagreement between Bayesian and classical statistics approaches. Wasserman [29], among many others, past and present, is a frequent critic on the issues surrounding the choice of selecting priors, by summarising all types of priors (including *informative* and *non-informative*) as weakly informative. Each of these informative and non-informative prior beliefs can have an effect on the final assessment in different ways. This set of beliefs can be preferred for a number of reasons:

- Informative and empirical; the data from similar studies are used to support the previous beliefs based on the data.
- Informative and non-empirical; there is some reason behind giving certain data more weight than others.
- Informative and domain-knowledge; despite the lack of conclusive evidence, experts know some facts are more reliable than others.
- Non-informative; data speaks for itself.

Furthermore, it is argued that the prior introduces too much subjectivity into the modeling process, skewing the results. There are, however, a number of responses to these criticisms:

- Priors can be non-informative (diffuse), meaning they give little or no information during the estimation process;
- Maximum likelihood (maximum likelihood) estimates are a type of a maximum a posteriori probability (MAP) estimation that is a subset of the Bayesian framework in which all prior distributions are uniform;
- Priors are usually asymptotically irrelevant, meaning that as N grows larger, the prior has less of an impact on the outcome.
- Instead of assuming naivety and testing a null hypothesis that is assumed to be false anyway, priors allows research to build on previous research.

This leads to the obvious questions here: "What to choose?" and "What class of models?"

Articles within the last decade have proposed prior distributions for classical (non-hierarchical) logistic and linear regression models [30], normal-gamma regression models [31],

models used for machine learning [32] and many more; along with newer practices such as objective Bayesian analysis for high-dimensional models [26] and deep learning neural networks [27], which are rapidly growing areas of research. The norm in these articles is to continue using the same (or similar) prior until it goes wrong. While this may seem like a practical approach, it is an extremely informal way of going through the model selection process.

Traditionalist would suggest going for the awkward choice, that is, either the non-informative (or diffuse) prior which will usually provide decent results for any given parameter value in strongly informative situations, or the informative prior which is usually the most unrealistic case. The flat (uniform) prior, usually denoted as a beta distribution with parameters set to 1, i.e.

$$\text{Beta}(1,1) \to \pi_0(\theta) \propto 1, \tag{1.3}$$

which is a prior distribution that assigns equal probability to all possible parameter values, is a popular option for a non-informative prior. Instinctively, this makes sense, and flat priors on the regression parameter are not informative in some cases, such as linear regression. This essentially spreads the density across the entire real positive line. It does not, of course, integrate into 1 and is an improper prior. There is also a weakly informative prior, which possesses some information about a parameter and an aim to regulate inferences logically.

However, in a world where everyday decisions are virtually Bayesian and inference has becoming increasingly easier for more complex models with various parameters and larger data-sets, what's needed is an *in-between informative prior*. The best prior distributions are those that accommodate computational restrictions by controlling, conveying, and standardizing variables.

### 1.5.4 Bayesian computation

Bayesian methods are well known for being computationally intensive, especially involving complex models with multiple variables (e.g., failure rate) which can support additional modelling to combine information; and as such, suitable models have historically allowed the integration of other sources of information. Moreover, because the marginal distribution of the data from Bayes' theorem is often a complicated integral, computing it can be challenging. Because the parameter is independent of the denominator, equation (1.2) can be used, which is easier to simulate from, and a Bayesian analysis' goal is to simulate values based on the posterior distribution.

Often, it may take several days to conduct a full analysis on complex models using Bayes' theorem. The need for advanced algebraic methods and an understanding of probability theory have also discouraged the use of applying this method. Nevertheless, it is essential to conduct decision analysis (e.g., minimising expected loss or maximising expected payout) based on parameter estimation techniques (e.g., MAP or ML solutions).

To tackle problems arising in Bayesian sampling methods, Geman and Geman [33] introduced what is now call Gibbs sampling, which was the first time that MCMC sampling methods had been properly applied to Bayesian problems. Therefore, they have created a technique by which simulating conditional distributions is the same as simulating joint distributions.

#### *1.5.4.1 MCMC methods*

Bayesian data analysis based on MCMC methods has become the choice method for examining and interpreting data in almost all scientific disciplines. Efficient and effective methodologies based on Monte Carlo are constantly being researched and discussed [34]. MCMC methods comprise of a general class of algorithms, which samples and estimates by simulation, to obtain an expectation of a statistic in a complex model from a probability distribution. Markov processes are the basis for general methods of MCMC simulation used to simulate sampling from complex probability distributions, which have been widely used in Bayesian inference, for instance, in hierarchical MCMC for Bayesian system reliability [35]. The last two initials in MCMC stems from the need for the random numbers to drive the Markov process with Monte Carlo integration. The random samples are statistically independent in modern Monte Carlo integration, while they are correlated in MCMC.

Bootstrap analysis, Gibbs sampling and MCMC sampling are often used, and although seemingly simple, are computationally draining and requires several hours on a personal computer [36, 37]. Other well-known examples of Monte Carlo random walk methods include the Metropolis–Hastings (MH) algorithm, slice sampling, Multiple-try Metropolis, Reversible-jump and the Hamiltonian (or Hybrid) Monte Carlo (HMC); see more in [34, 35]. The Gibbs sampler is a special case of the MH algorithm since it takes samples from the full conditional distributions of parameters. These methods have also paved the way for the development of techniques for investigating PPDs, which I discuss here as an extension of the estimation algorithm.

### 1.5.5 Posterior predictive distributions

Other statistics, in addition to descriptive statistics, can be generated from an MCMC-generated sample to make inferences about other important quantities of interest and to evaluate model fit. One such quantity is the posterior predictive distribution (PPD). The PPD is the probability distribution of unobserved values based on observed values that provides some insight into what the next observation might look like given the prior, whereas Bayes' theorem in (1.2) computes the posterior. This distribution is defined as a mixture of the new observation's PDF and the posterior, and it is given by:

$$\pi_*(y_{\text{new}}|\mathcal{D}) = \int_\Theta g(y_{\text{new}}|\theta)\pi_*(\theta|\mathcal{D})\, d\theta, \tag{1.4}$$

where $g(y_{\text{new}}|\theta)$ is the PDF of a new observation $y_{\text{new}}$ from the sampling distribution. The PPD takes into account two types of uncertainty for $\theta$:

- sampling uncertainty about $y_{\text{new}}|\theta$;
- parametric uncertainty about $\theta$.

This makes it a more powerful tool and distinguishes it from classical predictive distributions, which rely on a single best estimate for $\theta$.

### 1.5.6 Bayesian hierarchical modeling

Bayesian hierarchical modeling (BHM) is a statistically rigorous way to draw conclusions based on many observations about unknown parameters. The model is written in several levels that use the Bayesian approach to estimate the parameters of the posterior distribution. The sub-models combine to form the Bayesian hierarchical model, and Bayes' theorem in (1.2) is applied with the observed data to account for all the present uncertainty. This leads to the multi-staged posterior distribution, as additional evidence is obtained on the prior distribution ([38]).

In general, BHM uses two important concepts in the derivation of the posterior distribution: hyperparameters, which are parameters of the prior distribution, and hyperpriors, the distribution of hyperparameters. For instance, if the marginal likelihood included multiple influencing factors (say, temperature $\theta_1$, environmental settings $\theta_2$ and material types $\theta_3$), then $\mathbb{P}(\mathcal{D})$ would compose of multiple integrals. This typically leads unto a multi-stage BHM with hyperparameters $\theta_i \in \Theta$ and hyperprior distribution $\pi_0(\theta_i)$ for $i = 1, 2, \ldots, N$, where $N$ is the number of factors. Finite exchangeability for all parameters is assumed under BHM. This means that the joint probability distribution of the observed data $\mathcal{D}$ and

the influencing factors $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N)$ does not change when the positions are changed for different permutations.

#### *1.5.6.1 A two-stage BHM*

For a simple case, the two-stage BHM can be described by setting the following inputs:

$$\mathcal{D}|\theta_1, \theta_2 \sim \mathcal{L}(\mathcal{D}|\theta_1, \theta_2), \quad \theta_1, \theta_2 \sim \pi_0(\theta_1, \theta_2), \quad \theta_2 \sim \pi_0(\theta_2).$$

The joint posterior distribution in a two-stage BHM is therefore:

$$\pi_*(\theta_1, \theta_2|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta_1, \theta_2)\pi_0(\theta_1, \theta_2), \quad \mathbb{P}(\mathcal{D}) = \iint_\Theta \mathcal{L}(\mathcal{D}|\theta_1, \theta_2)\pi_0(\theta_1, \theta_2)\, d\theta_1 d\theta_2. \quad (1.5)$$

By the law of conditional probability, the prior of $\pi_0(\theta_1, \theta_2)$ can be expressed as $\pi_0(\theta_1|\theta_2)\, \pi_0(\theta_2)$ or $\pi_0(\theta_2|\theta_1)\, \pi_0(\theta_1)$. When either $\theta_1$ or $\theta_2$ is not required for a specific measurement, marginalizing one or the other results in

$$\pi_*(\theta_1|\mathcal{D}) = \int_\Theta \pi_*(\theta_1, \theta_2|\mathcal{D})\, d\theta_2, \quad \pi_*(\theta_2|\mathcal{D}) = \int_\Theta \pi_*(\theta_1, \theta_2|\mathcal{D})\, d\theta_1. \quad (1.6)$$

### 1.5.7 Bayesian point estimation

Bayesian point estimation focuses on the summary statistics of its posterior distribution, namely: the *posterior mean*, which minimizes the posterior expected loss for a given cost function, i.e. a squared-error cost function, the *posterior median*, which minimizes the posterior risk for the absolute-value cost function, and the *posterior mode*, which is easy to obtain via the maximum a posteriori (MAP) estimate. The posterior's statistics of central tendency are pivotal in Bayesian decision making.

Significant research has been made in regards to MAP estimation, namely for hidden Markov models and several useful parametric densities generally used in SMART systems, i.e. automatic speech recognition and natural language processing; for instance, see [39]. MAP probability usually comes up in Bayesian setting, typically, as a contrast to the ML estimation. Because, as the name suggests, it works on a posterior distribution, not only the likelihood. Even for complex problems where ML estimation fails, the MAP estimator is well-known for its well-defined asymptotic properties. For simpler problems, the ML estimator is consistent and agrees with the MAP estimator [40].

Recall, that the posterior distribution $\pi_*(\theta|\mathcal{D})$ contains current knowledge about some unknown quantity $\theta$. This means I can use posterior distribution to obtain point or interval estimates of $\theta$. A well known method to find point estimates is to choose a value for $\theta$ that maximises the posterior. This method is called the MAP estimation and is used to obtain a

point estimate of an unknown quantity on the basis of previous information, which equals the mode of the posterior distribution (this is the highest point on a PDF, see discrete and continuous cases in Figure 1.2).

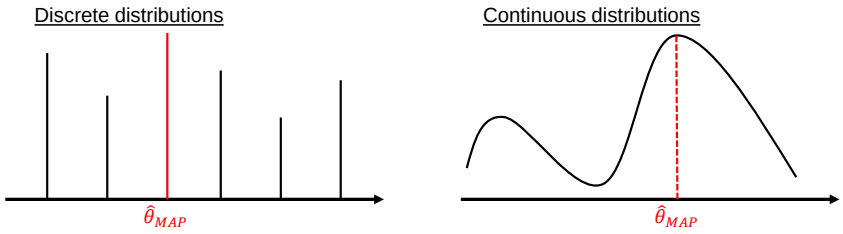



**Figure 1.2:** The MAP estimate of $\theta|\mathcal{D}$ is the value of $\theta$ that maximizes the posterior density $\pi(\theta|\mathcal{D})$, given by $\hat{\theta}_{MAP}$.

From Bayes' theorem in (1.2), one could obtain the posterior from likelihood and prior, without having to be concerned with usually scaling constants, since the primary intent is optimisation, so proportionality works. Thus, the estimate

$$\hat{\theta}_{MAP} = \arg\max_{\theta} \pi_*(\theta|\mathcal{D}) = \arg\max_{\theta} \mathcal{L}(\mathcal{D}|\theta)\pi_0(\theta), \tag{1.7}$$

is the MAP estimate of $\theta$. The maximising value can be obtained by differentiating with respect to $\theta$, then solving for $\theta$ (and checking the maximisation criteria). Note that (1.7) coincides with the ML estimate when the prior $\pi_0(\theta)$ is a constant (i.e. a uniform density).

Additional Bayesian point estimators, such as the minimum message length (MML) point estimator, based on Bayesian information theory (the study of quantification, storage, and communication of information), have been used, but they do not require extensive use of the posterior distribution. Furthermore, Bayesian point estimation frequently employs computational methods previously discussed in Section 1.5.4, such as the Markov chain Monte Carlo method (MCMC).

### 1.5.8 Overview of current FMEA

The current FMEA (Failure Mode and Effects Analysis) framework is an inductive and bottom-up reasoning tool for failure analysis and is used in many areas of engineering (i.e. quality engineering including reliability and safety). This tool was developed by reliability engineers in the late 1950s to analyse problems that might arise from malfunctioning of military systems [8]. As a quantitative analysis, a systematic approach to the use of FMEA consists of identifying ways in which failure can occur ("failure mode"), then identifying the causes, effects of each failure mode, evaluating the corresponding risks, and then taking the necessary steps to minimise the risks and consequences. By including a criticality analysis,

the FMEA can be extended to an FMECA. In the FMECA, the criticality of each failure mode is calculated using a RPN according to three risk measures:

- Severity $\sim$ rates the severity of the potential effect of the failure;
- Occurrence $\sim$ rates the probability that the failure will occur;
- Detection $\sim$ rates the probability that the problem will be detected before it reaches the system or user.

Each risk measure (severity, occurrence and detection) is typically quantified using rating scales (e.g., from 1 to 5, or from 1 to 10), with the higher number representing the higher risk. On a five-point Occurrence scale, for example, a rating of 5 means that the failure is most likely to occur, while a rating of 1 indicates that the failure is extremely unlikely to occur. The company or research team defines the basic rating descriptions and standards to match the products or processes being assessed. The RPN for each failure mode is determined by multiplying each risk measure, i.e.

$$\text{RPN} = \text{Severity} \times \text{Occurrence} \times \text{Detection}. \tag{1.8}$$

As an example, Table 1.4 shows a sample five-point scale for severity, occurrence and detection risk measures.

| | Severity | | Occurrence | | Detection | |
|---|---|---|---|---|---|---|
| **Rating** | **Effect** | **Description** | **Probability** | **Description** | **Probability** | **Description** |
| **1** | Negligible | Very little or no effect on the customer. | Remote | Failure is extremely unlikely or non-existent with very low probability. | Certain | A possible cause/mechanism and failure mode would almost certainly be detected by Design Control. |
| **2** | Minor | Very minor effect on the customer. | Low | Few failures are likely. | High | A high chance the Design Control will pick up on a possible cause/mechanism and failure mode. |
| **3** | Moderate | A moderate effect on the customer. Requires attention. | Moderate | Occasional failure is expected. | Moderate | A moderate chance that the Design Control will pick up on a possible cause/mechanism and failure mode. |
| **4** | Critical | A critical effect on the customer. Requires significant attention. | High | High number of failures. | Low | A low chance the Design Control will pick up on a possible cause/mechanism and failure mode. |
| **5** | Catastrophic | A catastrophic effect on the customer and occurs without warning. Failure mode is noncompliant with quality regulations. | Inevitable | Failure is certain to happen with very high probability. | Undetectable | A possible cause/mechanism and failure mode will almost certainly be undetected by Design Control; or there is no detection available. |

**Table 1.4:** Rating for severity, occurrence and detection.

Another result useful to FMECA is criticality which is found by multiplying severity by occurrence, i.e.

$$\text{Criticality} = \text{Severity} \times \text{Occurrence}. \tag{1.9}$$

Equations (1.8) and (1.9) give relative priority to the failure modes in the order they should be addressed. The RPN is implemented and explored in greater detail in chapters 2 and 3.

The outcomes of an FMEA/FMECA are usually reported in a form (examples of forms can be seen in later case studies, e.g. chapter 2-4). There are standards available that provide guidance for conducting a proper FMEA/FMECA (e.g., [41]). Consider FMEA/FMECA as a generic form, regardless of whether or not it includes a criticality analysis. Figure 1.3 depicts a flowchart detailing the FMEA process.

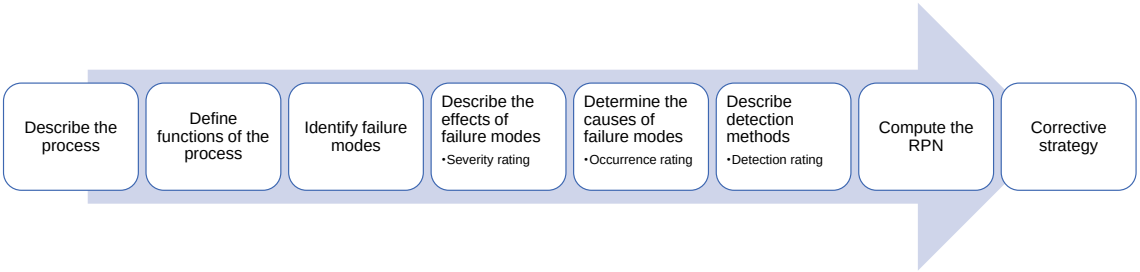



**Figure 1.3:** Flow chart describing the FMEA procedure.

Despite the significant contribution that conventional FMEA has made to industry around the world, its flaws have never been overlooked, and it has been criticized for a variety of reasons. More information on the shortcomings of traditional FMEA will be addressed in the following section.

### 1.5.9 Limitations of current FMEA

The method of obtaining RPNs is the most widely criticized shortcoming of FMEA. RPN is a critical component of FMEA and the primary method for assessing the probability of failure modes. However, since RPN is the product of risk measures, as seen in (1.8), its validity and rationality, as well as the result it produces, are frequently challenged [42–46]. Some of the common reasons are:

i If completely different combinations of risk measures are used, the RPN value can be the same even though the risks are completely different. On a ten-point scale, two incidents with S, O, and D scores of 10, 2, 2 and 8, 5, 1 have the same RPN values of 40, despite the fact that they pose completely different risks.

ii The distribution of RPNs from 1 to 1000 is not uniform, which makes understanding the sense of variations between RPNs difficult.

iii Since risk values are also quantified based on the expertise of the team members conducting the study, different assessment outcomes can be obtained when the same failure mode is analyzed by different FMEA teams. This makes the resulting RPN

difficult to quantify precisely and objectively due to the subjective nature of choosing risk values from 1 to 10.

iv The number of parameters that can be used in a traditional FMEA is limited.

As a result, it is concluded that RPNs used in conventional FMEA may not provide a clear risk assessment. Moreover, the RPN's ability to compare and incorporate risk data is severely limited.

### 1.5.10 Improvements to FMEA

Several methods have been taken thus far in order to make FMEA a more effective method for risk prioritization. Some of the most common approaches to improving conventional FMEA include incorporating other variables such as costs into the risk assessment process or combining FMEA with another failure analysis system such as the FTA.

For example, Rhee [47] proposed systematic use of empirical evidence for life cost-based FMEA. According to this method, the system's availability can be calculated using empirical data, which includes downtime and failure frequency. Furthermore, using empirical evidence, loss time can be calculated, allowing for the calculation of failure cost. Later on, von Ahsen [48] suggested a cost-oriented approach to develop the FMEA process from an economic standpoint, which takes into account the failure costs associated with both externally and internally detected faults to completely cover a financial risk assessment. A case study with an automobile supplier demonstrated that the new approach is more beneficial than traditional FMEA.

Outside of traditional FMEA, a combined approach using FMEA and FTA has been featured. The most widely used tools in the industry, by far, are FMEA and FTA [49]. FTA is a top-down analysis that visually represents a failure path or failure chain and is based on the principle of Boolean logic, which allows for the construction of a set of True/False statements. It employs logic gates and events to model how the component states contribute to the overall state of the system. The aim of FTA is to efficiently detect causes of failure and reduce risks before they happen. The following logic gates are widely used in FTA: (1) OR-gate, (2) AND-gate and (3) inhibit or conditional gate. FTA's symbols are defined in depth in [50]. Moreover, a strictly qualitative FTA can be transformed into a quantitative FTA by incorporating quantitative component reliability data (e.g., failure rates). Using boolean algebra, such a quantitative FTA can be used to calculate the system's reliability. A comprehensive description of mostly quantitative FTAs is also given in Hamda [51].

Some authors argue that FTA and FMEA should be used together. For example, Bertsche [52] claims that due to the different starting points of both methods, this can increase the number of failure modes found (e.g. FMEA from the bottom up vs. FTA from the top down). Even so, conducting both analyses would be time consuming and could result in a lack of concentration on the most important components of the system, which is what a failure analysis is supposed to identify.

To summarize, several different approaches to improving FMEA have been proposed, and many more in the literature (e.g., identifying the most critical failure causes or by incorporating a fuzzy belief rule approach with Bayesian networks [53, 54]). Although some of them have proven to be effective in removing the flaws of conventional FMEA, others need further investigation.

## 1.6 Thesis structure

This chapter introduced the research's purpose, as well as the overall research goal and related objectives. A background and literature review was provided and is also given in the introduction portions of subsequent chapters. The thesis is divided into five chapters and the remaining four are organised as follows.

The first case study is covered in chapter 2, where I use a FMEA and a modified SEIR epidemic model to perform a Bayesian risk assessment of the ongoing COVID-19 pandemic. The second case study is covered in chapter 3, in which I present a Bayesian risk assessment of the CMM measurement process using an augmented form error model and FMECA. The third case study is covered in chapter 4, where I present a Bayesian risk assessment to failure prediction using an autoregressive form error model, Gibbs sampling and an integrated FTA/FMEA. Analytical and numerical methods are used in all studies to emphasize the practical implications of my findings, and the proposed methodologies can be used in future COVID-19 like pandemics and wide quality engineering. Finally, in chapter 5, I provide an overview of the discussion as well as some concluding remarks and suggestions for future work.

# Chapter 2

# A Bayesian Risk Assessment of the COVID-19 Pandemic using a Modified SEIR Epidemic Model

## 2.1 Introduction

### 2.1.1 Overview

The coronavirus disease 2019 (COVID-19) is an infectious disease and global pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and has been designated a Public Health Emergency of International Concern by the World Health Organization (WHO) [6]. The 1918-19 H1N1 influenza pandemic was the last time the world responded to an imminent global disease outbreak on the size of the current COVID-19 pandemic with no exposure to vaccines. The COVID-19 disease was first detected in Wuhan, China in 2019, and has since spread worldwide, leading to a coronavirus outbreak of 2019–20 [55]. The virus is suspected to have an animal origin by spillover infection and was first transmitted to humans in Wuhan, China, in November or December 2019 and became a major outbreak by early January 2020. The COVID-19 pandemic hit the United Kingdom (UK) at the end of January 2020. There is often a delay between the onset of symptoms and correct diagnosis. The most critical things are the prompt diagnosis and identification of the infected, and the number of confirmed patients. Common symptoms include coughing, shortness of breath and fever, while less common symptoms may include muscle pain, sputum problems, and sore throat. Although most cases have mild symptoms,

some cases have progressed to extreme pneumonia and multi-organ failure.

Due to the unexpected and ongoing nature of COVID-19, the main interest is to evaluate using a Bayesian inference approach and a modified SEIR model (susceptible-exposed-infectious-recovered) for the transmission dynamics of COVID-19. Here, the effective reproduction number ($\mathcal{R}_t$) is estimated based on laboratory-confirmed cases and death data using Bayesian inference and analyse the impact of the community spread of COVID-19 across the UK. The value of $\mathcal{R}_t$ models the average number of infections caused by a case of an infectious disease in a population that includes not only susceptible people. A FMEA (Failure mode and effects analysis) is then applied to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic. As the focus of FMEA is COVID-19 infections, failure modes are taken as positive cases. Therefore, the model is applied to COVID-19 data to demonstrate the effectiveness of control interventions by reducing the value of $\mathcal{R}_t$. This study will show that, combined with FMEA and conditional analysis, Bayesian modelling will help to expand and provide responses to this outbreak. A summary of this case study is presented in Figure 2.1.

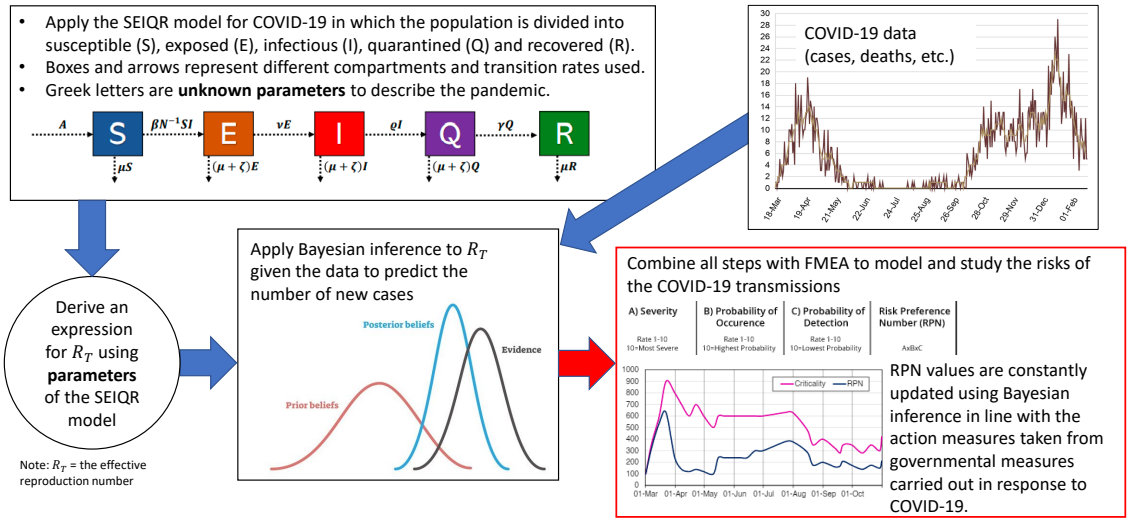



**Figure 2.1:** Summary of the first case study methodology.

### 2.1.2 Coverage and data sources

Real-time data on the number of laboratory-confirmed cases, recoveries and deaths due to COVID-19 have been derived from a variety of authoritative sources, such as the Official Websites of Ministries of Health or other Government Agencies and the Social Media Pages of Government Authorities. Most of the data can also be accessed via *Worldometer* (https://www.worldometers.info/coronavirus/#countries), which is a

reference website and has been cited in more than 10,000 published books and 6,000 technical journal articles and has been named one of the best free reference websites of the American Library Association (ALA), the oldest and largest library organisation in the world. Data on hospitalisation for the average length of stay in hospitals in many countries around the world can be found in the OECD (https://data.oecd.org/healthcare/length-of-hospital-stay.htm). Data on infections and testing, hospital resource use, mask use and social distancing (as well projections and forecasting) is provided by the Institute for Health Metrics and Evaluation (https://covid19.healthdata.org/). The data on these websites is collected and analysed regularly around the clock, 24 hours a day and 7 days a week.

The next step is to follow the standard routine of dynamic modeling by focusing on some of the most affected countries in Europe (i.e., UK, France, Spain, Netherlands, Germany and Portugal). It is important to know the daily and total number of lab-confirmed positive cases, recoveries, and deaths. The latest cases data for the United Kingdom (as of 31 August 2020) is available for selected regions throughout the country, while the latest death data is limited to the nations of the United Kingdom (i.e. England, Northern Ireland, Scotland, and Wales). Figure 2.2 presents the number of lab-confirmed daily cases with COVID-19 for some of the infected countries in Europe, recorded between 1 March and 31 October, inclusive (data source: https://ourworldindata.org/coronavirus/). The recovery data for Spain and the United Kingdom is unavailable to the public.

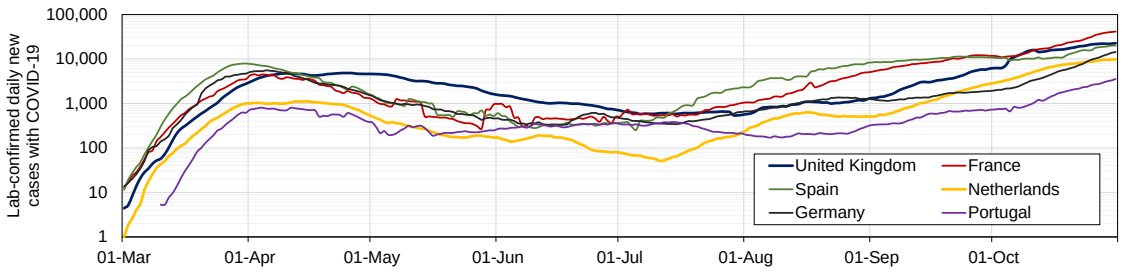



**Figure 2.2:** A line graph for the number of lab-confirmed daily cases with COVID-19 for some of the infected countries in Europe. The number of cases are displayed using a logarithmic scale to base 10. Coverage: 1 March to 31 October.

## 2.2 Model and methods

### 2.2.1 A modified SEIR model adapted to COVID-19

The classical SEIR model assumes that births and deaths are not related to factors other than the disease itself, and is an extension to the classical SIR (susceptible-infectious-recovered) model [56]. Each compartment in the SEIR model represents:

- $S \sim$ absolute number of susceptible individuals (those who may potentially contract the disease). Without further information, this group is represented by the whole population.
- $E \sim$ absolute number of exposed individuals (those who have been infected but are not yet infectious).
- $I \sim$ absolute number of infective individuals (those capable of transmitting the disease).
- $R \sim$ absolute number of recovered individuals (those who have become immune).

The SEIR epidemic model has been widely used in large populations to study the dynamics of infectious diseases when there is an incubation period during which individuals have been infected but are not yet infectious themselves. For this period, the individual is in the exposed state $E$. This may be called a latent phase, and may be infectious, partially infectious or not infectious yet. In the classical SEIR model, compartment $R$ also includes the number of people who died from the disease.

Several authors have developed different models of the classical SEIR taking into account the complexities of the disease in order to make the model as realistic as possible. To generalise the SEIR model to describe the COVID-19 pandemic in several countries, I adopt an additional compartment: $Q$ (quarantined), which adds a passage for the fraction of infectious individuals into the quarantined compartment. A characteristic feature of this model is that the total population in a given region $N$ is equal to the sum of all compartments:

$$S(t) + E(t) + I(t) + Q(t) + R(t) = N. \tag{2.1}$$

Equation (2.1) holds if the number of births $A$ are balanced with death rates $\mu$ during the time span of the disease (e.g. $A = \mu N$). These quantities are expressed in units of births and deaths per 1,000 individuals per year, respectively. All individuals in each compartment are subjected to a natural death. I define the reciprocal $\mu^{-1}$ as the average life expectancy

of a country's population in a given year. This value applies to typical human deaths (e.g. due to natural death, usual influenza, fatalities, etc.) and is not linked to COVID-19.

Initially, assume that the population at time $t = 0$ only includes susceptibles, so that, $S(0) \approx N$, $E(0) \geq 0$, $I(0) \geq 0$, $Q(0) = 0$ and $R(0) = 0$. There is a disease free equilibrium (DFE) with

$$(S(0), E(0), I(0), Q(0), R(0)) = (N, 0, 0, 0, 0). \tag{2.2}$$

The number of individuals in $I$ tends to zero as $t \to \infty$, which ensures that $\lim_{t\to\infty} S(t) =: S_\infty > 0$. Individuals in state $R$ have no further role to play in the epidemic, and it is assumed that infected individuals are mutually independent and shows an exponential growth, characteristic of any epidemic's initial stages [57].

### 2.2.2 Transition rates of vital dynamics

The exponential distribution plays a significant role as the probability distribution that underlies the time spent in a compartment (or state), which is fundamental for Bayesian modelling (later sections will explain this). As such, these states can be seen as a flow diagram in which the boxes and the arrows represent the different compartments and transition rates of vital dynamics used for COVID-19 (a graphical illustration of the transmission model for COVID-19 can be seen in Figure 2.3).

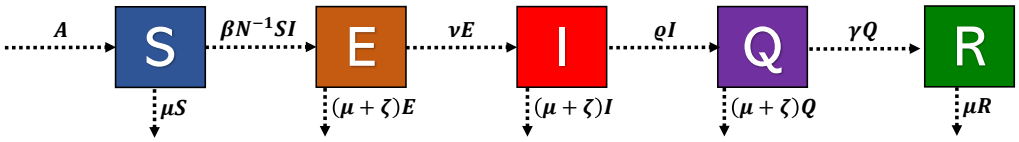



**Figure 2.3:** A modified SEIR transmission model in which the boxes and the arrows represent the different compartments and transition rates used for COVID-19.

The modified SEIR model for COVID-19 described in Figure 2.3 can also be expressed by the following set of nonlinear ordinary differential equations (ODEs), which evolves according to:

$$S'(t) = A - (\beta N^{-1} I(t) + \mu) S(t), \tag{2.3}$$

$$E'(t) = \beta N^{-1} S(t) I(t) - (\nu + \zeta + \mu) E(t), \tag{2.4}$$

$$I'(t) = \nu E(t) - (\varrho + \zeta + \mu) I(t), \tag{2.5}$$

$$Q'(t) = \varrho I(t) - (\gamma + \zeta + \mu) Q(t), \tag{2.6}$$

$$R'(t) = \gamma Q(t) - \mu R(t), \tag{2.7}$$

where $\beta$ is the infection (or contact) rate, $\nu$ is the average incubation rate, $\varrho$ is the average hospitalisation rate, $\gamma$ is the recovery rate and $\zeta$ is death rate due to COVID-19.

The transition between each state can be explained as follows. Transmission of diseases occur through interactions between susceptible and exposed persons in the home, workplace, school, hospitals or at random in the neighbourhood, with the latter depending on the spatial distance between interactions. Any such contact results in the susceptible contracting the virus and thus becoming exposed and then infected, leading to transition $S \to E$ with probability $S(t)/N$ and population $\beta I(t)$ which exposes $\beta^{-1}$ new individuals per day. When exposed individuals transition from $E \to I$, the probability is 1 since all exposed individuals will become infected, the population is $E(t)$ and the latency rate is $\nu$ with a mean incubation period of $\nu^{-1}$.

The infected individuals eventually transition from $I \to Q$ which matches the "active confirmed cases" as documented in many official databases and reports. In fact, an infected person is not quarantined immediately because the authorities have often been unable to test enough people while keeping pace with the spread of the disease. The rate at which the number of infectious individuals move into the $Q$ compartment is given by $\varrho$, which is the inverse of the mean time needed to quarantine an infected patient. Recovered individuals do not return to compartment $S$ as long-term immunity is assumed, but it remains to be seen if patients recovered from COVID-19 can produce antibodies and gain lifelong immunity.

In most transitions, the duration is typically explained by an exponential distribution. The need to quarantine is driven by the number of confirmed cases and is necessary to significantly and effectively reduce the spread of infection. Individuals in compartments $E$, $I$ or $Q$ may die from the virus with rate of $\zeta$, respectively.

It's worth noting that the proposed modified SEIR model excludes the vaccination effect due to the research's timeliness: vaccinations began in December 2020, and this study was completed by then. It would be interesting to change this model to add vaccination and compare the outcomes to the original version of the model in future research. This concept can be further expanded in many model realisations with varying post-vaccinated susceptibility, depending on vaccine efficiency, etc.

### 2.2.3 Deriving $\mathcal{R}_t$ using the Next Generation Method

The next-generation matrix (NGM) approach was introduced in 1990 by Diekmann *et al.* [58] and is, in such situations, a general method of deriving $\mathcal{R}_t$, encompassing any scenario in which the population is divided into independent, disjoint variables. This

technique assumes that the transmission probabilities between compartments are constant and that the distribution of the time of residence in each compartment is exponentially distributed. The NGM approach has been expanded on in numerous articles [59–61].

In this section, the steps required to find the NGM operator in matrix notation are outlined [62], and this approach is then applied to the proposed modified SEIR model adapted to COVID-19.

Let $\boldsymbol{F}$ be the matrix of transmissions of new infections in the infectious compartments, and let $\boldsymbol{V}$ be the matrix of individuals entering and leaving the infectious classes (i.e. matrix transitions). Assuming that both matrices meet the conditions of $\mathcal{R}_t$ for compartmental models at the DFE from (2.2), with $S = N$, then I can form a NGM operator from the ODEs of the infectious classes, given by $\boldsymbol{F}\boldsymbol{V}^{-1}$ where $\boldsymbol{V}^{-1}$ is the inverse $\boldsymbol{V}$. If infected states are denoted by indices $i$ and $j$, then entry $F_{ij}$ is the rate at which individuals in state $j$ transmit to individuals in state $i$. In other words, $F_{ij}$ is equal to zero if no new cases generated by a infectious individual in state $j$ can be in infected in state $i$ immediately after infection. The value of $\mathcal{R}_t$ is obtained from the spectral radius (dominant eigenvalue) of $\boldsymbol{F}\boldsymbol{V}^{-1}$ [63].

For populations that transmit the virus, an expression for $\mathcal{R}_t$ can be derived from two cases: one with quarantine and one without it.

#### *2.2.3.1 2-dimensional case: $\varrho = 0$*

In this example, I only model the exposed $E$ and infected $I$ classes. Assuming that the DFE conditions are met at matrices $\boldsymbol{F}$ and $\boldsymbol{V}$, then the infectious class dynamics for $\mathcal{R}_t$ is given by

$$\mathcal{R}_t = \frac{\beta\nu}{(\nu+\zeta+\mu)(\gamma+\zeta+\mu)}. \tag{2.8}$$

*Proof.* From the ODEs of the SEIR model, there are two infectious compartments: exposed $E$ (state 1) and infected $I$ (state 2). For this $2 \times 2$ system, I find that the appearance of new infections and individuals entering and leaving the infectious classes depends on:

$$\boldsymbol{F} = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix} \nu+\zeta+\mu & 0 \\ -\nu & \gamma+\zeta+\mu \end{pmatrix}.$$

The inverse of $\boldsymbol{V}$ is simple since I am using a $2 \times 2$ matrix, therefore:

$$\begin{aligned}\boldsymbol{V}^{-1} &= \frac{1}{(\gamma+\zeta+\mu)(\nu+\zeta+\mu)} \begin{pmatrix} \gamma+\zeta+\mu & 0 \\ \nu & \nu+\zeta+\mu \end{pmatrix} \\ &= \begin{pmatrix} (\nu+\zeta+\mu)^{-1} & 0 \\ \nu\big[(\nu+\zeta+\mu)(\gamma+\zeta+\mu)\big]^{-1} & (\gamma+\zeta+\mu)^{-1} \end{pmatrix}.\end{aligned}$$

Thus, $\boldsymbol{F}\boldsymbol{V}^{-1}$ simplifies to:

$$\boldsymbol{F}\boldsymbol{V}^{-1} = \begin{pmatrix} \frac{\beta\nu}{(\nu+\zeta+\mu)(\gamma+\zeta+\mu)} & \frac{\beta}{(\gamma+\zeta+\mu)} \\ 0 & 0 \end{pmatrix}.$$

The dominant eigenvalue of the above matrix is the largest value between first (top-left) and fourth (bottom-right) entries. Since the fourth entry is zero, then $\mathcal{R}_t$ is obtained from the first entry, thus the proof is complete. □

#### *2.2.3.2 3-dimensional case: $\varrho \neq 0$*

For the second case, I take into account all three infectious compartments: $E$, $I$ and $Q$. Once again, assuming that the DFE conditions are met at matrices $\boldsymbol{F}$ and $\boldsymbol{V}$, then the infectious class dynamics for $\mathcal{R}_t$ may also be given by

$$\mathcal{R}_t = \frac{\beta\nu\varrho}{(\nu+\zeta+\mu)(\varrho+\zeta+\mu)(\gamma+\zeta+\mu)}. \tag{2.9}$$

*Proof.* From the ODEs of our SEIR model, I now consider all three infectious compartments: $E$, $I$ and $Q$. Regarding the necessary ODEs for each compartment at the DFE, I obtain:

$$\boldsymbol{F} = \begin{pmatrix} 0 & 0 & \beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix} \nu+\zeta+\mu & 0 & 0 \\ -\nu & \varrho+\zeta+\mu & 0 \\ 0 & -\varrho & \gamma+\zeta+\mu \end{pmatrix}.$$

Hence, the NGM operator for $\boldsymbol{F}\boldsymbol{V}^{-1}$ under this $3 \times 3$ matrix is given by:

$$\begin{aligned}\boldsymbol{F}\boldsymbol{V}^{-1} &= \begin{pmatrix} 0 & 0 & \beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\nu+\zeta+\mu} & 0 & 0 \\ \frac{\nu}{(\nu+\zeta+\mu)(\varrho+\zeta+\mu)} & \frac{1}{\varrho+\zeta+\mu} & 0 \\ \frac{\nu\varrho}{(\nu+\zeta+\mu)(\varrho+\zeta+\mu)(\gamma+\zeta+\mu)} & \frac{\varrho}{(\varrho+\zeta+\mu)(\gamma+\zeta+\mu)} & \frac{1}{\gamma+\zeta+\mu} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\beta\nu\varrho}{(\nu+\zeta+\mu)(\varrho+\zeta+\mu)(\gamma+\zeta+\mu)} & \frac{\beta\varrho}{(\varrho+\zeta+\mu)(\gamma+\zeta+\mu)} & \frac{\beta}{\gamma+\zeta+\mu} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

The dominant eigenvalue is equal to $\mathcal{R}_t$ and is obtained from the first entry. Thus, the proof is complete. □

Since classical statistics are empirical and see probability as something that has to do with a limiting frequency based on an observable proportion; it will then be more important to see how the subjective nature of the Bayesian method is implemented by using a *priori* belief to describe a distribution of probabilities on the possible values of the unknown model parameters of the epidemic. Knowing the value of $\mathcal{R}_t$ makes this quantity very important for estimating the proportion of a population to be quarantined to avoid a further outbreak. In this case, the main interest will be to estimate some of these unknown parameters, such as $\beta$, $\gamma$, and $\zeta$, which will allow for prediction of the outbreak size from $\mathcal{R}_t$ and application of control methods to reduce it. Other transition parameters (i.e. $\nu$, $\varrho$ and $\mu$) will be estimated empirically from known sources.

### 2.2.4 The FMEA procedure to the community spread of COVID-19

The earliest form of risk analysis involved identifying all potential risks without taking into account the likelihood of them occurring. In the 1940s, structured risk analysis became popular as a risk assessment tool. Traditional risk analysis (TRA) approaches were plagued by poorly defined steps, high levels of uncertainty, and decision-making problems in the process. As a result, in the late 1940s, the US Armed Forces introduced FMEA which is a significant improvement over TRA and acts as a proactive tool for identifying, evaluating, and preventing process failures. In the late 1950s, reliability engineers further improved this method to analyze issues that could arise from military device malfunctions. Today, there are a wide variety of risk analysis methods from which to choose: FMEA and its extension FMECA,

In the case of pandemics, I will show that the FMEA can be used to assess the impacts of various potential failures in order to determine which prevention measures are most needed. Therefore in this chapter, I extend these ideas to analyse the COVID-19 pandemic in the United Kingdom.

Figure 2.4 illustrates the setup for the proposed FMEA procedure.

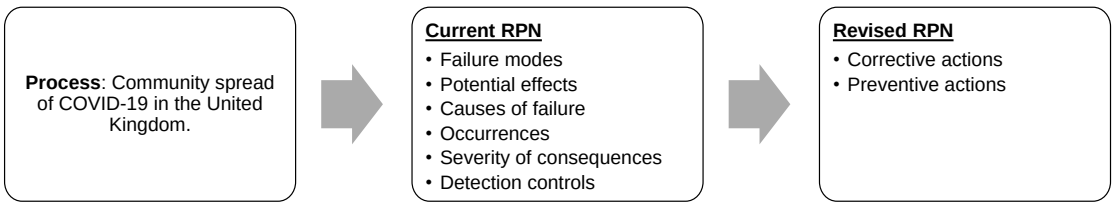



**Figure 2.4:** FMEA mind map for the COVID-19 pandemic in the United Kingdom.

The proposed process is given as the community spread of COVID-19 in the United Kingdom. The failure modes consist of imported infected cases, local infected cases, infected items and isolated infected sites in different parts of the United Kingdom. Potential effects include higher mortality rates due to COVID-19, increased isolation effects on the well-being of individuals, higher reproduction numbers and a negative impact on the economy. The following is then assessed:

- How severe are the consequences of these failures? I denote these as the severity rating on a scale from 1 to 10, where 1 is insignificant (or non-existent) and 10 is catastrophic.
- What is the likelihood of infected cases who tests positive? I denote this as the occurrence rating on a scale from 1 to 10, where 1 is remote and 10 is inevitable
- What are the chances of detection of infected cases by the available control measures? I denote this as the detection rating on a scale from 1 to 10, where 1 means a control measure is highly certain to detect it and 10 means highly impossible (or no control exists).

The risk priority number (RPN) in (1.8) can then be used to get an overall rating for each failure mode by combining each of the results above.

Equations (1.8) and (1.9) give relative priority to the failure modes in the order they should be addressed. A high RPN means that there are a lot of new cases and deaths and a corrective action may be needed for the occurrence, severity or detection of individual high ratings. This corrective action should be extended to containment and prevention methods such as national lockdown, asymptomatic testing, sanitization of cities and towns, self-isolation and "stay at home" practises, social distancing, regular hand washing at regular intervals provided by the National Health Service (NHS) or other local health centres in accordance with government requirements, and so on. Applying these methods should significantly reduce the previous RPN value to a smaller value closer to 1. If the RPN reaches 1, then the pandemic has likely ended.

## 2.3 Bayesian method on infectious periods

This section discusses how the Bayesian approach can be applied to the proposed modified SEIR model for infectious periods. In many epidemic cases, a distribution from the exponential family, namely the Poisson and Gamma distributions, fits the shape of the

pandemic well. The proposed model consists of two layers: the likelihood of the data and the inference from prior to posterior on $\mathcal{R}_t$. The Bayesian method is therefore engaged in transforming time series of case numbers into probability distributions.

### 2.3.1 Dynamic Poisson cases

Suppose I may estimate the infectious period to be exponentially distributed with mean $\gamma^{-1}$ and assume it to be constant throughout the duration of the COVID-19 pandemic. I may be interested in calculating $\mathcal{R}_t$, which provides a value on each day in the interval $t \in [1, T]$. I want to see today's posterior to be updated from yesterday's prior. In order to use the Bayes theorem, I must select a prior distribution. The flat or uniform prior is a very common choice for a prior distribution. A flat prior assumes that all probabilities on $f(\mathcal{R}_t)$ are equally likely. Since $\mathcal{R}_t$ could take any value from 0 to infinity, most of our beliefs about $f(\mathcal{R}_t)$ will come through data observation, i.e. the likelihood $\mathcal{L}$. In this case, the density function is uniformly distributed and therefore the prior function is proportional to 1 (or proportional to some constant $c$) which therefore leaves the density function to be more flexible, and general, without violating any principles of probability.

Suppose now that the rate of $j$ new cases per day can be described by a time-dependent Poisson distribution with parameter $\theta_t \equiv f(\mathcal{R}_t)$. By using Bayes' theorem from (1.2) and assuming a flat prior, the posterior probability of $\theta_T$ given seeing $j$ new cases at time $T$ is updated to

$$\pi_*(\theta_T|j_T, \boldsymbol{j}, \boldsymbol{\theta}) \propto \mathcal{L}(j_T, \boldsymbol{j}, \boldsymbol{\theta}|\theta_T) = \prod_{t=1}^{T} \frac{\theta_t^{j_t} e^{-\theta_t}}{j_t!}, \tag{2.10}$$

where $\boldsymbol{j} = \{j_1, j_2, \ldots, j_{T-1}\}$ and $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_{T-1}\}$. For the sake of brevity, I define $\mathcal{D} := \{j_T, \boldsymbol{j}, \boldsymbol{\theta}\}$. If $\theta_t \equiv \theta$, then the posterior PDF is $C_T \theta^{\sum_t j_t} e^{-T\theta}$ and $\theta_T|\mathcal{D} \sim \text{Gamma}(\sum_t j_t - 1, T)$ distributed. Another useful representation of the above can be made by taking the natural log on both sides of the equation:

$$\ln \pi_*(\theta_T|\mathcal{D}) \propto \sum_{t=1}^{T} \big\{ j_t \ln(\theta_t) - \theta_t \big\}. \tag{2.11}$$

Equation (2.11) is used in the subsequent sections.

### 2.3.2 Bayesian inference on Bettencourt & Ribeiro's $\theta$

Bettencourt & Ribeiro [64] derived a relationship linking $\theta_t$, $\gamma$, $j_{t-1}$ and $\mathcal{R}_t$ together by:

$$\theta_t = j_{t-1} \exp\left(\gamma(\mathcal{R}_t - 1)\right) =: j_{t-1} \Psi_1(\mathcal{R}_t) \tag{2.12}$$

Since the $\mathcal{R}_t$ value is dynamic and is more closely linked to recent values than older ones, I can follow Systrom's method [65] by considering the last $(w+1)$ days, that is $t \in [T-w, T]$. From the above, I can derive a posterior distribution for $\mathcal{R}_T$. By taking the last $(w+1)$ points in (2.11) and adapting Bettencourt & Ribeiro's $\theta$ to that equation, the posterior probability is therefore

$$\begin{aligned}\pi_*(\mathcal{R}_T|\mathcal{D}) &\propto \exp\left\{\sum_t \Big(j_t\big(\ln(j_{t-1}) + \gamma(\mathcal{R}_t - 1)\big) - j_{t-1}\exp\left(\gamma(\mathcal{R}_t - 1)\right)\Big)\right\}, \\ &\propto \exp\left\{\sum_t j_t \ln(j_{t-1}) + \sum_t j_t\gamma\mathcal{R}_t - \sum_t j_t\gamma - \sum_t j_{t-1}\exp\big(\gamma(\mathcal{R}_t - 1)\big)\right\}.\end{aligned} \tag{2.13}$$

From (2.13), I find that $\exp\{\sum_t j_t \ln(j_{t-1})\}$ and $\exp\{\sum_t j_t\gamma\}$ are independent from $\mathcal{R}_T$. Since I am interested in $\mathcal{R}_T$, these terms may be taken out of the equation to become one with proportionality, i.e.

$$\begin{aligned}\pi_*(\mathcal{R}_T|\mathcal{D}) &\propto \exp\left\{\sum_t j_t \ln(j_{t-1})\right\}\exp\left\{\sum_t j_t\gamma\right\}\exp\left\{\gamma\sum_t j_t\mathcal{R}_t - \sum_t j_{t-1}\exp\big(\gamma(\mathcal{R}_t - 1)\big)\right\}, \\ &\propto \exp\left\{\gamma\sum_t j_t\mathcal{R}_t - \sum_t j_{t-1}\exp\big(\gamma(\mathcal{R}_t - 1)\big)\right\},\end{aligned} \tag{2.14}$$

Now, if I expand the summations in (2.14), I find that each $\mathcal{R}_t$ term for $t = T-w, \ldots, T-1$ is independent from $\mathcal{R}_T$. Therefore these excess terms can also be removed from the equation to become one with proportionality, thus reducing the posterior PDF to a more compact form:

$$\pi_*(\mathcal{R}_T|\gamma, j_{T-1}, j_T) = C_T^{-1}\exp\left\{j_T\gamma\mathcal{R}_T - j_{T-1}e^{\gamma(\mathcal{R}_T-1)}\right\}, \tag{2.15}$$

where $C_T$ is the constant of integration after integrating the exponential function with respect to $\mathcal{R}_T$. Hence, $C_T$ is equal to

$$C_T = \int_0^\infty \exp\left\{\gamma j_T\mathcal{R}_T - j_{T-1}\Psi_1(\mathcal{R}_T)\right\}d\mathcal{R}_T = \frac{\Gamma(j_T, j_{T-1}e^{-\gamma})}{\gamma(j_{T-1}e^{-\gamma})^{j_T}},$$

where $\Gamma(a, b)$ is the upper incomplete gamma function with $\Gamma(a, b) \to 0$ if $b \to \infty$.

### 2.3.3 Updated $\theta$ for the COVID-19 pandemic

To find an expression that accounts for the evolution of new infections after each day in terms of epidemiological observables, I discretize the infectious ODE equations in (2.4), (2.5) and (2.6), for the change in daily number of cases between $t-1$ and $t$. Note that the cumulative number of exposures, cases and actively confirmed cases up to time $t$, $CE(t)$,

$CI(t)$ and $CQ(t)$, respectively. Epidemic reports most commonly give the frequency of cumulative infected cases, which are reported daily and given by $CI(t) - CI(t-1) = \Delta CI(t) \approx CI'(t)$. The change in daily cumulative exposures and quarantines are also given by $\Delta CE(t) \approx CE'(t)$ and $\Delta CQ(t) \approx CQ'(t)$. I find that $CE(t)$, $CI(t)$ and $CQ(t)$, between $t-1$ and $t$, obeys the following equations:

$$CE'(t) = E(t) \approx \beta S(t) I(t)/N, \tag{2.16}$$

$$CI'(t) = I(t) \approx \nu E(t), \tag{2.17}$$

$$CQ'(t) = Q(t) \approx \varrho I(t). \tag{2.18}$$

Therefore, by substituting (2.16) into (2.17), the change in the cumulative number of daily cases updates to:

$$CI'(t) \approx \nu E(t) \approx \beta\nu S(t) I(t)/N. \tag{2.19}$$

The RHS of equation (2.19) is exact if $S(t)/N$ is constant in the period $[t-1, t]$. Since $S(t)/N \approx 1$ and $\mathcal{R}_0 \approx \mathcal{R}_t$, the ODE in (2.5) which models the change in $I(t)$ updates to:

$$\begin{aligned} I'(t) &\approx \big\{\beta\nu - (\varrho + \zeta + \mu)\big\} I(t), \\ &= \Big\{(\varrho + \zeta + \mu)\Big((\nu + \zeta + \mu)(\gamma + \zeta + \mu)\mathcal{R}_t/\varrho - 1\Big)\Big\} I(t), \\ &= \big\{\phi_1 \mathcal{R}_t - \phi_2\big\} I(t), \end{aligned} \tag{2.20}$$

where $\phi_1 = \phi_2(\nu + \zeta + \mu)(\gamma + \zeta + \mu)/\varrho$ and $\phi_2 = \varrho + \zeta + \mu$. To obtain an equation which accounts for the change in daily new infections, I must integrate (2.20) between $t-1$ and $t$, which gives:

$$I(t) = I(t-1) \exp\big\{\phi_1 \mathcal{R}_t - \phi_2\big\} =: I(t-1)\Psi_2(\mathcal{R}_t). \tag{2.21}$$

Now, by rewriting the equation in terms of $\theta_t$ and $j_{t-1}$, I obtain $\theta_t = j_{t-1}\Psi_2(\mathcal{R}_t)$. Once again, by taking the last $(w+1)$ points in (2.11) and adapting the $\theta$ expression into our $\theta_t$ equation, the posterior probability is therefore:

$$\pi_*(\mathcal{R}_T | \mathcal{D}) = C_T^{-1} \exp\Big\{\phi_1 j_T \mathcal{R}_T - j_{T-1}\Psi_2(\mathcal{R}_T)\Big\}, \tag{2.22}$$

where

$$C_T = \frac{\Gamma\left(j_T, j_{T-1} e^{-\phi_2}\right)}{\phi_1 \exp\left\{-\phi_2 j_T + j_T \ln\left(j_{T-1}\right)\right\}}.$$

### 2.3.4 Prediction

It may also be useful to have a better representation of the uncertainty in $\mathcal{R}_t$ via our posterior. Suppose I am interested in the next case $j_{T+1}$. I can use the posterior predictive

distribution $\pi_*(j_{T+1}|\mathcal{D})$ which is the distribution of a new case $j_{T+1}$, marginalized over the posterior, i.e.

$$\pi_*(j_{T+1}|\mathcal{D}) = \frac{\exp\left\{\phi_2 j_{T+1} + j_T \ln(j_{T-1}) - \Delta^+ j_{T+1} \ln(\Delta^+ j_T) - \ln(j_{T+1}!)\right\}}{\Gamma(j_T, j_{T-1}\exp\{-\phi_2\}) / \Gamma(\Delta^+ j_{T+1}, \Delta^+ j_T \exp\{-\phi_2\})}, \tag{2.23}$$

where $\Delta^+ j_T = j_T + j_{T-1}$.

*Proof.* The posterior predictive distribution of $j_{T+1}$, given the data of $\{j_T, j_{T-1}, \phi_1, \phi_2\}$, marginalised over the posterior, is formally solved using:

$$\pi_*(j_{T+1}|\mathcal{D}) = \int_0^\infty p(j_{T+1}|\mathcal{R}_T, j_T, \phi_1, \phi_2)\pi_*(\mathcal{R}_T|j_T, j_{T-1}, \phi_1, \phi_2)\, d\mathcal{R}_T. \tag{2.24}$$

Recall that the posterior distribution takes the form of that seen in (2.22) and that the PDF governing the distribution of $j_{T+1}|\mathcal{R}_T, j_T, \phi_1, \phi_2$ is a time-dependent Poisson with rate $j_T \exp(\phi_1 \mathcal{R}_T - \phi_2)$ and has a density function equal to:

$$p(j_{T+1}|\mathcal{R}_T, j_T, \phi_1, \phi_2) = \exp\{j_{T+1} \ln(j_T) + j_{T+1}(\phi_1 \mathcal{R}_T - \phi_2) - j_T \Psi_2(\mathcal{R}_T) - \ln(j_{T+1}!)\}.$$

By combining the PDFs into (2.24) and ignoring the non-$\mathcal{R}_T$ terms, we have:

$$\begin{aligned}
\pi_*(j_{T+1}|\mathcal{D}) &= C_{T+1} \int_0^\infty \exp\left\{\phi_1 \Delta^+ j_{T+1} \mathcal{R}_T - \Delta^+ j_T \Psi_2(\mathcal{R}_T)\right\} d\mathcal{R}_T, \\
&= C_{T+1} \frac{1}{\phi_1} E_{1-\Delta^+ j_{T+1}}\left[\Delta^+ j_T e^{-\phi_2}\right], \\
&= C_{T+1} \frac{1}{\phi_1} \left(\Delta^+ j_T e^{-\phi_2}\right)^{-\Delta^+ j_{T+1}} \Gamma\left(\Delta^+ j_{T+1}, \Delta^+ j_T e^{-\phi_2}\right),
\end{aligned}$$

where $\Delta^+ j_T = j_T + j_{T-1}$ is the one-step summation function, $C_{T+1}$ is a constant equal to

$$C_{T+1} = \frac{\phi_1 \exp\left\{-\phi_2 \Delta^+ j_{T+1} + j_T \ln(j_{T-1}) + j_{T+1} \ln(j_T) - \ln(j_{T+1}!)\right\}}{\Gamma(j_T, j_{T-1}\exp\{-\phi_2\})},$$

and $E_n[x]$ is an exponential integral function equal to $x^{n-1}\Gamma(1-n, x)$. Putting everything together, we obtain the same result in (2.23). Thus, the proof is complete. □

The data $\mathcal{D}$ is represented by the history of $\{j_T, j_{T-1}, \phi_1, \phi_2\}$. The posterior predictive distribution has the same mean as the posterior distribution, but a greater variance which takes into account the additional "sampling uncertainty" since I am drawing new data points. Often the $\pi_*(j_{T+1}|\mathcal{D})$ form can be obtained directly, but it is always simpler to analyze $\pi_*(j_{T+1}|\mathcal{D})$ using Markov chain Monte Carlo (MCMC) methods.

## 2.4 Analysis, results and discussion

In this section, I will test our Bayesian methods above applied to the coronavirus outbreak in Europe, namely, a model validation with selected countries (France, Germany, Italy and Portugal), and a Bayesian analysis according to the pandemic situation in the United Kingdom, to see what sort of outcomes may be observed. Recall that $\beta$ is the probability of transmitting the disease multiplied by the number of contacts per unit time. Essentially, the reduction of $\beta$ (and therefore $\mathcal{R}_t$) means that the peak decreases in intensity but shifts to later periods for $\mathcal{R}_t$ greater than 1.

Data from the UK COVID Symptom Study indicates that while most individuals recover from COVID-19 within two weeks, one in 10 individuals may still have symptoms after three weeks and some may suffer for months. The current official estimated range for the recovery period is within 14 days ($1/14 \leq \gamma \leq 1$), whereas the mean period of incubation observed appears to be 5-6 days ($1/6 \leq \nu \leq 1/5$) [66, 67]. In this case, $\gamma$ and $\nu$ could vary significantly among patients. For mathematical simplicity, I set the recovery and incubation rate to be the same and equal to 1/5 (unless stated otherwise).

The results include parameter estimation, model validation, fatality analysis, the impact of governmental protective measures and outbreak forecasting. Integrals and other computations were computed in Mathematica and MATLAB, and converted to graphs and tables in Microsoft Excel.

### 2.4.1 Parameter analysis and model validation for the COVID-19 pandemic for Western European countries

Here, by comparing the Bayesian posterior averages with the actual data, I aim to justify the modified SEIR model using the posterior functions in (2.15) and (2.22) for forecasting the COVID-19 outbreak across selected countries in Western Europe. Example calculations for model (2.15) are given. A real-time count of (smoothed) new cases is extracted from Roser *et al.* [68] in real-time. Here, I set $\gamma = 1/5$ and $T = 1, \ldots, 4$ to represent the period of 1-4 March 2020. For simplicity, if $T = 2$, then the posterior PDF in (2.15) reduces to $\pi_*(\mathcal{R}_2|\gamma, \boldsymbol{j}) = C_2^{-1} \exp\left\{j_2\gamma(\mathcal{R}_2 - 1) - j_1 e^{\gamma(\mathcal{R}_2-1)}\right\}$, where $C_2$ is a constant. Here, the value of $j_1$ and $j_2$ represents the number of cases (after data smoothing) in times 1 and 2, respectively (i.e. the first two days of March). In this example, the focus is on the COVID-19 pandemic in two countries: Germany and Italy. For the case of Italy, it was observed that $j_1 = 149.86$ and $j_2 = 222.43$, and so $C_2 = 2.97131 \times 10^{-59}$. The posterior mean,

median, mode and variance are therefore found to be $\mathbb{E}[\mathcal{R}_2|\mathcal{D}] = 2.963$, $\mathbb{Q}_{0.5}[\mathcal{R}_2|\mathcal{D}] = 2.975$, $\mathbb{M}[\mathcal{R}_2|\mathcal{D}] = 2.967$ and $\mathbb{V}[\mathcal{R}_2|\mathcal{D}] = 0.113$, respectively, and the posterior probability that the epidemic will drop below 1, given the last two days, is $1.86401 \times 10^{-8}$. The empirical $\mathcal{R}_t$ is the result of (2.12) when $\theta_t = j_t$ and $\mathcal{R}_t = \gamma^{-1} \ln(j_t/j_{t-1}) + 1$. In this case, the empirical result of $\mathcal{R}_2$ is equal to 2.975, which is strongly consistent with the posterior averages.

Table 2.1 presents the number of new cases in Germany and Italy over four days with the respective posterior means, medians, modes and other vital statistics. The posterior averages are in agreement with the empirical result.

| Time $t$ | Today's cases $j_t$ | Yesterday's cases $j_{t-1}$ | Empirical result of $\mathcal{R}_t$ | Constant on integration $\mathcal{C}_t$ | Posterior Mean of $\mathcal{R}_t$ | Posterior Median of $\mathcal{R}_t$ | Posterior Mode of $\mathcal{R}_t$ | Posterior Var. of $\mathcal{R}_t$ | Posterior SD of $\mathcal{R}_t$ | $\mathbb{P}(\mathcal{R}_t < 1)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| (a) Community spread of COVID-19 in Germany: Early March | | | | | | | | | | |
| 1 Mar 2020 | 13.71 | 6.00 | 5.133 | $3.12 \times 10^{-1}$ | 4.948 | 5.128 | 5.007 | 1.870 | 1.367 | 0.0040 |
| 2 Mar 2020 | 16.29 | 13.71 | 1.859 | $4.02 \times 10^{-6}$ | 1.953 | 1.869 | 1.908 | 1.076 | 1.037 | 0.2025 |
| 3 Mar 2020 | 20.29 | 16.29 | 2.098 | $3.50 \times 10^{-7}$ | 2.093 | 2.097 | 2.078 | 0.996 | 0.998 | 0.1512 |
| 4 Mar 2020 | 25.57 | 20.29 | 2.158 | $7.01 \times 10^{-9}$ | 2.124 | 2.160 | 2.125 | 0.857 | 0.926 | 0.1225 |
| (b) Community spread of COVID-19 in Italy: Early March | | | | | | | | | | |
| 1 Mar 2020 | 149.86 | 124.43 | 1.930 | $1.07 \times 10^{-53}$ | 1.913 | 1.930 | 1.919 | 0.167 | 0.409 | 0.0146 |
| 2 Mar 2020 | 222.43 | 149.86 | 2.975 | $2.97 \times 10^{-59}$ | 2.963 | 2.975 | 2.967 | 0.113 | 0.336 | $1.86 \times 10^{-8}$ |
| 3 Mar 2020 | 258.14 | 222.43 | 1.745 | $2.99 \times 10^{-96}$ | 1.735 | 1.744 | 1.738 | 0.097 | 0.312 | $1.04 \times 10^{-2}$ |
| 4 Mar 2020 | 311.43 | 258.14 | 1.938 | $9.60 \times 10^{-111}$ | 1.930 | 1.938 | 1.933 | 0.080 | 0.284 | $7.06 \times 10^{-4}$ |

**Table 2.1:** Empirical and posterior summary statistics for the community spread of COVID-19 pandemic using model (2.15). Coverage: 1-4 March 2020 in Germany and Italy.

The posterior PDF of this distribution, as well as three follow up cases in Italy, $j_3 = 258.14$, $j_4 = 311.43$, and $j_5 = 384.14$ with respective $\mathcal{R}_T$ for $T = 3, 4$ and 5, have been illustrated in Figure 2.5.

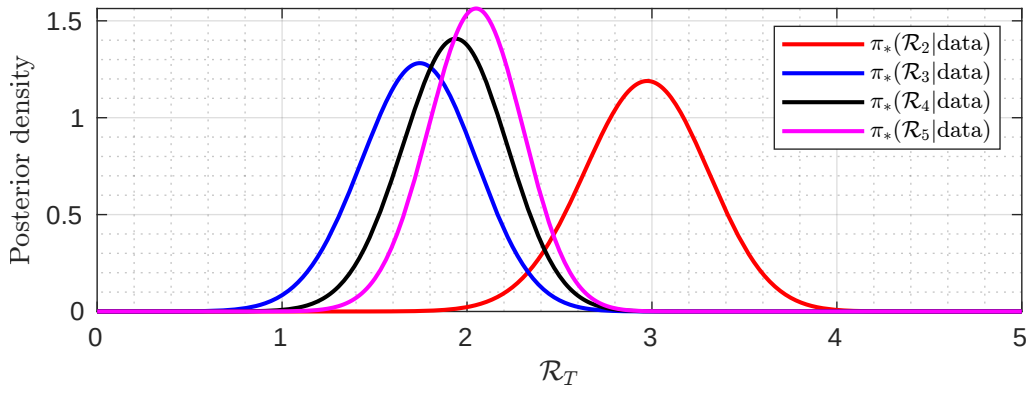



**Figure 2.5:** Real-time data for the variation of $\pi_*(\mathcal{R}_T|\mathcal{D})$ for $\mathcal{R}_T$ over 5 days with $\gamma = 1/5$. Coverage: Italy from 1-5 March 2020.

The equation derived in (2.21) is an expression accounting for the evolution of new infections. Example calculations can be seen in Figure 2.6 for the variation of $\Psi_1(\mathcal{R}_t)$ and $\Psi_2(\mathcal{R}_t)$ for $\mathcal{R}_t$ with empirically estimated parameters describing COVID-19 outbreak in the United Kingdom. I find that $\Psi_2(\mathcal{R}_t)$ is in strong agreement with $\Psi_1(\mathcal{R}_t)$. These parameter values are related to different methods and should only be seen as a qualitative benchmark.

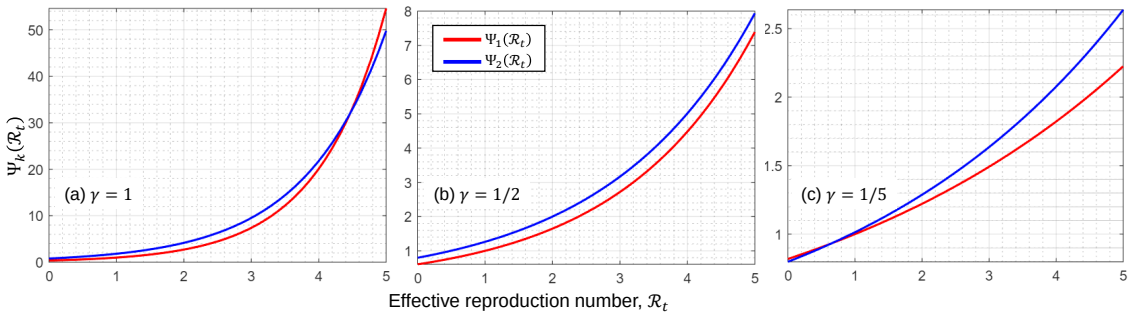



**Figure 2.6:** The variation of $\Psi_1(\mathcal{R}_t)$ and $\Psi_2(\mathcal{R}_t)$ for $\mathcal{R}_t$ with empirically estimated parameters describing COVID-19 outbreak: $\gamma = \{1, 1/2, 1/5\}$ (both models), $\nu = 1/5$, $\varrho = 1/10$, $\zeta = 0.116$ and $\mu = 9.4/1000$. Coverage: 1-11 September 2020 in the United Kingdom.

Example calculations on (2.22) are given in Table 2.2, which presents the number of new cases in four countries over three different periods. The respective posterior means, medians, modes, variances and 90% credibility intervals are given. For each period, I set $\nu = 1/5.2$, $\varrho = 1/7$, $\gamma = 1/5$ and $\zeta = 0.15$. I then matched the empirical mortality rate to each country (e.g. $\mu = 9.4/1000$ in the United Kingdom) so that $\phi_1 \approx 7(0.293+\mu)(0.342+\mu)(0.350+\mu)$. The posterior averages are strongly in line with the empirical outcome. The empirical result of $\mathcal{R}_t$ is the solution of (2.21) and is approximately equal to the posterior mode.

| Region | Dates $(T-1, T)$ | Cases $j_T$ | Cases $j_{T-1}$ | $\sum_{i=1}^{T} j_T$ | Empirical $\mathcal{R}_T$ | $\mathbb{E}[\mathcal{R}_T]$ | $\mathbb{Q}_{0.5}[\mathcal{R}_T]$ | $\mathbb{M}[\mathcal{R}_T]$ | $\mathbb{V}[\mathcal{R}_T]$ | 90% Credibility |
|---|---|---|---|---|---|---|---|---|---|---|
| United Kingdom | 22-23 Aug 2020 | 1,022.4 | 992.3 | 327,643 | 1.195 | 1.198 | 1.203 | 1.215 | 0.130 | (0.593, 1.784) |
| | 6-7 Sep 2020 | 1,812.1 | 1,630.3 | 352,451 | 1.483 | 1.483 | 1.486 | 1.493 | 0.074 | (1.029, 1.923) |
| | 12-13 Sep 2020 | 3,001.4 | 2,760.9 | 370,930 | 1.430 | 1.436 | 1.438 | 1.443 | 0.047 | (1.078, 1.788) |
| France | 22-23 Aug 2020 | 718.9 | 708.7 | 280,459 | 1.186 | 1.182 | 1.183 | 1.182 | 0.020 | (0.951, 1.411) |
| | 6-7 Sep 2020 | 1,294.3 | 1,261.9 | 367,174 | 1.369 | 1.364 | 1.368 | 1.378 | 0.109 | (0.813, 1.898) |
| | 12-13 Sep 2020 | 1,809.6 | 1,691.1 | 402,893 | 1.377 | 1.375 | 1.378 | 1.385 | 0.078 | (0.910, 1.827) |
| Portugal | 22-23 Aug 2020 | 210.1 | 204.0 | 55,597 | 1.219 | 1.220 | 1.223 | 1.229 | 0.065 | (0.796, 1.634) |
| | 6-7 Sep 2020 | 355.7 | 356.4 | 60,507 | 1.111 | 1.108 | 1.109 | 1.113 | 0.038 | (0.783, 1.426) |
| | 12-13 Sep 2020 | 481.0 | 479.4 | 63,983 | 1.127 | 1.129 | 1.130 | 1.132 | 0.028 | (0.850, 1.403) |
| Germany | 22-23 Aug 2020 | 1,233.0 | 1,223.9 | 234,494 | 1.138 | 1.133 | 1.138 | 1.147 | 0.110 | (0.579, 1.671) |
| | 6-7 Sep 2020 | 1,202.6 | 1,173.4 | 253,626 | 1.201 | 1.198 | 1.203 | 1.213 | 0.113 | (0.636, 1.743) |
| | 12-13 Sep 2020 | 1,349.0 | 1,354.7 | 261,737 | 1.079 | 1.091 | 1.096 | 1.105 | 0.100 | (0.562, 1.606) |

**Table 2.2:** Posterior summary statistics for the COVID-19 pandemic using model (2.22). Coverage: United Kingdom, France, Portugal and Germany regions, updated on 14 September 2020.

### 2.4.2 Assessing the impact of English government protective measures to COVID-19

The main objective of government protection policies and lockdowns is to reduce the rates of infection so as to prevent the health care system from being overloaded by too many serious cases at the same time. In order to research the impact of the protection measures adopted by the United Kingdom government, I compare the confirmed (relative) infection rate, the cumulative infections in a population $CI(t)$, the empirical reproduction number $\mathcal{R}_t$, and

the posterior averages of $\mathcal{R}_t$, at different time periods with and without the measures being implemented. The NHS is a publicly funded health care system in the United Kingdom. The primary responsibility for public health and healthcare in the United Kingdom rests with the Department of Health and Social Care (DHSC). The periodic infection rate in the United Kingdom and how it varies over the course of time between early March and late October are shown in Table 2.3. The (relative) infection rate is expressed as the daily change in infections as a percentage of the value of the indicator in the earlier periods, i.e.

$$\text{Relative infection rate} = \frac{\Delta I(t)}{I(t-1)}, \tag{2.25}$$

where $\Delta I(t) = I(t) - I(t-1)$. The comparison of the estimated rates of infection in the corresponding periods before and after the introduction of preventive measures highlights the significance and reliability of measures such as social distancing and lockdowns in managing and slowing down the spread of COVID-19. As of 31 October 2020, there has been 989,745 and 42,592 confirmed cases of infection and death, respectively.

| | No measures applied. | | First round of lockdown applied across the UK. | | | | First round of lockdown measures lifted. | International travel quarantine applied. | Primary age children return to school post lockdown. | Non-essential retail reopened following guidelines. | Travel corridors opens and pubs, restaurants and bars reopened. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Time interval $(t_i, t_j)$ | Mar. 1-7 | Mar. 8-22 | Mar. 23-31 | Apr. 1-15 | Apr. 16-30 | May 1-10 | May 11-21 | May 22-31 | June 1-14 | June 15-30 | July 1-23 |
| Daily change in infections | 21.7% | 16.5% | 8.3% | -3.4% | -0.1% | -6.5% | -1.9% | -4.5% | -1.0% | -5.5% | 1.5% |
| Infections at $t_{j-1}$ | 28.29 | 673.57 | 2,555.57 | 4,586.00 | 4,614.57 | 3,659.00 | 2,624.14 | 1,692.00 | 1,016.43 | 697.29 | 632.71 |
| Infections at $t_j$ | 34.43 | 784.57 | 2,767.00 | 4,428.29 | 4,610.14 | 3,419.86 | 2,574.14 | 1,615.57 | 1,006.71 | 658.86 | 642.14 |
| Cumulative infections | 271 | 7,736 | 29,681 | 97,052 | 167,150 | 206,234 | 235,727 | 254,390 | 270,285 | 284,888 | 297,659 |
| Empirical $\mathcal{R}_t$ | 1.983 | 1.763 | 1.397 | 0.825 | 0.995 | 0.662 | 0.904 | 0.769 | 0.952 | 0.717 | 1.074 |
| Posterior prob of $\mathcal{R}_t \leq 1$ | 13.0% | 9.8% | 9.9% | 77.5% | 51.4% | 89.9% | 62.8% | 72.1% | 53.3% | 63.9% | 43.6% |
| Posterior mean of $\mathcal{R}_t$ | 1.984 | 1.764 | 1.404 | 0.813 | 0.990 | 0.650 | 0.894 | 0.779 | 0.973 | 0.846 | 1.121 |
| Posterior median of $\mathcal{R}_t$ | 1.986 | 1.773 | 1.407 | 0.814 | 0.991 | 0.648 | 0.895 | 0.768 | 0.959 | 0.800 | 1.098 |
| Posterior mode of $\mathcal{R}_t$ | 2.015 | 1.793 | 1.413 | 0.813 | 0.989 | 0.649 | 0.900 | 0.760 | 0.950 | 0.705 | 1.077 |

| | Mandatory face coverings in shops across England. | | | | "Rule of six" applied in England. Increased fines up to £10,000 added shortly afterwards. | | Tighter restrictions are put in place across the UK, with pubs ordered to close at 10pm. | | A new three-tier system for COVID restrictions in England. | | Preparing for second round of lockdown in Nov 2020. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Time interval $(t_i, t_j)$ | July 24-31 | Aug. 1-15 | Aug. 16-31 | Sep. 1-13 | Sep. 14-18 | Sep. 19-21 | Sep 22-30 | Oct 1-11 | Oct 12-20 | Oct 21-29 | Oct 30-31 |
| Daily change in infections | 1.9% | 8.4% | 8.3% | 8.7% | 2.1% | 2.3% | 5.5% | 2.1% | 4.1% | -1.3% | 2.5% |
| Infections at $t_{j-1}$ | 542.29 | 970.43 | 1,164.43 | 2,760.86 | 3,285.71 | 3,597.71 | 5,769.71 | 15,504.71 | 16,956.14 | 22,147.57 | 22,124.71 |
| Infections at $t_j$ | 552.71 | 1,051.71 | 1,260.71 | 3,001.43 | 3,353.71 | 3,679.00 | 6,086.43 | 15,832.43 | 17,646.29 | 21,863.71 | 22,678.14 |
| Cumulative infections | 302,301 | 316,367 | 334,467 | 365,174 | 381,614 | 394,257 | 446,156 | 590,844 | 741,212 | 942,275 | 989,745 |
| Empirical $\mathcal{R}_t$ | 1.095 | 1.402 | 1.397 | 1.418 | 1.102 | 1.112 | 1.267 | 1.105 | 1.199 | 0.936 | 1.124 |
| Posterior prob of $\mathcal{R}_t \leq 1$ | 42.1% | 21.4% | 19.6% | 8.0% | 36.1% | 34.1% | 9.8% | 40.4% | 30.7% | 58.2% | 36.3% |
| Posterior mean of $\mathcal{R}_t$ | 1.157 | 1.400 | 1.395 | 1.426 | 1.099 | 1.109 | 1.274 | 1.099 | 1.195 | 0.927 | 1.118 |
| Posterior median of $\mathcal{R}_t$ | 1.129 | 1.404 | 1.400 | 1.429 | 1.101 | 1.111 | 1.275 | 1.100 | 1.198 | 0.927 | 1.121 |
| Posterior mode of $\mathcal{R}_t$ | 1.099 | 1.418 | 1.413 | 1.434 | 1.107 | 1.116 | 1.109 | 1.109 | 1.207 | 0.933 | 1.128 |

**Table 2.3:** Comparisons of the daily rate of infection, the overall cumulative infected population, the empiric reproduction number and the posterior averages to the reproduction number at the end of the various time periods with and without the United Kingdom protective measures in 2020.

The forecast of the situation in the United Kingdom, based on the Bayesian method (Figure 2.7), appears reasonably positive, as the (posterior median) $\mathcal{R}_t$ values follows a quartic[1] trend. Table 2.3 and Figure 2.7 are in strong agreement with each other and represents the most probable scenario. According to a statistical study conducted in

[1] A polynomial of order 4.

September 2020 [69], the number of positive cases doubles every seven to eight days, with a total of 3,539 new coronavirus cases were recorded in the United Kingdom on 10 September. Scientists have estimated the empirical $\mathcal{R}_t$ value to be at 1.7 which appears to agree with our posterior estimates.

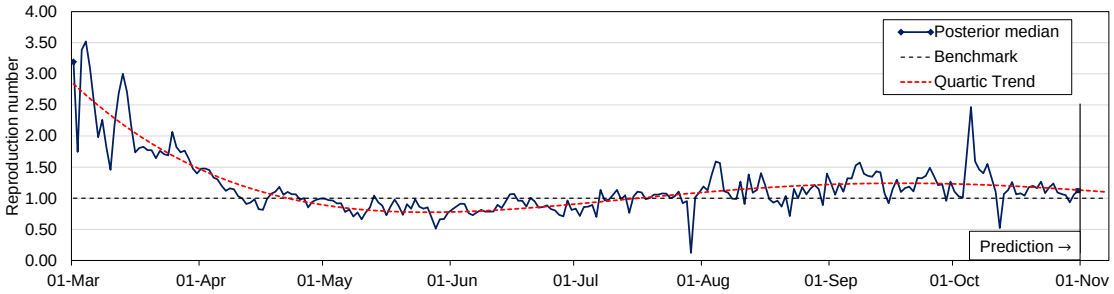



**Figure 2.7:** The continuous red dotted lines refer to the predicted evolution in 7 days according to the modified SEIR model. Coverage: United Kingdom region, updated on 31 October 2020.

In countries without public health interventions, such as Sweden, the demand for an intensive care unit (ICU) is estimated to be much higher than the intensive care capacity in other western countries, resulting in a much larger, predicted $\mathcal{R}_t$ [70]. In the United Kingdom, the NHS is reported to have used half as many ICU beds as France, Belgium and other badly-hit European nations during height of pandemic in May 2020. This is equivalent to 50 infected patients attached to the ventilators for every million people in mid-April. In Belgium, where there were deaths at a similar rate to the United Kingdom at the time, the figure was approximately 111 per million people. During the same week, France treated 104 people per million in intensive care [71]. With 2.7 hospital beds per 1,000 citizens, the United Kingdom has fewer hospital beds than other European nations, compared with an European Union (EU) average of 5.2 [72].

### 2.4.3 Fatality analysis

In order to slow down the outbreak and therefore prevent an increase in deaths, social distances and other preventive measures were implemented in the United Kingdom around 23 March 2020. The daily deaths in the United Kingdom as well as the relative shift in deaths are shown in Figure 2.8. The relative change in deaths mirrors equation (2.25) and is given by $\Delta D(t)/D(t-1)$ where $\Delta D(t) = D(t) - D(t-1)$.

Another formula of interest in epidemiology is the naïve case fatality rate (CFR), which is the proportion of COVID-19 cumulative deaths over the total number of confirmed cases over a specified time period. This formula provides an overview of the severity of the disease

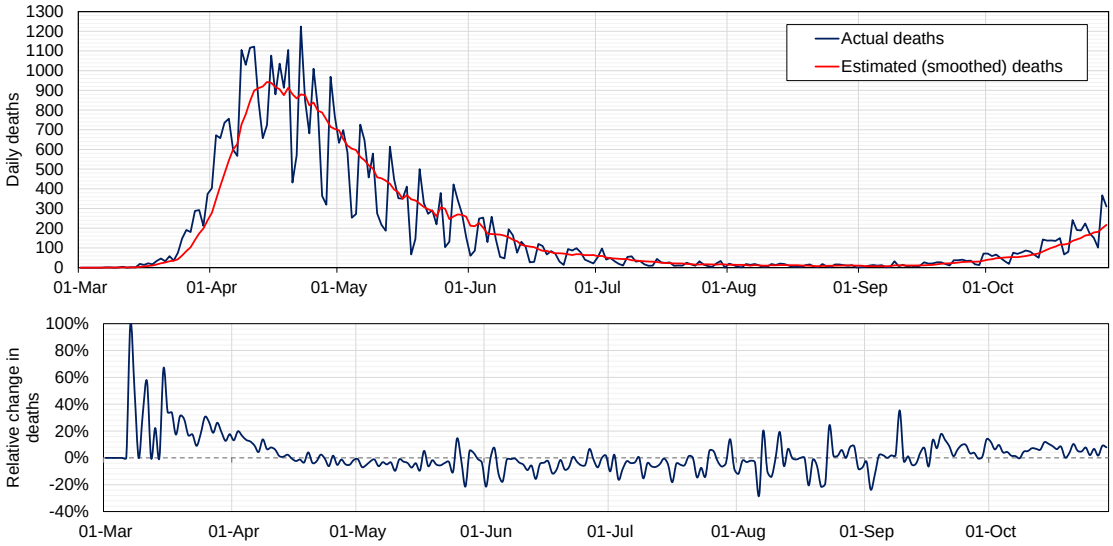



**Figure 2.8:** The daily deaths (top) and the relative shift in deaths (bottom) in the United Kingdom.

and is given as:

$$\text{CFR} = D(t)/I(t). \tag{2.26}$$

For this study, I apply the formula in (2.26) and present a fatality analysis that is important for government protection decision-making. Simple calculations using data recorded in the United Kingdom indicate that on 16 April, the CFR initially increased sharply and peaked at an all-time high of 21.5%, and has since decreased with varying degrees of drops (with a few significant increases around June and July). The CFR rate fell below the "safety net" (2% level) for the first time on 4 August and has since remained below that level. The daily CFR rate in the United Kingdom is 0.3 percentage points lower than the global average as of 31 October 2020, at 1.0% and 1.3%, respectively.

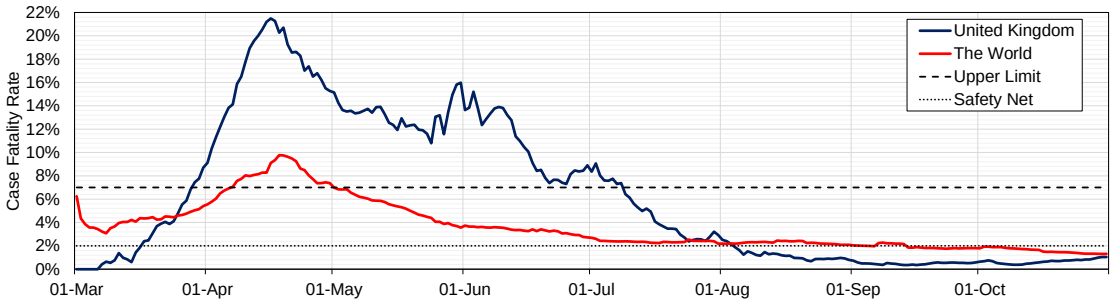



**Figure 2.9:** Case fatality rate in the United Kingdom and the world calculated using (2.26).

Infection, reproduction and fatality rates (see Table 2.3 and Figure 2.8) depend on the preventive measures in place, the number of fatalities may rise due to a large number of infected individuals if the government safety measures are ignored. The rate of infection has been higher in care homes than in the community and there has been significant regional

variation in the severity of the outbreak. This, in fact, reduces the number of ICU beds available as the number of intensive cases rises drastically.

### 2.4.4 FMEA of the community spread of COVID-19 in the United Kingdom

The number of COVID-19 cases, fatalities and reproduction numbers have an impact on the population's well-being and on the economy as a whole. An FMEA could be used to shed some light on the direction the government has taken and inform the future decision-making. Control measures such as national lockdown or self-isolation are effective in reducing the Likelihood of Detection (i.e. the probability for the failure to reach the individual) and thus reducing the severity of case fatalities and decreasing the Rate of Occurrence of positive infections in order to prevent failures. To demonstrate the use of an FMEA, a 10-point severity, occurrence and detection scale to represent the community spread of COVID-19 in the United Kingdom is presented in Table 2.4.

| Ranking | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Severity | Insignificant | Very minor | Minor | Low | Moderate | Moderately high | High | Very high | Extremely high | Critical |
| Criteria S: Case fatality rate | ≈ 0% | < 0.1% | 0.1-0.3% | 0.3-0.5% | 0.5-1.0% | 1.0-1.5% | 1.5-2.0% | 2.0-2.5% | 2.5-3.0% | > 3% |
| Rate of Occurrence | Very remote | Remote | Very low | Low | Moderately low | Moderate | Moderately high | High | Very high | Extremely high |
| Criteria O: Expected number of individuals infected from a single infectious person (using posterior median of the reproduction number) | ≈ 0 | < 0.1 | 0.1-0.2 | 0.2-0.4 | 0.4-0.7 | 0.7-1.0 | 1.0-1.3 | 1.3-1.6 | 1.6-2.0 | > 2 |
| Likelihood of Detection/Prevention | Failure prevented | Failure likely to be prevented | | Control measures are working | | | Low levels of detection/prevention | | Not likely to detect | No detection opportunity |
| Criteria D: Control measures (affecting the whole population) and rate of detection/prevention | Vaccine is produced with 90%+ success rate | Nationwide lockdown applied in full force | Nationwide lockdown gradually applied | Self-isolation and "stay at home" practices | Practice social distancing | Sanitization of cities and towns | Mandatory face coverings | Regular hand washing | Prevention measures not practiced regularly | None |

**Table 2.4:** A ten point severity, occurrence and detection scale with a list of criterion to reflect the COVID-19 pandemic in the United Kingdom. Updated on 31 October 2020.

Each criterion is ranked from 1 (insignificant / remote / failure prevented) to 10 (critical / extremely high / no detection opportunity). For the analysis of the community spread of COVID-19 in the United Kingdom (from March to October 2020), I matched Criteria S (CFR rate) to Figure 2.9, Criteria O (posterior median of $\mathcal{R}_t$) to Figure 2.7, and Criteria D (control methods) to the government policies defined in Table 2.3 for the given time intervals. The RPN and criticality values are calculated using equations (1.8) and (1.9), respectively. For the sake of brevity, I scaled criticality by a factor of 10 to match the RPN value. A graphical summary of our findings on the severity, occurrence and detection of COVID-19 throughout the United Kingdom (from 1 March to 31 October) is provided in Figure 2.10.

In our FMEA, I focused on the COVID-19 infections and thus the failure mode is taken

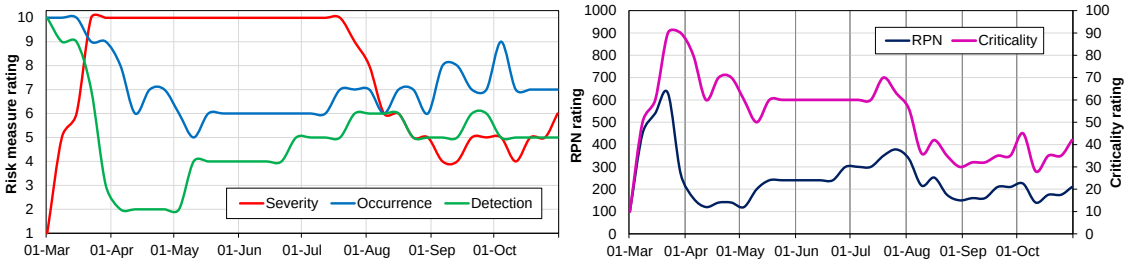



**Figure 2.10:** A visual summary of a 10-point FMEA rating (left) and the criticality/RPN (right). Results reflect the United Kingdom COVID-19 pandemic by following the timeline on Table 2.3 and the evaluation criteria on Table 2.4.

as positive cases. In addition, the process phase, cause of failures, current control measures and risk measures are summarised in Table 2.5, with the subsequent government responses and re-evaluated risk measures are presented in Table 2.6.

| Process | Cause of failure | Current measures | Sev | Occ | Det | RPN* |
|---|---|---|---|---|---|---|
| First cases<br>(22 Jan to 27 Feb) | • Failure to shut down international airlines due to imported infections from outside the UK. | • None | 1 | 10 | 10 | 100 |
| First deaths and early spread<br>(28 Feb to 23 Mar) | • No social distancing for more than 15 minutes at distances of less than 2 metres. | • Hand washing (not practiced regularly)<br>• Social distancing (not practiced regularly) | 9 | 9 | 7 | 567 |
| First lockdown<br>(23 Mar to 4 May) | • Delayed lockdown against the rising pandemic. | • All of the above.<br>• Nationwide lockdown applied in full force. | 10 | 9 | 3 | 270 |
| Easing down lockdown measures and continued restrictions<br>(5 May to 31 May) | • Failure to practice safety measures once travel corridors opens and pubs, shops, restaurants and bars had reopened.<br>• A vast number of sharing of potentially contaminated spaces. | • Sanitization of cities and towns<br>• Self-isolation and "stay at home" practices<br>• Regular hand washing and social distancing<br>• Wearing face masks | 10 | 6 | 2 | 120 |
| Continued restrictions and local lockdowns<br>(1 June to 5 Sep) | • Failure to follow policies and protocols set by the government. | • Self-isolation and "stay at home" practices<br>• Regular sanitization, social distancing and wearing face masks | 10 | 6 | 4 | 240 |
| Resurgance<br>(6 Sep to 31 Oct) | • Overcrowed restaurants during the month-long "Eat Out to Help Out" scheme in August.<br>• Reopening of schools in England, Wales and Northern Ireland for the autumn term. | • Same as the above.<br>• Curfews set across the country.<br>• Tighter restrictions are put in place across the UK. | 5 | 8 | 5 | 200 |

**Table 2.5:** Summary of the process phase, cause of failures, current control and risk measures. The subsequent government responses and re-evaluated risk measures are discussed in Table 2.6. *Risk measures calculated by averaging over the first 7 days.

It is estimated that some of the main causes of the failure were government responses, namely, failure to shut down international airlines due to imported infections from outside the UK (22 Jan – 27 Feb), no social distancing for more than 15 minutes at distances of less than 2 metres (28 Feb – 23 Mar), delayed lockdown against the rising pandemic (23 Mar – 4 May), failure to practice safety measures once travel corridors opens and pubs, shops, restaurants and bars had reopened (5 May – 31 May), failure to follow policies and protocols set by the government (1 June – 5 Sep) and overcrowded restaurants during the month-long 'Eat Out to Help Out' scheme and the re-opening of schools (6 Sep – 31 Oct). The effect of failure for each process phase is discussed in the following paragraphs.

The failure to shut down international airlines affected the United Kingdom critically in

the first process phase, as the national risk level for COVID-19 was raised from “very low” to “low” by Public Health England (PHE) on 22 January. As a result of this failure, British nationals had to move from Wuhan to quarantine at Arrowe Park Hospital. All airports in the United Kingdom were required to provide written guidance to unwell travellers. As no current measures were in place, the first two cases of COVID-19 were confirmed in the United Kingdom on 31 January.

In the second process phase, failure modes were the first deaths and early spread of the disease across the country. The effect of these failures has led to the slow spread of COVID-19 across the country. The number of confirmed cases rose to 23 on 29 February, after 10,483 people had been tested and two of the most recent cases had recently returned from Italy, while the third had returned from Asia. Subsequently, the United Kingdom’s death toll was 55, with the number of cases of COVID-19 passing 1,500 by 16 March. The increased severity was due to the very high death rates that occurred.

In the third phase, the national lockdown came to full effect across the country and all non-essential shops, libraries, places of worship, playgrounds and outdoor gyms were closed by 23 March. The government’s failure to respond strongly to the rising pandemic had raised the risk level from "moderate" to "high". Too much attention was also given to hospital deaths, not including those in care homes or the home of a person. The effect of these previous failures has led to a large increase in deaths, bringing the total number of deaths to 1,019 by 28 March; 17,089 individuals tested positive. Following the lockdown, there was a significant increase in anxiety and depression among the UK population and the number of people infected in the hospital exceeded 10,000 by 31 March. In addition, less than a week later, the death toll was more than 5,000 and the total number of reported cases was almost 52,000 on 6 April. An additional 823 deaths were reported on 21 April, which amounted to a total of 17,337, a sharp increase in previous days, but many of them were related to deaths that occurred in previous days and weeks, some of which date back to March.

Following the easing of lockdown measures and continued restrictions in the fourth phase, the government’s scientific guidance suggested that the $\mathcal{R}_t$ number had risen marginally from 0.5–0.9 at the end of the lockdown to 0.7–1.0 on 15 May, closer to the point at which infections would begin to escalate exponentially. Fittingly, these values are in line with our posterior median estimates of 0.66–0.96 and 0.88–0.94, respectively. Restrictions continued and local lockdowns were introduced in the fifth phase as infection and death rates continued to rise. Wearing face masks were made compulsory for public transport (15

June), in shops and supermarkets in Scotland (10 July) and England (24 July) as well as indoor settings, such as cinemas and places of worship (31 July). These regulations may have significantly contributed to the reduction of the failure associated with the practise of safety measures once travel corridors and non-essential venues were reopened. However, the daily number of new cases exceeded 1,000 on 9 August for the first time since June, rising from 1,062 to 310,825, but it is not apparent whether the increase is due to higher infection rates or higher test volumes.

In the last phase, many of the main causes of the failure were government responses from the fifth phase, namely overcrowded restaurants during the month-long 'Eat Out to Help Out' scheme in August and the re-opening of schools in England, Wales and Northern Ireland for the autumn term. The effect of these failures is likely to have led to a significant increase in cases since September. I find that an additional 2,988 cases were reported on 6 September, the highest number since 22 May, with an estimated a posterior median of $\mathcal{R}_t$ of 1.32. As government researchers warn that the virus is widespread across the nation, the $\mathcal{R}_t$ number floated between 1.1–1.4 on 18 September, which is consistent with our posterior median estimate of 1.10–1.30. There were 6,178 new cases on 23 September, the highest daily reported since 1 May, and the posterior median of $\mathcal{R}_t$ was 1.32. Cases continue to rise and, three weeks later, on 16 October, there were 27,900 new cases in England, a 60% increase over the previous week, while the reported $\mathcal{R}_t$ was between 1.3–1.5 and our posterior median estimate of $\mathcal{R}_t$ was between 1.26–1.55. Finally, the United Kingdom crossed a million cases of COVID-19 on 31 October, taking 21,915 cases to 1,011,6600.

To summarise, the use of the SEIQR (susceptible-exposed-infectious-quarantined-recovered) model over the SEIR has been justified by comparing the previous posterior model in (2.15) with the improved model in (2.22) for forecasting the COVID-19 outbreak across selected countries in Western Europe, by comparing the Bayesian posterior averages with the actual data. The equations, $\Psi_1(\mathcal{R}_t)$ and $\Psi_2(\mathcal{R}_t)$, are functions of the posterior models in (2.15) and (2.22), respectively. Equation $\Psi_2(\mathcal{R}_t)$ is in strong agreement with $\Psi_1(\mathcal{R}_t)$, which is effectively seen as a qualitative benchmark for describing COVID-19 outbreak in the United Kingdom (see Figure 2.6). I have then shown that the posterior estimates of the effective reproduction number for the COVID-19 pandemic in the United Kingdom, under the new model, strongly agrees with the deduced empirical values (as seen in Figure 2.7). I then used our FMEA of the community spread of COVID-19 in the United Kingdom and calculated the initial and revised RPN values, where each risk measure was ranked from 1 (insignificant / remote / failure prevented) to 10 (critical /

extremely high / no detection opportunity). For the analysis of the community spread of COVID-19 in the United Kingdom (from March to October 2020), the risk measures ($S$, $O$ and $D$) were matched according to the following. Criteria $S$ (CFR rate) to Figure 2.9, Criteria $O$ (posterior median of $\mathcal{R}_t$) to Figure 2.7, and Criteria $D$ (control methods) to the government policies defined in Table 2.3 for the given time intervals. As a result, the initial and revised RPN values seemed to accurately represent the government's contribution to reducing infection rates. For example, in the fourth phase, the government's scientific guidance indicated that the reproduction number had risen from 0.5–0.9 at the end of the lockdown to 0.7–1.0 on 15 May, closer to the point at which infections will begin to exponentially escalate. Fittingly, these figures corresponded to our posterior median estimates of 0.66–0.96 and 0.88–0.94 for the SEIQR model, respectively.

## 2.5 Concluding remarks

In this work, a Bayesian inference for the modified SEIR model with an additional state (self-quarantined $Q$) was considered, which characterizes the pandemic of COVID-19, focusing in particular on $\mathcal{R}_t$ in cases where the infectious period is exponentially distributed and where the posterior densities are extended from a Poisson-gamma mixture. The choice of exponential infectious period is used because it provides a natural analogue to the deterministic differential equation models, in which the constant removal rate corresponds to the exponentially distributed infectious period.

Bayesian inference and FMEA were combined together to study the risks of COVID-19 infections and to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic. The Bayesian model and FMEA are applied to the COVID-19 data showing the effectiveness of the interventions adopted to control the pandemic by reducing the $\mathcal{R}_t$ of COVID-19. In the FMEA, the focus was on COVID-19 infections and therefore the failure mode was taken as positive cases. The process phases, cause of failures, current measures and current risk ratings are discussed, and subsequent government action measures are presented with re-assessed risk ratings.

The results have shown that the combination of Bayesian inference, compartmental modelling and FMEA are effective to model and study the risks of the COVID-19 transmissions, able to evaluate quantitatively the action measures and identify the lessons learned from the impacts of governmental measures and actions carried out in response to the community spread of COVID-19 in the United Kingdom. Here, the use of the SEIQR model

over the SEIR was justified for forecasting the COVID-19 outbreak across selected countries in Western Europe, by comparing the Bayesian posterior averages with the actual data. The SEIQR model was found to have a high level of agreement with SEIR, which is used as a qualitative benchmark for describing the COVID-19 outbreak in the United Kingdom. It was demonstrated that the new model's posterior estimates of the effective reproduction number for the COVID-19 pandemic in the United Kingdom closely match the deduced empirical values. Fittingly, the results have shown that the RPN values (using the posterior medians) under the SEIQR model corresponded agree with the previous models.

Although the methodology presented here is intended to aid scientific research on Bayesian inference and risk assessment in the current COVID-19 pandemic using FMEA and the modified SEIR model, it can also be applied to future pandemics and other quality engineering applications.

| Process | Governmental response and effects | Sev | Occ | Det | RPN* |
|---|---|---|---|---|---|
| First cases<br>(22 Jan to 27 Feb) | • Travelers returned from Hubei, Iran, and certain regions of South Korea to the UK to self-isolate, even if they had no symptoms (25 Feb).<br>• Health Protection Scotland had set up an incident management team with full contact tracing for delegates who had tested positive (27 Feb). | 1 | 10 | 9 | 90 |
| First deaths and early spread<br>(28 Feb to 23 Mar) | • Government published its action plan for dealing with COVID-19 (3 Mar).<br>• PM announced that £46 million had been spent on research into coronavirus vaccines and rapid diagnostic tests (6 Mar).<br>• Chancellor of the Exchequer presented the first budget of the Johnson Government, which included £30 billion in measures to protect the economy from coronavirus (11 Mar).<br>• Chief Medical Officers raised the national risk level from "moderate" to "high" (12 Mar).<br>• Vice-President (US) announced an extension of its travel ban on coronavirus to include the UK (14 Mar).<br>• PM advised against "non-essential" travel and communication, working from home and avoiding visits as pubs, clubs or theatres (16 Mar). Four days later, all cafes, pubs and restaurants were ordered to close, except for take-away food (20 Mar).<br>• All non-essential shops, libraries, places of worship and outdoor gyms closed, and police were given authority (23 Mar). | 10 | 9 | 3 | 270 |
| First lockdown<br>(23 Mar to 4 May) | • Health Protection Regulations 2020 applied: sweeping restrictions legally enforceable.<br>• Contactless payment limit for in-store expenditure was increased from £30 to £45 (1 Apr).<br>• Foreign Secretary announced a three-week extension of the lockdown measures as the number of confirmed cases exceeded 100,000 (16 Apr).<br>• Vaccine development was underway as Matt Hancock announced the Government's £42.5m clinical trial plan to be conducted by Imperial College London and Oxford University (21 Apr).<br>• A study involving 20,000 households in England, coordinated by the Office for National Statistics, tracked the progress of COVID-19 towards a better understanding of infection and immunity (23 Apr). | 10 | 6 | 2 | 120 |
| Easing down lockdown measures and continued restrictions<br>(5 May to 31 May) | • The Government updated its coronavirus message from "Stay at Home, Protect the NHS, Save Lives" to "Stay Alert, Control the Virus, Save Lives" and announced a new "COVID-19 alert level system" ranging from green (level one to red (level five) (10 May).<br>• Loss of smell and loss of taste were added to the list of symptoms of COVID-19 (18 May).<br>• Following an agreement between the government and the Swiss pharmaceutical company Roche, the NHS provided a COVID-19 antibody test to check whether someone had the virus (21 May).<br>• The Government laid out new quarantine rules for travellers to the UK that required them to self-isolate from 8 June for 14 days (22 May).<br>• Contact tracing systems went live in England (NHS Test and Trace) and Scotland (Test and Protect) (28 May). | 10 | 6 | 4 | 240 |
| Continued restrictions and local lockdowns<br>(1 June to 5 Sep) | • Wearing face masks were made compulsory on public transport (15 Jun).<br>• Lowered the COVID-19 Alert Level from Level 4 (severe risk, high transmission) to Level 3 (substantial risk, general circulation) (19 Jun).<br>• Trialed a new coronavirus saliva test (22 Jun).<br>• Following an increase in COVID-19 cases in Leicester, stricter lockdown measures for the city were reintroduced (29 Jun).<br>• Wearing face masks were made compulsory in shops and supermarkets in Scotland (Jul 10) and England (Jul 24) as well as indoor settings, such as cinemas and places of worship (31 Jul).<br>• Introduced the month-long "Eat Out to Help Out" scheme (3 Aug).<br>• Increased the number of tests from 28,000 in England to 150,000 by October (19 Aug).<br>• Reopened schools in England, Wales and Northern Ireland for the autumn term (1 Sep). | 4 | 8 | 5 | 160 |
| Resurgance<br>(6 Sep to 31 Oct) | • Health Secretary informed the House of Commons that the "sharp rise" in new infections is "concerning" and a sign that the disease "remains a threat" (8 Sep).<br>• The "rule of six" came into force in England (14 Sep).<br>• The NHS contact-tracing app covering England and Wales was released (24 Sep).<br>• Fines of up to £10,000 levied for people in England who failed to self-isolate themselves (28 Sep).<br>• PM unveiled a new three-tier system of restrictions in England (14 Oct).<br>• PM announced a second lockdown for England to prevent what he calls the "medical and moral disaster" of the NHS from 5 November to 2 December, then England will return to the tier system (31 Oct). | 5 | 6 | 5 | 150 |

**Table 2.6:** Summary of subsequent government responses and re-evaluated risk measures. The initial cause of failures, current control and risk measures are discussed in Table 2.5. *Risk measures calculated by averaging over the last 7 days.

# Chapter 3

# A Bayesian Risk Assessment of the CMM Measurement Process using an Augmented Form Error Model

This chapter is concerned about whether a product meets its specifications based on its form error. Form error, that is, the departure of a manufactured part for its design or ideal shape, is a key characteristic to be assessed in quality engineering in manufacturing. In practice, form errors are usually estimated from coordinate measurements involving only a finite number of measured points and the form error for the complete workpiece surface has to be inferred on the basis of these measurements. This chapter presents a probabilistic form error model based on symmetric unimodal distributions. Bayesian inference is used to identify influence factors associated with the measurement process, such as form error, environmental influences, operator errors and random effects. A risk assessment is then performed by combining Bayesian inference, FMECA (Failure mode, effects, and criticality analysis) and conformity testing, to quantify and minimise the risk of wrong decisions. I present an innovative way of applying FMECA by assessing the risks of a CMM (coordinate measuring machine) process needed for improved product conformity testing. In the FMECA, I focused on CMM measurement strategy and so critical/security characteristic deviations, assembly part positioning errors, noise problems, and incorrect CMM calibration are taken to be significant failure modes that can occur. The model is applied to CMM form error data. The study will show that the combination of Bayesian inference, conformity assessment and FMECA is effective in modelling and studying the risks associated with CMM measurements, leading to the quantitative evaluation of the action

measures and the identification of improvement measures and actions taken in response to critical failure nodes of a CMM, and thus improve reliability of future measurements for product conformity assessment. The methods proposed here are meant to be practical without complex computing and can be further extended to general measurement processes and wide quality engineering.

## 3.1 Introduction

### 3.1.1 Overview

Form errors are the marginal differences between the measured data and reference nominal surface. Based on form error data and a product specification, the relationship between conformity testing and making decisions is established. A CMM is used to obtain measurement data and the form error is then computed. The assessment of form error involves two computational steps: calculating some measure of the distance of $d_i(\boldsymbol{x}_i, \boldsymbol{a})$ of a point $\boldsymbol{x}_i$ to the ideal geometry parametrized by parameters $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$ and then adjusting $\boldsymbol{a}$ so that some aggregate measure $D(\boldsymbol{a})$ of distance is minimised; see e.g., [17, 18, 73–80].

For example, a circle can be parametrised by circle centre coordinated $(x_0, y_0)^\top$ and radius $r_0$. Circularity is a two-dimensional tolerance that governs the overall shape of a circle ensuring that it is not too oval or oblong (e.g. a rotating shaft or a bearing). This type of measurement that is very common and is used in all forms of production and it aims to improve the roundness quality of an item. The (orthogonal, signed) distance $d(\boldsymbol{x}_i, \boldsymbol{a})$ from a point $\boldsymbol{x}_i = (x_i, y_i)^\top$ for a circle specified by $\boldsymbol{a} = (x_0, y_0, r_0)^\top$ is given by

$$d(\boldsymbol{x}_i, \boldsymbol{a}) = [(x_i - x_0)^2 + (y_i - y_0)^2]^{1/2} - r_0.$$

For form error expressions for other standard geometric shapes, see [17, 81].

Two aggregate measures of distance are commonly used:

- least squares (LS): $D_{\mathrm{LS}}(\boldsymbol{a}) = \sum_i d_i^2(\boldsymbol{x}_i, \boldsymbol{a})$,
- Chebyshev, minimax, or minimum zone (MZ): $D_{\mathrm{MZ}}(\boldsymbol{a}) = \max_i |d_i(\boldsymbol{a})|$.

A probe schematic and a visual example for estimating the form error by LS fitting from a CMM are provided in Figure 3.1 and Figure 3.2, respectively.

A third proposed measure is based on the $L_1$ norm:

$$D_{\mathrm{L1}}(\boldsymbol{a}) = \sum_i |d_i(\boldsymbol{a})|,$$

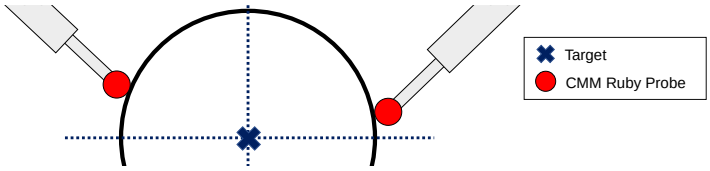



**Figure 3.1:** A CMM probe schematic for measuring roundness.

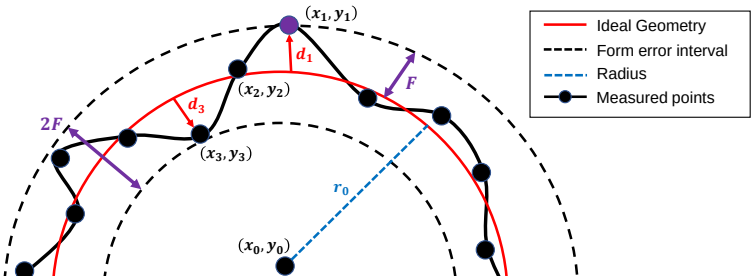



**Figure 3.2:** A visual example a circular artefact with centre $(x_0, y_0)$, radius $r_0$ and several measured points $\boldsymbol{x}_i$ consisting of $m$ discrete points for estimates $\boldsymbol{d}$. Form generated by LS fitting.

which has the potential advantage to being much less sensitive to the presence of outlying data points that can arise in coordinate metrology due to dust particles or cracks. Since the form error of a part is usually defined as the maximum departure (distance) from ideal geometry, the MZ criterion is often preferred as it represents the form error directly based on a discrete representation of the workpiece via the measured coordinated $\boldsymbol{x}_i$. These aggregate measures will depend nonlinearly on the parameters $\boldsymbol{a}$ so that iterative optimisation techniques are required to minimise them. In general, algorithms for nonlinear least squares optimisation are much more straightforward to implement those for nonlinear Chebyshev or $L_1$ optimisation. For problems that are linear in the parameters $\boldsymbol{a}$, there are effective algorithms for both the Chebyshev and $L_1$ problems; see e.g. see for example, [82–87] and the nonlinear associated problem can sometimes be addressed through solving a sequence of linearised problems [88].

In more recent years, Dong *et al.* [89] improved the data quality assessment technique and showed it is possible to compensate for the dynamic form errors of CMMs, while Wu *et al.* [90] conducted an error analysis and accuracy assessment of the reconstructed surfaces assessing the accuracy of the CMM device. Form tolerances associated with a part generally state that the form error $F$ should not exceed some tolerance $T$. In terms of a discrete representation $X = \boldsymbol{x}_{1:m} = \{\boldsymbol{x}_i, i = 1, \ldots, m\}$ of the part, I can interpret tolerance requirement as

$$\hat{F}(X) = \hat{F} = \min_{\boldsymbol{a}} \max_{i} |d_i(\boldsymbol{x}_i \boldsymbol{a})| \leq T,$$

where $\hat{F}$ is the estimate of the form error $F$ derived from the measured data $\boldsymbol{x}_{1:m}$. Hence,

a strategy for assessing part conformity is to gather coordinate measurements $\boldsymbol{x}_{1:m}$, apply some computational algorithm to determine the estimate $\hat{F}$ and pass the part if $\hat{F} \leq T$ and fail it otherwise[1]. However, this approach does not take into account the fact that $\hat{F}$ is determined from measurements and is subject to measurement uncertainty, e.g. see Joint Committee for Guides in Metrology (JCGM) 100 [91]. Hence, there is possibility that $\hat{F} < F$ or $\hat{F} > F$ and that a wrong decision will be made (false acceptance or false rejection).

A general approach to assessing conformity on the basis of an estimate is described in JCGM 106 [92]. The idea is that the information about the measurand, in the case $F$, is encoded in a probability distribution $p(F)$ and the decision can be made on the basis of assessing $\mathbb{P}(F \geq T)$. In a Bayesian setting, the data $X$ allows us to evaluate posterior distribution $\pi(F|X)$, given a model of the measurement system, and hence assess $\mathbb{P}(F \geq T)$. The approach in JCGM 106 can be generalised using Bayesian decision theory [93] to take into account the difference in costs associated with passing a nonconforming part (often termed the consumer's risk) and rejecting a conforming part (the producer's risk) via cost functions [94]. In [95], a cost matrix was used to quantify the different costs of making decisions that were wrong in different ways (e.g. passing a nonconforming part or failing a conforming part) to derive an optimal decision strategy that minimised the expected costs in the long run. In the context of statistical process control [96], it is usually assumed that the random effects associated with the measurement system are independent from one measurement to the next.

Forbes [95] discusses the impact of a systematic effect common to all measurements on the possible risks associated with conformity assessment. The estimate $\hat{F}(X)$ of the form error derived from data $X$ is affected by uncertainties associated with the coordinate measurements $X$ but also by sampling effects relating to the fact that only a finite set of points is used to represent the complete surface leading to a potentially significant underestimation of the true form error. Forbes [19] uses a Bayesian approach to determine a posterior distribution for $F$ on the basis of a finite sample from a rectangular distribution with independent sampling or Gaussian distribution with independent or correlated sampling where the spatial correlation associated with the sampling is modelled using Gaussian

[1]Note that is not necessary to determine the parameters $\boldsymbol{a}$ that minimise the Chebyshev measure in order to pass a part according to this strategy; it is only necessary to find a set of parameters $\boldsymbol{a}$, e.g., using a least squares optimisation, such that the tolerance is met. In order to fail a part according to this strategy, the Chebyshev solution is required. However, if I can find a subset $X_1 \subset X$ of the data for which $F(X_1) > T$, then necessarily $\hat{F}(X) \geq \hat{F}(X_1) > T$

processes. This chapter aims to extend the analysis associated with a risk and conformity assessment.

In this chapter, I present an augmented form error model to estimate the form error quantity $F$ based on the CMM measurement data, combined with environmental, human and random effects, identified by Bayesian inference. The proposed model is described using symmetric unimodal distributions (i.e., Gaussian and uniform distributions). A risk assessment is then performed by combining Bayesian inference, FMECA and conformity testing, to quantify and minimise the risk of wrong decisions. The objectives of this chapter are:

- To study the risks of the CMM measurement process and to evaluate the effectiveness of the action measures taken to manage product conformity rates by combining Bayesian inference and FMECA together;
- To assess the impact of corrective measures to the CMM measurement process by comparing the empirical and posterior statistics of the form error parameter $F$ at different failure modes with and without the measures being implemented;
- To demonstrate the impact of preventive measures taken to monitor and minimize measurement errors, resulting in more reliable product conformity scores.
- To show how these ideas can be implemented to general measurement processes and in wider quality engineering applications.

The study will show that the combination of Bayesian inference, conformity assessment and FMECA is effective in modelling and studying the risks of CMM measurements, leading to the quantitative evaluation of the action measures and the identification of improvement measures and actions taken in response to critical failure nodes of a CMM, and thus improve reliability of future measurements for product conformity assessment.

### 3.1.2 Chapter structure

This chapter is organised as follows. In section 3.2, I define the form error and propose a new (augmented) form error model which takes environmental, human and random effects into account. In section 3.3, I describe the form error model, which uses a uniform distribution with unknown boundaries to generate a posterior distribution using Bayesian inference. I then build on from the first posterior model by considering an extended posterior model that takes into account the likely random effects of form errors as well as real form error

data. The second model is then further updated using the proposed augmented model. In section 3.4, I generate expressions for conformity probability and specific risks associated with conformity assessment using the provided posterior distribution from the previous section. In section 3.5, I present a framework which combines FMECA with the methods and calculations in the previous sections using an *integrated approach*. In section 3.6, I present the findings which follow directly from the previous sections. The analysis begins with a description of the CMM measurement process and a form error model validation. Then, utilizing the integrated approach, a critical FMECA risk assessment is performed to identify relevant failure modes/causes to the CMM measurement process, estimate unknown FMECA risk parameters and obtain the criticality/RPN values. In section 3.7, a discussion on the calculated criticality/RPN values is provided. Finally, in section 3.8, I present the concluding remarks for this chapter. A summary of this case study is presented in Figure 3.3.

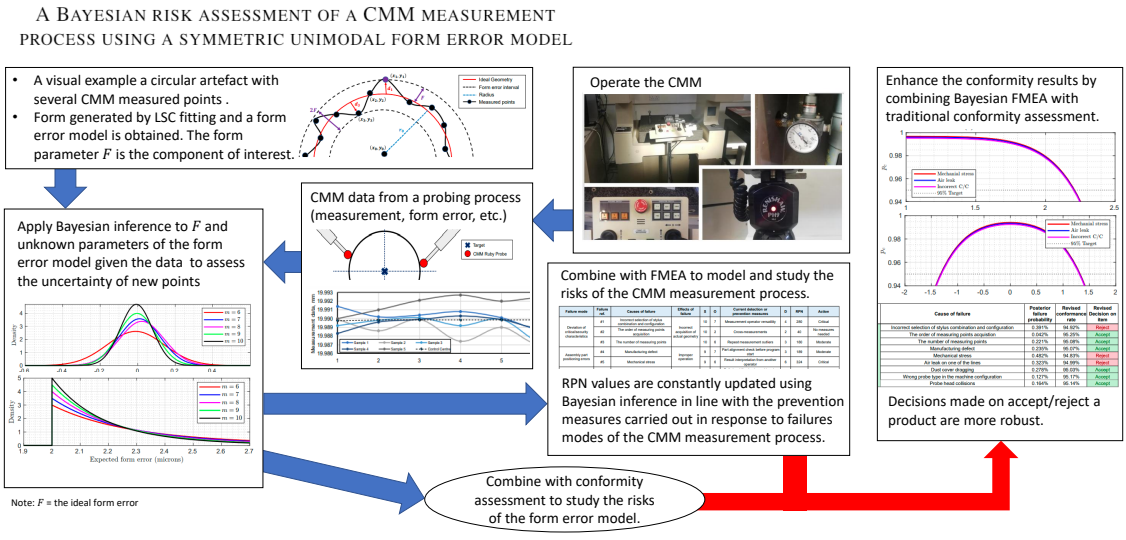



**Figure 3.3:** Summary of the second case study methodology.

## 3.2 Model and preliminaries

Using coordinates data, form error estimates are derived using the following model

$$\boldsymbol{x}_i = \boldsymbol{s}_i + d_i \boldsymbol{n}_i + \boldsymbol{\epsilon}_i, \quad \boldsymbol{\epsilon}_i \in \mathcal{N}(\boldsymbol{0}, \sigma_\epsilon^2 I), \quad i = \{1, 2, \ldots, m\}, \tag{3.1}$$

where $\boldsymbol{x}_i$ are the measured coordinated, $\boldsymbol{s}_i \equiv \boldsymbol{s}_i(\boldsymbol{u}_i, \boldsymbol{a})$ is the point on an ideal surface, $\boldsymbol{n}_i$ is the normal vector to the ideal surface at $\boldsymbol{s}_i$, $d_i$ is the form error at $\boldsymbol{s}_i$ and $\boldsymbol{\epsilon}_i$ is a random effect associated with the measurement system, representing a sum of effects that

change over a very short period of time and model the repeatability element of the CMM. The uncertainty associated with the measurement is represented by the standard deviation $\sigma_\epsilon$. (I can also extend the model to include a systematic effect $\boldsymbol{e}_i$ associated with the measurement system.) Determining $\boldsymbol{a}$ by minimising some aggregate measure $D(\boldsymbol{a})$ related to the orthogonal distances $d(\boldsymbol{x}_i, \boldsymbol{a})$, I determine estimates $\hat{\boldsymbol{a}}$ of $\boldsymbol{a}$. To the first order,

$$f_i = d(\boldsymbol{x}_i, \hat{\boldsymbol{a}}) \approx d_i + \epsilon_i, \quad \epsilon_i \in \mathcal{N}(0, \sigma_\epsilon^2), \quad i = \{1, 2, \ldots, m\}, \tag{3.2}$$

so that the residual distances $f_i = d(\boldsymbol{x}_i, \hat{\boldsymbol{a}})$ are estimates of the form errors $d_i$. If the model for the form errors is that $d_i \in \mathcal{N}(0, \sigma_d^2)$ then, approximately $f_i \in \mathcal{N}(0, \sigma_d^2 + \sigma_\epsilon^2)$. The above model in (3.2) related to the estimation of form error in the presence of random effects associated with the measurement system. Below I extend the model to account for other effects.

### 3.2.1 The proposed model

To generalise the form error model to describe the CMM measurement process, I extend the fitted form error model in (3.2) by introducing two new parameters: environmental error ($\theta_i$) and human error ($h_i$). A CMM's performance can be influenced by a variety of external environmental factors [97]. Common environmental errors include the effects of:

- Temperature (continual heat switching or overheating of machine tools);
- Humidity (harmful effects of moisture and the corrosion of items and measurement devices);
- Trapped dust (small particles trapped in the CMM's bearing surface can cause breakdown during operation or inaccurate measurements);
- Oil (leakage can lead to deterioration of CMM and future measurements);
- Vibration (criteria levels of whether the CMM is capable of withstanding).

Human errors are specified in ISO 14224 [98] and are associated with various failure causes. The definition used in ISO 14224 is the "*discrepancy between the human action taken or omitted and that intended*". Given the concept, it's also unclear how to improve on it, particularly in terms of understanding how human performance affects equipment failures. Is it permissible in a standard focusing on equipment dependability performance to attempt to link failure causes and human errors? Selvik and Bellamy [99] provided a

lengthy response to this question and the following three errors were proposed as the most common sources of human errors when conducting the measurement process:

- Operating error (operation/maintenance-related);
- Maintenance error (operation/maintenance-related);
- Documentation error (management-related).

These factors represent additional sources of variation that increase the measurement uncertainty associated with the estimates of the surface parameters $\boldsymbol{a}$ and estimates of the form error $F$. These factors may contribute poor results, decision costs and (a never-ending supply of) user dissatisfaction. When all of these factors are considered, the new augmented model can be extended to:

$$f_i = d_i + \theta_i + h_i + \epsilon_i, \tag{3.3}$$

with $\theta_i$ and $h_i$ associated with Gaussian distributions: $\theta_i \sim \mathcal{N}(0, \sigma_\theta^2)$ and $h_i \sim \mathcal{N}(\mu_h, \sigma_h^2)$. Human error is defined in this thesis as failures caused by humans during the CMM measuring process, such as dirt on the measured item, a probe that is not calibrated, and so on. In general, because no experiment can be perfect, there will always be some level of error, so $h_i > 0$. Furthermore, the observed residual distances $f_i$ associated with a surface fit will be approximately distributed by a Gaussian, $f_i \in \mathcal{N}(0, \sigma^2)$, with

$$\sigma^2 = \sigma_d^2 + \sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2. \tag{3.4}$$

The model in (3.3) involving a response $f_i$ that depends additively on a number of factors allows for a measurement system analysis approach based on the analysis of variance [20–23] in which experiments are undertaken with a number of factors are held constant (e.g. the same operator) in order to quantify the influence of the factors through deriving estimates of $\sigma_\theta$, etc. These extra uncertainties can be calculated using an analytical method to CMM uncertainties based on the Guide to the Expression of Uncertainty in Measurement (GUM) [92]. These uncertainties are classed as Type B uncertainties and may include uncertainty in temperature measurement, uncertainty related with stylus modifications, uncertainty in material expansion coefficient knowledge, and so on [100]. Type A standard uncertainties, which are typically used to quantify the repeatability or randomness of a measurement process, are closely linked to human-related errors [100]. I can give an example of how different environmental factors (temperature effects associated with the CMM and workpiece) can be combined and modelled in terms of

$$\theta_i \in \mathcal{N}(0, \sigma_\theta^2), \quad \sigma_\theta^2 = \sum_{k=1}^{N} c_k^2$$

where $c_k$ is the standard uncertainty associated with the $k$-th effect. Similarly, I can also give an example of how different human factors (e.g., operational and maintenance effects associated with the CMM measurement process) can be used to estimate $\sigma_h$, i.e.

$$\sigma_h \approx \frac{1}{m-1}\sum_{i=1}^{m}(h_i - \overline{h}), \tag{3.5}$$

where $m$ is the number of measured points, $h_i = f_i - f_i'$ is the difference between the form errors measured before and after any human-related failures (e.g., dirt on item, probe not calibrated, etc) associated with the CMM measurement process, and $\overline{h} = \frac{1}{m}\sum_{i=1}^{m} h_i$.

### 3.2.2 Contributing uncertainties to $c_k$

I want to calculate the size of all contributing uncertainties of $c_k$ (for $k = 1, \ldots, 4$) from the following contributors using the methods in [100], namely:

- The coefficient of thermal expansion of the workpiece ($c_1$).
- The temperature deviation of the workpiece ($c_2$).
- The coefficient of thermal expansion of CMM scale ($c_3$).
- The temperature deviation of CMM scale ($c_4$).

Here, each of these standard uncertainties are given by

$$c_1 = u_\alpha L\Delta T/\sqrt{3}, \quad c_2 = u_T L\alpha/\sqrt{3}, \quad c_3 = u_S L\Delta T/\sqrt{3}, \quad c_4 = u_\nu L\nu/\sqrt{3}, \tag{3.6}$$

where $u_\alpha$ is the contribution to the expansion coefficient of the component material, $L$ is the length of the material, $\Delta T$ is the deviation of the workpiece from 20°C, $u_T$ is the contribution of error from the CMM's workpiece thermometer, $\alpha$ is the workpiece thermometer, $u_S$ is the contribution to the expansion coefficient of the CMM's scale material, $\nu$ is the CMM's scales expansion coefficient and $u_\nu$ is contribution of the possible error of the CMM's scale thermometer. The NPL guide [100] provides default values for some of these parameters: $u_T = 0.1$, $\alpha = 11.7 \times 10^{-6}$, $u_S = 1 \times 10^{-6}$, $\nu = 7.8 \times 10^{-6}$ and $u_\nu = 0.1$.

## 3.3 Bayesian method on the form error

In this section, the form error estimates $f_i$ for $i = 1, 2, \ldots, m$ associated with points on the ideal surface are assumed to be drawn from a uniform distribution whose parameters are related to the form error boundary parameter $F > 0$ (more specifically, $F$ is the boundary

of the form error estimates as seen in Figure 3.2). The aim is to use Bayesian inference to obtain a posterior distribution for $F$ based on observed form error estimates $f_i$. Conformity and risk calculations associated with the generated posterior distributions are presented in section 3.4.

### 3.3.1 Form errors associated with a uniform distribution

Let $f_i \in (-\infty, \infty)$ for $i = 1, \ldots, m$ be draws from a continuous uniform distribution in the interval $[-F, F]$, with PDF:

$$g(f_i|F) \equiv U(f_i|-F, F) = \begin{cases} (2F)^{-1} & f_i \in [-F, F], \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

The symmetric nature of this distribution ensures that the mean, median and skewness are zero, and that the variance of this distribution is explained from $\sigma_f^2 = F^2/3$. Form errors depend on the data of $f_i$ and each $f_i$ has uncertainty with the measurements and is the difference between the measured quantity and its true value.

Since the form error data is given by $U(f_i|-F, F)$, then the likelihood function of associated with the model for the form error data is

$$\mathcal{L}(\boldsymbol{f}|F) = \prod_{i=1}^{m} U(f_i|-F, F) \propto 1/F^m, \quad f_i \in [-F, F]. \tag{3.8}$$

By treating $F$ as scale parameter, a suitable non-informative prior PDF for $F$ is $\pi_0(F) = 1/F$. Then by Bayes' theorem, the combined posterior PDF results to a Pareto distribution:

$$\pi(F|m, F_0) = \begin{cases} mF_0^m/F^{m+1} & F \geq F_0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.9}$$

where $F_0 = \max(|f_1|, \ldots, |f_m|)$. The Pareto is a skewed distribution with heavy or "slowly decaying" tails, it places significantly higher probabilities as more measured points are observed. The parameter $F_0$ is the lower bound of the Pareto (which in turn, is the mode of the distribution). Figure 3.4 graphs (3.9) for various values of $m$ for the case of $F_0 = 2$ μm. Table 3.1 shows that the posterior mean, median and percentiles approaches $F_0$ as $m$ increases. Similarly, the standard deviation decreases towards zero as $m$ increases. It is interesting to note that the posterior median is always less or equal to the posterior mean for all values of $m$. This is due to the Pareto distribution's unique nature.

In order to predict future form error numbers, it is better to avoid estimating $F$ and

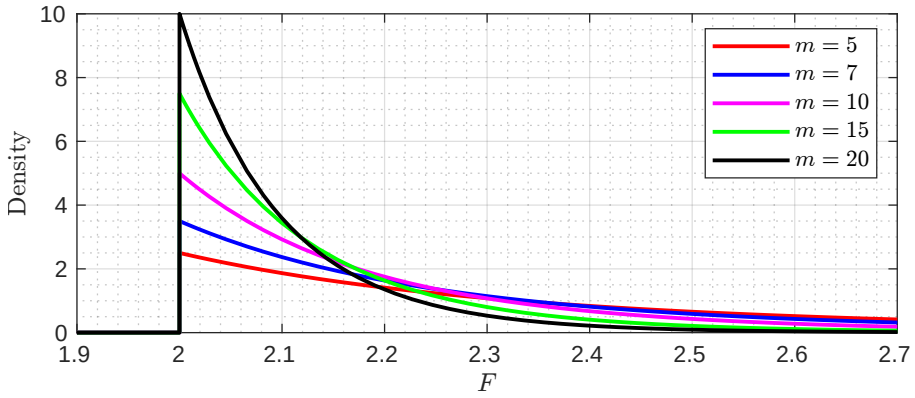



**Figure 3.4:** Pareto distribution for various levels of $m$ starting with a limiting value of $F_0 = 2$ µm. The peak rises and the tail cuts off earlier as more samples are added.

**Table 3.1:** Pareto statistics for the posterior mean, median, standard deviation and some upper percentiles for the case of $F_0 = 2$ µm.

| **Measured points** $m$ | **5** | **6** | **7** | **8** | **9** | **10** | **12** | **15** | **20** | **25** | **30** | **40** | **50** | **75** | **100** | **200** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Posterior mean** | 2.50 | 2.40 | 2.33 | 2.29 | 2.25 | 2.22 | 2.18 | 2.14 | 2.11 | 2.08 | 2.07 | 2.05 | 2.04 | 2.03 | 2.02 | 2.01 |
| **Posterior median** | 2.30 | 2.24 | 2.21 | 2.18 | 2.16 | 2.14 | 2.12 | 2.09 | 2.07 | 2.06 | 2.05 | 2.03 | 2.03 | 2.02 | 2.01 | 2.01 |
| **Posterior SD** | 0.65 | 0.49 | 0.39 | 0.33 | 0.28 | 0.25 | 0.20 | 0.15 | 0.11 | 0.09 | 0.07 | 0.05 | 0.04 | 0.03 | 0.02 | 0.01 |
| **95th percentile** | 3.64 | 3.30 | 3.07 | 2.91 | 2.79 | 2.70 | 2.57 | 2.44 | 2.32 | 2.25 | 2.21 | 2.16 | 2.12 | 2.08 | 2.06 | 2.03 |
| **99th percentile** | 5.02 | 4.31 | 3.86 | 3.56 | 3.34 | 3.17 | 2.94 | 2.72 | 2.52 | 2.40 | 2.33 | 2.24 | 2.19 | 2.13 | 2.09 | 2.05 |

instead use the predictive density, which is defined as:

$$\pi(f_{\text{new}}|\mathcal{D}) = \int_{F_0^*}^{\infty} U(f_{\text{new}}|-F, F)\pi(F|\mathcal{D})\, dF = \begin{cases} \frac{m}{2(m+1)F_0} & \text{if } f_{\text{new}} < F_0, \\ \frac{mF_0^m}{2(m+1)f_{\text{new}}^{m+1}} & \text{if } f_{\text{new}} \geq F_0, \end{cases} \tag{3.10}$$

where $F_0^* = \max(F_0, f_{\text{new}})$. Figure 3.5 shows the effects of plotting a Pareto distribution's predictive density after extracting a new point from a uniform distribution. It's worth noting that the predictive density is divided into two parts: a constant $\pi(f_{\text{new}}|\mathcal{D}) = A$ and a reciprocal component $\pi(f_{\text{new}}|\mathcal{D}) = B/f_{\text{new}}^k$, where $A$, $B$ and $k$ are constants, and that the value of the mode acts as a vertical asymptote to the reciprocal part. If the number of measured points increases, the reciprocal function approaches the asymptote, implying that the probability of obtaining a new measurement of $f_{\text{new}}$, given $F_0$, decreases.

### 3.3.2 Incorporating uncertainty from random effects

In this section, I present a method to obtain a slightly more extensive posterior model, by taking into consideration the probable random effects of form errors. The calculations in the previous section assumed that the observed residual distances are accurate estimates of the surface's form error estimates $f_i \in U(-F, F)$ and were free from some random effects.

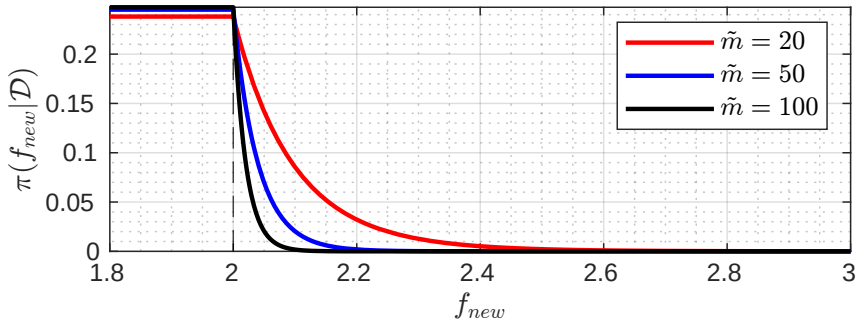



**Figure 3.5:** Pareto predictive density using $\pi(f_{\text{new}}|\mathcal{D})$ for the case of $F_0 = 2$ and various levels of $m$.

Suppose now that $f_i \approx d_i + \epsilon_i$ where $d_i \in U(-F, F)$ and $\epsilon_i$ are random effects associated with a Gaussian distribution $\mathcal{N}(0, \sigma_\epsilon^2)$. Random errors are measurement errors that result in inconsistent quantifiable values when repeated measurements are taken. Most of the JCGM [92] methodology is based on the premise that the underlying error distribution is Gaussian. Why are I assuming a Gaussian distribution? Simply because this is the distribution that represents or approximates what I often see in the physical universe.

As such, I shall model the measurements of CMM data resulting from random errors added to produce an ideal geometric form. I can use Bayesian inference to obtain a posterior distribution for the convolution of $d_i + \epsilon_i$. The result of the sum of two independent continuous random variables has PDF:

$$\begin{aligned} g(z|F, m, \sigma_\epsilon) &= \int_{-F}^{F} U(t| - F, F)\, \mathcal{N}(z - t|0, \sigma_\epsilon^2)\, dt, \\ &= \frac{1}{2F} \int_{-F}^{F} \frac{1}{\sigma_\epsilon \sqrt{2\pi}} \exp\left(\frac{-(z-t)^2}{2\sigma_\epsilon^2}\right) dt, \\ &= \frac{1}{4F}\left[\text{erf}\left\{\frac{F+z}{\sigma_\epsilon\sqrt{2}}\right\} + \text{erf}\left\{\frac{F-z}{\sigma_\epsilon\sqrt{2}}\right\}\right], \end{aligned} \tag{3.11}$$

where $\text{erf}(y)$ is the Gauss error function, given by:

$$\text{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^z \exp\left(-t^2\right) dt, \quad t \in [-y, y].$$

By taking the likelihood of (3.11) after observing $m$ measured points and setting a non-informative prior of $\pi_0(F) \propto 1/F$, the combined posterior PDF is therefore:

$$\pi(F|\boldsymbol{d}, F_0, m, \sigma_\epsilon) = \frac{C_I}{F^{m+1}} \prod_{i=1}^{m} \left[\text{erf}\left\{\frac{F+d_i}{\sigma_\epsilon\sqrt{2}}\right\} + \text{erf}\left\{\frac{F-d_i}{\sigma_\epsilon\sqrt{2}}\right\}\right], \quad F \geq F_0, \tag{3.12}$$

and zero everywhere else. Here, $F_0 = \max\{|d_1|, \ldots, |d_m|\}$ and $C_I$ is the integrating factor to ensure the PDF integrates to one. This distribution reflects the measurement uncertainty affiliated with the residual distances examined, due to the randomness of $\sigma_\epsilon$. The value of $C_I$ can be computed very easily using *Mathematica's* `1/NIntegrate[]` or *MATLAB's*

`1./int()` command. Due to complicated integrals, summations or products, obtaining a posterior distribution in many cases is not possible, and thus the posterior distribution does not have a closed form. Numerical integration is used to compute the value of $C_I$, moments, quantiles and other common statistics associated with this PDF. Alternatively, I can sample from the posterior and draw distributions across the parameters using well-known techniques like the MCMC without having to worry about computing the evidence.

### 3.3.3 Incorporating environmental and human errors

In this section, I now take into account the effects of additional errors to obtain an even more extensive posterior mode. I shall model the form errors of CMM measurement data using Bayesian inference for the convolution of $d_i + \theta_i + h_i + \epsilon_i$. Since the three Gaussian random variables has a resulting PDF of $\mathcal{N}(z|\mu_h, \sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)$, then the PDF of the convolution of a uniform and three Gaussian random variables is:

$$\begin{aligned} g(z|F, m, \boldsymbol{\mu}, \boldsymbol{\sigma}) &= \int_{-F}^{F} U(t|-F, F)\, \mathcal{N}(z-t|\mu_h, \sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)\, dt, \\ &= \frac{1}{4F}\left[\operatorname{erf}\left\{\frac{F + z + \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}}\right\} + \operatorname{erf}\left\{\frac{F - z - \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}}\right\}\right]. \end{aligned} \tag{3.13}$$

By taking the likelihood of (3.13) after observing $m$ measured points and setting a non-informative prior of $\pi_0(F) \propto 1/F$, the combined posterior PDF is therefore:

$$\pi(F|\boldsymbol{d}, F_0, m, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \frac{C_I}{F^{m+1}} \prod_{i=1}^{m}\left[\operatorname{erf}\left\{\frac{F + d_i + \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}}\right\} + \operatorname{erf}\left\{\frac{F - d_i - \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}}\right\}\right], \tag{3.14}$$

for $F \geq F_0$ and $\pi(F|\boldsymbol{d}, F_0, m, \boldsymbol{\mu}, \boldsymbol{\sigma})$ is zero everywhere else.

### 3.3.4 Form errors associated with a Gaussian

Alternatively, one could use a Gaussian distribution to represent the form error. In statistics, the Gaussian distribution is one of the most commonly used. It's also popular to use Bayesian inference and conjugate priors to estimate its parameters. Most of the results can be obtained in closed form thanks to the use of conjugate priors.

Now suppose I consider a Bayesian estimation of a univariate Gaussian's mean, with the variance assumed to be known. If $\boldsymbol{f} = (f_1, f_2, \ldots, f_m)^\top$ are accurate form error estimates observed from a Gaussian distribution with mean $\mu$ and (known) standard deviation $\sigma$ equal to (3.4), then the PDF of this distribution is:

$$g(f_i|\mu, \sigma) \equiv \mathcal{N}(f_i|\mu, \sigma^2) \propto \exp\left[-\frac{(f_i - \mu)^2}{2\sigma^2}\right]. \tag{3.15}$$

According to Bayesian inference, I can derive a posterior probability from two components: a prior probability and a likelihood function derived from a statistical model for the observed data. Since the form error data is given by $\mathcal{N}(f_i|\mu,\sigma^2)$, then the likelihood function of associated with the model for the form error data is

$$\mathcal{L}(\boldsymbol{f}|\mu,\sigma) = \prod_{i=1}^{m} \mathcal{N}(f_i|\mu,\sigma^2) \propto \exp\left[-\frac{\sum_{i=1}^{m}(f_i-\mu)^2}{2\sigma^2}\right]. \tag{3.16}$$

If I assume a Gaussian prior of $\pi_0(\mu) = \mathcal{N}(\mu|\mu_0,\sigma_0^2)$, then the combined posterior distribution is also Gaussian distributed:

$$\begin{aligned}\pi(\mu|\boldsymbol{f},\mu_m,\sigma_m) &= \prod_{i=1}^{m} \mathcal{N}(f_i|\mu,\sigma^2)\mathcal{N}(\mu|\mu_0,\sigma_0^2),\\ &\propto \exp\left[-\frac{\sum_{i=1}^{m}(f_i-\mu)^2}{2\sigma^2}\right] \times \exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right]\\ &\propto \exp\left[-\frac{(\mu-\mu_m)^2}{2\sigma_m^2}\right] \sim \mathcal{N}(\mu|\mu_m,\sigma_m^2).\end{aligned} \tag{3.17}$$

Here, $\mu_m$ and $\sigma_m^2$ are hyper-parameters of the posterior distribution and are given by:

$$\mu_m = \sigma_m^2\left(\frac{\mu_0}{\sigma_0^2} + \frac{m\overline{f}}{\sigma_d^2+\sigma_\theta^2+\sigma_h^2+\sigma_\epsilon^2}\right), \tag{3.18}$$

$$\sigma_m^2 = \frac{\sigma_0^2(\sigma_d^2+\sigma_\theta^2+\sigma_h^2+\sigma_\epsilon^2)}{m\sigma_0^2+\sigma_d^2+\sigma_\theta^2+\sigma_h^2+\sigma_\epsilon^2}, \tag{3.19}$$

where $m\overline{f} = \sum_{i=1}^{m} f_i$. Figure 3.6 shows the effects of updating the Gaussian distribution in (3.17) sequentially using the hyperparameters in (3.18). As more measurements are taken, I can see that the posterior density narrows towards a form error mean of zero.

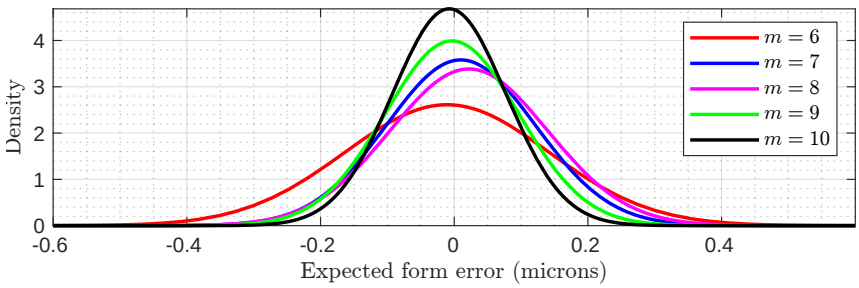



**Figure 3.6:** Sequentially updating a Gaussian mean starting with $\mu_0 = 0$, $\sigma_0 = 1$ and $\sigma \approx 0.9501$ using real form error data. Notice how the posterior becomes narrower as more points are collected.

The Gaussian distribution itself is an example of a symmetric unimodal distribution since the normal curve has one local maximum (peak) at the mean, median and mode which makes it a suitable distribution for the form error. However, it may not capture the form error data precisely, as the probability of observing the form error is almost guaranteed to be consistent and uniform. For measurements whose best estimate is within the usual range of standard uncertainty, this fraction of the probability can be substantial, so a Gaussian PDF may not always be the best choice.

## 3.4 Conformity assessment

The main goal of conformity assessment is to ensure that a measurement result is utilized to determine whether or not a particular item conforms with a set of requirements such as ISO's geometrical product criteria [101]. Due to probable system failures or when the item itself does not meet the established standards, the CMM may not always record accurate information. In conformity assessment, I wish to assign tolerance limits for a quantity to control the risks associated with making a wrong decision [92]. Presently, the most common form of product quality inspection is to use the specification zone over the required tolerance zone $[T_L, T_U]$ as a criterion and assess the quality of each product based on its calculated value (i.e., the interval containing all acceptable values). When the measurement results are within the tolerance limit, the item is recognized as accepted. The items are rejected if the measurement results surpass the tolerance limit. There is a possibility of uncertainty in the test results due to the effects of measurement uncertainty.

In the following sections, the Pareto posterior distribution in (3.9) and the Augmented posterior distribution in (3.14) will be used to better understand the conformity method.

### 3.4.1 Conformity and specific risk calculations

I assume a production line manufactures circular parts that are measured, leading to form error estimates that are assumed to follow a Pareto-like distribution. To assess whether the form error complies with a specification, the measured value is compared to a tolerance interval. The conformity rate, denoted by $p_c$, is an important metric for determining a process's capability. This is defined as the probability of an item falling within a one-sided tolerance interval from $[0, T]$, where $T$ is the upper limit. For a Pareto distribution, defined by $m$ and $F_0$,

$$p_c = \int_{F_0}^{T} \pi(F|m, F_0)\, dF = 1 - (F_0/T)^m. \tag{3.20}$$

For the augmented model in (3.14), specified by $m$, $F_0$, $\boldsymbol{d}$, $\mu$ and $\sigma$,

$$p_c = \int_{F_0}^{T} \frac{C_I}{F^{m+1}} \prod_{i=1}^{m} \left[ \operatorname{erf}\left\{ \frac{F + d_i + \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}} \right\} + \operatorname{erf}\left\{ \frac{F - d_i - \mu_h}{\sqrt{2(\sigma_\theta^2 + \sigma_h^2 + \sigma_\epsilon^2)}} \right\} \right] dF. \tag{3.21}$$

If the item does not conform with the specification, then the probability of non-conformity is $1 - p_c$. Note that the tolerance intervals in (3.20) and (3.21) are actually $[F_0, T]$ instead of $[0, T]$. This is because the Pareto and augmented models are defined for $F \geq F_0$.

JCGM 106 [92] published guidelines for such an assessment addressing two types of specific risks. Only binary decision criteria are evaluated, which means the item must either

be deemed as accepted or rejected. As such, these risks can be obtained as a function of $p_c$. The first is the specific consumer's risk, denoted here as $R_c^S = 1 - p_c$, which is the probability that an accepted item will be non-conforming. The second is the specific producer's risk, $R_p^S = p_c$, which is the probability that a rejected item will be conforming. Both risks are extremely important for a manufacturer to consider when structuring their measuring procedure in order to strike an appropriate balance between the risk of rejecting conforming items and the risk of accepting non-conforming items. The calculation of conformity probabilities allows design of experiment questions to be answered, e.g., how many measurement points $m$ are required to reduce risks to a specified level.

### 3.4.2 Measurement process capability index

Another useful indicator is the measurement process capability index $C_m$. This is a useful metric that characterizes the measurement's quality in relation to a tolerance criterion. This metric compares the process's actual spread (typically stated as a standard deviation) to the process's permissible spread (typically computed as the complete range of the specification set by the consumer). This is defined by

$$C_m = \frac{T_U - T_L}{2U}, \tag{3.22}$$

where $U = 2\sigma$ is the expanded uncertainty. The conventional choice of $2U$ is reinforced by the widespread preference for a coverage interval of approximately 95% [92]. Furthermore, the measurement capacity index can be linked to the conformity probability. Hence, the following form can be used to combine the Pareto-like conformity rate, given by $p_c = \int_{F_0}^{T} \pi(F|\mathcal{D})\, dF$, with (3.22):

$$p_c = \Pi(F_0 + 2C_m U|\mathcal{D}) - \Pi(T - 2C_m U|\mathcal{D}), \tag{3.23}$$

where $p_c$ is now determined by $\mathcal{D} = \{\boldsymbol{f}, \boldsymbol{\mu}, \boldsymbol{\sigma}\}$, $F_0$, $T$, $U$ and $C_m$ and $\Pi(\cdot)$ is the posterior cumulative distribution function (CDF).

### 3.4.3 Decision rules for accepting/rejecting a product

When specifying the rules for accepting or rejecting a product, it is important to define the method for accounting for measurement uncertainty in relation to the measurement result and the specified requirements [92]. Accepting or rejecting an item when the measurement result is near to a tolerance limit may lead to a wrong decision and unfavorable outcomes. There are two types of incorrect decisions: when an item is accepted as conforming but

may be non-conforming (false positive), and when an item is rejected as non-conforming but may be conforming (false negative). Altogether, there are four types of decisions in a conformity assessment with a binary decision rule as seen in Figure 3.7 for a single tolerance limit $T$ equal to acceptance limit $A$ in a set coverage interval (i.e. $A = T$):

- Valid acceptance: Item is accepted and conforms. This outcome of the conformity assessment test results in a conforming item being accepted (e.g. see case (I), Figure 3.7).
- False acceptance: Item is accepted but does not conform. This is an incorrect decision because the cost of such an error is often borne by a consumer who acknowledges the item as fit for its purpose and acts accordingly (e.g. see case (IV), Figure 3.7).
- False rejection: Item is rejected but conforms. This is an incorrect decision because the cost associated with such an error is often borne by a manufacturer who can not sell an item that has failed a conformity test (e.g. see cases (II-III), Figure 3.7).
- Valid rejection: Item is rejected and does not conform. This outcome of the conformity assessment test results in a conforming item being rejected (e.g. see case (V), Figure 3.7).

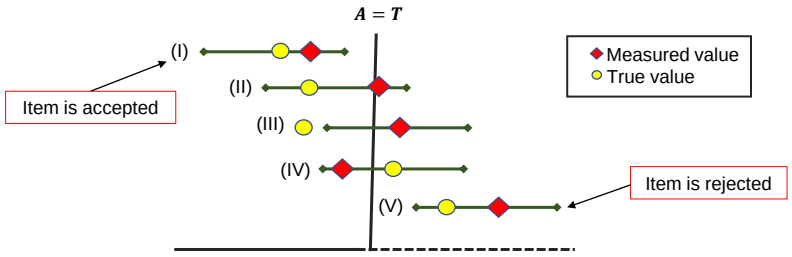



**Figure 3.7:** Cases (I) and (V) shows that the measured value and true value responses are within the regions of valid acceptance and valid rejection, respectively. Cases (II) and (IV) are examples of false rejections and false acceptance, respectively, due to incorrect decisions made. In case (III), the true measurement value is outside the coverage interval and the response is characterized by a false rejection.

A decision rule based on simple acceptance and the measured value provided by a symmetric unimodal distribution (such as a Gaussian) will result in a probability of accepting a non-conforming item or rejecting a conforming item of up to 50% if the measurement results coincides with a tolerance limit [92]. To avoid these high probabilities, it is recommended that acceptance limits be offset from tolerance limits using a conformity decision strategy known as *guard banding*. The guard band was first specified in ISO 14253-1 [101], which

establishes an acceptance decision rule for demonstrating specification compliance with the goal of minimising consumer's risk. It is frequently described using an expanded uncertainty parameter $U$ with a coverage factor of 2 and a coverage probability of 95%. Mathematically, the expanded uncertainty parameter can be expressed as $U = |T - A| = 2\sigma$ which is used to define the area between the tolerance limit $T$ and the acceptance limit $A < T$ in the case of a one-sided tolerance interval, where $\sigma$ is the standard uncertainty represented by the standard deviation parameter. Minimising the producer's risk is established from $A > T$ (see Figure 3.8). Reducing the consumer risk is an important approach for increasing conformity rates. When using a binary decision rule to assess conformity, acting to minimize the consumer's risk always increases the producer's risk.

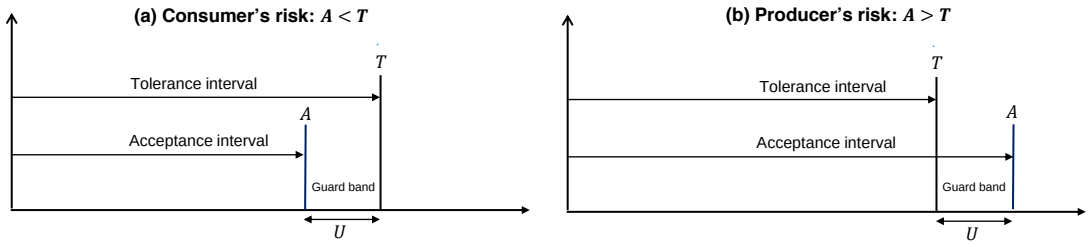



**Figure 3.8:** Binary conformity assessment for form errors in order to minimise: (a) consumer's risk; (b) producer's risk.

For the purposes of performing a conformity assessment, decision rules applied to a conformity assessment of one or two side tolerance limits are often applicable for a variety of statistical distributions.

## 3.5 Failure modes, effects, and criticality analysis

### 3.5.1 The FMECA procedure

FMECA was one of the earliest systematic failure analysis tools. The main goal of an FMECA risk assessment is to ensure that all possible failure modes/causes have been evaluated, as well as their effects on the system's operational success and to assess the current detection/corrective measures in place. After that, the risk priority number (RPN) is determined by assigning each failure mode a severity ($S$), occurrence ($O$), and detection ($D$) rating. The RPN is calculated from the product of these three parameters. Each parameter has a range of values, but the most common are 1-5 or 1-10, with 1 being the lowest (and most desirable) value. In general, the lower the RPN, the lower the risk, and the greater the RPN, the higher the risk, hence the goal is to keep the RPN as low as

possible. Failures that result in a high RPN are critical and given top priority [42–46, 102]. Control measure detection is examined in the final stage to reduce measurement errors and improve the reliability of making the right decision. Comparing the RPN for the initial and revised values, respectively, may reduce the risk associated with a corrective measure. Although the classic FMECA has a good ability to assess system safety, there are numerous limitations that have been documented in the literature [44]. A FMECA approach will be applied to the CMM measurement process and the risk measures will be rated on a scale of 1 to 10.

### 3.5.2 The integrated approach

There are three primary steps to the integrated approach that this study's methodology is based on. The first step is all about getting ready for failure analysis and control. For the CMM measurement process, trained persons or experts were identified at this step. Furthermore, the primary FMs and FCs that result from significant measurement errors during the CMM measurement process, as well as the current detection measures in place, were identified. The second step is the implementation phase. I want to estimate the FMECA risk parameters here ($S$, $O$ and $D$). Using the augmented posterior model for the form error, values for $S$ and $D$ can be calculated by computing the specific consumer's risk and the posterior standard deviation, respectively. A Bayesian network is used to calculate the values for the remaining risk parameter $O$, which is a probabilistic approach. In this approach, each FM/FC can be considered as a random event when trying to determine the values of $O$ [103]. In this chapter's results/findings phase (section 3.6), the Bayesian network procedure is detailed. The final step is to integrate all the calculated risk parameters for each FM/FC and calculate the RPNs that result. Figure 3.9 depicts the flow chart for this integrated approach. The following section is a detailed analysis of the findings and results.

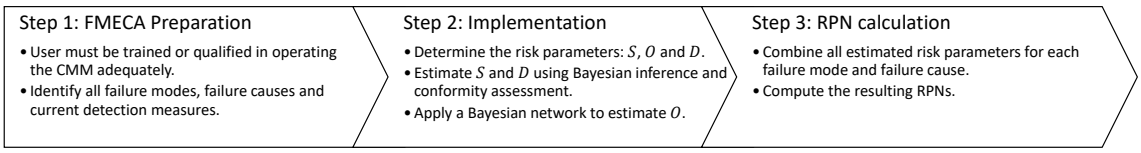



**Figure 3.9:** A flow chart for the integrated approach.

## 3.6 Results and findings

As a follow-up to the previous sections, this section presents the main results and findings of this chapter. These include a description of the CMM measurement process, a form error model validation and a critical risk assessment using the integrated approach for the Bayesian FMECA. The form error model is calibrated using data from a CRYSTA-Apex S, a high-precision CMM that offers a maximum permissible length measurement of 1.7 μm. It also includes a temperature correction system that ensures measurement accuracy at temperatures ranging from 16 to 26 degrees Celsius. The combined scale and workpiece temperature compensation scheme in this system checks the temperature before transmitting the measurement result to the controller and corrects it to the value that would be measured at 20 degrees Celsius. This outperforms prior CMM systems that just compensated for the scale temperature. In the model, the combined temperature sensor and workpiece is approximated by $\sigma_\theta$ using the Type B contribution method in (3.6). Similarly, I use (3.5) to find that human errors can vary significantly from 0.2-0.5 μm (dust on item) to 0.9-1.3 μm (mechanical stress and CMM making noise). I assume that the users performing the form error measurement have received adequate training to detect CMM faults and other related issues. Furthermore, the standard deviation associated with the random effects of the CMM measurement process has been set to $\sigma_\epsilon = 0.1$ μm. Integrals and other computations were performed in Mathematica and MATLAB, and converted to graphs and tables in Microsoft Excel. Table 3.2 summarizes each form error model based on the posterior density obtained (if analytically available), which is used in this section's discussion. The findings here will describe the impact of Bayesian inference and FMECA on a product's conformity based on its specifications.

**Table 3.2:** Summary of heavy-tailed form error models. *Only requires the largest absolute form error value, e.g. $F_0 = \max|f_i|$. **Extended model (uniform form errors with random effects). ***Augmented model (uniform form errors with environmental, human and random effects).

| Model | Section | Likelihood | Posterior PDF | Parameters |
|---|---|---|---|---|
| Pareto | 3.3.1 | Uniform | $mF_0^m/F^{m+1}$ | $m$, $F_0$ |
| Extended** | 3.3.2 | Uniform-Gaussian mixture (convolution) | See (3.12) | $m$, $F_0$, $\sigma_\epsilon$, $\boldsymbol{d}$ |
| Augmented*** | 3.3.3 | Uniform-Gaussian mixture (convolution) | See (3.14) | $m$, $F_0$, $\boldsymbol{d}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ |

### 3.6.1 Description of the CMM measurement process

The measuring process for the CMM setup is broken down into several parts, as shown in Figure 3.10. The first step is to ensure that the operator has the necessary training and experience to execute the measurements. The operator must be able to operate the appropriate CMM software/hardware. The next stage is CMM initialization of the software/hardware. For the CMM measuring procedure, a companion software (Mitutoyo PartManager) is used. The workpiece is then placed on the CMM's surface, with a thermometer touching it to determine its temperature. The CMM must then be calibrated to confirm that its geometry and moving parts are in good working order and that it is safe to measure with. After the calibration is complete, it is important to select what needs to be measured by the CMM. In this case study, a circular steel alloy item will be assessed at room temperature. A probe calibration test is then conducted to confirm that the workpiece is positioned accurately in the X, Y, and Z planes. The probe must be properly inserted and the workpiece should be (mostly) dust free as well as free of odd finger prints. A 5-point circle measurement can then be performed using the CMM to determine the workpiece's estimated geometry. It's critical to use the CMM controller's "Go To" (or similar) feature to record exact coordinates. Because the same conditions will be used throughout the measurement method, repeating the measurement will be faster and more reliable. Use the software's "CMM repeat mode" after all of the measurement algorithms have been saved. This macro does automatic measurement based on the measurement algorithm. The CMM will halt and/or restart the procedure if there are significant failures when in repeat mode. All measurement errors (related to the probe, the mechanics, the measurement approach, or the environment) are digitally stored using the CMM companion software.

### 3.6.2 Form error model validation

In this section, I compare each form error model using their posterior densities which are summarised in Table 3.2. I am interested in understanding the uncertainty behavior of the posterior densities obtained from the actual error data obtained from the CMM after measuring the product. Figure 3.11 shows posterior densities of three models (Pareto, Extended and Augmented) using real form error data. The following 10 form error measurements taken from the data points are 0.74, -0.46, -0.26, 0.14, -0.56, -1.86, 0.04, 0.04, 1.84 and 0.34 μm. Table 3.3 presents the posterior mean, median, standard deviation, lower and upper quartiles, extreme upper percentiles, a 95% credibility interval (and a conditional probability) for the observed form error data fitted to each of the three models (Pareto,

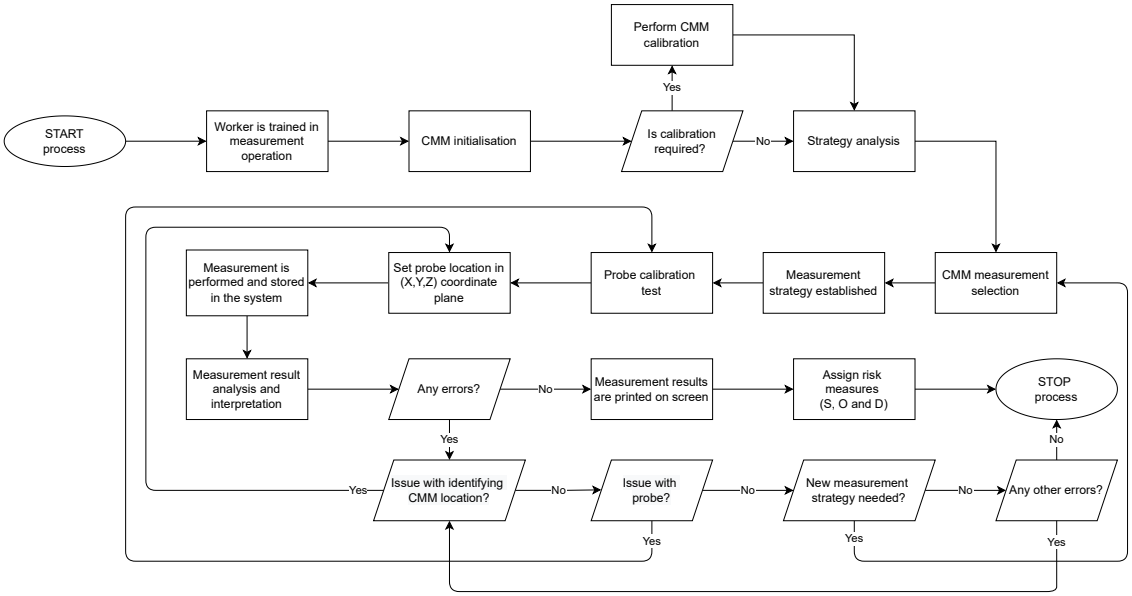



**Figure 3.10:** A flow chart for a CMM measurement process.

Extended and Augmented). The capability index and $\sigma_m$ is computed using (3.22) and (3.19), respectively. Here, I find that all three distributions are consistent. For the observed data, the minimum absolute form error is $F_0 = 1.86$ µm. The Pareto has the narrowest peak, while the Extended and Augmented models have a shape and pattern that is quite similar. The Extended model has a slightly broader peak than the Augmented and this reflects the measurement uncertainty due to $\sigma_\epsilon$ only. The Extended and Augmented models have a combined measurement uncertainty of 1.0 and 1.5 µm. The means of the Extended and Augmented models are 2.182 and 2.130 µm, respectively, which are higher than the Pareto mean of 2.067 µm. Similarly, the standard deviations for both distributions are greater than the Pareto, at 0.327 and 0.376 µm, respectively, compared to 0.231 µm. A basic distinction between the three models is that the Pareto ignores measurement uncertainty since it is not data-dependent and there is no random component. Since the Extended and Augmented models are both data-dependent, they reflect measurement uncertainty and reality better and should therefore be used over the Pareto. The Pareto model has an advantage over the other two since it is easier and quicker to calculate and provides a nice lower bound on the distribution of form errors. The form error distribution's exact shape is likely to adopt an Extended or Augmented posterior distribution. Although all three distributions take little or no account of the probability values when $F < F_0$, I find that the Augmented model takes the greatest amount of measurement uncertainty into consideration to provide a the smoothest and longest Pareto-tailed cut-off.

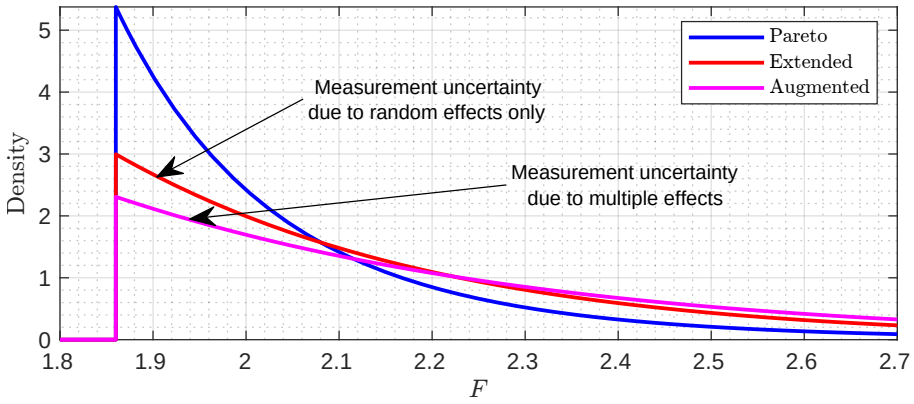



**Figure 3.11:** Posterior densities (Pareto, Extended and Augmented) for the observed form error for the case of $m = 10$ and $F_0 = 1.86$ µm.

**Table 3.3:** Posterior statistics (using Pareto, Extended and Augmented) for the observed form error data for the case of $m = 10$ and $F_0 = 1.86$ µm. The capability index is $C_m = 3.87$.

| Model | Mean | Median | SD | Lower Q. | Upper Q. | 90th Per. | 95th Per. | 99th Per. | 95% Cred. Int. |
|---|---|---|---|---|---|---|---|---|---|
| Pareto | 2.067 | 1.994 | 0.231 | 1.914 | 2.137 | 2.342 | 2.510 | 2.948 | (1.865, 2.690) |
| Extended | 2.182 | 2.084 | 0.327 | 1.953 | 2.304 | 2.596 | 2.821 | 3.349 | (1.868, 3.051) |
| Augmented | 2.230 | 2.118 | 0.376 | 1.969 | 2.368 | 2.700 | 2.959 | 3.605 | (1.870, 3.229) |

### 3.6.3 A critical risk assessment using FMECA

One of the main approaches for identifying failure modes (FMs) and prioritizing possible risk according to failure causes (FCs) is to conduct a risk analysis using FMECA. The top FCs and FMs are calculated based on severity, occurrence, and detection ratings by analyzing the FMECA output. In the critical risk assessment using Bayesian FMECA, assume that a production line manufactures circular parts that are measured, resulting in form error estimates that follow a Pareto-like distribution. The measured value is compared to a tolerance interval to determine whether the form error complies with the specification. For this analysis, I matched the following risk measures:

- $O \sim$ What are the likely causes of failure from a CMM measurement process? I determine this using a posterior probability from a Bayesian network (BN) framework (see subsubsection 3.6.3.2).
- $S \sim$ What is the magnitude of the end-user risk? I determine this using the specific consumer's risk $R_c^S$.
- $D \sim$ What are the chances that current detection measures will prevent the FM/FC? I determine this using the posterior standard deviation of the augmented model in (3.14).

The FCs, FMs and detection methods are adapted from the available literature [104, 105]. The values of $O$, $S$ and $D$ are given according to a 10-point occurrence scale in Table 3.4.

**Table 3.4:** Occurrence, severity and detection rating/criteria associated with the CMM measurement process.

| **Occurrence rating,** $O$ | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Occurrence: FC probabilty** | *Remote* | *Extremely unlikely* | *Very unlikely* | *Unlikely* | *Low* | *Moderately low* | *Moderate* | *Likely* | *Very likely* | *Frequent* |
| **Posterior probability of FC occurring** | <2 % | <5% | <7.5% | <10% | <12.5% | <15% | <20% | <25% | <50% | ≥ 50% |
| **Severity rating,** $S$ | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Severity: Risk level** | *Remote* | *Very low* | *Low* | *Moderately low* | *Moderate* | *Moderately high* | *High* | *Very high* | *Extreme* | *Severe* |
| **Specific consumer's risk (%)** | <2% | <5% | <7.5% | <10% | <12.5% | <15% | <17.5% | <20 % | <25% | ≥ 25% |
| **Detection rating,** $D$ | **10** | **9** | **8** | **7** | **6** | **5** | **4** | **3** | **2** | **1** |
| **Detection probability** | *Remote* | *Very Low* | *Low* | *Unlikely* | *Occasional* | *Moderately likely* | *Likely* | *Very likely* | *Almost certain* | *Certain* |
| **Posterior SD of form errors (µm)** | <0.6 | <0.7 | <0.8 | <0.9 | <1.0 | <1.1 | <1.2 | <1.3 | <1.4 | ≥ 1.4 |

### *3.6.3.1 Identification of failure modes/causes*

The most failure modes (FMs) and their failure causes (FCs) occur immediately after the "measurement result analysis and interpretation" step. According to the trained users, experts and the available literature [104, 106], four main FMs that affect the CMM measurement process were identified. The FMs are probe failure (FM1), measurement performance failure (FM2), mechanical failure (FM3) and environmental failure (FM4). There were also eleven FCs identified. The FCs are item displaced slightly or greatly (FC1/2), dust on time (FC3), mechanical stress (FC4), wrong probe type used (FC5), valve not set to required pressure level (FC6), temperature set too high or too low (FC7/8), CMM making noise (FC9), probe not fitted properly (FC10) or probe not calibrated (FC11).

### *3.6.3.2 A Bayesian network to determine parameter O*

The BN technique is a probability-theory-based visual tool. A directed acyclic graph is utilized in this method to show the number of variables and their dependent links [107]. The BN is a concept that consists of nodes and arrows that is widely used to explain uncertainty [108]. I used a BN which groups FCs and FMs together in order to anticipate when or where the proportion of failure is likely to occur next (with a given probability). The construction of a BN on the GeNIe software for the FMECA is shown in Figure 3.12. The FCs (blue) are the parent nodes in this case, whereas the FMs (purple) are the children nodes. The resulting occurrence probability is represented by the last node (yellow). Every node has two states: "state 0" (non-failure) and "state 1" (failure). Each FM (child node) is coupled with a conditional probability table that displays the states of its FC (parent nodes). The main idea here is to obtain and update the probability of a given FM for the observed FC. The model can address design of experiment questions about the presence of

an FM in the presence of an FC, such as "What is the probability of a mechanical failure, given that the CMM is making noise?"

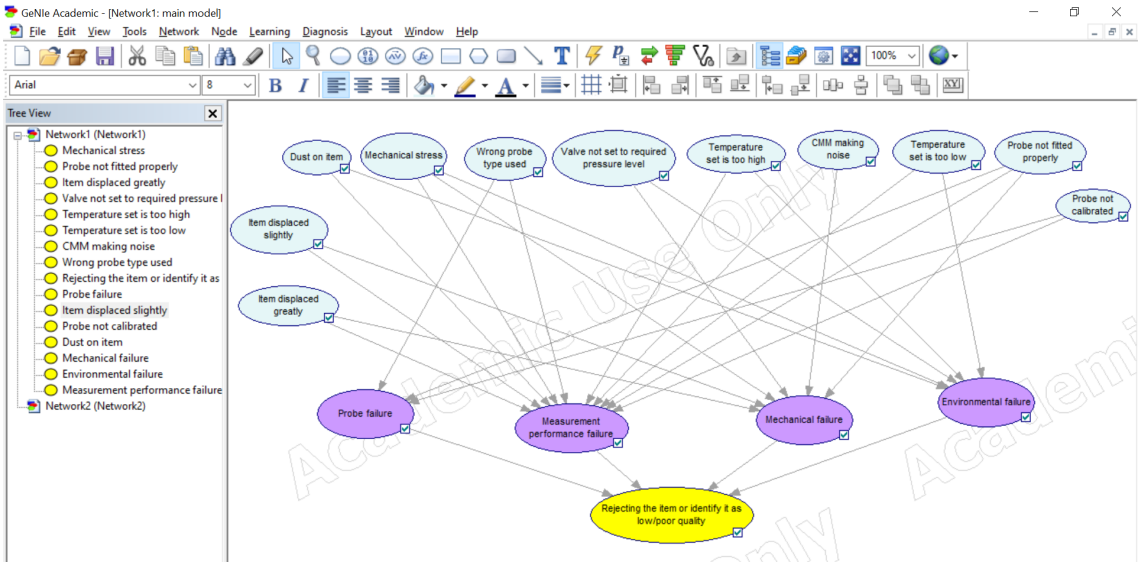



**Figure 3.12:** A Bayesian network on the GeNIe software for the FCs (blue), FMs (purple) and rejection result (yellow).

The calculating process for the posterior probabilities of FC given their respective FMs is shown in the following formula:

$$\pi(\text{FC}|\text{FM}) = \frac{\mathbb{P}(\text{FM}|\text{FC})\pi_0(\text{FC})}{\sum_{\text{FM}} \mathbb{P}(\text{FM}|\text{FC})\pi_0(\text{FC})}. \tag{3.24}$$

where $\pi_0(\cdot)$ and $\mathbb{P}(\cdot)$ are the prior and conditional probabilities, respectively. The prior probability for the FC is calculated by taking the number of occurrences of the $j$-th FC as a fraction of the total number of occurrences over all FCs, i.e.

$$\pi_0(\text{FC}_j) = n(\text{FC})_j / \sum_{k=1}^{11} n(\text{FC})_k.$$

The conditional probabilities are derived by dividing the frequency of FCs by the sum of the FM groups in which they occur:

$$\mathbb{P}(\text{FM}_i|\text{FC}_j) = n(\text{FM}_i|\text{FC}_j) / \sum_{k \in \text{FM}} n(\text{FM}_i|\text{FC}_k),$$

Table 3.5 shows the implications of using prior and conditional probabilities to calculate the posterior probabilities for FCs given a FM. The top row are the FMs with respective prior and posterior probabilities for each FC. The combined FM accumulates all four probabilities for convenience. From this table, the combined posterior probabilities show that "probe not fitted properly" (34.8%, $O = 9$), "wrong probe type used" (19.9%, $O = 7$) and "probe not calibrated" (12.0%, $O = 5$) are the three leading FCs throughout the CMM measurement process.

**Table 3.5:** A critical risk assessment to rating the "occurrence of failure" likelihood using a BN approach to FMECA. Probabilities for FC/FM associated with a CMM measurement process are calculated according to (3.24) using expert judgments.

| | Mechanical | | Probe | | Measurement | | Environmental | | Combined | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Failure cause** | **Prior** | **Posterior** | **Prior** | **Posterior** | **Prior** | **Posterior** | **Prior** | **Posterior** | **Prior** | **Posterior** | **Priority** | $O$ |
| Probe not fitted properly | 5.4% | 7.5% | 9.2% | 19.0% | 7.9% | 8.4% | N/A | N/A | 22.5% | 34.8% | 1 | 9 |
| Wrong probe type used | N/A | N/A | 7.0% | 11.1% | 8.1% | 8.8% | N/A | N/A | 15.2% | 19.9% | 2 | 7 |
| Probe not calibrated | N/A | N/A | 6.0% | 7.9% | 5.5% | 4.1% | N/A | N/A | 11.5% | 12.0% | 3 | 5 |
| Mechanical stress | 4.3% | 4.8% | N/A | N/A | 6.0% | 4.7% | 1.9% | 0.5% | 12.2% | 10.1% | 4 | 5 |
| CMM making noise | 3.6% | 3.4% | N/A | N/A | 4.0% | 2.1% | N/A | N/A | 7.6% | 5.5% | 5 | 3 |
| Temperature is set too low | N/A | N/A | N/A | N/A | 1.0% | 0.1% | 6.3% | 5.3% | 7.3% | 5.4% | 6 | 3 |
| Temperature is set too high | N/A | N/A | N/A | N/A | 0.9% | 0.1% | 5.5% | 4.1% | 6.4% | 4.2% | 7 | 2 |
| Dust on item | N/A | N/A | N/A | N/A | 1.1% | 0.2% | 4.7% | 2.9% | 5.7% | 3.1% | 8 | 2 |
| Item displaced greatly | 2.7% | 1.9% | N/A | N/A | 1.6% | 0.4% | N/A | N/A | 4.3% | 2.2% | 9 | 1 |
| Valve not set to required pressure level | 1.1% | 0.3% | N/A | N/A | N/A | N/A | 2.9% | 1.1% | 4.0% | 1.4% | 10 | 1 |
| Item displaced slightly | 2.2% | 1.2% | N/A | N/A | 1.1% | 0.2% | N/A | N/A | 3.2% | 1.4% | 11 | 1 |

### *3.6.3.3 Determining parameters $S$ and $D$*

I now wish to find the parameter values for the $S$ and $D$ parameters in the second stage. These estimations were made using the specific consumer's risk and the posterior standard deviation of the form error model, respectively. The experts computed the form error measurements under controlled conditions (to assure the validity of the results) for each of the eleven FCs found in Table 3.5 and reported their findings in Table 3.6. For simplicity, each FC is numbered from 1 to 11 using the "priority" index from Table 3.5. The contributions $c_j$ (mm) for $j = 1, \ldots, 4$ are calculated using (3.6).

**Table 3.6:** Observed form error measurements $f_i$ (µm) and the change in form error assessments $h_i$ (µm) after repeated experiments for $i = 1, \ldots, 5$. Experiments were conducted using a CRYSTA-Apex V544 CNC CMM in a metrology lab. Contributions $c_j$ (mm) for $j = 1, \ldots, 4$ are evaluated using (3.6). Parameters: $m = 5$ and $\sigma_\epsilon = 0.1$ µm *Determined from numerous hands-on experiments.

| Priority | Failure cause | $O$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $h_1$ | $h_2$ | $h_3$ | $h_4$ | $h_5$ | Temp. (°C) | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Probe not fitted properly. | 9 | -2.25 | -1.88 | 0.46 | -1.37 | 2.84 | 0.32 | 0.19 | -0.09 | -0.19 | 0.23 | 20.9 | 3.89E-04 | 4.68E-05 | 9.00E-08 | 3.12E-05 |
| **2** | Wrong probe type used. | 7 | -0.31 | 1.65 | -3.30 | 1.25 | -0.60 | 0.03 | 0.15 | 0.53 | 0.00 | 0.10 | 20.6 | 2.59E-04 | 4.68E-05 | 6.00E-08 | 3.12E-05 |
| **3** | Probe not calibrated. | 5 | -1.83 | 1.97 | 1.37 | -1.39 | 2.94 | -0.05 | -0.04 | 0.09 | -0.04 | 0.12 | 20.6 | 2.59E-04 | 4.68E-05 | 6.00E-08 | 3.12E-05 |
| **4** | Mechanical stress. | 5 | 1.78 | 0.91 | -2.88 | -1.46 | -0.13 | -0.36 | -0.08 | -0.06 | -0.31 | 0.03 | 20.7 | 3.02E-04 | 4.68E-05 | 7.00E-08 | 3.12E-05 |
| **5** | CMM making noise. | 3 | -0.65 | 1.58 | -3.13 | 0.79 | 1.37 | -0.02 | 0.02 | -0.04 | 0.10 | -0.18 | 20.7 | 3.02E-04 | 4.68E-05 | 7.00E-08 | 3.12E-05 |
| **6** | Temperature is set too low. | 3 | 2.49 | 2.92 | -1.58 | -1.05 | -0.47 | -0.55 | 0.47 | 0.16 | -0.15 | -0.01 | 17.8 | 9.50E-04 | 4.68E-05 | 2.20E-07 | 3.12E-05 |
| **7** | Temperature is set too high. | 2 | -2.45 | -0.34 | 1.38 | 2.51 | 0.39 | -0.29 | 0.04 | -0.01 | 0.30 | -0.01 | 23.1 | 1.34E-03 | 4.68E-05 | 3.10E-07 | 3.12E-05 |
| **8** | Dust on item. | 2 | -0.63 | 1.24 | -0.37 | 1.12 | 1.85 | -0.09 | -0.06 | -0.06 | -0.22 | 0.02 | 20.6 | 2.59E-04 | 4.68E-05 | 6.00E-08 | 3.12E-05 |
| **9** | Item displaced greatly. | 1 | -2.83 | 1.03 | 1.85 | -1.23 | 2.17 | -0.08 | -0.01 | -0.13 | 0.02 | 0.04 | 20.9 | 3.89E-04 | 4.68E-05 | 9.00E-08 | 3.12E-05 |
| **10** | Valve not set to required pressure level. | 1 | 0.82 | 0.83 | 1.75 | -0.79 | 0.91 | -0.14 | 0.05 | -0.37 | 0.14 | 0.15 | 20.7 | 3.02E-04 | 4.68E-05 | 7.00E-08 | 3.12E-05 |
| **11** | Item displaced slightly. | 1 | -1.02 | 0.78 | 2.23 | -1.04 | -1.16 | 0.07 | 0.07 | 0.24 | 0.22 | 0.21 | 20.7 | 3.02E-04 | 4.68E-05 | 7.00E-08 | 3.12E-05 |

I can now estimate the parameters of $S$ and $D$ using the specific consumer's risk and posterior standard deviation of the form error model, respectively, based on the findings from Table 3.6. The criticality/RPN values are also subsequently calculated by the estimated $S$, $O$, and $D$ parameters. The corrective strategies are determined according to the outcome of the criticality/RPN values. A good measurement strategy will strike a balance between the

cost of reducing measurement uncertainty and the benefit of knowing the true value of the measurand with greater certainty. The calculation of criticality/RPN values allows design of experiment questions to be answered, e.g., how many measurement points $m$ are required to reduce risks to a specified level, or what correction methods are in place to reduce the detection level. Therefore, it is critical to understand the criticality of failures and current corrective measures in place during the CMM measuring process before assessing the RPN value.

A comprehensive breakdown on obtaining these values are given in Table 3.7, which also summarizes the decisions (accept, conditionally accept and reject) made in response to the FCs associated with the CMM measurement process. I find that the posterior standard deviation closely matches the empirical outcome. According to the decision rules, I accept an item if the risk is less than 5%. This corresponds to a severity rating of $S = 1$ or 2. I may also conditionally accept (item is accepted but subject to defects) if the risk is less than 10% ($S = 3$ or 4), and I reject an item otherwise ($S \geq 5$). These ratings reflect the severity of the decisions made, but this could easily be expanded to include many more decisions. As a result, in the next section, I'll look at a discussion on the criticality/RPN values.

**Table 3.7:** Criticality/RPN results of the CMM measurement process which follows from Table 3.6. Parameters: $m = 5$, $T = 5$ µm and $\sigma_\epsilon = 0.1$ µm. *Accept an item if $R_c^S \leq 5\%$, conditionally accept if $R_c^S \leq 10\%$, or reject otherwise.

| Priority | Failure cause | $O$ | $F_0$ | $\mu_h$ | $\sigma_h$ | $\sigma_\theta$ | $\sigma$ | $p_c$ | $R_c^S$ | Decision* | $S$ | Post Mean | Post SD | $D$ | $C$ | RPN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Probe not fitted properly. | 9 | 2.84 | 0.089 | 0.219 | 0.519 | 1.103 | 88.8% | 11.2% | Reject | 5 | 3.877 | 1.132 | 7 | 45 | 315 |
| **2** | Wrong probe type used. | 7 | 3.30 | 0.160 | 0.214 | 0.441 | 1.007 | 83.2% | 16.8% | Reject | 7 | 4.298 | 1.177 | 7 | 49 | 343 |
| **3** | Probe not calibrated. | 5 | 2.94 | 0.016 | 0.084 | 0.441 | 1.050 | 88.0% | 12.0% | Reject | 5 | 3.962 | 1.132 | 7 | 25 | 175 |
| **4** | Mechanical stress. | 5 | 2.88 | -0.155 | 0.167 | 0.469 | 0.974 | 89.7% | 10.3% | Reject | 5 | 3.846 | 1.093 | 6 | 25 | 150 |
| **5** | CMM making noise. | 3 | 3.13 | -0.025 | 0.102 | 0.469 | 1.001 | 86.0% | 14.0% | Reject | 6 | 4.118 | 1.147 | 7 | 18 | 126 |
| **6** | Temperature is set too low. | 3 | 2.92 | -0.016 | 0.375 | 0.771 | 1.272 | 87.6% | 12.4% | Reject | 5 | 3.964 | 1.149 | 7 | 15 | 105 |
| **7** | Temperature is set too high. | 2 | 2.51 | 0.005 | 0.211 | 0.905 | 1.255 | 93.4% | 6.6% | Cond. Acc. | 3 | 3.464 | 1.026 | 6 | 6 | 36 |
| **8** | Dust on item. | 2 | 1.85 | -0.084 | 0.088 | 0.441 | 0.669 | 98.9% | 1.1% | Accept | 2 | 2.453 | 0.690 | 2 | 4 | 8 |
| **9** | Item displaced greatly. | 1 | 2.83 | -0.031 | 0.074 | 0.519 | 1.101 | 90.3% | 9.7% | Cond. Acc. | 4 | 3.785 | 1.084 | 6 | 4 | 24 |
| **10** | Valve not set to required pressure level. | 1 | 1.75 | -0.032 | 0.221 | 0.469 | 0.670 | 99.1% | 0.9% | Accept | 1 | 2.331 | 0.662 | 2 | 1 | 2 |
| **11** | Item displaced slightly. | 1 | 2.23 | 0.161 | 0.083 | 0.469 | 0.830 | 97.0% | 3.0% | Accept | 2 | 2.991 | 0.860 | 4 | 2 | 8 |

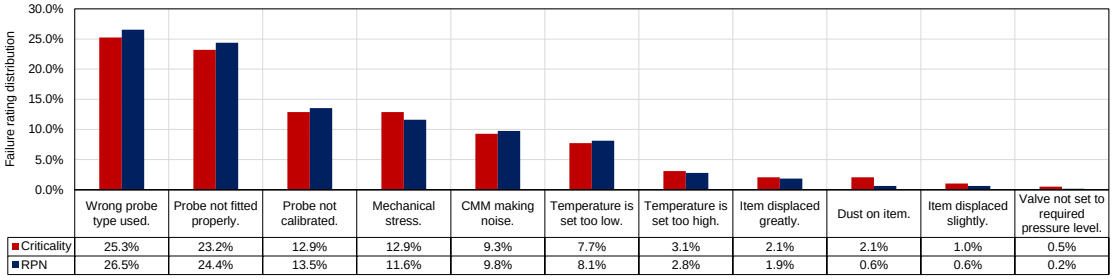
## 3.7 Discussion on the criticality/RPN values

The criticality value, which is derived from the product of the occurrence and severity ratings, is a very valuable indication for determining which failure causes should be prioritized. I prioritise FCs by the highest criticality value. I find that the following are some of the most significant failures causes that can affect the measurement process: "wrong probe type used"

($C = 49$), "probe not fitted properly" ($C = 45$) and "probe not calibrated / mechanical stress" ($C = 25$ each). The RPN should be taken into account as well because it combines the current criticality level with the detection measures in place. I find that low detection levels rapidly raise the overall RPN. It is crucial to minimize the criticality and increase the likelihood of preventing the failure in order to considerably reduce or anticipate future measurement failures. I find that three of the most significant FCs with the highest RPN are also: "wrong probe type used" (RPN = 343), "probe not fitted properly" (RPN = 315) and "probe not calibrated" (RPN = 175 each). However, "mechanical stress" has a lower RPN to criticality ratio then "probe not calibrated" due to differences in the estimated $D$ parameter for both (i.e. 150/25 vs 175/25), so the highest RPNs are not always the same as those with the highest criticality. A graphical summary of the findings on probable FCs, ranked by the highest criticality/RPN values, are given in Figure 3.13. Note that I can't compare criticality and RPN on the same graph because they're measured on a scale of 1-100 and 1-1000, respectively. To get around this, I took the j-th criticality/RPN for the j-th FC and divided it by the overall criticality/RPN. The formulae are given below:

$$C_j^{\%} = C_j / \sum_{k=1}^{n(\mathrm{FC})} C_k, \quad \mathrm{RPN}_j^{\%} = \mathrm{RPN}_j / \sum_{k=1}^{n(\mathrm{FC})} \mathrm{RPN}_k,$$

where $n(\mathrm{FC})$ is the number of FCs.




| | Wrong probe type used. | Probe not fitted properly. | Probe not calibrated. | Mechanical stress. | CMM making noise. | Temperature is set too low. | Temperature is set too high. | Item displaced greatly. | Dust on item. | Item displaced slightly. | Valve not set to required pressure level. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Criticality | 25.3% | 23.2% | 12.9% | 12.9% | 9.3% | 7.7% | 3.1% | 2.1% | 2.1% | 1.0% | 0.5% |
| RPN | 26.5% | 24.4% | 13.5% | 11.6% | 9.8% | 8.1% | 2.8% | 1.9% | 0.6% | 0.6% | 0.2% |

**Figure 3.13:** Risk rating distribution, ranked by the highest criticality/RPN. This is an extension to Table 3.7.

Overall, employing a mix of distinct Bayesian techniques (i.e., a Bayesian network approach to occurrence calculation, a Bayesian conformity assessment approach to severity calculation and using the posterior standard deviation for detection calculation) to compute criticality/RPN values is innovative and highly convenient, and it extends existing FMECA methods. With more information given to the FMs and FCs, the review team can study the FMECA worksheet, conformity rates, risks, criticality and RPN. In order to reduce risk, the review team can identify potential CMM improvements by assessing the key reasons of failure using the methods and results described above. Design improvements, engineering safety features, safety devices, warning devices, and procedures/training can all help to

lower the risk. As a result of the updated measures, the severity, occurrence, and detection values are reduced, and the RPN value is re-evaluated. The findings of the FMECA review and the data collected can be stored in a database that can be used to analyze similar measurement processes in the future. In this manner, future failures can be avoided. As a result, quality management is enhanced by assessing the risk measures by using this approach.

## 3.8 Conclusion

This chapter was concerned about whether a product meets its specifications by adapting a proposed form error model identified by Bayesian inference to a conformity-based FMECA. Most conformity assessment problems assume that the uncertainty associated with a quantity is Gaussian or Gaussian-like. For form error assessment, the associated distribution is far from a Gaussian. Through Bayesian inference, important posterior distributions were derived (Pareto, Extended and Augmented models) by considering the likelihood of measurement data and a non-informative prior PDF (e.g., a reciprocal function or a Gaussian prior), to describe the proposed form error model. The Pareto is the outcome of the likelihood taking no measurement uncertainty into consideration and the form error estimates being uniformly distributed, resulting in an accurate representation of the model. The Extended model is the result of a uniform-Gaussian mixture distribution by integrating measurement uncertainty from the likely random effects of form errors. The Augmented model improves on the Extended model by accounting for additional sources of measurement uncertainty, such as human and environmental errors. The methodologies considered in this chapter ignore the fact that surface fitting determines form error estimates at the surface points, and so they are subject to a correlated effect associated with the fitting.

A novel way of combining conformity assessment with FMECA was then presented. For the analysis of the FMECA of the CMM measurement process, three risk measures were matched accordingly:

- Occurrence rating, $O$, to the posterior probability of FCs occurring using a Bayesian network approach;
- Severity rating, $S$, to the specific consumer's risk;
- Detection rating, $D$, to the posterior standard deviation of the form error model.

These three values were used to compute the criticality/RPN which prioritises risks/failures

with the highest values. The probe-related FMs were found to be the most significant, followed by measurement, mechanical, and then environmental. I were also interested in predicting when and where the related FCs will occur. I found that the four most significant FCs are: "wrong probe type used" ($C = 49, \mathrm{RPN} = 343$), "probe not fitted properly" ($C = 45, \mathrm{RPN} = 315$), "probe not calibrated" ($C = 25, \mathrm{RPN} = 175$) and "mechanical stress" ($C = 25, \mathrm{RPN} = 150$). The calculated values for criticality and RPN confirmed that probe failures were extremely likely to be the major FM during the measurement procedure as a result of the research.

Results have shown that an integrated approach to Bayesian FMECA (using Bayesian networks, conformity assessment, and posterior statistics) is effective in modeling and studying the risks of CMM measurements. Contributions in this chapter include using heavy tail distributions (e.g. Pareto-like distributions) to assess conformity of items at the form error's upper limit and using non-Gaussian distributions to evaluate specific consumer risks, expanding the possibility of using various probability distributions to achieve desired outcomes. As a result, action steps are quantitatively evaluated, and improvement actions taken in response to critical FMs of a CMM are identified. Furthermore, the methods presented here contribute significantly to the extensive literature in JCGM [92] regarding item conformity assessment.

# Chapter 4

# A Bayesian Risk Assessment of the CMM Measurement Process using an Autoregressive Form Error Model and Gibbs Sampling

## 4.1 Introduction

### 4.1.1 Overview and related works

Recall from chapter 3 that the classical form error model assumes that I can obtain form error estimates $f_i$ from an artefact's ideal geometry about the on-machine probe coordinate data from (at least) two components: orthogonal distances $d_i \in \mathcal{N}(0, \sigma_d^2)$ and the (usually) unavoidable random error $\epsilon_i \in \mathcal{N}(0, \sigma_\epsilon^2)$, modelled in (3.2). This model related to the estimation of form error in the presence of random effects associated with the measurement system. It is critical for manufacturers to assess the effectiveness of CMMs (coordinate measuring machines) by ensuring that the measuring process runs smoothly. For conformity or risk analysis, predicting form errors after measuring an item is critical. Furthermore, the main interest here is in extending the model to account for autoregressive (AR) lags using failure prediction methods and a time series model.

Failure prediction is required for predictive maintenance due to its capacity to reduce failure incidents and maintenance costs. Consistent with modeling assumptions, work has been done in applying the Bayesian framework to various fields, such as model updating,

reliability analysis, modal analysis, inverse modeling, risk detection, and improving the performance of construction projects [109–112]. The Bayesian network was employed by Abu Samah *et al.* [113] to forecast failure probabilities. While the research appeared to be fascinating, they did not release the data set they used, making it difficult to duplicate or compare their proposed model to existing machine learning algorithms. Currently, the most common methodologies for failure prediction are mathematical and statistical modeling, which are based on physical models and machine learning techniques [114].

The most widely used tools in the industry, by far, are FMEA (Failure mode and effects analysis) and FTA (Fault tree analysis) [49]. FTA is a top-down analysis that uses the notion of Boolean logic to generate a set of True/False assertions to visually represent a failure path or failure chain. It uses logic gates and events to simulate how component states affect the overall state of the system, which is extremely beneficial for failure prediction. The goal of FTA is to quickly identify failure causes and mitigate risks before they occur. The OR-gate, AND-gate, and inhibit or conditional gate are all commonly used logic gates in FTA [50]. Furthermore, by including quantitative component reliability data (e.g., failure rates), a strictly qualitative FTA can be turned into a quantitative FTA. A quantitative FTA can be used to calculate the system's reliability using boolean algebra. Hamda [51] provides a detailed description of largely quantitative FTAs. Some authors argue that FTA and FMEA should be used together. For example, Bertsche [52] claims that due to the different starting points of both methods, this can increase the number of failure modes found (e.g. FMEA from the bottom up vs. FTA from the top down). Even so, completing both analyses would be time intensive and could lead to a lack of focus on the most critical system components, which is what a failure analysis is designed to reveal. In light of the literature, I propose an integrate FTA/FMEA approach to the Bayesian AR model in this study to predict significant failures throughout the CMM measurement process (this is further discussed in this chapter).

Time series models have been used for prediction in all sectors of study for decades [115]. From linear to non-linear regression, models such as AR, moving average (MA), and exponential smoothing, as well as a slew of others, are available. A famous time series model termed autoregressive integrated moving average (ARIMA) was created by Box and Jenkins. These techniques have been effectively implemented in a variety of fields, including data centers, complex industrial systems, transportation networks, and healthcare, to forecast system failure [116, 117].

A popular model which performs similarly to the classic AR models is the median

autoregressive (MAR) model which is a special case of quantile autoregressive (QAR) models used for predicting the conditional quantile $Q_p$ where $p \in (0, 1)$ in which the median regression occurs when $p = 1/2$. The median regression model is equivalent to the mean regression if the conditional distribution is symmetric; otherwise, they are different. The mean regression model can be thought of as a summation of all the quantile effects, therefore QAR provides a thorough examination of how future data given past data are related. Quantile regression has a lot of applications in risk measurement, and QAR is just an extension of the work done in Yu and Moyeed [118], where the idea of Bayesian quantile regression using a likelihood function based on the asymmetric Laplace distribution was introduced. Bayesian MAR and QAR models, in which the AR coefficients can be represented as monotone functions of a single, scalar random variable, are of particular interest. In addition, simple examples of AR modelling using Bayesian inference can be found online (i.e., using Julia with Jupyter notebook for plots and outputs [119]). However, there has been a lack of Bayesian inference in regards to forecasting data points of a measurement process given model parameters and/or decision framework. Specifically, there is no available literature on an AR model based Bayesian method that can identify both modal parameters and their uncertainties for a coordinate measuring machine (CMM) process that can be used efficiently for quality engineering.

The aim of this study is to create an accurate model for predicting future measurement failure trends by focusing on a measurement criterion such as form error. In this section, I propose a new model and method for forecasting form errors for decision making. The proposed model is described using an AR model with Bayesian inference for parameter estimation of form error data. The main goal is to approximate the coefficients' posterior distribution ($\alpha$, $\beta_1$, ..., $\beta_q$ and $\sigma_\epsilon^2$). I must marginalize each parameter of this posterior distribution related to each coefficient in order to determine the posterior distribution. This requires calculating marginal distributions, which can be difficult to do analytically in practice. This is where Gibbs sampling comes in handy. The Gibbs sampler is an example of Markov chain Monte Carlo (MCMC) that allows one to estimate the joint marginal distribution using draws from the conditional distribution. Following this, I can then build a form error forecast from the algorithm's posterior density. This makes the forecast modelling extremely useful in making decisions. The algorithms and coding has all been done on RStudio.

The study will show that the combination of Bayesian inference, AR modelling, an integrated FTA/FMEA and Gibbs sampling is effective in modelling and studying the

predictions of CMM measurements, and thus improve reliability of future measurements for product conformity assessment.

### 4.1.2 Coverage and data collection

Real-time CMM form error data was collected using 500 points on the outer surface of a circular workpiece from the CMM's automated circle measurement function.

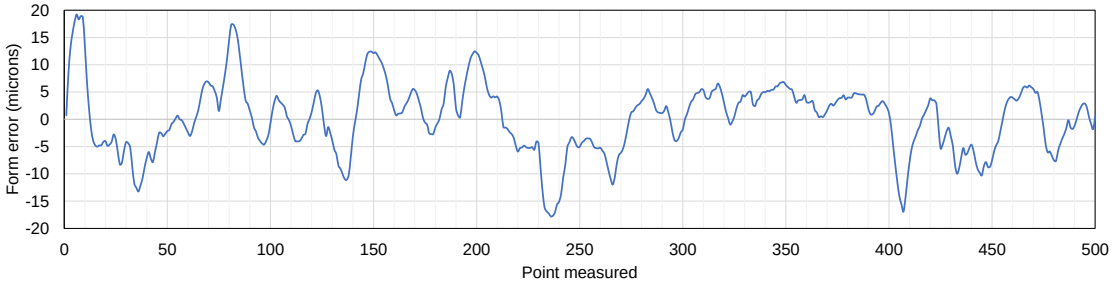



**Figure 4.1:** Empirical form error data obtained for 500 points using the scanned feature of a CMM. The observed radius of the workpiece is 19.9823 mm. 1 micron is equal to 1/1000 mm. Item measured: a smooth circular metal plate. Measured points were taken using a Renishaw SP25M probe on a CMM.

The plot shows that the form error distribution is highly volatile about zero, with several highly correlated neighboring points and symmetrically uniform movements with some randomness. The mean for the sample form error is approximately zero, with approximately evenly spaced ranges between maximum and minimum form error points, indicating that the distribution of form errors is highly uniform with a symmetrical structure. I can associate the randomness of such movements with a number of different symmetric unimodal distributions to describe the form error distribution such as an uniform or Gaussian distribution.

## 4.2 Model and methods

### 4.2.1 The proposed model

In this section, I propose a new model for forecasting form errors. To begin, I propose an AR model with Bayesian inference for parameter estimation of form error data. This model is an alternative to the classic mean-based AR models. An AR model with $q$ lag conditions is referred to as the AR($q$) model. The lag parameter $q$ specifies how many measured points should be used at an instant. For a given form error model denoted by $f_m$, with $m$ measured points, I get the following regression structure for an AR($q$) model:

$$f_m = \alpha + \beta_1 f_{m-1} + \beta_2 f_{m-2} + \cdots + \beta_q f_{m-q} + \epsilon_m =: \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{f} + \epsilon_m, \tag{4.1}$$

where $\alpha$ is a mean/median centred constant, $\beta_i$ are parameters of the model, $\epsilon_i$ is the random error, and $\boldsymbol{\beta} = [\alpha, \beta_1, \ldots, \beta_q]$, $\boldsymbol{f} = [1, f_{m-1}, \ldots, f_{m-q}]$ are vectors of unknown coefficients and form error data points, respectively. The superscript T denotes the transpose. The AR model assumes that the random error is subjected to a white noise process with mean zero and variance $\sigma_\epsilon^2$, i.e, $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

Regression forecasts are frequently obtained using AR models. This is because each successive point observed by a CMM resembles a time-series moving around a mean value of zero has a strongly correlated structure, and so this is possible. Regression, often known as a regression model, is a statistical technique for modeling the conditional expectation of the dependent variable given a vector of independent variables, i.e. $\mathbb{E}[f_m|\boldsymbol{f}]$. This can be used to determine the dependency or effect, as well as to make predictions. I am interested in the forecast of point-varying probability of an item exceeding a given tolerance level, and the calculation of $p$ given $Q$ such that $p = \mathbb{P}(f_m \leq Q|\boldsymbol{f})$.

### 4.2.2 Bayesian inference of AR parameters

The main goal is to approximate the coefficients' posterior distribution: $\alpha$, $\beta_1$, ..., $\beta_q$ and $\sigma_\epsilon^2$. I can describe the posterior distribution of the above model in a more succinct way using Bayesian inference:

$$\pi(\boldsymbol{\beta}, \sigma_\epsilon^2|\boldsymbol{f}) \propto \mathcal{L}(\boldsymbol{f}|\boldsymbol{\beta}, \sigma_\epsilon^2)\pi_0(\boldsymbol{\beta}, \sigma_\epsilon^2), \tag{4.2}$$

where $\mathcal{L}(\boldsymbol{f}|\boldsymbol{\beta}, \sigma_\epsilon^2)$ is the likelihood of the data and $\pi_0(\boldsymbol{\beta}, \sigma_\epsilon^2)$ is the prior distribution. When deriving the likelihood of (4.1), it may be easier to write:

$$\epsilon_m = f_m - \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{f}. \tag{4.3}$$

Therefore, I can use (4.3) to define the likelihood function according to a Gaussian distribution:

$$\mathcal{L}(\boldsymbol{f}|\boldsymbol{\beta}, \sigma_\epsilon^2) = (2\pi\sigma_\epsilon^2)^{-1/2} \exp\left(\frac{(f_m - \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{f})^2}{2\sigma_\epsilon^2}\right). \tag{4.4}$$

I can define the priors for the Bayesian analysis of $\boldsymbol{\beta}$ by using a Gaussian distribution with a zero-mean vector $\boldsymbol{\mu}_0 = \mathbf{0} = (0, 0, 0)^{\mathrm{T}}$ governed by a single variance parameter $\sigma_\epsilon^2$, such that $\boldsymbol{V}_0 = \sigma_\epsilon \boldsymbol{I}$ where $\boldsymbol{I}$ is the identity matrix. I have also used a scaled inverse chi-squared (SICS) distribution as a prior to model the initial variance parameter $\sigma_\epsilon^2$ which are parametrised by the two quantities, $m_0 > 0$ and $\tau_0^2 > 0$, representing the number of chi-squared degrees of freedom and the scaling parameter, respectively. Specifically, I

denote these priors by

$$\pi_0(\boldsymbol{\beta}, \sigma_\epsilon^2) = \pi_0(\boldsymbol{\beta}|\sigma_\epsilon^2)\,\pi_0(\sigma_\epsilon^2),$$

$$\pi_0(\boldsymbol{\beta}|\sigma_\epsilon^2) \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{V}_0) = \mathcal{N}(\mathbf{0}, \sigma_\epsilon \boldsymbol{I}), \tag{4.5}$$

$$\pi_0(\sigma_\epsilon^2) \sim \text{Scaled-Inv-}\chi^2(m_0, \tau_0^2) \equiv \text{IG}(m_0/2, m_0\tau_0^2/2). \tag{4.6}$$

It should be noted that the inverse gamma (IG) distribution is a special case of the SICS distribution, which uses a different parametrization to define the same data distribution. Since the IG distribution is provided by more programming languages and statistical packages than the SICS distribution, I'll use it instead. In any case, both distributions are suitable priors to use because they only support positive numbers and can also be used as a conjugate prior in Bayesian statistics. For computation simplicity, I set $\zeta_0 = m_0\tau_0^2$.

The posterior distribution of the parameters based on the data is proportional to the likelihood function multiplied by the prior distribution of the coefficients. I'll assume that both the likelihood function and the prior distribution are Gaussian, which means that the posterior will be Gaussian as well [120]. A posterior update to (4.5) and (4.6) are outlined in the implementation phase.

To obtain individual posterior distributions, i.e. $\pi(\alpha|\boldsymbol{f}), \ldots, \pi(\sigma_\epsilon^2|\boldsymbol{f})$, I must marginalize each parameter of this posterior distribution related to each coefficient in order to determine the posterior distribution. This requires calculating marginal distributions, which can be difficult to do analytically in practice. The Gibbs sampling method is applied here. The Gibbs sampler is an example of MCMC that allows one to estimate the joint marginal distribution using draws from the conditional distribution. However, I must first understand the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the series, which are required to determine the order of AR terms, before selecting candidate AR models for time series analysis and forecasting. The process and findings are described in the sections that follow.

### 4.2.3 Analysing ACF and PACF Plots

To completely define the AR model, I must choose a value for $q$ that best fits the observed form error data while also ensuring that the fundamental assumptions of this class of models are true [121]. The most critical assumption is that the error term is subjected to a White Noise Process, i.e. $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$. I'll characterize the data's ACF and PACF to find the best model for the form error. The technique would be to choose models that closely align the trends of the sample autocorrelation function (SACF) and sample partial

autocorrelation function (SPACF) with the theoretical model's ACF and PACF. The word "auto" comes from the fact that I am computing the correlation of a sequence of measured points with the values of that same series.

The theoretical definition of the $q$-th order ACF, denoted by $\rho_q$, is given by:

$$\rho_q = \frac{Cov(f_m, f_{m-q})}{Var(f_m)} = \frac{\gamma_q}{\gamma_0}. \tag{4.7}$$

In practice, the SACF or the correlogram is obtained from an observed form error sample:

$$\hat{\rho}_q = \frac{\hat{\gamma}_q}{\hat{\gamma}_0} = \frac{\frac{1}{M}\sum_{i=m+1}^{M}(f_i - \bar{f})(f_{i-q} - \bar{f})}{\frac{1}{M}\sum_{i=m+1}^{M}(f_i - \bar{f})^2}, \tag{4.8}$$

where $M > m$ is a predicted measured point. The PACF is defined as the last coefficient in the $q$-th order autoregression in an AR($q$) model, i.e. $\beta_q$, for $q \geq 1$. The main idea is that the PACF measures the additional correlation between $f_{i-1}$ and $f_{i-q+1}$ after removing the linear dependence of the intermediate lags. The ratio of matrix determinants gives the explicit formulas for $\beta_q$ as a function of the autocorrelations, $\rho_q$, as follows:

$$\beta_q = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{q-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{q-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{q-1} & \rho_{q-2} & \rho_{q-3} & \cdots & \rho_1 & \rho_q \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{q-2} & \rho_{q-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{q-3} & \rho_{q-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{q-1} & \rho_{q-2} & \rho_{q-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \tag{4.9}$$

I can get the SPACF by doing the following:

1. Load the CMM data as CSV file.
2. Estimate AR models, i.e., from (4.1), in stages using orders 1, 2, ..., and pick out the last coefficient at each point using OLS.
3. Replace the $\rho_i$ with their sample equivalents $\hat{\rho}_i$ using (4.9).

I start by importing the form error data into RStudio (a R-based integrated development environment). The ACF and PACF plots for the training data are shown in Figure 4.2. In general, the following findings in the ACF and PACF plots indicate an autoregressive model:

- A positive autocorrelation at lag 1.

- The ACF plot steadily decreases to zero.
- At some lag $k$, the PACF cuts off sharply. I should use this value of $k$ in the AR($q$) model since it equals the value of $q$.

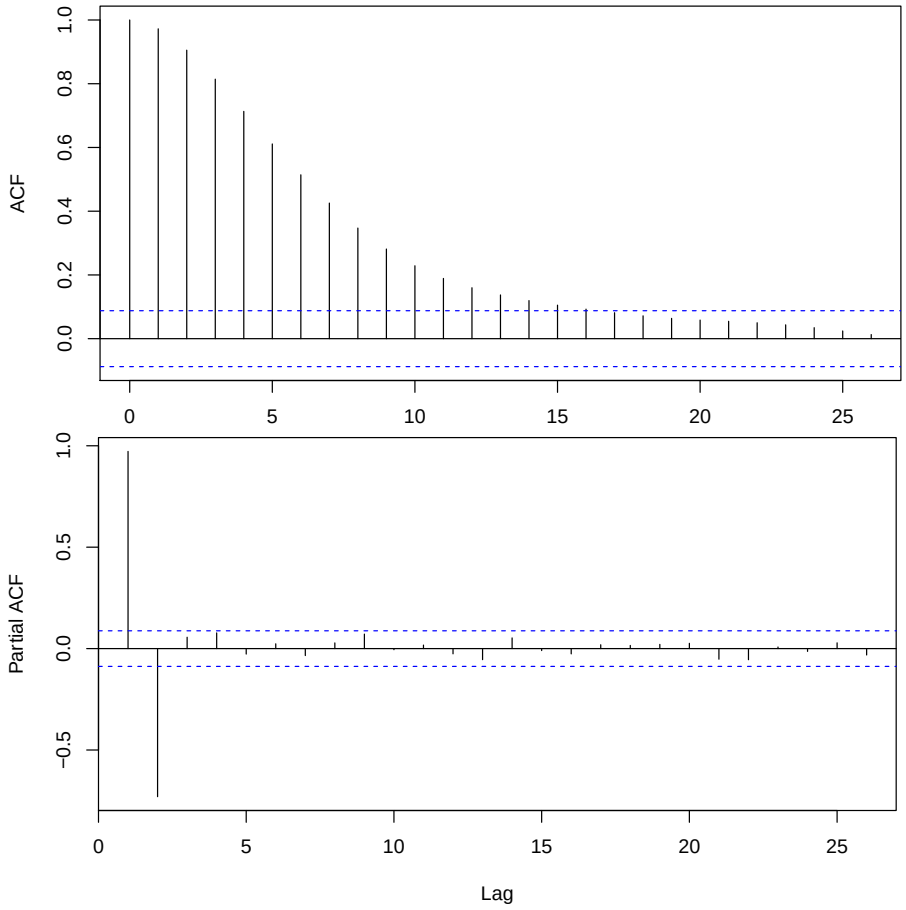



**Figure 4.2:** ACF and PACF plots sampled from real form error data using 500 consecutive measured points from a CMM's scanned process.

It can be seen that the ACF and PACF plots meet the requirements in the three points above. Furthermore, since the PACF plot cuts off at the second lag, the value of $q$ I obtain for the AR($q$) model is $q = 2$ implying that the form error model follows an AR(2) process with equation:

$$f_m = \alpha + \beta_1 f_{m-1} + \beta_2 f_{m-2} + \epsilon_m, \quad \epsilon_m \sim \mathcal{N}(0, \sigma_\epsilon^2). \tag{4.10}$$

Finally, to ensure that an AR(2) process is stable, then $\alpha$, $\beta_1$ and $\beta_2$ should satisfy the following stationarity conditions:

$$\beta_2 < 1 \pm \beta_1 - \alpha, \quad \beta_2 > \alpha - 1. \tag{4.11}$$

#### 4.2.3.1 *Stationarity conditions for an AR(2) process*

For the proof to (4.11), I'll use the lag or backshift operator $B$ to quickly extract a range of results relevant to AR processes. The lag operator allows us to write a general AR($q$) process in a more concise form:

$$\beta(B)d_m = \epsilon_m, \tag{4.12}$$

where $\epsilon_m$ is a standard white noise process and $\beta(B)$ is

$$\beta(B) = 1 - \sum_{i=0}^{q} \beta_i B^i, \quad \beta_0 =: \alpha. \tag{4.13}$$

*Proof.* I study the AR(2) process's MA representation to extract the stationarity conditions. To do so, I must take the following steps:

1. Factor the polynomial $\beta(B)$.
2. Use the lag operator property in (4.13).

Consider the expected value of an AR(2) process described by the following characteristic equation:

$$\mathbb{E}[\beta(B)d_m] = 1 - \alpha - \beta_1 B - \beta_2 B^2. \tag{4.14}$$

If $B$ is the root of the characteristic equation, then the output obtained by rewriting the above is $\mathbb{E}[\beta(B)d_m] = 0$. I can derive the stationarity triangle of an AR(2) using this representation, which states that an AR(2) is stable if the AR polynomial's roots of $\beta(B)$ are outside the unit circle, i.e. $|B| > 1$. Another criterion for AR(2) stability is that all roots of are contained within the unit circle, i.e. $|\lambda| = |B^{-1}| < 1$, with characteristic equation $(1-\alpha)\lambda^2 - \beta_1\lambda - \beta_2 = 0$. Note that the quadratic formula can be used to find the roots of the mean equation:

$$\lambda_{1,2} = \frac{\beta_1 \pm \sqrt{\beta_1^2 + 4(1-\alpha)\beta_2}}{2(1-\alpha)}. \tag{4.15}$$

The larger $\beta_i$ value is bounded by the larger root of $\lambda_{1,2} < 1$, i.e.

$$\begin{aligned}
\frac{\beta_1 + \sqrt{\beta_1^2 + 4(1-\alpha)\beta_2}}{2(1-\alpha)} &< 1 \\
\beta_1 + \sqrt{\beta_1^2 + 4(1-\alpha)\beta_2} &< 2(1-\alpha) \\
\beta_1^2 + 4(1-\alpha)\beta_2 &< (2(1-\alpha) - \beta_1)^2 \\
\beta_2 &< 1 - \beta_1 - \alpha.
\end{aligned}$$

Similarly, I find that $\beta_2 < 1 + \beta_1 - \alpha$. If $\lambda_i$ is a complex number, then $\beta_1^2 - 4(1-\alpha)\beta_2 < 0$, therefore

$$\lambda_{1,2} = \frac{\beta_1 \pm i\sqrt{-\left(\beta_1^2 + 4(1-\alpha)\beta_2\right)}}{2(1-\alpha)},$$

where $i = \sqrt{-1}$ is an imaginary number. A complex number's squared absolute value is equal to the square of the real part plus the square of the imaginary part, thus:

$$|\lambda|^2 = \frac{\beta_1^2}{4(1-\alpha)^2} + \frac{-\left(\beta_1^2 + 4(1-\alpha)\beta_2\right)}{4(1-\alpha)^2} = -\beta_2/(1-\alpha).$$

The AR(2) process is stable if $|\lambda| < 1$ and thus $\beta_2 > \alpha - 1$ is satisfied. □

## 4.3 An integrated approach to FTA/FMEA modelling

I use the Bayesian AR model and the integrated Fault tree analysis (FTA) and Failure mode and effects analysis (FMEA) approach in this study to predict significant failures throughout the CMM measurement process. Failure modes (FMs) are a direct result of the faults identified in the FTA process, and failure causes (FCs) are assumed to be mutually independent in an FMEA. The design of the integrated failure prediction model is shown in Figure 4.3. This set of pre-selected parameters is gathered on a regular basis to see if their values may be used to anticipate CMM measurement failures. The most significant FMs have come from probe, mechanical, measurement performance and environmental, all of which have an impact on the form error estimations. In the following subsections, I will describe the steps required to perform the integrated failure analysis method.

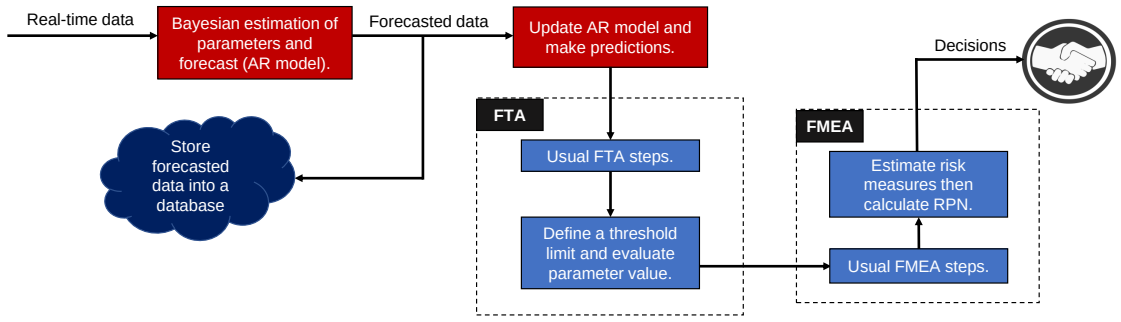



**Figure 4.3:** Failure prediction using Bayesian AR modelling and an integrated FTA/FMEA approach.

### 4.3.1 FTA steps

A FTA evaluates each tree node's state using the concept of detectability index [122]. At measured point $m$, $\hat{f}_m$ is a predicted value produced from the associated parameters of the AR model. The predictors' outputs are converted to binary form using thresholding values and then input into the FTA. The fault tree analyzes the probability of the CMM measurement process failing using binary data. I denote $T$ as the threshold tolerance limit of actual and predicted data. If the actual or predicted form error (e.g., $f_i$ or $\hat{f}_i$) surpasses the threshold, the values given into the tree has a detectability index of $\delta_i = 1$; otherwise, $\delta_i = 0$. This is illustrated in the following:

$$\delta_i = \begin{cases} 1, & |f_i| > T, \\ 0, & |f_i| \leq T, \end{cases} \quad \hat{\delta}_i = \begin{cases} 1, & |\hat{f}_i| > T, \\ 0, & |\hat{f}_i| \leq T, \end{cases} \quad \text{for } i = 1, \ldots, m. \tag{4.16}$$

The main idea is to combine (4.16) with a confusion matrix, which could be used to assess the forecast quality of the proposed Bayesian AR model. In general, a confusion matrix is a performance metric for machine learning classification problems with two or more classes as output [123, 124]. There are four classifications in the matrix of predicted and actual values, namely: $TP$, $TN$, $FP$ and $FN$. For example, the interpretation of $TP$ is that I predicted positive and it turned out to be true. The values of each classification is given below:

$$TP_i = \begin{cases} 1, & \delta_i = \hat{\delta}_i = 1, \\ 0, & \text{otherwise}, \end{cases} \qquad TN_i = \begin{cases} 1, & \delta_i = \hat{\delta}_i = 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$FP_i = \begin{cases} 1, & \{\delta_i = 0 \cup \hat{\delta}_i = 1\}, \\ 0, & \text{otherwise}, \end{cases} \qquad FN_i = \begin{cases} 1, & \{\delta_i = 1 \cup \hat{\delta}_i = 0\}, \\ 0, & \text{otherwise}, \end{cases}$$

Each sub-scripted classification is summed from $i = 1$ to $i = m$ to form the whole classification, e.g. $\sum_{i=1}^{m} TP_i = TP$. These classifications are extremely helpful and can be used to assess a variety of rates. For example, the error, accuracy, prevalence, prediction,

sensitivity and specificity rates can all be calculated as shown:

$$\text{Error rate} = \frac{FP + FN}{TP + FP + TN + FN}, \tag{4.17}$$

$$\text{Accuracy rate} = 1 - \text{Error rate}, \tag{4.18}$$

$$\text{Prevalence rate} = \frac{TP + FP}{TP + FP + TN + FN}, \tag{4.19}$$

$$\text{Prediction rate} = \frac{TP}{TP + FP}, \tag{4.20}$$

$$\text{Sensitivity rate} = \frac{TP}{TP + FN}, \tag{4.21}$$

$$\text{Specificity rate} = \frac{TN}{TN + FP}. \tag{4.22}$$

Some of these measures are extremely useful in the estimation of the unknown parameters of the FMEA risk measures, namely: the severity ($S$), occurrence ($O$) and detection ($D$) ratings. This is further discussed and explained the in the following section.

### 4.3.2 FMEA steps

Aside from the conventional FMEA stages that are outlined and explained in the preceding chapters, additional tasks that should be completed include employing an approach to derive the risk parameters: $S$, $O$ and $D$ for the Bayesian AR model, which are then multiplied together to obtain the RPN [10]. The following is matched according to the confusion matrices classification meeasures:

- Failure severity $S \sim$ Estimated from the error rate in (4.17);
- Failure occurrence $O \sim$ Estimated from the prevalence rate in (4.19).
- Failure detection $D \sim$ Estimated from the prediction rate in (4.20).

Table 4.1 depicts an integrated FTA/FMEA spreadsheet with all relevant data from the recommended methodologies, concluding with suggested preventive measures. Each of these formulas is compared to a severity, occurrence, and detection scale that ranges from 1 to 10. The results section includes a detailed table that highlights the FMEA's ratings and criteria.

**Table 4.1:** An integrated FTA/FMEA form.

| | | | FTA | | | | Classification measures | | | | | FMEA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # | Failure mode | Failure cause | $TP$ | $FP$ | $TN$ | $FN$ | Error (%) | Prevalence (%) | Prediction (%) | Sensitivity (%) | Specificity (%) | $S$ | $O$ | $D$ | RPN | Priority | Action |
| 1 | | | | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | | | | |

## 4.4 Implementation

### 4.4.1 Gibbs Sampling to derive posterior estimates

Now that I've described the form error model using an AR(2) process, I can see how it works in practice. In this section, I will construct some new algorithms and formulas for implementing linear regression using the Gibbs sampler. Using a Bayesian method, I can obtain vital posterior statistics for each unknown parameter then use the model to predict the volatile of CMM measurements. I can accomplish this by using the R code to calculate conditional distributions and collect samples from the posterior distribution within a Gibbs sampling framework, making the entire process smooth and seamless.

The aim is to approximate the marginal posterior distribution of each variable from (4.2) to describe the form error model. In order to obtain these marginals, initialise starting values for the variables:

$$\text{Initial} = \left[\alpha^{(j)}, \beta_1^{(j)}, \beta_2^{(j)}, \sigma_\epsilon^{2(j)}\right], \quad j = 0.$$

Then, sample the first variable $\alpha^{(j)}$ based on the current values of the other variables by setting $j = j + 1$ with the initial $j = 0$. This is then repeated by sampling the second variable $\beta_1^{(j)}$ based on all of the others, and so on until all of the variables (e.g., $\beta_2^{(j)}$, $\sigma_\epsilon^{2(j)}$) have been sampled and posterior densities have been obtained.

$$\alpha^{(1)} \sim \pi\left(\alpha|\beta_1^{(0)}, \beta_2^{(0)}, \sigma_\epsilon^{2(0)}\right), \quad \beta_1^{(1)} \sim \pi\left(\beta_1|\alpha^{(0)}, \beta_2^{(0)}, \sigma_\epsilon^{2(0)}\right),$$
$$\beta_2^{(1)} \sim \pi\left(\beta_2|\alpha^{(0)}, \beta_2^{(0)}, \sigma_\epsilon^{2(0)}\right), \quad \sigma_\epsilon^{2(1)} \sim \pi\left(\sigma_\epsilon^2|\alpha^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}\right).$$

At this stage, the Gibbs sampling algorithm has completed one iteration ($j = 1$). The samples from the conditional distributions converge to the joint marginal distributions as I replicate these procedures a large number of times, e.g. $j = N$ times. With all other parameters fixed, the result is often a known density form, i.e.

$$\alpha^{(N)} \sim \pi\left(\alpha|\beta_1^{(N-1)}, \beta_2^{(N-1)}, \sigma_\epsilon^{2(N-1)}\right), \quad \beta_1^{(N)} \sim \pi\left(\beta_1|\alpha^{(N-1)}, \beta_2^{(N-1)}, \sigma_\epsilon^{2(N-1)}\right),$$
$$\beta_2^{(N)} \sim \pi\left(\beta_2|\alpha^{(N-1)}, \beta_2^{(N-1)}, \sigma_\epsilon^{2(N-1)}\right), \quad \sigma_\epsilon^{2(N)} \sim \pi\left(\sigma_\epsilon^2|\alpha^{(N-1)}, \beta_1^{(N-1)}, \beta_2^{(N-1)}\right).$$

This makes many analyses far simpler. The expected value of the retained draws can be thought of as an approximation of the mean of the posterior distribution after sampling many times.

### 4.4.2 Main steps

This section highlights the main steps for applying Bayesian Linear Regression to the form error model, which include:

1. Generating matrices and other preliminary functions based on the Bayesian AR(2) model to store form error data model parameters;
2. Generating a results matrix and specifying priors for the model parameters;
3. Creating a model mapping the training inputs to the training outputs;
4. Creating a Gibbs sampling algorithm to draw samples from the posterior distribution.

The end result will be marginal posterior densities for the parameters, as well as projected form error point placements, which will be examined to see what is going on. Appendix A.2 contains many pieces of code that explain how Bayesian linear regression is used in RStudio to estimate parameters and execute the necessary forecasts.

#### *4.4.2.1 Phase 1: Generating matrices and other preliminary functions*

The first step is to load the form error data $\boldsymbol{d}$ (e.g., from a suitable CSV file containing the data) and specify the number of rows ("nrow"), columns ("ncol"), variables ("nvar") and lags $q$. From here, the next steps are to define a function, i.e. $F_1$, from three components: form error data, number of lags used, and the intercept parameter $\alpha$, then to generate a matrix $\boldsymbol{M}_1$ which depends on the size of the data and lags, and is equal to $(m-1)\times(q+1)$.

The final steps are to generate a $q \times q$ sized matrix $\boldsymbol{M}_2$, a function $F_2$ of the beta coefficients, i.e. $\boldsymbol{\beta} = [\beta_1, \beta_2]$, and a $q \times (q-1)$ identity matrix. Matrix $\boldsymbol{M}_2$ converts the model's coefficient vector into a mixed matrix of $\beta_i$, ones and zeros. By expressing the matrix in this way, it can later be used to determine the model's stability, which is an important part of the Gibbs algorithm.

Since I have approximately $m = 500$ data entries, $q = 2$ lags and a constant, the dimension of matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are $499 \times 3$ and $2 \times 2$, respectively. Both matrices are given by:

$$\boldsymbol{M}_1 = \begin{pmatrix} 1 & f_2 & f_1 \\ 1 & f_3 & f_2 \\ \vdots & \vdots & \vdots \\ 1 & f_m & f_{m-1} \end{pmatrix} \quad \boldsymbol{M}_2 = \begin{pmatrix} \beta_1 & \beta_2 \\ 1 & 0 \end{pmatrix} \tag{4.23}$$

#### *4.4.2.2 Phase 2: Setting up priors*

In this step, priors are set up to obtain approximate posterior distributions for the model parameters. To begin, denote the initial prior distributions for $\boldsymbol{\beta}$ and $\sigma_\epsilon^2$, and specify them according to (4.5) and (4.6), respectively. For simplicity, select $m_0$ and $\zeta_0$ to be equal to 1. Increasing the value of $\zeta_0$ gives us a broader distribution, with the $\boldsymbol{\beta}$ becoming more likely to take on larger values. Robustness checks have been performed to see where changing the original prior values can have a major impact on the output posterior distribution. Thanks to the many iterations of simulation, changing the values has little effect on the posterior distribution.

#### *4.4.2.3 Phase 3: Mapping input to output matrices*

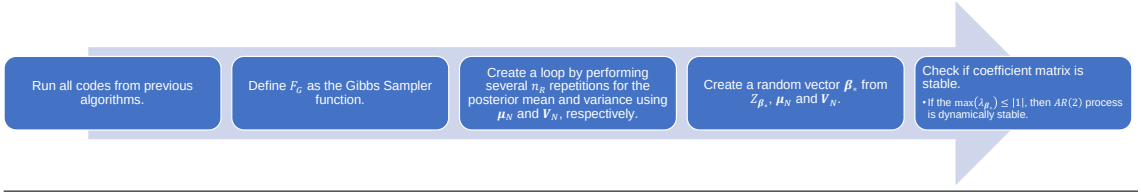
In this step, two new matrices are defined to keep track of the findings, namely, $\boldsymbol{M}_3$ and $\boldsymbol{M}_4$. The first output matrix, $\boldsymbol{M}_3$, holds all of the repetitions and has rows equal to the sampler's number of repetitions (which is set to 10,000). The dimensions of this matrix depends on the number of repetitions and the number of coefficients (including $\alpha$ plus $\sigma_\epsilon$), thus this is expected to be a large matrix of size $10000 \times 4$.

The second output matrix, $\boldsymbol{M}_4$, records the results of the predictions. To make a prediction, I'll need the last $w < m$ measurable points. As a result, the number of columns in $\boldsymbol{M}_3$ will be equal to the number of repetitions assigned times the number of projected calculated points times, i.e. a matrix of size $10000 \times 15$.

#### *4.4.2.4 Phase 4: Gibbs Sampling*

The first step of Gibbs Sampling is outlined in Algorithm 1, and the Gibbs sampler function is defined as $F_G \equiv F_G(\boldsymbol{M}_1, \boldsymbol{f}, \boldsymbol{\mu}_0, \boldsymbol{V}_0, \sigma_\epsilon^2, m_0, \zeta_0, \boldsymbol{M}_3, \boldsymbol{M}_4)$. Algorithm 1 requires all codes to be run, including the matrix $\boldsymbol{M}_1$, which is just a vector $\boldsymbol{f}$ preceded by two periods and a column of ones. I'll also need all of the previously established priors, as well as the number of times I'll iterate the algorithm, and the two output matrices, $\boldsymbol{M}_3$ and $\boldsymbol{M}_4$.

**Algorithm 1** Implementation of Gibbs Sampling (Phase I)




The loop in the third block (Algorithm 1) is the most important aspect to consider.

Since a conjugate Gaussian prior distribution was chosen, the posterior would be Gaussian as well, this is where all of the main computations take place. The normalization coefficient for this distribution can be found by completing the square of the exponential component and then using the standard result for a normalized Gaussian. In this case, the resulting posterior distribution for the given prior in (4.5) updates to:

$$\pi(\boldsymbol{\beta}|\boldsymbol{f}, \sigma_\epsilon^2) \sim \mathcal{N}(\boldsymbol{\mu}_N, \boldsymbol{V}_N), \tag{4.24}$$

where

$$\boldsymbol{\mu}_N = \boldsymbol{V}_N \left( \boldsymbol{V}_0^{-1}\boldsymbol{\mu}_0 + \frac{1}{\sigma_\epsilon^2}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{f} \right), \quad \boldsymbol{V}_N = \left( \boldsymbol{V}_0^{-1} + \frac{1}{\sigma_\epsilon^2}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{M}_1 \right)^{-1}. \tag{4.25}$$

Bishop's Chapter 3 [125] contains a similar proof of the formulation and derivation to the posterior mean and variance in (4.25), but with slightly different notation.

*Proof.* Let's assume I have a Gaussian marginal distribution $p(\boldsymbol{x})$ and a Gaussian conditional distribution $p(\boldsymbol{y}|\boldsymbol{x})$, with the mean of $p(\boldsymbol{y}|\boldsymbol{x})$ being a linear function of x and its covariance being independent of x. I'll use the following marginal and conditional distributions:

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1}) \tag{4.26}$$

$$p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{L}^{-1}) \tag{4.27}$$

where $\boldsymbol{\mu}$, $\boldsymbol{A}$ and $\boldsymbol{b}$ are parameters governing the means, and $\boldsymbol{\Lambda}$ and $\boldsymbol{L}$ are precision matrices. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors with dimensions $M$ and $D$, then the matrix $\boldsymbol{A}$ has size $D \times M$. I start by calculating the joint distribution over x and y. To do so, I define $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}}$ and then consider the joint distribution's log function:

$$\begin{aligned} \ln p(\boldsymbol{z}) &= \ln p(\boldsymbol{x}) + \ln p(\boldsymbol{y}|\boldsymbol{x}) \\ &= -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\boldsymbol{x} - \boldsymbol{\mu}) - \frac{1}{2}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) + \text{constant} \end{aligned} \tag{4.28}$$

Given that this is a quadratic function of the components of $\boldsymbol{z}$, $p(\boldsymbol{z})$ must be a Gaussian distribution. It's worth noting that the quadratic form of a Gaussian distribution is derived from the exponent part of a multivariate Gaussian PDF, which is useful for calculating the mean and variance. It is easy to solve such problems by recalling that the exponent in a general Gaussian distribution $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ can be written as:

$$-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = -\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \text{constant}. \tag{4.29}$$

By identifying the linear terms in (4.28), I can find the mean of the Gaussian distribution over z, which are given by:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{\mu} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{b} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{b} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \boldsymbol{\Lambda}\boldsymbol{\mu} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{b} \\ \boldsymbol{L}\boldsymbol{b} \end{pmatrix}. \tag{4.30}$$

Using the previous result in (4.29), obtained by completing the square over the quadratic form of a multivariate Gaussian, I find that the mean of z is:

$$\mathbb{E}[\boldsymbol{z}] = \boldsymbol{R}^{-1} \begin{pmatrix} \boldsymbol{\Lambda\mu} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Lb} \\ \boldsymbol{Lb} \end{pmatrix}, \tag{4.31}$$

where

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{\Lambda} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{LA} & -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{L} \\ -\boldsymbol{LA} & \boldsymbol{L} \end{pmatrix}. \tag{4.32}$$

Thus, the mean of $\boldsymbol{z}$ is

$$\mathbb{E}[\boldsymbol{z}] = \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{A\mu} + \boldsymbol{b} \end{pmatrix} \tag{4.33}$$

Similarly, I consider the second order terms in (4.28) to determine the precision of this Gaussian, which can be written as:

$$-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\Lambda} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{LA})\boldsymbol{x} - \frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Ly} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{LAx} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ly} \tag{4.34}$$

$$= -\frac{1}{2}\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}^{\mathrm{T}} \boldsymbol{R} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \tag{4.35}$$

Thus, taking the inverse of matrix $\boldsymbol{R}$ yields the variance matrix, i.e.

$$\mathbb{V}[\boldsymbol{z}] = \boldsymbol{R}^{-1} = \begin{pmatrix} \boldsymbol{\Lambda}^{-1} & \boldsymbol{\Lambda}^{-1}\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A\Lambda}^{-1} & \boldsymbol{L}^{-1} + \boldsymbol{A\Lambda}^{-1}\boldsymbol{A}^{\mathrm{T}} \end{pmatrix}. \tag{4.36}$$

Using $\mathbb{E}[\boldsymbol{z}]$ and $\mathbb{V}[\boldsymbol{z}]$, I can calculate the mean and variance of the marginal distribution $p(\boldsymbol{y})$ from:

$$\mathbb{E}[\boldsymbol{y}] = \boldsymbol{A\mu} + \boldsymbol{b} \tag{4.37}$$

$$\mathbb{V}[\boldsymbol{y}] = \boldsymbol{L}^{-1} + \boldsymbol{A\Lambda}^{-1}\boldsymbol{A}^{\mathrm{T}}. \tag{4.38}$$

Finally, by combining $\mathbb{E}[\boldsymbol{z}]$ and $\mathbb{V}[z]$ with a method from Bishop's Chapter 2, I find that the expressions required for the mean and variance of the conditional probability $p(\boldsymbol{x}|\boldsymbol{y})$ are given by:

$$\mathbb{E}[\boldsymbol{x}|\boldsymbol{y}] = \mathbb{V}[\boldsymbol{x}|\boldsymbol{y}]\{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{L}(\boldsymbol{y} - \boldsymbol{b}) + \boldsymbol{\Lambda\mu}\} \tag{4.39}$$

$$\mathbb{V}[\boldsymbol{x}|\boldsymbol{y}] = (\boldsymbol{\Lambda} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{LA})^{-1}. \tag{4.40}$$

□

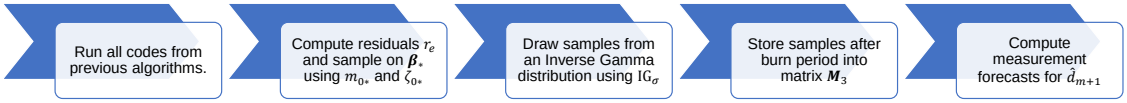
Since the posterior distribution is Gaussian, its mean, median and mode are the same. Following that, a random draw from the correct distribution must be performed. A standard

normal distribution can be used to sample a random vector $Z_\beta = Z(\boldsymbol{\beta}|0,1)$ and transform it from a Gaussian distribution in (4.24) using the formula below:

$$\boldsymbol{\beta}_* = \boldsymbol{\mu}_N + \left(Z_\beta \, \boldsymbol{V}_N^{1/2}\right)^{\mathrm{T}}. \tag{4.41}$$

The formula in (4.41) indicates to multiply the random vector from the standard normal distribution by the square root of the posterior variance matrix and then add it to the posterior mean matrix to get the random vector $\boldsymbol{\beta}_*$ from the conditional posterior distribution. Lastly, a stability check is required to check the robustness of (4.41). This leads to the second (and final) stage of Gibbs Sampling is outlined in Algorithm 2.

**Algorithm 2** Implementation of Gibbs Sampling (Phase II)




According to Algorithm 2, I can compute the residuals $\boldsymbol{f} - \boldsymbol{M}_1 \boldsymbol{\beta}_*$ which is useful to find an update to the prior distribution specified in (4.6). This component is essential in order to sample from the correct IG distribution conditional on $\boldsymbol{\beta}_*$ with $m_{0*}/2$ degrees of freedom and scaled parameter $\zeta_{0*}/2$. In this case, the resulting posterior distribution for the given prior in (4.6) updates to:

$$\pi(\sigma_\epsilon^2|\boldsymbol{f}) \sim \mathrm{IG}(m_{0*}, \zeta_{0*}), \tag{4.42}$$

where the updated (posterior) hyper-parameters for the correct distribution of $\sigma_\epsilon^2$ are given by:

$$m_{0*} = m_0 + m, \quad \zeta_{0*} = \zeta_0 + (\boldsymbol{f} - \boldsymbol{M}_1 \boldsymbol{\beta}_*)^{\mathrm{T}} (\boldsymbol{f} - \boldsymbol{M}_1 \boldsymbol{\beta}_*). \tag{4.43}$$

The algorithm can now sample $m_{0*}$ random variables from an alternative standard normal distribution $Z_\sigma$ and then apply the following modifications:

$$\mathrm{IG}_\sigma = \frac{\zeta_{0*}}{Z_\sigma \, Z_\sigma^{\mathrm{T}}}, \tag{4.44}$$

where $\mathrm{IG}_\sigma$ is a draw from the correct Inverse Gamma distribution. The first four steps of Algorithm 2 are given in the following code.

#### *4.4.2.5 Phase 5: Forecasts*

Algorithm 2's final phase is concerned with computing measurement forecasts. Here, I assign a $\hat{\boldsymbol{f}}$ matrix to store the predictions for the next $n$ calculated points after the $m$-th

point. In general, I can write the equation for a one-step prediction as

$$\hat{f}_{m+1} = \alpha_* + \beta_{1*}\hat{f}_m + \beta_{2*}\hat{f}_{m-1}, \tag{4.45}$$

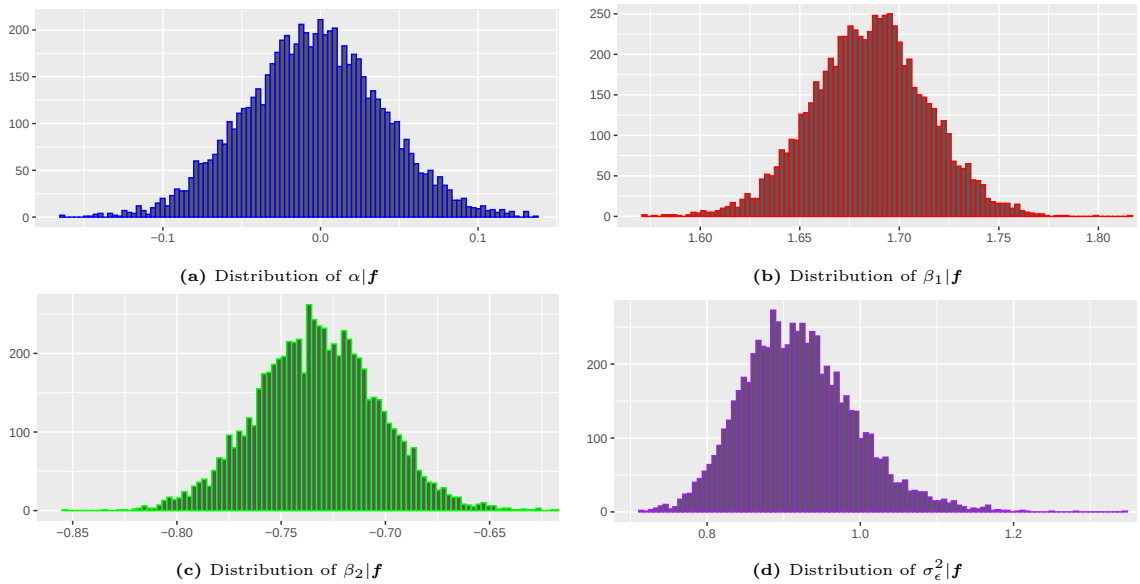
where $\{\alpha_*, \beta_{1*}, \beta_{2*}\} \in \boldsymbol{\beta}_*$. To summarise, I'll need a square matrix of size $(n+q)$ to cover the number of points I want to forecast in the AR model. This is explained in the next bit of code:

I can now define a new variable $R_2$ which represents the "updated" results list taken from the output of Gibbs sampling method. I can then disregard a certain number of samples (during the burn-in period). The rationale for this is that the stationary distribution of the Markov chain is the desired joint distribution over the variables, but it can take a long time to get there; subsequent samples are not independent of one another, but create a Markov chain with some correlation. After the burn phase, the findings are updated further into $R_2$, revealing the posterior distribution's coefficients and forecasts.

## 4.5 Results

### 4.5.1 Analysis of the AR(2) model

The algorithm's results can be examined. I used `qplot` to plot the posterior distributions of $\alpha$, $\beta_1$, $\beta_2$ and $\sigma_\epsilon^2$ in Figures 4.4a to 4.4d. Considering that I used a conjugate prior and likelihood function, it is understandable that their shape is very similar to a Gaussian distribution.




**(a)** Distribution of $\alpha|\boldsymbol{f}$

**(b)** Distribution of $\beta_1|\boldsymbol{f}$

**(c)** Distribution of $\beta_2|\boldsymbol{f}$

**(d)** Distribution of $\sigma_\epsilon^2|\boldsymbol{f}$

**Figure 4.4:** Posterior distributions of $\alpha|\boldsymbol{f}$, $\beta_1|\boldsymbol{f}$, $\beta_2|\boldsymbol{f}$ and $\sigma_\epsilon^2|\boldsymbol{f}$ for an AR(2) model.

Each of these variables has also been estimated to obtain its posterior statistics (mean, median, and quantiles), which is an estimate of each distribution's true posterior statistics. These estimates are summarised in Table 4.2. As a result of the Gaussian structure of the parameters, the posterior means and medians are nearly equal, with mean/median centred parameter $\alpha$ approximately zero. The quantiles are fairly close together, suggesting the model's robustness. All recorded values are given in μm.

**Table 4.2:** Posterior statistics for the parameters of the AR(2) form error model.

| | $\alpha \mid \boldsymbol{f}$ | $\beta_1 \mid \boldsymbol{f}$ | $\beta_2 \mid \boldsymbol{f}$ | $\sigma_\epsilon^2 \mid \boldsymbol{f}$ |
|---|---|---|---|---|
| Posterior Mean | −0.00443 | 1.68505 | −0.73234 | 0.91973 |
| Posterior Median | −0.00418 | 1.68555 | −0.73259 | 0.91469 |
| Posterior SD | 0.04321 | 0.02888 | 0.02904 | 0.07335 |
| Posterior 5% Quantile | −0.0759 | 1.64 | −0.780 | 0.809 |
| Posterior 95% Quantile | 0.0671 | 1.73 | −0.685 | 1.05 |

Figure 4.5 shows the plot with forecast for the next 250 measurement points, which appear to closely follow the pattern established by the CMM measurement process. The credibility intervals are relatively narrow, indicating that the model is quite robust, as it allows for form errors of less than 5 microns. The error bands are calculated using 25-75% credibility intervals in the R framework:

```
error_bands <- colQuantiles(forecasts,prob = c(0.25,0.75))
```

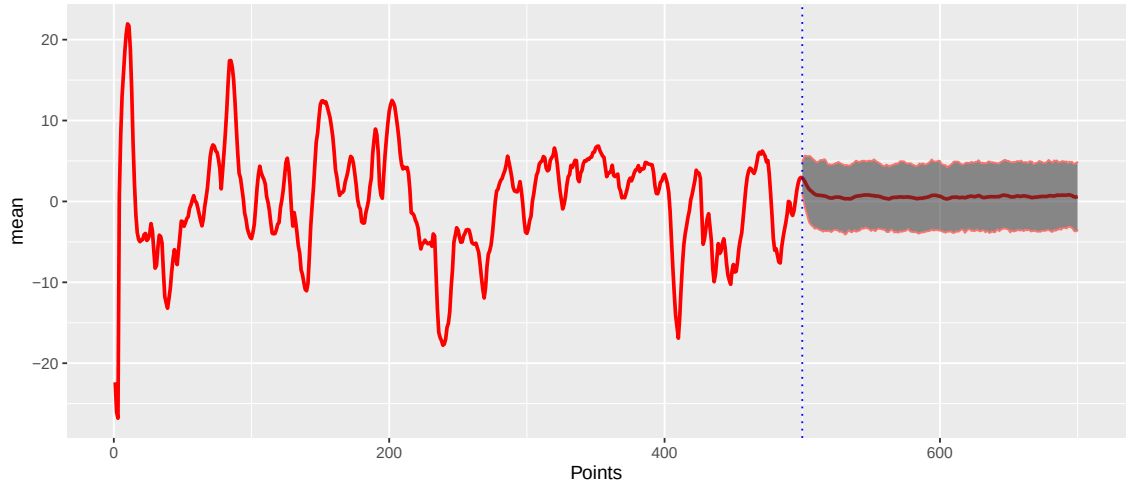



**Figure 4.5:** Bayesian AR(2) form error forecast model using the posterior mean form error (μm). Points measured: 1-500. Points predicted: 501-750. Credibility bands: 25-75%.

### 4.5.2 Failure analysis

This section will utilize the integrated FTA/FMEA approach first outlined in section 4.3 for the purposes of analyzing the relevant FMs/FC of the CMM measurement process. The following is a step-by-step explanation of the proposed analysis.

1. From the FTA, define a form error threshold and then calculate the four classification measures: true positive ($TP$), false positive ($FP$), false negative ($FN$), and true negative ($TN$).

2. Calculate all of the classification rates using (4.17)-(4.22) and compare them to empirical methods to see how effective they are.

3. Using the calculated classification measures, estimate all risk parameters for the FMEA and compute the resulting RPN. The criteria for occurrence, severity and detection ratings associated with the CMM measurement process under the Bayesian AR(2) model is given in Table 4.3.

**Table 4.3:** Occurrence, severity and detection ratings associated with the CMM measurement process under the Bayesian AR(2) model.

| **Occurrence rating** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Occurrence probability** | *Remote* | *Extremely unlikely* | *Very unlikely* | *Unlikely* | *Low* | *Moderately low* | *Moderate* | *Likely* | *Very likely* | *Frequent* |
| **Criteria: Prevalence rate associated with the Bayesian AR(2) model** | <1% | <2% | <3% | <5% | <7.5% | <10% | <15% | <20% | <25% | ≥ 25% |
| **Severity rating** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Severity rate** | *Remote* | *Very low* | *Low* | *Moderately low* | *Moderate* | *Moderately high* | *High* | *Very high* | *Extreme* | *Extremely high* |
| **Criteria: Error rate associated with the Bayesian AR(2) model** | <0.1% | <0.2% | <0.5% | <1% | <1.5% | <2% | <2.5% | <3% | <5% | ≥ 5% |
| **Detection rating** | **10** | **9** | **8** | **7** | **6** | **5** | **4** | **3** | **2** | **1** |
| **Detection probability** | *Remote* | *Very Low* | *Low* | *Unlikely* | *Occasional* | *Moderately likely* | *Likely* | *Very likely* | *Almost certain* | *Certain* |
| **Criteria: Prediction rate associated with the Bayesian AR(2) model** | ≤ 50% | >50% | >55% | >60% | >65% | >70% | >80% | >90% | >95% | >99% |

A confusion matrix is used to evaluate the fault prediction's accuracy. $TP$, $FP$, $FN$ and $TN$ are the four classification measures in the matrix. I set the tolerance limit to $T = 10$ µm. In Table 4.4, the four classification measures are provided alongside a sample of observed and predicted form error values obtained from the Gibbs sampler.

Of the initial 500 points measured, I found that $TP = 29$ points were correctly classified as exceeding the threshold and $FP = 5$ points were not. Similarly, $TN = 465$ points were correctly classified as below below the threshold and only $FN = 1$ point was not. Overall, the fault state had a high accuracy rate of 98.8% and a prediction rate of 85.3%, respectively. The prediction rate at capturing faults is significantly higher than empirical rates in other

**Table 4.4:** Sample of observed vs predicted dataset using the Bayesian AR(2) model.

| $i$ | $f_i$ | $\hat{f}_i$ | $\delta_i$ | $\hat{\delta}_i$ | $TP_i$ | $TN_i$ | $FP_i$ | $FN_i$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.740 | 0.740 | 0 | 0 | 0 | 1 | 0 | 0 |
| 2 | 8.170 | 8.170 | 0 | 0 | 0 | 1 | 0 | 0 |
| 3 | 13.000 | 14.140 | 1 | 1 | 1 | 0 | 0 | 0 |
| 4 | 15.730 | 16.838 | 1 | 1 | 1 | 0 | 0 | 0 |
| 5 | 17.823 | 17.901 | 1 | 1 | 1 | 0 | 0 | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 497 | 0.420 | 2.381 | 0 | 0 | 0 | 1 | 0 | 0 |
| 498 | - 0.820 | 0.063 | 0 | 0 | 0 | 1 | 0 | 0 |
| 499 | - 1.800 | - 0.774 | 0 | 0 | 0 | 1 | 0 | 0 |
| 500 | 0.600 | - 1.517 | 0 | 0 | 0 | 1 | 0 | 0 |

articles that combined FTA with the classical AR model (e.g. 53% in Chalermarrewong *et al.* [115]). The prevalence rate, which describes the frequency with which failures occur, was calculated to be 6.8%. Sensitivity and specificity are the percentage of positives and negatives, respectively, that are successfully identified, and in this example, 96.7% and 98.9% were correctly detected. All six classification measures are statistical indicators of a binary classification test's performance that are frequently utilized and hence highly valuable in ensuring the model's robustness.

The integrated FTA/FMEA results will now be examined. The high accuracy rate implies that the error (or misclassification) rate is only 1.2% which equates to a severity score of $S = 5$ on a 10 point scale, indicating a moderate risk. The prevalence rate corresponds to an occurrence value of $O = 5$, indicating that the likelihood of failure is somewhere between "unlikely" and "possibly" feasible. Furthermore, on a scale of one to ten, the prediction rate correlates to a detection value of $D = 4$, showing that prediction quality is very likely to detect failures. When failures do arise, the prediction model will know ahead of time, and the loss of computation and data will almost certainly be avoided, thanks to the high accuracy of detecting the failure cause. When all three risk factors are combined, the RPN produces a score of 100. For different FMs/FCs, RPN values will range from 1 to 1000. Assigning a threshold RPN value to categorise FMs/FCs is quite helpful for engineers. A FM/FC with RPN = 100, for example, can be classified as requiring "minor corrective action", whereas a FM/FC with RPN = 200 is classified as requiring "critical corrective action".

The failure analysis will now be extended to match the study provided in subsection 3.6.3,

in which 11 FCs and 4 FMs were identified with the support of peer experts and domain experts. Here, I convert the FMEA-generated FCs and FMs into the FTA tree-structure. The FTA tree structure of the provided method is shown in Figure 4.6. The OR-gate, which is frequently used in FTA and is explained in depth in [50], is the roundish element in Figure 4.6.

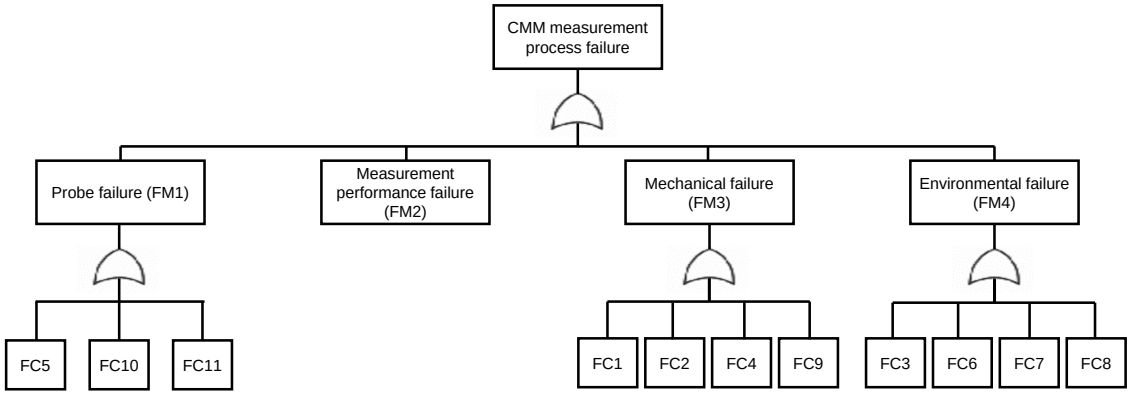



**Figure 4.6:** FTA of the CMM measurement process (top level).

The CMM measurement process failure is specified as a FTA's top event. The second layer is determined by their FMs, which are probe failure (FM1), measurement performance failure (FM2), mechanical failure (FM3) and environmental failure (FM4). The last is determined by their respective FCs, which are item displaced slightly or greatly (FC1/2), dust on time (FC3), mechanical stress (FC4), wrong probe type used (FC5), valve not set to required pressure level (FC6), temperature set too high or too low (FC7/8), CMM making noise (FC9), probe not fitted properly (FC10) or probe not calibrated (FC11). Finally, all FM-to-FC pathways analysed previously in Table 3.5 with the highest posterior rates are kept for the FTA, while the rest are removed (i.e. the posterior probabilities of observing FM1, FM2, and FM3 given FC10 were 19.0%, 8.4%, and 7.5%, respectively, so I assign a connecting branch between FM1-to-FC10 and remove FM2/3-to-FC10). A filled out integrated FTA/FMEA form for significant FMs/FCs associated with a CMM measurement process is presented Table 4.5.

## 4.6 Discussion and concluding remarks

This chapter was concerned with using an integrated FTA/FMEA approach with a Bayesian AR model using form error data for identifying, evaluating and prioritising risks. The importance of applying the form error model to analyze and anticipate CMM measurement failures was investigated in this research and confirmed by a literature study. I estimated the

**Table 4.5:** A filled out integrated FTA/FMEA form for significant FMs/FCs associated with a CMM measurement process.

| | | FTA | | | | Classification measures | | | | | FMEA | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Failure mode** | **Failure cause** | **TP** | **FP** | **TN** | **FN** | **Error (%)** | **Prevalence (%)** | **Prediction (%)** | **Sensitivity (%)** | **Specificity (%)** | $S$ | $O$ | $D$ | $C$ | **RPN** | **Priority** | **Corrective action** |
| FM1 | FC5 | 80 | 9 | 411 | 0 | 1.8% | 17.8% | 89.9% | 100.0% | 97.9% | 6 | 8 | 4 | 48 | 192 | 4 | High |
| FM1 | FC10 | 70 | 11 | 416 | 3 | 2.8% | 16.2% | 86.4% | 95.9% | 97.4% | 8 | 8 | 4 | 64 | 256 | 1 | Critical |
| FM1 | FC11 | 63 | 8 | 427 | 2 | 2.0% | 14.2% | 88.7% | 96.9% | 98.2% | 7 | 7 | 4 | 49 | 196 | 2 or 3 | High |
| FM3 | FC1 | 5 | 1 | 491 | 3 | 0.8% | 1.2% | 83.3% | 62.5% | 99.8% | 4 | 2 | 4 | 8 | 32 | 10 | Not needed |
| FM3 | FC2 | 14 | 2 | 484 | 0 | 0.4% | 3.2% | 87.5% | 100.0% | 99.6% | 3 | 4 | 4 | 12 | 48 | 9 | Not needed |
| FM3 | FC4 | 52 | 5 | 440 | 3 | 1.6% | 11.4% | 91.2% | 94.5% | 98.9% | 6 | 7 | 3 | 42 | 126 | 7 | Minor |
| FM3 | FC9 | 49 | 7 | 442 | 2 | 1.8% | 11.2% | 87.5% | 96.1% | 98.4% | 6 | 7 | 4 | 42 | 168 | 5 or 6 | Moderate |
| FM4 | FC3 | 19 | 1 | 477 | 3 | 0.8% | 4.0% | 95.0% | 86.4% | 99.8% | 4 | 4 | 3 | 16 | 48 | 8 | Not needed |
| FM4 | FC6 | 11 | 1 | 487 | 1 | 0.4% | 2.4% | 91.7% | 91.7% | 99.8% | 3 | 3 | 3 | 9 | 27 | 11 | Not needed |
| FM4 | FC7 | 59 | 7 | 433 | 1 | 1.6% | 13.2% | 89.4% | 98.3% | 98.4% | 6 | 7 | 4 | 42 | 168 | 5 or 6 | Moderate |
| FM4 | FC8 | 55 | 8 | 434 | 3 | 2.2% | 12.6% | 87.3% | 94.8% | 98.2% | 7 | 7 | 4 | 49 | 196 | 2 or 3 | High |

coefficients' posterior distribution by using Gibbs sampling which is an example of MCMC that allows one to estimate the joint marginal distribution using draws from the conditional distribution. I then built a form error forecast from the algorithm's posterior density. The results have shown that the integrated approach to AR modelling and Gibbs sampling is effective in studying the risks and predictions of CMM measurements, and thus improve reliability of future measurements for product conformity assessment. The Bayesian AR(2) model confirmed that it is capable of anticipating form errors, as the forecast was within the desired measurement (5 μm). The model was also found to have a high fit accuracy rate (98.8%), making it suitable for the study, as well as high prediction (85.3%), sensitivity (96.7%), and specificity (98.9%), indicating that the classification measures combined with Bayesian analysis are extremely useful in assessing the model's robustness and forecasting ability. The findings of the investigation also supported the proposed Bayesian FTA/FMEA approach.

In conclusion, this chapter presents a nice introduction to Bayesian regression in metrology by intuitively connecting the form error model to an AR process with perfectly reasonable results. There are several other types of models I could use to extend this model to get a more reliable forecast, such as MA, ARIMA, state-space or multi-model linear regression (among others) [126–128], which can adapt a variety of other variables (such as human, environmental and random effects). The proposed framework, I believe, might be used to increase the overall efficiency of a CMM measurement process with some fine-tuning.

# Chapter 5

# Discussion and conclusion

## 5.1 General discussion

This thesis develops an approach to failure and risk analysis that combines Bayesian inference with Failure Mode and Effects Analysis (FMEA) as a way to evaluate the effectiveness of the action measures taken to reduce the likelihood of making wrong decisions. I have shown that Bayesian statistics and an integrated failure analysis (using FMEA/FMECA (Failure mode and effects analysis/Failure mode, effects, and criticality analysis) or a combined FTA (Fault tree analysis) and FMEA approach) can be used to estimate risk measures (severity $S$, occurrence $O$ and detection $D$ ratings) in all three case studies effectively. These case studies demonstrated the effectiveness of combining Bayesian statistics with improved failure analysis methods, which had been severely lacking in the literature. For instance, posterior statistics were used to estimate failure severity, e.g., as seen in case studies 1-3, current detection measures for COVID-19 imposed by the government were used to estimate the detection rating (e.g, the first case study), a Bayesian network approach was used to estimate failure occurrences (e.g., the second case study), and classification measures from a confusion matrix were used to assess all three risk measures (e.g., the third case study). In addition, thanks to the estimation of unknown parameters in the models used (e.g., the modified SEIR or the augmented/autoregressive form error models), I could have used a variety of criteria to effectively assess the risk measures of the integrated Bayesian failure analysis. The following were the primary research questions that needed to be answered for this thesis:

1. What are the best ways to accurately determine FMs/FCs in complex manufacturing, measurement systems, or epidemics (among other things)?

2. Are there any ways to improve risk assessment methods by using Bayesian inference together with failure analysis methods?

3. What are the options for assessing the effectiveness of the corrective measures taken to minimize the risk and to improve product quality or process efficiency?

The following subsections summarize the answers to these questions from their respective case studies.

### 5.1.1 A discussion on the first case study

In the first case study (see chapter 2), I have combined Bayesian inference and FMEA to study the risks of COVID-19 infections and to evaluate the effectiveness of the action measures taken to manage the COVID-19 pandemic. The Bayesian model and FMEA are applied to the COVID-19 data showing the effectiveness of the interventions adopted to control the pandemic by reducing the $\mathcal{R}_t$ of COVID-19. Since the FMEA focused on COVID-19 infections, the FM was taken to represent the positive (infected) cases. The process phases, cause of failures, current measures and current risk ratings are discussed, and subsequent government action measures are presented with re-assessed risk ratings. The results have shown that the combination of Bayesian inference, compartmental modelling and FMEA are effective to model and study the risks of the COVID-19 transmissions, able to evaluate quantitatively the action measures and identify the lessons learned from the impacts of governmental measures and actions carried out in response to the community spread of COVID-19 in the United Kingdom. Using the Bayesian posterior distributions with actual data, it was demonstrated here that SEIQR was a more effective model than SEIR for forecasting the outbreak of COVID-19 in Western Europe.

Additionally, I found that SEIQR model was in strong agreement with SEIR, which can be thought of as a qualitative measure of the UK's COVID-19 outbreak. After that, a comparison of the posterior estimates of the effective reproduction number in the United Kingdom, under the new model, closely matches the deduced empirical values (as seen in Figure 2.7). A FMEA was then performed for the community spread of COVID-19 in the United Kingdom, and each risk measure was ranked from 1 (insignificant / remote / failure prevented) to 10 (critical / extremely high / no detection opportunity). Three risk measures were examined to assess the risk of COVID-19 spreading in the UK (from March to October 2020): the CFR rate ($S$), the posterior median of the effective reproduction number ($O$), and the current control measures employed by UK government policies ($D$),

as seen in Figure 2.9, Figure 2.7 and Table 2.3, respectively.

The initial and revised RPN values seemed to reflect the government's contribution to reducing infection rates accurately. During the fourth phase, for instance, it was found that the government's scientific guidance indicated that the reproduction numbers had risen from 0.5–0.9 at the end of the lockdown to 0.7–1.0 on 15 May, getting closer to the point at which infections will exponentially increase. For the SEIQR model, these figures correspond to the posterior median estimates of 0.66–0.96 and 0.88–0.94, respectively (indicating an excellent fit). The methods and results presented here can contribute to the research on Bayesian inference and risk assessment associated with the current COVID-19 pandemic, but they can also serve as a baseline for future pandemics as well as for wider failure analysis and quality engineering applications.

### 5.1.2 A discussion on the second case study

In the second case study (see chapter 3), I then considered a Bayesian risk assessment of a coordinate measuring machine (CMM) measurement process using an augmented form error model and an integrated FMECA approach. The failure analysis methods have been greatly extended. The criticality value, which is derived from the product of $O$ and $S$, is a very valuable indication for determining which failure causes should be prioritized. The main novelty here was to present an innovative way of combining conformity assessment with FMECA to the form error model. Ultimately, this was achieved. The three risk measures were matched together in the analysis of the FMECA for the CMM measurement process: $O$ to the posterior probability of FCs occurring using a Bayesian network approach, $S$ to the specific consumer's risk, and $D$ to the posterior standard deviation of the form error model. These three values were used to compute the criticality/RPN which prioritises risks/failures with the highest values. The most significant FMs came from probes and measurements, followed by mechanical and environmental factors. Among the most significant failure causes that can affect the measurement process are: "wrong probe type used", "probe not fitted properly" and "probe not calibrated / mechanical stress".

The RPN was taken into account as well because it combines the current criticality level with the detection measures in place. It is crucial to minimize the criticality and increase the likelihood of preventing the failure in order to considerably reduce or anticipate future measurement failures. The three FCs with the highest RPN were also "wrong probe type used", "probe not properly fitted" and "probe not calibrated". Despite this, "mechanical stress" has a lower RPN to criticality ratio than "probe not calibrated" due to differences

in the estimated $D$ parameter for both, so the highest RPNs are not necessarily those with the highest criticality (see Figure 3.13). RPN and criticality can't be compared on the same graph since they are measured according to scales of 1-1000 and 1-100, respectively. The solution was to divide the $j$-th criticality/RPN for the $j$-th FC by the overall criticality/RPN. As a result, the findings of the FMECA review and the data collected can be stored in a database that can be used to analyze similar measurement processes in the future. In this manner, future failures can be avoided.

### 5.1.3 A discussion on the third case study

In the third case study (see chapter 4), I considered an integrated FTA/FMEA approach with a Bayesian autoregressive (AR) model using form error data for identifying, evaluating and prioritising risks. The importance of applying the form error model to analyze and anticipate CMM measurement failures was investigated in this research. I estimated the parameters of the AR(2) model by using Gibbs sampling, an example of Markov chain Monte Carlo (MCMC) that allows one to estimate the joint marginal distribution by drawing from the conditional distribution. A form error forecast was then constructed from the posterior density of the algorithm. FMEA-generated FCs and FMs from the second case study were converted to FTA tree structures (see Figure 4.6). The CMM measurement process failure was specified as a FTA's top event. The second layer was determined by their FMs, which were probe failure (FM1), measurement performance failure (FM2), mechanical failure (FM3) and environmental failure (FM4). The last was determined by their respective FCs, which were item displaced slightly or greatly (FC1/2), dust on time (FC3), mechanical stress (FC4), wrong probe type used (FC5), valve not set to required pressure level (FC6), temperature set too high or too low (FC7/8), CMM making noise (FC9), probe not fitted properly (FC10) or probe not calibrated (FC11). Finally, all FM-to-FC pathways analysed previously in Table 3.5 with the highest posterior rates were kept for the FTA, while the rest were removed (i.e. the posterior probabilities of observing FM1, FM2, and FM3 given FC10 were 19.0%, 8.4%, and 7.5%, respectively, so I assign a connecting branch between FM1-to-FC10 and remove FM2/3-to-FC10).

The FMEA outputs were then examined. Risk measures were evaluated using the error rate ($S$), the prevalence rate ($O$) and the prediction rate ($D$). The results had shown that the integrated approach to AR modelling and Gibbs sampling is effective in studying the risks and predictions of CMM measurements, and thus improve reliability of future measurements for product conformity assessment. For a given measurement, the fairly high

accuracy rate (e.g., 1.2% for a given measurement) indicated a low/moderate risk. Similarly, the prevalence and prediction rates indicated that the likelihood of failure is somewhere between "unlikely" and "possibly" feasible, or that the prediction quality was very likely to detect failures. When failures do arise, the prediction model will know ahead of time, and the loss of computation and data will almost certainly be avoided, thanks to the high accuracy of detecting the failure cause. Assigning a threshold RPN value to categorise FMs/FCs was quite helpful (as determined by the experts). Similarly, FM/FCs that have a RPN of 100 may need "minor corrective action", whereas FM/FCs with a RPN of 200 may need "critical corrective action". In addition, the Bayesian AR(2) model had high accuracy and prediction rates when using the classification measures from a confusion matrix, which compared well with empirical results and validated the suggested Bayesian FTA/FMEA model. For this reason, anticipating FMs and FCs based on an integrated FTA/FMEA approach and an autoregressive form error model can be highly beneficial during the CMM measurement process.

## 5.2 Conclusive remarks and reiterating the contributions

The creation of new knowledge within the context of a specific field of study is an essential criterion for receiving a PhD. The contributions made in this thesis are reiterated in the following paragraphs.

For contributions specific to the first case study, Bayesian inference, FMEA and a modified SEIR model has been utilised to study the risks of COVID-19 infections and to evaluate the effectiveness of the actions taken to manage the COVID-19 pandemic by combining together. The risk measures, $S$ and $O$ parameters, were evaluated using the CFR rate and posterior median of the effective reproduction number, respectively. Parameter $D$ was estimated using the current detection measures in place. Thus, the evaluation of the effectiveness of English government protective measures against COVID-19 was conducted by comparing the empirical and posterior statistics of the effective reproduction number at different times with and without the measures. The methods/findings obtained were in line with empirical results, and, thus, can be used for future pandemic simulations in which individuals in different compartments interact.

For contributions specific to the second case study, I investigated the risks and identified significant observable FMs/FCs of the CMM measurement process in order to effectively

assess the corrective/preventive measures in place. Additional sources of uncertainty (e.g., random, environmental, human, etc.) were implemented as additive factors, resulting in more reliable product conformity rates that matched a Pareto-like distribution. In addition, the FMECA risk parameters, $S$ from the severity of the specific consumer's risk, $O$ from the posterior probability of failures associated with the CMM measurement process under a Bayesian network, and $D$ from the detectability of the posterior standard deviation of the proposed form error model, were estimated using the integrated Bayesian approach.

For contributions specific to the third case study, a Bayesian risk assessment that incorporates a MCMC algorithm (i.e. Gibbs sampling), an autoregressive form error model, and an integrated FTA/FMEA was shown to be particularly effective in predicting failure and making decisions. The risk measures $S$, $O$, and $D$, which were critical for determining the RPN and making decisions, were assessed using classification measures derived from a combination of the integrated FTA/FMEA and a confusion matrix.

Thus, it has been shown that the methods applied to each case study can be easily prolonged or applied to other cases studies in the research of Bayesian statistics and failure analysis. For example, the third case study supplemented and extended the ideas and methods described in the first and second case studies, respectively, by extending the time-series approach of updating daily data from the first case study; transforming the augmented form error model from the second case study into a time-series equivalent (i.e., an autoregressive form error model). As a result of the hard work put in to obtain this PhD, the methodology presented in the three case studies can be used to contribute to scientific research on Bayesian inference and risk assessment in epidemiology and metrology using FMEA, as well as other disciplines and applications in quality engineering.

## 5.3 Future works

In reviewing my case studies, I realised that I concentrated too much on the CMM case studies and not enough on the COVID-19 modelling (2 vs 1 case studies). The compartmental model employed in chapter 2 was based on the proposed modified SEIR model, however due to the research's timing, it eliminates the vaccination effect: vaccinations started in December 2020, and this study was completed by then. In future research, it would be interesting to modify this model to include a vaccinated ($V$) compartment and compare the results to the original version of the model. This approach can be further developed in a variety of model implementations, with varied post-vaccinated susceptibility based

on vaccine efficiency, for example. This project is fascinating enough to warrant a whole chapter.

The suggested compartmental model would take into account six stages of infection: susceptible ($S$), exposed ($E$), infectious ($I$), quarantined ($Q$), removed ($R$), and vaccinated ($V$). Initially, a mathematical analysis would be performed to demonstrate the suggested model's non-negativity, boundedness, epidemic equilibrium, existence, and uniqueness, as well as the effective reproduction number. Finally, I'd combine this modelling approach with an improved FMEA to model the risks of COVID-19 transmissions due to quarantining and the various effects of vaccinations, allowing me to quantitatively evaluate action measures and identify lessons learned from the impacts of governmental measures and actions taken in response to the community spread of COVID-19 in the United Kingdom.

In addition, I could also greatly enhance the reliability of the results from chapter 3 by considering the effects of combining multiple failure analysis methods with Bayesian inference (e.g., the Bow–tie method (BT), Design Review by Failure Mode (DRBFM), Hazard Analysis (HA), Fault Tree Analysis (FTA) and its extension Event Tree Analysis (ETA), What-if/Checklist, etc) to perform risk assessments in a variety of case studies, and thus compare the results to the original version of the models. It would also be interesting to compare the results to the original in chapter 4 using other time series methods, such as an adaptive Moving Average (MA) or Auto Regressive Integrated Moving Average (ARIMA), which can adapt a variety of other variables (e.g., additive effects in the case of the CMM measurement process). Overall, these ideas can easily justify adding a few more chapters to this thesis or possibly publishing more journal articles.

# Bibliography


[1] M. D. Russell and T. A. Jur, "Engineering analysis of failure: A determination of cause method," *Journal of Failure Analysis and Prevention*, vol. 17, no. 1, pp. 8–14, 2017. DOI: 10.1007/s11668-016-0224-9.

[2] M. Gul, M. Yucesan, and E. Celik, "A manufacturing failure mode and effect analysis based on fuzzy and probabilistic risk analysis," *Applied Soft Computing*, vol. 96, 2020. DOI: 10.1016/j.asoc.2020.106689.

[3] N. Boodhun and M. Jayabalan, "Risk prediction in life insurance industry using supervised learning algorithms," *Complex amp; Intelligent Systems*, vol. 4, no. 2, pp. 145–154, 2018. DOI: 10.1007/s40747-018-0072-1.

[4] M. A. Moktadir, S. M. Ali, S. Kusi-Sarpong, and M. A. Shaikh, "Assessing challenges for implementing industry 4.0: Implications for process safety and environmental protection," *Process Safety and Environmental Protection*, vol. 117, pp. 730–741, 2018. DOI: 10.1016/j.psep.2018.04.020.

[5] D. Lee and D. Choi, "Analysis of the reliability of a starter-generator using a dynamic Bayesian network," *Reliability Engineering amp; System Safety*, vol. 195, p. 106 628, 2020. DOI: 10.1016/j.ress.2019.106628.

[6] WHO, *Naming the coronavirus disease (COVID-19) and the virus that causes it*, Feb. 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.

[7] S. Boral, I. Howard, S. K. Chaturvedi, K. McKee, and V. Naikan, "An integrated approach for fuzzy failure modes and effects analysis using fuzzy ahp and fuzzy mairca," *Engineering Failure Analysis*, vol. 108, p. 104 195, 2020. DOI: 10.1016/j.engfailanal.2019.104195.

[8] USDOD, *Procedures for performing a failure mode effects and criticality analysis*, 1949. [Online]. Available: https://elsmar.com/pdf_files/FMEA%5C%20Examples/milstd1629.pdf.

[9] E. Bozdag, U. Asan, A. Soyer, and S. Serdarasan, "Risk prioritization in failure mode and effects analysis using interval type-2 fuzzy sets," *Expert Systems with Applications*, vol. 42, no. 8, pp. 4000–4015, 2015. DOI: 10.1016/j.eswa.2015.01.015.

[10] P. Bhattacharjee, V. Dey, and U. Mandal, "Risk assessment by failure mode and effects analysis (FMEA) using an interval number based logistic regression model," *Safety Science*, vol. 132, no. 8, 2020. DOI: 10.1016/j.ssci.2020.104967.

[11] A. E. P. Villa, *Artificial neural networks and machine learning– ICANN 2012 22nd International Conference on Artificial Neural Networks, Lausanne, Switzerland, September 11-14, 2012, Proceedings. Part I*. Springer, 2012.

[12] J. A. Palacios and V. N. Minin, "Gaussian process-based Bayesian nonparametric inference of population size trajectories from gene genealogies," *Biometrics*, vol. 69, no. 1, pp. 8–18, 2013. DOI: 10.1111/biom.12003.

[13] A. Carbonari, M. Vaccarini, and A. Giretti, "Bayesian networks for supporting model based predictive control of smart buildings," *Dynamic Programming and Bayesian Inference, Concepts and Applications*, 2014. DOI: 10.5772/58470.

[14] A. Dale, "Thomas Bayes, an essay towards solving a problem in the doctrine of chances (1764)," *Landmark Writings in Western Mathematics 1640-1940*, pp. 199–207, 2005. DOI: 10.1016/b978-044450871-3/50096-6.

[15] T. Baldacchino, S. R. Anderson, and V. Kadirkamanathan, "Computational system identification for Bayesian narmax modelling," *Automatica*, vol. 49, no. 9, pp. 2641–2651, 2013. DOI: 10.1016/j.automatica.2013.05.023.

[16] M. Papananias, T. E. McLeay, M. Mahfouf, and V. Kadirkamanathan, "A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing," *Computers in Industry*, vol. 105, pp. 35–47, 2019. DOI: 10.1016/j.compind.2018.10.008.

[17] A. B. Forbes, "Least squares best fit geometric elements," *Algorithms for Approximation II*, pp. 311–319, 1990. DOI: 10.1007/978-1-4899-3442-0_28.

[18] A. B. Forbes and H. D. Minh, "Form assessment in coordinate metrology," *Approximation Algorithms for Complex Systems*, vol. 3, pp. 69–90, Dec. 2010. DOI: 10.1007/978-3-642-16876-5_4.

[19] A. B. Forbes, "Uncertainty associated with form assessment in coordinate metrology," *International Journal of Metrology and Quality Engineering*, vol. 4, no. 1, pp. 17–22, 2013. DOI: 10.1051/ijmqe/2012032.

[20] ——, "Measurement systems analysis and statistical process control," *Advanced Mathematical and Computational Tools in Metrology VII*, 2006. DOI: 10.1142/9789812774187_0015.

[21] ISO, *5725-1:1994 - Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions*, Jul. 2018. [Online]. Available: https://www.iso.org/standard/11833.html.

[22] ——, *5725-2:2019 - Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*, Dec. 2019. [Online]. Available: https://www.iso.org/standard/69419.html.

[23] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. Wiley, 2019.

[24] Y. Sverdlik, *The world's 10 fastest supercomputers - in pictures*, Jul. 2018. [Online]. Available: https://www.datacenterknowledge.com/supercomputers/world-s-10-fastest-supercomputers-pictures.

[25] B. P. Carlin and T. A. Louis, *Bayesian methods for data analysis*. Chapman Hall/CRC, 2009.

[26] G. Consonni, D. Fouskakis, B. Liseo, and I. Ntzoufras, "Prior distributions for objective Bayesian analysis," *Bayesian Analysis*, vol. 13, no. 2, pp. 627–679, 2018. DOI: 10.1214/18-ba1103.

[27] A. Chernoded, L. Dudko, I. Myagkov, and P. Volkov, "Deep learning neural networks and Bayesian neural networks in data analysis," *EPJ Web of Conferences*, vol. 158, p. 06 008, 2017. DOI: 10.1051/epjconf/201715806008.

[28] D. S. Sivia, *Data analysis: a Bayesian tutorial*. Clarendon Press, 2004.

[29] L. Wasserman, "Why isnt everyone a Bayesian?" *Springer Series in Statistics The Science of Bradley Efron*, pp. 260–267, 1986. DOI: 10.1007/978-0-387-75692-9_13.

[30] A. Gelman, A. Jakulin, M. Pittau, and Y. Su, "A weakly informative default prior distribution for logistic and other regression models," *The Annals of Applied Statistics*, vol. 2, no. 4, pp. 1360–1383, 2008. DOI: 10.1214/08-aoas191.

[31] J. Griffin and P. Brown, "Inference with normal-gamma prior distributions in regression problems," *Bayesian Analysis*, vol. 5, no. 1, pp. 171–188, 2010. DOI: 10.1214/10-ba507.

[32] J. Wang, Y. Dong, K. Zhang, and Z. Guo, "A numerical model based on prior distribution fuzzy inference and neural networks," *Renewable Energy*, vol. 112, pp. 486–497, 2017. DOI: 10.1016/j.renene.2017.05.053.

[33] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984. DOI: 10.1109/tpami.1984.4767596.

[34] J. K. Kruschke, "Markov Chain Monte Carlo," *Doing Bayesian Data Analysis*, pp. 143–191, 2015. DOI: 10.1016/b978-0-12-405888-0.00007-6.

[35] A. G. Wilson, "Hierarchical Markov Chain Monte Carlo (MCMC) for Bayesian system reliability," *Wiley StatsRef: Statistics Reference Online*, 2014. DOI: 10.1002/9781118445112.stat04234.

[36] M. Mcallister, "Formulating quantitative methods to evaluate fishery-management systems: What fishery processes should be modelled and what trade-offs should be made?" *ICES Journal of Marine Science*, vol. 56, no. 6, pp. 900–916, 1999. DOI: 10.1006/jmsc.1999.0547.

[37] T. Schweder, "Integrative fish stock assessment by frequentist methods: Confidence distributions and likelihoods for bowhead whales," *Scientia Marina*, vol. 67, no. S1, pp. 89–97, 2003. DOI: 10.3989/scimar.2003.67s189.

[38] J.-P. Fox, "Bayesian hierarchical response modeling," *Bayesian Item Response Modeling*, pp. 31–44, 2010. DOI: 10.1007/978-1-4419-0742-4_2.

[39] G. Goh and D. Dey, "Bayesian map estimation using Gaussian and diffused-gamma prior," *Canadian Journal of Statistics*, vol. 46, no. 3, pp. 399–415, Jun. 2018. DOI: 10.1002/cjs.11458.

[40] T. Ferguson, "Functions of the sample moments," *A Course in Large Sample Theory*, pp. 44–50, 1996. DOI: 10.1007/978-1-4899-4549-5_7.

[41] B. S. Blanchard and W. J. Fabrycky, *Systems engineering and analysis*. Pearson/Prentice Hall, 2011.

[42] M. Catelani, L. Ciani, and M. Venzi, "Failure modes, mechanisms and effect analysis on temperature redundant sensor stage," *Reliability Engineering amp; System Safety*, vol. 180, pp. 425–433, 2018. DOI: 10.1016/j.ress.2018.08.013.

[43] S. Moslem, M. Gul, D. Farooq, E. Celik, O. Ghorbanzadeh, and T. Blaschke, "An integrated approach of Best-Worst Method (BWM) and Triangular Fuzzy Sets for evaluating driver behavior factors related to road safety," *Mathematics*, vol. 8, no. 3, p. 414, 2020. DOI: 10.3390/math8030414.

[44] V. Başhan, H. Demirel, and M. Gul, "An FMEA-based topsis approach under single valued neutrosophic sets for maritime risk evaluation: The case of ship navigation safety," *Soft Computing*, vol. 24, no. 24, pp. 18 749–18 764, 2020. DOI: 10.1007/s00500-020-05108-y.

[45] H. Zhang, Y. Dong, J. Xiao, F. Chiclana, and E. Herrera-Viedma, "Personalized individual semantics-based approach for linguistic failure modes and effects analysis with incomplete preference information," *IISE Transactions*, vol. 52, no. 11, pp. 1275–1296, 2020. DOI: 10.1080/24725854.2020.1731774.

[46] M. Yucesan, M. Gul, and E. Celik, "A holistic FMEA approach by fuzzy-based Bayesian network and best-worst method," *Complex amp; Intelligent Systems*, vol. 7, no. 3, pp. 1547–1564, 2021. DOI: 10.1007/s40747-021-00279-z.

[47] S. J. Rhee and K. Ishii, "Life cost-based FMEA using empirical data," *Volume 3a: 8th Design for Manufacturing Conference*, 2003. DOI: 10.1115/detc2003/dfm-48150.

[48] A. von Ahsen, "Cost-oriented failure mode and effects analysis," *International Journal of Quality amp; Reliability Management*, vol. 25, no. 5, pp. 466–476, 2008. DOI: 10.1108/02656710810873871.

[49] G. Cristea and D. Constantinescu, "A comparative critical study between FMEA and FTA risk analysis methods," *IOP Conference Series: Materials Science and Engineering*, vol. 252, p. 012 046, 2017. DOI: 10.1088/1757-899x/252/1/012046.

[50] M. Stamatelatos, W. Vesely, J. Dugan, J. Fragola, and J. Railsback, *Nasa: Fault tree handbook with aerospace applications*, 2002. [Online]. Available: http://www.mwftr.com/CS2/Fault%5C%20Tree%5C%20Handbook_NASA.pdf.

[51] M. Hamada, *Bayesian reliability*. Springer, 2011.

[52] B. Bertsche, A. Schauz, and K. Pickard, *Reliability in automotive and mechanical engineering: determination of component and system reliability*. Springer, 2011.

[53] S. Rastayesh, S. Bahrebar, F. Blaabjerg, D. Zhou, H. Wang, and J. Dalsgaard Sørensen, "A system engineering approach using FMEA and Bayesian network for risk analysis-a case study," *Sustainability*, vol. 12, no. 1, p. 77, 2019. DOI: 10.3390/su12010077.

[54] C. Wan, X. Yan, D. Zhang, Z. Qu, and Z. Yang, "An advanced fuzzy bayesian-based FMEA approach for assessing maritime supply chain risks," *Transportation Research Part E: Logistics and Transportation Review*, vol. 125, pp. 222–240, 2019. DOI: 10.1016/j.tre.2019.03.011.

[55] D. S. Hui *et al.*, "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - the latest 2019 novel coronavirus outbreak in Wuhan, china," *International Journal of Infectious Diseases*, vol. 91, pp. 264–266, 2020. DOI: 10.1016/j.ijid.2020.01.009.

[56] R. M. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford University Press, 2010.

[57] I. Nesteruk, "Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China," 2020. DOI: 10.1101/2020.02.12.20021931.

[58] O. Diekmann, J. Heesterbeek, and J. Metz, "On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations," *Journal of Mathematical Biology*, vol. 28, no. 4, 1990. DOI: 10.1007/bf00178324.

[59] C. Castillo-Chavez, Z. Feng, and W. Huang, "On the computation of $R_0$ and its role on global stability," *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction The IMA Volumes in Mathematics and its Applications*, pp. 229–250, 2002. DOI: 10.1007/978-1-4757-3667-0_13.

[60] M. Jit and P. White, "Economic analysis of interventions against infectious diseases," *Infectious Disease Epidemiology*, pp. 243–256, 2016. DOI: 10.1093/med/9780198719830.003.0017.

[61] P. J. White, "Mathematical models in infectious disease epidemiology," *Infectious Diseases*, 2017. DOI: 10.1016/b978-0-7020-6285-8.00005-8.

[62] O. Diekmann and J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Wiley, 2000.

[63] P. V. D. Driessche and J. Watmough, "Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, no. 1-2, pp. 29–48, 2002. DOI: 10.1016/s0025-5564(02)00108-6.

[64] L. M. A. Bettencourt and R. M. Ribeiro, "Real time Bayesian estimation of the epidemic potential of emerging infectious diseases," *PLoS ONE*, vol. 3, no. 5, 2008. DOI: 10.1371/journal.pone.0002185.

[65] K. Systrom, *The metric we need to manage COVID-19*, Apr. 2020. [Online]. Available: http://systrom.com/blog/the-metric-we-need-to-manage-covid-19/.

[66] Worldometer, *Coronavirus incubation period:* 2020. [Online]. Available: https://www.worldometers.info/coronavirus/coronavirus-incubation-period/.

[67] C. S. Study, *How long does COVID-19 last?* Jun. 2020. [Online]. Available: https://covid.joinzoe.com/post/covid-long-term.

[68] M. Roser, H. Ritchie, E. Ortiz-Ospina, and J. Hasell, *Coronavirus pandemic (covid-19) - statistics and research*, Mar. 2020. [Online]. Available: https://ourworldindata.org/coronavirus.

[69] I. C. London and I. Mori, *Real-time assessment of community transmission findings*, Sep. 2020. [Online]. Available: https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/real-time-assessment-of-community-transmission-findings/.

[70] H. Sjödin *et al.*, "COVID-19 healthcare demand and mortality in Sweden in response to non-pharmaceutical mitigation and suppression scenarios," *International Journal of Epidemiology*, vol. 49, no. 5, pp. 1443–1453, 2020. DOI: 10.1093/ije/dyaa121.

[71] V. Chalmers, *NHS used half as many ICU beds during COVID-19 crisis in the spring as France*, Oct. 2020. [Online]. Available: https://www.dailymail.co.uk/news/article-8879633/NHS-used-HALF-ICU-beds-Covid-19-crisis-spring-France.html.

[72] NHS, "The NHS and healthcare in Europe," *European Heart Journal*, vol. 39, no. 44, pp. 3923–3924, 2018. DOI: 10.1093/eurheartj/ehy691.

[73] G. Anthony *et al.*, "Reference software for finding Chebyshev best-fit geometric elements," *Precision Engineering*, vol. 19, no. 1, pp. 28–36, 1996. DOI: 10.1016/0141-6359(96)00005-0.

[74] P. T. Boggs, R. H. Byrd, and R. B. Schnabel, "A stable and efficient algorithm for nonlinear orthogonal distance regression," *SIAM Journal on Scientific and Statistical Computing*, vol. 8, no. 6, pp. 1052–1078, 1987. DOI: 10.1137/0908085.

[75] P. T. Boggs, R. H. Byrd, J. R. Donaldson, and R. B. Schnabel, "User's reference guide for ODRPACK," 1989. DOI: 10.6028/nist.ir.89-4103.

[76] K. Carr and P. Ferreira, "Verification of form tolerances part I: Basic issues, flatness, and straightness," *Precision Engineering*, vol. 17, no. 2, pp. 131–143, 1995. DOI: 10.1016/0141-6359(94)00017-t.

[77] ——, "Verification of form tolerances part ii: Cylindricity and straightness of a median line," *Precision Engineering*, vol. 17, no. 2, pp. 144–156, 1995. DOI: 10.1016/0141-6359(94)00018-u.

[78] N. Venkaiah and M. Shunmugam, "Evaluation of form data using computational geometric techniques - Part I: Circularity error," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 7-8, pp. 1229–1236, 2007. DOI: 10.1016/j.ijmachtools.2006.08.010.

[79] ——, "Evaluation of form data using computational geometric techniques - Part II: Cylindricity error," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 7-8, pp. 1237–1245, 2007. DOI: 10.1016/j.ijmachtools.2006.08.011.

[80] G. A. Watson, *Approximation theory and numerical methods*. New York, 1997.

[81] A. Gosavi and E. Cudney, "Form errors in precision metrology," Ph.D. dissertation, A Survey of Measurement Techniques, 2011, pp. 1–26.

[82] N. N. Abdelmalek, "On the discrete linear $L_1$ approximation and $L_1$ solutions of overdetermined linear equations," *Journal of Approximation Theory*, vol. 11, no. 1, pp. 38–53, 1974. DOI: 10.1016/0021-9045(74)90037-9.

[83] ——, "Chebyshev solution of overdetermined systems of linear equations," *BIT*, vol. 15, no. 2, pp. 117–129, 1975. DOI: 10.1007/bf01932684.

[84] I. Barrodale and F. D. Roberts, "An efficient algorithm for discrete $l_1$ linear approximation with linear constraints," *SIAM Journal on Numerical Analysis*, vol. 15, no. 3, pp. 603–611, 1978. DOI: 10.1137/0715040.

[85] ——, "Solution of the constrained $L_1$ linear approximation problem," *ACM Transactions on Mathematical Software*, vol. 6, no. 2, pp. 231–235, 1980. DOI: 10.1145/355887.355896.

[86] G. A. Watson, "A method for the Chebyshev solution of an overdetermined system of complex linear equations," *IMA Journal of Numerical Analysis*, vol. 8, no. 4, pp. 461–471, 1988. DOI: 10.1093/imanum/8.4.461.

[87] K. Madsen, "An algorithm for minimax solution of overdetermined systems of non-linear equations," *IMA Journal of Applied Mathematics*, vol. 16, no. 3, pp. 321–328, 1975. DOI: 10.1093/imamat/16.3.321.

[88] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the nonlinear case," *The Computer Journal*, vol. 12, no. 1, pp. 63–68, 1969. DOI: 10.1093/comjnl/12.1.63.

[89] C. Dong, C. C. Zhang, B. Wang, and G. Zhang, "Prediction and compensation of dynamic errors for coordinate measuring machines," *Journal of Manufacturing Science and Engineering*, vol. 124, no. 3, pp. 509–514, Aug. 2002. DOI: 10.1115/1.1465435. [Online]. Available: https://doi.org/10.1115/1.1465435.

[90] X. Wu, G. Yu, G. X. Li, and D. B. Shan, "Error analysis and precision evaluation of reconstructed surfaces in reverse engineering," *Key Engineering Materials*, vol. 460-461, pp. 581–586, 2011. DOI: 10.4028/www.scientific.net/kem.460-461.581.

[91] JCGM, "100:2008: Evaluation of measurement data: Guide to the expression of uncertainty in measurement," *Guide to the Expression of Uncertainty in Measurement*, Oct. 2008. [Online]. Available: https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.

[92] ——, "106:2012: Evaluation of measurement data: The role of measurement uncertainty in conformity assessment," *Guide to the Expression of Uncertainty in Measurement*, Oct. 2012. [Online]. Available: https://www.bipm.org/utils/common/documents/jcgm/JCGM_106_2012_E.pdf.

[93] J. O. Berger, *Statistical decision theory and Bayesian analysis*. Springer, 2011.

[94] L. R. Pendrill and H. Källgren, "Decision-making with uncertainty in attribute sampling," *Advanced Mathematical and Computational Tools in Metrology VII*, pp. 212–220, 2006. DOI: 10.1142/9789812774187_0021.

[95] A. B. Forbes, "Measurement uncertainty and optimized conformance assessment," *Measurement*, vol. 39, no. 9, pp. 808–814, Nov. 2006. DOI: 10.1016/j.measurement.2006.04.007.

[96] ISO, *11462-1:2001: Guidelines for implementation of statistical process control (spc) - part 1: Elements of spc*, May 2019. [Online]. Available: https://www.iso.org/standard/33381.html.

[97] CMMXYZ, 2021. [Online]. Available: https://blog.cmmxyz.com/blog/5-factors-that-can-affect-your-cmms-handling.

[98] ISO, *14224:2016: Petroleum, petrochemical and natural gas industries: Collection and exchange of reliability and maintenance data for equipment*. 2016. [Online]. Available: https://www.iso.org/standard/64076.html#:~:text=ISO%5C%2014224%5C%3A2016%5C%20provides%5C%20a,operational%5C%20life%5C%20cycle%5C%20of%5C%20equipment..

[99] J. T. Selvik and L. J. Bellamy, "Addressing human error when collecting failure cause information in the oil and gas industry: A review of iso 14224:2016," *Reliability Engineering amp; System Safety*, vol. 194, p. 106 418, 2020. DOI: 10.1016/j.ress.2019.03.025.

[100] D. Flack, *Good practice guide no. 130: Co-ordinate measuring machine task-specific measurement uncertainties*, Sep. 2013. [Online]. Available: https://www.npl.co.uk/special-pages/guides/gpg130_machine.

[101] ISO, "14253-1:1998 - Geometrical product specifications (GPS) - Inspection by measurement of workpieces and measuring equipment - Part 1: Decision rules for verifying conformity or nonconformity with specifications," 2017.

[102] J. Park, C. Park, and S. Ahn, "Assessment of structural risks using the fuzzy weighted euclidean FMEA and block diagram analysis," *The International Journal of Advanced Manufacturing Technology*, vol. 99, no. 9-12, pp. 2071–2080, 2018. DOI: 10.1007/s00170-018-1844-x.

[103] M. Mohammadi and J. Rezaei, "Bayesian best-worst method: A probabilistic group decision making model," *Omega*, vol. 96, p. 102 075, 2020. DOI: 10.1016/j.omega.2019.06.001.

[104] Renishaw, *Troubleshooting*, 2021. [Online]. Available: https://www.renishaw.com/cmmsupport/knowledgebase/en/troubleshooting--27710.

[105] M. Bulgaru, L. Fulea, and V. Bocanet, "Development of a FMEA analysis on the coordinate measuring process," *XI-th International Scientific Conference Coordinate Measuring Technique*, Apr. 2014.

[106] L. G. Fulea, M. Bulgaru, M. Borzan, and V. Bocanet, "Contributions regarding the application of FMEA analysis to measuring with a coordinate measuring machine," *Acta Technica Napocensis*, 2015th ser., vol. 58, no. 1, Mar. 2015.

[107] G. Corriveau, R. Guilbault, A. Tahan, and R. Sabourin, "Bayesian network as an adaptive parameter setting approach for genetic algorithms," *Complex amp; Intelligent Systems*, vol. 2, no. 1, pp. 1–22, 2016. DOI: 10.1007/s40747-016-0010-z.

[108] W. Premchaiswadi, *Bayesian networks*. InTech, 2012.

[109] H.-F. Lam and J. Yang, "Bayesian structural damage detection of steel towers using measured modal parameters," *Earthquakes and Structures*, vol. 8, no. 4, pp. 935–956, 2015. DOI: 10.12989/eas.2015.8.4.935.

[110] H. Sun and O. Büyüköztürk, "Probabilistic updating of building models using incomplete modal data," *Mechanical Systems and Signal Processing*, vol. 75, pp. 27–40, 2016. DOI: 10.1016/j.ymssp.2015.12.024.

[111] S. He and C.-T. Ng, "Guided wave-based identification of multiple cracks in beams using a Bayesian approach," *Mechanical Systems and Signal Processing*, vol. 84, pp. 324–345, 2017. DOI: 10.1016/j.ymssp.2016.07.013.

[112] Z. Hassan, A. M. Ibrahim, and H. Naji, "Evaluation of legislation adequacy in managing time and quality performance in Iraqi construction projects - A Bayesian decision tree approach," *Civil Engineering Journal*, vol. 4, no. 5, p. 993, 2018. DOI: 10.28991/cej-0309151.

[113] A. Abu-Samah, M. Shahzad, E. Zamai, and A. Said, "Failure prediction methodology for improved proactive maintenance using Bayesian approach," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 844–851, 2015. DOI: 10.1016/j.ifacol.2015.09.632.

[114] Q. Fan and H. Fan, "Reliability analysis and failure prediction of construction equipment with time series models," *Journal of Advanced Management Science*, pp. 203–210, 2015. DOI: 10.12720/joams.3.3.203-210.

[115] T. Chalermarrewong, T. Achalakul, and S. C. See, "Failure prediction of data centers using time series and fault tree analysis," *2012 IEEE 18th International Conference on Parallel and Distributed Systems*, 2012. DOI: 10.1109/icpads.2012.129.

[116] S. Ho, M. Xie, and T. Goh, "A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction," *Computers amp; Industrial Engineering*, vol. 42, no. 2-4, pp. 371–375, 2002. DOI: 10.1016/s0360-8352(02)00036-0.

[117] Y. Zhou, "Failure trend analysis using time series model," *2017 29th Chinese Control And Decision Conference (CCDC)*, 2017. DOI: 10.1109/ccdc.2017.7978640.

[118] K. Yu and R. A. Moyeed, "Bayesian quantile regression," *Statistics amp; Probability Letters*, vol. 54, no. 4, pp. 437–447, 2001. DOI: 10.1016/s0167-7152(01)00124-9.

[119] S. Shah, *Bayesian inference for AR(p) models*, Aug. 2019. [Online]. Available: https://towardsdatascience.com/bayesian-inference-for-ar-models-73ff916101c9.

[120] K. P. Murphy, *Conjugate Bayesian analysis of the Gaussian distribution*, Oct. 2007. [Online]. Available: https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf.

[121] M. Masum, *Identifying AR and MA terms using ACF and PACF plots in time series forecasting*, Aug. 2020. [Online]. Available: https://towardsdatascience.com/identifying-ar-and-ma-terms-using-acf-and-pacf-plots-in-time-series-forecasting-ccb9fd073db8.

[122] T. Kohda and K. Inoue, "Fault-tree analysis considering latency of basic events," *Annual Reliability and Maintainability Symposium. 2001 Proceedings. International Symposium on Product Quality and Integrity (Cat. No.01CH37179)*, 2001. DOI: 10.1109/rams.2001.902437.

[123] C. Liu, P. Frazier, and L. Kumar, "Comparative assessment of the measures of thematic classification accuracy," *Remote Sensing of Environment*, vol. 107, no. 4, pp. 606–616, 2007. DOI: 10.1016/j.rse.2006.10.010.

[124] D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Mining*, vol. 14, no. 1, 2021. DOI: 10.1186/s13040-021-00244-z.

[125] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer New York, 2016.

[126] D. v. Bergh *et al.*, "A tutorial on Bayesian multi-model linear regression with Bas and jasp," *Behavior Research Methods*, vol. 53, no. 6, pp. 2351–2371, 2021. DOI: 10.3758/s13428-021-01552-2.

[127] X. Wang, Y. Yue, and J. J. Faraway, "Bayesian linear regression," *Bayesian Regression Modeling with Inla*, pp. 39–69, 2018. DOI: 10.1201/9781351165761-3.

[128] L. Roque, *The first step in Bayesian time series - linear regression*, Sep. 2020. [Online]. Available: https://towardsdatascience.com/the-first-step-in-bayesian-time-series-linear-regression-89a64b826a7e.

# Appendix A

# Coding in R

## A.1 Plotting the ACF and PACF

The following code is used to load the form error data using a CSV file and then produce a plot of the ACF and PACF for Figure 4.2.

```
# Load CSV form error data
da=read.csv('CMM_data.csv',header=T)
# Check the 1st row of the data
da[1,]
# Plot the ACF and PACF of form error
acf(da[,2]); pacf(da[,2])
```

## A.2 Implementation of a Bayesian approach to AR($q$)

The several pieces of code below demonstrates how we used Bayesian linear regression in R to estimate parameters, determine marginal posterior densities, and forecast form error point positions.

### A.2.1 Step 1: Defining functions and generating matrices

```
# Define F1 and generate M1
F1 <- function(data,q,alpha){
  nrow <- as.numeric(dim(data)[1])
  nvar <- as.numeric(dim(data)[2])
```

```
  dmatrix <- as.matrix(data, ncol = nvar)
  M1 <- embed(dmatrix, q+1)
  M1 <- M1[,(nvar+1):ncol(M1)]
  if(alpha == TRUE){
    M1 <-cbind(rep(1,(nrow-q)),M1)
  }
  d = matrix(dmatrix[(q+1):nrow(dmatrix),])
  nvar2 = ncol(M1)
  return = list(d=d,M1=M1,nvar2=nvar2,nrow=nrow)
}

# Define F2 and generate M2
F2 <- function(beta){
  k = nrow(beta) - 1
  M2 <- matrix(0, nrow = k, ncol = k)
  #insert identity matrix
  M2[2:k, 1:(k-1)] <- diag(1, nrow = k-1, ncol = k-1)
  betaT <- t(beta[2:(k+1), 1:1])
  #Insert coeffcients along top row
  M2[1:1,1:k] <- betaT
  return(M2)
}
```

### A.2.2 Step 2: Generating results and setting up priors

```
# Generate results matrices
R1 = list()
R1 <- F1(d, q, TRUE)
M1 <- R1$M1
d <- R1$d
nrow <- R1$nrow
nvar <- R1$nvar

# Initialise Priors for Beta and Sigma distributions
```

```
mu0 <- c(rep(0, nvar))
mu0 <- as.matrix(mu0, nrow = 1, ncol = nvar)
V0 <- diag(1,nvar)
m0 = 1 # prior number of points / degrees of freedom
tau0 = 1 # prior scale parameter
zeta0 = m0*tau0^2
sigma2 = 1  # initial value for variance
```

### A.2.3 Step 3: Output matrices

```
# Set up some parameters and new matrices
RP = 10000 # number of repetitions
forecastpoints = 15 # number of predicted points
M3 = matrix(0, nrow = RP, ncol = nvar + 1)
colnames(M3) <- c('alpha', 'beta1', 'beta2', 'sigma')
M4 <- matrix(0, nrow = RP, ncol = forecastpoints)
```

### A.2.4 Steps 4 and 5: Gibbs sampling and forecasting

```
gibbs_sampler <- function(M1,d,mu0,V0,sigma2,m0,zeta0,RP,M3,M4
    ){
  for(i in 1:RP){
    if (i %% 1000 == 0){
      print(sprintf("Interaction: %d", i))
    }
    # Let NM be the posterior mean:
    NM = solve(solve(V0) + as.numeric(1/sigma2) * t(M1) %*% M1
        ) %*%
      (solve(V0) %*% mu0 + as.numeric(1/sigma2) * t(M1) %*% d)
    # Let NV be the posterior variance:
    NV = solve(solve(V0) + as.numeric(1/sigma2) * t(M1) %*% M1
        )

    # Check for stability
    stability_check = -1
```

```
14      while(stability_check < 0){
15        Bstar <- NM + t(rnorm(q+1) %*% chol(NV))
16
17        # Check: not stationary for 3 lags
18        b = F2(Bstar)
19        ee <- max(sapply(eigen(b)$values,abs))
20        if( ee<=1){
21          stability_check=1
22        }
23      }
24      # Compute residuals
25      resids <- d- M1%*%Bstar
26      m1 = nrow(d)
27      m2 = m0 + m1
28      zeta1 = zeta0 + t(resids) %*% resids
29
30      # Keeps samples after burn period
31      M3[i,] <- t(matrix(c(t(Bstar),sigma2)))
32
33      # Draw samples from the Inverse Gamma
34      # Let z~N(0,1) such that z=rnorm(m1,1), then:
35      z0 = t(rnorm(m2,1)) %*% rnorm(m2,1)
36      sigma2 = zeta1/z0
37
38      # keeps samples after burn period
39      M3[i,] <- t(matrix(c(t(Bstar),sigma2)))
40
41      # compute 2 year forecasts
42      fhat = rep(0,forecastpoints)
43      end = as.numeric(length(d))
44      fhat[1:2] = f[(end-1):end,]
45      cfactor = sqrt(sigma2)
46      M1_mat = c(1,rep(0,q))
47      for(m in (q+1):forecastpoints){
```

```
48        for (lag in 1:q){
49          # Update M1 matrix with q lags
50          M1_mat[(lag+1)] = fhat[m-lag]
51        }
52        # Use M1 matrix to forecast fhat
53        fhat[m] = M1_mat %*% Bstar + rnorm(1) * cfactor
54      }
55      M4[i,] <- fhat
56    }
57    return = list(M3,M4)
58  }
59
60  # After the burn phase, the findings are updated further into
       R2, revealing the posterior distribution's coefficients and
       forecasts. The next lines of code explains this.
61  R2 <- gibbs_sampler(M1,d,mu0,V0,sigma2,m0,zeta0,RP,M3,M4)
62  burn = 3000
63  coef <- R2[[1]][(burn+1):RP,]
64  forecasts <- R2[[2]][(burn+1):RP,]
```