Robust analysis of uncertainty in scientific assessments

Raices Cruz, Ivette

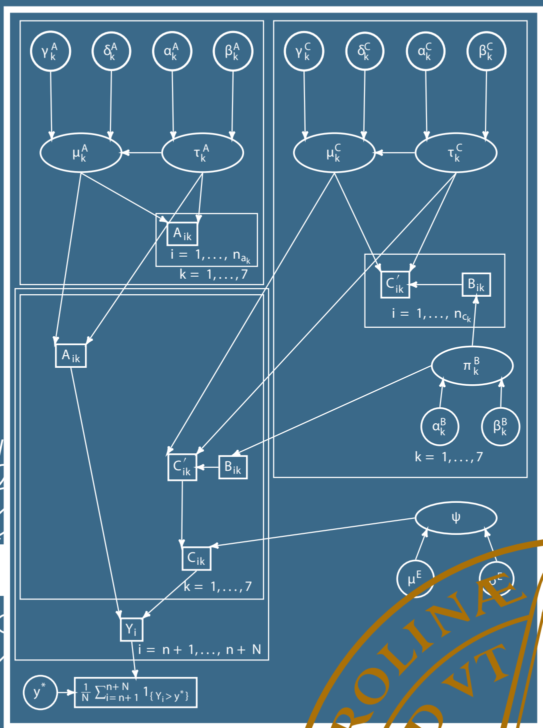
2021

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Robust analysis of uncertainty in scientific assessments

IVETTE RAICES CRUZ
CENTRE FOR ENVIRONMENTAL AND CLIMATE SCIENCE | LUND UNIVERSITY



Robust analysis of uncertainty in scientific assessments

I. Iterative importance sampling with Markov chain Monte Carlo sampling in robust Bayesian analysis.

II. A suggestion for the quantification of precise and bounded probability to quantify epistemic uncertainty in scientific assessments.

III. A robust Bayesian bias-adjusted random effects model in evidence synthesis.

IV. Treating uncertainty and variability in the effect of diversified farming on biodiversity and crop yields using Bayesian network meta-analysis and portfolio analysis.

Printed by Media-Tryck, Lund 2021 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Faculty of Science
Centre for Environmental and Climate Science

ISBN 978-91-8039-102-3

9 789180 391023

Robust analysis of uncertainty in scientific assessments

Robust analysis of uncertainty in scientific assessments

Ivette Raices Cruz

DOCTORAL DISSERTATION
by due permission of the Faculty of Science, Lund University, Sweden.
To be defended in the Blue Hall, Ecology Building, Sölvegatan 37, Lund University
On Friday 17th December 2021 at 13:00,
for the degree of Doctor of Philosophy in Environmental Science

Faculty opponent
Professor Mark E. Borsuk

Duke University
United States of America

Organization
LUND UNIVERSITY

Centre for Environmental and Climate Science

Document name
DOCTORAL DISSERTATION

Date of disputation
2021-12-17

Sponsoring organization

Author(s)
Ivette Raices Cruz

Title
Robust analysis of uncertainty in scientific assessments.

Abstract
Uncertainty refers to any limitation in knowledge. Identifying and characterizing uncertainty in conclusions is important to ensure transparency and avoid over or under confidence in scientific assessments. Quantitative expressions of uncertainty are less ambiguous compared to uncertainty expressed qualitatively, or not at all. Subjective probability is an example of a quantitative expression of epistemic uncertainty, which combined with Bayesian inference makes it possible to integrate evidence and characterizes uncertainty in quantities of interest. This thesis contributes to the understanding and implementation of robust Bayesian analysis as a way to integrate expert judgment and data into assessments and quantify uncertainty by bounded probability. The robust Bayesian framework is based on sets of probability for epistemic uncertainty, where precise probability is seen as a special case. This thesis covers applications relevant for scientific assessments, including evidence synthesis and quantitative risk assessment.

Paper I proposes to combine two sampling methods: iterative importance sampling and Markov chain Monte Carlo (MCMC) sampling, for quantifying uncertainty by bounded probability when Bayesian updating requires MCMC sampling. This opens up for robust Bayesian analysis to be applied to complex statistical models. To achieve this, an effective sample size of importance sampling that accounts for correlated MCMC samples is proposed. For illustration, the proposed method is applied to estimate the overall effect with bounded probability in a published meta-analysis within the Collaboration for Environmental Evidence on the effect of biomanipulation on freshwater lakes.

Paper II demonstrates robust Bayesian analysis as a way to quantify uncertainty in a quantity of interest by bounded probability, and explicitly distinguishes between epistemic and aleatory uncertainty in the assessment and learn parameters by integrating evidence into the model. Robust Bayesian analysis is described as a generalization of Bayesian analysis, including Bayesian analysis through precise probability as a special case. Both analyses are applied to an intake assessment.

Paper III describes a way to consider uncertainty arising from ignorance or ambiguity about bias terms in a quantitative bias analysis by characterizing bias with imprecision. This is done by specifying bias with a set of bias terms and use robust Bayesian analysis to estimate the overall effect in the meta-analysis. The approach provides a structured framework to transform qualitative judgments concerning risk of biases into quantitative expressions of uncertainty in quantitative bias analysis.

Paper IV compares the effect of different diversified farming practices on biodiversity and crop yields. This is done by applying a Bayesian network meta-analysis to a new public global database from a systematic protocol on diversified farming. A portfolio analysis calibrated by the network meta-analyses showed that uncertainty about the mean performance is large compared to the variability in performance across different farms.

Key words
expert knowledge, robust Bayesian analysis, scientific assessments, subjective probability, uncertainty analysis

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

ISBN
978-91-8039-102-3 (print)
978-91-8039-101-6 (pdf)

Recipient's notes

Number of pages
171

Price

Security classification

Signature

Date 2021-11-09

Robust analysis of uncertainty in scientific assessments

Ivette Raices Cruz

LUND UNIVERSITY

Cover illustration by Anette Raices Cruz

Faculty of Science
Centre for Environmental and Climate Science

ISBN: 978-91-8039-102-3 (print)
ISBN: 978-91-8039-101-6 (pdf)

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

A mis padres, mi hermana y Daniel

Contents

List of papers

This thesis is based on the following papers, referred to in the text by their Roman numbers:

I **Iterative importance sampling with Markov chain Monte Carlo sampling in robust Bayesian analysis**
Ivette Raices Cruz, Johan Lindström, Matthias Troffaes and Ullrika Sahlin
Submitted to Computational Statistics and Data Analysis

II **A suggestion for the quantification of precise and bounded probability to quantify epistemic uncertainty in scientific assessments**
Ivette Raices Cruz, Matthias Troffaes and Ullrika Sahlin
Journal of Risk Analysis (Under Review)

III **A robust Bayesian bias-adjusted random effects model in evidence synthesis**
Ivette Raices Cruz, Johan Lindström, Matthias Troffaes and Ullrika Sahlin
Submitted to Statistics in Medicine

IV **Treating uncertainty and variability in the effect of diversified farming on biodiversity and crop yields using Bayesian network meta-analysis and portfolio analysis**
Ivette Raices Cruz, Henrik G. Smith, Martin Stjernman, Sarah Jones and Ullrika Sahlin
Manuscript

Author contributions

Co-authors are abbreviated as follows:
Ivette Raices Cruz (IRC), Ullrika Sahlin (US), Matthias Troffaes (MT), Johan Lindström (JL), Henrik G. Smith (HS), Martin Stjernman (MS) and Sarah Jones (SJ).

Paper I: Iterative importance sampling with Markov chain Monte Carlo sampling in robust Bayesian analysis

Conceptualization and methodology: IRC, US, MT and JL. Writing original draft: IRC. Coding: IRC, MT and US. Formal analysis and visualization: IRC. Writing - review and editing: IRC, US, MT and JL.

Paper II: A suggestion for the quantification of precise and bounded probability to quantify epistemic uncertainty in scientific assessments

Conceptualization: US. Methodology: IRC, US and MT. Writing - original draft: IRC. Coding: IRC, US and MT. Formal analysis and visualization: IRC. Writing - review and editing: IRC, US and MT.

Paper III: A robust Bayesian bias-adjusted random effects model in evidence synthesis

Conceptualization: US and IRC. Methodology: IRC, US, MT and JL. Writing - original draft: IRC. Coding: IRC. Formal analysis and visualization: IRC. Writing - review and editing: IRC, US, MT and JL.

Paper IV: Treating uncertainty and variability in the effect of diversified farming on biodiversity and crop yields using Bayesian network meta-analysis and portfolio analysis

Conceptualization: IRC and HS. Data: SJ. Methodology: IRC, US and MS. Writing - original draft: IRC. Coding: IRC and US. Formal analysis and visualization: IRC. Writing - review and editing: IRC, US, HS, MS and SJ.

Popular summary

Modification of the environment to satisfy the needs of society is causing significant harm to the ecosystems and human well-being. Several assessments have been conducted to evaluate changes in a ecological community in relation to climate change and exposure to chemicals; the effect of sustainable agroecological practices on conservation of species, the effect of management interventions to reduce contamination levels on the air, land and water; and establishing safety thresholds of consumption and thereby support decision making.

There is uncertainty in these types of assessments arising from high variability in a population, scarce data and/or measurement errors, however decisions still need to be made. To take this uncertainty into account, it is necessary to assess, characterize and communicate uncertainty in the results of an assessment, to inform decisions properly and avoid over or under confidence in the conclusion. How and what to communicate depends on how uncertainty is expressed (quantitatively, qualitatively or not at all). For example, in the Intergovernmental Panel on Climate Change (IPCC) 2018 report, uncertainty is described using verbal and quantitative expressions together (e.g. very likely indicates 90-100% probability). Uncertainty can also be expressed by subjective probability (a person's degree of belief). For instance, a person might feel that there is a 20% probability that it will rain tomorrow or that there is a 20 - 30% probability that it will rain tomorrow. Characterizing uncertainty is important to describe a range of possible things that may affect the outcome and make better and informed decisions. Organizations such as European Food Safety Authority (EFSA), United States Environmental Protection Agency (US EPA) and the World Health Organization (WHO) have developed guidance and recommendations about characterizing uncertainty in scientific assessment and communicating uncertainty to decision makers and the public.

Developments in computer science and statistics, have led to an increase of useful tools and models to assess and quantify uncertainty in these types of assessments, yielding to more accurate estimates and conclusions, and facilitating integration of uncertainty quantification and communication to decision makers. This thesis focuses on probabilistic approaches (Bayesian analysis and robust Bayesian analysis) that allow to combine prior knowledge with available data, make predictions and use expert knowledge.

This thesis contributes to the advance of methods for quantifying uncertainty using probabilities and shows through case studies how they can be applied in scientific assessments including evidence synthesis and quantitative risk assessments, for instance, to assess the effect of biomanipulation to improve water quality in eutrophic lakes (Paper I), the efficacy of a drug in the treatment of a disease (Paper III), the effect of multiple diversified farming practices on biodiversity and crop yields (Paper IV) and the probability of exceeding the aluminium tolerability safety threshold via consumption of chocolate products (Paper II).

Introduction

Environmental risk assessment is the process for identifying and evaluating the nature, magnitude, and likelihood of possible environmental risks to inform decisions (Institute of Medicine, 2013). For instance, assessing changes in an ecological community (population) in relation to a contamination level, climate change, introduction and spread of an invasive species and exposure to chemicals (Institute of Medicine, 2013). Environmental risk assessments contribute to identify populations at risk and thus, provide recommendations to decision makers such as prioritization of resources and locations and regulation of specific chemicals.

Unfortunately, knowledge is often limited to inform decisions on how to manage environmental systems, but since actions are required and processes are not reversible, decision makers have to make decisions under uncertainty (Institute of Medicine, 2013; National Research Council, 2009). To this end, knowledge and methods to identify, characterize and communicate uncertainty are important to enhance transparency and increase trust in risk assessments (Fischhoff and Davis, 2014; Sahlin et al., 2021).

In this thesis, I discuss ways to express and characterize uncertainty in scientific assessments and provide examples of probabilistic approaches (Bayesian analysis and robust Bayesian analysis) relevant for scientific assessment, including evidence synthesis and quantitative risk assessment.

This chapter provides a general overview to evidence synthesis, quantitative risk assessments, scientific assessments and uncertainty analysis as a basis for the analyses and discussions in this thesis.

Evidence synthesis

Evidence based decision making is a formalized process to compile scientific evidence to support decisions. Synthesizing evidence often involves the following frameworks: systematic reviews and weight of evidence (Higgins et al., 2019; Suter et al., 2020). These frameworks are useful for assembling and making inference from several studies.

According to the Cochrane handbook, a systematic review attempts to gather all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question (Higgins et al., 2019). Systematic reviews use explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review. Systematic reviews involve comprehensive search strategies for identifying all relevant evidence that meet the eligibility criteria (DeLuca et al., 2008). Systematic reviews aim to provide reliable synthesis of evidence by using methods that increase repeatability, reproducibility and reduce bias to support decision making (Higgins et al., 2019; Mikolajewicz and Komarova, 2019).

Organizations such as the Cochrane Collaboration and the Collaboration for Environmental Evidence (CEE) produce relevant and up to date systematic reviews to support health and environmental decision making respectively (CEE, 2013; Higgins et al., 2019). A common statistical method for analyzing and summarizing the results of systematic reviews is meta-analysis. It estimates an intervention effect by combining results from different studies that have a similar research question (same outcome measures) (Higgins et al., 2019).

On the other hand, weight of evidence gathers information from different types of evidence (e.g. expert knowledge, lab and field data, chemical and biological measurements) which requires the application of methods able to combine all these types of information (e.g. expert elicitation, qualitative and quantitative methods) for making inference (Suter et al., 2020). Although, weight of evidence is commonly seen in the context of risk assessments, features of systematic reviews are often used to integrate and make inference from evidence in risk assessments (Suter et al., 2020).

Evidence is seldom conclusive and therefore, it is necessary to deal with uncertainty. Failing to acknowledge uncertainty could result in bad decisions (Fischhoff and Davis, 2014).

Quantitative risk assessments

Quantitative Risk Assessment (QRA) (also referred to as Probabilistic Risk Assessment (PRA)) is an approach for evaluating events, hazards or impacts in the future (Aven, 2011). It is often associated to safety-related issues. QRA comprises risk analysis, risk assessment and risk management.

QRA involves the identification of hazards or threats, analysis of their causes and consequences, and description of risk (Aven, 2011; Aven et al., 2018). QRA typically addresses questions such as: i) what could go wrong?, ii) how likely is it to happen? and iii) what are the consequences if it does happen?

A QRA is usually done by specifying an assessment model, integrating information from experts and data as well as considering sources of uncertainty in the analysis. QRA often provides a quantitative estimate of risk using probabilities and expected values (e.g. probability of exceeding a tolerable or acceptable threshold). QRA cannot eliminate risk, but it can provide support to risk management and decision making processes (Aven, 2011).

Scientific assessments

Scientific assessment is the process of using scientific evidence and reasoning to answer a question or estimate a quantity of interest (EFSA et al., 2018). This term is used by EFSA as a generalization of risk assessment with a specific decision making focus. A scientific assessment aims to evaluate scientific or technical knowledge by gathering and compiling evidence from multiple sources: data, models, assumptions and expert judgments to characterize uncertainty in available information. Among the many desirable characteristics of scientific assessments we find: quality, consistency, impartiality, transparency and openness, and fitness-for-purpose (Deluyker, 2017).

Scientific assessments include, but are not limited to: weight of evidence analyses; meta-analysis; health, safety, or ecological risk assessments; integrated assessment models and exposure assessments (EFSA et al., 2018; Higgins et al., 2019; Institute of Medicine, 2013). These assessments can be seen as a scientific approach to produce knowledge to answer specific problems sometimes under non-ideal conditions (i.e without high quality data coming from randomized controlled trials or controlled experiments). Scientific assessments use principles of evidence based decision making, quantitative risk analysis and uncertainty analysis. Scientific assessments typically need to adapt their methods to quantify uncertainty in the available information (EFSA et al., 2018). Incorporating uncertainty in scientific assessments is an area for research in several fields (Cox Jr., 2012; Maier et al., 2008), which could be transferred to evidence synthesis.

Uncertainty

Uncertainty refers to all types of limitations in available knowledge (Burgman, 2005; Lindley, 2006; O'Hagan et al., 2006). Uncertainty is associated to assessors' limited knowledge at the time of an assessment (Burgman, 2005; Lindley, 2006; O'Hagan et al., 2006; O'Hagan, 2019). A distinction is made between uncertainty coming from inherent randomness, natural variability and stochastic uncertainty (i.e. aleatory uncertainty) and uncertainty related to knowledge based, for example due to some level of ignorance, or incomplete knowledge of the system under study (i.e. epistemic uncertainty). Aleatory uncertainty cannot be reduced with further knowledge while epistemic uncertainty may be reduced with improved understanding. In the literature, aleatory and epistemic uncertainty are often referred to as variability and uncertainty respectively (Institute of Medicine, 2013; Helton et al., 2004). For example, sources of epistemic uncertainty are measurement and systematic errors, models (i.e. parameter uncertainty); and sources of aleatory uncertainty are natural variation in bodyweight and height in a population, weather variability, as well as flipping a coin and predicting either heads or tails.

Uncertainty can be expressed by probability, however, there could also be uncertainty regarding the probability values themselves (for example: when a hierarchical model has been specified). In this case two levels of uncertainty can be distinguished: first- and second-order uncertainty. Second-order uncertainty is often expressed by probability distributions over first-order probability distributions (Hansson, 2008).

Moreover, van der Bles et al. (2019) propose to categorize epistemic uncertainty into two levels: direct and indirect uncertainty. Indirect uncertainty concerns the quality of the evidence and strength of knowledge used in an assessment whereas direct uncertainty refers to how is uncertainty in the assessment outcome expressed (e.g. probability distribution, probability interval). This distinction is useful for understanding the implications of indirect uncertainty about knowledge bases on the conclusion of the assessment as well as for deciding how to treat direct uncertainty (van der Bles et al., 2019). Indirect uncertainty is often presented as a list of caveats about the different sources of evidence in assessments.

Characterizing and communicating uncertainty is important to ensure that decision makers do not place too much or too low confidence in the conclusion of an assessment (Fischhoff and Davis, 2014). To accomplish this, an uncertainty analysis (i.e. the process of identifying and characterizing uncertainty about a quantity of interest) should be conducted (EFSA et al., 2018). Organizations responsible for scientific assessments have developed guidance and recommendations about characterizing epistemic uncertainty in scientific assessment and communicating uncertainty to decision makers and the public (EFSA et al., 2018, 2019; Institute of Medicine, 2013; FAO and WHO, 2021).

EFSA's guidance on uncertainty analysis presented the main steps of an uncertainty analysis as: 1) identifying uncertainty affecting the assessment, 2) prioritizing uncertainty within the assessment, 3) dividing the uncertainty analysis into parts, 4) ensuring assessment questions and/or quantities of interest are well-defined, 5) characterizing uncertainty in each part, 6) combining uncertainty from different parts, 7) characterizing overall uncertainty (including both quantitative and qualitative uncertainty sources, i.e. uncertainties that can and cannot be quantified respectively) and 8) reporting and communicating the uncertainty analysis (EFSA et al., 2018).

Uncertainty analysis should embrace methods able to deal with different sources and types of uncertainty. There is not one method that fits all problems, therefore what method to use depends on what sources of uncertainty are considered, how uncertainty is described as well as the context of the decision (e.g. severity of the problem at hand and the time frame within which a decision needs to be taken). Methods for uncertainty analysis face a potential trade-off between resource requirement (i.e. time constraint, simplicity) and scientific rigor (i.e. methods that transparently characterize uncertainty in all steps of an uncertainty analysis) (EFSA et al., 2018). For those reasons, uncertainty analysis needs to be planned on a case-by-case basis. Uncertainty analysis is being developed to embrace quality of knowledge, subjective aspects of risk, decision context and stakeholders (Sahlin and Troffaes, 2017).

Characterization and quantification of uncertainty

Uncertainty can be described through qualitative expressions of uncertainty (i.e. verbal terms or ordinal categories). A word or phrase can have different meanings to different people (Morgan, 2014). Thus, qualitative expressions of uncertainty by themselves are ambiguous and therefore, for clarity they should be accompanied by a quantitative definition or scale (EFSA et al., 2018). For this reason, in scientific assessments, it is recommended to express epistemic uncertainty about a quantity of interest in a quantitative way using subjective probability (either precise or bounded) (EFSA et al., 2018). A subjective probability represents someone's degree of belief that a statement is true now (or will be true at a specified time in the future) given his/her current knowledge (Hampton et al., 1973; Lindley, 2006; O'Hagan and West, 2013; Singpurwalla and Wilson, 2008). It is recommended to complement a quantitative measure of uncertainty (e.g. a probability) with judgments of the strength of knowledge (Aven et al., 2018; Aven, 2020).

Subjective probability is a key concept to the approach presented in this thesis. Precise probability indicates a single probability value or distribution whereas bounded or imprecise probability refers to a pair of lower and upper probability values or a set of probability distributions (Walley, 1991).

Probabilistic methods allow integration and propagation of multiple sources of uncertainty through a model and summarize the impact of uncertainty on decisions (Apostolakis, 1990; Lindley, 2006; O'Hagan, 2012). Once uncertainty in model inputs has been characterized, it is then propagated through the model to characterize uncertainty in model outputs. Simulation based methods such as one and two dimensional Monte Carlo simulations (1D-MC and 2D-MC) have been used in risk assessments for propagating uncertainty where 1D-MC considers only epistemic uncertainty and 2D-MC accounts for both, aleatory and epistemic uncertainty (Cohen et al., 1996; EFSA et al., 2018; Helton, 1997; Nauta, 2000). An advantage of the 2D-MC is that maintains a distinction between aleatory and epistemic uncertainty when epistemic uncertainty is expressed by a precise probability (US EPA, 2011; Helton, 1997; Nauta, 2000; O'Hagan and West, 2013).

Methods for uncertainty quantification that rely on precise and bounded probability using Bayesian and robust Bayesian frameworks have also been proposed respectively (Burgman, 2005; Cox Jr., 2012; Lindley, 2006; O'Hagan et al., 2006; Ferson et al., 2003; Helton et al., 2004; Troffaes and Cooman, 2014; Walley, 1991). These two approaches are presented in more details later.

Quantification of uncertainty is usually done and summarized by estimating statistical quantities of interest such as expectations and percentiles (median, probability intervals) as well as bounds on probability or expectation which often relies on propagation of uncertainty methods. The motivation for research in this area is to argue for, and propose, solutions to integrate uncertainty analysis in evidence based decision making. This requires to expand current frameworks for evidence synthesis with analyses of robustness to uncertainty. Specifically, in this thesis, the focus is on the characterization of uncertainty by bounded probability using robust Bayesian analysis (i.e. Bayesian inference over a set of priors) which includes standard Bayesian analysis as an special case where uncertainty is quantified by precise probability.

Aims

The aim of this thesis is to contribute to evidence synthesis and quantitative risk assessments in scientific assessments by developing the methodology for robust Bayesian analysis, and applying and comparing different ways to quantify uncertainty in existing assessments. The specific objectives are to:

- Apply and evaluate methods for uncertainty analysis in scientific assessments (Paper II).
- Contribute to methods to quantify epistemic uncertainty by bounded probability (Paper I, Paper III).
- Contribute to the methodology of uncertainty analysis using Bayesian analysis and robust Bayesian analysis with applications in scientific assessments including evidence synthesis and quantitative risk assessment (Paper I-IV).

Theory and Methods

In this thesis, I used different research approaches: research done via case studies (Paper I-III) and research done using empirical studies (Paper IV). Specifically, in Paper I and Paper III, I used data from already published meta-analyses that have been conducted using traditional methods (classical approach). To illustrate and motivate the methodology, I reanalyzed the data, but within a robust Bayesian framework. In Paper II, a report comparing and demonstrating a probabilistic approach to quantify uncertainty was used to support the development of methods for uncertainty analysis at EFSA. In Paper IV, I extracted data from a global database to compare the effect of diversified farming practices on biodiversity and crop yields.

My research has emerged from practical problems in evidence synthesis and risk assessments. Next, I describe the Bayesian and robust Bayesian frameworks for quantifying uncertainty by precise and bounded probability respectively.

Bayesian analysis

The Bayesian approach provides a consistent framework to quantify uncertainty about model parameters and to make predictions of future events. A key feature is the use of subjective probability for quantifying uncertainty by setting a full probabilistic model. Bayesian analysis is a method of statistical inference that allows to combine prior knowledge in the form of probability with evidence (observed data). The Bayesian framework quantifies uncertainty about unknown parameters, θ, by treating them as random variables. This requires the specification of a prior distribution $p(\theta)$ which characterizes uncertainty about θ (i.e. epistemic uncertainty) before the data is observed. The prior distribution is updated with observed data, $\mathbf{x} = (x_1, \ldots, x_n)$ using Bayes' rule yielding the posterior distribution, $p(\theta|\mathbf{x})$ which is:

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) \cdot p(\theta) \tag{1}$$

where $p(\mathbf{x}|\theta)$ is the likelihood.

The prior can, but does not have to, be expressed with a parametric probability distribution, $p(\theta|t)$, (i.e. probability density function with a fixed number of parameters) (Cox, 2006). Parameters of parametric prior distributions are called hyperparameters. The use of hyperprior distributions, $p(t)$ to characterize uncertainty in prior distributions is known as hierarchical Bayesian modeling (Gelman et al., 2013).

A simple hierarchical Bayesian model is then:

$$p(\theta, t|\mathbf{x}) \propto p(\mathbf{x}|\theta) \cdot p(\theta|t) \cdot p(t). \tag{2}$$

As an example, variability in bodyweight of individuals of a certain age can be modeled by a normal distribution with unknown mean μ and known variance σ^2. Uncertainty about parameter μ can be modeled by a normal distribution with known mean μ_0 and variance σ_0^2:

$$\begin{aligned} x_i|\mu &\sim \mathrm{Normal}(\mu, \sigma^2), \\ \mu &\sim \mathrm{Normal}(\mu_0, \sigma_0^2). \end{aligned}$$

Dependencies between variables, parameters and hyperparameters can be represented through a Directed Acyclic Graph (DAG) (Figure 1).

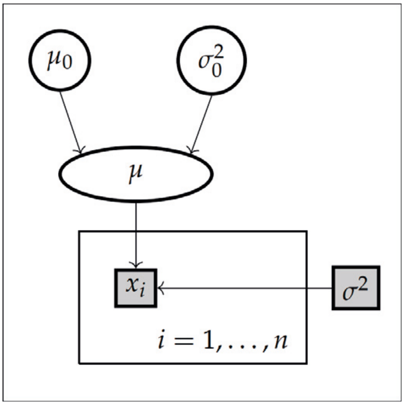
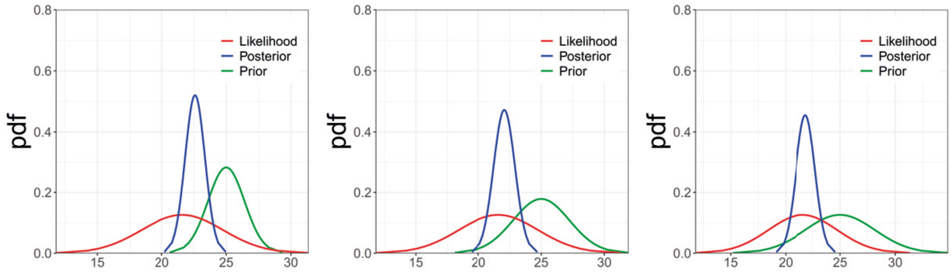


Figure 1: DAG of the hierarchical Bayesian model. Parameters are represented by ellipses, hyperparameters by circles and observed variables by grey squares. Plates indicates repeated cases.

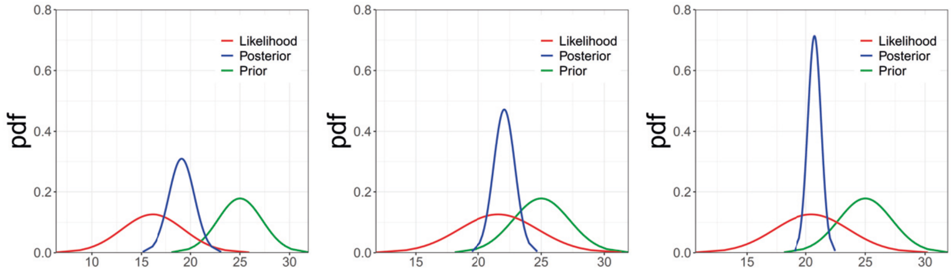
Bayesian statistical methods are often criticized as being subjective, specifically, due to the subjective nature of the choice of prior (i.e. driven by assessors' experience and expert knowledge) (Gelman, 2008; Reich and Ghosh, 2019). Prior distributions play an important role in Bayesian models (Gelman, 2008), however, different priors can be specified for the same model, and consequently obtain different posteriors which can lead to different conclusions (Figure 2). Moreover, when data is scarce (i.e. small sample size), the prior has more impact on posterior summaries (Figure 3). Therefore, priors should be carefully specified and it is also recommended to conduct a sensitivity analysis comparing the effect of different priors on the posterior distribution (Berger, 1990; Roos et al., 2015).



(a) $\mu_0 = 25, \sigma_0^2 = 2$ **(b)** $\mu_0 = 25, \sigma_0^2 = 5$ **(c)** $\mu_0 = 25, \sigma_0^2 = 10$

Figure 2: Effect of different priors (different hyperparameter values) on the posterior, n = 12. Normal prior density, normal likelihood (with unknown mean and known variance) and normal posterior density.

As more data are gathered, the impact of the prior is diminished and the posterior is shaped more like the likelihood (Figure 3).



(a) n = 4 **(b)** n = 12 **(c)** n = 30

Figure 3: Effect of data (sample size) on the posterior. Normal prior density, normal likelihood (with unknown mean and known variance) and normal posterior density.

A prior distribution is conjugate to the likelihood if the resulting posterior distribution belongs to the same family of distributions as the prior (Gelman et al., 2013). Conjugate priors are important in exact Bayesian inference. The conjugacy property is convenient because the analytical form of the posterior distribution

tion is known. However, the use of conjugate priors is limited in real applications because it is not always possible to specify conjugate priors for a specific problem (for example, it does not reflect prior knowledge or there is no conjugate prior for the specified model).

When the analytical form of the posterior distribution is unknown, numerical methods are needed for estimating posterior summaries such as expected values and credible intervals of parameters of interest. A common method is Markov chain Monte Carlo (MCMC) sampling which can handle high-dimensional probability distributions and therefore useful in Bayesian inference (Gelman et al., 2013; Liu, 2008). An important characteristic of MCMC sampling is that draws correlated samples from the posterior distribution.

Robust Bayesian analysis

Robust or robustness is typically used to refer to methods that can cope with uncertainty (Moallemi et al., 2020). Within the Bayesian framework, there are different types of robust analysis, for instance i) seen as Bayesian sensitivity analysis towards the choice of priors (Berger, 1990), ii) use of distributions with heavier tails thereby inference is less sensitive to outliers (Rosa et al., 2003), and the one in this thesis iii) Bayesian inference over a set of distributions (Bernard, 2005; Walley et al., 1996).

Robust Bayesian analysis provides a way to consider the impact of the choice of prior on uncertainty in relevant quantities by specifying a set of possible prior distributions (bounded probability) instead of a precise prior probability (Berger, 1990; Troffaes and Cooman, 2014; Walley, 1991). Consequently, a set of posterior distributions is derived from a set of prior distributions. This thesis focuses on the impact of prior distributions, but similar issues are raised when considering doubts about the specification of the likelihood of a given parametric model or dealing with imprecise data (Benavoli and Ristic, 2011; Cattaneo and Wiencierz, 2012).

This type of robust Bayesian analysis can be seen as an extension of Bayesian analysis which relies on statistical principles for inference within the theory of imprecise probability (Walley, 1991). From a technical point of view, this procedure is closely related to robust Bayesian inference. However, the degree of indeterminacy (imprecision) from limited information, ambiguity or ignorance, the use and interpretation of the resulting bounded probability goes beyond a simple sensitivity analysis and robustness analysis (Coolen et al., 2011). An example of this type of robust Bayesian analysis can be found in (Sahlin et al., 2021) where a robust Bayesian decision analysis is performed to solve an environmental risk management problem under severe uncertainty and value ambiguity. Moreover, Rinderknecht et al. (2012) used imprecise (bounded) probability to characterize ambiguity in probability distributions elicited from experts and discussed their implications for environmental decision support.

A set of priors represents prior beliefs measured through bounds. It reflects expert's uncertainty in the chosen prior or a range of prior judgments from several experts. For example, sets of prior distributions have been used to resolve prior-data conflicts (Walter et al., 2007; Walter and Augustin, 2009). Specifying sets of priors to model prior uncertainty involves some compromises among the following i) the calculation of the lower and upper bound should be as easy as possible; ii) the set should contain as many 'reasonable' priors as possible; iii) the set should correspond to easily elicitable prior distributions (Berger, 1990). Robust Bayesian analysis is useful to address severe uncertainty arising from low quality in knowledge, when data are scarce, for combining several sources of information and for dealing with expert elicitation (Sahlin et al., 2021).

In robust Bayesian analysis, one is interested in estimating a quantity of interest (e.g. expectations, percentiles or probability intervals) with uncertainty quantified by bounded probability (Berger, 1990; Troffaes and Cooman, 2014; Walley, 1991). For instance, the lower bound on expectation of a function f with respect to a set of posterior distributions, $\mathcal{M}$, is expressed as

$$\underline{\mathrm{E}}(f) := \inf_{p \in \mathcal{M}} \int f(x)p(x)dx. \tag{3}$$

Here a set of posterior distributions is derived from a set of prior distributions. If the difference between upper and lower bound is small enough then the indeterminacy in the prior is not relevant (Berger, 1990).

In practice, computations are not always easy since they require the estimation of supremum and infimum of quantities of interest which involve calculating complicated integrals. In the case of conjugate models, bounds on probability or expectation may be directly computed. Otherwise, it is required to apply Monte Carlo methods or MCMC sampling where the quantities of interest are approximated.

Results and Discussions

In this thesis, I have contributed to the methodology of robust Bayesian analysis for quantifying uncertainty by bounded probability with applications in scientific assessments. In general my research outputs are models and methods motivated by the need to distinguish between aleatory and epistemic uncertainty in scientific assessments, integrate subjective judgment with statistical inference and improve the methodology for uncertainty quantification.

My main contributions are i) an expression for the effective sample size of importance sampling that accounts for correlated MCMC samples, thereby allowing to combine iterative importance sampling with MCMC sampling for estimating an expected value with uncertainty quantified by bounded probability in robust Bayesian analysis (Paper I); ii) a framework, robust Bayesian analysis, that combines the principles of Bayesian inference and the theory of imprecise probability, which allows for epistemic uncertainty to be quantified by bounded probability (Paper II); iii) robust Bayesian bias analysis, as an alternative to consider ambiguity or ignorance about bias terms in meta-analysis (Paper III); iv) a probabilistic uncertainty analysis and a decision analysis applied to an evidence synthesis about the effect of diversified farming practices on biodiversity and crop yields. (Paper IV). More details are given below.

Paper I

Importance sampling is a technique within Monte Carlo methods that can be used for estimating expectations by weighting samples drawn from a proposal 'alternative' distribution which it is easier to generate samples from (Owen, 2013). An iterative version of importance sampling has been proposed for estimating expectation bounds using independent samples (Troffaes, 2017, 2018).

In this paper, we extended the use of iterative importance sampling proposed by Troffaes (2018) to models that require MCMC sampling. We combined iterative importance sampling with MCMC sampling for estimating expectation bounds in robust Bayesian analysis. To achieve this, it was needed to derive a new expression for the effective sample size, ESS, (i.e a measure of efficiency) of importance sampling which accounts for the correlation in the MCMC samples (the standard effective sample size relies on independent and identically distributed (iid) samples and therefore it is not possible to use). Yielding to the following estimate:

$$\mathrm{ESS} \approx \frac{\mathrm{ESS}_{\mathrm{MCMC}}}{N} \cdot \mathrm{ESS}_{\mathrm{IS}}. \tag{4}$$

The derived expression consists on the standard expression for the effective sample size of importance sampling $\mathrm{ESS}_{\mathrm{IS}}$ times the effective sample size of MCMC divided by the number of samples, $\frac{\mathrm{ESS}_{\mathrm{MCMC}}}{N}$, which accounts for a reduction in the effective sample size due to correlated MCMC samples. Other difference with respect to the method introduced in (Troffaes, 2018) is the requirement of a large sample size when updating via MCMC sampling to guarantee that the optimization is done using a reliable sample. In practice, we suggest setting the target effective sample size first and then specifying how much larger the effective sample size should be using MCMC sampling relative to the target effective sample size. In the application, we used a 20% greater, but, other values could have been used as well.

We also described how a set of prior distributions can be specified in robust Bayesian analysis using an approach similar to prior predictive check. The set of priors is here specified by different hyperparameter values.

Iterative importance sampling with MCMC sampling was applied to estimate the expected overall effect with uncertainty quantified by bounded probability in a previously published meta-analysis on the effect of biomanipulation with a random effects model. The conclusion of a positive effect from the original meta-analysis (done in a classical statistical framework) is confirmed to be robust to uncertainty associated with prior specification.

We tested the method's performance by using two target effective sample size, 5 000 and 10 000, as well as by changing the degree of prior-data conflict in the selected sets of priors. Our recommendation is to use the method when the set of priors does not have conflict with data (i.e. there is not prior-data conflict). Although, sets of priors allow to express prior knowledge more cautiously, expectation bounds are sensitive to the choice of sets of priors (Walter and Augustin, 2009). Iterative importance sampling with MCMC sampling offers more flexibility in robust Bayesian analysis, but it needs to be evaluated on more complex models to further assess its potential and limitations.

Paper II

This paper was motivated by the recommendation of EFSA to use quantitative expressions of epistemic uncertainty in scientific assessments (EFSA et al., 2018) and to some extend by the ongoing debate in the risk analysis community about the use of probability only as an expression of aleatory uncertainty and the use of other expressions of uncertainty for expressing epistemic uncertainty (Aven, 2010, 2020; Ferson and Ginzburg, 1996; Helton et al., 2004). We argued that any expression of uncertainty has pros and cons which should be recognized in the assessment.

Bounded probability is often presented using p-boxes (i.e. a pair of lower and upper cumulative distribution functions), where epistemic uncertainty is expressed by intervals and aleatory uncertainty by a probability distribution, or where there is no distinction made between the two types of uncertainty, e.g. using predictive distributions, or where there is no step to integrate evidence, e.g. uncertainty is given by expert judgment and there is no updating. In this paper, bounded probability was presented within the framework of robust Bayesian analysis instead of using p-boxes.

For this, we reviewed and discussed a mathematical framework for the quantification of epistemic uncertainty by precise and bounded probability using Bayesian and robust Bayesian analysis. The framework meets two important requirements for scientific assessment, i) the possibility to distinguish aleatory from epistemic uncertainty and ii) a systematic principle to integrate evidence to the assessment.

We also discussed propagation of uncertainty via two dimensional Monte Carlo simulation. We showed different representations of uncertainty at variable and parameter levels depending on how is uncertainty quantified (either by Bayesian analysis, probability bounds analysis or robust Bayesian analysis). For instance, when parameter uncertainty (i.e. epistemic) is characterized by a precise probability, an interval and a bounded probability then, uncertainty in variable can be characterized by 2D-distribution, a p-box and 2D-distributions for every distribution in the set of probability distributions for the parameter respectively (Figure 1, Paper II). However, we argued that robust Bayesian analysis, does not have a standard way to visualize uncertainty at the variable level in a manner that clearly separate aleatory and epistemic uncertainty similar to the 2D-distribution in the precise probability case. We highlighted the importance of communicating the results of the assessment in a way that shows that uncertainty has been quantified.

Bayesian and robust Bayesian analysis were applied to an intake assessment about aluminium intake (chronic toxicity) by consumption of chocolate and cocoa products for children (Schendel et al., 2018). The intake assessment contains several continuous assessment variables, which are combined to estimate a quantity of interest, (i.e. the frequency of exceeding the tolerable weekly intake of aluminium of a child in a target population). For doing this, a hierarchical Bayesian model was specified and later extended to a robust Bayesian model by specifying a set of priors. Both Bayesian and robust Bayesian analysis include the possibility, although with small probability, of exceeding the weekly intake safety threshold. We also presented how results can be communicated using quantitative and verbal expressions.

Paper III

Quantitative bias analysis is a statistical method that acknowledges differences in study quality associated with the design and conduct of studies (i.e. risk of bias) in meta-analysis (Lash et al., 2009, 2014). Quantitative bias analysis requires the meta-analysis model to be extended with bias adjustments (e.g. additive or proportional adjustments of study specific errors in the model) (Turner et al., 2009; Spiegelhalter and Best, 2003; Verde, 2021), and additional expert judgment (Turner et al., 2009; van der Bles et al., 2019; Rhodes et al., 2020) on bias terms.

We proposed a way to consider uncertainty, arising from ambiguity or ignorance, about bias terms by modeling bias in a bias-adjusted random effects model with imprecision. We described how to transform qualitative judgments about risk of bias (using the Cochrane's risk of bias table) into a set of bias terms (bounded probability) to be incorporated in a meta-analysis. Then, the difference between the bounds on the overall effect is only associated to how we specify the risk of bias.

We illustrated a robust Bayesian bias analysis on a bias-adjusted random effects model with data from a published meta-analysis from the Cochrane Collaboration where the risk of bias table (i.e. a qualitative assessment of the quality of evidence) is available (Lopez-Olivo et al., 2015). The Cochrane's risk of bias table contains six risk of bias domains which are assessed separately and classified in three categories: low, unclear and high risk of bias (Higgins et al., 2019).

The paper outlined some considerations for establishing the rate between studies and constructing the set of study qualities based on the three categories. The process was described and illustrated step by step for each Cochrane's risk of bias table domain. It also included an example of how multiple domains can be combined, thereby giving a full description of possible cases.

For comparison, both unadjusted and robust Bayesian bias-adjusted random effects models were applied for each risk of bias domain. The estimated overall effect and specific study effect were displayed using forestplots for each bias domain (Appendix, Paper III). The results from robust Bayesian analysis, were added to the standard forestplots which showed bounds on the expected overall effect, the lower 2.5th percentile and the upper 97.5th percentile (Figure 2, Paper III).

The difference between bounds (degree of imprecision) of the expected overall effect varied when adjusting for different risk of bias domains (Figure 3, Paper III). As expected, there was more (less) imprecision in the estimated overall effect when all studies had an unclear (low) risk of bias, Domain 1-2 (Domain 5-6).

Adjusting for bias may reveal important aspects to consider when framing a conclusion in evidence synthesis. In this example, adjusting for bias did not have a large impact on the estimated overall effect of the meta-analysis (Lopez-Olivo et al., 2015) (Figure 3, Paper III). Thus, the evidence in favor of the treatment group from the meta-analysis remained strong after adjusting for bias.

The use of robust Bayesian bias-adjustment modeling opens up to include more studies into the meta-analysis. It also allows to consider the influence of indirect uncertainty in the specification of direct uncertainty. More experience and training with robust Bayesian bias adjustment analysis needs to be gained in terms of real decision-making. The proposed approach contributes to bridge the gap between qualitative and quantitative expressions of uncertainty.

Paper IV

Agricultural intensification is one of the main drivers of global biodiversity loss (Emmerson et al., 2016; Tilman et al., 2002; Wittwer et al., 2021). To mitigate this problem, diversified farming practices, such as agroforestry, crop rotation and embedded natural, have been proposed as more sustainable strategies for land use and management and can contribute to reduce the negative impact of agricultural intensification on biodiversity loss (Beillouin et al., 2021; Rosa-Schleich et al., 2019; Sánchez et al., 2021; Wittwer et al., 2021).

We compared the effect of different diversified farming practices with respect to their effect on biodiversity and crop yields. For biodiversity, we selected the organism groups: birds, insects (natural enemies) and below ground organisms; and two common biodiversity metrics: species abundance and species richness. For the analysis, data was extracted from a global database of the effects of diversified farming practices on biodiversity and crop yields (Jones et al., 2021). The comparisons were made using Bayesian network meta-analyses (an extension of pairwise meta-analysis) which allows to compare three or more interventions in a single coherent analysis of all relevant studies by generating direct and indirect (based on a common comparator) comparisons of interventions (Dias et al., 2018; Dias and Caldwell, 2019; Hu et al., 2020). We did not find a clear effect of diversified farming because between study heterogeneity was large, which can be explained by the heterogeneity in the data.

In addition, we compared alternative allocations of farming practices across a set of hypothetical farms. The decision analysis model was calibrated on network meta-analysis models where biodiversity was measured by species richness and contained multiple diversified farming practices. The decision analysis using portfolio theory was conducted considering uncertainty about parameters of the Bayesian network meta-analysis models, which were later propagated into the utility functions. Bayesian network meta-analysis has the potential to be useful to calibrate the portfolio analysis, however, due to the large variability and uncertainty in the network meta-analysis, the paper showed the need to include more random effects.

The conclusion from this paper highlighted challenges in the use of evidence based decision making in environmental policy, as there was no diversified farming practice that clearly performed better than the other.

Decision making context

Papers I-IV have used different studies to estimate a quantity of interest. Although these types of assessments are very important, it is necessary to put them into a decision context to inform decisions and provide scientific advise.

For example, (National Research Council, 2009) proposed a framework for risk-based decision making that comprises three phases: 1) problem formulation and scoping, 2) planning and conduct of risk assessment and 3) risk management. This framework highlights the importance of the communication between risk assessors and decision makers to identify and specify an assessment question, what the assessment model represents, how to summarize the model and how to distinguish between aleatory and epistemic uncertainty. The task of risk assessors is mainly to answer the assessment question whereas the task of decision makers is to make decisions based on uncertainty about the conclusion of the assessment.

Three different types of decisions are also discussed in (Fischhoff and Davis,

2014) which distinguish between decisions about action threshold, with fixed options and about potential options. These types of decisions are closely related to the assessment question and how uncertainty has been characterized. Decision makers often choose the alternative with less uncertainty (Gärdenfors and Sahlin, 1988) without relying on decision rules.

To support and guide decision making, decision theories and approaches have been developed, for example, theories relying on maximization (minimization) of expected utility (loss) (e.g. Bayesian decision theory) (Parmigiani and Inoue, 2009; Reich and Ghosh, 2019) or maximization (minimization) of minimal (maximal) expected effect (e.g. within a robust Bayesian framework) (Walley, 1991). Scenario based approaches have also been proposed to cope with uncertainty where possible alternative scenarios are considered in the assessment (Parmigiani and Inoue, 2009).

In summary, the results of this thesis provide important insights as well as methodological developments on the quantification of uncertainty by subjective probability in scientific assessments. Furthermore, the results can directly be used in initiatives for evidence synthesis, not limited to the field of environmental sciences.

Environmental science perspective

Value of modeling uncertainty

Environmental science is an interdisciplinary field which focuses on understanding how environmental systems (i.e. natural systems under influence of human activity) function and finding solutions to environmental problems. Although, there is a lot of uncertainty associated with these systems, for instance: natural variability in the biological processes; limited or imperfect data due to observation errors and/or incomplete understanding of environmental systems (model structure and dynamics), decisions still need to be made. To take this uncertainty into account, it is required to characterize, assess and convey uncertainty (Fischhoff and Davis, 2014).

What to communicate is closely related to how is uncertainty expressed. For example, in case a quantitative expression is chosen, uncertainty may be expressed by a subjective probability (either precise or bounded). These two measures of uncertainty are advantageous as they allow for propagation and continuous learning when new data become available. The Bayesian approach formalizes the use of subjective information, including personal judgments from experts, for characterizing uncertainty about parameter estimates and for making predictions. This is an useful feature because, in many risk assessments, available information is limited and sometimes even fully subjective. In addition, the Bayesian framework allows to make predictions conditional on data which contributes to better support management decisions (Maier et al., 2008).

Value of expert knowledge

Expert knowledge is a valuable source of information to support decisions in scientific assessments where data is typically scarce (Cook et al., 2010; Drescher and Edwards, 2019). Sometimes, expert judgments may be the only basis for an informed management decision, without any additional empirical evidence (Drescher et al., 2013; Martin et al., 2012). For instance, in (EFSA (AHAW) et al., 2021), expert knowledge is used to quantify uncertainty in the conclusion of an assessment regarding swine fever, by looking just at the evidence.

Expert knowledge provides a credible source of information for supporting management decisions, modeling species distributions and assessing the detection probability, prevalence, and risk of establishment of an invasive species (Burgman, 2005; Drescher et al., 2013; Kuhnert, 2011; Martin et al., 2012). Expert knowledge can provide information about model parameters and thus help to characterize uncertainty in models (Drescher et al., 2013). For example, in (EFSA (BIOHAZ) et al., 2020), expert knowledge is used to quantify uncertainty in parameters (model inputs) of a stochastic model for Salmonella detection sensitivity, and then propagates it within the model, without using Bayesian inference to learn from data.

Elicitation of expert knowledge in the form of a (subjective) probability is relevant in Bayesian analysis for specifying informative prior distributions (O'Hagan et al., 2006; O'Hagan, 2019). The use of expert knowledge in ecology and conservation biology is broad and diverse. It can be used to inform ecological models and to fill potential knowledge gaps or pitfalls, that could result, especially with limited supporting evidence. For example, Murray et al. (2009) used informative priors based on expert knowledge with field data in a species distribution model within the Bayesian framework. Martin et al. (2005) elicited and used expert knowledge (as prior) to inform an ecological model concerning the effect of grazing on birds using the Bayesian approach. Barons et al. (2018) estimated the probability of good pollinator abundance under weather, disease, and habitat conditions using structured expert judgment. Bayesian modeling with informative priors based on expert knowledge can provide a useful "bridge" for ecologists, from purely conceptual models to statistical models that are calibrated to observed data (Low-Choy et al., 2009).

Conclusions

This thesis dealt with methods to quantify uncertainty by probability such as Expert Knowledge Elicitation, Bayesian and robust Bayesian analysis. Bayesian and robust Bayesian analysis provide a rigorous methodology to quantify uncertainty by precise and bounded probability to support decision making under uncertainty in evidence synthesis. Both frameworks are useful for modeling decision problems and learning from experience. In practice, the use of bounded or precise probability depends on several factors such as the nature of available information, the amount of risk involved, the aim of the analysis, as well as computational requirements.

This thesis has contributed to the methodology of robust Bayesian analysis by developing and extending methods to quantify uncertainty by bounded probability. It has also evaluated and improved the methodology of uncertainty analysis in scientific assessments by making use of Bayesian and robust Bayesian analysis. This thesis has shown the potential of robust Bayesian analysis in frameworks for evidence based decision making where uncertainty arise from low quality data, scarce data, imprecise information or disagreement among experts. It has also showed how to quantify uncertainty and propagate it into a decision analysis.

Future work will explore the performance of iterative importance sampling with MCMC sampling in more complex models and how to extend Bayesian network meta-analysis with additional random effects. To conclude, there is still a need for research testing the usefulness, applicability and reliability of these methods for uncertainty analysis in evidence synthesis and evaluate how to fit them in evidence based decision making.

References

Apostolakis, G. (1990). The concept of probability in safety assessment of technological systems. *Science*, 250(4986):1359–1364.

Aven, T. (2010). On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk Analysis*, 30(3):354–360.

Aven, T. (2011). *Quantitative Risk Assessment: The Scientific Platform*. Cambridge University Press.

Aven, T. (2020). Bayesian analysis: Critical issues related to its scope and boundaries in a risk context. *Reliability Engineering & System Safety*, 204:107209.

Aven, T., Andersen, H. B., Cox, T., Droguett, E. L., Greenberg, M., Guikema, S., Kröger, W., McComas, K., Renn, O., Thompson, K. M., and Zio, E. (2018). Risk analysis: Fundamental principles.

Barons, M. J., Hanea, A. M., Wright, S. K., Baldock, K. C. R., Wilfert, L., Chandler, D., Datta, S., Fannon, J., Hartfield, C., Lucas, A., Ollerton, J., Potts, S. G., and Carreck, N. L. (2018). Assessment of the response of pollinator abundance to environmental pressures using structured expert elicitation. *Journal of Apicultural Research*, 57(5):593–604.

Beillouin, D., Ben-Ari, T., Malézieux, E., Seufert, V., and Makowski, D. (2021). Positive but variable effects of crop diversification on biodiversity and ecosystem services. *Global Change Biology*, 27(19):4697–4710.

Benavoli, A. and Ristic, B. (2011). Classification with imprecise likelihoods: A comparison of TBM, random set and imprecise probability approach. In *14th International Conference on Information Fusion, Chicago, IL*, pages 1–8.

Berger, J. O. (1990). Robust Bayesian analysis: sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25:303–328.

Bernard, J.-M. (2005). An introduction to the imprecise Dirichlet model for Multinomial data. *International Journal of Approximate Reasoning*, 39(2):123–150.

Burgman, M. (2005). *Risks and Decisions for Conservation and Environmental Management*. Cambridge University Press.

Cattaneo, M. E. and Wiencierz, A. (2012). Likelihood based imprecise regression. *International Journal of Approximate Reasoning*, 53(8):1137–1154.

CEE (2013). Guidelines for systematic review and evidence synthesis in environmental management. version 4.2.

Cohen, J. T., Lampson, M. A., and Bowers, T. S. (1996). The use of two-stage monte carlo simulation techniques to characterize variability and uncertainty in risk analysis. *Human and Ecological Risk Assessment: An International Journal*, 2(4):939–971.

Cook, C. N., Hockings, M., and Carter, R. B. (2010). Conservation in the dark? The information used to support management decisions. *Frontiers in Ecology and the Environment*, 8(4):181–186.

Coolen, F. P. A., Troffaes, M. C. M., and Augustin, T. (2011). *Imprecise Probability*, pages 645–648. Springer Berlin Heidelberg.

Cox, D. (2006). *Principles of Statistical Inference*. Cambridge University Press.

Cox Jr., L. A. (2012). Confronting deep uncertainties in risk analysis. *Risk Analysis*, 32(10):1607–1629.

DeLuca, J. B., Mullins, M. M., Lyles, C. M., Crepaz, N., Kay, L., and Thadiparthi, S. (2008). Developing a comprehensive search strategy for evidence based systematic reviews. *Evidence Based Library and Information Practice*, 3(1):3–32.

Deluyker, H. (2017). Is scientific assessment a scientific discipline? *EFSA Journal*, 15(11):e15111, 51 pp.

Dias, S., Ades, T., Welton, N. J., Jansen, J., and Sutton, A. J. (2018). *Network Meta-Analysis for Decision-Making*. Wiley, United States.

Dias, S. and Caldwell, D. M. (2019). Network meta-analysis explained. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 104(1):F8–F12.

Drescher, M. and Edwards, R. C. (2019). A systematic review of transparency in the methods of expert knowledge use. *Journal of Applied Ecology*, 56(2):436–449.

Drescher, M., Perera, A. H., Johnson, C. J., Buse, L. J., Drew, C. A., and Burgman, M. A. (2013). Toward rigorous use of expert knowledge in ecological research. *Ecosphere*, 4(7):83.

EFSA, Committee, S., Benford, D., Halldorsson, T., Jeger, M. J., Knutsen, H. K., More, S., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Schlatter, J. R., Silano, V., Solecki, R., Turck, D., Younes, M., Craig, P., Hart, A., Goetz, N. V., Koutsoumanis, K., Mortensen, A., Ossendorp, B., Germini, A., Martino, L., Merten, C., Mosbach-Schulz, O., Smith, A., and Hardy, A. (2018). Scientific opinion on the principles and methods behind EFSA's guidance on uncertainty analysis in scientific assessment. *EFSA Journal*, 16(1):e05122.

EFSA, Hart, A., Maxim, L., Siegrist, M., Goetz, N. V., da Cruz, C., Merten, C., Mosbach-Schulz, O., Lahaniatis, M., Smith, A., and Hardy, A. (2019). Guidance on communication of uncertainty in scientific assessments. *EFSA Journal*, 17(1):5520.

EFSA (AHAW), Nielsen, S. S., Alvarez, J., Bicout, D. J., Calistri, P., Canali, E., Drewe, J. A., Garin-Bastuji, B., Rojas, J. L. G., Schmidt, C. G., Herskin, M., Michel, V., Ángel Miranda Chueca, M., Padalino, B., Pasquali, P., Sihvonen, L. H., Spoolder, H., Ståhl, K., Velarde, A., Viltrop, A., Winckler, C., Gubbins, S., Stegeman, J. A., Antoniou, S.-E., Aznar, I., Broglia, A., Lima, E., der Stede, Y. V., Zancanaro, G., and Roberts, H. C. (2021). Assessment of the control measures of the category a diseases of animal health law: Classical swine fever. *EFSA Journal*, 19(7):e06707.

EFSA (BIOHAZ), Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bolton, D., Bover-Cid, S., Chemaly, M., Davies, R., Cesare, A. D., Herman, L., Lindqvist, R., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Álvarez Sánchez, J., Blagojevic, B., Fürst, P., Garin-Bastuji, B., Jensen, H. E., Paulsen, P., Baert, K., Barrucci, F., Broglia, A., Georgiadis, M., Hempen, M., and Hilbert, F. (2020). Evaluation of public and animal health risks in case of a delayed post-mortem inspection in ungulates. *EFSA Journal*, 18(12):e06307.

Emmerson, M. C., Morales, M. B., Oñate, J. J., Batáry, P., Berendse, F., Liira, J., Aavik, T., Guerrero, I., Bommarco, R., Eggers, S., Pärt, T., Tscharntke, T., Weisser, W., Clement, L., and Bengtsson, J. (2016). Chapter Two -

How agricultural intensification affects biodiversity and ecosystem services. In Dumbrell, A. J., Kordas, R. L., and Woodward, G., editors, *Large-Scale Ecology: Model Systems to Global Perspectives*, volume 55 of *Advances in Ecological Research*, pages 43–97. Academic Press.

FAO and WHO (2021). *Microbiological risk assessment - Guidance for food.* Microbiological Risk Assessment Series No. 36. Rome.

Ferson, S. and Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2):133–144. Treatment of Aleatory and Epistemic Uncertainty.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D., and Sentz, K. (2003). Constructing probability boxes and dempster-shafer structures.

Fischhoff, B. and Davis, A. L. (2014). Communicating scientific uncertainty. *Proceedings of the national academy of sciences of the United States of America*, 111(4):13664–13671.

Gärdenfors, P. and Sahlin, N. (1988). *Decision, Probability and Utility: Selected Readings.* Cambridge University Press.

Gelman, A. (2008). Objections to bayesian statistics. *Bayesian Analysis*, 3(3):445–450.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition.* Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Hampton, J., Moore, P., and Thomas, H. (1973). Subjective probability and its measurement. *Journal of the Royal Statistical Society. Series A (General)*, 136(1):21–42.

Hansson, S. O. (2008). Do we need second-order probabilities? *Dialectica*, 62(4):525–533.

Helton, J. C. (1997). Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. *Journal of Statistical Computation and Simulation*, 57(1-4):3–76.

Helton, J. C., Johnson, J. D., and Oberkampf, W. L. (2004). An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering & System Safety*, 85(1):39–71.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., and (editors), V. A. W. (2019). *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019).*

Hu, D., O'Connor, A. M., Wang, C., Sargeant, J. M., and B.inder, C. (2020). How to conduct a Bayesian network meta-analysis. *Frontiers in Veterinary Science*, 7:271.

Institute of Medicine (2013). *Environmental Decisions in the Face of Uncertainty: Chapter 6.* The National Academies Press.

Jones, S. K., Sánchez, A. C., Juventia, S. D., and Estrada-Carmona, N. (2021). A global database of diversified farming effects on biodiversity and yield. *Scientific Data volume*, 8(212).

Kuhnert, P. M. (2011). Four case studies in using expert opinion to inform priors. *Environmetrics*, 22(5):662–674.

Lash, T. L., Fox, M. P., and Fink, A. K. (2009). *Applying Bias Analysis to Epidemiologic Data.* New York, NY: Springer.

Lash, T. L., Fox, M. P., MacLehose, R. F., Maldonado, G., McCandless, L. C., and Greenland6, S. (2014). Good practices for quantitative bias analysis. *International Journal of Epidemiology*, 43(6):1969–1985.

Lindley, D. V. (2006). *Understanding Uncertainty.* John Wiley & Sons, Inc.

Liu, J. S. (2008). *Monte Carlo Strategies in Scientific Computing.* Springer, New York.

Lopez-Olivo, M. A., Amezaga-Urruela, M., McGahan, L., Pollono, E. N., and Suarez-Almazor, M. E. (2015). Rituximab for rheumatoid arthritis. *Cochrane Database of Systematic Reviews*, 1(CD007356).

Low-Choy, S., O'Leary, R., and Mengersen, K. (2009). Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. *Ecology*, 90(1):265–277.

Maier, H., Ascough, J., Wattenbach, M., Renschler, C., Labiosa, W., and Ravalico, J. (2008). Chapter five: Uncertainty in environmental decision making: Issues, challenges and future directions. *Developments in Integrated Environmental Assessment*, 3:69–85.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Choy, S., McBride, M., and Mengersen, K. (2012). Eliciting expert knowledge in conservation science. *Conservation Biology*, 26(1):29–38.

Martin, T. G., Kuhnert, P. M., Mengersen, K., and Possingham, H. P. (2005). The power of expert opinion in ecological models using bayesian methods: Impact of grazing on birds. *Ecological Applications*, 15(1):266–280.

Mikolajewicz, N. and Komarova, S. V. (2019). Meta-analytic methodology for basic research: A practical guide. *Frontiers in Physiology*, 10:203.

Moallemi, E. A., Elsawah, S., and Ryan, M. J. (2020). Robust decision making and epoch–era analysis: A comparison of two robustness frameworks for decision-making under uncertainty. *Technological Forecasting and Social Change*, 151:119797.

Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences*, 111(20):7176–7184.

Murray, J. V., Goldizen, A. W., O'Leary, R. A., McAlpine, C. A., Possingham, H. P., and Choy, S. L. (2009). How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? a case study using brush-tailed rock-wallabies petrogale penicillata. *Journal of Applied Ecology*, 46(4):842–851.

National Research Council (2009). *Science and Decisions: Advancing Risk Assessment*. The National Academies Press, Washington, DC.

Nauta, M. J. (2000). Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology*, 57(1):9–18.

O'Hagan, A. (2012). Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. *Environmental Modelling & Software*, 36(10):35–48.

O'Hagan, A. (2019). Expert knowledge elicitation: Subjective but scientific. *The American Statistician*, (S1):69–81.

O'Hagan, A., Buck, C., Daneshkhah, A., and Eiser, J. (2006). *Uncertain Judgments: Eliciting Experts' Probabilities*. Chichester: Wiley.

O'Hagan, A. and West, M. (2013). *The Oxford Handbook of Applied Bayesian Analysis*. Oxford University Press, Oxford.

Owen, A. B. (2013). Monte Carlo theory, methods and examples.

Parmigiani, G. and Inoue, L. Y. T. (2009). *Decision Theory: Principles and Approaches*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd.

Reich, B. J. and Ghosh, S. K. (2019). *Bayesian Statistical Methods*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Rhodes, K. M., Savović, J., Elbers, R., Jones, H. E., Higgins, J. P. T., Sterne, J. A. C., Welton, N. J., and Turner, R. M. (2020). Adjusting trial results for biases in meta-analysis: combining data-based evidence on bias with detailed trial assessment. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(1):193–209.

Rinderknecht, S. L., Borsuk, M. E., and Reichert, P. (2012). Bridging uncertain and ambiguous knowledge with imprecise probabilities. *Environmental Modelling & Software*, 36:122–130. Thematic issue on Expert Opinion in Environmental Modelling and Management.

Roos, M., Martins, T. G., Held, L., and Rue, H. (2015). Sensitivity Analysis for Bayesian Hierarchical Models. *Bayesian analysis*, (2):321–349.

Rosa, G. J., Padovani, C. R., and Gianola, D. (2003). Robust linear mixed models with normal/independent distributions and bayesian mcmc implementation. *Biometrical Journal*, 45(5):573–590.

Rosa-Schleich, J., Loos, J., Mußhoff, O., and Tscharntke, T. (2019). Ecological-economic trade-offs of diversified farming systems – A review. *Ecological Economics*, 160:251–263.

Sahlin, U. and Troffaes, M. (2017). A note on efsa's ongoing efforts to increase transparency of uncertainty in scientific opinions. *Journal of Risk Research*, pages 1–8.

Sahlin, U., Troffaes, M. C., and Edsman, L. (2021). Robust decision analysis under severe uncertainty and ambiguous tradeoffs: An invasive species case study. *Risk Analysis*.

Schendel, T., Jung, C., Lindtner, O., and Greiner, M. (2018). Guidelines for uncertainty analysis: Application of the respective documents of efsa and bfr for exposure assessments. *EFSA Journal*, pages 1–113.

Singpurwalla, N. D. and Wilson, A. G. (2008). Probability, chance and the probability of chance. *IIE Transactions*, 41(1):12–22.

Sánchez, A. C., Estrada-Carmona, N., Juventia, S. D., and Jones, S. K. (2021). The impact of diversified farming practices on terrestrial biodiversity outcomes and agricultural yield worldwide: A systematic review protocol. *Methods and Protocols*, 4(8).

Spiegelhalter, D. J. and Best, N. G. (2003). Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine*, 22(23):3687–3709.

Suter, G., Nichols, J., Lavoie, E., and Cormier, S. (2020). Systematic review and weight of evidence are integral to ecological and human health assessments: They need an integrated framework. *Integrated Environmental Assessment and Management*, 16(5):718–728.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418:671–677.

Troffaes, M. (2017). A note on Imprecise Monte Carlo over Credal Sets via Importance Sampling. *ISIPTA*.

Troffaes, M. C. and Cooman, G. D. (2014). *Lower previsions*. Wiley series in probability and statistics. John Wiley and Sons.

Troffaes, M. C. M. (2018). Imprecise Monte Carlo simulation and iterative importance sampling for the estimation of lower previsions. *International Journal of Approximate Reasoning*, 101:31–48.

Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S., and Thompson, S. G. (2009). Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society. Serie A, (Statistics in Society)*, (1):21–47.

US EPA (2011). *Exposure Factors Handbook (Final Report): Chapter 2*.

van der Bles, A. M., van der Linden, S., Freeman, A. L., Mitchell, J., Galvao, A. B., Zaval, L., and Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *The Royal Society*.

Verde, P. E. (2021). A bias-corrected meta-analysis model for combining, studies of different types and quality. *Biometrical Journal*, 63(2):406–422.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London.

Walley, P., Gurrin, L., and Burton, P. (1996). Analysis of clinical data using imprecise prior probabilities. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 45(4):457–485.

Walter, G. and Augustin, T. (2009). Imprecision and prior-data conflict in generalized Bayesian inference. *Journal of Statistical Theory and Practice*, 3(1):255–271.

Walter, G., Augustin, T., and Peters, A. (2007). Linear regression analysis under sets of conjugate priors. *ISIPTA'07-Society for Imprecise Probability: Theories and Applications*, pages 445–455.

Wittwer, R. A., Bender, S. F., Hartman, K., Hydbom, S., Lima, R. A. A., Loaiza, V., Nemecek, T., Oehl, F., Olsson, P. A., Petchey, O., Prechsl, U. E., Schlaeppi, K., Scholten, T., Seitz, S., Six, J., and van der Heijden, M. G. A. (2021). Organic and conservation agriculture promote ecosystem multifunctionality. *Science Advances*, 7(34):eabg6995.

Acknowledgements

I would like to express my sincere gratitude to my supervisors **Ullrika**, **Matthias**, **Johan**, **Henrik** and **Nils Eric** for their continuous support at every stage of the research project. I have learnt a lot from them.

In particular, I am deeply grateful to **Ullrika** not only for sharing her expertise in the topic of uncertainty, but also for her kindness and patience. I also thank her for encouraging me and further pushing my limits and ambitions. To her lovely family goes my sincere appreciation for the warm welcoming when I first came to Sweden as well as for the many family trips. I look forward to our future collaborations and adventures.

I would like to offer my special thanks to **Matthias** for welcoming me into the community of imprecise probability and for being a great host during my stay in Durham, UK.

I want to acknowledge my supervisors **Johan**, **Nils Eric** and **Henrik** for sharing their knowledge in Bayesian modeling, evidence based medicine decision making and interdisciplinary studies, respectively.

I also say thank you to my co-authors **Martin** and **Sarah** for their valuable comments and feedback.

I want to thank my master thesis advisor **Vivian Sistasch Vega** for teaching me Statistics and for being such an inspiration. She was the one who most encouraged me to pursue a PhD. I deeply appreciate and respect her friendship.

Thanks to everyone at CEC for making this such a nice and friendly working environment. In particular, thanks to **Pål-Axel, Natascha** and **Maria Hansson** for their support through my PhD education.

Thank you to **Åsa-Katrin, Irene, Annika, Ylva** and **Helena** for helping me with all the administrative support related to my PhD studies.

Thanks to **all CEC PhD students** for interesting discussions, shared moments in courses, fikas, excursions and retreats. Big thanks to my officemates **Terese, William, Dmytro** and **Pedro** for nice chats and discussions.

Special thanks to my dearest friends: **Maria**, **Tristan**, **Leidys**, **Øystein** and **Hamid**. My time in Lund would not have been the same without all of you.

A **Luis, Milay, Diego y Diana**. Muchas gracias por hacerme sentir parte de su familia, ustedes son mi familia cubana en Suecia.

Mis amigas de Cuba: **Adriana, Amanda, Claudia, Laura, Anisleydi, Chabela, Silvia y Yolaida** por llenar mi vida de alegría, entusiasmo, por siempre creer en mí y en nuestra amistad.

A mi familia le agradezco infinitamente todo el apoyo. En particular: a mi tía **Sonia** por su preocupación y cariño. A mis padres, **Ivonne y Manuel**, y mi hermana **Anette** muchas gracias por motivarme a ser cada día mejor, asi como por su amor incondicional. La distancia entre nosotros ha sido lo más difícil.

A **Daniel** le agradezco el apoyo, la confidencia, el amor y todas las cosas que nos quedan por vivir juntos.