DISS. ETH NO. 20047

CONTRIBUTIONS TO THE USE OF IMPRECISE SCIENTIFIC KNOWLEDGE IN DECISION SUPPORT

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

SIMON LUKAS RINDERKNECHT

M.Sc. Major Mathematics, Minor Earth Sciences
Université de Genève CH

B.Sc. Mathematics
Université de Neuchâtel CH

Date of birth
08. February 1979
citizen of Wallisellen ZH

accepted on the recommendation of
Prof. Dr. Peter Reichert, examiner
Prof. Dr. Marco Zaffalon, co-examiner
Prof. Dr. Mark E. Borsuk, co-examiner
Prof. Dr. Hans R. Künsch, co-examiner

2011

DISS. ETH NO. 20047

CONTRIBUTIONS TO THE USE OF IMPRECISE SCIENTIFIC KNOWLEDGE IN DECISION SUPPORT

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

SIMON LUKAS RINDERKNECHT

M.Sc. Major Mathematics, Minor Earth Sciences
Université de Genève CH

B.Sc. Mathematics
Université de Neuchâtel CH

Date of birth
08. February 1979
citizen of Wallisellen ZH

accepted on the recommendation of
Prof. Dr. Peter Reichert, examiner
Prof. Dr. Marco Zaffalon, co-examiner
Prof. Dr. Mark E. Borsuk, co-examiner
Prof. Dr. Hans R. Künsch, co-examiner

2011

To my family.

Intentionally left blank for a handwritten dedication.

Preface

This doctoral thesis at hand is only a small additional contribution to the already existing scientific findings. It became possible with the help and great assistance of many people that are explicitly mentioned in the acknowledgements. They know that I know that — and vice versa. However, I hope this work may be a useful tessera in the mosaic that depicts the area of its according expertise.

I would like to take the opportunity such a preface offers to emphasize something that has not necessarily or directly to do with this thesis, namely, the perpetual importance of:

Simultaneous Authenticity in Thinking, Feeling and Doing

correct — wrong — think about it
good — bad — feel it
everything — nothing — do it

Because – at the end – one has to take a decision. I always was fascinated by the interaction of the triplet *esprit*, *corps* and *âme*. Awareness of how *Körper*, *Geist* and *Seele* play together may be treated in many contexts such as in ethics, philosophy, music or also decision analysis – amongst others.

I had many interesting discussions with colleagues and friends about what is called here *Simultaneous Authenticity in Thinking, Feeling and Doing*. It must be something that is strong and it seems to me, personally, that it always was - and always is - worth it to pay some attention to that triplet.

Dübendorf, October 2011 Simon L. Rinderknecht

Zusammenfassung der Doktorarbeit

Das Ziel der Entscheidungstheorie ist es, Entscheidende dabei zu unterstützen, eine Entscheidungsalternative zu finden, die zur bestmöglichen Erfüllung ihrer Ziele führt. Das braucht (1) die quantitative Bewertung möglicher Konsequenzen in Funktion messbarer Systemattribute, (2) die Vorhersage von Wahrscheinlichkeitsverteilungen dieser Attribute für alle Entscheidungsalternativen und (3) das Erstellen einer Rangliste aller Entscheidungsalternativen durch die Kombination von (1) und (2). Um quantitative Vorhersagen zu erhalten, ist es oft unumgänglich, Wissen über Einflussfaktoren, Modellparameter oder Attributwerte von Fachexperten zu erheben. Leider kann die Charakterisierung subjektiver Überzeugungsgrade, in einem Bayesianischen Kontext verstanden, oft mit Ambiguität behaftet sein. Wir schlagen den Gebrauch von unpräzisen Wahrscheinlichkeiten vor, um Ambiguität in Bezug auf Unsicherheit zu beschreiben. In diesem robusten Bayesianischen Konzept evaluierten wir diverse Wahrscheinlichkeitsverteilungsklassen. Wir befanden die *Dichte-Verhältnis-Klasse* als die geeignetste: (a) Sie kann intersubjektives Wissen, welches den aktuellen Stand der Wissenschaft und der Technik beschreibt, adäquat beschreiben, was typischerweise in naturwissenschaftlichen Modellierungen gebraucht wird, und (b) sie weist einzigartige konzeptionelle Vorteile auf, welche unten benannt werden. Die *Dichte-Verhältnis-Klasse* ist, nebst ihren vorteilhaften Eigenschaften, für ihre nicht einfache Erhebung bekannt.

Um letzteres Problem anzugehen, entwickelten wir, erstens, eine Konstruktionsmethode für *Dichte-Verhältnis-Klassen*, welche es den Fachexperten erlaubt, Intervalle für Quantile oder Intervalle für Wahrscheinlichkeiten anzugeben. Um die Anwendung und Praktizierbarkeit der Methode möglichst einfach zu gestalten, erweiterten wir zu diesem Zweck etablierte Erhebungsmethoden. Zweitens, um mehr Einsicht und eine quantitative Beschreibung der Ambiguität zu erlangen, führten wir allgemein formulierte Metriken ein, welche auf beliebige Wahrscheinlichkeitsverteilungsklassen anwendbar sind. Die Metriken messen die Ambiguität, relativ zu einem im Voraus gewählten Vertrauensniveau, in Bezug auf wichtige und spezifische Eigenschaften von Verteilungen wie zum Beispiel die Weite der Verteilung, die Verteilungsform oder die Position des Modalwertes. Drittens, wir zeigten, dass die *Dichte-Verhältnis-Klasse* (i) invariant unter Bayesianischer Parameterinferenz ist, (ii) invariant unter Marginalisierung ist, (iii) invariant unter Propagation durch ein deterministisches Modell ist, und (iv) falls durch ein stochastisches Modell propagiert, wiederum in einer *Dichte-Verhältnis-Klasse* eingebettet ist, welche auch grösser als die propagierte Ur-*Dichte-Verhältnis-Klasse* sein kann. Die Invarianzeigenschaften machen die *Dichte-Verhältnis-Klasse* konzeptionell einzigartig und erlauben konsistente, sequenziell iterierbare Bayesianische Lernprozesse.

Wir machten auch einen Vorschlag, wie man die oben genannten Punkte numerisch implementiert und entwickelten hierfür ein generisch ausbaubares R-Software Packet (kostenlos erhältlich), welches (I) eindimensionale *Dichte-Verhältnis-Klassen* bestimmt, so wie es in der vorgeschlagen Erhebungsmethode gemacht wird, und welches (II) die vorgeschlagenen Metriken für *Dichte-Verhältnis-Klassen* berechnet. Schlussendlich wurden alle Implementationsschemen exemplarisch in einer Studie auf ein deterministisches Periphyton-Modell mit einem additiven stochastischen Fehler angewandt. Diese Studie illustriert die Auswirkungen von unpräzisem Vorwissen auf die Bayesianische Parameterschätzung und die Modellvorhersage.

Summary of the Ph.D. Thesis

The goal of decision theory is to support decision makers in finding alternatives that lead to the best possible fulfillment of their objectives. This requires (1) the quantitative valuation of possible outcomes as a function of measurable system attributes, (2) the prediction of probability distributions of these attributes for each alternative, and (3) the ranking of all alternatives by combining (1) and (2). To obtain such quantitative predictions it is often necessary to elicit knowledge about influence factors, model parameters or attribute values from subject matter experts. Unfortunately, the characterization of subjective degrees of belief, in the Bayesian context, can be ambiguous. We suggest the use of imprecise probabilities to describe ambiguity of uncertainty. In this robust Bayesian concept we evaluated diverse classes of probability distributions. We found the *Density Ratio Class* the most adequate: (a) being able to adequately represent intersubjective knowledge, describing the state-of-the-art knowledge of science and technology, which is typically needed in environmental modeling and (b) having unique conceptual properties that are discussed below. Apart from the advantageous class properties, the *Density Ratio Class* is known to be difficult to elicit.

To address this last point, we developed a method for constructing *Density Ratio Classes* based on intervals of quantiles or probabilities elicited from experts. To enhance the method's accommodation and practicability we extended established elicitation techniques to become applicable for this purpose. Second, to get deeper insight and a quantitative description of the ambiguity, we introduced generally formulated metrics, applicable to any type of class of probability distributions. The metrics measure, relative to a previously chosen credibility level, the ambiguity of important specific probability distribution attributes such as the width, the shape and the position of the mode. Third, we showed that the *Density Ratio Class* is (i) invariant under Bayesian updating, is (ii) invariant under marginalization, is (iii) invariant if propagated through a deterministic model, and is (iv) embedded again

into a *Density Ratio Class* that can be larger than the set of propagated distributions of the original class if the model is stochastic. These invariance properties make the class unique with regard to conception and allows for a consistent sequential Bayesian learning process.

We also made a proposition of how to numerically implement all the points mentioned before and developed a generically extendable R software package (freely available) that (I) numerically fits ready-to-use one-dimensional *Density Ratio Classes* according to the proposed elicitation method and (II) calculates the proposed metrics for *Density Ratio Classes*. Finally, we illustrated the steps required for considering imprecision by an exemplary application to a simple, deterministic periphyton model with an additive stochastic error term. This demonstrates the effect of imprecise prior knowledge on parameter estimates and model predictions.

Contents

Chapter 1

Introduction

1.1 Rational Decision Analysis and Support

In rational decision support it is fundamental to analytically structure the decision problem by clearly separating objectives, a decision maker or a stakeholder group would like to achieve, from predictions of the outcomes of the decision alternatives. The overall objective the decision maker(s) would like to achieve is hierarchically divided into complementary and more concrete sub-objectives. The degree of fulfillment of the objectives is then quantified by a value or utility function as a function of measurable attributes. The value function specifies directly the degree of fulfillment of the objectives on a scale between zero and unity; the utility function additionally considers risk attitudes and has to be elicited by asking preferences between lotteries of outcomes rather than certain outcomes. The hierarchical structuring of the objectives helps constructing such a function.

Rational decision support requires the combination of (1) the prediction of probability distributions of outcomes and (2) the elicitation of a utility function as a function of the outcomes. Alternatives are then ranked according to expected utilities (Von Neumann and Morgenstern 1944; Savage 1954; Berger 1985; von Winterfeldt and Edwards 1986; Clemen and Reilly 2001; French and Ríos Insua 2000; Eisenführ and Weber 2003). These theoretical concepts are also successfully applied in practice.

It is well known that individual and group behavior may violate rational choice theory (Allais 1953; Ellsberg 1961), and alternative behavioral theories have been suggested (Kahneman and Tversky 1979). Nevertheless, for supporting decisions that must be justifiable e.g. to authorities or the society, the framework of rational decision making can be very useful.

1.2 The Mathematical Framework

In model-based decision support, especially with regard to environmental systems, a model can only approximate the complex real system and thus leads always to uncertain predictions. Therefore, uncertainty has to be rigorously evaluated and carefully taken into account. Sources of uncertainty include: (i) non-deterministic, potentially stochastic behavior of the true system - referred to as *aleatory uncertainty* and (ii) lack of knowledge about the true system, its mathematical representation, and specific parameter values - referred to as *epistemic uncertainty* (Parry 1996; Walker et al. 2003; Refsgaard et al. 2007). Probability theory has long been the well-accepted framework for describing aleatory uncertainty. However, epistemic uncertainty is often the dominant source of uncertainty in environmental modeling (Ayyub and Klir 2006). It was shown by Keynes (1921), de Finetti (1931), Ramsey (1931), Cox (1946) and others that probability theory is also appropriate for describing epistemic uncertainty. We argue, independently of whether uncertainty is *aleatory*, or *epistemic*, that probability is the natural mathematical construct for describing uncertainty (Rinderknecht et al. 2012, 3). The motivation to treat aleatory and epistemic uncertainty in one and the same mathematical framework becomes even stronger considering the fact that aleatory uncertainty can turn into epistemic uncertainty. This is for instance the case if a random event has taken place but its outcome is not yet observed. In this situation, having a unique mathematical framework for both kinds of uncertainty also avoids unnecessary problems of inconsistency between mathematical formalisms. Furthermore, if individuals quantify their subjective degrees of belief with the aid of bets on lotteries between which they are indifferent, and if their beliefs are consistent in the sense of avoiding sure loss, then the resulting knowledge quantifications are consistent with the axiomatic foundation of probability theory (Box and Tiao 1973; de Finetti 1974; Howson and Urbach 1989; Seidenfeld et al. 1995; Kadane and O'Hagan 1995; Kadane et al. 1996). The logical framework for parameter inference and model prediction, consistent with the reasonings above, is then Bayesian statistics (Rinderknecht et al. 2011; Rinderknecht et al. 2012; Rinderknecht et al. 2011, 2.1 3.1 4.1).

1.3 Ambiguity: Reason for Robustification

Uncertain quantities needed for predictive modeling and decision support are often elicited from subject matter experts. It is common to use single, precise probability distributions for this purpose although the trustworthiness is questionable. Inaccuracies in elicitation procedures, misrepresentation of elicitation results, problems in expressing an individual's beliefs quantitatively, different perception of information by different individuals, or disagreement between experts lead to so called *ambiguity* in the probabilistic quantification of knowledge (Ellsberg 1961; Frisch and Baron 1988; O'Hagan and Oakley 2004; Rinderknecht et al. 2012, 3). Furthermore, the relevance of ambiguity becomes even stronger if (potentially ambiguous) *intersubjective knowledge*, characterizing the current state of knowledge of the scientific community (Gillies 1991; Rinderknecht et al. 2012, 3), shall be specified. We are interested to identify, describe, and seek to reduce ambiguity as a different aspect of uncertainty apart from the probabilistically quantified uncertainty. Several methods were proposed in the past for separating ambiguity in the choice of a probability specification from the uncertainty contained within the specification itself. Amongst others, the use of second-order probabilities or hierarchical probability models (Draper 1995) is a possible attempt to address this problem. However, having in mind that second-order uncertainty results e.g. from the problem that an expert has to express her or his beliefs precisely in the form of a probability distribution, it does not seem realistic that the knowledge about the second-order distribution (which is an even more abstract concept) can then be expressed precisely (Rinderknecht et al. 2012, 3). Our preferred method to characterize ambiguity is to replace the precise single probability distribution with a set of distributions. We developed three metrics of imprecision, relative to a credibility level, with regard to specific attributes of a distribution such as the width, shape and the mode, applicable to any type of probability set definition to enhance assessment and communication of ambiguity in quantified knowledge (Rinderknecht et al. 2012, 3.3). The set-based robustness concept is an extension of the conventional probability theory and the literature refers to as imprecise probability theory (Walley 1991; Caselton and Luo 1992, http://www.sipta.org). In the context of imprecise probability theory, conventional Bayesian statistics extend to what is called *robust Bayesian statistics* (Ríos Insua and Ruggeri 2000; Berger 1994).

1.4 An Overview of Selected Classes of Probability Distributions and its Definitions

As discussed above, we prefer the replacement of a single precise probability distribution by a set of probability distributions - if ambiguity is to be described. Different concepts have been proposed for specifying nonparametric sets of probability distributions. In Rinderknecht et al. (2011, 2) we evaluated five common classes of probability distributions. Since these classes are not presented in detail in the following chapters, we briefly review their definitions. These classes, for reasons of simplicity presented in their one-dimensional form, are the *Probability Box*, $\Gamma^{PB}_{L,U}$; the *Quantile Class*, $\Gamma^{QC}_{l,u}$; the ϵ-*Contamination Class*, Γ^{EC}_{ϵ}; the *Density Bounded*

Class, $\Gamma^{DB}_{l,u}$; and last but not least the *Density Ratio Class*, $\Gamma^{DR}_{l,u}$. These classes are graphically illustrated in Figure 1.4. Note, all of these classes can be used for improper probability densities in addition to proper densities. As we think that there is no reasonable behavioral interpretation of improper densities, we will focus on class specifications that only contain proper densities.

1.4.1 The Probability Box

Suppose L and U are cumulative distribution functions (CDFs) and $L(\theta) \leq U(\theta)$ for all θ in the parameter space. The set $\Gamma^{PB}_{L,U}$ of CDFs F

$$\Gamma^{PB}_{L,U} = \{\text{CDFs } F : L(\theta) \leq F(\theta) \leq U(\theta) \ \forall \theta\} \tag{1.1}$$

is called a *Probability Box* or *p-box.* If the random variable Θ belongs to the class, $L(\theta)$ and $U(\theta)$ are the lower and upper bounds for the probability of $\Theta \leq \theta$. This gives the defining functions L and U a particularly simple interpretation. The *Probability Box* is discussed more extensively in Ferson et al. (2003) and references cited therein.

1.4.2 The Quantile Class

Let the parameter space be partitioned into m disjoint pieces $I_1 \cup ... \cup I_m$. For $i \in \{1, ..., m\}$, let L_i and U_i satisfy $0 \leq L_i \leq U_i$ and $\sum_i^m L_i \leq 1 \leq \sum_i^m U_i$. Then the *Quantile Class* is defined as the set $\Gamma^{QC}_{L,U}$ of probability densities (PDFs) f

$$\Gamma^{QC}_{L,U} = \{\text{PDFs } f : L_i \leq \int_{I_i} f(\theta) \, \mathrm{d}\theta \leq U_i \quad \forall i \in \{1, ..., m\}\}. \tag{1.2}$$

In words, *Quantile Classes* are defined by placing upper an lower bounds on the probability that a parameter value θ lies within each of a finite number of subsets. The *Quantile Class* is discussed more extensively in Lavine (1991a) and Moreno and Pericchi (1993b).

1.4.3 The ϵ-Contamination Class

For a fixed $\epsilon \in [0, 1]$, a fixed PDF f_0 and a given class G of PDFs the ϵ-*Contaminated Class* $\Gamma^{EC}_{f_0,G,\epsilon}$ is defined as:

$$\Gamma^{EC}_{f_0,G,\epsilon} = \{f = (1 - \epsilon) f_0 + \epsilon g : g \in G\}. \tag{1.3}$$

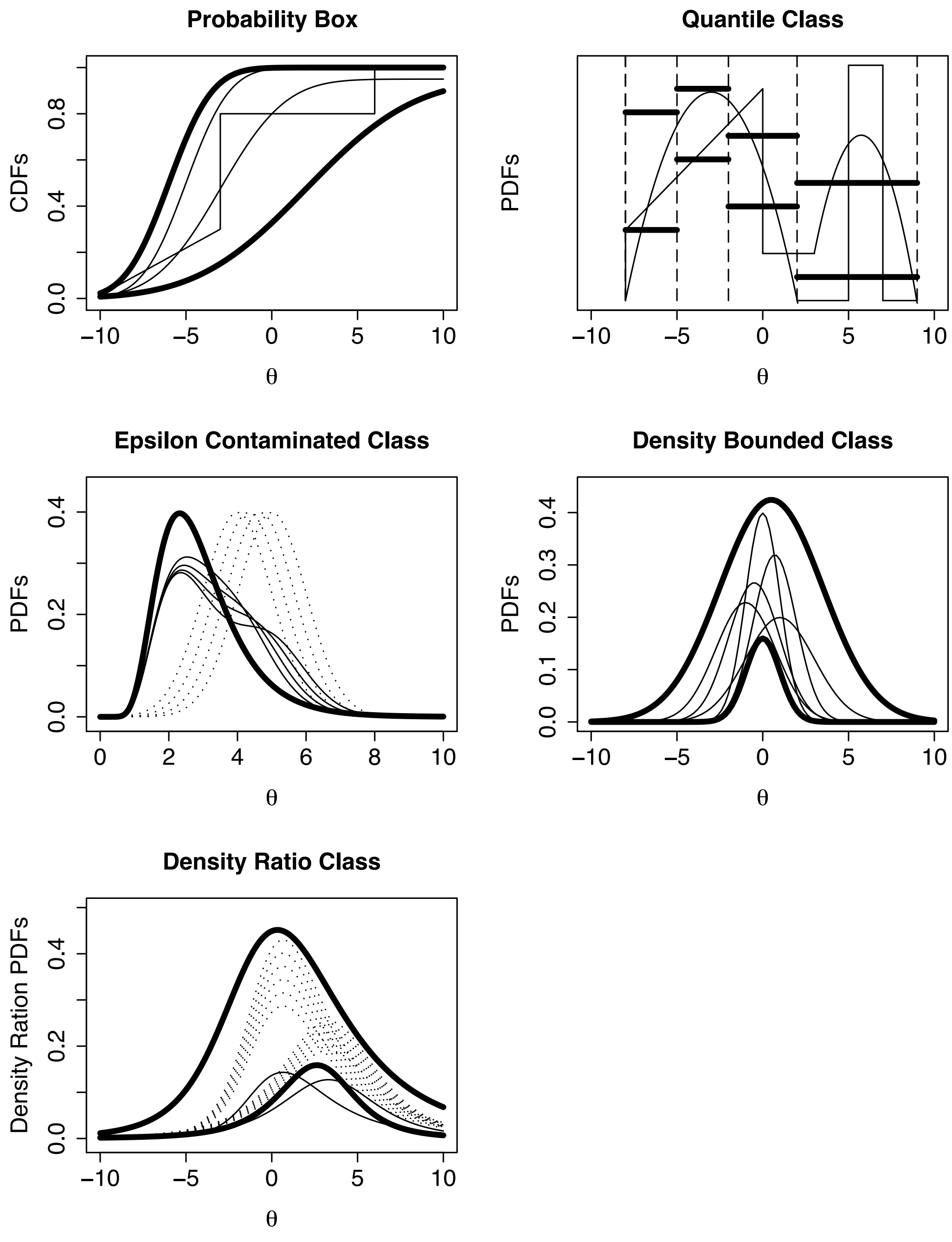

Figure 1.1: Graphical illustrations of the five classes of distributions: *Probability Box*, $\Gamma^{PB}_{L,U}$; *Quantile Class*, $\Gamma^{QC}_{L,U}$; *ϵ-Contamination Class*, $\Gamma^{EC}_{G,\epsilon}$; *Density Bounded Class*, $\Gamma^{DB}_{l,u}$; *Density Ratio Class*, $\Gamma^{DR}_{l,u}$. Thick solid lines indicate functions that define the class, thin solid lines represent examples of probability densities (PDF) or cumulative distribution functions (CDF) belonging to the class, and dotted lines represent contamination distributions for Γ^{EC}_{ϵ} and unnormalized densities for $\Gamma^{DR}_{l,u}$.

The fixed PDF f_0 is interpreted as the base distribution, ϵ is the assessed amount of uncertainty in the base distribution, and G is the class of contaminations considered. The *ϵ-Contaminated Class* was introduced by Huber (1973) and has been studied by Berger and Berliner (1986), Berger (1985), Berger (1990), Pericchi and Walley (1991) and Walley (1991).

1.4.4 The Density Bounded Class

The *Density Bounded Class* is defined as the set $\Gamma^{DB}_{l,u}$ of PDFs f:

$$\Gamma^{DB}_{l,u} = \{\text{PDFs } f : l(\theta) \leq f(\theta) \leq u(\theta);\ \forall \theta\} \tag{1.4}$$

where l and u are two bounded non-negative functions, such that $l(\theta) \leq u(\theta)$ for all θ, $\int u(\theta)d\theta \leq 1$ and $1 \leq \int u(\theta)d\theta < \infty$. In words, a density bounded set consists of all normalized densities between upper and lower bounds. This class is discussed in Lavine (1991a) and Wasserman and Kadane (1992).

1.4.5 The Density Ratio Class

For uncertain continuous parameters $\boldsymbol{\theta} \in M \subset \mathbb{R}^n$, the *Density Ratio Class*, $\Gamma^{DR}_{l,u}$, is defined as the set of probability density functions

(1.5)

$$\Gamma^{\text{DR}}_{l,u} := \left\{ \hat{f}(\boldsymbol{\theta}) \ = \frac{f(\boldsymbol{\theta})}{\int f(\boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'} \ \middle|\ l(\boldsymbol{\theta}) \leq f(\boldsymbol{\theta}) \leq u(\boldsymbol{\theta})\ \ \forall \boldsymbol{\theta}\,, \int l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} > 0\,, \int u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} < \infty \right\},$$

where l and u are lower and upper bounded non-negative, non-normalized densities. These upper and lower densities bound the shapes of the not normalized probability densities in the class. If $l(\theta) > 0\ \forall\ \theta$, this definition is equivalent to

$$\Gamma^{DR}_{l,u} := \left\{ f(\theta) \ = \frac{g(\theta)}{\int g(\theta')\,\mathrm{d}\theta'} \ \middle|\ \frac{g(\theta)}{g(\theta')} \leq \frac{u(\theta)}{l(\theta')}\ \forall\ \theta, \theta' \right\}. \tag{1.6}$$

As $\Gamma^{DR}_{l,u} = \Gamma^{DR}_{\lambda l, \lambda u}$ for any $\lambda > 0$, we can normalize one of the bounds, l or u, respectively. In contrast to the previous class that provides bounds on densities, the focus of the density ratio class is on bounding shapes that then require normalization to become probability density functions. DeRobertis and Hartigan (1981) introduced the *Density Ratio Class* under the name of "intervals of measures"; Berger (1990) called the class *Density Ratio Class* as it bounds ratios of densities as clarified by the second definition above. Using the same shape for $l = f_0$ and $u = kf_0$ leads to the special case $\Gamma^{DR}_{f_0,k}$ that may be thought of as a neighborhood around the target prior f_0 (Wasserman 1992b).

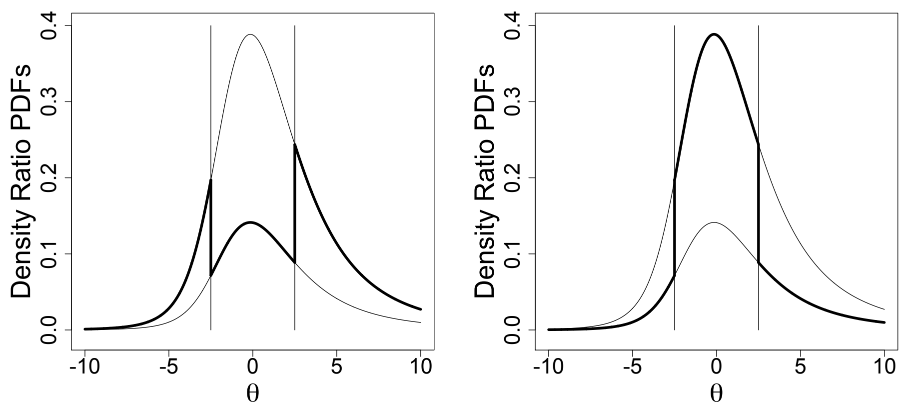

Figure 1.2: Lower and upper bounds (thin curves) and the not necessarily normalized densities (thick curves) of a *Density Ratio Class* leading to the lower, $\underline{P}(A)$ (left), and to the upper, $\overline{P}(A)$ (right), probability of the value to be in the interval $A = [-2.5, 2.5]$.

The lower and upper probabilities, $\underline{P}$ and $\overline{P}$, for a random variable that is characterized by the *Density Ratio Class*, $\Gamma_{l,u}^{\mathrm{DR}}$, to take a value within a subset A of its domain are given by

(1.7)
$$\underline{P}(A) = \frac{\int_A l \; d\theta}{\int_A l \; d\theta + \int_{A^c} u \; d\theta}$$

and

(1.8)
$$\overline{P}(A) = \frac{\int_A u \; d\theta}{\int_A u \; d\theta + \int_{A^c} l \; d\theta}$$

where A^c is the complement of A (see Figure 1.4.5). The proof for these equations is given in 4.2.1.

1.5 Motivation for the Density Ratio Class

We believe that there are at least five desirable properties of classes of probability distributions: (1) variety of shapes and absence of 'unreasonable' shapes, (2) invariance under Bayesian updating, (3) tractability of Bayesian updating, (4) ease of assessment, and (5) invariance under marginalization. In Rinderknecht et al. (2011, 2.2.3) we discuss in detail how well each of the non parametric classes, described above in 1.4, fulfils these properties and conclude that the *Density Ratio Class*, $\Gamma_{l,u}^{DR}$, has clear conceptual advantages over the other classes. In particular, its invariance under updating and marginalization makes it the only class that can be used to describe a consistent sequential Bayesian learning process.

A sixth desirable property of a class of probability distributions is the (6) invariance if propagated through a model. In Rinderknecht et al. (2011, 4) we study this explicitly with regard to predictions of models with parameters that have the form of a *Density Ratio Class* and show also how this can be numerically implemented. We have found, under mild regularity conditions, that the *Density Ratio Class* is invariant if propagated through a deterministic model. Predictions that are based on stochastic models with model parameters defined by a *Density Ratio Class* are naturally embedded into a *Density Ratio Class* that can be larger than the set of propagated distributions.

Hence, we evaluate the *Density Ratio Class* to be the most versatile with regard to conception. The main disadvantage has been found to be the practical difficulty of its assessment — which invited us to develop a practicable method that fits *Density Ratio Classes* to elicited quantile or probability intervals (Rinderknecht et al. 2011, 2).

1.6 The Structure of this Book

Chapter 2 contains the published paper Rinderknecht et al. (2011). We propose an elicitation method for the *Density Ratio Class* that makes use of already known and established elicitation methods - as far as this was possible. To facilitate the elicitation of the *Density Ratio Class*, a software package was written and made freely available for the R statistical programming environment. One may download the package `fitDRC.R` on http://cran.r-project.org/ subject to the terms of agreement.

Chapter 3 contains the published paper (Rinderknecht et al. 2012). We make an argument for using the mathematical concept referred to as imprecise probabilities to represent epistemic, subjective or intersubjective knowledge, as possibly revealed through a process of expert elicitation. We suggest to replace a precise probability distribution with a set of distributions allowing for a continuous degree of imprecision, such that the concept can form a bridge between total ignorance and precisely characterized risk. Three metrics of imprecision, applicable to any type of probability set definition, are developed to enhance assessment and communication of ambiguity in quantified knowledge. Specific forms of the metrics are derived for the *Density Ratio Class*. Three brief case studies are used to illustrate application of the imprecise probability concept and the three developed metrics to represent expert opinion.

Chapter 4 contains the paper (Rinderknecht et al. 2011) that is ready for submission. We mathematically show how Bayesian inference, marginalization and model predictions are made with *Density Ratio Class* priors (or posteriors). We also show the corresponding numerical implementation schemes. We apply our findings to a simple deterministic periphyton model that has an additive stochastic error in order to study the effects of imprecise prior knowledge on parameter estimates and predictions.

Finally, in Chapter 5 we draw our conclusions and give a brief outlook in the form of further research questions.

Chapter 2

Eliciting Density Ratio Classes

Simon L. Rinderknecht[1,2], Mark E. Borsuk[3] and Peter Reichert[1,2]

Published in the International Journal of Approximate Reasoning (2011), 52 (6), 792-804. doi:10.1016/j.ijar.2011.02.002

[1] Eawag, Swiss Federal Institute of Aquatic Science and Technology, CH-8600 Dübendorf, Switzerland.
[2] ETH Zürich, Institute of Biogeochemistry and Pollutant Dynamics (IBP), CH-8092 Zürich, Switzerland.
[3] Dartmouth College, Thayer School of Engineering, Hanover, NH 03755-8000, USA.

Abstract

The probability distributions of uncertain quantities needed for predictive modeling and decision support are often elicited from subject matter experts. However, experts are often uncertain about quantifying their beliefs using precise probability distributions. Therefore, it seems natural to describe their uncertain beliefs using sets of probability distributions. There are various possible structures, or *classes*, for defining set membership of continuous random variables. The *Density Ratio Class* has desirable properties, but there is no established procedure for eliciting this class. Thus, we propose a method for constructing *Density Ratio Classes* that builds on conventional quantile or probability elicitation, but allows the expert to state intervals for these quantities. Parametric shape functions, ideally also suggested by the expert, are then used to bound the nonparametric set of shapes of densities that belong to the class and are compatible with the stated intervals. This leads to a natural metric for the size of the class based on the ratio of the total areas under

upper and lower bounding shape functions. This ratio will be determined by the characteristics of the shape functions, the scatter of the elicited values, and the explicit expert imprecision, as characterized by the width of the stated intervals. We provide some examples, both didactic and real, and conclude with recommendations for the further development and application of the *Density Ratio Class*.

Keywords

Probability assessment; probability elicitation; expert elicitation; elicitation of vague knowledge; subjective probabilities; imprecise probabilities; decision theory; robust Bayesian statistics; quantile elicitation; *Density Ratio Class*.

2.1 Introduction

Rational choice theory requires consideration of the probabilities of all possible outcomes of decision alternatives, in addition to valuations of these outcomes, in order to provide decision makers with a ranking of alternatives based on expected utilities (Von Neumann and Morgenstern 1944; Savage 1954; Berger 1985; von Winterfeldt and Edwards 1986; Clemen and Reilly 2001; French and Ríos Insua 2000; Eisenführ and Weber 2003). In many contexts, the uncertainty of outcomes is not dominated by aleatory uncertainty, due to randomness of a system, but by epistemic uncertainty, due to lack of precise knowledge of system behavior. In such cases, scientific knowledge is often elicited from experts in the form of their subjective degrees of belief in the outcomes. If we decide to quantify such subjective degrees of belief with the aid of bets on lotteries between which a person is indifferent, and if we require these beliefs to be consistent in the sense of avoiding sure loss, then the mathematical construct for describing and updating such beliefs must be Bayesian probability (de Finetti 1974; Howson and Urbach 1989; Seidenfeld et al. 1995; Kadane and O'Hagan 1995; Kadane et al. 1996). It is well known that individual and group behavior may violate rational choice theory (Allais 1953; Ellsberg 1961), and alternative behavioral theories have been suggested (Kahneman and Tversky 1979). Nevertheless this framework can still be useful for prescribing the transparent use of scientific knowledge in support of societal decisions (Reichert et al. 2007; Corsair et al. 2009). This makes elicited probability distributions an essential element of societal decision support and makes it important to develop techniques to carefully use and improve this instrument.

A significant problem in the application of rational choice theory is that in many cases probability distributions characterizing knowledge and beliefs are themselves uncertain. This may also be the cause of some of the apparent violations of rationality mentioned above (Ellsberg 1961). The problem of uncertainty in probability distributions is often especially relevant for priors in Bayesian inference procedures (Berger 1985). For this reason, we are interested in the robustness of the results to modifications in subjective uncertainties or beliefs. This can be done systematically by employing a set of distributions that contains all those considered to be feasible. Depending on the degree of ambiguity about a particular distribution, such a set can contain a large variety of shapes or can simply contain those shapes in the neighborhood of a given distribution. This concept has been addressed under the topic of "imprecise probability" (Walley 1991; Caselton and Luo 1992, http://www.sipta.org). It can also be seen as generalizing "robust Bayesian statistics" (Berger 1994; Ríos Insua and Ruggeri 2000) but it has been applied in other fields as well (Huber and Strassen 1973; Cattaneo 2008; Fierens et al. 2009; Maturi et al. 2010).

As it is not evident how to extend interval probabilities of discrete random variables to continuous random variables, we focus in this paper on the latter case that is of high practical relevance. A variety of structures or classes have been proposed to characterize sets of distributions of continuous random variables. In Section 2.2, we suggest that the *Density Ratio Class* is, from a conceptual point of view, the most satisfying approach. However, this class is not currently well-developed with respect to methods of expert elicitation. Therefore, in section 2.3 we propose a practical procedure for eliciting this class. In section 2.4, we illustrate our procedure with didactical examples and with an application to published elicitation results to demonstrate the feasibility of the suggested approach. Finally, we discuss our findings in section 2.5.

2.2 The Density Ratio Class

2.2.1 Definition

For continuous uncertain quantities, the *Density Ratio Class* of probability density functions (PDFs), f ($f \geq 0$, $\int f(\theta)d\theta = 1$), is defined as the set

$$\Gamma^{\mathrm{DR}}_{l,u} = \{\text{PDFs } f : \exists c : l(\theta) \leq cf(\theta) \leq u(\theta)\ \forall\theta\} \tag{2.1}$$

where l and u are two bounded nonnegative functions (non-normalized densities) such that $l(\theta) \leq u(\theta)$ for all θ in the domain of the random variable and $\int u(\theta)d\theta < \infty$. We limit the class to proper densities (with $\int f(\theta)d\theta = 1$) as the behavioral interpretation of improper densities is questionable. If $l(\theta) > 0\ \forall\ \theta$, this definition is equivalent to:

$$\Gamma^{\mathrm{DR}}_{l,u} = \left\{ \text{PDFs } f : \frac{f(\theta)}{f(\theta')} \leq \frac{u(\theta)}{l(\theta')} \ \forall\, \theta, \theta' \right\} \tag{2.2}$$

The focus of the *Density Ratio Class* is on bounding shapes that then require normalization to become probability density functions. DeRobertis and Hartigan (1981) introduced the *Density Ratio Class* under the name of "intervals of measures"; Berger (1990) called the class the *Density Ratio Class*, as it bounds ratios of densities, as is clear from the second definition above.

As $\Gamma^{\mathrm{DR}}_{l,u} = \Gamma^{DR}_{\alpha l, \alpha u}$ for any $\alpha > 0$, we can normalize one of the bounds, l or u. With $f_l = l / \int l(\theta')d\theta'$ and $f_u = u / \int u(\theta')d\theta'$ we can thus write

$$\Gamma^{\mathrm{DR}}_{l,u} = \Gamma^{\mathrm{DR}}_{f_l, \kappa f_u} \tag{2.3}$$

where the ratio

$$\kappa = \frac{\int u(\theta)d\theta}{\int l(\theta)d\theta} \geq 1 \tag{2.4}$$

can be interpreted as a measure of the "size" of the class. For continuous densities, $\kappa = 1$ is only possible if $u = l$, and the class then reduces to the precise density $f = f_l = f_u$. The larger κ, the larger is the variety of possible shapes. Using the same shape for $l = f_0$ and $u = \kappa f_0$ leads to the special case $\Gamma^{DR}_{f_0,\kappa}$ that may be thought of as a neighborhood around the target prior f_0 (Wasserman 1992b).

2.2.2 Bounds on Probabilities

Following from equation (2.1), the lower and upper probabilities, $\underline{P}$ and $\overline{P}$, for a random variable characterized by the *Density Ratio Class*, $\Gamma^{\mathrm{DR}}_{l,u}$, to take a value within a subset A of its domain are given by

$$\underline{P}(A) = \frac{\int_A l \ d\theta}{\int_A l \ d\theta + \int_{A^c} u \ d\theta} \tag{2.5}$$

and

$$\overline{P}(A) = \frac{\int_A u \ d\theta}{\int_A u \ d\theta + \int_{A^c} l \ d\theta} \tag{2.6}$$

where A^c is the complement of A.

As a special case of equations (2.5) and (2.6), we can calculate the lower and upper cumulative probabilities $P(\Theta < \theta)$, $\underline{F}_{l,u}(\theta)$ and $\overline{F}_{l,u}(\theta)$, respectively. Using the *Density Ratio Class* in the form $l = f_l$, $u = \kappa f_u$ (2.3), we get

$$\underline{F}_{f_l,\kappa f_u}(\theta) = \frac{F_l(\theta)}{F_l(\theta) + \kappa(1 - F_u(\theta))} \tag{2.7}$$

and

$$\overline{F}_{f_l,\kappa f_u}(\theta) = \frac{\kappa F_u(\theta)}{\kappa F_u(\theta) + (1 - F_l(\theta))} \tag{2.8}$$

where F_l and F_u denote the cumulative distribution functions of the normalized lower and upper densities, f_l and f_u, respectively. Note that $\underline{F}$ and $\overline{F}$ represent pointwise bounds of the cumulative distribution functions but are not in general cumulative distribution functions of a class member.

2.2.3 Properties

We do not intend to suggest that the *Density Ratio Class* uniquely represents the uncertain beliefs of an expert concerning the value of a quantity. However, we think that it has some properties that give it advantages over other classes used for this purpose. In particular, in this section we provide a comparison based on four desirable properties: (1) the ability to accommodate a variety of shapes and exclude "unreasonable" shapes, (2) invariance under Bayesian updating, (3) tractability of Bayesian updating and posterior expectations, and (4) ease of assessment. The other classes considered in this comparison are the *ϵ-Contamination Class* (Huber 1973), the *Probability Box* (Williamson and Downs 1990), the *Quantile Class* (Lavine 1991a; Moreno and Pericchi 1993b), and the *Density Bounded Class* (Lavine 1991b) (Berger 1994, gives an overview of imprecise probability classes). There may be other properties we do not focus on, such as invariance under scale transformations or other nonlinear mappings, that are important in certain applications.

Variety of Shapes and Exclusion of "Unreasonable" Shapes

Pericchi and Walley (1991) distinguish between a "reasonable class of prior densities" and a "class of reasonable prior densities." We believe that a class is reasonable if and only if it contains reasonable shapes. That is, it should contain a wide variety of shapes to provide sufficient degrees of freedom for representing ambiguity, but should not allow for highly aberrant shapes that would not be deemed reasonable by the expert.

To fulfill the criterion of containing a sufficient variety of shapes, we focus on non-parametric classes that contain at least a neighborhood of "interior" densities of the class. Sets for which the members are parametrized, or *Parametric Classes*, seem to be too limited with respect to such neighborhoods. In the case of the *Density Ratio Class* the lower and upper parameterized shapes, l and u, are only used to bound the non-parametric set of shapes.

Another important consideration is that the class allows a variety of tail behaviors, as such behaviors often fall outside the realm of past experience. By choosing an upper function, u, that has heavier tails than the lower function, l, the *Density Ratio Class* can contain densities that exhibit different tail behavior (Pericchi and Walley 1991). The ϵ-*Contamination Class* can also readily admit a range of different tail behaviors if the set of contaminating distributions is chosen to be large enough (e.g. if this set is the set of all probability distributions). However, with this choice, the chance of including unreasonable shapes is also large, especially for large values of ϵ (Borsuk and Tomassini 2005; Berger 1994; Walley 1991; Moreno and Cano 1991; Moreno and Pericchi 1993a).

We believe that for most elicited distributions of continuous quantities, point masses or extremely high peaks of densities may often not well represent the experts' opinions. In contrast to the *Probability Box* and the *Quantile Class* (Pericchi 1998), the *Density Ratio Class* excludes such probability distributions. It is worth noting, however, that the *Density Ratio Class* still allows multiple local maxima and minima that may not strictly represent the views of the expert. This situation is a common feature of non-parametric classes unless additional constraints are imposed. The ϵ-Contamination Class offers the opportunity of a specific choice of deviations from a reference distribution. This makes the choice of reasonable shapes more explicit.

Invariance under Bayesian Updating

A class is invariant under updating if the set of posteriors after Bayesian updating of a prior class (this means updating all members of the prior class) is again represented by a member of the same class (DeRobertis and Hartigan 1981; Wasserman 1992b). This property is required to make sequential updating possible within the same framework as more data become available. The *Density Ratio Class* is invariant under updating as well as marginalization. Wasserman (1992a) showed that under mild regularity conditions, this is the only class with these two properties. This makes it possible not only to use this class for sequential updating, but also to conveniently demonstrate the effect of the learning process on marginals.

Tractability of Bayesian Updating

Tractability of updating is the ease of calculating the posterior class and bounds of posterior expectations and other posterior quantities given the prior class and observed data. The invariance properties of the *Density Ratio Class* under Bayesian updating and marginalization provide the basis for tractable numerical implementation with this class. Specifically, Wasserman and Kadane (1992) showed that it is possible to derive numerical approximations to bounds of posterior expectations from a random sample of a single member of the *Density Ratio Class* under the special case with $f_l = f_u = f_0$, $\Gamma^{DR}_{f_0,\kappa}$. We are currently preparing a paper addressing the general case (Rinderknecht et al. 2011).

Kriegler and Held (2005) conclude that *Probability Boxes* fulfill the tractability criterion to some extent. Methods for calculating bounds on expectations for this class are given by Basu and DasGupta (1995). The *Quantile Class* is not sensible in higher dimensions unless shape constraints are introduced which make the model less tractable for updating (Pericchi 1998). However, it is relatively easy to calculate bounds of expectations for this class (Lavine 1991a). The ϵ-*Contamination Class* is popular because it is very tractable (Huber 1973; Berger and Berliner 1986; Sivaganesan and Berger 1989). It is more difficult to find posterior bounds for the *Density Bounded Class* than for the *Quantile Class* or the ϵ-*Contaminated Class* (Lavine 1991a).

Ease of Assessment

A class should be as easy to assess and interpret as possible (Berger 1994). Assessment in this context is the process of constructing the class by formalizing prior knowledge, often by expert elicitation. As experts generally recognize the ambiguity contained in their own beliefs, ranges of probabilities or quantiles are often as easy to elicit as exact values (Berger 1990). The problem is how to construct classes from such assessments.

As distributions are often elicited by asking the experts for quantiles or cumulative probabilities, the *Probability Box* and the *Quantile Class* are particularly simple to elicit by simply asking for ranges for these quantities that are consistent with the expert's beliefs. The elicitation of an ϵ-*Contamination Class* can consist of a conventional elicitation of the reference distribution f_0, followed by the class of contaminations G and the choice of the value of ϵ. Often, the contamination class G will be chosen to consist of all densities. In such a case, the elicitation is not much different from a conventional elicitation process for precise distributions. Therefore, the ϵ-*Contamination Class* is easy to work with and has a rather intuitively appealing interpretation (Borsuk and Tomassini 2005). Elicitation and interpretation of the *Density Bounded Class* and of the *Density Ratio Class* are more difficult. The main reason is that the bounds, l and u, do not have a similarly simple interpretation as for the defining quantities of the other classes. To our knowledge, no elicitation procedures have been published for these two classes.

Summary of Properties

The *Density Ratio Class* has clear conceptual advantages over the other classes. In particular, its invariance under updating and marginalization makes it the only class that can be used to describe a consistent sequential learning process. In addition, it excludes unreasonable shapes to a much better degree than the other classes. However, this comes with a requirement for additional information beyond simply elicited ranges of probabilities or quantiles. This leads to the primary disadvantage of the *Density Ratio Class*: the potential difficulty of its assessment. Consequently, the development of a practicable elicitation technique for the *Density Ratio Class* is the topic of this paper.

2.3 Eliciting Density Ratio Classes

It is relatively easy to elicit probability or quantile intervals by simply extending existing elicitation procedures for precise distributions. However, as demonstrated by the *Probability Box* and the *Quantile Class*, defining a class using such intervals as the only constraint leads to the inclusion of distributions with shapes that are unreasonable in most applications. This is because very high peaks and even point masses are compatible with such intervals. To uniquely specify a class with more stringent restrictions, more information must be provided. This cannot be accomplished without making the elicitation process somewhat more involved or making additional assumptions.

We propose a relatively straightforward method for deriving the *Density Ratio Class*. For simplicity, we concentrate on a one-dimensional random variable and benefit as much as possible from established elicitation procedures for precise probability distributions. Hence, our suggested procedure starts with the elicitation of quantile or cumulative probability intervals that represent the expert's ambiguity. To these intervals is added the selection of a parametric shape of the bounding densities, or two shapes differing, for example, in their tail behavior. We then construct the smallest *Density Ratio Class* that is bounded by the specified parametric shapes and yet contains the elicited quantile or cumulative probability intervals. This outer approximation' can be seen as a conservative representation of the expert's opinion, as the quantile intervals within the *Density Ratio Class* may be considerably larger than those specified by the expert. By potentially extending, rather than contracting, the smallest set of densities consistent with the experts' assertions, this assumption is consistent with the understanding that experts tend to be overconfident. If there is a large number of experts, it might be useful to eliminate the most extreme views by using a specified quantile of the interval endpoints instead of the minimum or maximum. If an expert is able to dispense with intervals by asserting precise CDF points, then the method relies on the scatter of these points around the optimized parametric shape to define the degree of imprecision of the class. The reader interested in viewing graphical representations of some examples is encouraged to look ahead to Figures 2.1, 2.3, 2.4 and 2.6.

2.3.1 Elicitation of Probability and Quantile Intervals

Generally, expert elicitation of the distribution of a continuous random variable θ relies on the quantile elicitation method (Wallsten and Budescu 1983; Cooke 2001; Chaloner 1996; Kadane and Wolfson 1998; Garthwaite et al. 2005; Jenkinson 2005; O'Hagan et al. 2006). According to this method, the analyst provides cumulative probabilities, $\{p_i\}_{i=1}^n$, and the expert estimates the corresponding quantiles, $\{\theta_i\}_{i=1}^n$ (i.e., elicitation of the inverse CDF). This procedure, first suggested by Winkler (1967), minimizes anchoring effects (Tversky and Kahneman 1974; Kynn 2008) that may be inherent in the probability elicitation procedure in which the analyst provides values of the random variable and the expert estimates the corresponding cumulative probabilities. We believe it is good practice to start by asking the expert for an overall interval of the random variable for which the expert is convinced that it is highly unlikely that the value θ could lie outside. It is important to clarify the precise meaning of the specified overall interval, i.e., the estimated probability of the random variable to be outside the interval (Alpert and Raiffa 1982; Lichtenstein et al. 1982). The analyst should be aware that experts tend to be overconfident in assessing such an interval.

After asking for the overall interval of the random variable, we recommend to follow the bisection method (see Garthwaite and Dickey (1985) for further technical details). That is, ask the expert for the median, the lower quartile, the upper quartile and more quantiles if necessary. However, based on our experience with elicitation, we have found that experts recognize that they cannot make probability judgments with absolute precision. Therefore, they often feel more comfortable if they are allowed to express their assessments as ranges. This also seems to be a natural way of eliciting information to construct imprecise probability distributions. Hence, in our procedure, the experts are allowed to specify intervals but precise estimates could be used as well. The interpretation of the intervals is that the entire range of the interval is consistent with the scope of their beliefs.

The result of this step for elicited quantiles is

QI: *For probabilities* $\{p_i\}_{i=1}^n$ *selected by the analyst, lower and upper bounds on the quantile* $\{\theta_i^l \leq \theta_i^u\}_{i=1}^n$ *are estimated by the expert.*

For elicited cumulative probabilities it is:

PI: *For values of the random variable,* $\{\theta_i\}_{i=1}^m$ *given by the analyst, lower and upper bounds on the cumulative probabilities* $\{p_i^l \leq p_i^u\}_{i=1}^m$ *are estimated by the expert.*

All endpoints of cumulative probability or quantile intervals can be summarized in a set of probability-quantile pairs, $\{p_i, \theta_i\}_{i=1}^{2(n+m)}$. This set typically contains only quantile intervals ($n > 0, m = 0$) or probability intervals ($n = 0, m > 0$); however, it can also contain intervals of both types ($n > 0, m > 0$) or precise point estimates.

According to our procedure, these intervals are only a partial representation of the expert's beliefs. The determination of parametric shapes is discussed next.

2.3.2 Choice of Parametric Shapes

A set of cumulative probability or quantile intervals does not adequately constrain the probability distributions that represent an expert's knowledge. For this reason, we next work with the expert to identify one or more parametric families of distributions, the shape of which can approximately represent his or her knowledge. Clearly, the parametric families of distributions should be compatible with the probability-quantile intervals elicited from the expert in the sense that they can be fitted to pass approximately through the quantile intervals. However, at this stage, only the allowable shapes, as represented by the parametric families of normalized densities f, are selected, not particular parameter values. Nevertheless, it is important to recognize that the choice of such a family of densities implies many assumptions about the distribution that the expert did not initially assert with the assessed intervals. Therefore, choosing the parametric families to represent the elicited data should be combined with a rather detailed discussion of what the expert deems to be reasonable and unreasonable shapes (see section 2.2.3). It should be kept in mind, however, that these families will only be used to define the bounding densities, not all members of the set, as we describe below.

There is no general rule for the choice of parametric families. However, the procedure is similar to selecting (or confirming) a parametric family for fitting a precise probability density to elicited points of a CDF (O'Hagan et al. 2006). Important properties are uni- versus multi-modality, skewness and bounds. Particular emphasis should be on tail behavior, as this is particularly difficult to capture with a small, discrete set of probability-quantile intervals. Three important cases to distinguish with respect to the domain of the elicited continuous random variable are (O'Hagan et al. 2006):

Unrestricted random variable: For a random variable Θ which can take any value, positive or negative, the most frequently used families are the *normal distribution* or the *Student t distribution*. The *t-distribution* is important in particular because it allows for heavier tails than the *Normal distribution*.

Random variable bounded on one side: For a random variable Θ which is bounded on one side, suitably shifted distributions of positive variables such as the *exponential*, *log-normal*, *gamma*, *inverse-gamma*, *chi-squared*, *Weibull* and the *F-distribution* are the most appropriate.

Bounded random variable: For a random variable Θ with bounded range, the most widely used distributions belong to the *beta family*. Other distributions used for bounded random variables are the bounded uniform, triangular and trapezoidal distributions.

If an expert is especially uncertain about the tail behavior of the distribution, then he or she should choose two families, f_1 and f_2, with different tail behavior. For an unrestricted random variable this means $f_1(\theta)/f_2(\theta) \to \infty$ as $|\theta| \to \infty$.

This elicitation step results in two families of parametric distributions representing the extremes of shapes (e.g. of tail behavior) compatible with the expert's beliefs. These two families will represent the normalized lower, $f_l(\theta, \psi_l)$, and the normalized upper, $f_u(\theta, \psi_u)$, densities partially characterizing the *Density Ratio Class*. Here, ψ_l and ψ_u are the parameters characterizing the densities of the parametric families. If their tail behavior is different, the one with heavier tails must be used as the upper normalized density, f_u.

If the expert is very confident in the set of reasonable shapes, the two families can be identical. If the expert is not confident at all in the choice of shapes, then multiple families can be used in various pairs to describe l and u, with parameters ψ_l and ψ_u and κ to be estimated empirically. The pair leading to the smallest class size, as measured by the value of κ, that accommodates the elicited intervals can then be maintained as the best description of the elicited *Density Ratio Class*. A method for estimating the values of ψ_l, ψ_u, and κ is discussed next.

2.3.3 Construction of a Density Ratio Class Based on Elicited Probability-Quantile Intervals and Parametric Shapes

From the elicitation steps described in sections 2.3.1 and 2.3.2 we get: (i) a discrete set of endpoints of cumulative probability or quantile intervals, $\{p_i, \theta_i\}_{i=1}^{2(n+m)}$, and (ii) two (or more) parametric families of distributions for the normalized lower, $f_l(\theta, \psi_l)$, and the normalized upper, $f_u(\theta, \psi_u)$, densities defining the shapes of the unnormalized densities, l and u that bound the *Density Ratio Class*.

Based on this information, we construct a *Density Ratio Class* as follows. We search for the set of parameter values (κ, ψ_l, ψ_u) that satisfies the following three conditions:

$$\begin{array}{ll} \text{(i)} & f_l(\theta,\psi_l) \le \kappa f_u(\theta,\psi_u)\ \forall\theta \\ \text{(ii)} & \underline{F}_{f_l,\kappa f_u}(\theta_i,\psi_l,\psi_u) \le p_i \text{ for } i=1,...,2(n+m) \\ & \overline{F}_{f_l,\kappa f_u}(\theta_i,\psi_l,\psi_u) \ge p_i \text{ for } i=1,...,2(n+m) \\ \text{(iii)} & \kappa \text{ should take the minimal value consistent with (i) and (ii)} \end{array} \tag{2.9}$$

(see equations 2.7 and 2.8 for the definitions of the functions $\underline{F}_{f_l,\kappa f_u}$ and $\overline{F}_{f_l,\kappa f_u}$). Condition (i) guarantees that the densities $l = f_l$ and $u = \kappa f_u$ correctly define a *Density Ratio Class*, condition (ii) guarantees consistency of the class with the elicited cumulative probability or quantile intervals in the sense of being an 'outer approximation'. Condition (iii) requires the class to be of minimum size compatible with conditions (i) and (ii) (see equation 2.4).

To facilitate the search for the parameter values fulfilling conditions (2.9), we derive a function that calculates the minimum value of κ fulfilling these conditions for arbitrary given values of the parameters ψ_l and ψ_u (see equation 2.11 below). The parameters ψ_l and ψ_u are then determined by numerically minimizing this function.

We first determine the smallest *Density Ratio Class* compatible with the two shapes $f_l(.,\psi_l)$ and $f_u(.,\psi_u)$ as a function of the first argument. For this purpose, we calculate

$$\lambda(\psi_l,\psi_u) = \sup_\theta \frac{f_l(\theta,\psi_l)}{f_u(\theta,\psi_u)} \tag{2.10}$$

Note that if this supremum does not exist, there does not exist a *Density Ratio Class* based on the shapes $f_l(.,\psi_l)$ and $f_u(.,\psi_u)$. If (for the parameter values ψ_l and ψ_u) the supremum does not exist because f_l has heavier tail(s) than f_u, then f_l and f_u can simply be interchanged. If (in the unrestricted case) one of the distributions has heavier tails for $\theta \to +\infty$ and the other for $\theta \to -\infty$ and this characteristic cannot be changed by choosing other parameter values for ψ_l and ψ_u, then the two distributions cannot be used to define a *Density Ratio Class*. The elicitation process for adequate shapes has then to be repeated. If $\lambda(\psi_l, \psi_u)$ exists, the class bound by $l = f_l(., \psi_l)$ and $u = \lambda(\psi_l, \psi_u) f_u(., \psi_u)$ is the smallest *Density Ratio Class* compatible with the shapes $f_l(., \psi_l)$ and $f_u(., \psi_u)$. It has now to be enlarged to contain the elicited quantile or cumulative probability intervals.

To guarantee that all elicited intervals are within the class, we substitute each of the $2(m+n)$ interval endpoints $\{p_i, \theta_i\}_{i=1}^{2(m+n)}$ of the $(m+n)$ elicited probability-quantile intervals into equations (2.7) and (2.8) and solve for κ:

$$\kappa_i(\psi_l, \psi_u) = \begin{cases} \dfrac{F_l(\theta_i, \psi_l)(1-p_i)}{p_i(1-F_u(\theta_i, \psi_u))} & \text{if} \quad p_i < \underline{F}_{f_l, \lambda(\psi_l, \psi_u) f_u}(\theta_i, \psi_l, \psi_u) \\ \lambda(\psi_l, \psi_u) & \text{if} \quad \begin{cases} \underline{F}_{f_l, \lambda(\psi_l, \psi_u) f_u}(\theta_i, \psi_l, \psi_u) \leq p_i \\ p_i \leq \overline{F}_{f_l, \lambda(\psi_l, \psi_u) f_u}(\theta_i, \psi_l, \psi_u) \end{cases} \\ \dfrac{p_i(1-F_l(\theta_i, \psi_l))}{F_u(\theta_i, \psi_u)(1-p_i)} & \text{if} \quad \overline{F}_{f_l, \lambda(\psi_l, \psi_u) f_u}(\theta_i, \psi_l, \psi_u) < p_i \end{cases} \tag{2.11}$$

The smallest *Density Ratio Class* based on the shapes $f_l(., \psi_l)$ and $f_u(., \psi_u)$ containing all elicited probability-quantile intervals is determined by the maximum of all these factors:

$$\kappa(\psi_l, \psi_u) = \max_i \{\kappa_i(\psi_l, \psi_u)\}_{i=1}^{2(m+n)} \tag{2.12}$$

To determine the class according to (2.9), the function $\kappa(\psi_l, \psi_u)$ (2.12) is now minimized over the parameters ψ_l and ψ_u. This then leads to the *Density Ratio Class* with bounding functions $l = f_l$ and $u = \kappa f_u$. The derivation procedure of this class is graphically illustrated in Figure 2.1.

If there are multiple candidates for the choice of parametric family for l and/or u, then the pair with the lowest estimated value of κ seems to be a reasonable choice for representing a *Density Ratio Class* that best corresponds to the elicited intervals. When the expert does not state intervals, but only probability-quantile points, these points can be considered as degenerate intervals with length zero. The method would remain the same, and the *Density Ratio Class* would be defined according to the scatter of points alone.

2.4 Examples

In this section, we illustrate the practicability of the suggested approach with two examples. In the first example, we use a synthetic case to investigate the sensitivity

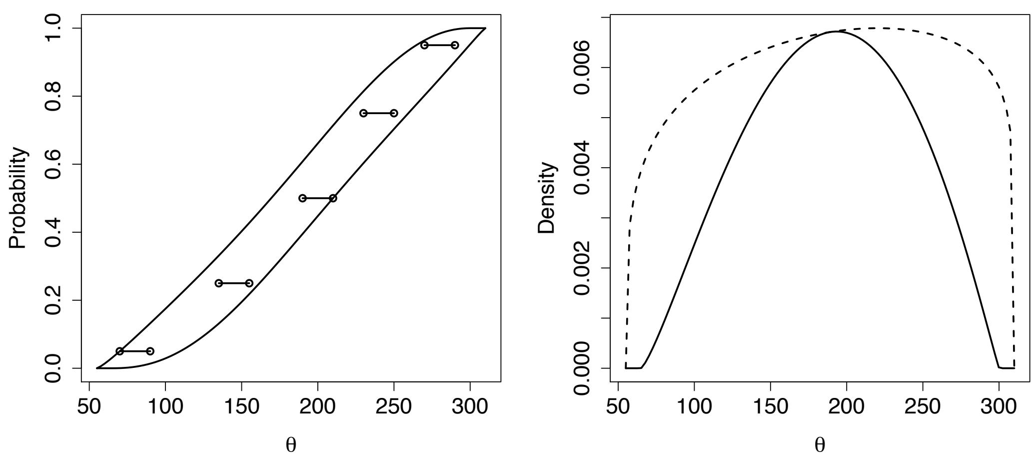

Figure 2.1: Left panel: Hypothetical elicited probability-quantile intervals for a quantity θ (horizontal lines bounded by markers) and lower and upper cumulative probability bounds of the resulting *Density Ratio Class*, $\underline{F}_{f_l,\kappa f_u}(\theta,\psi_l,\psi_u)$ and $\overline{F}_{f_l,\kappa f_u}(\theta,\psi_l,\psi_u)$ (solid lines). Right panel: normalized $l = f_l$ (solid line) and unnormalized $u = \kappa f_u$ (dashed) defining the *Density Ratio Class*. Hypothetical data: $p = \{0.05, 0.25, 0.50, 0.75, 0.95\}$, $\theta_{min} = \{70, 135, 190, 230, 270\}$, $\theta_{max} = \{90, 155, 210, 250, 290\}$, $f_l \sim Beta(2.29, 2.07)$ on $[65, 300]$, $f_u \sim Beta(1.24, 1.13)$ on $[55, 310]$ and $\kappa = 1.54$.

of the constructed *Density Ratio Class* to the difference in shape between the lower function, l, and the upper function, u, as well as to the width of assessed quantiles. In the second example, we demonstrate the application of our procedure to actual elicitation data from the literature (Borsuk et al. 2002).

2.4.1 The Sensitivity of Constructed Density Ratio Class to Shapes of l and u and to Width and Position of Elicited Quantile Intervals

Normal distributions are widely used for describing randomness in natural, social, and technical systems. However, empirical evidence shows that the rapid decrease of the *Normal density* in the tails may often be unrealistic. In such situations, empirical description of data can be improved by using distributions with heavier tails. The *Student t distribution* has heavier tails than a *Normal distribution* and approaches the *Normal distribution* as the degrees of freedom, *df*, approach infinity. With three or more degrees of freedom, the *Student t distribution* has finite variance (of $df/(df-2)$) and can be shifted and scaled to a distribution with any mean and standard deviation. For these reasons, we use a *Student t distribution* in our example to illustrate robustification of probability assessments using a *Density Ratio Class* with different shapes of l and u. When using two *Student t distributions*, l must always be the distribution with more degrees of freedom than u, otherwise we cannot find a finite constant λ with $f_u \leq \lambda f_l \; \forall \theta$ as required for construction of the *Density Ratio Class* (see equation 2.10).

Sensitivity to Shapes of l and u

Even when precise points are elicited for the probability-quantile pairs, the use of two different shapes for l and u will lead to a *Density Ratio Class* with a non-negligible degree of imprecision. To explore this issue, we assessed the sensitivity of constructed *Density Ratio Classes* to the choice of degrees of freedom of l and u. First, five quantile points with cumulative probabilities $p_1 = 0.05,\ p_2 = 0.25,\ p_3 = 0.5,\ p_4 = 0.75,\ p_5 = 0.95$ were derived from a *Student t distribution* with mean $\mu_S = 0$, standard deviation $\sigma_S = 1$ and 6 degrees of freedom. These synthetic data were then treated as "elicited" values to construct *Density Ratio Classes* using *Student t distributions* with a variety of degrees of freedom for l and u. Means of l and u were held to zero, while standard deviations and the factor κ were estimated according to the method described in section 2.3. The factor κ guarantees a proper *Density Ratio Class* ($u > l$ on the entire axis of θ) as well as coverage of all "elicited" quantiles within lower and upper cumulative probability bounds for the class. The resulting

Table 2.1: Sensitivity of κ and $\max \Delta P$ to changes in the degrees of freedom of *Student t distributions* used for l and u. Estimated standard deviations of l and u are also given.

df_l	df_u	sd_l	sd_u	κ	$\max \Delta P$
6	6	1.00	1.00	1.00	0.00
7	5	0.96	1.05	1.03	0.01
8	4	0.91	1.11	1.06	0.03
9	3	0.85	1.30	1.11	0.05

degree of imprecision can be quantified (indirectly) by the factor κ and (directly) by the maximum difference between upper and lower cumulative probability bounds for the class over all values of θ, $\max \Delta P = \sup_\theta \big(\underline{F}_{l,u}(\theta) - \overline{F}_{l,u}(\theta) \big)$. Table 2.1 shows the growth in imprecision as the difference between df_l and df_u grows.

Figure 2.2 summarizes the change in imprecision as a function of a broader range of df_l for four choices of df_u. The measure $\max \Delta P$ is increasing in df_l for fixed df_u, and is decreasing in df_u for fixed df_l. Two specific examples at the extremes are illustrated in Figure 2.3.

Sensitivity to Width and Position of Elicited Quantile Intervals

When intervals, rather than precise probability-quantile points, are elicited, the constructed *Density Ratio Class* can be expected to be sensitive to the position and width of the elicited interval. To explore this, we used the hypothetical elicited data sampled from a *Student t distribution* with six degrees of freedom for probabilities $p_1 = 0.025,\ p_2 = 0.25,\ p_3 = 0.5,\ p_4 = 0.75,\ p_5 = 0.975$ and superimposed intervals centered around various values of θ with various widths Δq. *Density Ratio Classes* were then constructed according to the method described in section 2.3. Here, the degrees of freedom of l and u were held at 8 and 4, respectively, while the means, standard deviations, and κ were estimated.

Figure 2.4 shows that, as expected, the imprecision of the *Density Ratio Class*, as measured by $\max \Delta P$, increases with the width of the elicited quantile interval. Additionally, when moving from the center to the tails of the distribution, there is first an increase and then a decrease of the size of the class. This phenomenon, as well as broader patterns of sensitivity to the width and position of quantile intervals, are detailed in Figure 2.5.

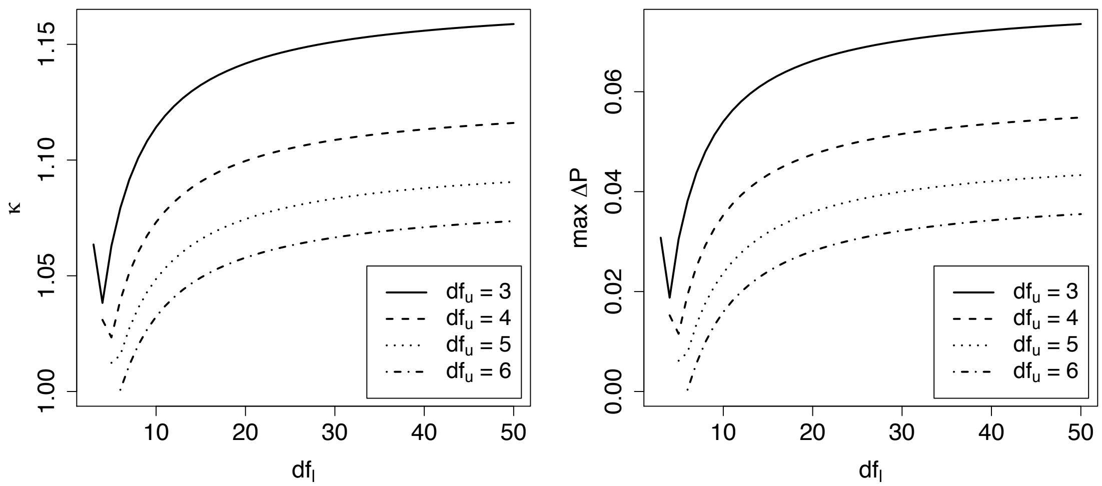

Figure 2.2: Sensitivity of imprecision in the *Density Ratio Class* constructed from hypothetical elicited probability-quantile pairs to a range of degrees of freedom $df_l \in \{df_u, ..., 50\}$ and $df_u \in \{3, ..., 6\}$ for the *Student t distributions* used to define l and u. As described in the text, means of l and u were held to zero while standard deviations and κ were estimated according to the method described in section 2.3. Imprecision is quantified (indirectly) by the factor κ (left panel) and (directly) by the maximum difference between upper and lower cumulative probability bounds $\max \Delta P$ (right panel). Two specific examples at the extremes are illustrated in Figure 2.3. Note that $\kappa = 1$ for $df_l = df_u = 6$ as the data points were derived from a *Student t distribution* with six degrees of freedom.

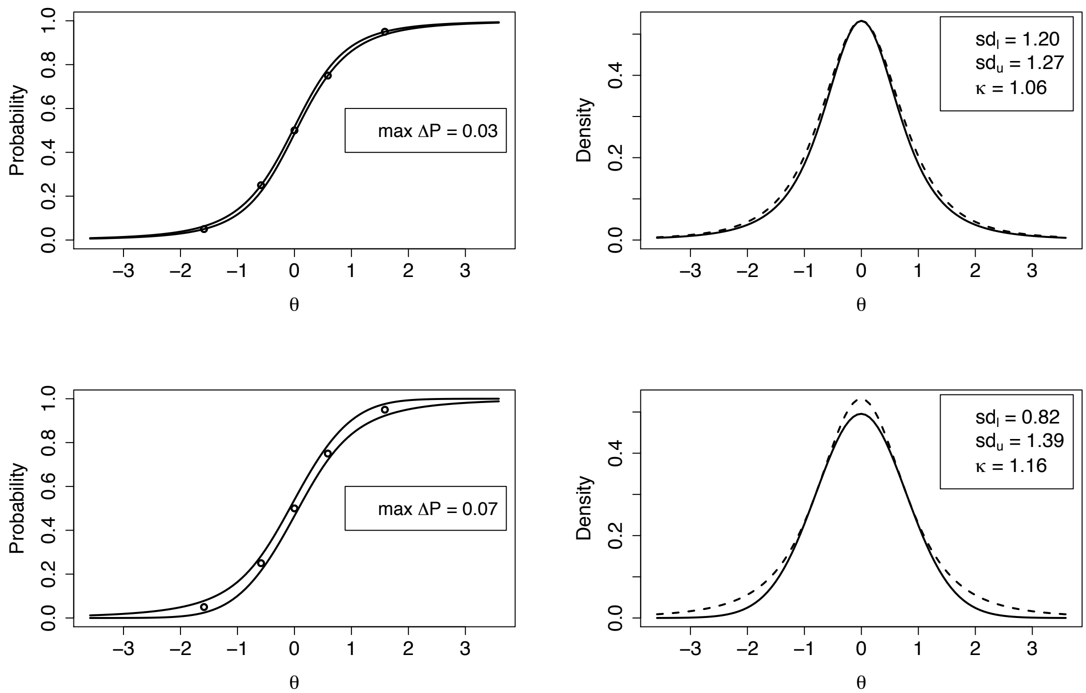

Figure 2.3: Examples of *Density Ratio Classes* constructed from hypothetical elicited probability-quantile pairs (points) using *Student t distributions* with identical (top panels: $df_l = 3 = df_u$) and vastly different (bottom panels: $df_l = 50, df_u = 3$) degrees of freedom for l and u. Left panels: Lower and upper cumulative probability bounds of the resulting *Density Ratio Class*, $\underline{F}_{f_l,\kappa f_u}(\theta, \psi_l, \psi_u)$ and $\overline{F}_{f_l,\kappa f_u}(\theta, \psi_l, \psi_u)$ (solid lines). Right panels: normalized $l = f_l$ (solid lines) and unnormalized $u = \kappa f_u$ (dashed lines) defining the *Density Ratio Class*. Means of l and u were held to zero, while standard deviations and κ were estimated according to the method described in section 2.3. Imprecision is quantified by the maximum difference between upper and lower cumulative probability bounds, $\max \Delta P$.

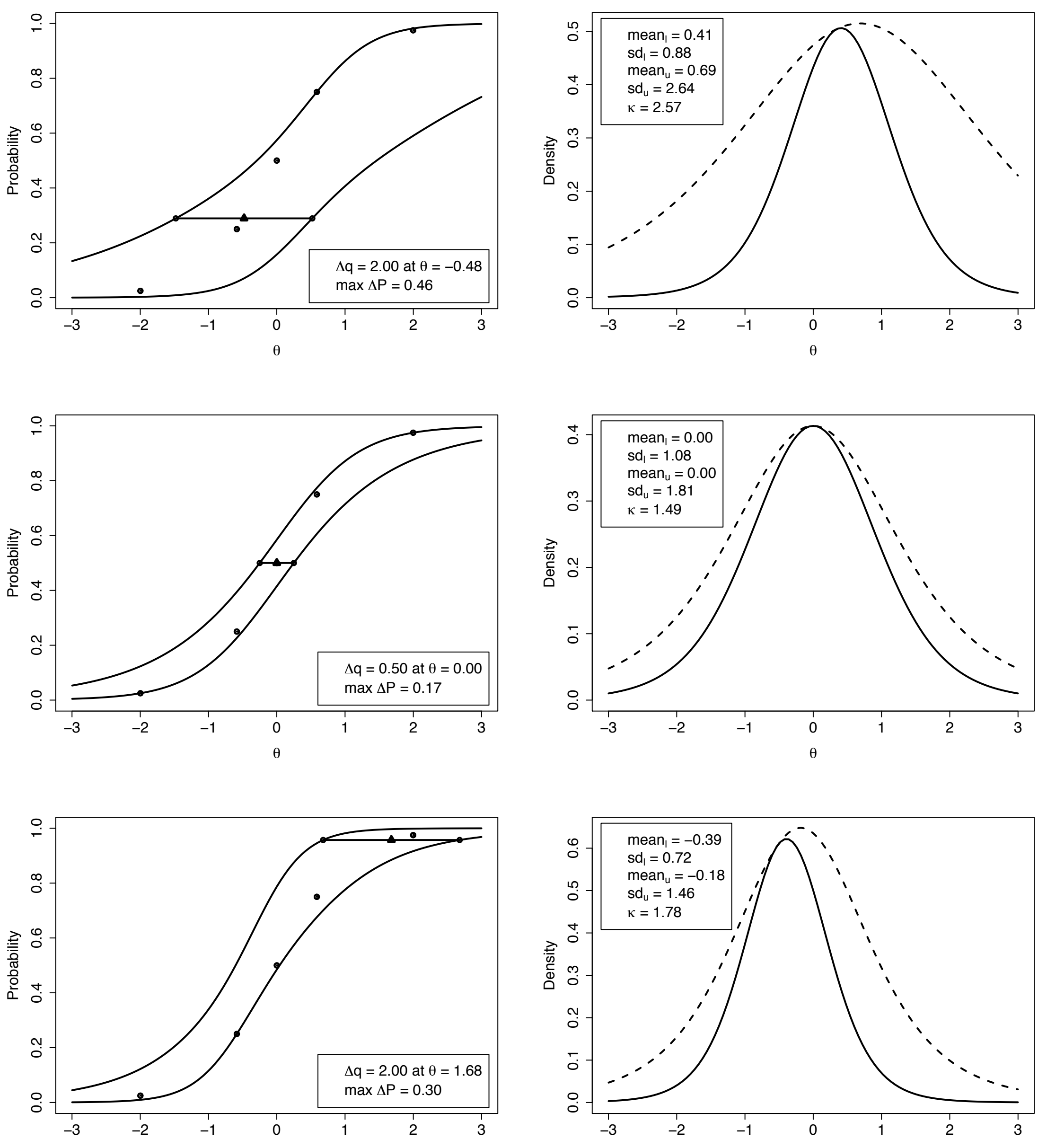

Figure 2.4: Examples demonstrating the sensitivity of constructed *Density Ratio Classes* to the width Δq and position θ of hypothetical elicited probability-quantile pairs (points) and hypothetical elicited quantile intervals (horizontal solid lines, the triangle marks the centre of the elicited quantile interval and corresponds to the quantile resulting from the *Student t distribution* with $df = 6$). We used *Student t distributions* for l and u with degrees of freedom held at 8 and 4, respectively, while the means, standard deviations, and κ were estimated according to the method of section 2.3. $\max \Delta P$ is reported as the measure of imprecision of the resulting *Density Ratio Class*. Left and right panels can be interpreted as in Figures 2.1 and 2.3.

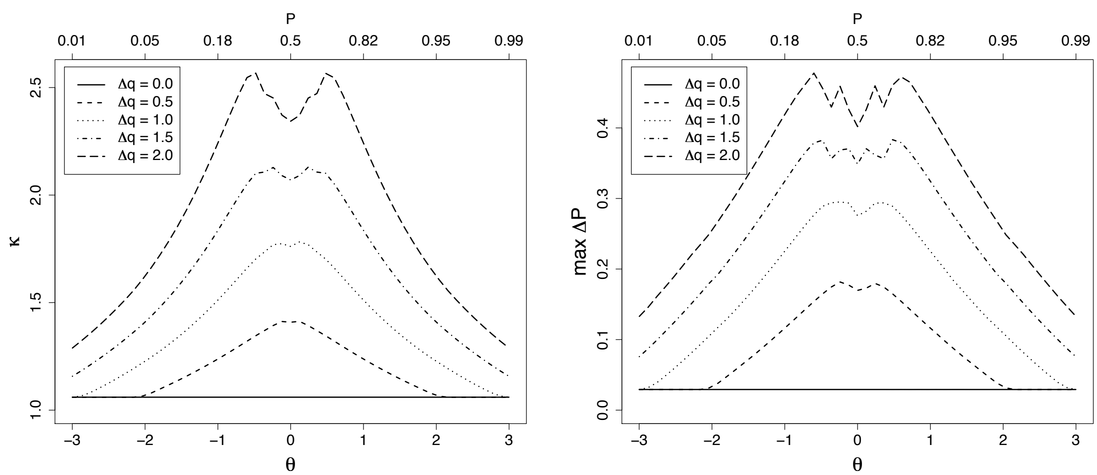

Figure 2.5: Detailed assessment of the sensitivity of constructed *Density Ratio Classes* to the width Δq and position (θ or P) of hypothetical elicited quantile intervals. As in Figure 2.4, we used *Student t distributions* for l and u with degrees of freedom held at 8 and 4, respectively. The means, standard deviations, and κ were estimated according to the method of section 2.3. The value of κ (left panel) and $\max \Delta P$ (right panel) are used to represent the imprecision of the constructed *Density Ratio Class*.

2.4.2 Example Based on Elicitation Data from the Literature

Borsuk et al. (2002) used an expert elicitation procedure to construct a model for the distribution of times-to-death of a clam species as a function of dissolved oxygen concentration in an estuary. Points on the cumulative distribution function of times-to-death were elicited for multiple oxygen concentrations from two experts simultaneously, using the fixed-probability protocol (Spetzler and Staël Von Holstein 1975). Because of either minor disagreement between the experts or shared imprecision, the experts stated intervals rather than points for many of the elicited quantiles. These intervals were recorded and reported, but not used, in the original study. Rather, Borsuk et al. used the midpoints of the intervals to fit precise distributions for each oxygen concentration. They fit a variety of parametric families to the elicited data and found that the *Lognormal* provided the best fit, but went on to use the *Log-Logistic* for computational convenience.

We construct a *Density Ratio Class* from the data of Borsuk et al. (2002) using the technique described in section 2.3. Because the experts did not assert a parametric family for l or u, we first log-transformed the elicited values and then fit every combination of the *Normal*, *logistic*, t_3, and t_{10} distributed families that would lead to a defined *Density Ratio Class*. (It would also be possible to empirically estimate the degrees of freedom parameters of the *Student t densities* characterizing l or u. For simplicity, we did not pursue this additional step for our example.) It should be noted that the logarithmic transformation has no consequence for the value κ.

A combination of the *Normal density* for l and the t_3 density for u provided the best fit, as measured by the lowest average value of κ across the four dissolved oxygen concentrations (Figure 2.6). This is a useful class representation that allows for a variety of tail behaviors. The imprecision contained in the elicited intervals is fairly symmetric on the log-scale, with means of l and u being similar for most cases. The imprecision is also fairly constant across differing oxygen concentrations, with similar relative values of the standard deviations l and u and similar values of κ and $\max \Delta P$.

2.5 Discussion

The comparison of various classes of probability distributions shows conceptual advantages for the *Density Ratio Class*. This class allows for a wide variety of shapes (including tail behaviors) without allowing very unreasonable shapes, such as high peaks and point masses. Additionally, based on some technical assumptions, Wasser-

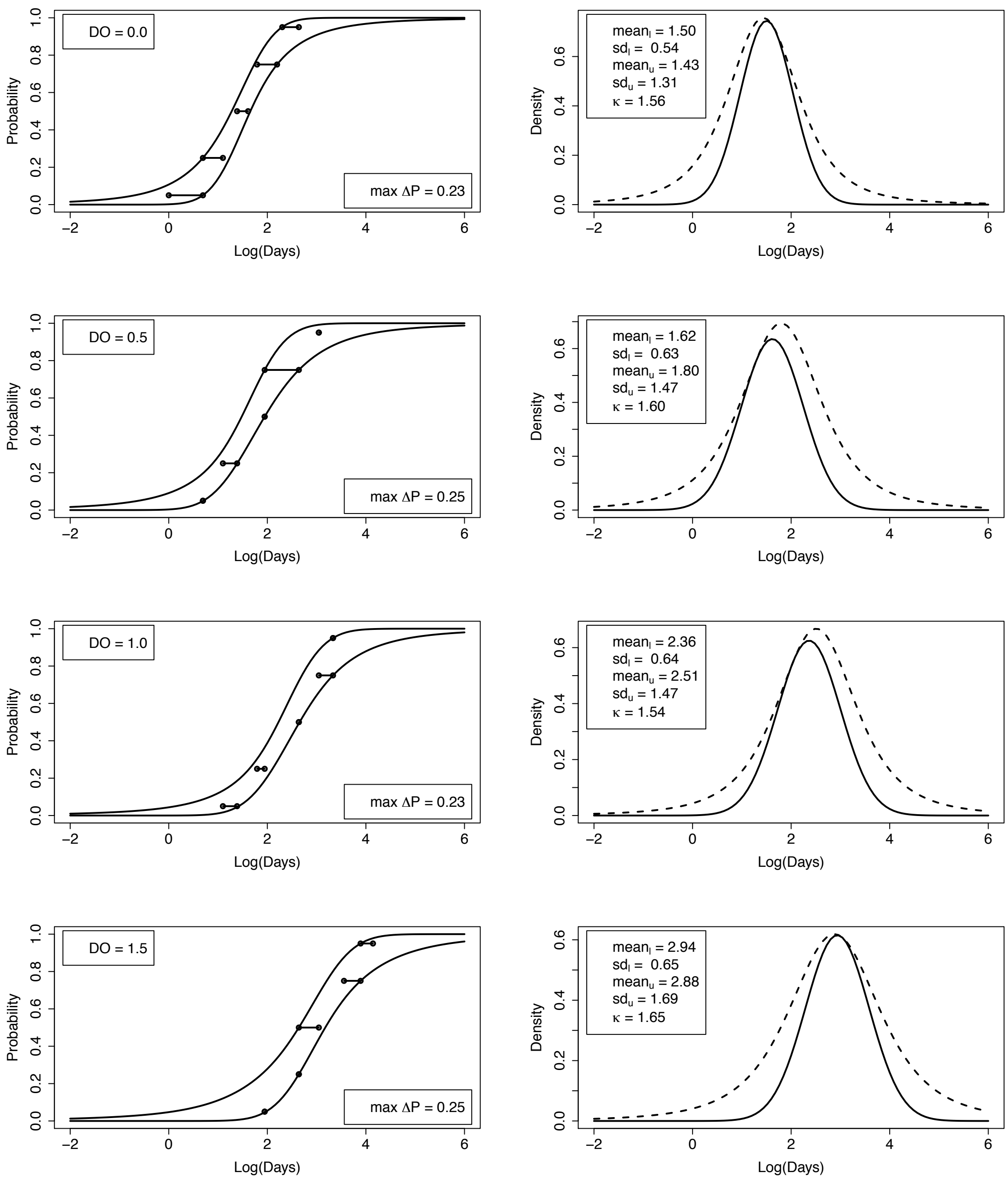

Figure 2.6: *Density Ratio Classes* constructed from the logarithm of assessed quantile intervals on the number of days corresponding to each specified cumulative mortality and ambient dissolved oxygen concentration (DO, [mgL^{-1}]) according to the data of Borsuk et al. 2002. In each case a *Normal density* was used for l and a *Student t density* with three degrees of freedom was used for u. Left and right panels can be interpreted as in Figures 2.1, 2.3 and 2.4.

man (1992a) showed that the *Density Ratio Class* is the only class that is invariant under Bayesian updating and marginalization. The former property is crucial for representing incremental learning within a consistent framework, the latter for visualizing important aspects of the class.

Of course, we do not claim that experts have internal representations of uncertainty that conform precisely with the *Density Ratio Class*. However, if an expert is able to specify quantile or cumulative probability intervals for a quantity, select one or two parametric families that are compatible with these intervals and his or her general beliefs about possible values of the quantity, and wants to allow neighboring shapes but exclude extreme departures, then a *Density Ratio Class* constructed according to our technique seems to yield a reasonable representation of his or her beliefs. These considerations seem to be fulfilled for many continuous random variables typically elicited from experts. However, it may be that additional practical experience will point to other representations.

Our suggested technique for constructing a *Density Ratio Class* is based on well established elements of conventional elicitation procedures, including elicitation of quantiles or cumulative probabilities and use of parametric shapes. Thus, the procedure should not be too unfamiliar to experts or analysts. In fact, experts may be even more comfortable with our proposed method because they can state their judgments as intervals rather than being required to assert precise values. Additionally, the parametric shapes that are selected are only employed to bound the *Density Ratio Class* and not to define the nonparametric members. In this way, robustification to deviations from parametric shapes is addressed. If the expert does want to specify precise quantiles (intervals of zero length), then class boundaries are derived from the scatter and systematic deviations of the assessed quantiles from the parametric shape. Finally, if the expert specifies precise quantiles that are compatible with a single parametric shape chosen as the bounding shape, then the constructed *Density Ratio Class* simplifies to a precise member of that family. This assures compatibility of results, as well as techniques, with conventional methods.

The size of the class constructed from our method is determined by three elements: (i) Differences in parametric shapes selected for bounding the shapes already introduce an element of imprecision to the class. (ii) Scatter of the elicited quantiles or incompatibility with the parametric shape(s) increase the size of the class. (iii) Finally, explicit imprecision, as specified by the expert using quantile or cumulative probability intervals, further increases the size of the class. This latter element can be particularly influential if cumulative probability intervals (rather than quantile intervals) are specified in the tails of the distribution.

There are also important interactions between the three elements of imprecision described above. Our analysis of the sensitivity of the *Density Ratio Class* to the shapes of the bounding densities showed that increasing the ambiguity of tail behavior leads to a comparable increase in imprecision in the central part of the distribution as well (see Fig. 2.3). We also found that equally wide elicited quantile intervals will have differing effects on overall imprecision depending on the shapes chosen for bounding the class, as well as the location of the elicited intervals (see Fig. 2.2 and 2.4). We make particular propositions for metrics characterizing the imprecision of *Density Ratio Classes* in Rinderknecht et al. (2012).

In the description of our technique, we focused on the one-dimensional case, but the *Density Ratio Class* can be readily defined for the multivariate case as well (DeRobertis and Hartigan 1981). However, the construction of elicitation techniques that appropriately capture covariance in an expert's knowledge about multiple uncertain quantities remains a challenge (Held et al. 2008).

When uncertainty is represented by imprecise probabilities, such as the *Density Ratio Class*, then conventional decision theory based on expected utility maximization may not provide a unique ranking of decision alternatives. This is because a *Density Ratio Class* for an uncertain quantity will usually translate into an interval for the resultant expected utility. If intervals on expectations overlap for two or more alternatives, then there is not an established Bayesian decision rule for choosing between them. Alternate criteria include maximum lower expected utility, maximum upper expected cost, or minimum upper regret. The need for careful consideration in adopting an appropriate secondary decision rule is one cost to be paid for the added descriptive capability provided by our proposed approach.

We are currently preparing an R package to facilitate elicitation and construction of the *Density Ratio Classes* (Rinderknecht et al. 2012). This package extends concepts implemented in the Sheffield Elicitation Framework (`http://www.tonyohagan.co.uk/shelf`) to *Density Ratio Classes*.

Acknowledgements

The authors would like to thank all anonymous reviewers for their useful comments. We gratefully acknowledge support from the Swiss National Science Foundation. Mark E. Borsuk was partially supported by the US EPA through a grant from the STAR program. This work has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

Chapter 3

Bridging Uncertain and Ambiguous Knowledge with Imprecise Probabilities

Simon L. Rinderknecht[1,2], Mark E. Borsuk[3] and Peter Reichert[1,2]

Published in the Thematic Issue on Expert Opinion in Environmental Management and Modelling in the Journal of Environmental Modelling & Software (2012), 36, 122-130. doi:10.1016/j.envsoft.2011.07.022

[1] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Systems
Analysis, Integrated Assessment and Modelling, CH-8600 Dübendorf, Switzerland
[2] ETH Zurich, Department of Environmental Sciences, Institute of Biogeochemistry and Pollutant Dynamics (IBP), CH-8092 Zürich, Switzerland
[3] Dartmouth College, Thayer School of Engineering, Hanover, NH 03755-8000, USA

Abstract

Model-based environmental decision support requires that uncertainty be rigorously evaluated. Whether uncertainty is *aleatory* or *epistemic*, we argue that probability is the natural mathematical construct for describing uncertainty in predictions used for decision-making. If expert knowledge is elicited using stated preferences between lotteries, and the experts are rational in the sense of avoiding sure loss, then the resulting knowledge quantifications will be consistent with the axiomatic foundation of probability theory. This idea can be extended to the description of *intersubjective* *knowledge* when the intent is to characterize the state of knowledge of the scientific community. Many methods for probability elicitation have been reported, but there is nearly always some degree of *ambiguity* in translating elicited quantities into probabilistic description. This would include: any lack of fit of a particular distributional form to elicited data; incertitude in the elicited data themselves; and/or disagreement in the elicited data across multiple experts. By replacing a precise probability distribution by a set of distributions, the mathematical concept of *imprecise probabilities* provides a means for representing this ambiguity. In this way, imprecise probabilities can form a bridge between total ignorance and precisely characterized risk by allowing for a continuous degree of imprecision to represent ambiguity. We introduce three metrics to describe the relative ambiguity of important attributes of probability distributions, namely their width, shape, and mode. These metrics are applicable to sets of distributions characterized by using any available method, and we derive the specific forms of these metrics for the *Density Ratio Class*, which we have found to have many desirable properties. Based on these metrics and on elicitation data from the literature, we use three examples to demonstrate the wide variety of ambiguity that can be present in elicited knowledge. Imprecise probabilities allow us to quantify this ambiguity and consider it in environmental decision-making. Our examples were implemented using a package we recently developed and made freely available for the R statistical programming environment.

Keywords

expert elicitation; subjective probabilities; intersubjective knowledge; interval probabilities; qualitative expertise; quantitative expertise; robust Bayesian inference; robust Bayesian statistics; quantile elicitation; imprecise probabilities; probability box; quantile class; *Density Ratio Class*.

3.1 Introduction

Nature's complexity and stochastic behavior imply that models of environmental systems are always approximations of reality and lead to uncertain predictions. Sources of uncertainty include: (i) non-deterministic, potentially stochastic behavior of the true system - referred to as *aleatory uncertainty* and (ii) lack of knowledge about the true system, its mathematical representation, and specific parameter values - referred to as *epistemic uncertainty* (Parry 1996; Walker et al. 2003; Refsgaard et al. 2007). In environmental modeling, epistemic uncertainty is often dominant (Ayyub and Klir 2006). Recognizing and quantifying both types of uncertainty is important because it allows modelers to effectively allocate their resources toward model improvement and allows decision-makers to assess the degree of confidence they can have in model predictions (Warmink et al. 2010).

Probability theory has long been the well-accepted mathematical framework for describing aleatory uncertainty. However, Keynes (1921), de Finetti (1931), Ramsey (1931), Cox (1946) and others have shown by the so-called *Dutch Book argument* that probability theory is also appropriate for describing epistemic uncertainty: when an individual's state of knowledge is quantified using stated preferences between lotteries (with the requirement that such preferences be consistent in the sense of avoiding sure loss), then the resulting knowledge quantifications adhere to the laws of probability. Additionally, since aleatory uncertainty becomes epistemic uncertainty once a random event has taken place and if its outcome is not yet observed, describing both kinds of uncertainty within the same mathematical framework avoids problems of inconsistency between mathematical formalisms. This is consistent with the viewpoint that Bayesian statistics is the logical framework for inference and prediction (de Finetti 1974; Howson and Urbach 1989; Seidenfeld et al. 1995; Kadane and O'Hagan 1995; Kadane et al. 1996).

Of course, in most cases, a modeler will not be entirely familiar with the current state of knowledge or opinion regarding the relevant uncertainties and so may seek outside expertise (Pollino et al. 2007; Reichert et al. 2007). The formal approach to obtaining expertise about an uncertain quantity within probability theory is referred to as *probability elicitation*, and a variety of pertinent approaches, guidelines, and cautions have been published (e.g., Meyer and Booker (1991), O'Hagan et al. (2006), James et al. (2010), and see also section 3.2 for further references). In the case of models being used to inform public decisions, a modeler might be interested in representing *intersubjective knowledge*, rather than the beliefs of individual experts. Intersubjective knowledge in such a context represents the current state of knowledge of the scientific community about an environmental system, its mathematical description, or specific parameter values. Arguments in favor of a mathematical formalism, such as probability theory, to represent both aleatory and epistemic uncertainty are even further strengthened in the case of intersubjective knowledge representation because of the need to maintain consistency and transparency. Note that the importance of an intersubjective interpretation of probabilities to describe scientific reasoning has already been discussed by Gillies (1991).

As outlined in the previous paragraph, there are convincing arguments for formulating epistemic, subjective and, especially, intersubjective knowledge by probabilities. However, inaccuracies in elicitation procedures, misrepresentation of elicitation results, problems in expressing an individual's beliefs quantitatively, different perception of information by different individuals, or disagreement between experts can all lead to uncertainty about the probabilistic quantification of knowledge

(O'Hagan and Oakley 2004). This type of uncertainty has been referred to as ambiguity (Ellsberg 1961; Frisch and Baron 1988). In particular, it has been discussed in decision sciences where ambiguity aversion (aversion to unknown probabilities) is distinguished from risk aversion (aversion to uncertainty that can be quantified probabilistically) (Einhorn and Hogarth 1985; Camerer and Weber 1992). As ambiguity is a different aspect of uncertainty than probabilistically quantified uncertainty, we are interested to identify, describe, and seek to reduce this particular form of uncertainty regardless of how much additional uncertainty may be embedded in the elicited probabilities themselves.

One method for separating ambiguity in the choice of a probability specification from the uncertainty contained within the specification itself is to replace the standard single probability distribution with a set of distributions. This is an extension of conventional probability theory and the literature broadly refers to it as *imprecise probability theory* (Walley 1991; Caselton and Luo 1992, http://www.sipta.org). In the context of imprecise probability theory, conventional Bayesian statistics extend to what is called *robust Bayesian statistics* (Ríos Insua and Ruggeri 2000; Berger 1994). Depending on the degree of ambiguity, a set of probability distributions can contain a large variety of shapes or can simply contain those shapes in a small *neighborhood* of a particular distribution. Multiple approaches, or *classes*, have been proposed to define membership in such sets (see references in section 3.2), and we believe it would be useful to have some standard metrics for describing the relative ambiguity contained in any particular set, independent of the approach taken to set specification.

In this paper, we propose metrics to describe the relative ambiguity contained in a set of distributions defined according to imprecise probability theory, and we apply these metrics to demonstrate the wide variety of ambiguity present across different application cases. The paper is structured as follows. In section 3.2, we briefly discuss probability elicitation and the relative merits of various classes of imprecise probabilities. In section 3.3, we introduce some general metrics to quantify the degree of relative ambiguity in any such class. In section 3.4, we implement these metrics for a particular class that we have found most useful, the *Density Ratio Class*. In section 3.5, we demonstrate the use of our metrics using elicitation data from the literature. We present three cases of differing degree of ambiguity in order to show the wide range present in actual elicitation results. In section 3.6, we discuss (1) our metrics of ambiguity relative to others, (2) the merits of using imprecise probabilities, relative to second-order probabilities and (3) some implications of using imprecise probabilities for environmental decision support. Finally, in section 3.7 we draw our conclusions. In the section Software Availability, we present our elicitation software written in R that is applicable to the *Density Ratio Class*.

3.2 Elicitation

A common technique for eliciting a probability distribution from an expert for a continuous quantity is to employ the *quantile method*. According to this method, the analyst provides a number of cumulative probabilities (e.g., 0.05, 0.25, 0.5, 0.75, 0.95) and the expert then estimates the corresponding quantiles of the uncertain quantity. This procedure, first suggested by Winkler (1967), minimizes anchoring and other biases that may be inherent in the "cumulative probability method", in which the modeler provides values of the uncertain quantity and the expert estimates the corresponding cumulative probabilities.

We and others (e.g. Borsuk et al. (2002), Kriegler et al. (2009)) have found that, whether the quantile or probability elicitation method is used, experts typically feel more comfortable if they are allowed to express their assessments as intervals. We take this as an indication that they cannot make probability judgments with absolute precision, and we have usually recorded the intervals as stated. However, using established techniques of summarizing expert elicitation results, we have then generally proceeded by choosing a precise probability distribution to represent the elicited data, ignoring the intervals (and the lack of fit to any individual data) except to check that the chosen distribution generally "passes approximately through" the data. But in cases as shown in our third example (see section 3.5.3) it seems important to consider the high uncertainty about probabilities explicitly (Einhorn and Hogarth 1985; Camerer and Weber 1992).

A natural way to use the intervals stated by an expert would be to interpret them as upper and lower bounds on the scope of their beliefs about the quantiles or cumulative probabilities being elicited. For example, Ferson and Hajagos (2004) take stated intervals to define the edges of a *p-box* presumed to contain a feasible set of cumulative distributions. This leads to a very broad set, including distributions with sharp peaks or point masses in their probability density functions. Destercke et al. (2008) further generalized the usual *p-boxes*. Alternative approaches to defining set membership include the ϵ-*Contamination Class* (Huber 1973), the *Quantile Class* (Lavine 1991a; Moreno and Pericchi 1993b), the *Density Bounded Class* (Lavine 1991b), and the *Density Ratio Class* (DeRobertis and Hartigan 1981; Berger 1990). Berger (1994) gives an overview of these classes, and Borsuk and Tomassini (2005) and Rinderknecht et al. (2011) discuss their relative merits in the context of environmental modeling. Although there are differences in their particular mathematical description, in theory any of these approaches to defining imprecise probabilities might be used to capture ambiguity in elicited distributions. We, therefore, next propose some general metrics for describing the relative ambiguity, or imprecision, of important attributes of an imprecise distribution that results from any particular set definition.

3.3 Metrics of Imprecision

There seem to be at least three important attributes of probability distributions for which we would be interested to quantify the ambiguity or imprecision: (i) the width of the distribution, (ii) the shape of the distribution within its range, and (iii) the position of the mode. We need metrics of ambiguity or imprecision about these attributes that are independent of the particular definition used to define the imprecise probability class. The focus on imprecision in specific attributes complements more general measures of imprecision of the whole distribution (Klir and Wierman 1998) that may be more difficult to interpret.

We develop metrics of imprecision that are expressed relative to a specified reference range (e.g., a credible interval) on the uncertain quantity in order to make the metrics comparable across different applications and to emphasize how much information is being added relative to a simple interval. It is important to note that high relative imprecision may be less practically important if overall uncertainty, as expressed by the reference range, is narrow.

3.3.1 Imprecision About the Width of a Variable

The width of an uncertain variable is typically characterized by a $1-\alpha$ credible interval which is denoted by $I_{1-\alpha}$ and defined as an interval that contains the value of the variable with a probability of $1-\alpha$. Typically, the significance level α is chosen to be 10%, 5%, or 1%. For a one-dimensional random variable $\mathbf{\Theta}$, such a credible interval is usually defined as being bounded by the $\alpha/2$ and $1-\alpha/2$ quantiles of the probability distribution of this random variable and hence one writes $I_{1-\alpha} = \left[Q_{\alpha/2}[\mathbf{\Theta}], Q_{1-\alpha/2}[\mathbf{\Theta}]\right]$. However, our metrics could easily be adapted relative to any other definition of credible interval, i.e., the highest probability density region of content $1-\alpha$.

If knowledge about $\mathbf{\Theta}$ is characterized by a set of probability distributions, we can determine the outer bounds on the $1-\alpha$ credible interval as

$$\overline{I}_{1-\alpha}[\mathbf{\Theta}] = \left[\overline{Q}_{\alpha/2}[\mathbf{\Theta}], \underline{Q}_{1-\alpha/2}[\mathbf{\Theta}]\right] \tag{3.1}$$

where $\overline{Q}_p[\mathbf{\Theta}]$ is the value, θ, at which $\overline{P}(\mathbf{\Theta} \leq \theta) = p$ and $\underline{Q}_p[\mathbf{\Theta}]$ is the value, θ, at which $\underline{P}(\mathbf{\Theta} \leq \theta) = p$. In these last expressions, $\overline{P}(.)$ is the supremum of the probabilities of the statement in its argument over the distributions belonging to the class, and $\underline{P}(.)$ is the infimum. In other words, we always have $\overline{Q}_{p_1}[\mathbf{\Theta}] \leq \underline{Q}_{p_2}[\mathbf{\Theta}]$ for $p_1 \leq p_2$ and $\overline{I}_{1-\alpha}[\mathbf{\Theta}]$ contains all $1-\alpha$ credible intervals defined by $\left[Q_{\alpha/2}[\mathbf{\Theta}], Q_{1-\alpha/2}[\mathbf{\Theta}]\right]$ within the set definition employed.

We can define a relative measure of imprecision, $r^{\text{width}}_{1-\alpha}$, on this interval by dividing the average width of the imprecise interval for each of the two endpoints by the length between the outer bounds on the credible interval given above:

$$r^{\text{width}}_{1-\alpha}[\boldsymbol{\Theta}] = \frac{\left(\underline{Q}_{\alpha/2}[\boldsymbol{\Theta}] - \overline{Q}_{\alpha/2}[\boldsymbol{\Theta}]\right) + \left(\underline{Q}_{1-\alpha/2}[\boldsymbol{\Theta}] - \overline{Q}_{1-\alpha/2}[\boldsymbol{\Theta}]\right)}{2\left(\underline{Q}_{1-\alpha/2}[\boldsymbol{\Theta}] - \overline{Q}_{\alpha/2}[\boldsymbol{\Theta}]\right)} \tag{3.2}$$

This relative measure of imprecision is zero for precise probability distributions and can reach values close to unity in cases of extreme imprecision such as e.g. no further knowledge than the random variable's range. However, even in cases of extreme imprecision, $\overline{Q}_{\alpha/2}[\boldsymbol{\Theta}]$ and $\underline{Q}_{1-\alpha/2}[\boldsymbol{\Theta}]$ are conservative (inclusive) estimates of the range of the variable at the probability level $1-\alpha$.

3.3.2 Imprecision About the Distribution of the Variable within its Range

Imprecision of a particular quantile at cumulative probability p can be characterized by the length of the quantile interval characterizing the imprecision in that quantile: $\underline{Q}_p[\boldsymbol{\Theta}] - \overline{Q}_p[\boldsymbol{\Theta}]$. Dividing the supremum of these imprecise quantile intervals (for all p between α and $1-\alpha$) by the length between the outer bounds of the $1-\alpha$ credible interval (see equation 3.1) leads to a non-dimensional relative measure of imprecision, $r^{\text{shape}}_{1-\alpha}$, about the distribution of $\boldsymbol{\Theta}$ within this interval:

$$r^{\text{shape}}_{1-\alpha}[\boldsymbol{\Theta}] = \frac{\sup_{\alpha/2\leq p\leq 1-\alpha/2}\left(\underline{Q}_p[\boldsymbol{\Theta}] - \overline{Q}_p[\boldsymbol{\Theta}]\right)}{\underline{Q}_{1-\alpha/2}[\boldsymbol{\Theta}] - \overline{Q}_{\alpha/2}[\boldsymbol{\Theta}]} \tag{3.3}$$

Again, if the shape of the distribution is precisely known, then $r^{\text{shape}}_{1-\alpha}[\boldsymbol{\Theta}]$ is zero, whereas in the case of extremely imprecise distributions, $r^{\text{shape}}_{1-\alpha}[\boldsymbol{\Theta}]$ is close to unity.

3.3.3 Imprecision About the Mode of the Variable

Imprecision of a particular characteristic, such as the mode, can be characterized by the length of the interval, $I_{\text{mode}} = \left[\underline{\theta}_m, \overline{\theta}_m\right] \subset \Theta$ considered to include that characteristic. Expressed relative to the length of $\overline{I}_{1-\alpha}$ we obtain the following non-dimensional measure:

$$r^{\text{mode}}_{1-\alpha}[\Theta] = \frac{\overline{\theta}_m - \underline{\theta}_m}{\underline{Q}_{1-\alpha/2}[\Theta] - \overline{Q}_{\alpha/2}[\Theta]}. \tag{3.4}$$

This measure is zero if the mode is precisely known and for typical shapes and small values of α it is below unity.

3.4 Implementation for the Density Ratio Class

As already mentioned, a variety of imprecise probability classes have been proposed. Rinderknecht et al. (2011) discuss the relative merits of the different classes and conclude that the *Density Ratio Class* has clear conceptual and practical advantages. In particular, the *Density Ratio Class*'s invariance under Bayesian updating and marginalization (Wasserman 1992a) makes it the unique class that allows for simultaneously describing a consistent sequential Bayesian learning process and conveniently conveying higher dimensional cases. A *Density Ratio Class* can contain a wide variety of not necessarily parametric shapes, including varied tail behavior. Sharp peaks and points masses that are likely to be considered unreasonable by an expert are typically excluded, which is a rather desirable feature for most environmental modeling applications. When parameters are described by *Density Ratio Classes*, then probabilistic predictions of even non-deterministic models can be calculated without a large increase of computational burden.

The main disadvantage of the *Density Ratio Class* has been its elicitation. However, Rinderknecht et al. (2011) suggest a practicable elicitation procedure for the *Density Ratio Class* to overcome most difficulties. After giving the definition of the *Density Ratio Class*, we briefly review this elicitation procedure. We then state special cases of our metrics for the *Density Ratio Class*.

3.4.1 Definition of the Density Ratio Class

The *Density Ratio Class* of normalized probability density functions (PDFs), f, is defined as the set

$$\Gamma^{DR}_{f_l,\kappa f_u} = \{\text{PDFs } f(\theta) : \exists c : f_l(\theta,\psi_l) \leq cf(\theta) \leq \kappa f_u(\theta,\psi_u) \leq \infty \;\forall\theta\} \tag{3.5}$$

where the bounding normalized densities, f_l and f_u (parameterized by ψ_l and ψ_u) describe the lower and upper shapes of the class. The factor κ obeys the following inequality: $1 \leq \sup_\theta\{f_l/f_u\} \leq \kappa < \infty$ and can be used as a direct characterization of class imprecision for given f_l and f_u. Put simply, a *Density Ratio Class*, $\Gamma^{DR}_{l,u}$, can be interpreted as a set of measures with unnormalized densities between specified unnormalized lower and upper bounds $l \leq u$. The set of probability densities is then obtained by normalizing these measures. Thus, the lower and upper bounds limit the shapes of allowable densities irrespective of their normalization. Note, by the definition of the *Density Ratio Class* it follows that $\Gamma^{DR}_{l,u} = \Gamma^{DR}_{\lambda l,\lambda u}$ for any $0 < \lambda < \infty$. Historically, DeRobertis and Hartigan (1981) introduced the *Density Ratio Class* under the name of *intervals of measures* whereas Berger (1990) called the class the *Density Ratio Class*, as it bounds ratios of densities.

3.4.2 Elicitation of the Density Ratio Class

Our elicitation concept for the *Density Ratio Classes* (Rinderknecht et al. 2011) can be summarized by three steps:

1. The analyst provides n cumulative probabilities and the expert states his or her beliefs by giving corresponding quantile intervals for each probability. The result of this step is a discrete set of endpoints of n quantile intervals, $\{p_i, \theta_i\}_{i=1}^{2n}$.
2. Possible parametric families of distributions for the normalized lower, $f_l(\theta, \psi_l)$, and the normalized upper, $f_u(\theta, \psi_u)$, densities defining the shapes of the unnormalized densities bounding the *Density Ratio Class* are discussed, evaluated and selected.
3. The smallest *Density Ratio Class* bounded by the specified parametric shapes and containing the elicited quantile intervals is constructed by choosing the parameters (ψ_l, ψ_u) to minimize κ.

Sometimes it is useful if the analyst transforms the parameter θ before applying the method described above. In this case, the analyst ideally shows the expert a larger set of possible shapes that include the transformation. This does not change the procedure specified above. Alternatively, if the expert is familiar with transforming probability distributions, the expert may suggest a transformation and a parametric distribution for the transformed variable in step 2. (It is important to note that the value of κ is independent of invertible θ-transformations.) A feedback session would then indicate whether the expert agrees with the result. This approach is exemplified in our example 3.

3.4.3 Metrics of Imprecision Applied to the Density Ratio Class

The metrics proposed to quantify relative imprecision in important characteristics of the distributions in section 3.3 can be calculated for any type of imprecise probabilities representable by a set of probability densities. In this section, we derive specific formulas for calculating these metrics for the *Density Ratio Class*.

In a *Density Ratio Class*, $\Gamma_{l,u}^{DR}$, where $l = f_l$ and $u = \kappa f_u$, the pointwise lower and upper cumulative probabilities $P(\Theta < \theta)$, $\underline{F}_{l,u}(\theta)$ and $\overline{F}_{l,u}(\theta)$, respectively, describing together the envelope of all possible cumulative distributions within the class, are given by

$$\underline{F}_{f_l,\kappa f_u}(\theta) = \frac{F_l(\theta)}{F_l(\theta) + \kappa(1 - F_u(\theta))} \tag{3.6}$$

and

$$\overline{F}_{f_l,\kappa f_u}(\theta) = \frac{\kappa F_u(\theta)}{\kappa F_u(\theta) + (1 - F_l(\theta))} \tag{3.7}$$

where F_l and F_u denote the cumulative distribution functions of the normalized lower and upper densities, f_l and f_u, respectively (Rinderknecht et al. 2011).

It is then straightforward to derive the formula for the credible interval presented in equation (3.1) for a *Density Ratio Class*:

$$\overline{I}^{\text{DRC}}_{1-\alpha}[\boldsymbol{\Theta}] = \left[\overline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2), \underline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2)\right]. \tag{3.8}$$

The first metric introduced in equation (3.2) is then given by:

$$r^{\text{width,DRC}}_{1-\alpha}[\boldsymbol{\Theta}] = \frac{\left(\underline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2) - \overline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2)\right) + \left(\underline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2) - \overline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2)\right)}{2\left(\underline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2) - \overline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2)\right)}. \tag{3.9}$$

Similarly, for equation (3.3) we obtain:

$$r^{\text{shape,DRC}}_{1-\alpha}[\boldsymbol{\Theta}] = \frac{\sup_{\alpha/2 \le p \le 1-\alpha/2}\left(\underline{F}^{-1}_{f_l,\kappa f_u}(p) - \overline{F}^{-1}_{f_l,\kappa f_u}(p)\right)}{\underline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2) - \overline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2)}. \tag{3.10}$$

Finally, imprecision in the mode according to equation (3.4) can be written as:

$$r^{\text{mode,DRC}}_{1-\alpha}[\Theta] = \frac{\max(\kappa f_u^{-1}(M_l)) - \min(\kappa f_u^{-1}(M_l))}{\underline{F}^{-1}_{f_l,\kappa f_u}(1-\alpha/2) - \overline{F}^{-1}_{f_l,\kappa f_u}(\alpha/2)}. \tag{3.11}$$

where $M_l = f_l(\text{mode}_l)$ and f_l^{-1} and f_u^{-1} denote the set-valued inverse lower and upper normalized distributions respectively with $f_\cdot^{-1}(d) = \{\theta \in \Theta : \ f_\cdot(\theta) = d\}$.

3.5 Examples

3.5.1 Largely uncertainty: Date of maximum periphyton biomass in a riverine ecosystem model

Schweizer (2007) developed a deterministic model with a stochastic error for river periphyton biomass recovery after a flood. In reduced form, this model can be expressed as

$$\boldsymbol{B}_{\Delta t_{\text{flood}}}(\boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{b}_{\Delta t_{\text{flood}}}(\boldsymbol{x}, \boldsymbol{\theta}) + \boldsymbol{Z}(\boldsymbol{\theta}), \tag{3.12}$$

where $\boldsymbol{b}_{\Delta t_{\text{flood}}}(\boldsymbol{x}, \boldsymbol{\theta})$ is the deterministically modeled biomass of periphyton, consisting multiplicatively of a Monod function, $\boldsymbol{m}(\boldsymbol{x}, \boldsymbol{\theta})$, limiting terms, $\boldsymbol{l}(\boldsymbol{x}, \boldsymbol{\theta})$, and a seasonality term, $\boldsymbol{s}(\boldsymbol{x}, \boldsymbol{\theta})$. Here, we focus on the model parameter describing the Julian day within the year at which the potential biomass is greatest, t_{jul}^{max}. It is used in the seasonality term, $\boldsymbol{s}(\boldsymbol{x}, \boldsymbol{\theta})$. From the literature, we know that the seasonal effect on periphyton for systems without other disturbance factors, such as hydrodynamics or grazing, is mainly driven by the seasonal variation of light and temperature (Biggs and Stokseth 1996). For the Swiss rivers originally represented by this model, the temperature maximum occurs at approximately Julian day 170 and the light maximum at Julian day 210 (at locations where shading by riparian vegetation is negligible). Our expert expects the maximum potential biomass to occur slightly after the maximum of light and temperature and estimated precise quantiles for standard cumulative probabilities. Therefore, any ambiguity is only the result of mismatches between the elicited values and the fitted distribution. The *Density Ratio Class* fit to the expert's estimates based on beta distributions in the interval $[0, 365.25]$ as bounding densities is shown in Figure 3.1.

3.5.2 Both uncertainty and ambiguity: Sensitivity of a clam species to low oxygen in an estuarine eutrophication model

Borsuk et al. (2002) constructed a model for the times-to-death of a clam species as a function of dissolved oxygen concentration in an estuary. Points on the cumulative distribution function of times-to-death for four different oxygen concentrations were elicited from two experts jointly. Because of either minor disagreement between the experts or shared imprecision, for many of the elicited quantiles the experts stated intervals rather than points. These intervals were recorded, but not used, in the original study. Rather, Borsuk et al. used the midpoints of the intervals to fit precise distributions for each oxygen concentration.

Rinderknecht et al. (2011) fit a *Density Ratio Class* to the data of Borsuk et al. (2002) using the technique described in section 3.4.2. After a log-transformation, a combination of the Normal density for l and the t_3 density for u provided the best fit, as measured by the lowest average value of κ across the four dissolved oxygen concentrations. Figure 3.2 shows the results for the lowest dissolved oxygen value. The ambiguity in this case is the result of both the fit of the distribution as well as the ambiguity contained in the experts' judgments themselves.

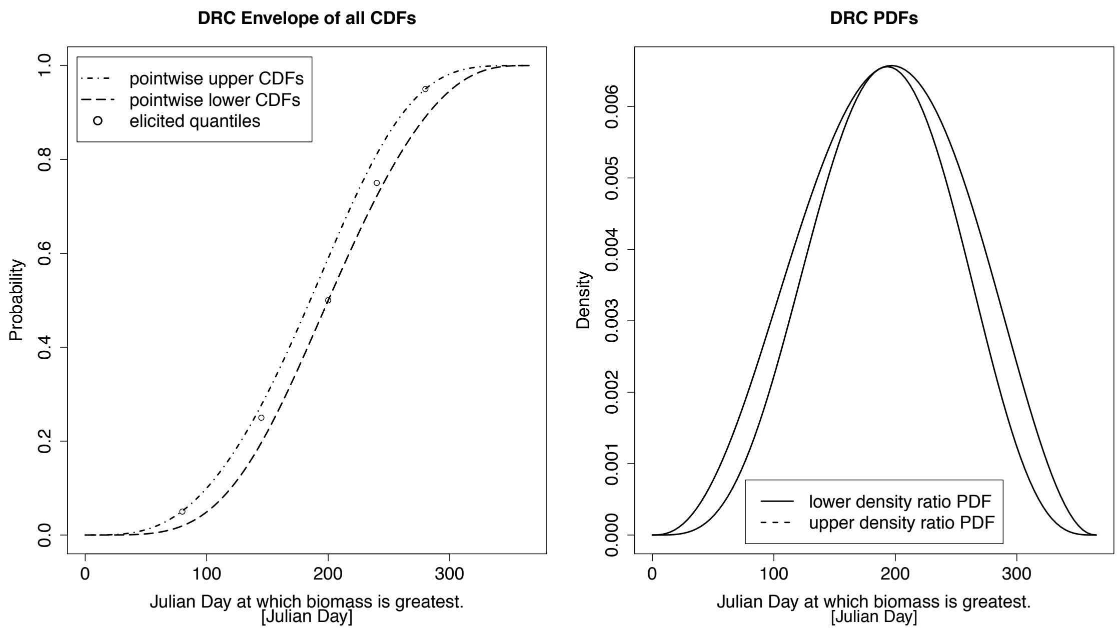

Figure 3.1: *Density Ratio Class*, $\Gamma^{DR}_{f_l,\kappa f_u}$ for the parameter t^{max}_{jul} [Julian day]. Cumulative probabilities are: $p = (0.05, 0.25, 0.5, 0.75, 0.95)$; lower quantile points coincide with upper quantile points: $q_{min} = q_{max} = (80, 145, 200, 240, 280)$. Left panel: quantile intervals with curves following equations (3.6) and (3.7). Right panel: *Density Ratio Class* densities normalized $l = f_l$ and unnormalized $u = \kappa f_u$ defining the class, where $\kappa = 1.20$, f_l: $Beta(4.95, 4.52)$, f_u: $Beta(3.58, 3.21)$ with the ranges $[L, U] = [0, 365.25]$.

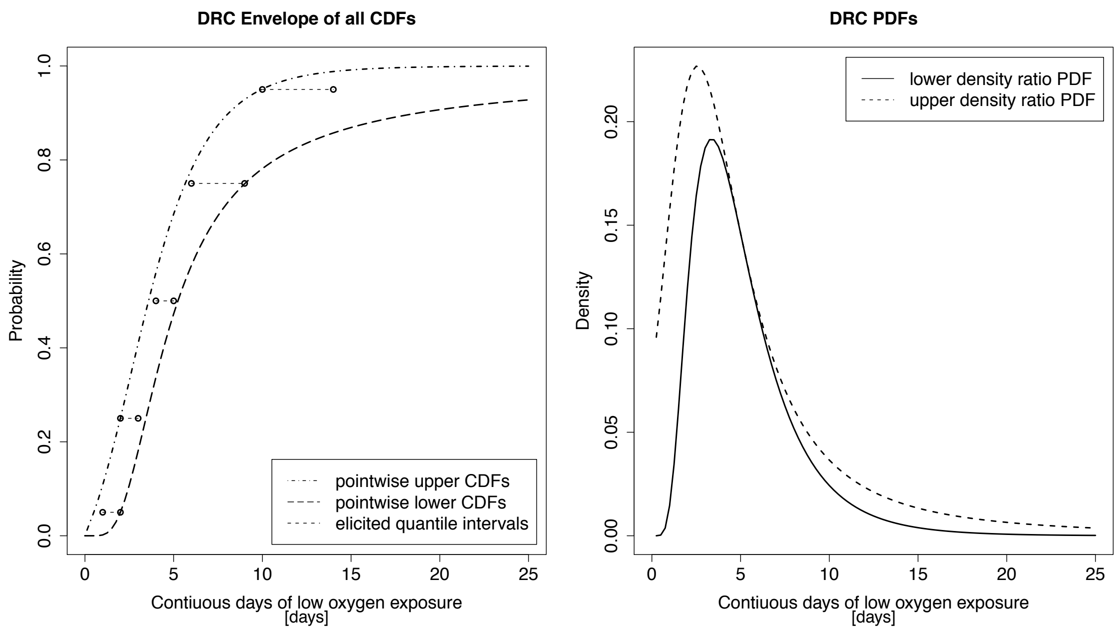

Figure 3.2: A *Density Ratio Class* was fitted to the logarithm of assessed quantile intervals, $q_{min} = ln(1, 2, 4, 6, 10)$ and $q_{max} = ln(2, 3, 5, 9, 14)$, on the number of days corresponding to each specified cumulative mortality $p = (0.05, 0.25, 0.50, 0.75, 0.95)$ and ambient dissolved oxygen concentration according to Borsuk et al. 2002. In this case a Normal density was used for l and a Student t density with three degrees of freedom was used for u. The factor $\kappa = 1.56$ and the distributional parameters are $N_l(\mu_l = 1.50, \sigma_l = 0.54)$, Student-$t_u(\mu_u = 1.44, \sigma_u = 1.31)$. In a second step, the *Density Ratio Class* was back transformed to the natural metric (days) resulting in the following distributional characteristics: the lower distribution has a mean of $\mu_l = 5.21$ days and a standard deviation of $\sigma_l = 3.03$ days, the upper distribution has a mean of $\mu_u = 9.09$ days and a standard deviation of $\sigma_u = 23.3$ days. Note that this transformation does not affect the value of κ. Left and right panels can be interpreted as in Figure 3.1.

3.5.3 Largely ambiguity: Economic damages caused by climate change

Nordhaus' (1994) survey of experts regarding the economic damages of climate change has frequently been used to parameterize the uncertainty of damages in integrated assessment models. However, to our knowledge, all past uses of the data have ignored the inherent ambiguity represented by experts' differing beliefs and by the fact that the survey did not elicit a specific probability distribution.

The elicited damages in the Nordhaus (1994) survey are in the form of loss of global economic output. Experts were asked to consider the uncertainty for two warming scenarios by providing estimates of the 10th, 50th, and 90th percentiles of the cumulative distribution of damages.

We used the data from the 6°C scenario to produce a *Density Ratio Class*. To protect against highly anomalous experts, the 5th and 95th percentiles of elicited values at each point in the cumulative distribution were used as absolute lower and upper bounds.

Since the data are very coarse (the quantile intervals largely overlap) we transformed θ according to $tr(\theta) = -(a/b^2)exp(-b\theta) + c\theta + (a/b^2)$, where a, b, c and the parameters for a lower and upper Gaussian distribution were estimated to minimize κ. The results are given in Figure 3.3. This example largely demonstrates the role of disagreement in generating ambiguity about intersubjective knowledge.

3.5.4 Comparison of Examples

Our three examples differ in the amount of ambiguity present in the elicited data. The values of κ can be used as a direct measure of imprecision for the *Density Ratio Classes* fit to the data. These numerically confirm what can be inferred graphically: the set constructed for example 3 is much less precise than those constructed for the other two examples (Table 1).

Unfortunately, the interpretation of κ is unique to the *Density Ratio Class*, and its value does not provide any detail concerning the ambiguity in specific attributes of the distribution. The metrics we propose give additional insight. For example, we can see that the imprecision about the width of the elicited variable is quite small for example 1 and is comparably large for examples 2 and 3. This is because of the general disagreement between the experts elicited for example 3 and the large difference in tail behavior between the upper and lower bounding densities for example 2. A similar conclusion can be drawn for imprecision about the distribution of the variable within its range; the set constructed for example 1 is the

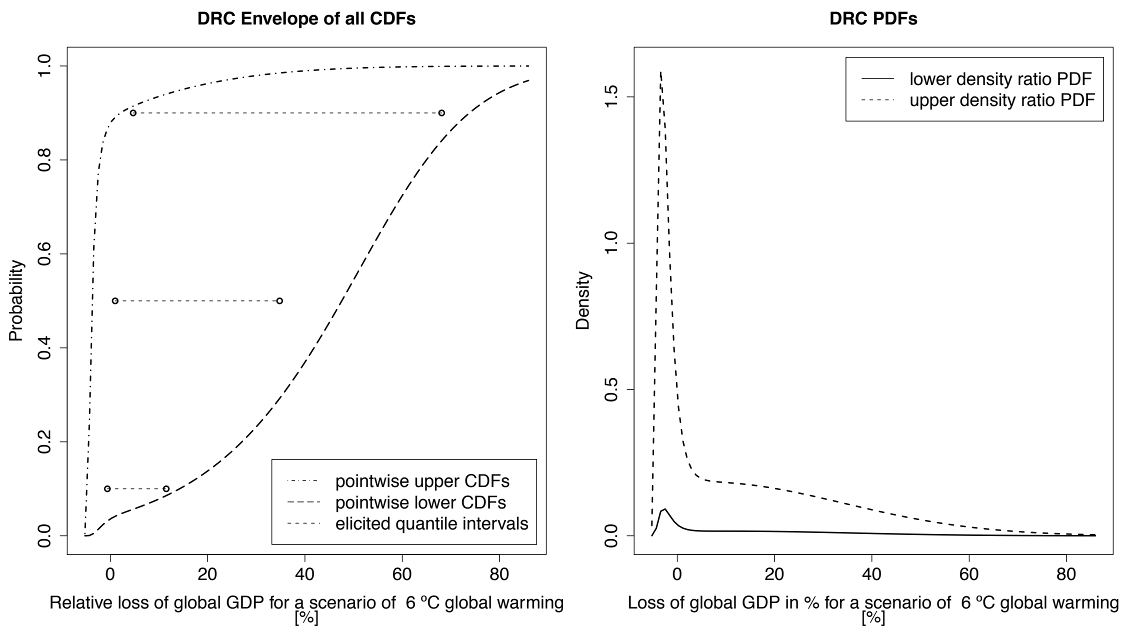

Figure 3.3: The variable θ was transformed according to $tr(\theta) = -(a/b^2)exp(-b\theta) + c\theta + (a/b^2)$ to minimize κ of the fitted *Density Ratio Class* $\Gamma_{N_l,\kappa N_u}$ for given cumulative probabilities, $p = (0.1, 0.5, 0.9)$, lower elicited quantile points, $q_{min} = (-0.6, 1.0, 4.7)$, and upper elicited quantile points, $q_{max} = (11.5, 34.8, 68.1)$. The numerical best estimates for the transformation and distributional parameters are: $a = 10.2$, $b = 0.67$, $c = 8.76$, $N_l(\mu_l = 115.8,\ \sigma_l = 220.3)$, $N_u(\mu_u = 72.7,\ \sigma_u = 249.7)$. This gives a factor $\kappa = 13.12$. The plot shows the back transformed *Density Ratio Class* with the following distributional characteristics: the lower distribution has a mean of $\mu_l = 15.4\%$ GDP and a standard deviation of $\sigma_l = 18.8\%$ GDP, the upper distribution has a mean of $\mu_u = 13.3\%$ GDP and a standard deviation of $\sigma_u = 19.4\%$ GDP. Left and right panels can be interpreted as in Figures 3.1 and 3.2.

Table 3.1: Summary of metrics for the three examples ($\alpha = 0.05$).

Example	κ	$r_{1-\alpha}^{\text{width,DRC}}$	$r_{1-\alpha}^{\text{shape,DRC}}$	$r_{1-\alpha}^{\text{mode,DRC}}$
1	1.20	0.08	0.09	0.02
2	1.56	0.43	0.84	0.03
3	13.12	0.35	0.77	0.47

most tightly constrained, while the set constructed for example 2 is less precise than the one for example 3. This is because the imprecision of example 2 is largely in the tails, rather than the center, of the distribution. Regarding the mode, however, the variable of example 1 is the most precisely defined, while example 3 remains the most imprecise. Clearly, different metrics give complementary information.

3.6 Discussion

We see three issues that require further discussion: (1) a comparison of our metrics of ambiguity relative to others that have been proposed in the literature, (2) the merits of using imprecise probabilities to describe ambiguity in elicitation results, relative to (precise) first and second-order probabilities or fuzzy distributions, and (3) the implications of using imprecise probabilities for environmental decision support. The third point will exemplify the bridging function we see imprecise probabilities serving between ambiguous and uncertain knowledge.

A variety of overall uncertainty measures have been described in the literature under the broad concepts of *probability theory*, *classical set theory*, *fuzzy set theory*, *possibility theory* and *evidence theory* (Klir and Wierman 1998). One may specify, for example, lower and upper bounds of a *(differential) entropy* by extending the concept of *Shannon's entropy* (Shannon 1948) to imprecise probabilities. Alternatively, one may calculate a metric deriving from *Dempster-Shafer theory* (Shafer 1976), such as the *Shafer continuous belief function*. Finally, for a *Density Ratio Class* specifically, the ratio between the surfaces of the upper and lower densities, $\kappa = \int u(\theta)\mathrm{d}\theta / \int l(\theta)\mathrm{d}\theta$, can serve as a reasonable overall measure of uncertainty.

Of course, the selection of a general metric depends on the underlying concept, the final purpose, and the particular case.

Such general measures as those mentioned above are characteristically not informative with regard to specific attributes of distributions (e.g., the mode). Our metrics, in contrast, while applicable to any imprecise probability class are attribute-specific: they characterize the degree of ambiguity in the width, the shape or the mode relative to a particular credibility interval. The dependence on a credibility level is hard to avoid since it acts as a reference for the distributional characteristics. However, from an analyst's perspective, this may be seen as a strength, as one can investigate how ambiguity changes with differing credibility levels.

Since our proposed metrics apply to all imprecise probability classes, different classes that are obtained on the basis of the same elicited data can be compared with regard to the three attributes: width, shape and mode. The most appropriate class can then be chosen according to the specific attribute of concern. Further, one could attempt to find the class that minimizes a chosen metric, while still being compatible with the elicited data. Finally, our attribute-specific metrics allow for deeper insight and a more specific description of the ambiguity represented by an imprecise probability class than a more general metric, as shown by the examples.

If, as we argued, it is reasonable to describe uncertain knowledge by probability distributions, then it also seems natural to describe uncertainty about the parameters of such distributions again by probability distributions. This leads to so-called second-order probabilities or hierarchical probability models (Draper 1995). However, having in mind that second-order uncertainty results e.g. from the problem that an expert has to express her or his beliefs *precisely* in the form of a probability distribution, it does not seem realistic that the knowledge about the second-order distribution (which is an even more abstract concept) can then be expressed precisely. Of course, imprecise probability distributions based on interval cumulative probabilities or quantiles also have this problem since the bounds of the intervals will not be known precisely (Howson and Urbach 1989). However, at least they allow for a wide variety of distributional shapes, and the bounds can be chosen to be on the cautious side in order to analyze the robustness of results to these distributions. O'Hagan (2012) warns that set-based robustness methods may be overcautious by not allowing the expert to distinguish between more or less likely members of the set. We believe that this is a particular concern for liberal set definitions such as p-boxes (Ferson and Hajagos 2004). However, when extreme density functions are excluded using parametric families to define bounds on a *Density Ratio Class*, ambiguity can be represented without necessarily being overcautious (see Rinderknecht et al. (2011) for more discussion of this point). Possibility theory (Zadeh 1978) based on fuzzy sets (Zadeh 1965) is an alternative approach to address the issue of describing uncertain knowledge (see e.g. Page et al. (2012). We prefer to use the probabilistic framework, extended to imprecision, because of its axiomatic foundation, the consistent formulation of conditional beliefs (which are very important e.g. for scenario analyses), and its compatibility with aleatory uncertainty as outlined in the introduction.

A decision-maker who wants to follow basic axioms of rationality should account for uncertainty by choosing the action that maximizes his or her expected utility. This requires assessment of probability distributions for the outcomes of all decision alternatives (Von Neumann and Morgenstern 1944; Savage 1954). When using imprecise rather than precise probability distributions, intervals of expected utilities are obtained instead of point values. For overlapping intervals, there exists no established decision rule to select the most appropriate decision alternative. The size of expected utility intervals depends on: (i) the degree of ambiguity in uncertain quantities and on (ii) the sensitivity of model predictions to these quantities. If the ambiguity and/or sensitivity are small, we may still get non-overlapping expected utility intervals that allow for a unique ranking of decision alternatives. However, in the case of high ambiguity and/or sensitivity, overlapping intervals indicate that there is not sufficient quantitative knowledge available to reach a decision among the considered alternatives. In this way, the quantification of ambiguity by imprecise probabilities makes it possible to bridge between decidable and non-decidable decision situations based on expected utility theory. In such situations decisions may rely on other principles such as the precautionary principle or reversibility of outcomes. This is in contrast to first or second-order probabilities which imply that there is always sufficient knowledge to obtain a unique ranking of decision alternatives.

3.7 Conclusion

Imprecise probabilities allow us to characterize the degree of ambiguity in probability distributions elicited from subject matter experts. Because of the variety of approaches taken to specifying imprecise probabilities, some generic metrics applicable to all approaches are required. Besides overall measures of imprecision of a class of distributions, we are interested in the imprecision of specific attributes of the class. In particular, important attributes are the width, shape and position of the mode of distributions. We propose three corresponding metrics and formulate them more specifically for the *Density Ratio Class*.

In previous analyses, we have found that the size of a *Density Ratio Class* constructed from elicitation data is determined primarily by three aspects: (i) the choice of parametric shapes selected for the bounding non-normalized densities, l and u, (ii) the scatter of the elicited data relative to the parametric shape(s), and (iii) the explicit imprecision, as specified by the expert(s) using quantile or cumulative probability intervals or by accounting for the disagreement across multiple experts. Our examples demonstrate all three of these contributions. The examples also show that, depending on the application, ambiguity can be rather small or very large. This indicates that its separate representation from uncertainty may be useful, in particular when used in decision support to analyze whether it may affect the generation of a unique expected utility ranking.

Information on the imprecision of probability distributions constructed by expert elicitation can be useful for a variety of purposes. For example, the degree of ambiguity or disagreement in *intersubjective knowledge* can be quantified. The modeler can then use this information to decide how much effort to expend on a more careful or extensive elicitation process. Alternatively, a different, or more refined, decomposition of variables might be considered to improve precision. Finally, even when imprecision cannot be practicably reduced, it is valuable to assess the implication for final results by propagating imprecision through the model. We are currently working on establishing methods of propagation for the *Density Ratio Class*.

Software Availability

The example results in section 3.5 were generated using our recently implemented software package for R (Ihaka and Gentleman 1996) that is able to calculate the *Density Ratio Class* for given quantile intervals according to the method described in section 3.4.2 (Rinderknecht, Borsuk, and Reichert 2011). Possible lower and upper densities are the Gaussian, Student-t, Logistic, Gamma, Weibull, F, Beta, Uniform, Log-Normal, Log-Student-t and the Log-Logistic. Additionally, two transformations for the variable θ are implemented: a simple logarithmic transformation, $tr_1(\theta) = log(\theta)$, and a general transformation of the form: $tr_2(\theta) = -(a/b^2)exp(-b\theta) + c\theta + (a/b^2)$ with $a > 0$, $b > 0$ and $c \geq 0$. This transformation has the following qualities: $tr_2(0) = 0$, $tr'_2(\theta) > 0$ and $tr''_2(\theta) < 0 \ \forall \theta \in \Theta$ where $tr''_2(\theta)$ drops to zero as $\theta \to \infty$. These two transformations allow the analyst to find a possibly better fit of the *Density Ratio Class* to given quantile intervals using the implemented densities. Templates from the example sections can be used to implement distributions and transformations that are not implemented in the standard package. The R package is freely downloadable at http://cran.r-project.org/ subject to the terms of agreement.

Acknowledgments

The authors would like to thank all anonymous reviewers for their useful comments and gratefully acknowledge support from the Swiss National Science Foundation. M.E.B. was partially supported by the US EPA through grant #RD-83366601 from the STAR program. This work has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

Chapter 4

The Effect of Imprecise Prior Knowledge on Bayesian Model Parameter Inference and Prediction

Simon L. Rinderknecht[1,2], Carlo Albert[1], Mark E. Borsuk[3], Nele Schuwirth[1], Hans R. Künsch[3] and Peter Reichert[1,2]

[1] Eawag, Swiss Federal Institute of Aquatic Science and Technology,
Dept. of Systems Analysis, Integrated Assessment and Modelling,
CH-8600 Dübendorf, Switzerland
[2] Institute of Biogeochemistry and Pollutant Dynamics (IBP), ETH Zurich,
CH-8092 Zürich, Switzerland
[3] Dartmouth College, Thayer School of Engineering,
Hanover, NH 03755-8000, USA
[4] Seminar for Statistics (SfS), ETH Zurich,
CH-8092 Zürich, Switzerland

Abstract

Prior probability distributions needed for Bayesian model analysis can rarely be formulated precisely based on existing knowledge. The concept of imprecise probabilities is an attempt to characterize this ambiguity in prior beliefs and the effect on model parameter inference and prediction. We elaborate on the *Density Ratio Class* of probability measures, which has important conceptual advantages over alternative classes for describing imprecise knowledge about continuous model quantities. In particular, we show that, under weak regularity conditions, the *Density Ratio Class* is invariant under (i) Bayesian inference, (ii) marginalization, and (iii) propagation through deterministic models. We also show that (iv) predictions of a stochastic model with parameters defined by a *Density Ratio Class* are naturally embedded in a *Density Ratio Class*. These invariance properties are desirable because they allow the process of sequential Bayesian learning and prediction with imprecise probabilities to proceed within a unified framework. They also minimize the computational burden of numerical implementation relative to using precise probabilities. Concepts and numerical methods are exemplified by application to a simple empirical ecological model. This example reveals the care required to select the model quantities bearing a measure of ambiguity, as there is the potential, in some cases, for model predictions to become ineffectually imprecise.

Keywords

intersubjective knowledge; interval probabilities; imprecise probabilities; robust Bayesian analysis; Density Ratio Class; Bayesian inference, marginalization and prediction.

4.1 Introduction

Bayesian statistical inference offers a mathematical framework to describe a learning process by combining prior knowledge with new data (Box and Tiao 1973; de Finetti 1974; Howson and Urbach 1989; Gelman et al. 2003). In this framework, prior knowledge is typically formulated with a precise probability distribution to describe the subjective belief of either an individual expert or the joint belief of several experts about the value of a specified variable or model parameter. In practice such belief statements are often ambiguous (Einhorn and Hogarth 1985; Camerer and Weber 1992). This is particularly the case if intersubjective belief is being expressed that is intended to represent the current state of knowledge of the scientific community (Gillies 1991; Rinderknecht et al. 2012). One way to take this ambiguity into account is to replace a single prior probability distribution by a set of distributions that spans the range of appropriate distributions. Many specifications of such sets of probability distributions over continuous variables, so-called classes of distributions or imprecise probabilities, have been proposed (Ríos Insua and Ruggeri 2000; Walley 1991; Berger 1994; Caselton and Luo 1992; Ríos Insua and Ruggeri 2000, http://www.sipta.org).

Despite this theoretical development, the concept of imprecise probabilities, which leads to a robustification of probability statements, is still very rarely applied. A reason for this may be that it is felt to be too difficult to implement. Difficulties could occur during elicitation, when updating priors with data, or when propagating imprecise distributions through models. To overcome the first of these potential difficulties, we developed an elicitation technique for the *Density Ratio Class*, which we believe to be the most satisfying class of probability distributions from a conceptual point of view (Rinderknecht et al. 2011). This technique was then applied to several case studies, to demonstrate that a wide range of ambiguity can occur in practical applications (Rinderknecht et al. 2012). Finally, in the present paper, we address the remaining potential obstacles by showing how Bayesian inference, marginalization, and model propagation with *Density Ratio Class* priors works and how it can be easily implemented numerically. The ease of these implementations relies on the resourceful definition of the *Density Ratio Class*.

The paper is structured as follows. Section 4.2 is dedicated to methodological development. Subsection 4.2.1 briefly reviews the *Density Ratio Class*. Next, we show in Subsection 4.2.2 how the *Density Ratio Class* can be used for Bayesian inference, in Subsection 4.2.3 how it can be marginalized, and in Subsection 4.2.4 how it can be propagated through a model to quantify prediction uncertainty. Section 4.3 discusses the numerical implementation of these tasks. In Section 4.4 we demonstrate the suitability of the approach through application to a simple empirical river periphyton model. Finally we draw our conclusions in Section 4.5.

4.2 Methods

4.2.1 Formulation of Imprecise Prior Knowledge as a Density Ratio Class

DeRobertis and Hartigan (1981) introduced the *Density Ratio Class* under the name of *Intervals of Measures*, whereas Berger (1990) later called the class the *Density Ratio Class*. Wasserman (1992a) asserted that, under mild regularity conditions, it is the only probability class to be invariant under Bayesian updating and marginalization. Update invariance is an important property, as it allows for the representation of sequential learning within a common framework. This gives an important advantage to the *Density Ratio Class* relative to other representations of imprecise probabilities. The *Density Ratio Class* also has the ability to accommodate a variety of density function shapes, while limiting 'unreasonable' shapes, such as sharp peaks or point masses that might not be deemed reasonable by an expert (Rinderknecht et al. 2011) (depending on the size of the class, weak or even strong oscillations are still possible).

For uncertain continuous parameters $\boldsymbol{\theta} \in M \subset \mathbb{R}^n$, the *Density Ratio Class* with lower bound $l \geq 0$ and upper bound $u \geq l$ is defined as the set of probability density functions

$$\boxed{\Gamma_{l,u}^{DR} := \left\{ \hat{f}(\boldsymbol{\theta}) \;=\; \frac{f(\boldsymbol{\theta})}{\int f(\boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'} \;\middle|\; l(\boldsymbol{\theta}) \leq f(\boldsymbol{\theta}) \leq u(\boldsymbol{\theta}) \;\; \forall \boldsymbol{\theta} \right\},} \tag{4.1}$$

where we assume that $0 < \int l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} \leq \int u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} < \infty$. The non-normalized densities l and u bound the shapes of the non-normalized probability densities in the class. The class then consists of the normalized densities that fulfill these shape restrictions. In this paper, we shall exclude improper densities since we consider their interpretation questionable (Rinderknecht et al. 2011). Note that the *Density Ratio Class* has the following property:

$$\Gamma_{l,u}^{DR} = \Gamma_{\lambda l, \lambda u}^{DR} \;\; \forall \lambda > 0 \; . \tag{4.2}$$

This implies that one of the "non-normalized" densities, l or u, can still be chosen to be normalized.

Following from (4.1), the lower and upper probabilities, $\underline{P}$ and $\overline{P}$, for a random variable characterized by the *Density Ratio Class*, $\Gamma_{l,u}^{\mathrm{DR}}$, to take a value within a subset A of its domain are given by

$$\underline{P}(A) = \inf_{\hat{f} \in \Gamma_{l,u}^{DR}} \int_A \hat{f}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = \frac{\int_A l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}}{\int_A l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} + \int_{A^c} u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}} \tag{4.3}$$

and

$$\overline{P}(A) = \sup_{\hat{f} \in \Gamma_{l,u}^{DR}} \int_A \hat{f}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = \frac{\int_A u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}}{\int_A u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} + \int_{A^c} l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}} \tag{4.4}$$

where A^c is the complement of A. The first of these equations follows as for any $\hat{f} \in \Gamma_{l,u}^{DR}$, $\int_A \hat{f}\,\mathrm{d}\boldsymbol{\theta}$ can be written in the form $\int_A f\,\mathrm{d}\boldsymbol{\theta} / \big(\int_A f\,\mathrm{d}\boldsymbol{\theta} + \int_{A^c} f\,\mathrm{d}\boldsymbol{\theta}\big)$ and $x/(x+y)$ is decreasing in y for fixed $x > 0$ and increasing in x for fixed $y > 0$. Note that the equation is obviously also true if either $\int_A l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = 0$ or $\int_{A^c} u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = 0$, and both integrals cannot be zero because of the condition $\int l\,\mathrm{d}\boldsymbol{\theta} > 0$. The second equation follows analogously.

In the following three Subsections we elaborate important properties about Bayesian inference, marginalization and prediction with the *Density Ratio Class*.

4.2.2 Bayesian Parameter Inference with Density Ratio Class Priors

The first property we discuss is the invariance of the *Density Ratio Class* under Bayesian inference, or *updating*. In Bayesian terminology, the *likelihood function*, $L(\boldsymbol{y}|\boldsymbol{\theta}) = p(\boldsymbol{y}|\boldsymbol{\theta})$, is the probability density of model results, $\boldsymbol{y}$, given the model parameters, $\boldsymbol{\theta}$. For statistical inference, we substitute observations for the argument $\boldsymbol{y}$ and are interested in the dependence of L on the parameters. For this reason, we simplify the notation in the following Sections to $L(\boldsymbol{\theta})$ and do not explicitly indicate the dependence on the observations, $\boldsymbol{y}$, which in the context of inference are assumed to be fixed. We will return to the full notation, $p(\boldsymbol{z}|\boldsymbol{\theta})$, where $\boldsymbol{y}$ is replaced by $\boldsymbol{z}$ to clarify that not observations $\boldsymbol{y}$ are substituted for the argument of the probability density function, in the context of probabilistic prediction in Subsection 4.2.4.

Let $\mathcal{D} = \{f(\boldsymbol{\theta}) \geq 0 \;\forall \boldsymbol{\theta} \in M \subset \mathbb{R}^n \,|\, 0 < \int_M f(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} < \infty\}$ be the set of all not necessarily normalized density functions and let $\mathcal{L}_{\mathcal{D}} = \{L(\boldsymbol{\theta}) \geq 0 \,|\, \int f(\boldsymbol{\theta})L(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} < \infty\,,\; f \in \mathcal{D}\}$. We introduce the operator Φ_L

$$\Phi_L[f](\boldsymbol{\theta}) = f(\boldsymbol{\theta})L(\boldsymbol{\theta}) \tag{4.5}$$

to map a not necessarily normalized prior density function $f \in \mathcal{D}$ to its not normalized posterior according to Bayes law for a likelihood function $L \in \mathcal{L}_{\mathcal{D}}$. A hat on top of an operator indicates composition with the normalization operator. Applied to the operator Φ_L this leads to

$$\hat{\Phi}_L[f](\boldsymbol{\theta}) = \frac{f(\boldsymbol{\theta})L(\boldsymbol{\theta})}{\int f(\boldsymbol{\theta'})L(\boldsymbol{\theta'})\,\mathrm{d}\boldsymbol{\theta'}} \quad . \tag{4.6}$$

Thus, this operator produces the posterior probability density function based on a not necessarily normalized prior, f, and a likelihood function, L. Note that for given L this operator is only defined for functions in $\mathcal{D}$ that fulfill $\int fL\,\mathrm{d}\boldsymbol{\theta} > 0$. Thus we make $\int lL\,\mathrm{d}\boldsymbol{\theta} > 0$ a requirement for the lower bound of the *Density Ratio Class* to be considered a prior class of the parameters used in L.

The first statement reflects the invariance property of the *Density Ratio Class* under Bayesian updating and consists of a constructive description of the posterior class:

$$\boxed{\hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Phi_L[l],\,\Phi_L[u]}} \;, \tag{4.7}$$

which implies that the lower and upper densities of the inferred class can be constructed by applying the operator Φ_L to the lower and upper densities of the prior class.

The proof of property (4.7) has two parts: first,

$$l \le f \le u \;\Rightarrow\; \Phi_L[l] \le \Phi_L[f] \le \Phi_L[u]\,, \tag{4.8}$$

which follows from $L \ge 0$, and second,

$$\Phi_L[l] \le g \le \Phi_L[u] \;\;\Rightarrow\;\; \exists\, f :\; l \le f \le u \text{ and } \Phi_L[f] = g\,. \tag{4.9}$$

For the second part, define the density function f as

$$f(\boldsymbol{\theta}) := \begin{cases} \dfrac{g(\boldsymbol{\theta})}{L(\boldsymbol{\theta})} & \text{if } L(\boldsymbol{\theta}) > 0 \\[2ex] l(\boldsymbol{\theta}) & \text{if } L(\boldsymbol{\theta}) = 0\,. \end{cases} \tag{4.10}$$

It follows then directly that $\hat{f} \in \Gamma^{DR}_{l,u}$ and that $\Phi_L[f] = g$ which completes the proof.

Note that a direct consequence of the definition of Φ_L and equation (4.7) is that the ratio of lower to upper densities is invariant under updating:

$$\frac{l}{u} = \frac{\Phi_L[l]}{\Phi_L[u]} \tag{4.11}$$

(we assume here that the set M on which the densities is defined, is restricted to those values of $\boldsymbol{\theta}$ for which u is strictly positive). Therefore, the set of posteriors does not shrink down to a single distribution, even if a large amount of data is available for inference.

A further property reveals that if one prior-posterior pair, $\{f, \Phi_L[f]\}$, $f > 0$ on $\{\boldsymbol{\theta} | u(\boldsymbol{\theta}) > 0\}$ is known, it is possible to derive its posterior, $\Phi_L[g]$, for any other given prior, g, as:

$$\Phi_L[g] = \frac{g}{f}\,\Phi_L[f]\;. \tag{4.12}$$

This property becomes obvious by applying the definition of the operator Φ_L (4.5). The properties (4.2), (4.7) and (4.12) can be combined to construct the posterior class from the update of a single strictly positive prior f:

$$\boxed{\hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\frac{l}{f}\,\Phi_L[f],\,\frac{u}{f}\,\Phi_L[f]} = \Gamma^{DR}_{\frac{l}{f}\,\hat{\Phi}_L[f],\,\frac{u}{f}\,\hat{\Phi}_L[f]}\;.} \tag{4.13}$$

Multiplication of l/f and u/f with the not necessarily normalized (see equation 4.2) posterior density of f results in the lower and upper densities, respectively, of the posterior *Density Ratio Class*. This is one of the core properties that make the use of imprecise probabilities tractable with a relatively modest increase in the computational burden. Note that choosing $f = u$ in equation (4.13) leads to the useful special case

$$\hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\frac{l}{u}\hat{\Phi}_L[u],\, \hat{\Phi}_L[u]} \,. \tag{4.14}$$

The second form of equation (4.13) and its special case (4.14) are of particular interest as they make it possible to use a normalized posterior which is usually easier to construct numerically (e.g. by applying a kernel density estimator to a posterior MCMC sample) at least if the dimension of the parameter space is not large. An alternative to this equation would be to switch to normalized posteriors directly in equation (4.7)

$$\boxed{\hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{r\,\hat{\Phi}_L[l]\,,\;\hat{\Phi}_L[u]}} \,, \tag{4.15}$$

where equation (4.2) has been used and r is the ratio of the normalizing constants of the posteriors:

$$r := \frac{\int \Phi_L[l](\boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'}{\int \Phi_L[u](\boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'} \,. \tag{4.16}$$

A further discussion about the use of these equations for numerical implementation schemes will be given in Section 4.3 together with a discussion of possibilities of numerically estimating the ratio r (which can be difficult in high dimensions).

4.2.3 Marginalization of the Density Ratio Class

The next properties concern marginalization of the *Density Ratio Class*. For a given nonnegative integrable function $f(\boldsymbol{\theta}) = f(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, we denote the marginalizing operator by

$$\Psi_1[f](\boldsymbol{\theta}_1) := \int f(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)\,\mathrm{d}\boldsymbol{\theta}_2 = f_1(\boldsymbol{\theta}_1) \,. \tag{4.17}$$

The first important property concerns marginalization invariance:

$$\boxed{\Psi_1[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Psi_1[l],\, \Psi_1[u]}} \,. \tag{4.18}$$

The proof of (4.18) has two parts. First,

$$\hat{f} \in \Gamma^{DR}_{l,u} \Rightarrow \Psi_1[\hat{f}] \in \Gamma^{DR}_{\Psi_1[l],\,\Psi_1[u]}\,, \tag{4.19}$$

which follows directly from the definition of the *Density Ratio Class*. Second,

$$\hat{g}_1 \in \Gamma^{DR}_{\Psi_1[l],\,\Psi_1[u]} \Rightarrow \exists\, \hat{f} \in \Gamma^{DR}_{l,u}:\ \Psi_1[\hat{f}] = \hat{g}_1\,, \tag{4.20}$$

which is less obvious to prove. From the left-hand side of equation (4.20) we have by definition of the *Density Ratio Class* a non-normalized g_1 such that $\Psi_1[l] \leq g_1 \leq \Psi_1[u]\ \ \forall\, \boldsymbol{\theta}_1$. Now, define f as follows

$$f(\boldsymbol{\theta}) := \begin{cases} \dfrac{\Psi_1[u]-g_1}{\Psi_1[u]-\Psi_1[l]}\, l \;+\; \dfrac{g_1-\Psi_1[l]}{\Psi_1[u]-\Psi_1[l]}\, u & \text{if } \Psi_1(l) < \Psi_1(u) \\ l & \text{if } \Psi_1(l) = \Psi_1(u)\,. \end{cases} \tag{4.21}$$

Thus, $\Psi_1[f] = g_1$. Consider now l in the form

$$l \;=\; \begin{cases} \dfrac{\Psi_1[u]-g_1}{\Psi_1[u]-\Psi_1[l]}\, l \;+\; \dfrac{g_1-\Psi_1[l]}{\Psi_1[u]-\Psi_1[l]}\, l & \text{if } \Psi_1(l) < \Psi_1(u) \\ l & \text{if } \Psi_1(l) = \Psi_1(u) \end{cases} \tag{4.22}$$

to conclude that $l \leq f$. Similarly, by writing u in a form analogously to equation (4.22) it follows that $f \leq u$ which completes the proof.

We now will concentrate on marginalization of posterior *Density Ratio Classes*. A combination of (4.7) and (4.18) leads directly to the marginalized posterior *Density Ratio Class*:

$$\boxed{\Psi_1 \circ \hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Psi_1 \circ \Phi_L[l],\ \Psi_1 \circ \Phi_L[u]}\quad .} \tag{4.23}$$

Update and marginalization invariance are extremely useful properties of the *Density Ratio Class*, as marginals are much more tractable than high-dimensional distributions and are important tools for communication. The invariance expressed by equation (4.23) allows us to communicate and represent priors, posteriors and posterior marginals all in the same way using *Density Ratio Classes*.

To further prepare for the numerical techniques discussed in Section 4.3, we derive expressions for posterior *Density Ratio Class* marginals based on normalized joint or marginal posterior densities. Applying equation (4.23) to equation (4.13) leads to the following form for the posterior *Density Ratio Class* marginals:

$$\boxed{\Psi_1 \circ \hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Psi_1\left[\frac{l}{f}\,\hat{\Phi}_L[f]\right],\ \Psi_1\left[\frac{u}{f}\,\hat{\Phi}_L[f]\right]}} \quad . \tag{4.24}$$

This equation allows us to derive posterior marginals from the normalized posterior of a single prior, f. In analogy to equation (4.14), the special case $f = u$ of this equation is of particular interest:

$$\Psi_1 \circ \hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Psi_1\left[\frac{l}{u}\,\hat{\Phi}_L[u]\right],\ \Psi_1\left[\hat{\Phi}_L[u]\right]} \quad . \tag{4.25}$$

Applying equation (4.23) to the posterior in the form of equation (4.15) leads to an alternative form for the posterior marginals:

$$\boxed{\Psi_1 \circ \hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{r\,\Psi_1 \circ \hat{\Phi}_L[l],\ \Psi_1 \circ \hat{\Phi}_L[u]}} \tag{4.26}$$

which relies on the ratio, r, of normalizing factors defined by equation (4.16). Again, we refer to Section 4.3 for a discussion of the use of these equations for numerical implementation of the *Density Ratio Class*.

Finally, there is a special case in which it can be avoided to calculate the ratio r of normalizing factors. If there exists a strictly positive function $q(\boldsymbol{\theta}_1)$, such that

$$\frac{l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}{u(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} = q(\boldsymbol{\theta}_1) \tag{4.27}$$

then one can use q to calculate the lower posterior measure from the upper. This can easily be seen by substituting equation (4.27) into equation (4.25):

$$\Psi_1 \circ \hat{\Phi}_L[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{q\,\Psi_1 \circ \hat{\Phi}_L[u]\ ,\ \Psi_1 \circ \hat{\Phi}_L[u]} \tag{4.28}$$

where we assume that u is strictly positive over all $\boldsymbol{\theta}$ (if this is not the case, we can restrict the set M to those values). Note that this applies only to the marginal corresponding to the component(s) $\boldsymbol{\theta}_1$. However, having obtained the lower and upper marginals of this component according to equation (4.28), we can combine equation (4.26) with equation (4.28) to get an estimate of the ratio r of normalizing factors from the relationship

$$r\,\Psi_1 \circ \hat{\Phi}_L[l] = q(\boldsymbol{\theta}_1)\,\Psi_1 \circ \hat{\Phi}_L[u] \quad . \tag{4.29}$$

This estimate can then be used to calculate the other marginal bounding densities using equation (4.26) as we will discuss in detail in Section 4.3. Note that the most distinctive case in which (4.27) is fulfilled is an independent combination of a *Density Ratio Class* prior for the component(s) $\boldsymbol{\theta}_1$ with a precise prior for the component(s) $\boldsymbol{\theta}_2$.

4.2.4 Prediction with the Density Ratio Class

To get an (imprecise) probabilistic description of model predictions based on a prior or posterior in the form of a *Density Ratio Class*, we have to propagate such a class through the model. We will distinguish the cases of a deterministic model and a stochastic model. In Subsection 4.2.4 we will show that under weak regularity conditions the predictions by a deterministic model will again be in the form of a *Density Ratio Class*. This is not anymore true for the predictions of a stochastic model. However, in Subsection 4.2.4 we will show that the predictions of a stochastic model are naturally embedded in a *Density Ratio Class* so that we at least can get conservative bounds of predictions by a *Density Ratio Class*. We will also show that this class can in fact be larger than the set of predicted densities obtained from propagating all elements of the *Density Ratio Class* of model parameters.

Prediction with a Deterministic Model Function

We consider the case of a deterministic model, given by a smooth function

$$\begin{array}{rlcl} \boldsymbol{g}: & M \subset \mathbb{R}^n & \rightarrow & \mathbb{R}^m \\ & \boldsymbol{\theta} & \mapsto & \boldsymbol{y} := \boldsymbol{g}(\boldsymbol{\theta}) \end{array} \tag{4.30}$$

and we show that, under weak regularity conditions, if the model parameters are defined by a *Density Ratio Class* then predictions also have the form of a *Density Ratio Class*.

If $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is a density function of the parameters, and the square root of the determinant of the product of the Jacobian of the model function, $\boldsymbol{g}$, with its transpose

$$\Im_{\boldsymbol{g}}(\boldsymbol{\theta}) = \sqrt{\det\Big(J_{\boldsymbol{g}}(\boldsymbol{\theta}) \cdot J_{\boldsymbol{g}}(\boldsymbol{\theta})^T\Big)} \tag{4.31}$$

is not equal to zero except on a set of Lebesgue measure zero, then, according to the coarea formula of geometric measure theory (cf. formula (5.3.28) in Stroock (1999)), the model predictions, $\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{\Theta})$, have the following density:

$$\Xi_{\boldsymbol{g}}[f](\boldsymbol{y}) := f_{\boldsymbol{Y}}(\boldsymbol{y}) = \int_{\{\boldsymbol{\theta} |\, \boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{y},\, \Im_{\boldsymbol{g}}(\boldsymbol{\theta}) \neq 0\}} \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\Im_{\boldsymbol{g}}(\boldsymbol{\theta})} \, \mathrm{d}\sigma_{\boldsymbol{y}}(\boldsymbol{\theta}) \quad . \tag{4.32}$$

In this equation, $\mathrm{d}\sigma_{\boldsymbol{y}}$ is the surface (Hausdorff) measure on $\{\boldsymbol{\theta} \,|\, \boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{y}\} = \boldsymbol{g}^{-1}(\boldsymbol{y})$ and $\Im_{\boldsymbol{g}}$ is defined in equation (4.31).

The class of model predictions is obtained by applying the transformation given by equation (4.32) to the non-normalized lower and upper measures defining the *Density Ratio Class* of model parameters:

$$\boxed{\Xi_{\boldsymbol{g}}[\Gamma^{DR}_{l,u}] = \Gamma^{DR}_{\Xi_{\boldsymbol{g}}[l],\, \Xi_{\boldsymbol{g}}[u]} \,.} \tag{4.33}$$

By the definition of the *Density Ratio Class* there exists for each random variable $\boldsymbol{\Theta}$ in the class a non-normalized density function $f_{\boldsymbol{\Theta}}$ such that $l_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \leq f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \leq u_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is true for all $\boldsymbol{\theta} \in M \subset \mathbb{R}^n$. By writing this statement in terms of equation (4.32), it follows for all $\boldsymbol{y} \in \mathbb{R}^m$

$$\underbrace{\int_{\boldsymbol{g}^{-1}(\boldsymbol{y})} \frac{l_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\Im_{\boldsymbol{g}}(\boldsymbol{\theta})} \, \mathrm{d}\sigma_{\boldsymbol{y}}(\boldsymbol{\theta})}_{l_{\boldsymbol{Y}}(\boldsymbol{y})} \leq \underbrace{\int_{\boldsymbol{g}^{-1}(\boldsymbol{y})} \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\Im_{\boldsymbol{g}}(\boldsymbol{\theta})} \, \mathrm{d}\sigma_{\boldsymbol{y}}(\boldsymbol{\theta})}_{f_{\boldsymbol{Y}}(\boldsymbol{y})} \leq \underbrace{\int_{\boldsymbol{g}^{-1}(\boldsymbol{y})} \frac{u_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\Im_{\boldsymbol{g}}(\boldsymbol{\theta})} \, \mathrm{d}\sigma_{\boldsymbol{y}}(\boldsymbol{\theta})}_{u_{\boldsymbol{Y}}(\boldsymbol{y})} \,. \tag{4.34}$$

This demonstrates that the class defined by the lower and upper measures $l_{\boldsymbol{Y}}(\boldsymbol{y})$ and $u_{\boldsymbol{Y}}(\boldsymbol{y})$ contains the propagated *Density Ratio Class* of the parameters.

To show the converse, namely, for each $f_{\boldsymbol{Y}}$ satisfying $l_{\boldsymbol{Y}}(\boldsymbol{y}) \leq f_{\boldsymbol{Y}}(\boldsymbol{y}) \leq u_{\boldsymbol{Y}}(\boldsymbol{y})$, we can find a density $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \in \Gamma^{\mathrm{DR}}_{l_{\boldsymbol{\Theta}},u_{\boldsymbol{\Theta}}}$ such that $f_{\boldsymbol{Y}}(\boldsymbol{y}) = \int_{\boldsymbol{g}^{-1}(\boldsymbol{y})} \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\Im_{\boldsymbol{g}}(\boldsymbol{\theta})} \mathrm{d}\sigma_{\boldsymbol{y}}(\boldsymbol{\theta})$, define the convex form

$$f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) := \begin{cases} \dfrac{u_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) - f_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta}))}{u_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) - l_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta}))} l_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) + \dfrac{f_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) - l_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta}))}{u_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) - l_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta}))} u_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) & \\ & \text{if } \; l_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) < u_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) \\ l_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) & \text{if } \; l_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) = u_{\boldsymbol{Y}}(\boldsymbol{g}(\boldsymbol{\theta})) \end{cases} \tag{4.35}$$

Easy calculations show that this function satisfies the required conditions which concludes the proof.

Prediction with a Stochastic Model

We consider the case of a stochastic model that maps each $\boldsymbol{\theta} \in M$ to a random variable $\boldsymbol{Z}$

$$\boldsymbol{\theta} \mapsto \boldsymbol{Z}(\boldsymbol{\theta}) \quad , \tag{4.36}$$

where the respective probability densities of the random variables $\boldsymbol{Z}(\boldsymbol{\theta})$ are given by the likelihood function which we now write in the detailed notation

$$\hat{f}_{\boldsymbol{Z}(\boldsymbol{\theta})}(\boldsymbol{z}) = p(\boldsymbol{z}|\boldsymbol{\theta}) \quad . \tag{4.37}$$

For a given not necessarily normalized density $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ we define the operator Ξ_p to express the propagated density $f_{\boldsymbol{Z}(\boldsymbol{\Theta})}(\boldsymbol{z})$ as follows

$$\Xi_p[f](\boldsymbol{z}) := \int_{M \subset \mathbb{R}^n} p(\boldsymbol{z}|\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \,\mathrm{d}\boldsymbol{\theta} \quad . \tag{4.38}$$

Note that the case of a deterministic model with function $\boldsymbol{g}$ discussed in Subsection 4.2.4 can be viewed as a special case of a stochastic model with a likelihood function $p(\boldsymbol{z}|\boldsymbol{\theta}) = \delta\big(\boldsymbol{z} - \boldsymbol{g}(\boldsymbol{\theta})\big)$. If a not necessarily normalized density $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is bounded by the lower density $l_{\boldsymbol{\Theta}}$ and upper density $u_{\boldsymbol{\Theta}}$ in the sense that the inequality $l_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \leq f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \leq u_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ holds $\forall\, \boldsymbol{\theta}$, we can prove:

$$\boxed{\Xi_p\left[\Gamma^{DR}_{l,\,u}\right] \subseteq \Gamma^{DR}_{\Xi_p[l],\,\Xi_p[u]}\ .} \tag{4.39}$$

For a non-normalized f fulfilling the inequality $l \leq f \leq u$, for all $\boldsymbol{\theta}$, it follows immediately

$$\Xi_p[l](\boldsymbol{z}) \leq \Xi_p[f](\boldsymbol{z}) \leq \Xi_p[u](\boldsymbol{z}) \tag{4.40}$$

for all $\boldsymbol{z}$ which means

$$l_{\boldsymbol{Z}(\boldsymbol{\Theta})}(\boldsymbol{z}) \leq f_{\boldsymbol{Z}(\boldsymbol{\Theta})}(\boldsymbol{z}) \leq u_{\boldsymbol{Z}(\boldsymbol{\Theta})}(\boldsymbol{z}) \tag{4.41}$$

and thus concludes the proof.

A simple one dimensional example, wherein the distribution of Z is independent of θ, already strengthens the intuition that the converse inclusion is not generally true: In that case, $\Xi_p[\Gamma^{DR}_{l,u}]$ consists of a single distribution whereas $\Xi_p[l] = \int l_\Theta(\theta)\,\mathrm{d}\theta$ and $\Xi_p[u] = \int u_\Theta(\theta)\,\mathrm{d}\theta$, so the *Density Ratio Class* that is bounded by $\Xi_p[l]$ and $\Xi_p[u]$ contains infinitely many densities.

We also provide a counterexample with a non-trivial likelihood function. To start, consider the real valued triangular probability density function

$$\hat{f}_X(x) := \begin{cases} 1 - |x| & \text{if } |x| \le 1 \\ 0 & \text{if } |x| > 1 \end{cases} \tag{4.42}$$

and derive its characteristic function, $\phi(\omega)$, defined as the expected value of $e^{i\omega X}$

$$\phi(\omega) = \int_{-\infty}^{\infty} e^{i\omega x}\, \hat{f}_X(x)\,\mathrm{d}x = \frac{\sin^2(\omega/2)}{(\omega/2)^2} = 2\frac{1-\cos(\omega)}{\omega^2} \ge 0\ \forall\, \omega. \tag{4.43}$$

Since $\phi(\omega)$ is the Fourier transform of $\hat{f}_X(x)$, we can get $\hat{f}_X$ back by the inverse Fourier transform

$$\hat{f}_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\phi(\omega)\,\mathrm{d}\omega \quad . \tag{4.44}$$

As a special case of this equation, we get

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\omega)\,\mathrm{d}\omega = \hat{f}_X(0) = 1. \tag{4.45}$$

which implies that $\phi(\omega)/(2\pi)$ is a probability density function.

Define now the likelihood function

$$p(z|\theta) = \phi(z-\theta) \tag{4.46}$$

with $\theta \in [0,1]$ and the *Density Ratio Class*, $\Gamma^{DR}_{1,c}$, for θ with a lower uniform density $l(\theta) = 1$ and an upper uniform density $u(\theta) = c > 1$ restricted to $[0,1]$. Using equation (4.39) we get for the lower and upper densities of the propagated class, $\Xi_p[\Gamma^{DR}_{1,c}] = \Gamma^{DR}_{l_Z, u_Z}$

$$l_Z(z) = \int_0^1 \phi(z-\theta)\cdot 1\,\mathrm{d}\theta \quad , \quad u_Z(z) = \int_0^1 \phi(z-\theta)\cdot c\,\mathrm{d}\theta = c\cdot l_Z(z)\ . \tag{4.47}$$

Note that $l_Z(z)$ is continuous over $\mathbb{R}$ and decays according to c'/z^2 if $z \to \pm\infty$ where c' is a constant. Therefore, we can find a c'' such that $l_Z(z) \le \frac{c''}{1+z^2} \le u_Z(z)$ for c large enough. Or in other words, the Cauchy-Distribution is included in the propagated *Density Ratio Class* $\Xi_p[\Gamma^{DR}_{1,c}]$. To accomplish the proof it is sufficient to show that there exists no function $h(\theta) \in \Gamma^{DR}_{1,c}$ such that $\int_0^1 \phi(z-\theta)\cdot h(\theta)\,\mathrm{d}\theta = \frac{c''}{1+z^2}$. This is obviously the case because the characteristic function of $\int_0^1 \phi(z-\theta)\cdot h(\theta)\,\mathrm{d}\theta$ writes

$$\int_{-\infty}^{\infty} e^{i\omega z}\int_0^1 \phi(z-\theta)h(\theta)\,\mathrm{d}\theta\mathrm{d}z = 2\pi\,\hat{f}_X(\omega)\cdot\int_0^1 e^{i\omega\theta}h(\theta)\,\mathrm{d}\theta \tag{4.48}$$

and hence is equal to zero for all $|\omega| > 1$ — meanwhile the characteristic function of the Cauchy-Distribution is equal to $e^{-|\omega|}$ that clearly is greater than $0 \, \forall \, \omega$. With this contradiction we have shown that the *Density Ratio Class* is not invariant if propagated through a stochastic model.

4.3 Numerical Implementation

4.3.1 Numerical Implementation of Bayesian Inference

Markov chain Monte Carlo (MCMC) techniques allow us to generate a sample of the posterior of any prior, $\hat{f} \in \Gamma_{l,u}^{DR}$ (for a standard MCMC procedure using e.g. the Metropolis sampling technique, see Metropolis et al. (1953), Gelman et al. (2003) or Gamerman and Lopes (2006)). Hence, according to equation (4.12) (note the importance of the condition $f > 0$), assigning such a sample, $\{\boldsymbol{\theta}_i\}_{i=1}^{N}$, with weights

$$w_i = \frac{\dfrac{g(\boldsymbol{\theta}_i)}{f(\boldsymbol{\theta}_i)}}{\displaystyle\sum_{j=1}^{N} \frac{g(\boldsymbol{\theta}_j)}{f(\boldsymbol{\theta}_j)}} \tag{4.49}$$

makes it a sample of the posterior of any other prior, g. Note that the $\{w_i\}_{i=1}^{N}$ may be very unbalanced and that the effective sample size decreases with increasing difference between the densities f and g.

According to equation (4.7) the set of all posteriors of the priors belonging to a *Density Ratio Class* is again a *Density Ratio Class*. We propose two methods to obtain a numerical approximation to this posterior *Density Ratio Class*: The first method is based on a sample of the posterior corresponding to a single prior and constructs the class using weights, whereas the second method is based on samples of the posteriors of the upper and lower density and uses the ratio (4.16) to construct the class.

The first method consists of constructing a Markov chain of the posterior corresponding to a single prior, f, satisfying $f(\boldsymbol{\theta}) > 0$ where $u(\boldsymbol{\theta}) > 0$, and calculating in a second step weighted samples of the posteriors of l and u applying (4.49) with $g = u$ and $g = l$, respectively (note that an obvious choice will be to propagate the upper measure $f = u$). These samples can be transformed to unweighted samples by resampling. As will be shown in Section 4.3.2, the non-normalized ratios l/f and u/f can then be used to construct the lower and upper marginals defining the posterior class. The weights in (4.49) for $g = l$ and $f = u$ will be unbalanced if $l(\theta)$

is close to $u(\theta)$ for some value θ, but l is much smaller than u near the mode of the θ_i's. On the other hand, if u is a multiple of l, then the weights are constant.

The second method consists of constructing Markov chains of the posteriors corresponding to the lower and upper bounds, l and u, of the *Density Ratio Class*. This leads directly to unweighted samples of these two posteriors. According to equation (4.15) to construct the posterior *Density Ratio Class*, we need additionally the ratio r of normalizing factors which can be approximated using a sample of the posterior of the prior corresponding to u (Gelman and Meng 1998). If $\{\boldsymbol{\theta}_i\}_{i=1}^N$ is such a sample:

$$r = \frac{\int \Phi_L[l]\,\mathrm{d}\boldsymbol{\theta}'}{\int \Phi_L[u]\,\mathrm{d}\boldsymbol{\theta}'} = \frac{\int \frac{l}{u}\,uL\,\mathrm{d}\boldsymbol{\theta}'}{\int uL\,\mathrm{d}\boldsymbol{\theta}'} = \int \frac{l}{u}\,\hat{\Phi}_L[u]\,\mathrm{d}\boldsymbol{\theta}' \approx \frac{1}{N}\sum_{i=1}^{N}\frac{l(\boldsymbol{\theta}_i)}{u(\boldsymbol{\theta}_i)}\ . \tag{4.50}$$

Again, the error in this estimate might be larger than the sample size suggests if the weights are unbalanced.

4.3.2 Numerical Implementation of Marginalization

Posteriors are often visualized by their low-dimensional marginals. According to (4.18), the set of posterior marginals is again a *Density Ratio Class*. Applying a kernel density estimator to a component of the samples of the posteriors of l and u leads to the normalized marginal densities

$$\Psi_1 \circ \hat{\Phi}_L[l]\ ,\ \Psi_1 \circ \hat{\Phi}_L[u] \quad . \tag{4.51}$$

Following the first method we generate a sample from a posterior of a single prior, f, $\{\boldsymbol{\theta}_i\}_{1=i}^N$, and obtain from (4.24) the approximation for the (non-normalized) lower bound of the marginal posterior *Density Ratio Class*

$$\Psi_1\left[\frac{l}{f}\hat{\Phi}_L[f]\right](\theta^1) \approx \frac{1}{Nh}\sum_{i=1}^{N}\frac{l(\boldsymbol{\theta}_i)}{f(\boldsymbol{\theta}_i)}K((\theta^1-\theta_i^1)/h) \tag{4.52}$$

where K is a kernel and h is a bandwidth. An analogous formula holds for the upper bound.

Following the second method, we generate two samples $\{\boldsymbol{\theta}_{i,l}\}_{i=1}^N$ and $\{\boldsymbol{\theta}_{i,u}\}_{i=1}^M$ from the posteriors corresponding to l and u respectively, and obtain from (4.26)

$$\Psi_1\big[\hat{\Phi}_L[u]\big](\theta^1) \approx \frac{1}{Mh}\sum_{i=1}^{M}K((\theta^1-\theta_{i,u}^1)/h) \tag{4.53}$$

a similar formula for l instead of u. The class can then be constructed by multiplying the lower normalized bound by the numerical estimate of r obtained from (4.50).

In the special case of a prior class fulfilling condition (4.27), we do not need the approximation of the ratio of the normalizing factors (4.50) but can directly apply equation (4.28) to calculate the marginal posterior class of the first component from a single sample of the posterior corresponding to the prior proportional to u. By seeking for the ratio r with the help of the over-determined system of linear equations given by equation (4.29),

$$r = q(\boldsymbol{\theta}_1)\, \frac{\Psi_1 \circ \hat{\Phi}_L[u]}{\Psi_1 \circ \hat{\Phi}_L[l]}, \tag{4.54}$$

we get an estimate of the ratio r that can be used to calculate the posterior classes of the other marginals. This can either be done by averaging or by fitting the ratio, r, of normalizing factors based on equation (4.54). If (4.27) holds, we can avoid the use of weights using the second method and equation (4.54). Otherwise the advantage of the second method seems to be small.

4.3.3 Numerical Implementation of Model Prediction

Given model parameters in the form of a *Density Ratio Class*, $\Gamma^{\mathrm{DR}}_{l_{\boldsymbol{\Theta}},u_{\boldsymbol{\Theta}}}$, we propagate a sample of the lower density, $\{\boldsymbol{\theta}_{i,l}\}_{i=1}^{N}$, and a sample of the upper density, $\{\boldsymbol{\theta}_{i,u}\}_{j=1}^{M}$, through the model, by drawing from the likelihood, which yields two samples of the respective normalized densities of predictions

$$\hat{\Xi}_{\cdot}[l_{\boldsymbol{\Theta}}],\ \hat{\Xi}_{\cdot}[u_{\boldsymbol{\Theta}}]. \tag{4.55}$$

Note that this is independent of whether the model is probabilistic or not.

To specify the *Density Ratio Class* embodying the entire set of distributions of predictions, reconsider it in the form of $\Gamma^{\mathrm{DR}}_{r'\,\hat{\Xi}_L[l_{\boldsymbol{\Theta}}],\hat{\Xi}_L[u_{\boldsymbol{\Theta}}]}$, analogous to the *Density Ratio Class* representation we used in (4.15), where r' is the ratio of normalizing factors

$$r' := \frac{\int \Xi_L[l](\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}}{\int \Xi_L[u](\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}} = \frac{\int l(\boldsymbol{\theta}) \int L(\boldsymbol{y}|\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\theta}}{\int u(\boldsymbol{\theta}) \int L(\boldsymbol{y}|\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\theta}} = \frac{\int l(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}}{\int u(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}} \tag{4.56}$$

that in turn can be numerically approximated similarly to (4.50). However, it is worthwhile noting that, if posteriors are propagated through the model, $r' = r$ as defined in (4.16). In other words:

$$\Xi_{\cdot}\left[\Gamma^{\mathrm{DR}}_{\Phi_L[l],\,\Phi_L[u]}\right] \subseteq \Gamma^{\mathrm{DR}}_{r\,\hat{\Xi}_p\left[\Phi_L[l]\right],\,\hat{\Xi}_p\left[\Phi_L[u]\right]}, \tag{4.57}$$

where l and u denote in this case the bounds of a prior *Density Ratio Class*.

Following the first method, non-normalized upper and lower bounds of the prediction *Density Ratio Class* are estimated from a sample of the posterior of any prior f. The lower bounding density is given by

$$\Xi_p\big[\Phi_L[l]\big](\boldsymbol{z}) \approx \frac{1}{N}\sum_{i=1}^{N}\frac{l(\boldsymbol{\theta}_i)}{f(\boldsymbol{\theta}_i)}p(\boldsymbol{z}|\boldsymbol{\theta}_i), \tag{4.58}$$

the upper by the same expression with substituting u for l.

Following the second method the normalized upper and lower bounds are estimated from two samples of posteriors for u and l, respectively, according to

$$\Xi_p\big[\hat{\Phi}_L[l]\big](\boldsymbol{z}) \approx \frac{1}{N}\sum_{i=1}^{N}p(\boldsymbol{z}|\boldsymbol{\theta}_{i,l}) \tag{4.59}$$

and the same expression with substituting u for l. In the latter case the ratio r needs to be estimated according to equations (4.50) or (4.54) and multiplied with the normalized lower bound to construct the class.

4.4 Application to a Simple Periphyton Model

We use a simple, empirical model of river periphyton dynamics to illustrate the suggested techniques. Periphyton is a part of benthic biofilms that consist of algae, bacteria, fungi, protozoa and polysaccharides and detritus attached to submerged surfaces in most aquatic ecosystems (Uehlinger 1991). It acts as a primary producer in running water, dominating the first levels of the trophic pyramid in many small and intermediate size rivers (Biggs 1994). Our periphyton model consists of a deterministic, algebraic model that describes the loss and recovery of periphyton in rivers based on the time of year, the time since the last flood, and typical values of flow velocity, water depth and median gravel size. Schweizer (2007) developed this model to describe periphyton dynamics in streams based on data sets that were collected in several previous studies at the Swiss rivers Necker (four sites) and Sihl (two sites), and the French river Garonne (two sites). Table 4.1 contains the references to the literature describing the corresponding surveys and investigations, including descriptions of the study sites, the sampling procedures and biomass measurements.

The model was chosen as an example for this paper because its empirical nature makes it difficult to establish unambiguous priors and its simple form makes it easy to run long Markov chains.

Table 4.1: Literature about the data sets used for the periphyton model. The letters A-H refer to the plots of predictions in Figures 4.2, 4.4 and 4.6.

Study Site	Author(s)	Panel in Figures
Sihl, upstream	Elber et al. (1992)	A
Sihl, downstream	Elber et al. (1996)	B
Necker at Achsäge, main channel	Uehlinger (1991)	C
Necker at Achsäge, gravel bar		D
Necker, side channel	Uehlinger et al. (1996)	E
Necker, downstream		F
Garonne at Gagnac	Boulêtreau et al. (2006)	G
Garonne at Aouach		H

4.4.1 Model Formulation

The Deterministic Part of the Model

The deterministic part of the model consists of a self-limiting Monod factor, a factor describing the limiting influence of stream velocity, water depth and gravel size, and a seasonality factor (Schweizer 2007):

(4.60)

$$
\begin{array}{llcll}
 & B_{\text{modelled}} & = & B_{\Delta t_{\text{flood}}} \cdot L_{v,h,d_{50}} \cdot S_{t_{jul}} & \text{modeled quantity} \\
\text{I.} & B_{\Delta t_{\text{flood}}} & = & \dfrac{\overline{B}_{max}\, \overline{k}_B\, \Delta t_{\text{flood}}}{\overline{B}_{max} + \overline{k}_B\, \Delta t_{\text{flood}}} & \text{Monod factor} \\
\text{II.} & L_{v,h,d_{50}} & = & e^{-\delta v}\, e^{-\gamma h}\, \dfrac{d_{50}}{k_{d_{50}} + d_{50}} & \text{limiting factor} \\
\text{III.} & S_{t_{jul}} & = & \max\left\{1 + \alpha\, cos\left(2\pi\, \dfrac{t_{jul} - t_{jul}^{max}}{365.25}\right), 0\right\} & \text{seasonality factor}
\end{array}
$$

A brief overview of the model parameters is given in Table 4.2. For times shortly after a flood, this model behaves as

$$
B_{\text{modelled}} \approx \overline{k}_B \cdot \Delta t_{\text{flood}} \cdot L_{v,h,d_{50}} \cdot S_{t_{jul}} \quad \text{if} \quad \Delta t_{\text{flood}} << \left(\frac{\overline{B}_{max}}{\overline{k}_B}\right) \tag{4.61}
$$

and for long times after the last flood, the model asymptotically approaches a biomass that depends only on the limiting factor $L_{v,h,d_{50}}$ and the seasonality factor $S_{t_{jul}}$

$$B_{\text{modelled}} \approx \overline{B}_{max} \cdot L_{v,h,d_{50}} \cdot S_{t_{jul}} \qquad \text{if} \quad \Delta t_{\text{flood}} >> \left(\frac{\overline{B}_{max}}{\overline{k}_B} \right) . \tag{4.62}$$

The Likelihood Function of the Model

As in Schweizer (2007), we used a normally distributed error model that is additive on a Box-Cox transformed scale (Box and Cox 1964) to address the heteroscedasticity of the results. This transformation is given by $((\text{data} + \lambda_2)^{\lambda_1} - 1)/\lambda_1$ where $\lambda_1 = 0.3$ and $\lambda_2 = 1$ gAFDM/m^2. The standard deviation of the additive error model was included as a parameter in the Bayesian inference process.

4.4.2 Prior Information about Model Parameters

In the case of our periphyton model, most parameters represent unobservable quantities. Therefore, to develop our prior we elicited the knowledge of an expert in mechanistic benthos modeling. Quantile intervals $\{[q_1^i, q_u^i]\}_{i=1}^5$ for probabilities $\{p_i\}_{i=1}^5 = \{0.05, 0.25, 0.5, 0.75, 0.95\}$ were elicited using the classic quantile elicitation method (Wallsten and Budescu 1983; Cooke 2001; Chaloner 1996; Kadane and Wolfson 1998; Garthwaite et al. 2005; Jenkinson 2005; O'Hagan et al. 2006), allowing the expert to answer using intervals, not only point values. These elicitation results are presented in Table 4.3. Further discussions about the appropriate distributional family for each parameter were conducted with the expert such that *Density Ratio Classes* could be deduced according to the method described by (Rinderknecht et al. 2011). The R software package `fitDRC` was used to support the elicitation. This package is freely available from http://cran.r-project.org/.

4.4.3 Bayesian Inference, Marginalization, and Prediction with Three Different Priors

Bayesian inference was based on Markov chain Monte Carlo using the Metropolis algorithm (Gelman et al. 2003; Gamerman and Lopes 2006) and the techniques described in Section 4.3. In what follows, we present three different cases showing the effect of imprecise prior knowledge on parameter estimates and model predictions.

Table 4.2: Overview of Model Parameters and Inputs.

Parameter	Description	Unit	use in model
$\overline{B}_{max}$	maximum (with respect to h, v and d_{50}) and mean (with respect to seasonality) asymptotic biomass after the last flood occurred	[gAFDM m^{-2}]	parameter
$\overline{k}_B$	maximum (with respect to h, v and d_{50}) and mean (with respect to seasonality) coefficient describing biomass (immediately) after the last flood occurred	[g m^{-2} d^{-1}]	parameter
t_{jul}^{max}	day within the year at which standing crop would be maximum for constant values of the other influence factors	[Julian Days]	parameter
$k_{d_{50}}$	grain size with half saturation for $\overline{k}_B$ and $\overline{B}_{max}$	[m]	parameter
γ	coefficient for limitation by h	[m^{-1}]	parameter
δ	coefficient for limitation by v	[m s^{-1}]	parameter
α	relative amplitude of the seasonal variation (relative to the mean)	[-]	parameter parameter
σ_{err}	standard deviation of the error model of observations	[$-$]	parameter
h	mean water depth	[m]	input
v	stream velocity	[m s^{-1}]	input
d_{50}	grain size (median diameter)	[m]	input
t_{jul}	julian day	[Julian Days]	input
Δt_{flood}	time after the last bed-moving flood	[Days]	input

Table 4.3: Elicited quantile intervals for standard probabilities $\{p_i\}_{i=1}^5 = [0.05, 0.25, 0.50, 0.75, 0.95]$.

name	q_1^l	q_1^u	q_2^l	q_2^u	q_3^l	q_3^u	q_4^l	q_4^u	q_5^l	q_5^u
B_{max}	20.0	60.0	50.0	100.0	100.0	140.0	160.0	210.0	300.0	400.0
k_B	0.75	1.0	2.0	3.0	4.0	5.0	7.0	10.0	18.0	25.0
t_{jul}^{max}	70.0	90.0	135.0	155.0	190.0	210.0	230.0	250.0	270.0	290.0
$k_{d_{50}}$	0.02	0.03	0.05	0.075	0.08	0.125	0.175	0.25	0.38	0.625
γ	0.3	0.7	0.9	1.5	1.0	3.0	3.0	4.5	7.0	20.0
δ	0.3	0.5	0.6	1.3	1.3	4.0	2.0	4.0	5.0	10.0

Table 4.4: Distributions fitted to midpoints of quantile intervals given in Table 4.3.

name	distribution and parameters
B_{max}	$LN(\mu = 147.38,\ \sigma = 108.38)$
k_B	$LN(\mu = 7.07,\ \sigma = 8.56)$
t_{jul}^{max}	$\beta(sh_1 = 2.45,\ sh_2 = 5.04)$ range $[0, 365.25]$
$k_{d_{50}}$	$LN(\mu = 0.169,\ \sigma = 0.195)$
γ	$LN(\mu = 3.81,\ \sigma = 7.02)$
δ	$LN(\mu = 2.74,\ \sigma = 2.61)$
α	$Unif(min = 0,\ max = 1)$
σ_{err}	$LN(\mu = 1.6,\ \sigma = 0.5)$

Table 4.5: Fitted *Density Ratio Classes* based on quantile intervals given in Table 4.3. The value κ expresses the ratio of normalizing constant of the upper distribution divided by the normalizing constant of the lower distribution.

name	distribution and parameters	κ
B_{max}	$LN_l(\mu = 161.57,\ \sigma = 97.91)$ $LN_u(\mu = 127.81,\ \sigma = 122.13)$	1.83
k_B	$LN_l(\mu = 6.246,\ \sigma = 5.634)$ $LN_u(\mu = 7.435,\ \sigma = 10.023)$	1.32
t_{jul}^{max}	$\beta_l(sh_1 = 5.476,\ sh_2 = 5.188)$ $\beta_u(sh_1 = 3.048,\ sh_2 = 2.566)$ range $[0, 365.25]$	1.44
$k_{d_{50}}$	$LN_l(\mu = 0.162,\ \sigma = 0.152)$ $LN_u(\mu = 0.17,\ \sigma = 0.27)$	1.51
γ	$LN_l(\mu = 3.279,\ \sigma = 3.228)$ $LN_u(\mu = 4.456,\ \sigma = 13.767)$	2.03
δ	$LN_l(\mu = 1.923,\ \sigma = 1.179)$ $LN_u(\mu = 2.583,\ \sigma = 3.483)$	1.81
α	$Unif(min = 0,\ max = 1)$ $Unif(min = 0, max = 1)$	1
σ_{err}	$LN_l(\mu = 1.6,\ \sigma = 0.5)$ $LN_u(\mu = 1.6,\ \sigma = 0.5)$	1

Precise Priors in All Marginals

As a first example, we used precise prior distributions in all marginals for Bayesian inference. Therefore, midpoints of the elicited quantile intervals were used. For each parameter, inverse cumulative distributions were fitted to these quantile points using non-linear least squares (Table 4.4). An independent combination of the fitted distributions results in a precise multivariate prior distribution. Figure 4.1 shows marginalized priors and posteriors for each component and Figure 4.2 shows the associated model predictions resulting from the precise posterior distribution.

Ambiguity in One Marginal

As a second example, we used an imprecise prior consisting of an independent combination of precise priors for all parameters listed in Table 4.4, except for the parameter t_{jul}^{max} which was in the form of a *Density Ratio Class*. The *Density Ratio*

Class of the imprecise parameter t_{jul}^{max}, see Table 4.5, was calculated based on the elicited quantile intervals summarized in Table 4.3. This prior fulfills the simplifying assumption given by equation (4.27) and thus makes it possible to derive the ratio of normalizing constants, r, from just one component, as described in Subsections 4.2.2 and 4.3.1. Figure 4.3 shows the corresponding priors and posteriors. By using equation (4.54), the ratio of normalizing constants is $r \approx 0.786$. By applying the numerical approximation given in (4.50) one gets $r \approx 0.785$ which is insignificantly smaller. The *Density Ratio Class* of the predictions is shown in Figure 4.4.

Ambiguity in All Marginals

As a third example, Table 4.5 lists the marginal *Density Ratio Classes* which were derived on the basis of the elicited quantile intervals given in Table 4.3. An independent combination of the upper and lower bounds of these one-dimensional *Density Ratio Classes* leads to a multivariate *Density Ratio Class*. It is not difficult to see that this is the smallest class which contains all product densities of members of the univariate classes. But marginalizing this class produces new classes which are larger than the univariate ones we started with. Figure 4.5 shows the corresponding priors and posteriors, and Figure 4.6 shows the corresponding plot of model predictions. We have used both methods from Section 4.3. Several calculations with different sample sizes have shown that with samples of size 800,000 we obtain estimates that contain enough information for our purposes although there remains still some uncertainty e.g. about the exact value of the ratio of normalizing constants. The *Density Ratio Class* of the priors and posteriors has a ratio of normalizing constants $r \approx 0.00108$ as calculated using equation (4.50). The inverse $r^{-1} = \kappa \approx 920$ even more clearly indicates the high ambiguity in this case.

4.4.4 Results and Discussion

As the model used in our example is largely empirical, there is little prior knowledge available concerning the parameter values. As a consequence, prior marginal distributions were rather wide, yet these narrowed substantially after Bayesian updating using measured data. This was the case independently of using a precise or an imprecise prior (see Figures 4.1, 4.3 and 4.5). However, taking into account more and more of the ambiguity expressed by the expert, makes these posterior classes larger and larger as seen by the increasing difference between lower and upper bounds of the marginal posterior classes (see Figures 4.1, 4.3 and 4.5). In the case of considering ambiguity in all parameters, this size becomes so large, that the lower measure of the marginal posterior class is hardly distinguishable from zero on a scale that is adjusted to display the upper measure (see Figure 4.5).

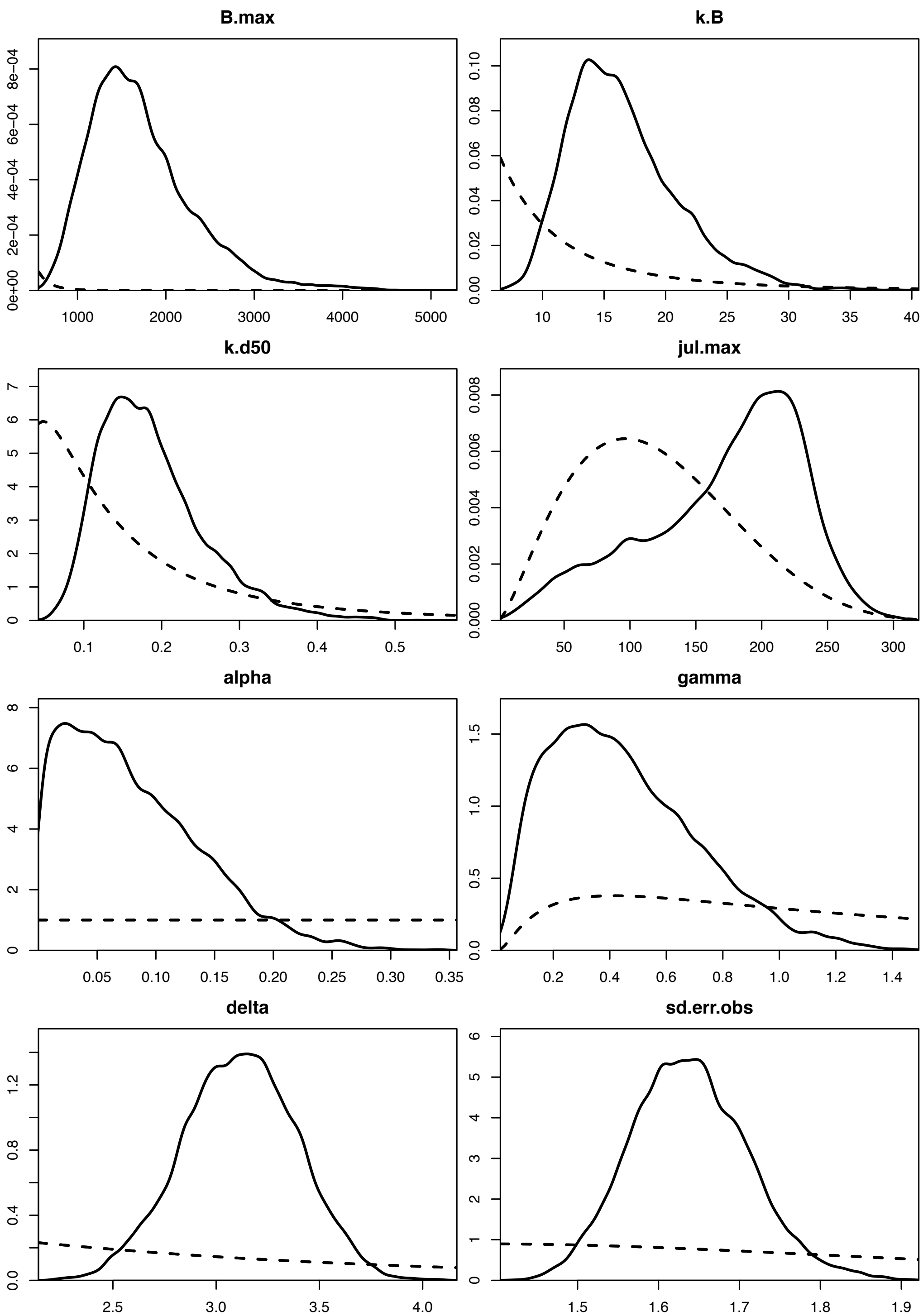

Figure 4.1: Marginalized posteriors (lines) after the Bayesian parameter inference with a precise prior (dotted) corresponding to the example discussed in 4.4.3.

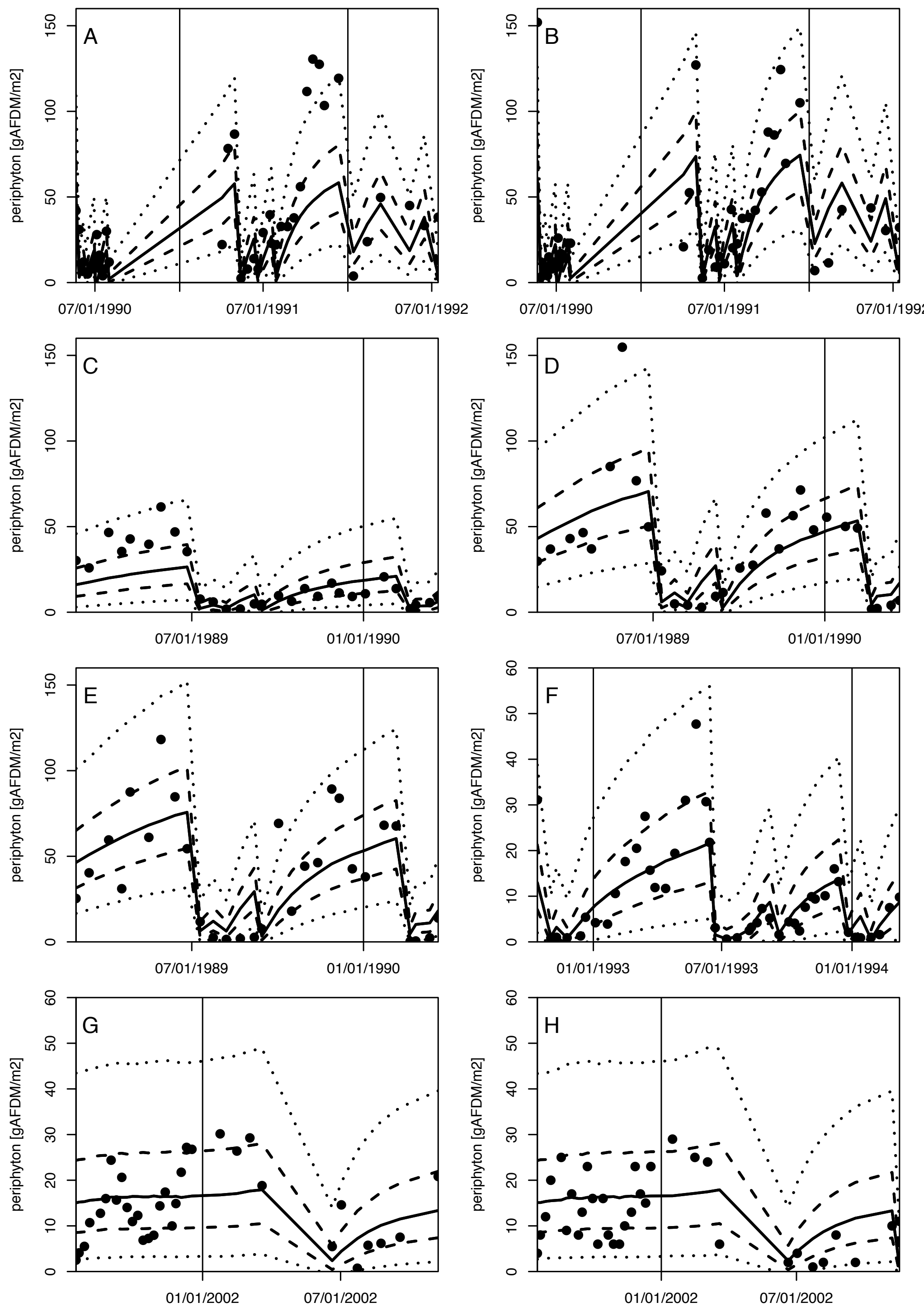

Figure 4.2: Measurements (dots) of periphyton biomass in gram ash free dry mass per square meter [gAFDM/m^2] and model predictions based on a precise prior showing quantiles (lines) for standard probabilities at each time. Each row corresponds to the entries of Table 1, respectively. The vertical lines indicate the beginning of a year.

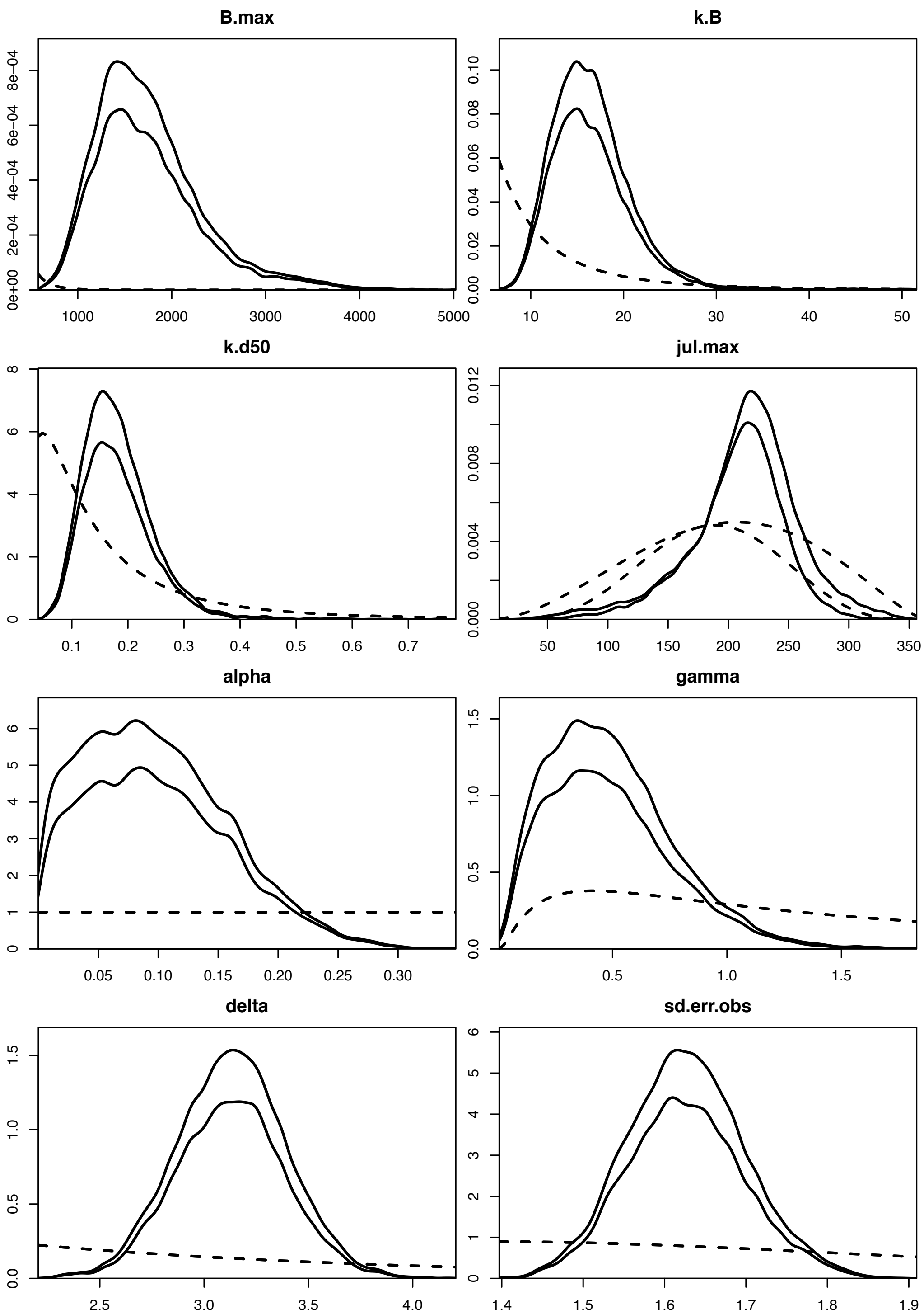

Figure 4.3: Marginalized lower and upper posteriors (lines) of the Bayesian parameter inference for the example in which only the prior of the parameter t_{jul}^{max} was imprecisely defined in the form of a *Density Ratio Class*. The prior marginals are plotted as dotted lines. The fitted ratio of normalizing constants used for the inference was $r \approx 0.837$.

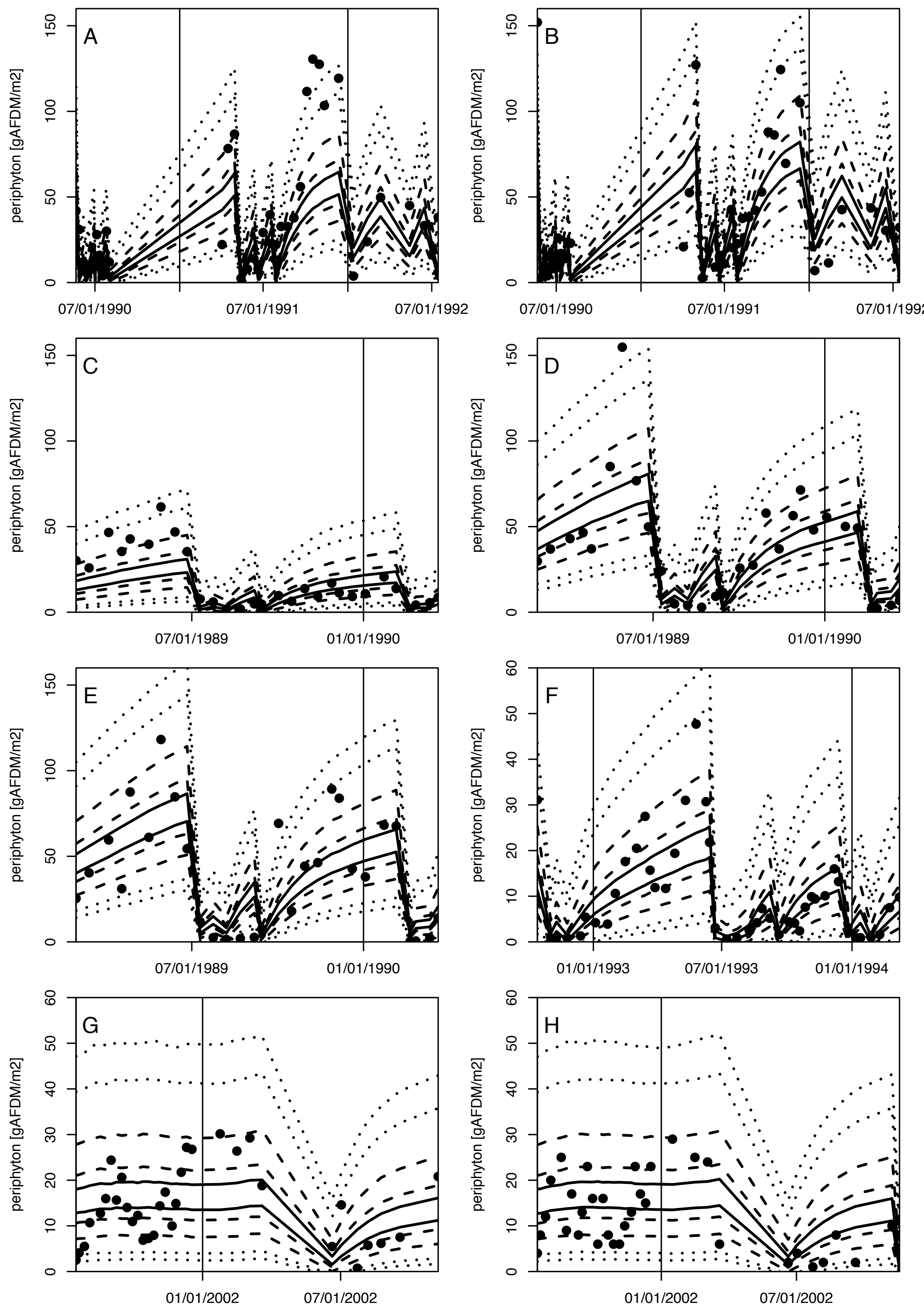

Figure 4.4: Imprecise predictions based on the posterior shown in Figure 4.3. Lower and upper quantiles for standard probabilities (lines) indicate the imprecise model prediction. The vertical lines indicate the beginning of a year.

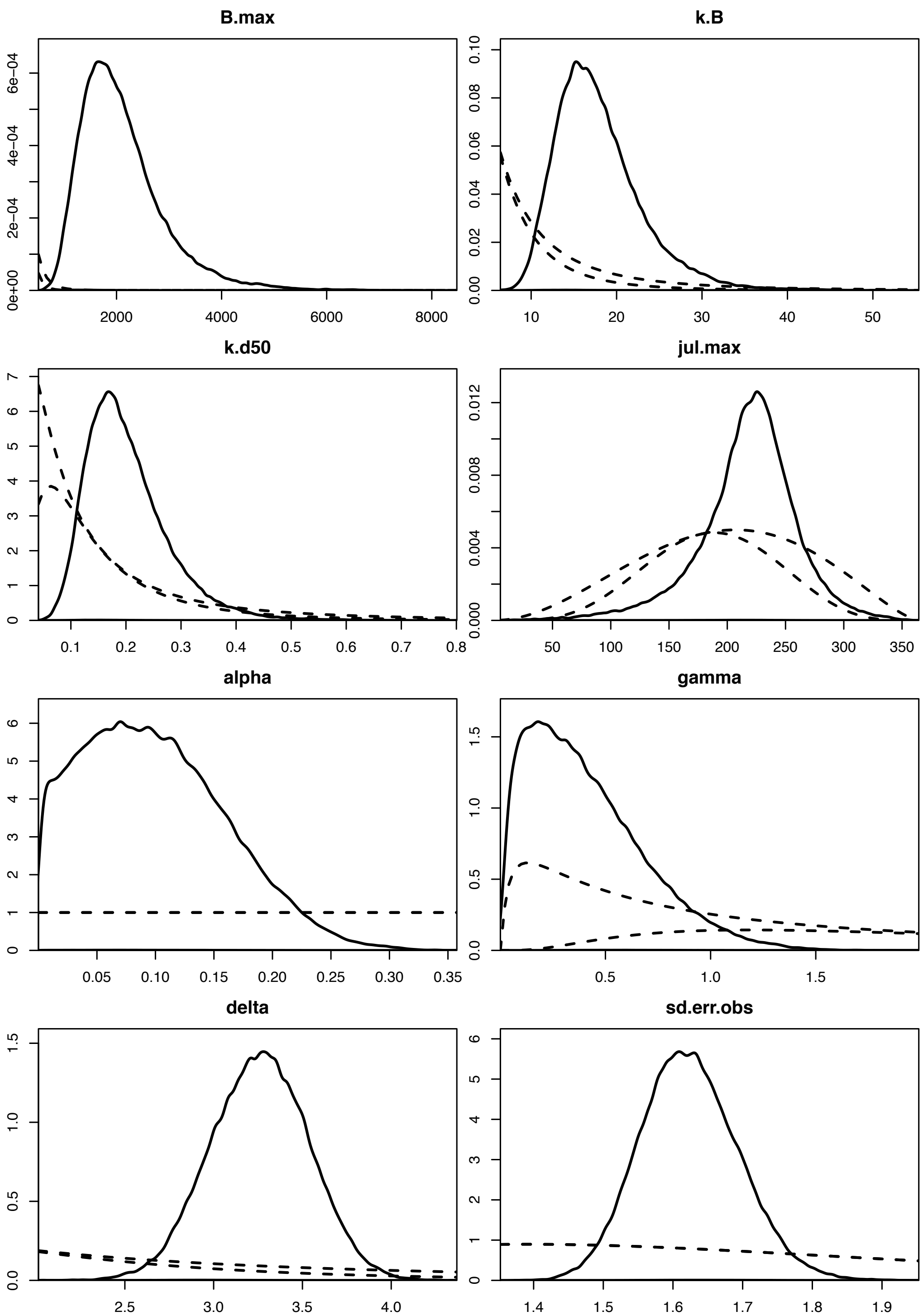

Figure 4.5: Upper posteriors (lines) of the Bayesian parameter inference for the example in which the multivariate prior (dotted) was imprecise. The lower measure of the marginal posterior class is hardly distinguishable from zero on the scale that is adjusted to display the upper measure. The approximated ratio of normalizing constants was $r \approx 0.00108$ according to (4.50).

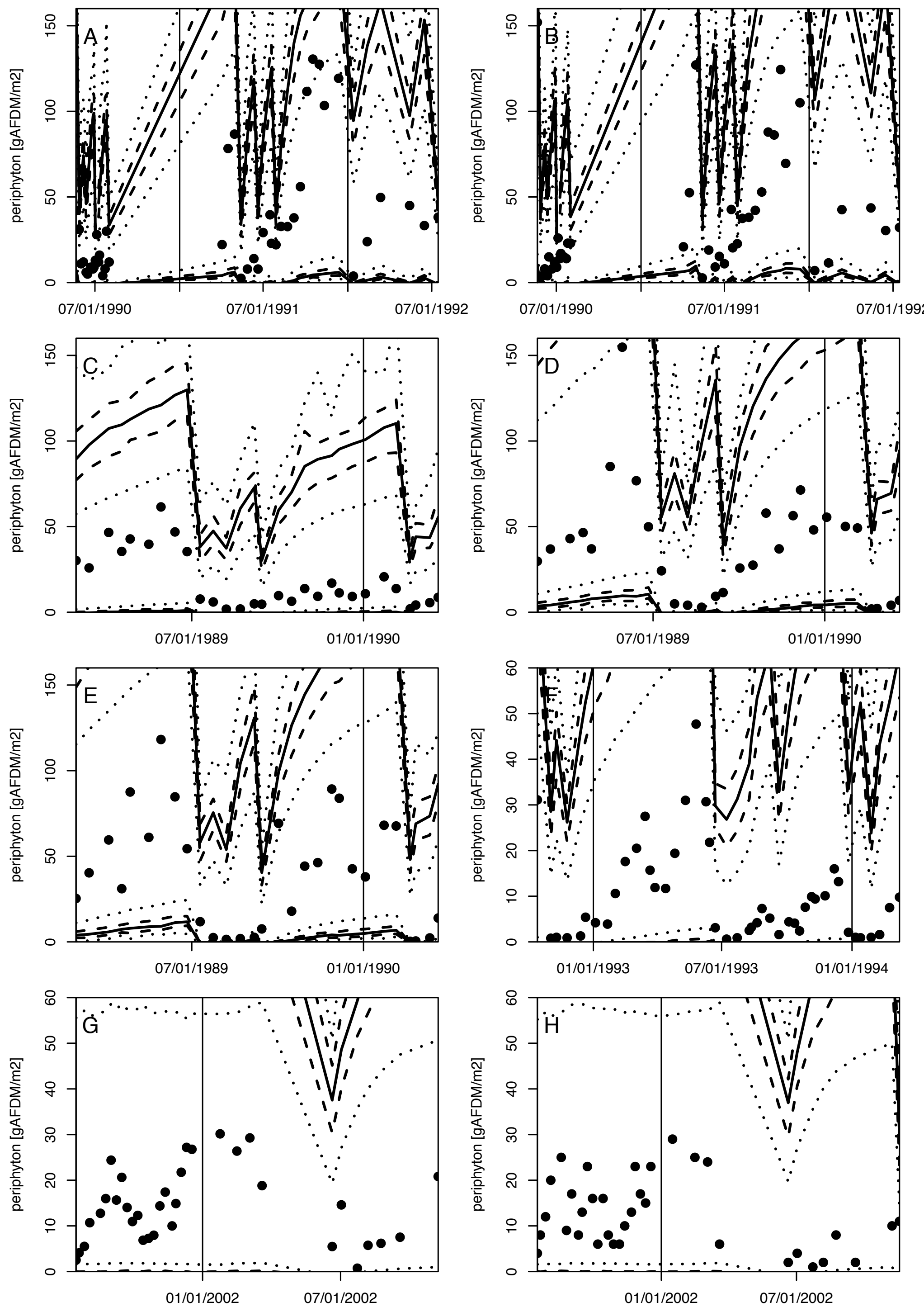

Figure 4.6: Model predictions based on the imprecise prior shown in Figure 4.5. Lower and upper quantiles for standard probabilities (lines) indicate the imprecise model prediction. The vertical lines indicate the beginning of a year.

As a consequence, there is increasing ambiguity in the corresponding model predictions (see Figures 4.2, 4.4 and 4.6). Already the consideration of ambiguity in one parameter leads to considerable ambiguity in model results (compare Figure 4.4 with Figure 4.2). When multiple parameters were allowed to have imprecise prior marginals, the ratio of normalization constants between the upper and lower densities of the prior and consequently of the posterior of parameters and predictions became very large (≈ 900), indicating a very high degree of ambiguity. Figure 4.6 demonstrates these results. It is remarkable, for example, that the upper 5-percentile has higher values than the lower 95-percentile.

This example demonstrates that ambiguity formulated as imprecision in prior marginals can be considered in Bayesian inference and the corresponding posterior class can be propagated to the model results. This is possible without significantly increasing the computational burden of the analysis. The sequence of examples also demonstrates that care is needed when constructing a prior class as the product of prior marginals. Ambiguity can easily become so large that the results indicate that no meaningful prediction can be made. This indicates the necessity of limiting the use of imprecise priors to those marginals for which the prior knowledge is really controversial. If this is true for many marginals, the likely result will be that prediction uncertainty bounds are very uncertain.

4.5 Discussion and Conclusions

Growing computing power and increasingly efficient algorithms have made Bayesian inference and prediction commonplace in science and engineering applications. However, in many cases, there is insufficient knowledge or experience to formulate precise prior distributions on model parameters. Alternately, there may be many experts, any one of whom may not admit to imprecision personally, but whose collective disagreement generates substantial ambiguity for the analyst. We have previously described methods for using expert elicitation of one or multiple experts to construct a *Density Ratio Class* (Rinderknecht et al. 2012), which we have found to be an attractive approach for characterizing imprecise probabilities. In particular, the fact that this class is invariant under Bayesian inference — and thereby can represent incremental learning in a consistent framework — makes it satisfying conceptually. The fact that it is also invariant under marginalization is convenient for visualization. In the present paper, we review the proofs of these two invariance properties, first noted by Wasserman (1992a). To our knowledge, our proof of *Density Ratio Class* invariance under propagation through a deterministic model (Section 4.2) is novel. We also show that when a stochastic model is being used a *Density Ra-*

tio Class can be constructed to represent a conservative estimate of the predictive distribution resulting from propagation through the model.

The invariance properties of the *Density Ratio Class* lead to significant advantages regarding numerical implementation of inference, marginalization, and model prediction, which we describe in detail in Section 4.3. In particular we show that, through judicious use of sample weights and ratios of normalizing constants, the MCMC samples generated using precise priors can be used to approximate the upper and lower bounds resulting from imprecise priors, thereby minimizing additional computational burden. This should open up many more opportunities for using the imprecise probabilities framework than have existed in the past.

We show by the third example in Section 4.4 that the number of parameters bearing imprecision has to be chosen carefully to avoid extreme ambiguity in results. Combining imprecision in marginals by constructing a joint class that contains all possible combinations of priors leads to very high imprecision in this joint class. It is an inconvenient and unavoidable 'feature' of this approach that marginalizing this class leads to a higher imprecision in the marginal than the one used to construct the joint distribution (see Subsection 4.4.3). Second-order probabilities can easily be combined with imprecise probabilities; it seems to be a reasonable strategy to use imprecise probabilities only for those marginals for which it seems to be difficult to specify a unique probability measure. This avoids "unnecessarily" extreme ambiguity in the results. On the other hand, extreme ambiguity is allowed by this framework. This is a very important additional element that is missing in decision support formulated probabilistically, even if second-order probabilities are used.

Extreme ambiguity as observed in our third example implies that model predictions have little practical value for decision support. For example, according to standard decision theory, a decision-maker who wants to follow basic axioms of rationality should account for uncertainty by choosing the action that maximizes his or her expected utility (Von Neumann and Morgenstern 1944; Savage 1954). When imprecise probability distributions are used to represent uncertainty, intervals of expected utilities result. Overlapping intervals imply that there is not sufficiently precise knowledge available to obtain a unique ranking among decision alternatives according to expected utility theory. In this way, imprecise probabilities provide a measure of the 'decidability' of a decision problem, as informed by model predictions. This is an important consideration that is missing in decision support formulated using only precise probabilities. For situations in which there is too much ambiguity to rely exclusively on expected utility theory, secondary criteria such as probabilities of improvement (Reichert and Borsuk 2005) or the precautionary principle (Foster et al. 2000; Gollier and Treich 2003) may be required.

4.6 Acknowledgements

This study was supported by the Swiss National Science Foundation. M.E.B. was partially supported by the US EPA STAR program. This work has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

Chapter 5

Conclusions and Further Research Needs

5.1 Conclusions

In rational decision analysis and decision support the decision making process is divided into the steps of (1) structuring objectives and quantifying their degree of fulfillment by a utility function that depends on measurable system attributes, (2) predicting the consequences of decision alternatives in the form of probability distributions of these attributes, and, finally, (3) ranking the alternatives according to decreasing values of expected utility. The derivation of probabilistic predictions often requires the elicitation of knowledge about influence factors, model parameters or directly predicted attributes from subject matter experts. This is especially the case in environmental modeling since epistemic uncertainty is often dominant. If we quantify subjective degrees of belief of experts with the aid of bets on lotteries between which a person is indifferent, and if we assume that the experts want to avoid sure loss, then the resulting knowledge quantifications will be consistent with the axiomatic foundation of probability theory and the mathematical construct for describing and updating such beliefs must be Bayesian statistics. Unfortunately, the expert(s) often are unable to characterize(s) her/his/their uncertain degrees of belief without ambiguity. Ambiguity becomes even more relevant if experts are asked to specify knowledge in terms of intersubjective knowledge, representing the current state of knowledge of the scientific community about an environmental system, its mathematical description, or specific parameter values. We therefore suggest to use imprecise probabilities that allow us to continuously characterize the degree of ambiguity, ranging from total ignorance to precise probability distributions. This extends our framework to robust Bayesian statistics. Many concepts have been proposed for specifying sets of probability distributions. The *Density Ratio Class* has unique conceptual advantages, as we have shown in Rinderknecht et al. (2011, 4): it is (i) invariant under Bayesian updating, (ii) invariant under marginalization, (iii) invariant if propagated through a deterministic model, and (iv) embedded again into a *Density Ratio Class* that can be larger than the set of propagated distributions of the initial class if the model is stochastic. In particular, the *Density Ratio Class*'s invariance under Bayesian updating and marginalization makes it the unique class that allows for simultaneously describing a consistent sequential Bayesian learning process and conveniently conveying higher dimensional cases. Furthermore, the class is nonparametric but naturally excludes extremely variable probability densities that seem unreasonable (such as densities with extreme peaks or even point masses). Because the *Density Ratio Class* is difficult to elicit, we developed an elicitation technique that fits a *Density Ratio Class* with given parametric bounds to elicited quantile or probability intervals. We do not claim that experts have internal representations of uncertainty that conform precisely with the *Density Ratio Class* but our technique seems to yield a reasonable representation of the expert's (possibly ambiguous) beliefs. For deeper insight and a more detailed description of the ambiguity represented by a class of probability distributions, we proposed three generic metrics applicable to any set specification. These metrics measure the relative, to a predefined credible level, ambiguity of specific characteristics of a probability distribution such as the width, the shape and the position of the mode.

We successfully applied – to parameters containing different degrees of uncertainty and ambiguity – (a) the elicitation technique presented in Rinderknecht et al. (2011, 2.3) and (b) the proposed metrics presented in Rinderknecht et al. (2012, 3.3). We hereby point to three exemplary cases originating from a wide range of application areas that are presented in Rinderknecht et al. (2012, 3.5.1, 3.5.2, 3.5.3).

Concerning the treatment of imprecise prior knowledge in the form of *Density Ratio Classes* with regard to its effects on parameter estimates and model predictions we successfully applied – without a substantial increase in computational burden compared to conventional methods – our proposed methods presented in Rinderknecht et al. (2011, 4) to a simple deterministic periphyton model that has an additive stochastic error. For complex models with a large number of parameters formulating ambiguity in all parameters can lead to a very high ambiguity in the predictions as we have shown, i.e., in the third example of our periphyton model in Rinderknecht et al. (2011, 4.4.3). Hence, when independent parameter distributions are merged to a joint distribution, we recommend to carefully select the parameters that need an ambiguous formulation by a *Density Ratio Class*. If this is the case for a large number of parameters, we have to accept that only highly ambiguous predictions can be made.

Finally, we developed a generically extendable R software package `fitDRC.R`, freely available on http://cran.r-project.org/ subject to the terms of agreement. The package (I) allows its users to construct ready-to-use one-dimensional smallest *Density Ratio Classes* for elicited probability-quantile points (or intervals) given a lower and upper (possibly transformed) distributional shape as presented in Rinderknecht et al. (2011, 2.3) and (II) it makes it possible to calculate the metrics as proposed in Rinderknecht et al. (2012, 3.3) for non transformed distributions. The most frequently used distributions (Normal, Student, Weibull, Log-Normal, Beta, Gamma, F, Uniform, Logistic) and transformations (Arctan, Tan, Log, Exponential, Dilation) are already implemented. If the desired distribution or transformation should not yet be implemented the generically extendable `fitDRC.R` package is easily customizable by simply modifying the standard templates from the example section.

5.2 Further Research Needs

Despite having made a step towards simplifying the use of *Density Ratio Classes* in modeling and decision support, there remain many important research questions to be addressed in future work. Some of these needs are listed below.

1. **Elicitation of Dependent Parameters**

   In the description of our proposed elicitation method, we focussed on one-dimensional parameters, but the *Density Ratio Class* can be readily defined for the multivariate case as well. It is well known that the construction of elicitation techniques that appropriately capture covariance in an expert's knowledge about multiple uncertain quantities is challenging. Examples of relevant research questions are:

   (a) How could knowledge on dependent multivariate parameters best be elicited in general?

   (b) How could covariances of multivariate parameters in the form of *Density Ratio Classes* best be elicited in particular?

2. **Application of Imprecise Probabilities in Decision Analysis and Formulation of Imprecise Utilities**

   When classes of probability distributions are used, then expected utilities result in intervals. These utility intervals may overlap or not. In the case of non-overlapping utility intervals, a unique ranking of decision alternatives is still possible and the selection of the most appropriate one is guaranteed. But in the case of overlapping utility intervals, there is, to our knowledge, no established decision rule for the selection of the most appropriate decision alternative. In addition, if ambiguity in probabilities is considered, it seems reasonable to consider ambiguity in preference quantification also. This would lead to imprecise utilities. This leads to the following relevant research questions:

(a) How can decisions best be supported if the expected utility intervals overlap?

(b) How can the theory of rational decision making best be extended to consider imprecise utilities?

3. **Treatment of Ambiguity**

We have found that the introduction of ambiguity in the description of knowledge about model parameters may cause a severe increase of ambiguity in model predictions which then has substantial consequences on model-based decision support. Relevant research questions in this field are:

(a) Is the combination of ambiguity by joining independent *Density Ratio Classes* too conservative and could a methodology be developed that would lead to a smaller increase in ambiguity?

(b) Could other concepts be combined with imprecise probabilities to reduce the increase in ambiguity?

4. **Marginalizing Classes of Combined Distributions**

For two univariate *Density Ratio Classes* with bounds l_1, u_1 and l_2, u_2, it makes sense to consider the smallest class which contains all product densities of members of the two classes. It is not difficult to see that this is the *Density Ratio Class* with bounds $l_1 l_2, u_1 u_2$ (take l_1 and l_2 normalized and consider densities f_i which are equal to l_i except on a small interval where $f_i = u_i$). Marginalizing this class produces new classes which are larger than the ones we started with.

(a) What are the theoretical and practical consequences of the fact described above?

(b) Is there a concept that eliminates the mentioned problem?

5. **Software Development**

   There may be various reasons why the use of classes of probability distributions did not become widespread in decision support. However, it is a fact that software tools that consider imprecise probabilities are scarce. One reason may be that the additional numerical burden was expected to be considerable. We demonstrated that the implementation of imprecise probabilities for the quantification of ambiguity does not lead to an overwhelming complexity or computational burden. This leads to questions regarding the optimal support of the application of imprecise probabilities by software tools:

   (a) Which software tools (with an appropriate interface) would be needed to further accommodate and facilitate the use of imprecise probabilities for predictions and decision analysis?

   (b) Are there existing software tools that could be extended for this purpose and how could such a software tool best be implemented?

Curriculum Vitae

1. Personal Data

Given name	SIMON Lukas
Last name	Rinderknecht
Date of birth	February 8, 1979
Place of birth	Baden, Switzerland
Place of citizenship	Wallisellen, ZH
Nationality	Swiss

2. Education

2011 Doctoral Thesis: *Contribution to the Use of Imprecise Scientific Knowledge in Decision Support* at the Department of System Analysis, Integrated Assessment and Modelling (Siam) of the Swiss Aquatic Research Institute (Eawag) and the Institute of Biogeochemistry and Pollutant Dynamics (IBP) of the ETH Zurich.

2007 Master Bi-Disciplinaire ès Sciences avec Majeure en Mathématiques et Mineure en Sciences de la Terre et de l'Environnement. *Numerical Analysis of the Surface Heat Flux of an active Volcano: Application to Villarrica, Chile* dans le Département de Minéralogie à l'Université de Genève.

2005 Bachelor ès Sciences en Mathématiques à l'Université de Neuchâtel.

1999 Naturwissenschaftliche Matura (Typus C) mit kleinem Latinum an der Aargauischen Kantonsschule in Baden.

Acknowledgments

People, I will never forget.

Supervisors of the Doctoral Thesis

Prof. Dr. Peter Reichert, Prof. Dr. Mark E. Borsuk, Prof. Dr. Hans R. Künsch.

Co-Workers

Prof. Dr. Martin Ackermann (Chairman of the Board of Examiners), Dr. Carlo Albert (Co-Author; Co-Assistant, Modelling of Aquatic Ecosystems, ETH Zurich; Co-Assistant, Environmental Systems Analysis, Eawag Summer School), Adelaida Arreaza (Master Student ETH Zurich), Anne Dietzel (Co-Ph.D. Student; Co-Assistant, Modelling of Aquatic Ecosystems, ETH Zurich; Co-Assistant, Environmental Systems Analysis, Eawag Summer School), Karin Ghilardi (Secretary of the Department Siam, Eawag; Member of Staff Council at Eawag), Dr. Tom Gonser (Member of Staff Council at Eawag), Dr. Lukas Indermaur (Co-Assistant, Modelling of Aquatic Ecosystems, ETH Zurich), Dr. Dmitri Kavetski (Guest Lecturer at Eawag Summer School), Dr. Hans Peter E. Kohler (Member of Staff Council at Eawag), Dr. Simone Langhans (Office-Mate), Gérard Mohler (IT-Support, Eawag), Philipe Périsset (IT-Support, Eawag), Raoul Schaffner (IT-Support, Eawag), Rudolf Schneebeli (Successor Member of Staff Council at Eawag), Florian Schnetzer (Office-Mate), Lisa Scholten (Office-Mate; Main Supervisor of our Master Student), Dr. Nele Schuwirth (Co-Autor; Co-Assistant, Modelling of Aquatic Ecosystems, ETH Zurich), Jacqueline Traber (Member of Staff Council at Eawag), Dr. Irene Wittmer (Co-Assistant, Modelling of Aquatic Ecosystems, ETH Zurich), Prof. Dr. Marco Zaffalon (independent co-examiner), Alois Zwyssig (Member of Staff Council at Eawag).

Institutions, Funds

The research group Systems Analysis, Surface Waters; *Siam* the Department of System Analysis, Integrated Assessment and Modelling; *Eawag* the Swiss Aquatic Research Institute; Thayer School of Engineering at the Dartmouth College; *IBP* the Institute of Biogeochemistry and Pollutant Dynamics; *D-UWIS* the Department of Environmental Sciences; *ETH Zurich* the Swiss Federal Institute of Technology; *SNSF* the Swiss National Science Foundation.

Family-Clan

Werner Josef Rinderknecht, Monika Maria Rinderknecht-Deiss, Fabian Daniel Rinderknecht, Marius Nicolas Rinderknecht, Johann Deiss (†), Katharina Deiss-Emmenegger (†), Rudolf Rinderknecht (†), Rösli Rinderknecht-Flück, Lina Flück (†).

† Rest in Peace.

Friends

Intentionally left due to interests of data security.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica 21*(4), 503–546.

Alpert, M. and H. Raiffa (1982). *A progress report on the training of probability assessors.*, pp. 294–305. Cambridge: Cambridge University Press.

Ayyub, B. M. and G. J. Klir (2006). *Uncertainty Modeling and Analysis in Engineering and the Sciences*. Chapman & Hall.

Basu, S. and A. DasGupta (1995). Robust Bayesian analysis with distribution bands. *Statistics and Decisions 13*, 333–349.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer.

Berger, J. and L. Berliner (1986). Robust Bayesian and empirical Bayes analysis with ϵ-contaminated priors. *Ann. Statist. 14*(2), 461–486.

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

Berger, J. O. (1994). An overview of robust Bayesian analysis. *Test 3*, 5–124.

Biggs, B. J. F. (1994). Responses of two trophic levels to patch enrichment along a New Zealand stream continuum. *New Zealand Journal of Marine and Freshwater Research 28*, 119–134.

Biggs, B. J. F. and S. Stokseth (1996). Hydraulic habitat suitability for periphyton in rivers. *Regulated Rivers-Research & Management 12*(2-3), 251–261.

Borsuk, M. E., S. P. Powers, and C. H. Peterson (2002). A survival model of the effects of bottom-water hypoxia on the population density of an estuarine clam (Macoma balthica). *Canadian Journal of Fisheries and Aquatic Sciences 59*(8), 1266–1274.

Borsuk, M. E. and L. Tomassini (2005). Uncertainty, imprecision, and the precautionary principle in climate change assessment. *Water Science and Technology 52*(6), 213–225.

Boulêtreau, S., F. Garabétian, S. Sauvage, and J.-M. Sánchez-Pérez (2006). Assessing the importance of a self-generated detachment process in river biofilm models. *Freshwater Biology 51*(5), 901–912.

Box, G. E. P. and D. R. Cox (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological) 26*(2), 211–252.

Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Wiley.

Camerer, C. and M. Weber (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty 5*, 325–370.

Caselton, W. F. and W. B. Luo (1992). Decision-making with imprecise probabilities: Dempster-Shafer theory and application. *Water Resources Research 28*(12), 3071–3083.

Cattaneo, M. E. G. V. (2008). Fuzzy probabilities based on the likelihood function. In D. Dubois, M. Lubiano, H. Prade, M. Gil, P. Grzegorzewski, and O. Hryniewicz (Eds.), *Soft Methods for Handling Variability and Imprecision, Advances in Soft Computing*, Volume 48, pp. 43–50. Springer.

Chaloner, K. (1996). The elicitation of prior distributions. In D. A. Berry and D. K. Stangel (Eds.), *Bayesian Biostatistics*, pp. 141–156. New York: Marcel Dekker.

Clemen, R. T. and T. Reilly (2001). *Making hard decisions with DecisionTools®*. Pacific Grove, CA: Duxbury Press.

Cooke, R. (2001). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. New York: Oxford University Press.

Corsair, H. J., J. Bassman Ruch, P. Q. Zheng, B. F. Hobbs, and J. F. Koonce (2009). Multicriteria decision analysis of stream restoration: Potential and examples. *Group Decision and Negotiation DOI 10.1007/s10726-008-9148-4*, 31.

Cox, R. (1946). Probability, frequency and reasonable expectation. *American Journal of Physics 14*(1), 1–13.

de Finetti, B. (1931). Sul significato soggettivo della probabilità. *Fundamenta Mathematicae 17*, 298–329.

de Finetti, B. (1974). *Theory of probability. A critical introductory treatment.*, Volume 1, 2. London: Wiley.

DeRobertis, L. and J. A. Hartigan (1981). Bayesian inference using intervals of measures. *The Annals of Statistics 9*(2), 235–244.

Destercke, S., D. Dubois, and E. Chojnacki (2008). Unifying practical uncertainty representations: I. generalized p-boxes - II. clouds. *International Journal of Approximate Reasoning 49*(3), 649–677.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological) 57*(1), 45–97.

Einhorn, H. J. and R. M. Hogarth (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review 92*(4), 433–461.

Eisenführ, F. and M. Weber (2003). *Rationales Entscheiden* (4 ed.). Berlin: Springer.

Elber, F., J. Hürlimann, and K. Niederberger (1992). Biologische Begleitung der Schwallversuche vom Sommer 1990 in der Sihl. *Wasser, Energie, Luft 84*(3), 42–50.

Elber, F., J. Hürlimann, and K. Niederberger (1996). Algenmonitoring als Grundlage für das Abflussmanagement in der Sihl. *Wasser, Energie, Luft 88*(3), 55–62.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Ferson, S. and J. G. Hajagos (2004). Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety 85*(1-3), 135–152.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz (2003). Constructing probability boxes and Dempster-Shafer structures. *Sandia Report SAND 2002-4015*, 1–143.

Fierens, P. I., L. C. Rêgo, and T. L. Fine (2009). A frequentist understanding of sets of measures. *Journal of Statistical Planning and Inference 139*, 1879–1892.

Foster, K. R., P. Vecchia, and M. H. Repacholi (2000). Science and the precautionary principle. *Science 288*, 979–981.

French, S. and D. Ríos Insua (2000). *Statistical decision theory.* London: Edward Arnold.

Frisch, D. and J. Baron (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making 1*, 149–157.

Gamerman, D. and H. F. Lopes (2006). *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference.* (2nd ed.). London: Chapman & Hall.

Garthwaite, P., J. Kadane, and A. O'Hagan (2005). Statistical methods for eliciting prior distributions. *Journal of the American Statistical Association 100*, 680–701.

Garthwaite, P. H. and J. M. Dickey (1985). Double- and single-bisection methods for subjecitve probability assessments in a location-scale family. *Journal of Econometrics 29*, 149–163.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2003). *Bayesian Data Analysis* (2nd ed.). London: Chapman & Hall.

Gelman, A. and X. L. Meng (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical Science 13*(2), 163–185.

Gillies, D. (1991). Intersubjective probability and confirmation theory. *British Journal for the Philosophy of Science 42*, 513–533.

Gollier, C. and N. Treich (2003). Decision-making under scientific uncertainty: the economics of the precautionary principle. *The Journal of Risk and Uncertainty 27*(1), 77—1030.

Held, H., T. Augustin, and E. Kriegler (2008). Bayesian learning for a class of priors with prescribed marginals. *International Journal of Approximate Reasoning 49*(1), 212–233.

Howson, C. and P. Urbach (1989). *Scientific reasoning: The Bayesian approach*. Open Court Pub Co.

Huber, P. J. (1973). The case of choquet capacities in statistics. *International Statistical Institute 4*, 181–188.

Huber, P. J. and V. Strassen (1973). Minimax tests and the Neyman-Pearson lemma for capacities. *The Annals of Statistics 1*(2), 251–263.

Ihaka, R. and R. Gentleman (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics 5*(3), 299–314.

James, A., S. Low Choy, and K. Mengersen (2010). Elicitator: An expert elicitation tool for regression in ecology. *Environmental Modelling & Software 25*(1), 129–145.

Jenkinson, D. (2005). The elicitation of probabilities - a review of the statistical literature. Technical report, Bayesian Elicitation of Expert's Probabilities (BEEP).

Kadane, J. B. and A. O'Hagan (1995). Using finitely additive probability: Uniform distributions on the natural numbers. *Journal of the American Statistical Association 90*(430), 626–631.

Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1996). Reasoning to a forgone conclusion. *Journal of the American Statistical Association 91*(435), 1228–1235.

Kadane, J. B. and L. Wolfson (1998). Experiences in elicitation. *Statistician 47*(1), 3–19.

Kahneman, D. and A. Tversky (1979). Prospect theory: Analysis of decision under risk. *Econometrica 47*(2), 263–291.

Keynes, J. M. (1921). *A Treatise on Probability*. Macmillan & Co., Ltd. London.

Klir, G. J. and M. J. Wierman (1998). *Uncertainty-Based Information*. New York: Physica Verlag.

Kriegler, E., J. W. Hall, H. Held, R. Dawson, and H. J. Schellnhuber (2009). Imprecise probability assessment of tipping points in the climate system. *Proceedings of the National Academy of Sciences 106*(13), 5041–5046.

Kriegler, E. and H. Held (2005). Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning 39*(2-3), 185–209.

Kynn, M. (2008). The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 171*(1), 239–264.

Lavine, M. (1991a). An approach to robust Bayesian analysis for multidimensional parameter spaces. *Journal of the American Statistical Association 86*(414), 400–403.

Lavine, M. (1991b). Sensitivity in Bayesian statistics: the prior and the likelihood. *Journal of the American Statistical Association 86*(414), 396–399.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips (1982). *Calibration of probabilities: the state of the art to 1980*. Cambridge University Press.

Maturi, T. A., P. Coolen-Schrijner, and F. P. Coolen (2010). Nonparametric predictive comparison of lifetime data under progressive censoring. *Journal of Statistical Planning and Inference 140*(2), 515 – 525.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics 21*(6), 1087–1092.

Meyer, M. A. and J. M. Booker (1991). *Eliciting and Analyzing Expert Judgment: A Practical Guide*. London: Academic Press Limited.

Moreno, E. and J. A. Cano (1991). Robust Bayesian analysis with ϵ-contaminations partially known. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 53*(1), 143–155.

Moreno, E. and L. R. Pericchi (1993a). Bayesian robustness for hierarchical ϵ-contamination models. *Journal of Statistical Planning and Inference 37*, 159–168.

Moreno, E. and L. R. Pericchi (1993b). Prior assessments for bands of probability measures: Empirical Bayes analysis. *Test 2*(1-2), 101–110.

Nordhaus, W. (1994). Expert opinion on climate change. *American Scientist 82*, 45–51.

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

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow (2006). *Uncertain Judgements. Eliciting Experts' Probabilities.* West Sussex, England: John Wiley.

O'Hagan, A. and J. E. Oakley (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety 85*, 239–248.

Page, T., A. L. Heathwaite, L. J. Thompson, L. Pope, and R. Willows (2012). Eliciting fuzzy distributions from experts for ranking conceptual risk model components. *Environmental Modelling & Software 36*, 19–34.

Parry, G. W. (1996). The characterization of uncertainty in Probabilistic Risk Assessments of complex systems. *Reliability Engineering & System Safety 54*, 119–126.

Pericchi, L. R. (1998). Sets of prior probabilites and Bayesian robustness. Documentation Section on the website of the Society for Imprecise Probability Theory and Applications (SIPTA).

Pericchi, L. R. and P. Walley (1991). Robust Bayesian credible intervals and prior ignorance. *International Statistical Review 58*(1), 1–23.

Pollino, C. A., O. Woodberry, A. Nicholson, K. Korb, and B. T. Hart (2007). Parameterisation and evaluation of a bayesian network for use in an ecological risk assessment. *Environmental Modelling & Software 22*(8), 1140–1152.

Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays.* London: Routledge and Kegan Paul Ltd.

Refsgaard, J. C., J. P. van der Sluijs, A. L. Højberg, and P. A. Vanrolleghem (2007). Uncertainty in the environmental modelling process - a framework and guidance. *Environmental Modelling & Software 22*(11), 1543 – 1556.

Reichert, P. and M. Borsuk (2005). Uncertainty in model predictions: does it preclude effective decision support? *Environmental Modelling & Software 20*, 991–1001.

Reichert, P., M. E. Borsuk, M. Hostmann, S. Schweizer, C. Spörri, K. Tockner, and B. Truffer (2007). Concepts of decision support for river rehabilitation. *Environmental Modelling & Software 22*(2), 188–201.

Rinderknecht, S. L., C. Albert, N. Schuwirth, M. E. Borsuk, H. R. Künsch, and P. Reichert (2011). The effect of imprecise prior knowledge on parameter estimates and model predictions. Submitted.

Rinderknecht, S. L., M. E. Borsuk, and P. Reichert (2011). Eliciting Density Ratio Classes. *International Journal of Approximate Reasoning 52*(6), 792–804.

Rinderknecht, S. L., M. E. Borsuk, and P. Reichert (2012). Bridging uncertain and ambiguous knowledge with imprecise probabilities. *Environmental Modelling & Software 36*, 122–130.

Ríos Insua, D. and E. Ruggeri, Fabrizio (2000). *Robust Bayesian Analysis*. New York: Springer.

Savage, L. (1954). *The Foundations of Statistics*. New York: John Wiley.

Schweizer, S. (2007). *Predicting the Consequences of River Rehabilitation Measures on Morphology, Hydraulics, Periphyton and on Invertebrates*. Diss. Nr. 16838, ETHZ, Zürich.

Seidenfeld, T., M. Schervish, and J. B. Kadane (1995). A representation of partially ordered preferences. *Annals of Statistics 23*(6), 2168–2217.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal 27*(3&4), pp. 379–423, 623–656.

Sivaganesan, S. and J. Berger (1989). Ranges of posterior measures for priors with unimodal contaminations. *Ann. Statist. 17*, 868–889.

Spetzler, C. S. and C.-A. S. Staël Von Holstein (1975). Probability encoding in decision analysis. *Management Science 22*(3), 340–358.

Stroock, D. W. (1999). *A concise introduction to the theory of integration* (3rd ed.). Birkhäuser.

Tversky, A. and D. Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science 185*(4157), 1124–1131.

Uehlinger, U. (1991). Spatial and temporal variability of the periphyton biomass in a prealpine river (Necker, Switzerland). *Archiv für Hydrobiologie 123*(2), 219–237.

Uehlinger, U., H. Bührer, and P. Reichert (1996). Periphyton dynamics in a flood-prone prealpine river: Evaluation of significant processes by modelling. *Freshwater Biology 36*(2), 249–263.

Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

von Winterfeldt, D. and W. Edwards (1986). *Decision analysis and behavioral research*. UK: Cambridge University Press.

Walker, W. E., P. Harremoës, J. Rotmans, J. P. van der Sluijs, M. B. A. van Asselt, P. Janssen, and M. P. Krayer von Krauss (2003). Defining uncetainty. a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment 4*(1), 5–17.

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

Wallsten, T. and D. Budescu (1983). Encoding subjective probabilities: A psychological and psychometric review. *Management Science 29*, 151–173.

Warmink, J., J. Janssen, M. Booij, and M. Krol (2010). Identification and classification of uncertainties in the application of environmental models. *Environmental Modelling & Software 25*(12), 1518–1527.

Wasserman, L. (1992a). Invariance properties of density ratio priors. *The Annals of Statistics 20*(4), 2177–2182.

Wasserman, L. (1992b). Recent methodological advances in robust Bayesian inference. in: Bayesian Statistics 4, Bernardo, J. M. and Berger, J. O. and Dawid, A. P. and Smith, A.F.M., (Eds.), Oxford University Press, 483-502 (with discussion).

Wasserman, L. and J. B. Kadane (1992). Computing bounds on expectations. *Journal of American Statistical Association 87*(418), 516–522.

Williamson, R. and T. Downs (1990). Probabilistic arithmetic. I. Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning 4*(2), 89–158.

Winkler, R. L. (1967). The assessment of prior distributions in Bayesian analysis. *Journal of American Statistical Association 62*(319), 776–880.

Zadeh, L. (1965). Fuzzy sets. *Information and Control 8*(3), 338–353.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems 1*(1), 3–28.