

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

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 unpräzise 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.