

Towards a comprehensive uncertainty assessment in environmental research and decision support

Peter Reichert

ABSTRACT

Uncertainty quantification is very important in environmental management to allow decision makers to consider the reliability of predictions of the consequences of decision alternatives and relate them to their risk attitudes and the uncertainty about their preferences. Nevertheless, uncertainty quantification in environmental decision support is often incomplete and the robustness of the results regarding assumptions made for uncertainty quantification is often not investigated. In this article, an attempt is made to demonstrate how uncertainty can be considered more comprehensively in environmental research and decision support by combining well-established with rarely applied statistical techniques. In particular, the following elements of uncertainty quantification are discussed: (i) using stochastic, mechanistic models that consider and propagate uncertainties from their origin to the output; (ii) profiting from the support of modern techniques of data science to increase the diversity of the exploration process, to benchmark mechanistic models, and to find new relationships; (iii) analysing structural alternatives by multi-model and non-parametric approaches; (iv) quantitatively formulating and using societal preferences in decision support; (v) explicitly considering the uncertainty of elicited preferences in addition to the uncertainty of predictions in decision support; and (vi) explicitly considering the ambiguity about prior distributions for predictions and preferences by using imprecise probabilities. In particular, (v) and (vi) have mostly been ignored in the past and a guideline is provided on how these uncertainties can be considered without significantly increasing the computational burden. The methodological approach to (v) and (vi) is based on expected utility theory, which extends expected utility theory to the consideration of uncertain preferences, and on imprecise, intersubjective Bayesian probabilities.

Key words | Bayesian inference, decision analysis, environmental decision support, expected utility, imprecise probabilities, stochastic models

Peter Reichert

Eawag: Swiss Federal Institute of Aquatic Science
and Technology,
8600 Dübendorf,
Switzerland
E-mail: peter.reichert@eawag.ch

INTRODUCTION

Scientific integrity principles and ethical concerns require scientists to be open about and proactively communicate the uncertainty in their predictions. In environmental decision support, predictions of the consequences of decision alternatives have to be assessed for the fulfilment of societal goals. This adds additional uncertainty to the decision support process, as elicited societal preferences are uncertain due to uncertainty of and temporal changes

in the preferences of individuals, different perceptions and values of different individuals in the society, and uncertainties induced by the parameterization of quantified preferences and the elicitation process.

Over the past few decades, many promising concepts and methodologies have been developed to address these uncertainties. Nevertheless, most environmental decision support processes address uncertainty only incompletely. To stimulate more comprehensive uncertainty analyses in the future, it is the goal of this paper to review and discuss techniques that can contribute to better considering uncertainty in environmental research and decision support and to outline a decision support procedure that

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

more comprehensively addresses uncertainty. In particular, the goal is to emphasize how the uncertainty of preferences and the ambiguity in specifying prior distributions can be included in a decision support process as these uncertainties, despite their high relevance, are often neglected in environmental decision support.

Many reviews on modelling environmental systems and on the modelling process exist. Most of these reviews emphasize the importance of quantifying uncertainty (e.g. Clark 2005; Refsgaard *et al.* 2007; Schuwirth *et al.* 2019). Other studies focus on expert or stakeholder involvement in model building (e.g. Voinov & Bousquet 2010; Krueger *et al.* 2012; Voinov *et al.* 2016) which also relates to extending structural diversity and considering uncertainty. This article builds on this literature, but focuses on useful methodologies for uncertainty quantification rather than on guidelines for the model building or decision support process.

In a modelling process that primarily focuses on increasing our understanding of the investigated system, we need:

- (a) a model that attempts to describe the underlying mechanisms of the observed behaviour of a system and makes it possible to test these model formulations ('hypotheses') with observed data.

To support environmental decisions, we need:

- (b1) a description of societal preferences about what should be achieved, ideally in quantitative terms expressed as functions of observable system attributes; and
- (b2) a model that is based on the current state of scientific knowledge and predicts the consequences of decision alternatives on output variables that are relevant for assessing the fulfilment of the societal goals (the attributes mentioned in b1).

Note that the requirements for models of category (a) and models of category (b2) are somewhat different, so that despite many synergies, a model designed for (a) is not always the best model for (b2). In particular, a model for decision support (b2) has to predict the output variables (attributes) used to quantify the preferences (b1) and it has to describe the dependence of these on input variables that distinguish the decision alternatives (Schuwirth *et al.* 2019).

In the following sections, conceptual aspects of uncertainty quantification and model building are discussed before moving on to models for predicting the consequences of decision alternatives and models for quantifying preferences. Then a decision support framework with comprehensive uncertainty quantification is described.

This is followed by a short section on numerical approaches and software to implement the suggested techniques. Finally, conclusions are drawn about the transfer of the suggested techniques to research and practice.

CONCEPTUAL ASPECTS OF UNCERTAINTY QUANTIFICATION AND MODEL BUILDING

The need for Bayesian techniques

The most important conceptual decision is on how to describe uncertain knowledge. Bayesian (epistemic) probabilities are the most straightforward choice as they provide a consistent framework for conditional beliefs and iterative learning and they are compatible with adopting randomness (aleatory probabilities) as part of the uncertain knowledge about future outcomes (Reichert *et al.* 2015).

As environmental decisions should be based on the best available current state of scientific knowledge, prior knowledge must be carefully elicited as intersubjective knowledge (Gillies 1991; Gillies 2000; Reichert *et al.* 2015). This is implemented in practice by using well-designed eliciting techniques (Morgan & Henrion 1990; Meyer & Booker 2001; O'Hagan *et al.* 2006; Rinderknecht *et al.* 2011) to elicit priors, combining the assessments from multiple experts, and by carefully documenting the use of prior information from the literature to get justifiable priors.

Beyond prior times likelihood – considering intrinsic uncertainty

The use of Bayesian probabilities to describe scientific knowledge is naturally linked to using Bayesian inference to describe a potentially iterative learning process from data. Bayesian inference is often introduced by the statement 'the posterior probability is proportional to the prior probability times the likelihood':

$$p(\theta|\mathbf{y}_{\text{obs}}, \mathbf{x}) \propto p(\theta) p(\mathbf{y}_{\text{obs}}|\theta, \mathbf{x}) \quad (1)$$

Here, θ are model parameters; \mathbf{y}_{obs} represent potentially observed states; $p(\mathbf{y}_{\text{obs}}|\theta, \mathbf{x})$ is the probabilistic model for observations conditional on the input, \mathbf{x} , and the parameter values (called likelihood function if viewed as a function of the parameters with actual observed values substituted for \mathbf{y}_{obs}); $p(\theta)$ is the probability distribution that quantifies the prior knowledge about the model parameters; and $p(\theta|\mathbf{y}_{\text{obs}}, \mathbf{x})$ is the probability distribution representing the updated,

posterior knowledge about the parameters given the observations. Although Equation (1) is correct in the right context (learning about model parameters), its interpretation ('prior times likelihood') obscures:

- (i) that in nearly all practical applications, the main part of the prior knowledge is formulated by the probabilistic model (likelihood), which is needed to define the meaning of the parameters (or may not have parameters);
- (ii) that one is usually interested in predicting future or conditional states, \mathbf{y}_{new} , (conditional on changes in influence factors or potential measures suggested as decision alternatives), $p(\mathbf{y}_{\text{new}}|\mathbf{y}_{\text{obs}}, \mathbf{x})$, rather than in model parameters (which are just an auxiliary tool for an intermediate step); and
- (iii) that, from a statistical perspective, there is often no fundamental difference between model parameters and unobserved states and Bayesian inference can be applied to condition the joint distribution of observed and unobserved model variables on those observed to get an update of those unobserved.

In the context above, it would be better to state: 'our joint prior knowledge of parameters and observed states is equal to the likelihood times the prior of its parameters; conditioning this joint probability on the observations leads to the posterior of the parameters'.

However, this does not cover the power of Bayesian analysis, as we can formulate much more powerful models that consider intrinsic uncertainty. An example is the consideration of uncertain parameters and internal states in a hierarchical model, e.g.:

$$\begin{aligned}
 p(\mathbf{y}_{\text{obs}}, \mathbf{y}, \mathbf{z}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\xi}|\mathbf{x}) \\
 &= p(\mathbf{z}|\mathbf{x}, \boldsymbol{\zeta})p(\boldsymbol{\zeta}) && \text{model of internal states } \mathbf{z} \\
 &\times p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{z}) && \text{model of true outputs } \mathbf{y} \\
 &\times p(\boldsymbol{\theta}|\boldsymbol{\psi})p(\boldsymbol{\psi}) && \text{model of parameters } \boldsymbol{\theta} \\
 &\times p(\mathbf{y}_{\text{obs}}|\mathbf{y}, \boldsymbol{\xi})p(\boldsymbol{\xi}) && \text{model of observed outputs } \mathbf{y}_{\text{obs}} \quad (2)
 \end{aligned}$$

This distribution formulates a joint distribution of unobserved internal states (\mathbf{z}), true outputs (\mathbf{y}), observed outputs (\mathbf{y}_{obs}), and parameters of the system model ($\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\psi}$) and of the observation model ($\boldsymbol{\xi}$) as a hierarchical model (see e.g. Clark 2005). The factorization of the joint distribution into conditionals given by Equation (2) decomposes the model into submodels for the investigated system, the parameters, and the observation process, as illustrated in Figure 1. This considerably facilitates model construction because the submodels are less complex as they describe more specific

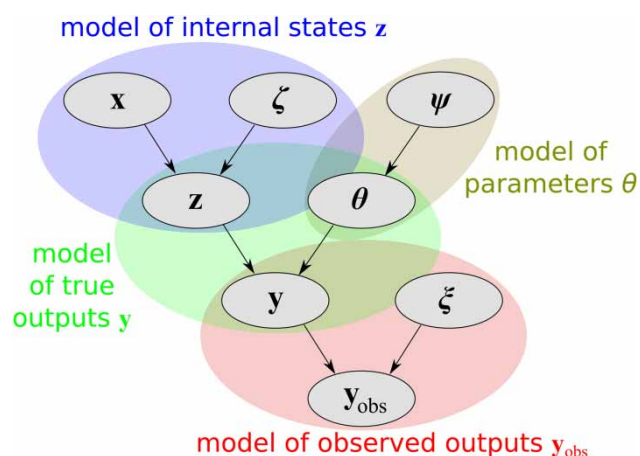


Figure 1 | Graphical illustration of the decomposition of a model into sub-models by the model structure given in Equation (2). Please refer to the online version to see this figure in colour: <http://dx.doi.org/10.2166/wst.2020.032>.

subsystems with a smaller number of input and output variables and parameters.

Model construction based on a graphical model as shown in Figure 1 and formalized by Equation (2) became prominent as 'Bayesian belief network modelling' (see e.g. Borsuk *et al.* 2004). However the underlying concept applies to all probabilistic models and is also known as 'hierarchical modelling' (with a slight variation in focus, see e.g. Clark 2005). Note that also 'structural equations modelling' (see e.g. Kline 2011) is based on very similar concepts (although initially limited to normal distributions). In the context of time-series models, again very similar model structures are known as 'state-space models' (see e.g. Künsch 2001), 'hidden Markov models' (see e.g. Künsch 2001) or 'dynamic Bayesian belief network models' (see e.g. Murphy 2002). Sometimes, those different terminologies for essentially the same underlying concepts can be confusing.

Imprecise probabilities

Due to sparse data, the use of prior information is often essential in environmental modelling and decision support. However, prior knowledge is often uncertain and it is important to analyse the robustness of conclusions regarding prior assumptions. Imprecise probabilities provide an ideal framework to do this. The concept of imprecise probabilities is to replace single probability distributions by sets of probability distributions to make results more robust regarding distributional assumptions (e.g. DeRobertis & Hartigan 1981; Berger 1990; Walley 1991; Rinderknecht *et al.* 2012, 2014). The following outline is based on the

density ratio class to formulate imprecise probabilities (DeRobertis & Hartigan 1981; Rinderknecht *et al.* 2012, 2014). This class is defined by constraining the shape of probability densities by a lower, l , and an upper, u , non-normalized probability density and then normalizing the shapes in between:

$$\Gamma_{l,u}^{\text{DR}} = \left\{ f(\theta) = \frac{\tilde{f}(\theta)}{\int \tilde{f}(\theta') d\theta'} \mid l(\theta) \leq \tilde{f}(\theta) \leq u(\theta) \forall \theta \right\} \quad (3)$$

This concept is illustrated in Figure 2. The left panel illustrates how shapes of non-normalized probability densities (green) are bounded by non-normalized upper (u , red) and lower (l , blue) non-normalized densities. The right panel illustrates the non-normalized density (green, dashed) with the highest probability in the interval $[\theta_1, \theta_2]$.

This class of probability distributions was chosen because of its invariance under Bayesian updating, under marginalization, and under propagation through deterministic functions (see Rinderknecht *et al.* 2014) which makes sequential learning possible within the same framework.

MODELS FOR UNDERSTANDING AND PREDICTION

Exploring data

Data science methodologies provide new opportunities for exploring data (see e.g. LeCun *et al.* 2015; or more specifically regarding application in ecology and water resources research, Peters *et al.* 2014; Shen 2018). Even as the main purpose of these methods is exploration and prediction without explicitly considering mechanisms, these methods can contribute to a primarily mechanistically oriented modelling process by:

- developing ‘prediction models’ for aspects for which understanding is not as important as for other aspects

(e.g. image recognition from remote sensing data or plankton species identification to compile input data for mechanistic models);

- identifying patterns to stimulate mechanistic model development (e.g. finding clusters of organisms that have similar properties or behave similarly);
- developing ‘black box’ models that serve as benchmarks to analyse the potential for the improvement of more sparsely parameterized mechanistic models (Kratzert *et al.* 2019a);
- trying to interpret the functioning of the model developed by applying machine learning techniques in the sense of ‘interpretable data science’ (e.g. Papernot & McDaniel 2018; Gilpin *et al.* 2019; Kratzert *et al.* 2019b);
- constraining the model or the learning algorithm to consider physical characteristics of the system to facilitate interpretation and learning or developing ‘hybrid models’ that bridge between mechanistic and data-based models (‘theory-guided data science’, e.g. Karpatne *et al.* 2017).

Constructing stochastic, mechanistic models

Wherever possible, models that are intended to predict beyond their calibration range should be designed to represent underlying mechanisms. Hardly any environmental system behaves deterministically at the level at which we observe its behaviour. Reasons for this can be true stochasticity resulting from quantum-mechanical processes, demographic stochasticity of birth and death processes, genetic stochasticity, or apparent non-deterministic behaviour caused by the limited temporal and spatial resolution of the initial state of the modelled system and of external driving forces. For these reasons, to get an appropriate description of system behaviour and of its uncertainty, we need stochastic, mechanistic models. Environmental stochasticity can be considered by making inputs and/or model parameters stochastic processes in time (see e.g. Reichert & Mieleitner

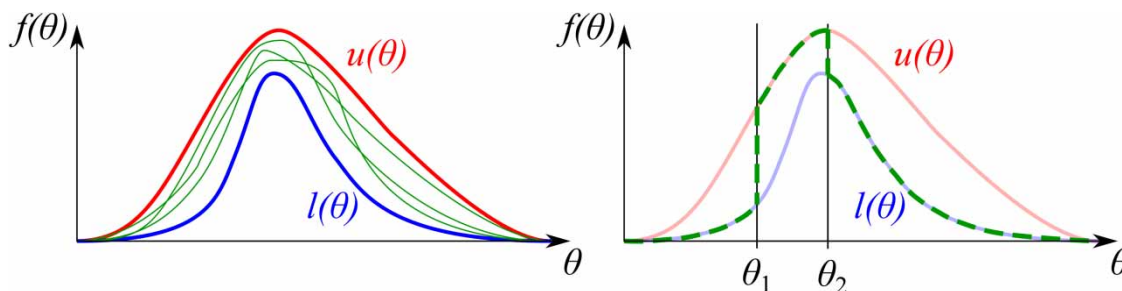


Figure 2 | One-dimensional illustration of the concept of imprecise (sets of) probabilities as implemented by the density ratio class. See text for more detailed explanations. Please refer to the online version to see this figure in colour: <http://dx.doi.org/10.2166/wst.2020.032>.

2009). Typically, considering stochasticity leads to hierarchical models that complement unobserved parameters with unobserved states as illustrated in Equation (2) and Figure 1. As shown by Chou & Greenman (2016) for a density-dependent, age-structured population model, it can be inconsistent if one tries to formulate deterministic models directly: they may not necessarily represent the development of the mean of a (more realistic) stochastic model. Ideally, stochastic modelling is combined with a multi-model approach to account for the consequences of structural uncertainty on predictions even better.

MODELS OF PREFERENCES

Multi-criteria decision analysis (MCDA)

The multi-criteria decision analysis (MCDA) methodology (Keeney & Raiffa 1976; Keeney 1992; Eisenführ *et al.* 2010) provides an ideal framework for modelling preferences for environmental decision support (see e.g. Reichert *et al.* 2015).

In this framework, preference modelling is based on constructing a ‘value function’ that quantifies the degree of fulfilment of the overall objective on a scale from 0 to 1 as a function of ‘attributes’ that characterize the system under analysis. Preference elicitation starts with breaking down the main objective hierarchically into sub-objectives that complementarily and exhaustively characterize the corresponding objective at the higher level. Such an objectives hierarchy can then be used to construct the value function of the overarching objective by constructing value functions of the lowest level objectives as functions of relevant system attributes, and value functions of higher-level objectives by constructing ‘aggregation functions’ depending on the

degrees of fulfilment (values) of the corresponding sub-objectives (Grabisch *et al.* 2011; Reichert *et al.* 2019). This leads to a value function of the main objective that, through the aggregation functions and the lowest-level value functions, depends on all attributes used to characterize the degrees of fulfilment of all lowest level sub-objectives. Equation (4) and Figure 3 show an example of the construction of the value function, v , for the main objective that depends through the value functions for lowest-level sub-objectives v_1 , v_2 , v_{3a} , and v_{3b} and the aggregation functions \tilde{v}_3 and \tilde{v} on the attributes y_1 to y_5 :

$$v(y_1, y_2, y_3, y_4, y_5) = \tilde{v}(v_1(y_1), v_2(y_2, y_3), \tilde{v}_3(v_{3a}(y_4), v_{3b}(y_5))) \quad (4)$$

Similarly to the case of the outcome prediction model (Equation (2) and Figure 1) this decomposition of the overall value function into value functions of lowest-level sub-objectives and aggregation functions at higher levels facilitates the construction of the overall value function by requiring less complex functions at each decomposition step.

The value function of the main objective can then be transformed into a ‘utility function’ by the consideration of risk attitudes (for more details, see Keeney & Raiffa 1976; Dyer & Sarin 1982; Eisenführ *et al.* 2010). Given the utility function, u , of the main objective and the probability distributions, $p_a(\mathbf{y})$, of the system attributes, \mathbf{y} , from probabilistically modelling the consequences of each alternative, a , the expected utilities, EU, can be calculated for all alternatives:

$$EU(a) = \int_{\mathbf{y}} u(\mathbf{y}) p_a(\mathbf{y}) d\mathbf{y}, \text{ where} \\ p_a(\mathbf{y}) = \int_{\theta} p(\mathbf{y}|\theta) p_a(\theta) d\theta \quad (5)$$

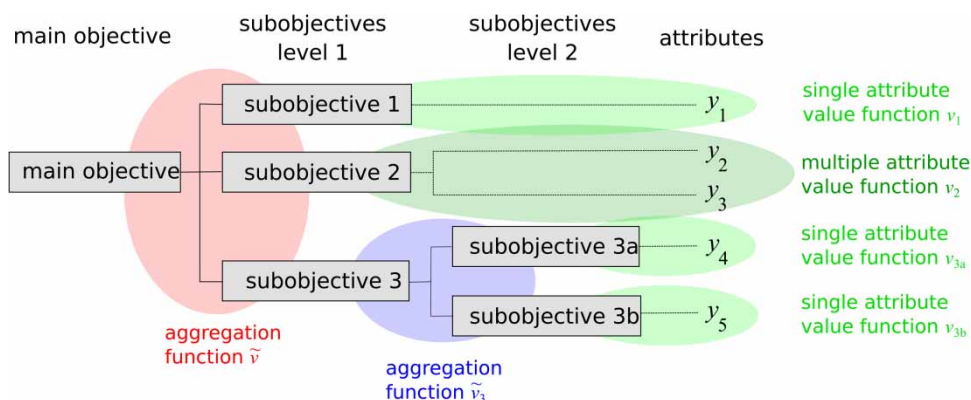


Figure 3 | Example of a simple objectives hierarchy with the corresponding value function given in Equation (3). Please refer to the online version to see this figure in colour: <http://dx.doi.org/10.2166/wst.2020.032>.

Here, $p_a(\theta)$ is the prior or posterior (if the prior was updated) parameter distribution of model parameters for alternative a and there may be additional integration across internal variables and states if the model is hierarchical. Decision support is then based on ranking the alternatives according to decreasing values of their expected utilities.

Expected expected utilities

This classical MCDA framework can be extended to ‘expected expected utilities’ to consider uncertainty in the elicited stakeholder preferences (Cyert & de Groot 1979; Boutilier 2003; Haag *et al.* 2019b). Here, the utilities, $u(\mathbf{y}, \boldsymbol{\theta})$, are parameterized with parameters, $\boldsymbol{\theta}$, the uncertainty of which is described by their distribution, $p(\boldsymbol{\theta})$. Utilities are only defined up to an affine transformation (e.g. Eisenführ *et al.* 2010). To compare utilities we need to ‘standardize’ all utilities of the uncertain set at extreme values (typically to 0 and 1 for the worst and best outcomes). If this is done, the alternatives can be ranked according to their expected expected utility, EEU (Boutilier 2003):

$$\begin{aligned} \text{EEU}(a) &= \int_{\boldsymbol{\theta}} \text{EU}(a, \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ &= \int_{\boldsymbol{\theta}} \left(\int_{\mathbf{y}} u(\mathbf{y}, \boldsymbol{\theta}) p_a(\mathbf{y}) d\mathbf{y} \right) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \end{aligned} \quad (6)$$

This framework leads to a unique ranking that considers the uncertain utilities. As outlined in the next section, combining this framework with imprecise probabilities makes it possible to assess the ambiguity of this ranking resulting from the ambiguity about the probability distributions used to quantify uncertainty of outcome predictions as well as values and utilities; this adds an essential new element to the analysis.

A COMPREHENSIVE FRAMEWORK FOR UNCERTAINTY ANALYSIS IN DECISION SUPPORT

By combining the ‘standard steps’ of decision support, such as problem definition, stakeholder analysis, elicitation of an objectives hierarchy, etc (see e.g. Reichert *et al.* 2015) with the techniques outlined above, a comprehensive uncertainty analysis can be performed by modifying the following three steps.

Elicitation and construction of value and utility functions

When using parameterized value functions and a parameterized conversion function to utilities, parameter estimation can be done based on elicited discrete choice selections (Hoyos 2010) or indifference replies (Haag *et al.* 2019a) using Bayesian inference with density ratio class priors. This leads to a posterior density ratio class of the value/utility parameters (Rinderknecht *et al.* 2014) that jointly describes uncertainty and the ambiguity about its quantification. If a multi-model approach was performed, either the best fitting value model can be selected or multiple models considered for further analysis.

Prediction of outcomes of decision alternatives

For predicting the outcomes of the decision alternatives, either density ratio class priors (for their elicitation, see e.g. Rinderknecht *et al.* 2011) can be used directly, or they can be updated by Bayesian inference if data is available. In the latter case, this also leads to a (posterior) density ratio class of the parameters (Rinderknecht *et al.* 2014). Again, a multiple model approach is useful for considering structural uncertainty.

Compilation of results

Evaluating the expected expected utilities in Equation (6) for imprecise probability distributions of $p_a(\mathbf{y})$ and $p(\boldsymbol{\theta})$ leads to an interval of expected expected utilities $\text{EEU}(a)$, for each alternative, a . These intervals for all alternatives lead to an incomplete ordering of the decision alternatives that reflects the ambiguity in addition to the uncertainty. However, this seems to be the best representation of knowledge and uncertainty and thus provides the best information to be communicated to decision makers.

NUMERICAL CHALLENGES AND SOFTWARE

Efficient numerical algorithms, mainly based on Markov Chain Monte Carlo (MCMC) techniques, such as Metropolis-Hastings sampling (see e.g. Gelman *et al.* 2013), have been developed to sample from posteriors in Bayesian inference for cases in which the likelihood function can easily be evaluated. This is no longer true for hierarchical models that require computationally much more demanding techniques, such as Gibbs sampling

(see e.g. Gelman *et al.* 2013), Particle Markov Chain Monte Carlo (PMCMC; see e.g. Andrieu *et al.* 2010), Hamiltonian Monte Carlo (HMC; see e.g. Betancourt 2017), or Approximate Bayesian Computation (ABC; see e.g. Beaumont 2010; Albert *et al.* 2015). Krapu & Borsuk (2019) provide an overview of software designed for this purpose.

On the other hand, the extension to imprecise probabilities based on the density ratio class does not add a lot of additional computational effort, as a (unweighted) sample of a single distribution from the class can easily be turned into a weighted sample of any other distribution from the class without having to do inference again (Rinderknecht *et al.* 2014). Still calculating the EEU-intervals requires the evaluation of expected utilities for a large sample of elements from the class.

CONCLUSIONS

An ideal framework for environmental decision support would be to combine mechanistic, stochastic models for prediction with uncertain utilities for quantifying preferences and to use imprecise probabilities for robustness analysis. We outlined how this can be done without significantly increasing the computational burden. It would be a very interesting next step to apply this procedure to an actual decision problem and to analyse the benefits and challenges of considering all sources of uncertainty. In particular, it would be interesting to explore whether decision makers are willing to accept the outcome in the form of an incomplete ranking based on intervals of expected utility rather than a unique ranking. Easy-to-apply software could significantly contribute to facilitating the application of the suggested approach in practice.

ACKNOWLEDGEMENTS

The development of this paper has benefited from discussions with many scientists over the last few years. I would like to mention in particular Nele Schuwirth, Simon Lukas Rinderknecht, Johanna Mieleitner, Carlo Albert, Simone Langhans, Fridolin Haag, and Judit Lienert. This paper is dedicated to my former PhD students Johanna Mieleitner (1975–2019) and Simon Lukas Rinderknecht (1979–2019), who contributed significantly to the development of some of the techniques discussed in this paper.

REFERENCES

- Albert, C., Künsch, H. R. & Scheidegger, A. 2015 *A simulated annealing approach to approximate Bayes computations. Statistics and Computing* **25**, 1217–1232. <https://doi.org/10.1007/s11222-014-9507-8>.
- Andrieu, C., Douce, A. & Holenstein, R. 2010 *Particle markov chain Monte Carlo methods. Journal of the Royal Statistical Society: Series B* **72** (Part 3), 269–342. <https://doi.org/10.1111/j.1467-9868.2009.00736.x>.
- Beaumont, M. 2010 *Approximate Bayesian computation in evolution and ecology. Annual Review of Ecology, Evolution, and Systematics* **41**, 379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>.
- Berger, J. O. 1990 *Robust Bayesian analysis: sensitivity to the prior. Journal of Statistical Planning and Inference* **25**, 303–328.
- Betancourt, M. 2017 *A Conceptual Introduction to Hamiltonian Monte Carlo. arXiv* 1701.02434. <https://arxiv.org/abs/1701.02434>.
- Borsuk, M. E., Stow, C. A. & Reckhow, K. H. 2004 *A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecological Modelling* **173**, 219–239. <http://doi.org/10.1016/j.ecolmodel.2003.08.020>.
- Boutilier, C. 2003 *On the foundations of expected utility. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc., Acapulco, Mexico*, pp. 285–290.
- Chou, T. & Greenman, C. D. 2016 *A hierarchical kinetic theory of birth, death and fission in age-structured interacting populations. Journal of Statistical Physics* **164**, 49–76. <http://doi.org/10.1007/s10955-016-1524-x>.
- Clark, J. S. 2005 *Why environmental scientists are becoming Bayesians. Ecology Letters* **8**, 2–14. <http://doi.org/10.1111/j.1461-0248.2004.00702.x>.
- Cyert, R. M. & de Groot, M. H. 1979 *Adaptive utility. In: Expected Utility Hypotheses and the Allais Paradox (M. Allais & O. Hagen, eds). D. Reidel, Dordrecht*, pp. 223–241.
- DeRobertis, L. & Hartigan, J. A. 1981 *Bayesian inference using intervals of measures. Annals of Statistics* **9** (2), 235–244.
- Dyer, J. S. & Sarin, R. K. 1982 *Relative risk aversion. Management Science* **28** (8), 875–886.
- Eisenführ, F., Weber, M. & Langer, T. 2010 *Rational Decision Making. Springer, Berlin, Germany*.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. 2013 *Bayesian Data Analysis*, 3rd edn. Chapman & Hall/CRC Press, Boca Raton, FL.
- Gillies, D. 1991 *Intersubjective probability and Confirmation theory. British Journal for the Philosophy of Science* **42**, 513–533.
- Gillies, D. 2000 *Philosophical Theories of Probability. Routledge, London, UK*.
- Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M. & Kagal, L. 2019 *Explaining explanations: an overview of interpretability of machine learning. https://arxiv.org/pdf/1806.00069.pdf*.

- Grabisch, M., Marichal, J.-L., Mesiar, R. & Pap, E. 2011 *Aggregation functions: means*. *Information Sciences* **181**, 1–22. <https://doi.org/10.1016/j.ins.2010.08.043>.
- Haag, F., Lienert, J., Schuwirth, N. & Reichert, P. 2019a *Identifying non-additive multi-attribute value functions based on uncertain indifference statements*. *Omega* **85**, 49–67. <http://doi.org/10.1016/j.omega.2018.05.011>.
- Haag, F., Reichert, P., Maurer, M. & Lienert, J. 2019b *Integrating uncertainty of preferences and predictions in decision models: an application to regional wastewater planning*. *Journal of Environmental Management* **252**, 109652. <http://doi.org/10.1016/j.jenvman.2019.109652>.
- Hoyos, D. 2010 *The state of the art of environmental valuation with discrete choice experiments*. *Ecological Economics* **69**, 1595–1603. <https://doi.org/10.1016/j.ecolecon.2010.04.011>.
- Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N. & Kumar, V. 2017 *Theory-guided data science: a new paradigm for scientific discovery from data*. *IEEE Transactions on Knowledge and Data Engineering* **29** (10), 2318–2331. <http://doi.org/10.1109/TKDE.2017.2720168>.
- Keeney, R. L. 1992 *Value-focused Thinking: A Path to Creative Decisionmaking*. Harvard University Press, Cambridge, Mass.
- Keeney, R. L. & Raiffa, H. 1976 *Decision with Multiple Objectives*. Wiley, New York.
- Kline, R. 2011 *Principles and Practice of Structural Equation Modeling*, 3rd edn. Guilford, New York, London.
- Krapu, C. & Borsuk, M. 2019 *Probabilistic programming: a review for environmental modellers*. *Environmental Modelling and Software* **114**, 40–48. <https://doi.org/10.1016/j.envsoft.2019.01.014>.
- Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S. & Nearing, G. S. 2019a *Toward improved predictions in ungauged basins: exploiting the power of machine learning*. *Water Resources Research* **55** (12), 11344–11354. <https://doi.org/10.1029/2019WR026065>.
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S. & Nearing, G. 2019b *Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets*. *Hydrology and Earth System Sciences* **23**, 5089–5110. <https://doi.org/10.5194/hess-23-5089-2019>.
- Krueger, T., Page, T., Hubacek, K., Smith, L. & Hiscock, K. 2012 *The role of expert opinion in environmental modelling*. *Environmental Modelling & Software* **36**, 4–18. <https://doi.org/10.1016/j.envsoft.2012.01.011>.
- Künsch, H. R. 2001 State space and hidden Markov models. Chapter 3. In: *Complex Stochastic Systems* (O. E. Barndorff-Nielsen, D. R. Cox & C. Klüppelberg, eds). Chapman & Hall/CRC, Boca Raton, FL.
- LeCun, Y., Bengio, Y. & Hinton, G. 2015 *Deep learning*. *Nature* **521**, 436–444. <http://doi.org/10.1038/nature14539>.
- Meyer, M. A. & Booker, J. M. 2001 *Eliciting and Analyzing Expert Judgment*. SIAM/ASA, Philadelphia/Alexandria.
- Morgan, M. G. & Henrion, M. 1990 *Uncertainty – A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge, UK.
- Murphy, K. P. 2002 *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD Dissertation, University of California Berkeley.
- O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garwite, P. H., Jenkinson, D. J., Oakley, J. E. & Rakow, T. 2006 *Uncertain Judgements: Eliciting Expert's Probabilities*. Wiley, Chichester, GB.
- Papernot, N. & McDaniel, P. 2018 *Deep k-nearest neighbors: towards confident, interpretable and robust deep learning*. <https://arxiv.org/abs/1803.04765>.
- Peters, D. P. C., Havstad, K. M., Cushing, J., Tweedie, C., Fuentes, O. & Villanueva-Rosales, N. 2014 *Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology*. *Ecosphere* **5** (6), 67. <https://doi.org/10.1890/ES13-00359.1>.
- Refsgaard, J. V., van der Sluijs, J. P., Hojberg, A. J. & Vanrolleghem, P. A. 2007 *Uncertainty in the environmental modelling process – A framework and guidance*. *Environmental Modelling & Software* **22**, 1543–1556. <https://doi.org/10.1016/j.envsoft.2007.02.004>.
- Reichert, P. & Mieleitner, J. 2009 *Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters*. *Water Resources Research* **45**, W10402. <http://doi.org/10.1029/2009WR007814>.
- Reichert, P., Langhans, S., Lienert, J. & Schuwirth, N. 2015 *The conceptual foundation of environmental decision support*. *Journal of Environmental Management* **154**, 316–332. <http://doi.org/10.1016/j.jenvman.2015.01.053>.
- Reichert, P., Niederberger, K., Rey, P., Helg, U. & Haertel-Borer, S. 2019 *The need for unconventional value aggregation techniques: experiences from eliciting stakeholder preferences in environmental management*. *EURO Journal on Decision Processes* **7**, 197–219. <https://doi.org/10.1007/s40070-019-00101-9>.
- Rinderknecht, S. L., Borsuk, M. E. & Reichert, P. 2011 *Eliciting density ratio classes*. *International Journal of Approximate Reasoning* **52**, 792–804. <http://doi.org/10.1016/j.ijar.2011.02.002>.
- Rinderknecht, S. L., Borsuk, M. E. & Reichert, P. 2012 *Bridging uncertain and ambiguous knowledge with imprecise probabilities*. *Environmental Modelling & Software* **36**, 122–130. <http://doi.org/10.1016/j.envsoft.2011.07.022>.
- Rinderknecht, S. L., Albert, C., Borsuk, M. E., Schuwirth, N., Künsch, H. R. & Reichert, P. 2014 *The effect of ambiguous prior knowledge on Bayesian model parameter inference and prediction*. *Environmental Modelling & Software* **62**, 300–315. <http://doi.org/10.1016/j.envsoft.2014.08.020>.
- Schuwirth, N., Borgwardt, F., Domisch, S., Friedrichs, M., Kattwinkel, M., Kneis, D., Kuemmerlen, M., Langhans, S. D., Martínez-López, J. & Vermeiren, P. 2019 *How to make ecological models useful for environmental management?* *Ecological Modelling* **411**, 108784. <https://doi.org/10.1016/j.ecolmodel.2019.108784>.
- Shen, C. 2018 *A transdisciplinary review of deep learning research and its relevance for water resources scientists*. *Water Resources Research* **54**, 8558–8593. <http://doi.org/10.1029/2018WR022643>.

Voinov, A. & Bousquet, F. 2010 [Modelling with stakeholders](#). *Environmental Modelling & Software* **25** (11), 1268–1281. <https://doi.org/10.1016/j.envsoft.2010.03.007>.

Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., Pierce, S. A. & Ramu, P. 2016

[Modelling with stakeholders – next generation](#).

Environmental Modelling & Software **77**, 196–220. <http://dx.doi.org/10.1016/j.envsoft.2015.11.016>.

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

First received 2 October 2019; accepted in revised form 21 January 2020. Available online 29 January 2020