Abstract: The practical impact of treatment of epistemic uncertainty on decision making was illustrated on two kinds of decisions from chemical regulation. First, regulatory strategies derived from a simplified decision model based on toxicity and persistence showed that regulated level of exposure is more conservative (safe) when uncertainty has been given a non-probabilistic treatment. Persistence and its uncertainty had been assessed by a Level II fugacity model for which input parameters had been quantified either by Bayesian probabilities, fuzzy numbers (non-probabilistic), or combinations of these (probability boxes). These findings are restricted to how we let decision makers respond to uncertainty in model predictions by the chosen set of decision rules. Further, the use of either treatment depends on the quality and quantity of background knowledge and the required level of detail on the assessment. In the absence of experimentally tested physicochemical endpoints, European chemical regulation REACH allows the use of non-testing strategies such as Quantitative Structure-Property Relationships (QSPR) to predict the required information. The second decision problem was to select which chemical substances to prioritize for experimental testing in order to strengthen the background knowledge for chemical regulation with respect to the uncertainty in QSPR predictions. We found that the value of reducing uncertainty, given by the expected gain in net benefit for society, was affected by its treatment and there were no consistent order of testing of the three compounds. However, value of information is a Bayesian probabilistic approach that, unless developed further, loose its interpretability under other treatments of uncertainty. The framework of a predictive model, risk model, decision model and value of information analysis provides a computational template for further evaluation of the effect of treatment of uncertainty on decision making.

Keywords: Predictive Uncertainty, Value of Information, Chemical Regulation, Quantitative Structure-Property Relationships

1. INTRODUCTION

1.1. Uncertainty in Decision Making
Decision making under uncertainty, with the goal to consider not only risk but also uncertainty following from lack of knowledge (epistemic uncertainty, ignorance), has raised the need of quantitative methods able to deal with different kinds of uncertainties (Kinzig et al., 2003, Aven, 2010b). Probabilistic approaches quantify uncertainty by probabilities, and are recommended when background knowledge is strong enough to allow uncertainty be given in terms of likelihoods or degrees of beliefs (Aven, 2010a). When uncertainty expresses our lack of knowledge, in situations of weak background knowledge it is suggested to give uncertainty a non-probabilistic treatment (Ferson and Ginzburg, 1996). For instance, uncertainty quantified as fuzzy numbers is a non-probabilistic and quantitative approach which gives all values within a specified interval with the same consideration (Zadeh, 1999). How to treat uncertainty is context dependent, and it may be that combinations of probabilistic and non-probabilistic approaches, such as Probability Bounds Analysis (PBA) (Ferson et al.), are the most flexible to meet varying strength of the background knowledge. By treatment of uncertainty we mean its quantification based on given background knowledge, its propagation in assessment models and corresponding rules for decision making.
1.2. Chemical Regulation

Chemical Safety Assessment (CSA) (Environmental Risk Assessment in the US EPA) aim to identify hazardous and safe compounds based on the ecotoxicological paradigm comparing exposure and effects (European Chemical Agency, 2008, National Research Council, 2009). Risk is quantified as the ratio of Predicted Environmental Concentration (PEC) and No Effect Concentration (or Lowest Effect Concentration) NOEC (LOEC). A CSA is done in three tiers, first a deterministic assessment, secondly an assessment consider worst case and uncertainty factors, and third an assessment where all uncertainties are quantified and propagated through the assessment in an uncertainty analysis. Risk assessments are carried out given the available background knowledge, which for CSA partly consist of experimental tests of relevant health or ecotoxicological endpoints of a chemical substance. It is common practice to add uncertainty factors for every kind of extrapolations from empirical data, but this approach has been criticized to add unnecessarily large uncertainty in risk assessments which may lead to unnecessary restrict chemical regulation. Quantitative methods to treat uncertainty are useful under various strength of background knowledge, such as in all tiers of a CSA.

According to the European legislation REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) all chemical substances produced on the European market with annual tonnage above 1 ton must have been registered before 2018 (EC, 2006). The information requirements for registration increase with the volume of production, but must at least include substances physicochemical information for human health and environmental hazard assessment. This means that information must be gathered on several tens of thousands of substances. The evaluation of these registrations is done substance-wise and will require many years to complete. In order to speed up risk assessment, cut down costs and reduce animal testing, non-testing strategies are allowed for CSA (Jaworska and Hoffmann, 2010). A non-testing strategy is to use predictions from Quantitative Structure-Activity (Property) Relationships (QSAR and QSPR) that predict by analogy saying that similar chemicals have similar chemical activities or properties (Cronin et al., 2003). Since non-testing strategies provide information at a lower strength compared to experimental data, a recommendation is for example to only allow predictive models that fulfill the OECD principles (OECD, 2006). Predictive chemical models can be used to screen for hazardous substances, predict input parameters in risk models, such as chemical-specific properties in environmental fate models, or to suggest prioritization of testing (Cronin et al., 2003).

Testing strategies are designed to provide support to a decision problem on which compounds to select for experimental testing. Experimental testing is costly and therefore tools for prioritization of which compounds to test first are useful (Jaworska and Hoffmann, 2010). A testing strategy can be designed to prioritize substances believed to be of high risk e.g. based on exposure based waiving or triggering. Another criterion is to prioritize compounds that are representative for the chemical space for training models to predict the remaining compounds, for which methods for experimental design are useful. A third criterion is to select compounds for which a reduction in uncertainty in relevant chemical properties, following from experimental testing, lead to the largest gains in net benefit to society (Value of information analysis) (Gabbert and Weikard, 2010). Without an experimental test information is either imperfect due to available prior knowledge of some kind, or information is totally absent. Therefore the strength of information, as determined by the extent of uncertainty, may vary between chemicals for which a testing strategy is to be developed. A value of information analysis is able to consider both risk and uncertainty at the same time, and may therefore diverge in results compared to prioritization of compounds for testing that solemnly are based on risk estimates.
Value of information analysis is a probabilistic method that jointly considers risk and uncertainty. A way to judge the value of reducing uncertainty would be useful even under non-probabilistic treatments of uncertainty, in order to address the value of reducing uncertainty under weak background knowledge, such as in early stages of risk assessment or when designing strategies to increase the strength of the background knowledge.

1.3. Objectives
Encouraged by the request to illustrate the practical impact of the treatment of epistemic uncertainty on decision making, examples from chemical regulation were used to demonstrate if such treatment had an impact on two kinds of decisions. First, regulatory strategies derived from a simplified decision model, parameterized for three highly environmental concern chemical compounds, made it possible to study how the regulated level of exposure of a chemical compound changed under alternative treatments of uncertainty. Second, which chemical compounds to prioritize for experimental testing in order to strengthen the background knowledge for chemical regulation, as judged by the value of reducing uncertainty was compared under alternative treatments of uncertainty in predicted key chemical-specific properties of environmental fate for the same three chemical compounds.

2. MODEL AND METHODS

2.1. Decision Model
Gabbert and Weikard (2010) proposed a decision model to rank chemicals according to “urgency to test”, considering weight of evidence on toxicity and persistence and uncertainty in these endpoints. To this end, they formulated a simple decision model for the regulatory exposure levels of a substance based on toxicity and persistence. Toxicity $\tau$ ($0 \leq \tau \leq 1$) is assumed to be a continuous variable that captures the damage potential of a substance. Damage $D$ are assumed to increase linearly with exposure and toxicity according to

$$D(e, \tau) = e \tau.$$  

(1)

Exposure is determined by several factors such as the production volume of the chemical, the way it is used, the safety measures and regulations adopted, and the persistence of the substance. Persistence $\rho$ ($0 \leq \rho \leq 1$) indicates how a substance will accumulate in the environment. Benefit from the use of a chemical is denoted by $B(e)$, where the benefit from not using the chemical is zero (i.e. $B(0) = 0$). Benefit is modeled such as safety measures, i.e. the reduction of exposure, reduces net benefits. In order to obtain analytical results, they achieved a decline in marginal benefits with exposure by modeling benefit as a quadratic function given by

$$B(e) = \beta e^2 (-e^2 + 2\varepsilon e),$$  

(2)

where $\beta$ is a scaling factor that reflects the social welfare that could be obtained from the use of the substance in the absence of regulations if we do not account for damages, and $\varepsilon$ is a parameter that captures the satiation level of the use of the substance (i.e. the level of exposure at which a substance would be used if it were non-toxic). Hence, $\varepsilon$ is the optimal level of exposure for a non-toxic substance or the unregulated level of exposure when damages are externalities.

The decision model was specified to consider exposure levels during two periods of regulation. For a persistent chemical the use of a chemical during the first period $e_1$ adds an amount $\rho e_1$ to the next period use $e_2$. The optimal exposure is then found by maximizing discounted net benefits

$$\max_{e_1, e_2} \{V = B(e_1) - D(e_1, \tau) + \delta(B(e_2) - D(\rho e_1 + e_2, \tau))\},$$  

(3)
where $\delta$ is a discount factor. The optimal regulation strategy for a persistent chemical is then under perfect information (i.e. when there is no uncertainty)

$$
e_{1}^{*} = e - \tau \varepsilon / 2\beta (1 + \delta \rho),$$  \hspace{1cm} (4)

and

$$
e_{2}^{*} = e - \tau \varepsilon / 2\beta,$$  \hspace{1cm} (5)

Here $\tau, \rho, \beta$ and $\varepsilon$ are substance-specific parameters.

### 2.2. Treatment of Uncertainty: Quantification, Propagation and Decision Rules

Uncertainty was treated using three quantitative approaches to quantify and propagate uncertainty which were Bayesian probability distributions and probabilistic calculations (BP), probability boxes and probability bounds analysis (PBA), and fuzzy numbers and fuzzy arithmetic (FA). Alternative decision rules are needed under non-probabilistic treatment (Borgelt et al., 2010). Under each treatment of uncertainty we decided to use the following rules for decision making

a) BP: Maximize expected value
b) PBA: Maximize lowest expected value $^1$

c) FN: Maximize lowest value for a specific possibility level.

Under uncertainty the considered social value of the use of a chemical with an unknown level of toxicity is

$$
V_U = DR_{e_1, e_2} [B(e_1) - D(e_1, \tau) + \delta (B(e_2) - D(\rho e_1 + e_2, \tau)) \mid f(\tau), g(\rho)],
$$  \hspace{1cm} (6)

where $f(\tau)$ and $g(\rho)$ are prior information in toxicity and persistence and $DR[-\text{prior}]$ is a general operator to find “optimal” levels of exposure ($e_1$ and $e_2$) for a given decision rule conditioned on the prior information. The chosen levels of exposure during the two regulation periods therefore depend on the way uncertainty is treated, both with respect to its quantification and the choice of decision rules.

### 2.3. Value of Information Analysis

Value of information analysis evaluates the value of testing (i.e. reducing uncertainty) in respect to how it improves regulation and increase social welfare (Gabbert and Weikard, 2010). In its original version the value of information is the expected gain in the social value from using a chemical if optimally regulated under perfect information ($V^P$), instead of regulation under uncertainty ($V^U$), i.e.

$$
VOI = V^P - V^U.
$$  \hspace{1cm} (7)

Here we attempt to extend the value of information to be able to work with other quantifications of uncertainty than probability distributions and expectations. The definition of the value of information depends on whether treatment of uncertainty goes according to a probability distribution, a p-box or a fuzzy number. In its original setting the value of information is an expectation assuming a risk neutral decision maker based upon a probabilistic quantification of uncertainty. For non-probabilistic quantifications we interpret the value of information as the largest gain in social welfare from reducing uncertainty seen over all possible values inside the non-probabilistic interval. This means that testing under a non-probabilistic method would occur

$^1$ Expectations are to be taken over all possible probability distributions within the probability box and not only the outer bounds. The resulting net benefit is an interval.
sooner compared to under a probabilistic treatment, since the smallest indication of an increase in value of information motivates testing. However, the modified VOI for non-probabilistic treatment of uncertainty have a less clear interpretation. The value to consider using a chemical under perfect information, i.e. after it has been tested, is

\[ V^p = DR(V^*(\tau, \rho) | f(\tau), g(\rho)) \]

where \( V^*(\tau, \rho) \) is the value you get from inserting Eq. 4 and 5 into Eq. 3, i.e. when a chemical is regulated at its optimal level of exposure. We performed value of information analysis (which included a decision analysis) on three polybrominated diphenyl ethers (PBDEs) which are chemical compounds of high environmental concern (Table 1). As this was a case-study to illustrate the impact of treatment of uncertainty, therefore, arbitrary values were assigned for some of the parameters. We followed the example in Gabbert and Weikard (2010) and let \( \beta = 1.5 \), \( \varepsilon = 1.2 \) and \( \delta = 1 \). Toxicity \( (\tau) \) had the same prior information either as the non-informative uniform prior or as a point estimate on the mean of the non-informative prior i.e. 0.5. In the first case the value of information express the gain in social welfare from reducing uncertainty in persistence and toxicity, while the latter is the gain from reducing uncertainty in persistence only. The only difference between the three evaluated compounds was therefore related to persistence \( (\rho) \) and its uncertainty.

Table 1. QSPR predictions for chemical properties and estimated half-lives at 25°C for the selected PBDEs.

<table>
<thead>
<tr>
<th>PBDE</th>
<th>log ( S^a ) (mol/L)</th>
<th>log ( V^p ) (Pa)</th>
<th>log ( K_{oc} ) (L/kg)</th>
<th>log ( \tau_{1/2(\text{atm})} ) (hours)</th>
<th>log ( \tau_{1/2(\text{w})} ) (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>congener</td>
<td>PRED SE</td>
<td>PRED SE</td>
<td>PRED SE</td>
<td>PRED SE</td>
<td>PRED PRED</td>
</tr>
<tr>
<td>BDE-47</td>
<td>-1.66 0.25</td>
<td>-3.52 0.16</td>
<td>4.34 0.56</td>
<td>2.55</td>
<td>3.66</td>
</tr>
<tr>
<td>BDE-99</td>
<td>-2.24 0.25</td>
<td>-4.37 0.16</td>
<td>4.58 0.56</td>
<td>2.82</td>
<td>3.93</td>
</tr>
<tr>
<td>BDE-153</td>
<td>-2.77 0.26</td>
<td>-5.22 0.17</td>
<td>4.81 0.56</td>
<td>3.09</td>
<td>4.15</td>
</tr>
</tbody>
</table>

\(^a\)Papa et al. (2009), \(^b\)Gramatica et al.(2007), \(^c\)Roy et al. (2011), \(^d\)Arnot et al.(2005).

\( S = \) water solubility, \( V_p = \) sub-cooled vapor pressure, \( K_{oc} = \) organic carbon-water partition coefficient, \( \tau_{1/2(\text{atm})} = \) degradation half-life in atmosphere, \( \tau_{1/2(\text{w})} = \) degradation half-life in water.

2.4. Uncertainty in Persistence

2.4.1. Environmental Fate Model

Prior information on persistence was derived from residence time given by an environmental fate model. More specifically, was overall residence time, i.e. the average time a chemical remains in a system after it has seized to be released, assessed by a level-II equilibrium criterion model (EQC) multimedia environmental model (Mackay, 2001). This fate model estimates the overall persistence of chemicals assuming the simplification that the concentration of a chemical is at a steady-state and equilibrium condition between the compartments is achieved. Therefore it does not matter into which compartment the chemical is introduced. Weight of evidence on persistence was derived from overall residence time by transforming it into a probability that a chemical particle remains in the system one year after the termination of the first regulation period. Assuming residence time \( (T) \) to be the expected value in an exponential distribution, the probability of a chemical particle remain one year after the period is \( \rho = \exp(-1/T) \). We further simplified the model such that the amount of chemicals to transfer from the first to the second period was \( \rho \epsilon_1 \).
2.4.2. QSPR Predictions

The multimedia fate model requires physical-chemical and environmental properties and individual-medium half-lives as input parameters. The environmental compartments properties (densities and organic carbon fractions) and volumes were taken from an evaluative environment previously reported by Mackay (2001), with a release of 1000 kg/h of the chemical into system (Mackay et al., 1996). For degradation half-life in water, the geometric mean using the linear regressions for BIOWIN, a part of EPISUITE™, outputs derived by Arnot et al. (2005) was considered. The half-life in water was assumed to be two times higher in soils and nine times higher in sediments respectively. The atmospheric half-life was predicted by $0.693/k_{\text{atm}}[\text{OH}]$, where $k_{\text{atm}}$ is the degradation rate constant in atmosphere, and $[\text{OH}]$ is the hydroxyl radical concentrations taken $10^6$ molecules per cm$^3$ (Öberg, 2005).

Uncertainty in environmental fate model was introduced by using information from predictive models, in this case Quantitative Structure-Property Relations (QSPRs), predictions of the chemical-specific input parameters (Table 1) instead of experimentally tested data. Uncertainty in organic carbon-water partition coefficient ($K_{oc}$), water-solubility ($S$), and vapor pressure ($V_p$) was the predictive uncertainty given the corresponding QSPR models (Table 1). With predictive uncertainty we mean an estimate of the predictive error (SE), that in combination with the predicted value (PRED) was used to quantify uncertainty as follows:

a) BP: A probability distribution as a $t$-distribution with the predicted value as its mean (mode) and estimated predictive error as the standard deviation with an appropriate degree of freedom given the size of the data set used to train and test the QSPR model;

b) PBA: A probability box by adding the uncertainty in the estimated predictive error for the $t$-distribution, which was approximately 10% of the predictive error; and
c) FN: A fuzzy number by considering the 2.5$^\text{th}$ and 97.5$^\text{th}$ percentiles of the predictive distribution in BP as the lower and upper bounds (corresponding to possibility level zero) and the point prediction as the best estimate (corresponding to possibility level one). Two levels of possibility were considered: 0.2 and 0.05.

Given a predictive model, in this case a linear regression, a probabilistic quantification follows given QSPR data of good enough quality and a model with verified statistical properties. The fuzzy approach is not meant to be derived from a probabilistic approach, but was included here for comparison. Since one of the aims was to evaluate the value of reducing uncertainty in QSPR predictions, and since predictive uncertainties for half-lives in different media for these specific compounds were missing, the uncertainties in half-lives were not considered here. RAMAS Risk
Calc version 4.0 (Applied Biomathematics Inc., Setauket, New York) was used for the propagation of uncertainty expressed as probability distributions, by discretizing the probability distributions, as fuzzy number, by fuzzy arithmetic, and of probability boxes, by a combinations of these. Other calculations were done with R (R Development Core Team, 2008).

3. RESULTS

3.1. Environmental Fate Model Predictions under Alternative Treatments of Uncertainty
Persistence of PBDEs is supposed to increase with the number of bromine atoms, which makes it easy to memorize the order of the compounds, BDE-47 lowest, BDE-99 intermediate and BDE-153 highest (Figure 1). The Bayesian cumulative probability distribution was bounded by the p-box from probability bounds analysis. A probabilistic treatment models the density under the probability distributions, and do not put any bounds on uncertainty, whereas fuzzy number and fuzzy arithmetic models the bound for uncertainty, which have larger width than the probability distribution.

Figure 2. Regulatory strategy for the use of a chemical during two time periods derived under alternative treatments of uncertainty (BP: Bayesian Probabilistic, PBA: Probability Bound Analysis, FN: Fuzzy Numbers), where toxicity is a) a point estimate and b) a non-informative prior.

3.2. Effects of Treatment of Uncertainty on Regulated Exposure Levels
It follows from the decision model the level of exposure always will be higher for the second period since persistence is not considered beyond this period (Figure 2). Since exposure level in the second period only depends on toxicity, the situation when toxicity is a point estimate will result in equal levels of exposure for alternative treatments of uncertainty during the second period (Figure 2 a). A non-probabilistic quantification by fuzzy number resulted in more conservative regulatory exposure levels compared to a probabilistic quantification (Figure 2 and 3). This was expected since the probabilistic decision rule does not include aversion to risk, whereas it is one of the bases for the non-probabilistic treatment. Adding a non-probabilistic treatment to a probabilistic treatment (i.e. from BP to PBA) resulted in more conservative exposure levels, but no as large as the purely non-probabilistic approach (Figure 2 and 3). The pattern was similar for the three evaluated compounds. The compound with the lowest persistence
(BDE 47) had the largest regulated total level of exposure compared to the other compounds. A change in possibility level from 0.2 to 0.05 in the fuzzy approach resulted in an equal or lower level of exposure for BDE 47 but not for the other compounds. Going from none to large uncertainty in toxicity (compare a to b in Figure 2 and 3) the difference between probabilistic and non-probabilistic treatments of uncertainty increased. This follows since the probabilistic method is risk neutral only considering what is expected, while the wider bounds on uncertainty in toxicity are considered in the non-probabilistic method.

Figure 3. Regulated amount of chemical given as the optimal level of exposure during two regulation periods considering social welfare and uncertainty in predictions, derived under different treatments of uncertainty, and where toxicity is a) a point estimate $\tau = 0.5$ and b) a non-informative prior.

Figure 4. Value of information over the three chemicals evaluated under alternative treatments of uncertainty, and where toxicity is a) a point estimate $\tau = 0.5$ and b) a non-informative prior. Probability Bound Analysis result in an interval on the value of information.
3.3. Effects of Treatment of Uncertainty on the Order of Chemical Testing
The gain in social welfare from performing additional testing on persistence only, assuming toxicity to be precisely known was relatively small (compare BP in Figure 4 a and b). This is because persistence had smaller influence in the decision model compared to toxicity. Naturally, using a non-informative prior on toxicity gives a higher value of information since more uncertainty is included in the model. The ordering of chemicals for testing according to their value of information changed with the treatment of uncertainty (Figure 4). The non-probabilistic approaches generated the same order chemical testing as the probabilistic method (BP) when only the uncertainty in persistence was considered (Figure 4 a).

4. CONCLUSION

The basic question to be answered was if matters for decision making how we choose to treat uncertainty. Using an example from chemical regulation, where uncertainty stemming from the use of predictive models instead of experimentally tested data, were given alternative treatments, we have shown that

- Treatment of uncertainty affected the order of which chemicals should be experimentally tested to reduce the uncertainty in the background knowledge for chemical regulation.
- Non-probabilistic treatment of uncertainty resulted in more conservative (read safe) chemical regulation.
- Differences between alternative treatments depended on the extent of uncertainty in the background knowledge.

These findings are restricted to how we let decision makers respond to uncertainty in model predictions by the chosen set of decision rules. For example, risk aversion in the probabilistic approach would lead to more conservative chemical regulation.

The framework of a predictive model – risk model – decision model – value of information analysis provide a computational platform for further evaluation of the effect of treatment of uncertainty on other decision problems or on a higher level of complexity. Increasing realism in individual models may give a better comparison between alternative treatments of uncertainty. Which method to use is context specific and depend on the strength of the background knowledge. If uncertainty comes with likelihoods then probabilistic approaches should be used. It therefore of less interest to compare alternative treatments of uncertainty, since which method to use is a decision taken before further knowledge is achieved, and not after.

Acknowledgements
This study was funded by the European Seventh Framework Programme (FP7) through the project CADASTER (grant agreement 212668). We are grateful to Tomas Öberg for his contribution.

References


Acknowledgements
This study was funded by the European Seventh Framework Programme (FP7) through the project CADASTER (grant agreement 212668). We are grateful to Tomas Öberg for his contribution.

References


The value of information analysis has been modified under a non-probabilistic treatment of uncertainty and therefore it is only the relative order of chemicals that can be studied.


