Chemical Hazard Assessment under Uncertainty

Evaluated by Robust Bayesian Evidence Synthesis

Yudiana Yang

Bachelor's thesis
2017:K13

Lund University

Faculty of Science
Centre for Mathematical Sciences
Mathematical Statistics

Centrum Scientiarum Mathematicarum

Abstract

Chemical hazard assessment is a procedure in chemical regulation to set which concentrations of chemical substances that are safe to an ecological system, i.e. where a majority of the species in the systems is not negatively affected by the substance. The Species Sensitivity Distribution (SSD) is an assessment model to estimate hazardous concentration based on toxicity data from representative species from an ecological system. In its current use, the hazard assessment using SSD approach is only to be applied based on enough large sample size of toxicity data, which in absence of large sample sizes leads to overly conservative regulation, and the choice of hazardous concentration is based on statistical properties of the SSD, without any explicit loss function.

In this thesis, the hazard assessment using SSD was implemented as a Bayesian evidence synthesis integrating different sources of evidence to calibrate the assessment model and propagate uncertainty to a decision model. The hazardous concentration was then the Bayes optimal decision that minimizes the undesired effects caused by the chemical substance. Loss was determined by a linear-exponential loss function, which captures conservatism in an effective way and is useful to find thresholds. The aim of this thesis was to evaluate the impact of uncertainty on the hazardous concentration due to small sample size, measurement errors in toxicity data and, since uncertainty was quantified by Bayesian probabilities, the choice of priors for the SSD as well.

The impact of uncertainty on the hazard concentration was evaluated by a simulation study. It was shown that choice of prior on the mean and spread of the SSD matters when sample size was less than 30 toxicity data. Quality in data, manifested as estimation errors of toxicity values, had a profound influence on choosing hazardous concentration. When the data was sufficiently large, choice of the priors and the quality of the data almost had no impact on choosing the optimal hazardous concentration. It was suggested that a higher value on a certain parameter for the loss function might compensate for the small data or data with poor quality. Several repetitions of simulation are needed to make result more reliable.

Populärvetenskaplig sammanfattning

Kemisk farobedömning används för att sätta gränsvärden på kemikalier. Dessa beslut utmanas av brist på data på hur giftigt ett ämne är för olika arter i ett system. Istället används, till en viss del, godtyckliga osäkerhetsfaktorer för att vara på den säkra sidan.

Em artkänslighetsfördelning är en bedömningsmodell för att beräkna den halt där ämnet har en godtagbar låg effekt. Detta gjordes genom att 1) implementera farobedömningen i en så kallad Bayesiansk evidenssyntes som kan väga ihop olika datakällor, 2) beskriva osäkerhet med hjälp av Bayesianska sannolikheter och ta den vidare till en beslutsanalys som 3) använder en linjär-exponentiell funktion för att beskriva förlust av att välja fel gränsvärde.

Detta sätt att föra farobedömningen undersökts med hjälp av en simuleringsstudie med artificiella data. Den visar att valet av gränsvärde (här det Bayeianska optimala beslutet) är känsligt för situationer med ont om data, och att det Bayesianska tillvägagångssättet kompenserar för detta genom att tilldela större osäkerhet till modellen. Studien visar på vilket sätt farobedömning påverkas av hur man beskriver artkänslighetsfördelningen innan man har sett data.

Detta förslag på att hantera osäkerhet vid farobedömning baserat på artkänslighetsfördelning är ett bidrag till nya metoder för osäkerhetsanalys. Här kan den användas för att sätta gränssvärden baserad på tillgängliga data och därmed undvika alltför sträng reglering av kemiska ämnen.

Popular summary

Chemical hazard assessment are made to select safety thresholds on chemicals. These decisions are challenged by lack of toxicity data for the different species in a system. Therefore, to some extent arbitrary, uncertainty factors are applied to be on the safe side.

The Species Sensitivity Distribution (SSD) is an assessment model to estimate the so called hazardous concentration, taking into account differences in toxicity data from representative species from an ecological system. This thesis study the impact of uncertainty due small sample size and varying quality in data on hazard assessment using the SSD approach. This was done by 1) implementing the hazard assessment as a Bayesian evidence synthesis integrating different sources of toxicity data to calibrate the SSD, 2) use Bayesian probability to quantify and propagate uncertainty to a decision analysis, where 3) the linear-exponential loss function was used to explicitly express loss from choosing the wrong threshold.

This assessment is evaluated in a simulation study using artificially generated toxicity data. It shows that the choice of safety threshold (here the Bayes optimal hazardous concentration) is sensitive to small data, and that the Bayesian approach account for this as wider uncertainty bounds. The study shows in what way hazard assessment applied on small data samples is sensitive to the prior belief of the SSD.

The suggested approach for hazard assessment using SSD is a contribution to novel methods for uncertainty analysis. It can help choosing safety thresholds based on available toxicity data, yet avoiding too conservative chemical regulation.

Acknowledgements

Thanks to Ullrika for your patience and encouragement.

Experiment!
Make it your motto day and night.
Experiment,
And it will lead you to the light.
The apple on the top of the tree
Is never too high to achieve,
So take an example from Eve,
Experiment!
Be curious,
Though interfering friends may frown.
Get furious
At each attempt to hold you down.
If this advice you always employ
The future can offer you infinite joy
And merriment,
Experiment
And you'll see!

—Cole Porter

What is true is already so.
Owning up to it doesn't make it worse.
Not being open about it doesn't make it go away.
And because it's true, it is what is there to be interacted with.
Anything untrue isn't there to be lived.
People can stand what is true,
for they are already enduring it.

—Eugene Gendlin

Contents

Glossary

BES Bayesian Evidence Synthesis.

DAG Directed Acyclic Graph.

HA Hazard Assessment.

HA-SSD Hazard assessment using Species Sensitivity Distribution.

HC Hazardous Concentration.

HPD Highest Posterior Density.

HPDI Highest Posterior Density Interval.

LHC Logarithimized Hazardous Concentration.

LINEX LINear-EXponential loss function.

NOEC No Observed Effect Concentration.

SSD Species Sensitivity Distribution.

1 Introduction

1.1 Background

Hazard Assessment (HA) is used in Chemical Safety Assessment to screen or classify chemical substances and set safety thresholds [1]. Whereas a risk assessment consider the combination of an exposure and an effect, while HA only considers the effect. The effect on a ecological system are evaluated by how toxic a chemical substance is for the different species in the system. Hazardous Concentration (HC) is the concentration at which an undesired effect emerges for the species in a ecological system, hence one does not want to have in the environment. The HC or Logarithimized HC (LHC) can be evaluated with respect to several species in the system forming a distribution.

The Species Sensitivity Distribution (SSD) is often used to model variability ecotoxicological data [2]. The SSD represents the distribution of different toxicity values for multiple species, as different species have varied sensitivity to a compound at a given concentration. Usually No Observed Effect Concentrations (NOEC) are used as the toxicity values.

A SSD is usually assumed to be a normal (Figure 1) or logistic distribution. The parameters of a SSD is estimated from a sample of toxicity data from different species representing the ecological system that is influenced by the chemical.

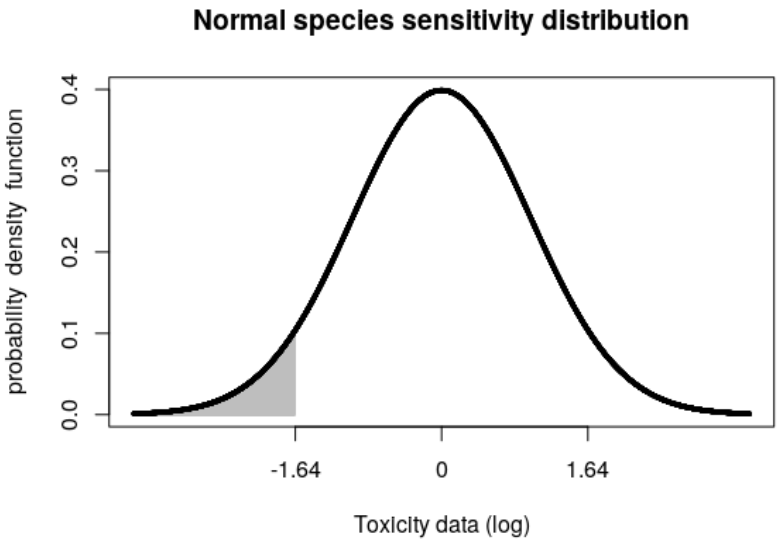

Figure 1: Standard normal species sensitivity distribution over the toxicity data on log-scale with LHC_p = -1.6449, i.e fifth percentile.

Using SSD to model variability reduces the hazard assessment to statistically assess the the p^{th} percentile of the SSD. For example, HC_5 is the highest concentration which would have a negative effect on 5 percent of the species in the environment.

Chemical HA using SSD face several problems such as:

1. Sample size

   In order to accurately determine the HC, a toxicity value for each species is required. The chemical regulations say that at least 10, preferably 20, data points are needed to use SSD. In many situations, researchers only have access to experimental toxicity data for up to few species. If so, the SSD approach is not suitable. Instead, the procedure is to take the lowest toxicity value and divide the it by 10 (so called safety factor) and use the result for setting a threshold. Another option is to use other types of information, e.g. *in silico* data (computer-aided derived data from models), or extrapolate from data of the same species but measured on tests.

2. Measurement errors

   The current form of HA using SSD doesn't take uncertainty in the toxicity data, e.g. measurement errors, extrapolation errors, or in silico prediction errors into account, nor uncertainty in the fit of the SSD. As a consequence, the impact of this uncertainty on the resulting HC threshold is unknown. The estimated HC could be higher or lower than the true HC. If the estimated HC is higher than the true HC, it means we underestimated the risk and more species are harmed than what is allowed. If the estimator is lower than the true HC, then the risk is overestimated which leads to overprotection.

3. Risk aversion

   From a decision maker's point of view, ignoring uncertainty creates an unknown risk aversion [3]. For example regulating chemical substances with unrealistically low HC can be problematic since a lot of resources are allocated to come up with a substitute, which may be as dangerous [4].

A SSD analysis can be considered as a meta-analysis since it uses toxicity estimates from multiple studies and integrates them to infer common parameters in the SSD, such as the mean and the standard deviation (μ and σ). The difference between a meta-analysis and other types of statistical analysis is that in a meta-analysis, differences in the quality of estimates (e.g. estimated standard errors) from independent sources of data are considered in the analysis.

Meta-analysis can be done with frequentist as well as Bayesian statistical principles. The Bayesian approach allows to integrate expert knowledge and data and quantify uncer-

tainty by probability. Expert knowledge can be used to construct the prior distribution [7].

Uncertainty quantified by Bayesian probability can be propagated through an assessment model and quantify uncertainty in the relevant output. This can in turn inform a decision model at hand. A Bayesian Evidence Synthesis (BES) is usually referring to a model/analysis where an assessment model is linked with a decision analysis, allowing the model to be calibrated by multiple sources of data and expert knowledge in a Bayesian framework. A BES is able to integrate different sources of evidence, e.g directly and indirectly relevant data or information, and propagate them through an assessment model to the assessment outputs [7, 8, 9].

The optimal decision is the Bayes optimal one, i.e. that minimizes expected loss. The expectation is taken over uncertainty quantified in the assessment output. In this case, it is the hazardous concentration that minimize the undesired effects caused by the chemical compound. A BES requires a loss function over failing to select the concentrations for which 5 percent of the species are protected to choose between. The need for a loss function has for example been identified by Hickey et al (2009) [5]. They suggests the linear-exponential (LINEX) loss function for HA using SSD. In this thesis, the HC is then the concentration which minimizes the expected loss using the LINEX loss function.

1.2 Hazard Assessment as a Decision Problem under Uncertainty

The goal of the hazard assessment is to set a HC threshold such that concentration of a compound is hazardous for not more than 5 percent of species in the system, HC_5.

A common procedure in meta-analysis is to consider error associated the estimates from different studies, using weighted pooled estimates of the main effect (e.g. forest plots) or weighted regression (if the analysis includes a common covariate). A similar approach can be applied on the HA by considering error in estimates of the toxicities. Today SSD do not usually consider estimation errors (for example, see [13]). Consideration of error in toxicities estimates will lead to stronger assessment, assuming the errors are known with good precision.

It is not obvious what is a good HC. It is certain that the choice may lead to over-protection or missing to protect species in the system. When the former is to prefer a non-symmetric loss function would be suitable. Hickey et al [5] used the so called LINEX loss function on a hazard assessment, where the decision is the optimal Bayes decision.

Uncertainty in loss comes from the difference between HC based on the actual SSD and the chosen HC. The actual SSD is defined by the true values of the parameters μ and σ, thus computing uncertainty in loss is done by quantifying uncertainty on the estimates of μ and σ. In this way, uncertainty is propagated to the decision alternatives, i.e different values on HC, and into the LINEX.

To sum it up, in this study the hazard assessment using the SSD (HA-SSD) is framed as a Bayesian Evidence Synthesis. A Bayesian approach to estimate the SSD and choose HC has been applied by Hickey et al, but they did not take into account uncertainty in toxicity data.

In this thesis, we turn the HA-SSD into a BES where uncertainty in toxicity data is quantified and use the LINEX loss function to set the Bayes optimal HC.

1.3 Goal of the Thesis

The aim of the thesis is to study how the hazardous concentration is influenced by:

- uncertainty due to small sample size
- uncertainty due to measurement errors in toxicity data, and
- uncertainty due to choice of prior for parameters in the species sensitivity distribution.

The impact of uncertainty on the chosen HC in a regulatory context will be studied by treating a HA-SSD as a BES, in which the HC is the Bayes optimal decision using the LINEX loss function.

In order to consider uncertainty factors mentioned above, we will run several simulations to create different artificial toxicity data using a design experiment based on uncertainty factors.

This thesis intends to answer the following questions:

1. What effect does sample size have on choosing a HC_p?
2. What effect does the quality of the toxicity data have on choosing a HC_p?
3. What effect does the prior, for example based on expert knowledge, have on choosing a HC_p?

The structure of this thesis is as follows. Section 1 is an introduction presenting an overview of the ideas underlying the thesis, where background and the problem are presented. Section 2 brings up the theory with further details of the ideas. Section 3 explains the methods used in this thesis. Section 4 presents the result. Discussion is then followed up in Section 5. Finally, conclusions are made in Section 6.

2 Theory

2.1 Definition of the Hazardous Concentration

Hazardous concentration, HC_p, is the concentration of a compound that is hazardous for some p percent of a set of species and is a measure used in chemical risk assessment. The p percent of species in the left tail of a probability distribution, could be set to an acceptable level as a way to assure that the compound would not harm the species more than desired. The probability distribution we will use in this thesis is the normal SSD.

Since the concentration of toxic compound is expressed on log-scale, we will use the logarithm of HC_p to express the hazard concentration threshold, LHC_p. The true LHC_p is estimated from the true values of the parameters, μ and σ. When the parameters are uncertain, LHC_p will be estimated from the posterior samples.

2.2 The Directed Acyclic Graph for the SSD

Our model consists of normal species sensitivity distribution (SSD) and prior distributions on the parameters for the SSD, μ and σ. The parameter μ has a normal distribution with hyperparameter μ_μ and the standard deviation σ_μ. The parameter σ will have different prior distributions that will be presented later.

Let t_k be the toxicity concentration on log-scale for each species and the total number of the species is K so k = 1, 2, ..., K. Let y_i be the toxicity data with $i = 1, 2, 3, ...$ The data y_i has the normal distribution with mean t_k and the standard deviation σ_{y_i} which represents the observational errors or data quality.

So our model, together with the prior distribution on σ and the hyperprior μ_μ, is a Bayesian hierarchical model and it is expressed best using directed acyclic graphs, DAGs (Figure 2) [8]. Quantities at the start of an arrow in the graph give rise to the quantities at the end of the arrow. They are called parents respective children. The quantities are also called nodes and in this study we are going to use 5 different kinds of nodes:

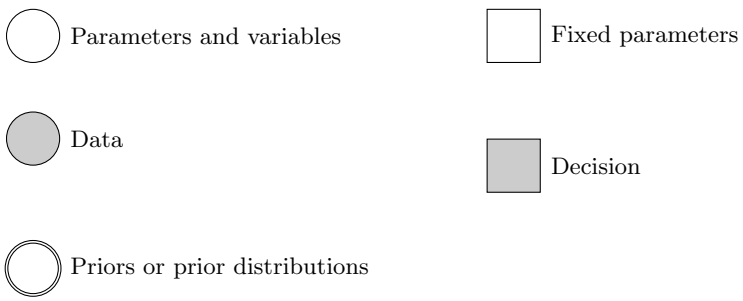

Figure 2 illustrates the DAG for the Bayesian hierarchical model.

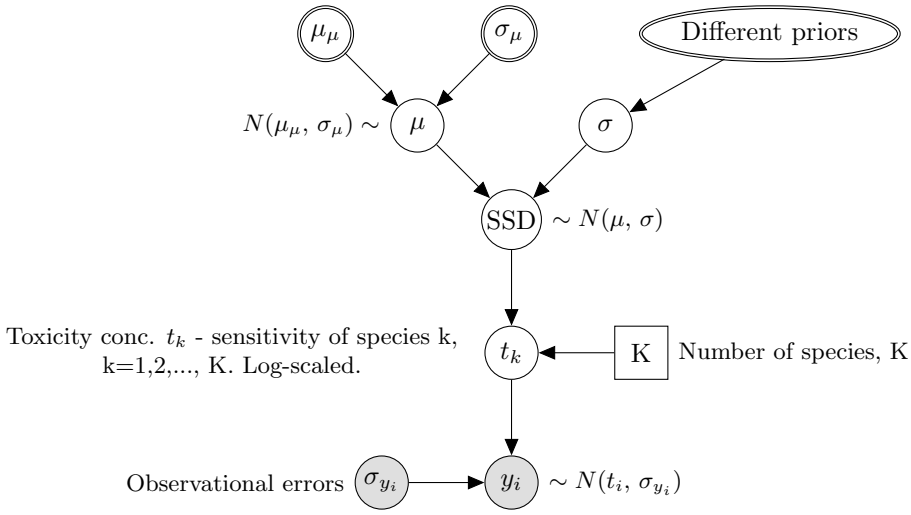

Figure 2: Directed acyclic graph for the Bayesian hierarchic model based on SSD distribution and the prior on σ. The endpoints are the toxicity data, y_i for $i = 1, 2, 3, ...$

For each species, there can be more than one estimates of toxicity available, e.g from different experiments and laboratories. For example, for species number 1 with expected toxicity of t_1 there can be several available toxicity data, say, y_1 and y_2, which are assumed to be conditionally independent on t_1. This study used one data point for each species, i.e $y_i = y_k$ (Figure 3).

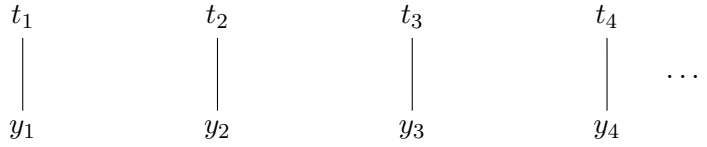

Figure 3: The relation between t_k and $y_i = y_k$ with k = 1, 2, 3, ..., K and i = 1, 2, 3, ... There are one toxicity endpoint per species. The endpoints follow a normal distribution with mean t_k.

2.3 Optimal Bayes Hazardous Concentration for a Given p and Loss

The notation used by Hickey et al will be followed from now on. Define a group of parameters $\boldsymbol{\theta} = (\mu, \sigma^2)$, and define $\mathbf{X}$ a vector containing the log-toxicity data for all species, y_1, y_2, ... Let $\bar{x}$ be the mean of the sample and s be the standard deviation of the sample. Then we can make the identities $\text{LHC}_p \equiv \psi_p(\boldsymbol{\theta}) = \mu - K_p\sigma$ and $\hat{\text{LHC}}_p \equiv \delta_p(\mathbf{X}) = \bar{x} - k_p s$, where K_p is the (1-p)th percentile of the normal distribution.

Let $\delta_p(\mathbf{X})$ be the decision rule and L a loss function. Then the Bayes optimal decision is:

$$\delta_p(\mathbf{X})^* = \underset{\delta(\mathbf{X})}{\arg\min}\, \mathbb{E}^{\boldsymbol{\theta}|\mathbf{X}} L(\delta(\mathbf{X}), \psi_p(\boldsymbol{\theta})), \tag{1}$$

where the expectation is taken with respect to the posterior sample of the parameters and we consider all the decision rules when minimizing. Correspondingly, k_p^* is the optimal decision for k_p, but we are more interested in $\delta_p(\mathbf{X})^*$ here. Using Bayes optimal decision, we can minimize the loss by setting $\delta_p(\mathbf{X})^*$ as our optimal decision.

The percentile, p, is usually to be set on 5 percent, but it is rather arbitrary and it could be set on lower or higher level. However, since 5 percent is the default, LHC_5 or HC_5 will be used in this study.

2.4 The Linear-Exponential (LINEX) Loss Function

A loss function is a function that measures so called cost associated with either over- or underestimating the true LHC_p. The linear-exponential, LINEX, loss function has a nice feature of increasing linearly loss on one side and exponentially on the other side:

$$L(LHC_p, \hat{LHC_p}) = \beta \left[\exp\left\{ \alpha \frac{\delta - \psi(\boldsymbol{\theta})}{\sqrt{\boldsymbol{\theta}^T \cdot \mathbb{I}}} \right\} - \alpha \frac{\delta - \psi(\boldsymbol{\theta})}{\sqrt{\boldsymbol{\theta}^T \cdot \mathbb{I}}} - 1 \right], \tag{2}$$

where δ is a decision of LHC_p. For simplicity, we will use $\beta = 1$ in this study. Since overestimating the true value of LHC_p, that is, harming or losing species more than we are ready for, we would like to "punish" the overestimation and a less severe punishment for underestimating the true LHC_p. LINEX loss function fulfills this purpose and using LINEX loss function enables conservatism in our hazard assessment in an effective way. Figure 4 describes how the loss is calculated from the parameters via LHC_p for given p, given loss parameter α and a given decision δ, i.e a chosen LHC_p.

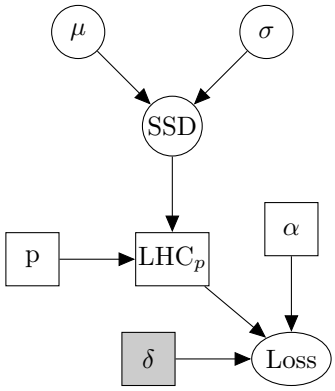

Figure 4: The DAG illustrates how a loss is calculated from μ and σ with p, α and δ being given.

For values on α, we will use three different values for comparison: 1, 2, and 2.5. Figure 5 shows the plots of LINEX with these different values. The higher value on α is, the steeper the LINEX loss function becomes, especially for overestimation of LHC$_p$.

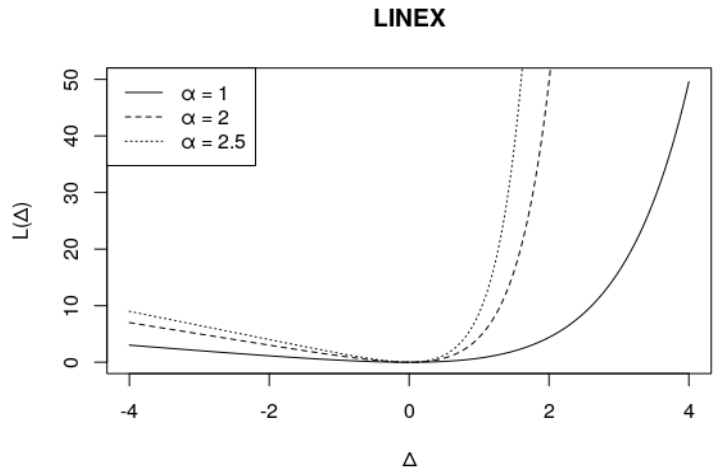

Figure 5: Loss according to the LINEX function L(Δ) where the difference between the chosen and the true threshold $\Delta = \delta_p(\mathbf{X}) - \psi_p(\boldsymbol{\theta})$ is re-weighted by the loss parameters $\beta = 1$ and $\alpha = 1$, 2 respectively 2.5.

Using Bayes optimal decision on LINEX and taking uncertainty from samples into account, we get the true value and a estimate of the LHC_p and the corresponding optimal decisions. Table 1 provides an overview over the true value of LHC_p, the estimate based on uncertain μ and σ and the optimal decisions.

	μ, σ known	μ, σ uncertain
Percentile in SSD	$\texttt{LHC}_p$	$\texttt{L}\hat{\texttt{H}}\texttt{C}_p$
Bayes optimal decision, δ	$\texttt{L}\hat{\texttt{H}}\texttt{C}_p^\alpha$	$\texttt{BL}\hat{\texttt{H}}\texttt{C}_p^\alpha$

Table 1: The different hazardous concentrations encountered in the simulation study are the true and estimated percentiles of the SSD, LHC_p, and the corresponding Bayes optimal decisions for a given loss parameter α, LHC_p^α

.

To see how well the Bayes optimal decision $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ is, we introduce a bias given by the difference between the Bayes optimal decisions, $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ based on the posterior sample and $\text{L}\hat{\text{H}}\text{C}_p^\alpha$ based on the known SSD parameters μ and σ, i.e. under perfect information:

$$\text{Bias} = \text{BL}\hat{\text{H}}\text{C}_p^\alpha - \text{L}\hat{\text{H}}\text{C}_p^\alpha \tag{3}$$

2.5 Bayesian evidence synthesis (BES)

We have a decision issue that involves choosing a concentration as threshold. To solve this we use Bayesian decision-making by minimizing the expected loss using the Bayes optimal decision. Putting it slightly different from Eq. (1) and using Bayes theorem, we have that:

$$\delta(\mathbf{X})^* = \arg\min_{\delta}(\mathbf{E}(L(\Delta)|\mathbf{X}) \tag{4}$$

where $\mathbf{E}(L(\Delta)|\mathbf{X}) = \mathbf{E}(L(\Delta)|\boldsymbol{\theta})P(\boldsymbol{\theta}|\mathbf{X})$. The posterior distribution $P(\boldsymbol{\theta}|\mathbf{X})$ is obtained from backwards Markov Chain Monte Carlo (MCMC) simulation, i.e the evidence from the data is propagated backwards to pass uncertainty to the parameters. When there is uncertainty in the parameters, the loss function $L(\Delta)|\boldsymbol{\theta}$ is calculated through forward MCMC simulation.

BES unifies the two calculations into a single process where the resulting posterior distributions from backward MCMC simulation are directly fed into the LINEX loss function and the decision analysis associated with the loss function, i.e Bayes optimal

decision [9]. The optimal Bayes decision, δ is then set as the LHC threshold. Figure 6 illustrates the BES.

Using the theory just presented, the integrated Bayesian approach on hazard assessment using the SSD and Bayes optimal decision with LINEX loss function will be applied in this thesis.

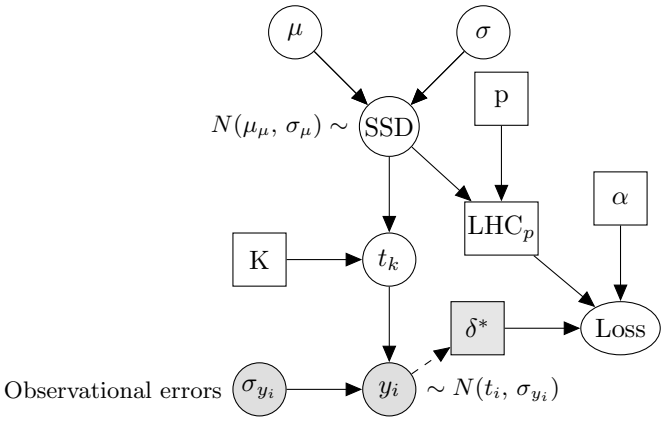

Figure 6: Bayesian hierarchic model, without details, shows how BES is structured. BES consists of SSD and Bayes optimal decision with LINEX.

3 Methods

3.1 Design of Simulation Experiment

Since there are different factors to consider and vary, we will use a factorial design for this study. For simulating toxicity data, we use 2 factors with 7 levels in total, namely the number of species and the quality of data:

- Number of species (data points): 5, 7, 10, 13 respective 20
- Data quality, σ_{y_i}: good respective poor

The standard deviation σ_{y_i} is chosen in relation to the standard deviation σ, which is chosen to be 1, and we assign uniform distributions to σ_{y_i} with two different sets of (hyper)parameters: U(0, 0.2) and U(0.2, 1). The former represents good quality on the data and the latter represents the poor one. Observe that the latter has a wider range and the lower limit is a bit higher compared with the former.

The other factors are the priors that will be used while fitting parameters μ and σ to the data with the help of Stan models. We choose normal prior on μ and we also want to vary the hyperparameters, μ_μ. σ_μ is simply set to 2. As for prior on the spread of the SSD, we choose three prior distributions for this study: uniform on the standard deviation and gamma and inverse-gamma distribution on the variance. Our choices of priors are as following:

- Prior on the spread: gamma (σ^2), inverse-gamma (σ^2) respective uniform distribution (σ)
- Prior on μ_μ: $\mu = 0$, $\mu + 4\sigma = 4$ respective $\mu - 4\sigma = -4$, with $\mu = 0$ and σ =1 as given earlier

The parameters for each of the prior distribution for the standard deviation will later be detailed as it deserves its own section. These priors, weak or informative, could be thought as a form of expert opinion and uncertainty.

So with the four factors and their levels, we get $5 \times 2 \times 3 \times 3 = 90$ combinations.

3.2 Simulating Artificial Toxicity Data

As mentioned earlier, for the purpose of the experiment, we need to simulate different sets of data. With the two factors given for simulating data in the previous section, we

get ten different data (five different numbers of species or data sizes times two kinds of quality). The ten different data form a dataset. Table 2 shows what a simulated artificial toxicity data looks like. Figure 3 together with Table 2 might help the reader to visualize a data.

The size of the data is rather random as between one to three toxicity values are assigned to each species, but it is obvious that as number of species grows, the more data points we get. Figure 7 illustrates an examples with three different data with increasing number of species.

Index, i	Species, k	Toxicity (log), y_i	Observational error, σ_{y_i}
1	1	y_1	σ_{y_1}
2	2	y_2	σ_{y_2}
3	3	y_3	σ_{y_3}
4	4	y_4	σ_{y_4}
5	5	y_5	σ_{y_5}
6	6	y_6	σ_{y_6}
7	7	y_7	σ_{y_7}
⋮	⋮	⋮	⋮

Table 2: An example on what a simulated toxicity data might look like. Observe that i = k since we simulate one data point per species.

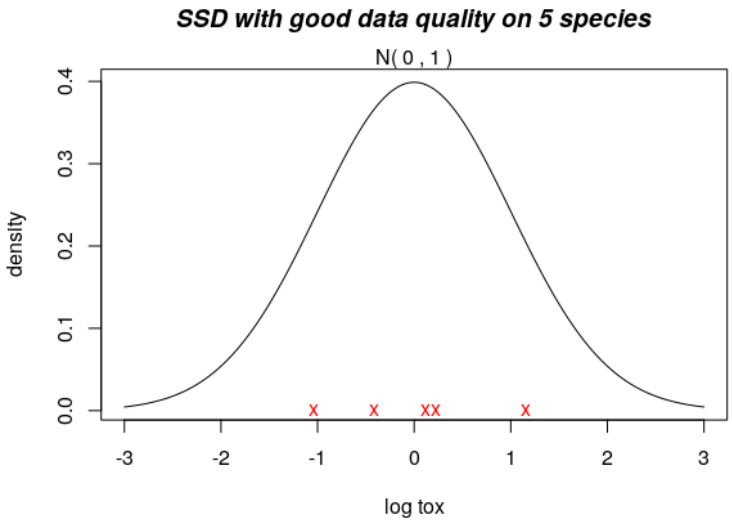

Figure 7: A sample of five toxicity values of good quality (red X marks) from a SSD with $\mu = 0$ and $\sigma = 1$.

3.3 Priors on the Standard Deviation, σ

When choosing prior distribution on the spread, there are several distributions to choose from. Wishart and inverse-Wishart distributions are used when the scale parameter (σ) is a covariance matrix. Due to convenience, we will not choose those. Gelman [10] recommended the non-informative uniform distribution on standard deviation as the starting point, which we will adopt in our study. We will also use gamma and inverse-gamma prior distribution on variance, although half t-distribution or half-Cauchy distribution might be better suggestions.

Gamma distribution is a rather familiar distribution and, just like inverse-gamma distribution, is one of the conjugate priors for normal likelihood. Due to the familiarity and the property of conjugate prior, those prior distributions are chosen.

The parameters for the priors were chosen to have a similar mode and range (Figure 8). Accordingly, hyperparameters were adjusted as follows: gamma prior distribution with shape = 3 and rate = 3 and inverse-gamma prior distribution with shape = 2 and scale = 2 are chosen. As for uniform prior distribution, the limits 0 and 2 are chosen. In short, the priors are:

- Uniform prior distribution on σ, U(0, 2)
- Gamma prior distribution on σ^2, G(3, 3)
- Inverse-gamma prior distribution on σ^2, IG(2, 2)

The gamma distribution allows the variance to be very close to zero, which has been shown to create problems with convergence [10]. To overcome this problem, it is possible to choose the inverse-gamma distribution which has a region near zero with very low density. In addition, the inverse-gamma distribution has a slightly thicker and longer tail than the one of the gamma distribution, which can be good or bad depending on the intention of the analysis.

Figure 8: 3 priors for the spread of SSD: U(0, 2), Gamma(3, 3) and Inv-Gamma(2, 2) in R.

3.4 Creating Models and Simulating Posterior Samples Using R and Stan

The HA-SSD is implemented as a Bayesian model in the program Stan which use Hamiltonian Monte Carlo sampling with the extension of a No U-Turn Sampler (NUTS). The simulation experiment and sampling is run in R using the package `rstan` as interface with Stan.

The posterior samples of the parameters μ, σ and LHC_p are stored from each data and choice of priors in the simulation experiment. There are three different prior distributions on σ and three different hyperparameters, μ_μ being 0, 4 respective -4, hence nine models are applied on each data. A sensitivity analysis is then performed on each combination of different data and the nine models. This type of analysis is also called robust Bayesian analysis.

The process to generate artificial data and sampling from posterior based on this data is repeated ten times in order to check how the result is affected by the variation in the different random-generated artificial dataset. Each iteration takes about ten minutes to complete on an ordinary laptop (Dell Latitude E7450).

3.5 Performance Measures

Besides comparing estimators of μ, σ and LHC_p to the corresponding true values, we will look at Bayes optimal decision $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias (the difference between the two Bayes optimal decisions, under perfect information respective under uncertainty). Bias

tells us how much well the application of the model would be. The smaller the absolute value of the bias is, the better.

To express uncertainty for estimators, a credible interval will be computed, specifically the 95% highest posterior density intervals, HPDI [11]. Tables with HPDI information tell us how well the model compared to the reality, that is, our simulated data.

The common diagnostics for checking convergence in MCMC are, among others, plotting traces of MCMC chains, plotting of variance and the autocorrelation. However, for convenience, we will mainly use a different way to check convergence, known as Gelman & Rubin's diagnostic, (split) potential scale reduction statistic, $\hat{R}$ [11]. If all chains converge to a equilibrium distribution, then split $\hat{R}$ will be one, otherwise greater than one. The R function `pairs()` will also be used, where red dots in plots indicate divergence transitions.

Another performance measure is n_{eff} which is a measure of effective sample size, ESS [11]. The ESS is the number of independent samples that has the same estimation power as the autocorrelated samples do. Since we use four chains, each with 10000 iterations which in turn consist of 5000 warmup draws and 5000 post-warmup draws, the total post-warmup draws is 20000. The closer to 20000 n_{eff} is, the better.

The process to generate data and sample from models with different priors is repeated ten times. After each iteration, averages of the lower and upper HPD limits for μ, σ respective LHC_p are computed along with the averages of the posterior means for μ, σ respective LHC_p. In an attempt to express uncertainty for the averages, the lowest and highest values of the posterior means, the lower HPD limits respective the five upper HPD limits are computed for μ, σ, LHC_p respectively.

For each model the following plots are stored: 1) plots with means of posterior samples with their HPDI vs the true values and 2) cumulative (posterior) distribution vs the true cumulative distribution (which is known from the true values of μ and σ).

4 Results

4.1 Convergence Properties of the Bayesian SSD

Divergent transitions is an indication of convergence problems and is here identified as red dots in the blue cloud representing the posterior sample of μ, σ and LHC_p (using `pairs()` in R) (Figure 9). Sampling based on data with 20 species has no divergent transitions in all iteration and model versions.

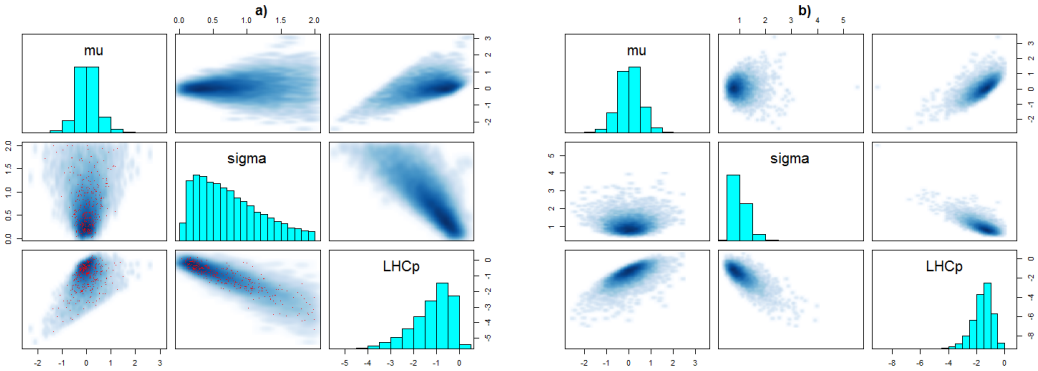

Figure 9: Posterior samples based on a) uniform prior and b) inv-gamma on σ and $\mu_\mu = 0$. The red dots indicate divergent transitions. Note the differences in the axes between a and b.

Divergences in the MCMC sampling are most common for the uniform prior on the spread of the SSD results in a lot of divergences in the MCMC chains (Table 3). The gamma prior in combination with poor data quality has poor convergence as well, whereas almost no divergent transitions are found for models with the inverse-gamma prior distribution on σ^2.

The $\hat{R}$ is 1 for the SSD parameters and the percentile LHC_p. The $\hat{R}$ statistic is not affected by poor quality and sample sizes. This result is found for both models with the gamma and the inverse-gamma prior distribution.

The effective sample size is lower for data with poor quality (Table 3). The uniform prior has relatively much smaller n_{eff} compared to the other priors. Performance of the other two is worse for data with poor quality and when sample size is five. ESS increases with sample size for the inverse gamma prior, whereas no strong corresponding pattern for the gamma prior is seen. Using an inverse gamma on the SSD spread results in a model more sensitive to the choice of prior on the SSD mean (seen by the higher n_{eff} for $\mu_\mu = 0$ compared $|\mu_\mu| = 4$).

Table 3: Number out of 10 iterations with divergent transitions during post-warmup and average effective sample size for the different priors on σ and μ_μ and data sets.

			Frequency of divergences			n_{eff}		
Prior on σ	Data quality	Sample size	$\mu_\mu = -4$	$\mu_\mu = 0$	$\mu_\mu = 4$	$\mu_\mu = -4$	$\mu_\mu = 0$	$\mu_\mu = 4$
Uniform	Good	5	2	3	3	16327	10955	10191
Uniform	Good	7	0	1	0	16169	10785	10369
Uniform	Good	10	0	0	0	160121	10728	10448
Uniform	Good	13	1	1	1	15855	11885	12313
Uniform	Good	20	0	0	0	15697	19461	19322
Uniform	Bad	5	9	10	7	15540	19461	19322
Uniform	Bad	7	7	7	8	15382	7082	6987
Uniform	Bad	10	4	6	6	15225	7096	6906
Uniform	Bad	13	2	2	2	15067	9159	8439
Uniform	Bad	20	0	0	0	14910	11008	10970
Gamma	Good	5	1	1	2	20000	20000	20000
Gamma	Good	7	0	0	0	20000	20000	20000
Gamma	Good	10	0	0	0	20000	20000	20000
Gamma	Good	13	0	0	0	20000	20000	20000
Gamma	Good	20	0	0	0	20000	20000	20000
Gamma	Bad	5	8	6	7	16275	17130	15365
Gamma	Bad	7	5	5	4	17185	17506	17250
Gamma	Bad	10	5	4	4	16230	16431	15944
Gamma	Bad	13	1	1	2	17303	17741	17979
Gamma	Bad	20	0	0	0	19295	19355	19402
Inv-Gamma	Good	5	0	0	0	14125	19468	16132
Inv-Gamma	Good	7	0	0	0	20000	20000	20000
Inv-Gamma	Good	10	0	0	0	20000	20000	20000
Inv-Gamma	Good	13	0	0	0	20000	20000	20000
Inv-Gamma	Good	20	0	0	0	20000	20000	20000
Inv-Gamma	Bad	5	1	0	1	12145	16124	12668
Inv-Gamma	Bad	7	0	0	0	14733	17592	15877
Inv-Gamma	Bad	10	0	0	0	16978	18281	17131
Inv-Gamma	Bad	13	0	0	0	17647	18339	18468
Inv-Gamma	Bad	20	0	0	0	17675	18086	18740

4.2 Impact of uncertainty on SSD parameters

Uncertainty in the SSD parameters and toxicity levels is influenced by the sample size and quality in data underlying the Bayesian calibration (example in Figure 10).Data with poor, compared to good, quality results in posterior samples with relatively wider 95% credible intervals on the marginal distribution of the SSD parameters (Table 10).

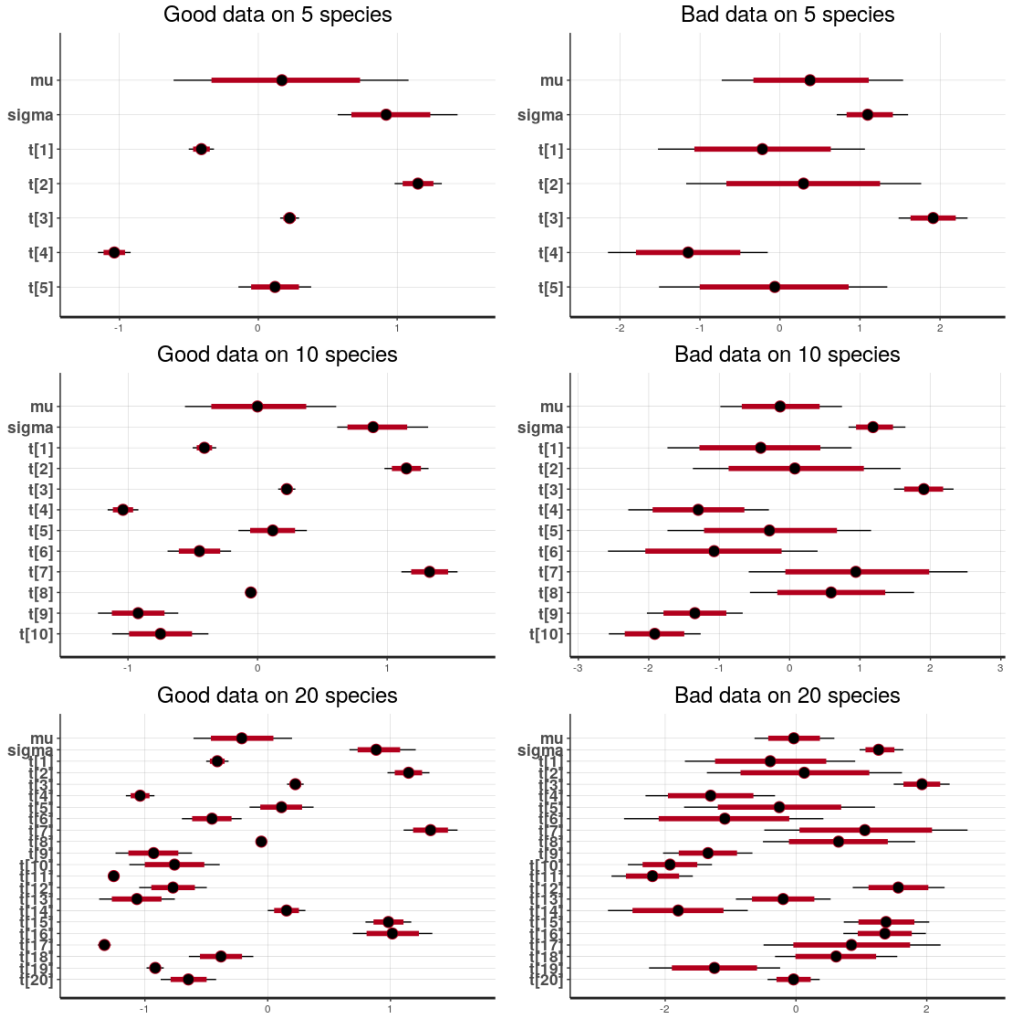

Figure 10: Posterior sample of parameters from a SSD calibrated on toxicity data of different sample sizes and with good versus poor quality, with a gamma prior on σ^2 and hyperparameter for μ, $\mu_\mu = 4$.

4.3 Accuracy of the Bayesian SSD

A model is here seen as accurate when the true values are within the bounds of the marginal marginal HPD intervals derived from the posterior sample (example in Figure 11). A summary of the HPDIs are in Appendix A. The true values on parameters and 5th percentile for the SSD are in most cases covered in all models (Figure 12).

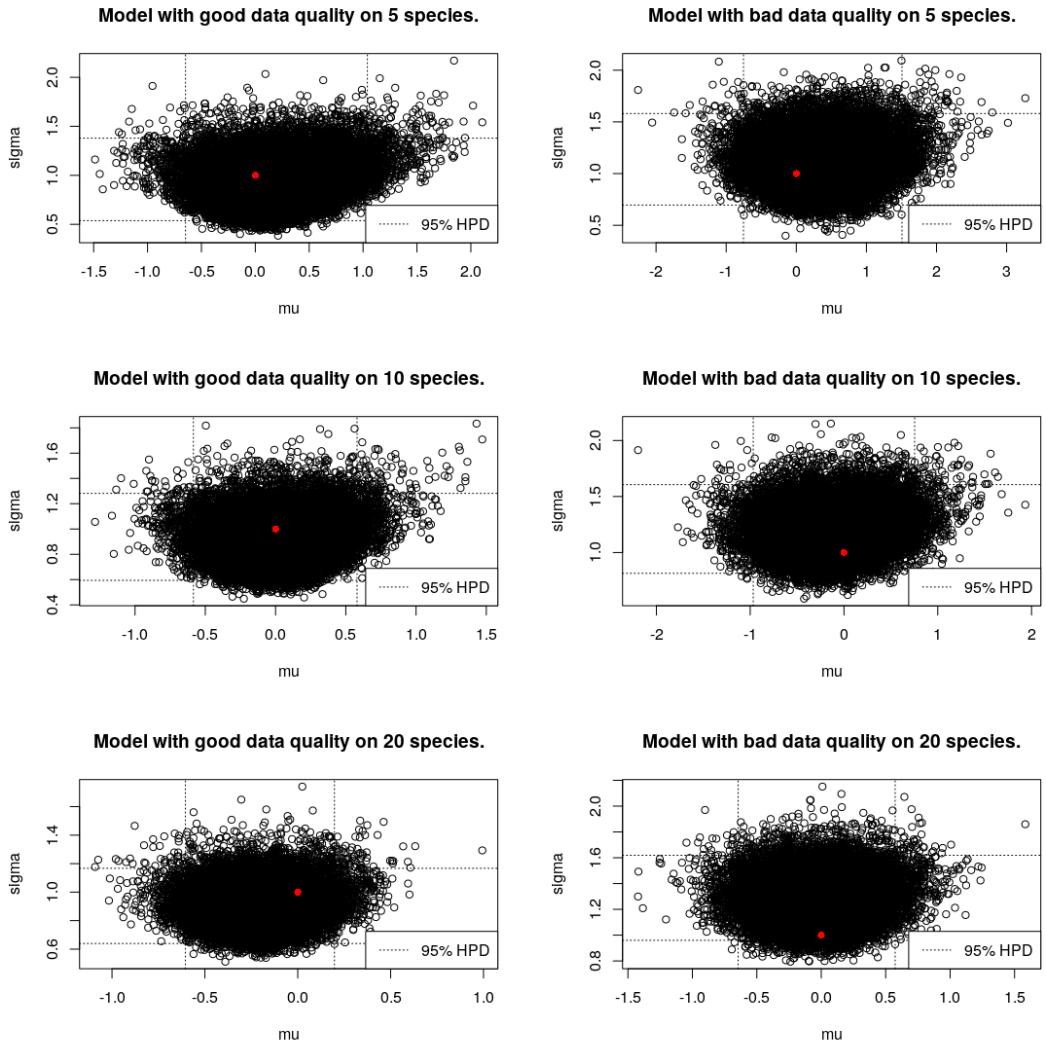

Figure 11: Posterior sample vs true value for different data from one iteration with a gamma prior on σ and hyperparameter $\mu_\mu = 4$. Dotted lines indicate the 95% HPDIs for marginals. The red dot represents the true value of the parameters, $\mu = 0$ and $\sigma = 1$.

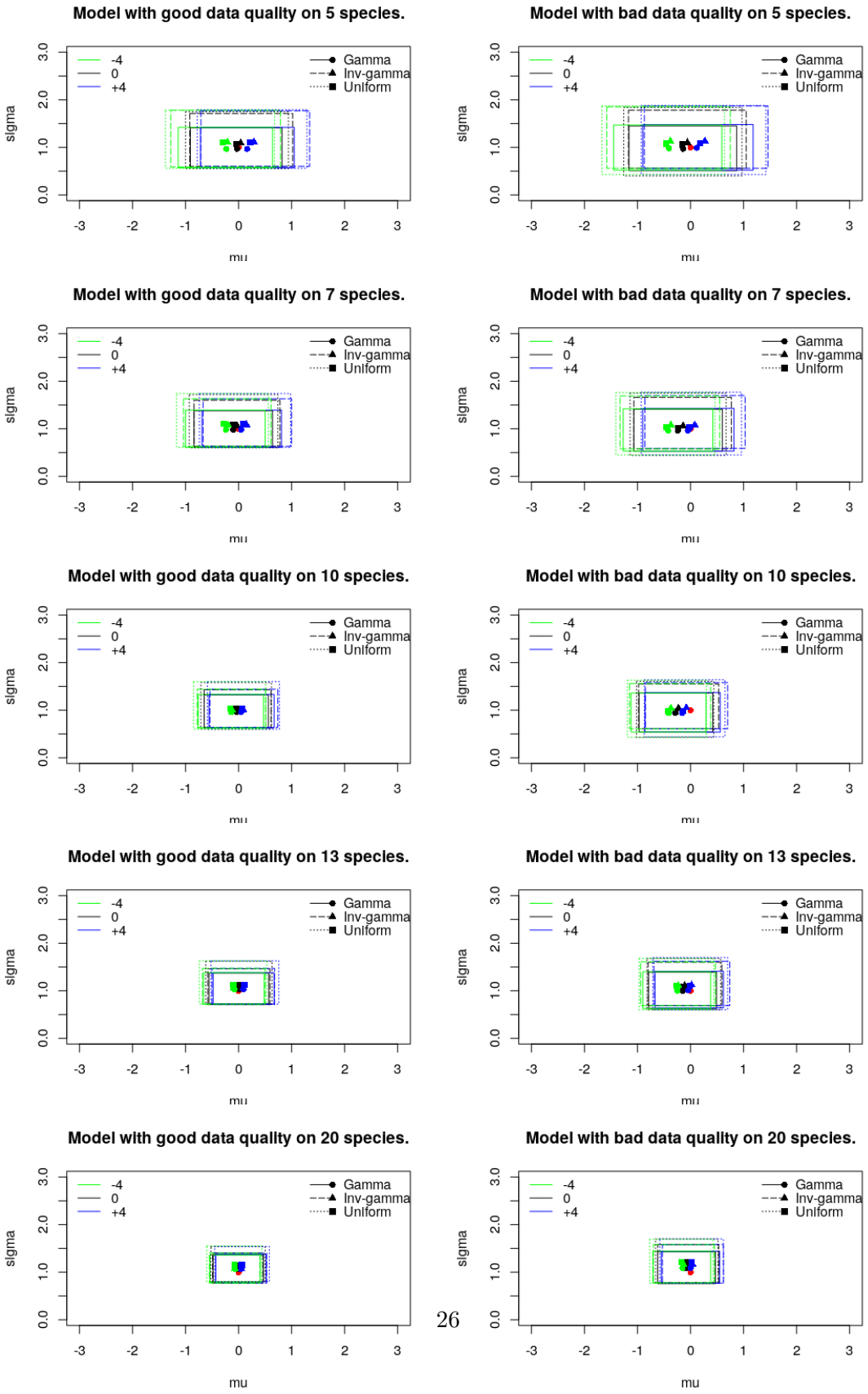

Figure 12: Plot of averaged posterior means, μ respective σ, with corresponding bounds of marginal HPDIs averaged over 10 iterations. Green displays models with $\mu_\mu = 4$, blue for models with $\mu_\mu = -4$ and black for models with $\mu_\mu = 0$. The red circle represents the true value of the parameters.

4.4 Impact of Uncertainty on the SSD

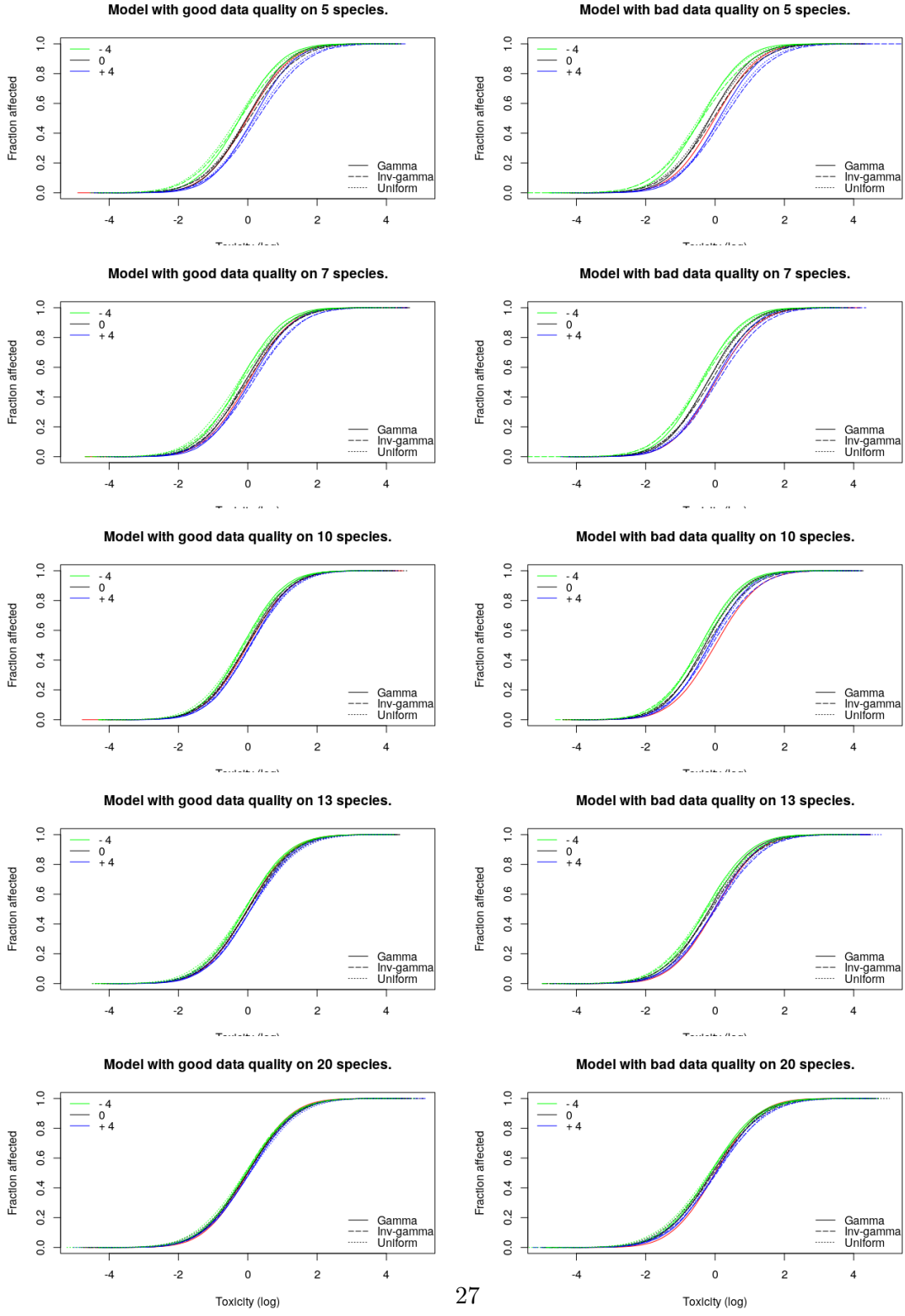

Figure 13: CDF plot of SSD with averaged posterior means of μ respective σ. Green displays models with $\mu_\mu = 4$, blue for models with $\mu_\mu = -4$ and black for models with $\mu_\mu = 0$. The red curve represents the true CDF.

The SSD is sensitive to choice of priors, sample size and data quality. The differences in cumulative distribution of the Bayesian SSD decreases going from top left to bottom right in Figure 13. For data with poor quality and 10 species there is, quite unexpectedly, a difference between the true CDF and sampled CDF that is both systematic and large at the same time. The choice of prior of the mean shifts the CDF upwards and downwards dependent the direction of the prior mean.

The 5^{th} percentile of the SSD is clearly sensitive to the choice of prior on μ for all sample sizes (Figure 14). The gamma prior resulted in the smallest differences between SSD variance between good and poor data quality.

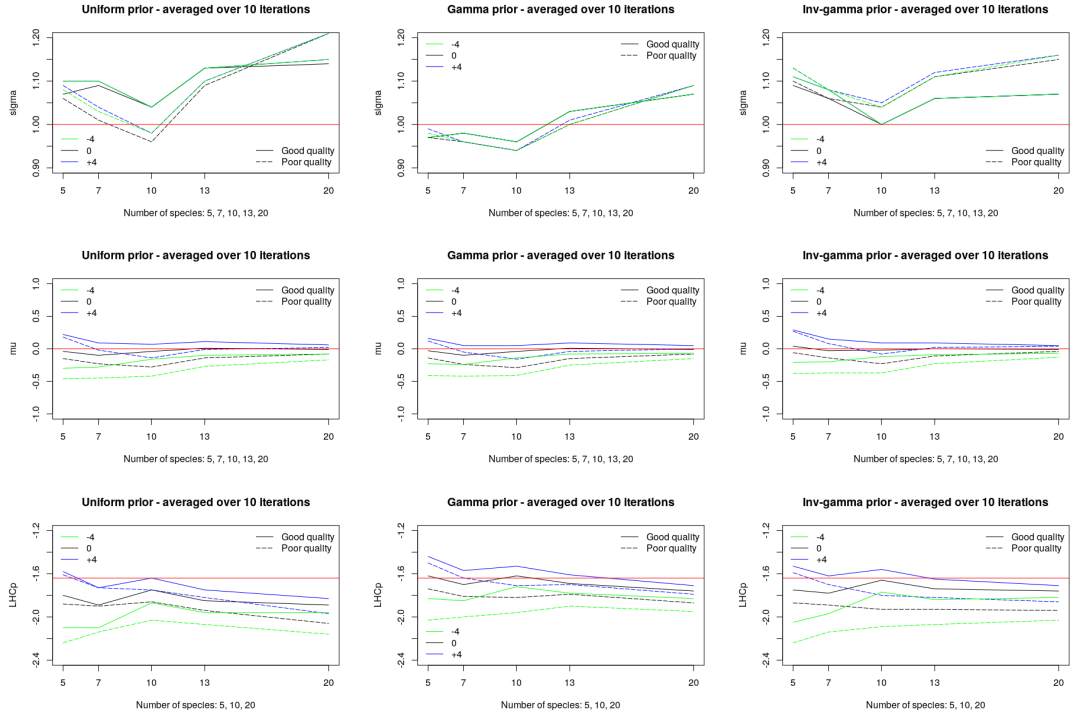

Figure 14: Averages of posterior means (μ, σ and LHC_p) vs data size for different data. The true values ($\mu = 0$, $\sigma = 1$ and $\text{LHC}_p = 1.64$) are added as red lines. Left: Uniform prior. Middle: Gamma prior. Right: Inverse-gamma prior. Note the differences in scales on the y-axes.

4.5 Bayes Optimal Hazardous Concentration and Bias

Bias in the hazard assessment is here defined as the difference between 1) the Bayes optimal hazardous concentration under uncertainty and 2) the optimal hazardous concentration under perfect information (Eq. (3)). The first, BLHC_p^{α}, is derived by minimizing the expected loss function, where the expectation is taken over the predictive posterior

sample. The second, LHC_p^α, is -1.64 for all the three different choices of α.

Here bias is increasing when the absolute difference becomes larger. Bias can be associated to overprotection or underprotection as follows: The presence of a bias with a Bayes optimal HC lower compared to the optimal HC under perfect information implies that more than 95 percent of species are protected. This corresponds in risk management terms to *err on the safe side* (i.e. overprotection). The shape of the LINEX loss function is chosen to give this effect. The opposite, i.e. to have a hazard concentration failing to protect at least 95 percent of the species, is something the hazard assessment seek to avoid (i.e. underprotection).

Bias increases (i.e. the Bayes optimal HC takes lower values) with a higher loss parameter α (Figure 15). This is expected since a higher α assigns more concern on values further away from the decision and shifts the optimal hazardous concentration under uncertainty downwards in relation to the optimal hazardous concentration under perfect information (Figure 5). The effect of the loss parameter on bias is consistent (Tables in Appendix B).

4.6 Impact of Uncertainty on Bias in Hazard Assessment

There is an impact from sample size and the quality of toxicity data on uncertainty in the SSD and the choice of hazardous concentration (Figure 14). This can be seen since there are differences in the HC due to the choice of priors evaluated here, also at the highest sample size. Everything else equal, HA based on data with poor quality results in more conservative (i.e. lower) choices of the hazardous concentration. This effect remains even as sample size increases to 20.

Generally, bias increases with a reduction in sample size. Thus, the Bayesian analysis compensate for low sample size by increasing uncertainty. Quite unexpectedly, in some cases bias went from small to large with increasing sample, especially for data with good quality, to the extent that the hazardous concentration were larger than the Bayes optimal concentration under perfect information (Figure 15). This exception is most prominent for the high prior on the SSD mean, $\mu_\mu = 4$, where the prior shifts the SSD upwards which counteracts the influence of higher uncertainty in the SSD.

The choice of prior for the SSD mean becomes less influential as sample size increases, and therefore it appears as if bias increases as samples size goes from 5 to 20 and sometimes from 10 to 20. This pattern is most strong for the uniform prior on the SSD spread and for the lowest loss parameter $\alpha = 1$. Thus, the influence by the choice of prior for SSD mean (Figure 15) is affecting the Bayes optimal HC in a systematic way where it may result in underprotection at low sample sizes.

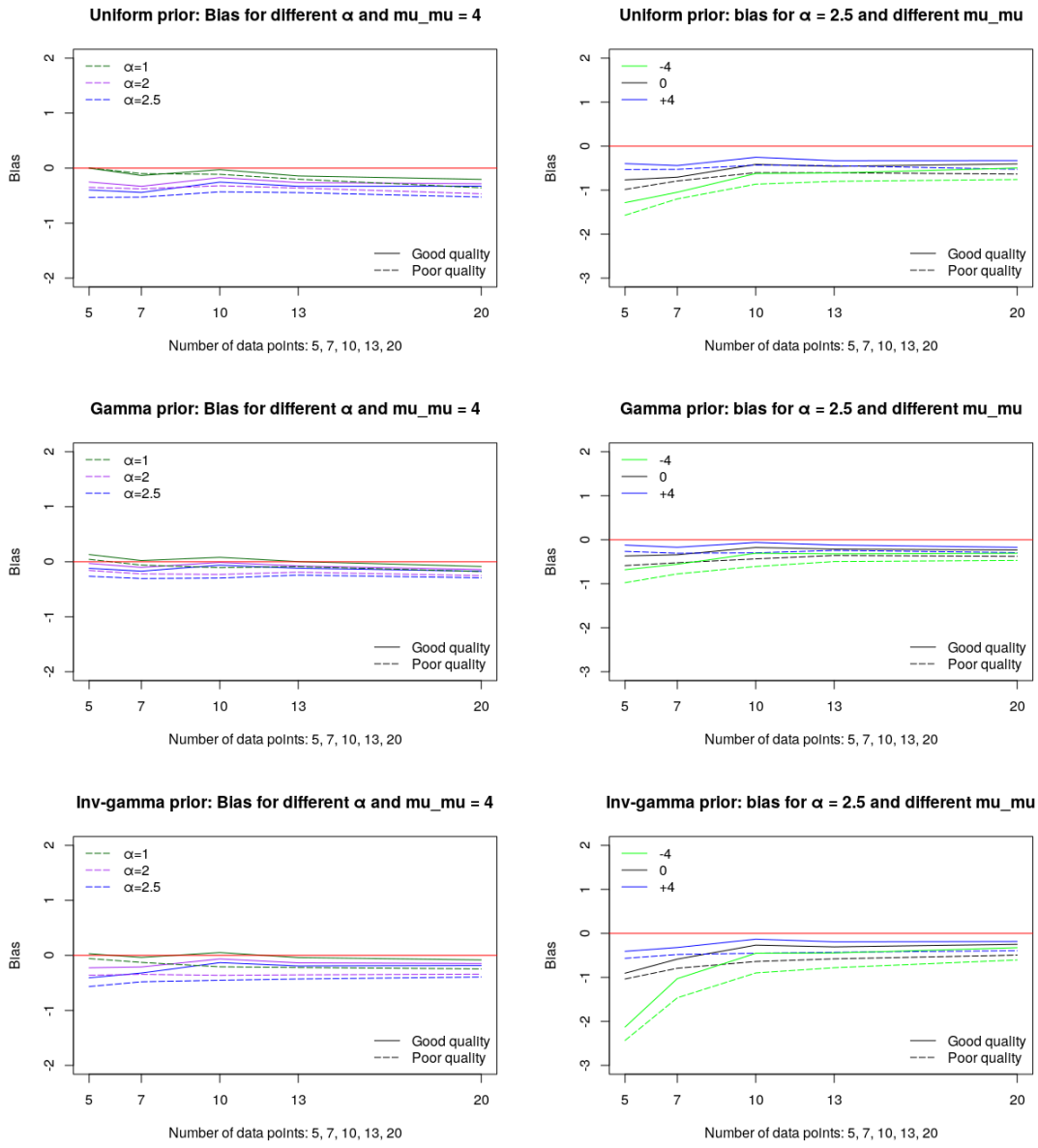

Figure 15: Right: Bias in Bayes optimal hazardous concentration averaged over 10 iterations for different α for $\mu = 4$. Left: Bias in Bayes optimal hazardous concentration averaged over 10 iterations for different μ_μ and $\alpha = 2.5$. Red solid line represents bias under perfect information.

Choosing an inverse-gamma as a prior for the SSD variance results in the most conservative Bayes optimal hazardous concentrations (Figure 14). The SSD variance is in general over-estimated, with the exception for intermediate sample sizes, even for good

quality data. This finding, together with more accurate estimation of the mean of the SSD (especially for data with high quality) can explain the raise in the 5th percentile. This results in turn in a low and even positive bias in the hazardous concentration of intermediate sample sizes.

Models with inverse-gamma prior were more sensitive to sample size compared to the other priors. Choosing an inverse-gamma prior distribution can be a deliberate way to underestimate the true hazardous concentration to get overprotection. On the contrary, if the aim is to avoid overprotecting species in the system, a gamma prior distribution could be used.

Since the hyperparameter μ_μ results in a systematic shift in bias (Figure 15), the largest over-protection is derived from using a combination of inverse-gamma prior distribution on the SSD variance and a low hyperparameter for the mean of the SSD (-4).

5 Discussion

5.1 A Bayesian Hazard Assessment

This thesis study the impact of uncertainty due to small sample size and varying data quality on the choice of a safety threshold. The choice of hazardous concentration (HC) in chemical hazard assessment was used as a case. Hazard assessment is the decision problem to find the HC_p such that is at least $1 - p\%$ of the species in a system is protected from a toxic compound. A Species Sensitivity Distribution (SSD) is used as a simple model of variability among the species in the ecological system which are to be protected.

This thesis takes inspiration from Hickey et al (2009) [5] and Aldenberg & Jaworska (2000) [6]. The latter were one of the first to promote considering uncertainty in the SSD parameters and to give lookup tables to calculate the HC. They showed that the HA problem (under the assumption of a Normal SSD) can be reduced to finding p^{th} percentile where uncertainty is treated using a non-central t-distribution [12]. Hickey et al [5] raised concern about the unknown level of risk aversion in HC, and recalled how a HA can be framed as a Bayesian decision problem. They let the HC be the Bayes optimal decision determined by the LINEX loss function. The LINEX loss function can help assessors to be more explicit about their level of risk aversion and have been used to set the appropriate sample size.

A Bayesian approach to quantify uncertainty in a HA using SSD has been done by others as well such as Fox [13]. In this thesis, the Bayesian HA-SSD is built as a Bayesian Evidence Synthesis (BES), where the HC is the Bayes optimal decision. An advantage with a BES is that all uncertainty is quantified in one calculation step. By all, we mean uncertainty in parameters as well as in decision nodes. Thus, the impact of changing a hyper parameter och data point is immediate, which makes the BES useful for novel ways of sensitivity analysis.

Sensitivity analysis can be applied to evaluate the influence of chosen decision parameters. For example, a higher value on the loss parameter, α, adjusts the hazardous concentration downwards and could therefore possibly be used to deliberately increase safety levels. Further studies, using real SSD data, would be useful to study to what extent and how the loss parameters could be used to account for small data sizes or data with poor quality.

5.2 Robust Analysis

Robust Bayesian analysis is often done to study if the choice of priors matters. Here, uniform prior distribution on the spread of the SSD is not a good choice due to high amount of divergent transitions and low effective sample size. Models with gamma prior

are less sensitive to data quality and sample size compared to models with the inverse-gamma prior.

The central tendency of the mean of the SSD, i.e. the hyperparameter μ_μ, has a large influence on the HC, especially when the data is small. The estimated mean of the SSD is systematically shifted upwards respectively downwards when μ_μ is set in the outer range of the true (but unknown) SSD, i.e. either 4 or -4. The difference compared to a μ_μ equal to the true value ($\mu_\mu = 0$) were smaller as data size increase.

In this study, robust analysis was applied to study the impact of some sources to uncertainty on a decision. The relative change in the sensitivity to the choice of priors was used as a measure of uncertainty from these sources, which were sample size and data quality. Extending this approach even further and keep all probability distributions from each individual Bayesian model (a.k.a. a generalisation of Bayesian analysis) is suggested to be a novel way for uncertainty analysis with respect to data issues in a hazard assessment, which deserves further evaluation and applications [14].

Another use of Robust Bayesian analysis is to use different priors to quantify uncertainty beyond the Bayesian approach. This is not done here.

5.3 The simulation study

The use of Robust Bayesian Evidence Synthesis was evaluated in a simulation study. The chosen simulation design was kept relatively small. Improvements can be made by running more than ten iterations to get a more reliable results. It is also possible to use different ways to generate artificial toxicity data. In this study, one data per species was used, whereas more than one data per species is occurring in realistic problems.

Another way to generate data is to manipulate real data. This would open up for an evaluation of the uncertainty analysis under more realistic conditions. In reality, obtaining toxicity data with known estimation errors are not always easy.

Then there is always the possibility to increase the number of steps in the MCMC simulation to improve convergence properties.

Other choices of priors for the SSD can be studied. It would be interesting to use other prior distributions on the spread of the SSD, such as half t-distribution or half-Cauchy distribution to see what the result would become. Another suggestion is to study the influence of hyperparameters for a given family of distributions on the priors for the SSD spread, e.g gamma distribution with different modes. In this study, only one choice of hyperparameters were used for each of the prior distributions on SSD spread.

6 Conclusions and Recommendations

This thesis introduces a novel approach for uncertainty analysis in hazard assessment using species sensitivity distributions (SSD) to account for species variability. A bit different from common practice, the assessment uses information of quality in toxicity values to better account for small data. This is done by treating the SSD as a meta analysis, weighing each species toxicity by known estimation errors in the toxicity values.

To better account for small differences the impact of uncertainty, the quantification of uncertainty in the inputs and outputs of the hazard assessment was made in one step. This was made possible by integrating the SSD meta analysis in the hazard assessment as a Bayesian Evidence Synthesis.

Finally, the robustness to small sample sizes can be evaluated by Robust Bayesian analysis. Here Robust Bayesian analysis was applied to evaluate the influence of different choices of priors on the outcome of the hazard assessment. Since the choice of prior is expected to have large influence under small sample sizes, robust analysis can show how sensitive an hazard assessment is to sample size without any reference SSD.

A simulation study showed that the hazard assessment is sensitive to sample size and quality in toxicity data. The Robust Bayesian Evidence Synthesis suggested is a possible complement to the use uncertainty factors, which is criticized to be overly conservative. In this framework, the choice of prior can be made to adjust safety levels informed by data.

The hazardous concentration is here shown to be sensitive to both the choice of prior for the mean and spread of the SSD. Bias in the hazard assessment were less sensitive to sample size with a gamma prior on the spread of the SSD, compared to an inverse-gamma distribution. Models with uniform prior distribution on the spread of the SSD had bad properties and is not recommended. The most conservative choice of prior is an inverse-gamma for the SSD spread and a low SSD mean.

References

[1] European Chemicals Agency (2011). Guidance on information requirements and chemical safety assessment Part B: Hazard Assessment. https://echa.europa.eu/documents/10162/13643/information_requirements_part_b_en.pdf/7e6bf845-e1a3-4518-8705-c64b17cecae8 (Accessed 17.06.06)

[2] Posthuma, L., Suter II, G.W., Traas, T.P. (2002). Species Sensitivity Distributions in Ecotoxicology. CRC Press.

[3] Hansson, S.O. (1999). Adjusting Scientific Practices to the Precautionary Principle. Human and Ecological Risk Assessment 5, 909-921.

[4] Löfstedt, R., Schlag, A. (2016). Risk-risk tradeoffs: what should we do in Europe?. Journal of Risk Research: Vol. 20, Iss. 8.

[5] Hickey, G.L., Craig, P.S., Hart, A. (2009). On the Application of Loss Functions in Determining Assessment Factors for Ecological Risk. Ecotoxicology and Environmental Safety 72, 293-300.

[6] Aldenberg, T., Jaworska, J.S. (2000). Uncertainty of the Hazardous Concentration and Fraction Affected for Normal Species Sensitivity Distributions. Ecotoxicology and Environmental Safety 46, 1-18.

[7] Ades, A.E., Sculper, M., Sutton, A., Abrams, K., Cooper, N., Welton, N., Lu, G. (2006). Bayesian Methods for Evidence Synthesis in Cost-Effectiveness Analysis. Pharmacoeconomics 24, 1-19.

[8] Jackson, C., Jit, M., Sharples, L., DeAngelis, D. (2015). Calibration of complex models through Bayesian evidence synthesis: a demonstration and tutorial. Med Decis Making 35, 148-161.

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

[10] Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. Bayesian Analysis 1, Number 3, 515-533.

[11] Kruschke, J. (2014). Doing Bayesian Data Analysis: a tutorial with R, JAGS, and Stan. Second edition. Academic Press.

[12] Owen, D.B. (1968). A survey of properties and applications of the noncentral t-distribution. Technom 10, 445-478.

[13] Fox, D.R. (2010). A Bayesian approach for determining the no effect concentration and hazardous concentration in ecotoxicology. Ecotoxicology and Environmental Safety, 73, 123-131.

[14] Montgomery, V. (2009). New statistical methods in risk assessment by probability bounds. Doctoral thesis, Durham University.

Appendix A Averages of the Ten Posterior Means and HPDI

Table A.1: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Uniform prior on σ and $\mu_\mu = 0$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.04 (-0.75, 0.71)	-1 (-1.58, -0.48)	0.94 (-0.11, 1.91)	0	5	good
σ	1.07 (0.53, 1.53)	0.55 (0.13, 1.05)	1.75 (1.15, 2)	1	5	good
LHC_p	-1.8 (-2.49, -1.41)	-3.31 (-4.07, -2.69)	-0.58 (-1.05, -0.11)	-1.64	5	good
μ	-0.1 (-0.92, 0.66)	-0.93 (-1.73, -0.31)	0.74 (-0.06, 1.61)	0	7	good
σ	1.09 (0.75, 1.44)	0.6 (0.34, 0.97)	1.72 (1.31, 2)	1	7	good
LHC_p	-1.89 (-2.72, -1.25)	-3.23 (-4.08, -2.38)	-0.77 (-1.63, -0.13)	-1.64	7	good
μ	-0.04 (-0.86, 0.56)	-0.71 (-1.44, -0.15)	0.62 (-0.25, 1.37)	0	10	good
σ	1.04 (0.8, 1.3)	0.6 (0.42, 0.81)	1.58 (1.27, 1.89)	1	10	good
LHC_p	-1.75 (-2.35, -1.16)	-2.86 (-3.49, -2.17)	-0.8 (-1.51, -0.29)	-1.64	10	good
μ	0.01 (-0.3, 0.75)	-0.62 (-1.07, 0.05)	0.63 (0.23, 1.41)	0	13	good
σ	1.13 (0.82, 1.51)	0.72 (0.5, 1.12)	1.62 (1.21, 2)	1	13	good
LHC_p	-1.85 (-2.73, -1.12)	-2.88 (-3.86, -2.02)	-0.96 (-1.67, -0.25)	-1.64	13	good
μ	-0.01 (-0.25, 0.35)	-0.52 (-0.87, -0.16)	0.5 (0.15, 0.89)	0	20	good
σ	1.14 (0.89, 1.47)	0.8 (0.6, 1.09)	1.54 (1.2, 1.92)	1	20	good
LHC_p	-1.89 (-2.55, -1.4)	-2.72 (-3.54, -2.15)	-1.14 (-1.65, -0.75)	-1.64	20	good
μ	-0.15 (-0.74, 0.21)	-1.26 (-2.08, -0.81)	0.97 (0.56, 1.33)	0	5	bad
σ	1.06 (0.58, 1.43)	0.4 (0.02, 0.84)	1.84 (1.5, 2)	1	5	bad
LHC_p	-1.88 (-3.09, -0.91)	-3.64 (-4.87, -2.59)	-0.42 (-1.47, 0.26)	-1.64	5	bad
μ	-0.23 (-0.83, 0.1)	-1.14 (-1.89, -0.64)	0.67 (0.22, 1.15)	0	7	bad
σ	1.01 (0.46, 1.4)	0.44 (0.03, 0.86)	1.75 (1.2, 2)	1	7	bad
LHC_p	-1.9 (-2.91, -0.81)	-3.39 (-4.46, -2.17)	-0.62 (-1.42, 0.12)	-1.64	7	bad
μ	-0.28 (-0.87, 0.16)	-1.02 (-1.64, -0.44)	0.44 (-0.07, 0.86)	0	10	bad
σ	0.96 (0.44, 1.44)	0.43 (0.05, 0.98)	1.61 (1.01, 2)	1	10	bad
LHC_p	-1.86 (-2.67, -0.87)	-3.15 (-4.09, -2.09)	-0.78 (-1.4, 0.13)	-1.64	10	bad
μ	-0.14 (-0.66, 0.29)	-0.82 (-1.3, -0.46)	0.57 (-0.04, 1.08)	0	13	bad
σ	1.09 (0.63, 1.5)	0.6 (0.19, 1.07)	1.68 (1.13, 2)	1	13	bad
LHC_p	-1.94 (-2.77, -1.13)	-3.1 (-3.97, -2.14)	-0.9 (-1.68, -0.33)	-1.64	13	bad
μ	-0.08 (-0.45, 0.36)	-0.66 (-1.06, -0.24)	0.52 (0.18, 1)	0	20	bad
σ	1.21 (0.77, 1.47)	0.77 (0.41, 1.05)	1.7 (1.16, 1.93)	1	20	bad
LHC_p	-2.06 (-2.63, -1.23)	-3.06 (-3.65, -2)	-1.15 (-1.64, -0.5)	-1.64	20	bad

Table A.2: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Uniform prior on σ and $\mu_\mu = 4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	0.22 (-0.6, 1.08)	-0.78 (-1.43, -0.15)	1.29 (0.05, 2.29)	0	5	good
σ	1.1 (0.55, 1.55)	0.56 (0.13, 1.05)	1.79 (1.22, 2)	1	5	good
LHC_p	-1.58 (-2.08, -1.2)	-2.93 (-3.52, -2.51)	-0.45 (-1.01, 0.12)	-1.64	5	good
μ	0.09 (-0.72, 0.89)	-0.74 (-1.6, -0.05)	0.98 (0.19, 1.9)	0	7	good
σ	1.1 (0.76, 1.46)	0.61 (0.34, 0.98)	1.74 (1.35, 2)	1	7	good
LHC_p	-1.73 (-2.57, -1.16)	-2.95 (-3.81, -2.19)	-0.71 (-1.53, -0.06)	-1.64	7	good
μ	0.07 (-0.77, 0.73)	-0.59 (-1.35, -0.02)	0.77 (-0.14, 1.57)	0	10	good
σ	1.04 (0.8, 1.3)	0.6 (0.43, 0.81)	1.6 (1.28, 1.88)	1	10	good
LHC_p	-1.64 (-2.28, -1.08)	-2.68 (-3.28, -2.04)	-0.76 (-1.47, -0.22)	-1.64	10	good
μ	0.11 (-0.23, 0.86)	-0.52 (-0.91, 0.22)	0.76 (0.3, 1.54)	0	13	good
σ	1.13 (0.82, 1.51)	0.72 (0.49, 1.11)	1.63 (1.23, 1.99)	1	13	good
LHC_p	-1.75 (-2.56, -1.06)	-2.73 (-3.65, -1.91)	-0.9 (-1.55, -0.19)	-1.64	13	good
μ	0.06 (-0.21, 0.42)	-0.45 (-0.73, -0.07)	0.58 (0.2, 0.97)	0	20	good
σ	1.15 (0.89, 1.48)	0.8 (0.6, 1.08)	1.54 (1.21, 1.92)	1	20	good
LHC_p	-1.83 (-2.45, -1.35)	-2.62 (-3.37, -2.03)	-1.1 (-1.56, -0.69)	-1.64	20	good
μ	0.18 (-0.29, 0.54)	-0.93 (-1.61, -0.58)	1.42 (1.12, 1.82)	0	5	bad
σ	1.09 (0.65, 1.45)	0.43 (0.04, 0.84)	1.88 (1.59, 2)	1	5	bad
LHC_p	-1.61 (-2.67, -0.82)	-3.17 (-4.21, -2.34)	-0.28 (-1.24, 0.37)	-1.64	5	bad
μ	-0.02 (-0.54, 0.36)	-0.93 (-1.58, -0.58)	0.96 (0.63, 1.54)	0	7	bad
σ	1.04 (0.51, 1.41)	0.45 (0.05, 0.86)	1.77 (1.27, 2)	1	7	bad
LHC_p	-1.73 (-2.67, -0.79)	-3.09 (-4.09, -2.04)	-0.53 (-1.28, 0.13)	-1.64	7	bad
μ	-0.14 (-0.7, 0.25)	-0.87 (-1.49, -0.33)	0.65 (0.14, 1.1)	0	10	bad
σ	0.98 (0.5, 1.45)	0.44 (0.06, 0.98)	1.64 (1.04, 2)	1	10	bad
LHC_p	-1.75 (-2.5, -0.75)	-2.95 (-3.83, -1.89)	-0.71 (-1.31, 0.35)	-1.64	10	bad
μ	-0.01 (-0.57, 0.45)	-0.71 (-1.19, -0.37)	0.7 (0.04, 1.26)	0	13	bad
σ	1.1 (0.62, 1.5)	0.6 (0.16, 1.07)	1.7 (1.11, 2)	1	13	bad
LHC_p	-1.82 (-2.62, -1.07)	-2.91 (-3.73, -2.02)	-0.81 (-1.59, -0.27)	-1.64	13	bad
μ	0.02 (-0.35, 0.46)	-0.58 (-1.01, -0.17)	0.62 (0.26, 1.08)	0	20	bad
σ	1.21 (0.77, 1.47)	0.77 (0.41, 1.05)	1.7 (1.16, 1.94)	1	20	bad
LHC_p	-1.97 (-2.53, -1.18)	-2.91 (-3.53, -1.91)	-1.09 (-1.59, -0.47)	-1.64	20	bad

Table A.3: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Uniform prior on σ and $\mu_\mu = -4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.3 (-0.91, 0.33)	-1.38 (-1.78, -1.06)	0.68 (-0.22, 1.48)	0	5	good
σ	1.1 (0.53, 1.56)	0.56 (0.12, 1.07)	1.78 (1.18, 2)	1	5	good
LHC_p	-2.1 (-2.98, -1.58)	-3.86 (-4.76, -2.96)	-0.71 (-1.34, -0.3)	-1.64	5	good
μ	-0.28 (-1.09, 0.43)	-1.17 (-2, -0.57)	0.56 (-0.24, 1.4)	0	7	good
σ	1.1 (0.76, 1.46)	0.61 (0.34, 0.99)	1.74 (1.36, 2)	1	7	good
LHC_p	-2.1 (-2.92, -1.35)	-3.59 (-4.45, -2.69)	-0.86 (-1.69, -0.26)	-1.64	7	good
μ	-0.16 (-0.95, 0.43)	-0.85 (-1.56, -0.36)	0.51 (-0.36, 1.23)	0	10	good
σ	1.04 (0.8, 1.3)	0.6 (0.42, 0.82)	1.6 (1.26, 1.9)	1	10	good
LHC_p	-1.87 (-2.44, -1.25)	-3.1 (-3.79, -2.37)	-0.86 (-1.57, -0.35)	-1.64	10	good
μ	-0.1 (-0.42, 0.63)	-0.74 (-1.24, -0.04)	0.53 (0.15, 1.31)	0	13	good
σ	1.13 (0.82, 1.51)	0.72 (0.5, 1.11)	1.63 (1.23, 1.99)	1	13	good
LHC_p	-1.96 (-2.89, -1.2)	-3.06 (-4.12, -2.21)	-1.01 (-1.79, -0.31)	-1.64	13	good
μ	-0.08 (-0.34, 0.29)	-0.6 (-0.92, -0.21)	0.44 (0.12, 0.82)	0	20	good
σ	1.15 (0.89, 1.48)	0.8 (0.61, 1.09)	1.55 (1.2, 1.92)	1	20	good
LHC_p	-1.96 (-2.67, -1.45)	-2.84 (-3.74, -2.25)	-1.19 (-1.75, -0.78)	-1.64	20	good
μ	-0.46 (-1.19, -0.1)	-1.67 (-2.58, -1.09)	0.64 (0.1, 1.03)	0	5	bad
σ	1.08 (0.61, 1.43)	0.43 (0.04, 0.87)	1.87 (1.56, 2)	1	5	bad
LHC_p	-2.24 (-3.55, -1.15)	-4.26 (-5.42, -3.18)	-0.57 (-1.72, 0.18)	-1.64	5	bad
μ	-0.45 (-1.12, -0.09)	-1.41 (-2.18, -0.8)	0.46 (-0.02, 0.76)	0	7	bad
σ	1.03 (0.48, 1.43)	0.45 (0.02, 0.88)	1.76 (1.23, 2)	1	7	bad
LHC_p	-2.14 (-3.22, -0.93)	-3.83 (-5.02, -2.49)	-0.74 (-1.64, 0.12)	-1.64	7	bad
μ	-0.42 (-1.04, 0.08)	-1.19 (-1.89, -0.55)	0.3 (-0.23, 0.66)	0	10	bad
σ	0.98 (0.49, 1.45)	0.43 (0.06, 0.97)	1.63 (1.05, 2)	1	10	bad
LHC_p	-2.03 (-2.92, -0.97)	-3.43 (-4.47, -2.26)	-0.84 (-1.57, 0.19)	-1.64	10	bad
μ	-0.27 (-0.77, 0.13)	-0.97 (-1.47, -0.55)	0.44 (-0.16, 0.95)	0	13	bad
σ	1.1 (0.64, 1.5)	0.6 (0.21, 1.08)	1.68 (1.13, 2)	1	13	bad
LHC_p	-2.07 (-2.96, -1.2)	-3.33 (-4.24, -2.25)	-0.97 (-1.78, -0.32)	-1.64	13	bad
μ	-0.17 (-0.55, 0.26)	-0.77 (-1.19, -0.37)	0.43 (0.08, 0.89)	0	20	bad
σ	1.21 (0.77, 1.47)	0.77 (0.41, 1.06)	1.69 (1.16, 1.94)	1	20	bad
LHC_p	-2.16 (-2.73, -1.28)	-3.2 (-3.79, -2.12)	-1.21 (-1.69, -0.55)	-1.64	20	bad

Table A.4: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Gamma prior on σ^2 and $\mu_\mu = 0$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.03 (-0.75, 0.73)	-0.91 (-1.51, -0.28)	0.82 (0.01, 1.69)	0	5	good
σ	0.97 (0.73, 1.24)	0.57 (0.31, 0.85)	1.41 (1.19, 1.66)	1	5	good
LHC_p	-1.62 (-2.12, -1.11)	-2.77 (-3.3, -2.39)	-0.59 (-1.2, 0.02)	-1.64	5	good
μ	-0.1 (-0.93, 0.67)	-0.84 (-1.66, -0.1)	0.64 (-0.19, 1.48)	0	7	good
σ	0.98 (0.8, 1.19)	0.61 (0.44, 0.81)	1.38 (1.2, 1.59)	1	7	good
LHC_p	-1.7 (-2.53, -1.07)	-2.71 (-3.5, -2.17)	-0.8 (-1.63, -0.12)	-1.64	7	good
μ	-0.04 (-0.87, 0.57)	-0.65 (-1.42, -0.09)	0.56 (-0.31, 1.27)	0	10	good
σ	0.96 (0.82, 1.11)	0.63 (0.49, 0.76)	1.32 (1.17, 1.49)	1	10	good
LHC_p	-1.62 (-2.31, -1.11)	-2.48 (-3.1, -1.94)	-0.83 (-1.57, -0.28)	-1.64	10	good
μ	0.01 (-0.3, 0.75)	-0.56 (-0.94, 0.16)	0.57 (0.19, 1.35)	0	13	good
σ	1.03 (0.83, 1.28)	0.72 (0.55, 0.95)	1.37 (1.16, 1.66)	1	13	good
LHC_p	-1.69 (-2.36, -1.01)	-2.49 (-3.24, -1.82)	-0.94 (-1.47, -0.23)	-1.64	13	good
μ	-0.01 (-0.25, 0.35)	-0.48 (-0.76, -0.11)	0.47 (0.15, 0.87)	0	20	good
σ	1.07 (0.88, 1.3)	0.8 (0.64, 1)	1.37 (1.16, 1.62)	1	20	good
LHC_p	-1.76 (-2.27, -1.36)	-2.45 (-3.08, -1.99)	-1.13 (-1.57, -0.74)	-1.64	20	good
μ	-0.14 (-0.78, 0.23)	-1.16 (-1.88, -0.79)	0.87 (0.29, 1.3)	0	5	bad
σ	0.97 (0.84, 1.12)	0.51 (0.36, 0.7)	1.45 (1.32, 1.6)	1	5	bad
LHC_p	-1.74 (-2.59, -1.28)	-3.07 (-4.01, -2.6)	-0.57 (-1.37, -0.06)	-1.64	5	bad
μ	-0.24 (-0.84, 0.1)	-1.07 (-1.75, -0.65)	0.6 (0.04, 1.03)	0	7	bad
σ	0.96 (0.78, 1.13)	0.53 (0.33, 0.71)	1.41 (1.26, 1.58)	1	7	bad
LHC_p	-1.81 (-2.56, -1.32)	-2.94 (-3.76, -2.37)	-0.77 (-1.44, -0.3)	-1.64	7	bad
μ	-0.29 (-0.87, 0.16)	-0.97 (-1.57, -0.48)	0.42 (-0.13, 0.79)	0	10	bad
σ	0.94 (0.73, 1.19)	0.54 (0.33, 0.8)	1.36 (1.17, 1.59)	1	10	bad
LHC_p	-1.82 (-2.46, -1.13)	-2.82 (-3.53, -2.13)	-0.93 (-1.52, -0.26)	-1.64	10	bad
μ	-0.15 (-0.67, 0.27)	-0.79 (-1.25, -0.42)	0.49 (-0.05, 0.96)	0	13	bad
σ	1 (0.75, 1.26)	0.63 (0.4, 0.91)	1.39 (1.14, 1.65)	1	13	bad
LHC_p	-1.79 (-2.42, -1.24)	-2.71 (-3.38, -2.07)	-0.96 (-1.55, -0.45)	-1.64	13	bad
μ	-0.08 (-0.44, 0.35)	-0.62 (-1, -0.18)	0.47 (0.12, 0.94)	0	20	bad
σ	1.09 (0.81, 1.27)	0.76 (0.5, 0.94)	1.43 (1.13, 1.61)	1	20	bad
LHC_p	-1.87 (-2.36, -1.29)	-2.67 (-3.17, -2.02)	-1.12 (-1.59, -0.67)	-1.64	20	bad

Table A.5: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Gamma prior on σ^2 and $\mu_\mu = 4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	0.16 (-0.6, 0.97)	-0.72 (-1.37, -0.03)	1.05 (0.2, 1.93)	0	5	good
σ	0.97 (0.74, 1.25)	0.57 (0.32, 0.85)	1.42 (1.22, 1.68)	1	5	good
LHC_p	-1.44 (-1.99, -0.87)	-2.51 (-2.97, -2.04)	-0.48 (-1.1, 0.16)	-1.64	5	good
μ	0.05 (-0.79, 0.83)	-0.69 (-1.52, 0.05)	0.81 (0.01, 1.66)	0	7	good
σ	0.98 (0.8, 1.2)	0.61 (0.44, 0.81)	1.39 (1.2, 1.6)	1	7	good
LHC_p	-1.57 (-2.4, -0.92)	-2.51 (-3.33, -1.9)	-0.71 (-1.53, 0.03)	-1.64	7	good
μ	0.05 (-0.79, 0.7)	-0.56 (-1.34, 0.03)	0.67 (-0.21, 1.41)	0	10	good
σ	0.96 (0.82, 1.12)	0.63 (0.49, 0.77)	1.33 (1.17, 1.49)	1	10	good
LHC_p	-1.53 (-2.24, -1.01)	-2.34 (-2.98, -1.84)	-0.77 (-1.52, -0.22)	-1.64	10	good
μ	0.09 (-0.24, 0.83)	-0.48 (-0.84, 0.24)	0.67 (0.26, 1.43)	0	13	good
σ	1.03 (0.83, 1.29)	0.72 (0.54, 0.96)	1.38 (1.16, 1.65)	1	13	good
LHC_p	-1.61 (-2.24, -0.92)	-2.39 (-3.13, -1.72)	-0.9 (-1.4, -0.17)	-1.64	13	good
μ	0.05 (-0.21, 0.41)	-0.43 (-0.72, -0.06)	0.52 (0.2, 0.91)	0	20	good
σ	1.07 (0.88, 1.3)	0.8 (0.63, 1)	1.37 (1.16, 1.62)	1	20	good
LHC_p	-1.71 (-2.19, -1.31)	-2.37 (-2.94, -1.91)	-1.09 (-1.45, -0.71)	-1.64	20	good
μ	0.12 (-0.49, 0.52)	-0.9 (-1.59, -0.56)	1.18 (0.66, 1.62)	0	5	bad
σ	0.99 (0.85, 1.13)	0.52 (0.35, 0.7)	1.48 (1.35, 1.63)	1	5	bad
LHC_p	-1.5 (-2.32, -1.03)	-2.71 (-3.63, -2.26)	-0.37 (-1.18, 0.16)	-1.64	5	bad
μ	-0.05 (-0.63, 0.3)	-0.9 (-1.57, -0.52)	0.82 (0.26, 1.3)	0	7	bad
σ	0.96 (0.78, 1.13)	0.53 (0.34, 0.71)	1.43 (1.27, 1.58)	1	7	bad
LHC_p	-1.64 (-2.36, -1.18)	-2.71 (-3.52, -2.22)	-0.65 (-1.33, -0.25)	-1.64	7	bad
μ	-0.16 (-0.74, 0.26)	-0.86 (-1.46, -0.4)	0.56 (0, 0.92)	0	10	bad
σ	0.94 (0.74, 1.2)	0.54 (0.35, 0.81)	1.37 (1.18, 1.61)	1	10	bad
LHC_p	-1.71 (-2.34, -1.03)	-2.64 (-3.34, -1.99)	-0.83 (-1.42, -0.2)	-1.64	10	bad
μ	-0.04 (-0.58, 0.4)	-0.67 (-1.16, -0.28)	0.62 (0.04, 1.12)	0	13	bad
σ	1.01 (0.75, 1.26)	0.64 (0.39, 0.92)	1.41 (1.13, 1.66)	1	13	bad
LHC_p	-1.7 (-2.3, -1.17)	-2.57 (-3.22, -1.98)	-0.88 (-1.45, -0.44)	-1.64	13	bad
μ	0 (-0.36, 0.44)	-0.55 (-0.92, -0.11)	0.55 (0.2, 1.01)	0	20	bad
σ	1.09 (0.81, 1.28)	0.77 (0.51, 0.96)	1.44 (1.13, 1.62)	1	20	bad
LHC_p	-1.79 (-2.28, -1.25)	-2.56 (-3.1, -1.93)	-1.06 (-1.53, -0.59)	-1.64	20	bad

Table A.6: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Gamma prior on σ^2 and $\mu_\mu = -4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.23 (-0.9, 0.48)	-1.14 (-1.71, -0.5)	0.64 (-0.16, 1.49)	0	5	good
σ	0.97 (0.73, 1.25)	0.57 (0.31, 0.84)	1.42 (1.2, 1.68)	1	5	good
LHC_p	-1.83 (-2.32, -1.38)	-3.09 (-3.72, -2.75)	-0.7 (-1.26, -0.12)	-1.64	5	good
μ	-0.24 (-1.07, 0.51)	-1 (-1.8, -0.3)	0.5 (-0.32, 1.34)	0	7	good
σ	0.98 (0.8, 1.2)	0.61 (0.44, 0.82)	1.39 (1.21, 1.6)	1	7	good
LHC_p	-1.85 (-2.67, -1.25)	-2.94 (-3.72, -2.39)	-0.87 (-1.7, -0.2)	-1.64	7	good
μ	-0.14 (-0.94, 0.46)	-0.76 (-1.5, -0.22)	0.47 (-0.37, 1.15)	0	10	good
σ	0.96 (0.82, 1.12)	0.63 (0.5, 0.77)	1.33 (1.18, 1.5)	1	10	good
LHC_p	-1.72 (-2.39, -1.22)	-2.63 (-3.23, -2.1)	-0.89 (-1.62, -0.33)	-1.64	10	good
μ	-0.08 (-0.37, 0.66)	-0.66 (-1.06, 0.07)	0.48 (0.12, 1.24)	0	13	good
σ	1.03 (0.83, 1.29)	0.72 (0.54, 0.96)	1.38 (1.15, 1.65)	1	13	good
LHC_p	-1.78 (-2.49, -1.1)	-2.62 (-3.45, -1.92)	-1 (-1.61, -0.3)	-1.64	13	good
μ	-0.07 (-0.32, 0.3)	-0.54 (-0.84, -0.17)	0.41 (0.12, 0.79)	0	20	good
σ	1.07 (0.88, 1.3)	0.79 (0.63, 1)	1.36 (1.15, 1.62)	1	20	good
LHC_p	-1.83 (-2.36, -1.4)	-2.53 (-3.18, -2.04)	-1.17 (-1.62, -0.78)	-1.64	20	good
μ	-0.41 (-1.07, -0.05)	-1.45 (-2.23, -1.06)	0.59 (-0.02, 0.97)	0	5	bad
σ	0.98 (0.83, 1.14)	0.52 (0.35, 0.71)	1.47 (1.33, 1.62)	1	5	bad
LHC_p	-2.03 (-2.92, -1.55)	-3.45 (-4.37, -2.94)	-0.73 (-1.6, -0.24)	-1.64	5	bad
μ	-0.42 (-1.05, -0.05)	-1.27 (-2, -0.87)	0.42 (-0.17, 0.8)	0	7	bad
σ	0.96 (0.78, 1.14)	0.53 (0.31, 0.72)	1.42 (1.24, 1.59)	1	7	bad
LHC_p	-2 (-2.79, -1.47)	-3.23 (-4.11, -2.62)	-0.88 (-1.61, -0.36)	-1.64	7	bad
μ	-0.41 (-1.01, 0.05)	-1.12 (-1.73, -0.63)	0.29 (-0.29, 0.69)	0	10	bad
σ	0.94 (0.74, 1.19)	0.54 (0.33, 0.81)	1.36 (1.16, 1.6)	1	10	bad
LHC_p	-1.96 (-2.61, -1.24)	-3.01 (-3.68, -2.31)	-1 (-1.58, -0.3)	-1.64	10	bad
μ	-0.25 (-0.76, 0.15)	-0.9 (-1.36, -0.55)	0.38 (-0.16, 0.83)	0	13	bad
σ	1 (0.76, 1.26)	0.63 (0.4, 0.89)	1.4 (1.14, 1.63)	1	13	bad
LHC_p	-1.9 (-2.56, -1.31)	-2.86 (-3.54, -2.21)	-1.02 (-1.63, -0.53)	-1.64	13	bad
μ	-0.15 (-0.52, 0.28)	-0.71 (-1.11, -0.29)	0.38 (0.03, 0.82)	0	20	bad
σ	1.09 (0.81, 1.28)	0.77 (0.51, 0.95)	1.44 (1.14, 1.62)	1	20	bad
LHC_p	-1.95 (-2.46, -1.35)	-2.77 (-3.34, -2.09)	-1.18 (-1.67, -0.7)	-1.64	20	bad

Table A.7: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Inverse-gamma prior on σ and $\mu_\mu = 0$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	0.04 (-0.66, 0.72)	-0.92 (-1.41, -0.39)	1.02 (0.15, 1.82)	0	5	good
σ	1.09 (0.84, 1.48)	0.6 (0.46, 0.82)	1.71 (1.32, 2.34)	1	5	good
LHC_p	-1.75 (-2.4, -1.36)	-3.19 (-4.25, -2.81)	-0.54 (-1.14, 0.05)	-1.64	5	good
μ	-0.03 (-0.92, 0.66)	-0.84 (-1.7, -0.2)	0.78 (-0.14, 1.51)	0	7	good
σ	1.06 (0.85, 1.34)	0.63 (0.51, 0.8)	1.6 (1.28, 2.02)	1	7	good
LHC_p	-1.78 (-2.61, -1.21)	-2.98 (-3.8, -2.38)	-0.73 (-1.64, -0.08)	-1.64	7	good
μ	-0.02 (-0.87, 0.57)	-0.65 (-1.45, -0.12)	0.61 (-0.3, 1.31)	0	10	good
σ	1 (0.84, 1.19)	0.64 (0.54, 0.75)	1.43 (1.21, 1.69)	1	10	good
LHC_p	-1.66 (-2.35, -1.17)	-2.61 (-3.21, -2.02)	-0.81 (-1.58, -0.27)	-1.64	10	good
μ	0 (-0.3, 0.75)	-0.58 (-1, 0.13)	0.59 (0.22, 1.36)	0	13	good
σ	1.06 (0.85, 1.42)	0.71 (0.57, 0.96)	1.46 (1.18, 1.96)	1	13	good
LHC_p	-1.74 (-2.58, -1.08)	-2.64 (-3.75, -1.86)	-0.94 (-1.52, -0.24)	-1.64	13	good
μ	-0.01 (-0.25, 0.35)	-0.49 (-0.74, -0.13)	0.46 (0.15, 0.86)	0	20	good
σ	1.07 (0.89, 1.39)	0.77 (0.64, 1)	1.4 (1.16, 1.82)	1	20	good
LHC_p	-1.76 (-2.42, -1.37)	-2.49 (-3.38, -2.04)	-1.09 (-1.59, -0.72)	-1.64	20	good
μ	-0.06 (-0.76, 0.24)	-1.17 (-1.92, -0.84)	1.05 (0.44, 1.39)	0	5	bad
σ	1.1 (0.95, 1.28)	0.56 (0.49, 0.67)	1.78 (1.54, 2.05)	1	5	bad
LHC_p	-1.87 (-2.82, -1.45)	-3.48 (-4.56, -2.93)	-0.48 (-1.42, -0.11)	-1.64	5	bad
μ	-0.14 (-0.55, 0.1)	-1.07 (-1.53, -0.74)	0.77 (0.42, 1.11)	0	7	bad
σ	1.06 (0.87, 1.28)	0.58 (0.49, 0.68)	1.66 (1.36, 2)	1	7	bad
LHC_p	-1.89 (-2.47, -1.48)	-3.25 (-3.89, -2.58)	-0.71 (-1.29, -0.41)	-1.64	7	bad
μ	-0.23 (-0.48, 0.16)	-0.98 (-1.24, -0.51)	0.53 (0.28, 0.84)	0	10	bad
σ	1.04 (0.82, 1.33)	0.61 (0.5, 0.81)	1.55 (1.23, 1.97)	1	10	bad
LHC_p	-1.93 (-2.49, -1.24)	-3.05 (-3.88, -2.23)	-0.93 (-1.28, -0.34)	-1.64	10	bad
μ	-0.11 (-0.44, 0.28)	-0.8 (-1.23, -0.45)	0.59 (0.29, 1.01)	0	13	bad
σ	1.11 (0.81, 1.4)	0.69 (0.49, 0.89)	1.6 (1.16, 1.98)	1	13	bad
LHC_p	-1.93 (-2.61, -1.32)	-2.99 (-3.88, -2.16)	-0.97 (-1.53, -0.53)	-1.64	13	bad
μ	-0.04 (-0.25, 0.36)	-0.62 (-0.8, -0.24)	0.53 (0.3, 0.93)	0	20	bad
σ	1.15 (0.83, 1.38)	0.78 (0.56, 0.96)	1.58 (1.15, 1.85)	1	20	bad
LHC_p	-1.94 (-2.4, -1.34)	-2.86 (-3.47, -2.04)	-1.12 (-1.47, -0.69)	-1.64	20	bad

Table A.8: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Inverse-gamma prior on σ^2 and $\mu_\mu = 4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.21 (-0.81, 0.38)	-1.28 (-1.85, -0.91)	0.79 (-0.06, 1.56)	0	5	good
σ	1.11 (0.84, 1.54)	0.59 (0.46, 0.81)	1.78 (1.33, 2.46)	1	5	good
LHC_p	-2.05 (-2.94, -1.75)	-3.8 (-5.45, -3.21)	-0.65 (-1.2, -0.05)	-1.64	5	good
μ	-0.2 (-1.08, 0.47)	-1.04 (-1.88, -0.44)	0.61 (-0.29, 1.33)	0	7	good
σ	1.08 (0.86, 1.36)	0.63 (0.5, 0.81)	1.63 (1.3, 2.08)	1	7	good
LHC_p	-1.97 (-2.78, -1.42)	-3.34 (-4.33, -2.57)	-0.82 (-1.73, -0.2)	-1.64	7	good
μ	-0.12 (-0.95, 0.44)	-0.78 (-1.55, -0.26)	0.51 (-0.38, 1.18)	0	10	good
σ	1 (0.85, 1.19)	0.63 (0.53, 0.75)	1.44 (1.22, 1.7)	1	10	good
LHC_p	-1.77 (-2.44, -1.27)	-2.83 (-3.47, -2.21)	-0.87 (-1.64, -0.31)	-1.64	10	good
μ	-0.09 (-0.4, 0.64)	-0.68 (-1.17, 0.03)	0.5 (0.15, 1.29)	0	13	good
σ	1.06 (0.85, 1.42)	0.71 (0.57, 0.94)	1.47 (1.18, 1.96)	1	13	good
LHC_p	-1.84 (-2.74, -1.17)	-2.79 (-4.05, -1.99)	-0.98 (-1.64, -0.26)	-1.64	13	good
μ	-0.07 (-0.29, 0.29)	-0.55 (-0.87, -0.16)	0.4 (0.09, 0.77)	0	20	good
σ	1.07 (0.89, 1.39)	0.77 (0.64, 1)	1.4 (1.17, 1.81)	1	20	good
LHC_p	-1.82 (-2.52, -1.42)	-2.58 (-3.52, -2.09)	-1.14 (-1.64, -0.76)	-1.64	20	good
μ	-0.38 (-1.11, -0.09)	-1.58 (-2.29, -1.26)	0.75 (0.1, 1.08)	0	5	bad
σ	1.13 (0.96, 1.34)	0.56 (0.48, 0.68)	1.85 (1.6, 2.19)	1	5	bad
LHC_p	-2.24 (-3.18, -1.8)	-4.16 (-5.15, -3.44)	-0.65 (-1.58, -0.27)	-1.64	5	bad
μ	-0.37 (-0.79, -0.07)	-1.33 (-1.79, -0.94)	0.55 (0.16, 0.85)	0	7	bad
σ	1.08 (0.88, 1.32)	0.58 (0.48, 0.7)	1.69 (1.38, 2.08)	1	7	bad
LHC_p	-2.14 (-2.71, -1.66)	-3.67 (-4.39, -2.94)	-0.83 (-1.4, -0.56)	-1.64	7	bad
μ	-0.37 (-0.61, 0.04)	-1.15 (-1.51, -0.68)	0.38 (0.12, 0.7)	0	10	bad
σ	1.04 (0.82, 1.34)	0.61 (0.48, 0.8)	1.56 (1.22, 1.97)	1	10	bad
LHC_p	-2.09 (-2.72, -1.38)	-3.31 (-4.25, -2.46)	-1.02 (-1.38, -0.44)	-1.64	10	bad
μ	-0.23 (-0.6, 0.15)	-0.94 (-1.39, -0.59)	0.47 (0.18, 0.88)	0	13	bad
σ	1.11 (0.81, 1.41)	0.69 (0.5, 0.9)	1.61 (1.18, 2)	1	13	bad
LHC_p	-2.07 (-2.79, -1.4)	-3.2 (-4.17, -2.3)	-1.06 (-1.65, -0.63)	-1.64	13	bad
μ	-0.13 (-0.32, 0.27)	-0.7 (-0.92, -0.3)	0.45 (0.23, 0.87)	0	20	bad
σ	1.16 (0.83, 1.38)	0.78 (0.55, 0.97)	1.58 (1.15, 1.88)	1	20	bad
LHC_p	-2.03 (-2.51, -1.39)	-2.98 (-3.64, -2.14)	-1.17 (-1.56, -0.73)	-1.64	20	bad

Table A.9: Averages of the ten posterior means (μ, σ and LHC_p) and averages of lower and upper HPD limits. In the table, they are called Posterior mean, Lower HPDI respective Upper HPDI. The parentheses display the lowest and highest values of the ten means. K= number of species. Inverse-gamma prior on σ^2 and $\mu_\mu = -4$.

	Posterior mean	Lower HPDI	Upper HPDI	True value	K	Quality
μ	-0.21 (-0.81, 0.38)	-1.28 (-1.85, -0.91)	0.79 (-0.06, 1.56)	0	5	good
σ	1.11 (0.84, 1.54)	0.59 (0.46, 0.81)	1.78 (1.33, 2.46)	1	5	good
LHC_p	-2.05 (-2.94, -1.75)	-3.8 (-5.45, -3.21)	-0.65 (-1.2, -0.05)	-1.64	5	good
μ	-0.2 (-1.08, 0.47)	-1.04 (-1.88, -0.44)	0.61 (-0.29, 1.33)	0	7	good
σ	1.08 (0.86, 1.36)	0.63 (0.5, 0.81)	1.63 (1.3, 2.08)	1	7	good
LHC_p	-1.97 (-2.78, -1.42)	-3.34 (-4.33, -2.57)	-0.82 (-1.73, -0.2)	-1.64	7	good
μ	-0.12 (-0.95, 0.44)	-0.78 (-1.55, -0.26)	0.51 (-0.38, 1.18)	0	10	good
σ	1 (0.85, 1.19)	0.63 (0.53, 0.75)	1.44 (1.22, 1.7)	1	10	good
LHC_p	-1.77 (-2.44, -1.27)	-2.83 (-3.47, -2.21)	-0.87 (-1.64, -0.31)	-1.64	10	good
μ	-0.09 (-0.4, 0.64)	-0.68 (-1.17, 0.03)	0.5 (0.15, 1.29)	0	13	good
σ	1.06 (0.85, 1.42)	0.71 (0.57, 0.94)	1.47 (1.18, 1.96)	1	13	good
LHC_p	-1.84 (-2.74, -1.17)	-2.79 (-4.05, -1.99)	-0.98 (-1.64, -0.26)	-1.64	13	good
μ	-0.07 (-0.29, 0.29)	-0.55 (-0.87, -0.16)	0.4 (0.09, 0.77)	0	20	good
σ	1.07 (0.89, 1.39)	0.77 (0.64, 1)	1.4 (1.17, 1.81)	1	20	good
LHC_p	-1.82 (-2.52, -1.42)	-2.58 (-3.52, -2.09)	-1.14 (-1.64, -0.76)	-1.64	20	good
μ	-0.38 (-1.11, -0.09)	-1.58 (-2.29, -1.26)	0.75 (0.1, 1.08)	0	5	bad
σ	1.13 (0.96, 1.34)	0.56 (0.48, 0.68)	1.85 (1.6, 2.19)	1	5	bad
LHC_p	-2.24 (-3.18, -1.8)	-4.16 (-5.15, -3.44)	-0.65 (-1.58, -0.27)	-1.64	5	bad
μ	-0.37 (-0.79, -0.07)	-1.33 (-1.79, -0.94)	0.55 (0.16, 0.85)	0	7	bad
σ	1.08 (0.88, 1.32)	0.58 (0.48, 0.7)	1.69 (1.38, 2.08)	1	7	bad
LHC_p	-2.14 (-2.71, -1.66)	-3.67 (-4.39, -2.94)	-0.83 (-1.4, -0.56)	-1.64	7	bad
μ	-0.37 (-0.61, 0.04)	-1.15 (-1.51, -0.68)	0.38 (0.12, 0.7)	0	10	bad
σ	1.04 (0.82, 1.34)	0.61 (0.48, 0.8)	1.56 (1.22, 1.97)	1	10	bad
LHC_p	-2.09 (-2.72, -1.38)	-3.31 (-4.25, -2.46)	-1.02 (-1.38, -0.44)	-1.64	10	bad
μ	-0.23 (-0.6, 0.15)	-0.94 (-1.39, -0.59)	0.47 (0.18, 0.88)	0	13	bad
σ	1.11 (0.81, 1.41)	0.69 (0.5, 0.9)	1.61 (1.18, 2)	1	13	bad
LHC_p	-2.07 (-2.79, -1.4)	-3.2 (-4.17, -2.3)	-1.06 (-1.65, -0.63)	-1.64	13	bad
μ	-0.13 (-0.32, 0.27)	-0.7 (-0.92, -0.3)	0.45 (0.23, 0.87)	0	20	bad
σ	1.16 (0.83, 1.38)	0.78 (0.55, 0.97)	1.58 (1.15, 1.88)	1	20	bad
LHC_p	-2.03 (-2.51, -1.39)	-2.98 (-3.64, -2.14)	-1.17 (-1.56, -0.73)	-1.64	20	bad

Appendix B Bayes Optimal Hazardous Concentration and Bias Averaged over 10 iterations

Table B.1: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter $\alpha = 1$, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Uniform prior on σ and hyperparameter $\mu_\mu = 0$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.886	-2.217	-2.412	-1.941	-2.398	-2.629
Bias	-0.241	-0.572	-0.767	-0.297	-0.753	-0.984
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.957	-2.207	-2.350	-1.928	-2.259	-2.437
Bias	-0.312	-0.562	-0.706	-0.283	-0.614	-0.792
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.787	-1.960	-2.061	-1.874	-2.116	-2.248
Bias	-0.142	-0.315	-0.416	-0.229	-0.472	-0.603
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.890	-2.027	-2.104	-1.973	-2.153	-2.250
Bias	-0.245	-0.383	-0.460	-0.328	-0.508	-0.605
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.915	-2.002	-2.049	-2.093	-2.215	-2.280
Bias	-0.271	-0.357	-0.404	-0.448	-0.570	-0.635

Table B.2: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Uniform prior on σ and hyperparameter $\mu_\mu = 4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.646	-1.898	-2.042	-1.640	-1.994	-2.177
Bias	-0.001	-0.253	-0.397	0.005	-0.350	-0.532
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.779	-1.975	-2.087	-1.744	-2.021	-2.172
Bias	-0.134	-0.330	-0.442	-0.099	-0.376	-0.527
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.673	-1.817	-1.899	-1.758	-1.965	-2.077
Bias	-0.028	-0.172	-0.254	-0.113	-0.320	-0.432
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.787	-1.908	-1.975	-1.847	-2.006	-2.091
Bias	-0.142	-0.263	-0.331	-0.202	-0.361	-0.446
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.850	-1.930	-1.973	-2.001	-2.111	-2.170
Bias	-0.205	-0.285	-0.328	-0.356	-0.466	-0.525

Table B.3: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Uniform prior on σ and hyperparameter $\mu_\mu = -4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-2.235	-2.687	-2.930	-2.348	-2.942	-3.215
Bias	-0.590	-1.042	-1.285	-0.703	-1.297	-1.570
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-2.189	-2.511	-2.693	-2.205	-2.626	-2.845
Bias	-0.544	-0.866	-1.048	-0.561	-0.981	-1.200
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.927	-2.142	-2.267	-2.060	-2.352	-2.510
Bias	-0.283	-0.497	-0.622	-0.415	-0.707	-0.865
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-2.006	-2.166	-2.256	-2.121	-2.332	-2.446
Bias	-0.361	-0.521	-0.611	-0.476	-0.687	-0.801
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.993	-2.090	-2.142	-2.196	-2.332	-2.404
Bias	-0.348	-0.445	-0.497	-0.552	-0.687	-0.759

Table B.4: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Gamma prior on σ^2 and hyperparameter $\mu_\mu = 0$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.704	-1.901	-2.014	-1.857	-2.100	-2.232
Bias	-0.059	-0.256	-0.369	-0.212	-0.455	-0.587
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.761	-1.906	-1.986	-1.882	-2.066	-2.167
Bias	-0.116	-0.261	-0.341	-0.237	-0.421	-0.522
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.658	-1.762	-1.820	-1.869	-2.006	-2.079
Bias	-0.013	-0.117	-0.175	-0.225	-0.361	-0.434
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.725	-1.812	-1.859	-1.832	-1.944	-2.003
Bias	-0.080	-0.167	-0.214	-0.188	-0.299	-0.358
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.788	-1.848	-1.880	-1.899	-1.980	-2.022
Bias	-0.143	-0.203	-0.235	-0.254	-0.335	-0.377

Table B.5: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Gamma prior on σ^2 and hyperparameter $\mu_\mu = 4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.514	-1.675	-1.765	-1.601	-1.803	-1.910
Bias	0.131	-0.030	-0.120	0.044	-0.158	-0.265
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.623	-1.747	-1.816	-1.705	-1.865	-1.950
Bias	0.022	-0.103	-0.171	-0.060	-0.220	-0.306
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.564	-1.656	-1.706	-1.752	-1.876	-1.942
Bias	0.081	-0.011	-0.062	-0.107	-0.231	-0.297
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.642	-1.721	-1.764	-1.733	-1.835	-1.888
Bias	0.003	-0.076	-0.119	-0.088	-0.190	-0.243
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.731	-1.787	-1.817	-1.822	-1.898	-1.938
Bias	-0.086	-0.143	-0.173	-0.177	-0.253	-0.293

Table B.6: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Gamma prior on σ^2 and hyperparameter $\mu_\mu = -4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.931	-2.181	-2.327	-2.157	-2.456	-2.620
Bias	-0.286	-0.537	-0.682	-0.513	-0.811	-0.975
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.923	-2.099	-2.200	-2.086	-2.303	-2.421
Bias	-0.278	-0.455	-0.556	-0.441	-0.658	-0.776
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.764	-1.884	-1.951	-2.010	-2.167	-2.253
Bias	-0.119	-0.239	-0.306	-0.365	-0.523	-0.608
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.816	-1.913	-1.965	-1.948	-2.073	-2.140
Bias	-0.172	-0.268	-0.320	-0.303	-0.428	-0.495
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.851	-1.917	-1.952	-1.982	-2.069	-2.115
Bias	-0.206	-0.272	-0.307	-0.337	-0.424	-0.470

Table B.7: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter α = 1, 2 respective 2.5. The hazardous concentration under perfect information $\text{LH}\hat{\text{C}}_p^\alpha$ is -1.64. Inverse-gamma prior on σ^2 and hyperparameter $\mu_\mu = 0$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.854	-2.224	-2.550	-2.008	-2.406	-2.681
Bias	-0.209	-0.579	-0.905	-0.363	-0.761	-1.036
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.848	-2.074	-2.229	-1.985	-2.257	-2.437
Bias	-0.204	-0.429	-0.585	-0.340	-0.613	-0.793
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.697	-1.831	-1.912	-1.991	-2.171	-2.285
Bias	-0.052	-0.186	-0.267	-0.346	-0.526	-0.640
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.774	-1.886	-1.951	-1.984	-2.135	-2.223
Bias	-0.130	-0.241	-0.307	-0.339	-0.491	-0.579
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.785	-1.855	-1.893	-1.977	-2.081	-2.139
Bias	-0.140	-0.210	-0.249	-0.332	-0.436	-0.494

Table B.8: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter $\alpha = 1$, 2 respective 2.5. The hazardous concentration under perfect information, $\text{LH}\hat{\text{C}}_p^\alpha$, is -1.64. Inverse-gamma prior on σ^2 and hyperparameter $\mu_\mu = 4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.615	-1.870	-2.053	-1.702	-2.005	-2.211
Bias	0.030	-0.225	-0.408	-0.057	-0.360	-0.567
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.681	-1.856	-1.965	-1.773	-1.990	-2.125
Bias	-0.036	-0.211	-0.320	-0.128	-0.345	-0.480
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.594	-1.709	-1.777	-1.853	-2.007	-2.098
Bias	0.051	-0.065	-0.132	-0.208	-0.363	-0.453
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.685	-1.782	-1.837	-1.865	-1.998	-2.074
Bias	-0.040	-0.137	-0.192	-0.221	-0.354	-0.430
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.727	-1.793	-1.829	-1.891	-1.986	-2.038
Bias	-0.083	-0.149	-0.185	-0.246	-0.341	-0.393

Table B.9: Bayes optimal hazardous concentration $\text{BL}\hat{\text{H}}\text{C}_p^\alpha$ and bias for loss parameter $\alpha = 1$, 2 respective 2.5. The hazardous concentration under perfect information, $\text{LH}\hat{\text{C}}_p^\alpha$, is -1.64. Inverse-gamma prior on σ^2 and hyperparameter $\mu_\mu = -4$ averaged over 10 iterations of the simulation study.

	Good data on 5 species			Bad data on 5 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-2.229	-3.029	-3.774	-2.452	-3.284	-4.078
Bias	-0.584	-1.384	-2.129	-0.807	-1.639	-2.433
	Good data on 7 species			Bad data on 7 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-2.063	-2.399	-2.673	-2.265	-2.710	-3.109
Bias	-0.418	-0.754	-1.028	-0.620	-1.065	-1.464
	Good data on 10 species			Bad data on 10 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.821	-1.990	-2.097	-2.160	-2.389	-2.543
Bias	-0.176	-0.345	-0.453	-0.515	-0.744	-0.899
	Good data on 13 species			Bad data on 13 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.879	-2.011	-2.091	-2.124	-2.308	-2.425
Bias	-0.234	-0.366	-0.446	-0.479	-0.663	-0.780
	Good data on 20 species			Bad data on 20 species		
	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 2.5$
$\text{BL}\hat{\text{H}}\text{C}_p$	-1.850	-1.927	-1.969	-2.069	-2.185	-2.249
Bias	-0.205	-0.282	-0.324	-0.425	-0.540	-0.604

Bachelor's Theses in Mathematical Sciences 2017:K13
ISSN 1654-6229

LUNFMS-4022-2017

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/