Ecotoxicology and Environmental Safety 73 (2010) 123–131

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv

A Bayesian approach for determining the no effect concentration and hazardous concentration in ecotoxicology

David R. Fox *

Australian Centre for Environmetrics, University of Melbourne, PO Box 4102, Parkville, Victoria, Australia

ARTICLE INFO

Article history:
Received 27 May 2009
Received in revised form
6 September 2009
Accepted 12 September 2009

Keywords:
Species sensitivity distribution
NOECs
Hypothesis testing
Ecosystem protection

ABSTRACT

This paper describes a Bayesian modeling approach to the estimation of the no effect concentration (NEC) and the hazardous concentration (HC_x) as an alternative to conventional methods based on NOECs – the no *observed* effect concentration. The advantage of the proposed method is that it combines a plausible model for dose–response data with prior information or belief about the model's parameters to generate posterior distributions for the parameters – one of those being the NEC. The posterior distribution can be used to derive point and interval estimates for the NEC as well as providing uncertainty bounds when used in the development of a species sensitivity distribution (SSD). This latter feature is particularly attractive and overcomes a recognized deficiency of the NOEC-based approach. Examples using previously published data sets are provided which illustrate how the NEC/HC_x estimation problem is re-cast and solved in this Bayesian framework.

1. Introduction

The species sensitivity distribution (SSD) is a cornerstone of modern ecotoxicology and provides a basis for establishing guidelines, trigger values, and limits on concentrations of hazardous chemicals in animals and the receiving environment. In the context of water quality, use of the SSD is underpinned by the well-validated belief that, at the community/assemblage level, aquatic species generally have different although predictable responses to increasing concentrations of physical–chemical toxicants. The familiar dose–response curves generated by laboratory experiments are used to estimate a variety of measures such as the LC_{50} (the concentration which is lethal to 50% of some defined population), the NEC (the maximum concentration which causes no adverse effect in a target organism), and EC_x (the concentration which affects *x*% of organisms in a dose–response experiment). Classical tools of statistical inference such as *t*-tests, ANOVA, and multiple comparison techniques are also widely used to estimate related statistical measures such as the no observed effect concentration (NOEC) and the lowest observed effect concentration (LOEC). One of the difficulties with conventional practice is that many of these statistics are being used interchangeably which, as argued by Fox (2009) not only creates problems of interpretation but obfuscates what is really being protected. A number of authors have denounced the ad-hoc procedures for setting safe environmental concentrations that are based on NOECs and LOECs and have argued for a more rigorous, model-based approach (Fox, 2009; Jager et al., 2006; Kooijman, 2006).

A vast literature has accumulated over the last 20 years in which the theoretical, computational, statistical, and socio-economic aspects associated with the identification of 'safe' concentrations have been discussed. Readers requiring more detailed background information on the use of SSDs and their application in ecotoxicology and within a regulatory framework will find the collection of papers in Posthuma et al. (2002) a useful starting point. A good review of the statistical issues associated with ecotoxicological risk assessment is provided by Van der Hoeven (2004) while more recently Fox (2006, 2008) reviewed the use of statistical methods in ecological risk assessment more generally.

The general problem addressed by this paper is as follows: how does one set a realistic threshold concentration on some contaminant or toxicant such that some arbitrary high fraction of all species will be protected provided environmental concentrations do not exceed the threshold? This is indeed not a new problem and has been studied extensively by many researchers (Wagner and Løkke, 1991; Kooijman et al., 1996; Aldenberg and Jaworska, 2000; Shao, 2000). Despite a plethora of models and incremental refinements, the numerous concerns with the NOEC-based procedures (Fox, 1999, 2006; Newman et al., 2000; Isnard et al., 2001; Pires et al., 2002; Verdonck et al., 2003) that underpin current practice in Australia, New Zealand, the United States, the Netherlands, and Denmark have not been extinguished and as recently noted by Newman (2008), the current ecotoxicological landscape is dominated by classical (i.e. frequentist) statistical methods.

Although a number of Bayesian papers have recently appeared in the ecotoxicological literature (Aldenberg and Jaworska, 2000;

* Fax: +61 3 8344 4616.
E-mail address: david.fox@unimelb.edu.au

0147-6513/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ecoenv.2009.09.012

Verdonck et al., 2001; Grist et al., 2006; Billoir et al., 2008; Hickey et al., 2008) the framework by and large remains outside the realm of conventional ecotoxicological practice. Readers wanting to learn more about Bayesian statistics should consult any of a number of good introductory texts such as Lee (2004) or McCarthy (2007). The statistical profession spent many years and devoted many journal pages to the debate over the legitimacy of the Bayesian paradigm. Early objectors strenuously refuted the notion of a 'prior' distribution, the incorporation of subjective assessment, and treatment of parameters as random quantities. Thankfully the old divisions between the 'Frequentist' and 'Bayesian' schools of thought have largely given way to a more pragmatic approach that accommodates multiple modes of statistical inference with the choice increasingly based on the notion of 'fit-for-purpose' rather than ideological or pedagogical constructs. Furthermore, the advent of high-powered desk-top computers and associated software such as WinBUGS (Lunn et al., 2000) has removed any lingering impediments to the Bayesian analysis of complex, real-world problems.

In light of these developments it is timely to revisit the role and place of Bayesian statistics in the context of determining hazardous concentrations for ecosystem protection. Some of the more serious limitations associated with conventional NOEC-based analyses center on the following: (i) the unknown (and perhaps unknowable) underlying distributional form for NOECs; (ii) the statistical method by which a NOEC is determined; (iii) the inability to represent uncertainty in the estimated NOEC; and (iv) the non-random selection of a small number of species (van der Hoeven, 1997; Crane and Newman, 2000). As will be demonstrated later in this paper, these issues are addressed through the use of posterior distributions to represent and describe uncertainty in the estimated no effect concentration (NEC). Uncertainty in the collection of NECs can be incorporated into a final estimate of a hazardous concentration to which a statement of 'confidence' (or the Bayesian analog *credibility*) can be attached.

Our starting point is a flexible and realistic model for the raw data generated by a dose–response experiment – this is consistent with the recommendations of Kooijman et al. (1996), Van der Hoeven (1997) and Van der Hoeven (2004). In the remainder of the paper we describe the procedure and illustrate its implementation with the use of previously published data sets.

2. A Bayesian model for the NEC

A number of models have been proposed to describe the dose–response relationship in ecotoxicological studies. We have adopted the model used by Pires et al. (2002) which relates the response (Y) to concentration (x) such that Y is constant from x=0 up to a threshold, γ and thereafter exhibits an exponential decay. It is important to note that the incorporation of γ in our model does not presuppose the existence of a threshold – it simply allows for one to be estimated if that is what the data suggests. Pires et al. (2002) assumed Y was discrete (numbers of individuals) and hence used a Poisson probability model to describe stochastic variation in Y. We relax this assumption and allow Y to be either discrete or continuous (for example, percent mortality) having arbitrary probability function $g_Y(\cdot)$. The complete model is defined by Eqs. (1) and (2):

$$Y_i \overset{d}{\sim} g_Y(\cdot) \tag{1}$$

$$E[Y_i|x_i] = \mu_i = \alpha \exp[-\beta(x_i - \gamma)I(x_i - \gamma)] \tag{2}$$

with

$$I(z) = \begin{cases} 1, & z > 0 \\ 0, & z \le 0 \end{cases}$$

$E[Y_i|x_i]$ denotes the mathematical expectation of Y_i conditional on a given concentration x_i; and the notation $\overset{d}{\sim}$ in Eq. (1) meaning 'is distributed as'.

Taken together, Eqs. (1) and (2) assume the response at the ith concentration, x_i follows some distribution $g_y(\cdot)$ having mean μ_i. The form of Eq. (2) generates a response curve as shown in Fig. 1.

The parameters α, β, and γ in Eq. (2) have intuitive interpretations:α is a 'basal' response – that is, the response at zero/low-dose concentrations; β controls the rate of decay in the response; and γ is the NEC. Given data $\{x_i,y_i\}$ our objective is to estimate the parameters α, β, γ. A conventional regression-based approach would do this by (i) assuming Y_i to be normally distributed with mean μ_i and some constant variance σ_ε^2, and (ii) use a least-squares (LS) or maximum likelihood (ML) criterion to find the 'best-fitting' parameter estimates.

The Bayesian formulation similarly requires specification of $g_y(\cdot)$ but in addition assumes the parameter vector $\underline{\Theta} = \{\alpha, \beta, \gamma\}$ is a random quantity to which is assigned a *prior* distribution, $p(\underline{\Theta})$.

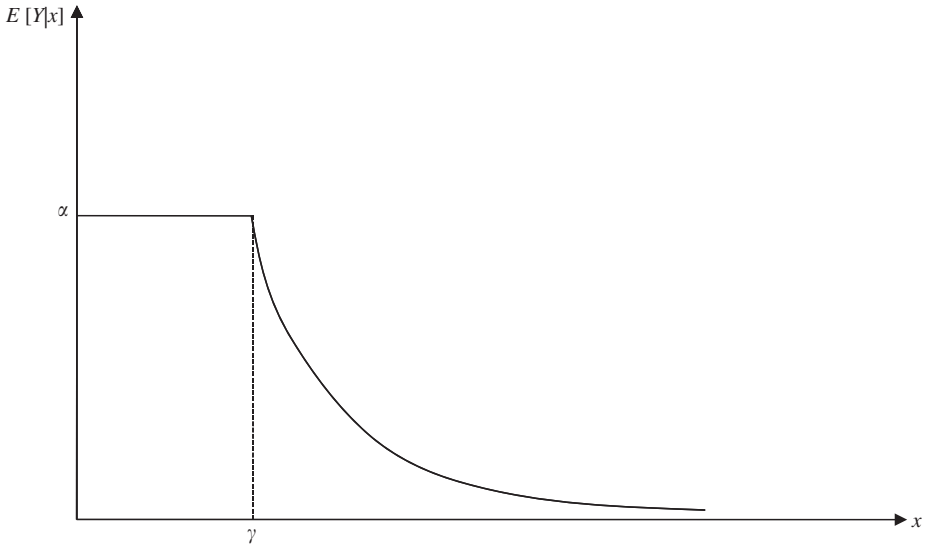

Fig. 1. Dose–response model given by Eq. (2).

The Bayesian method provides an updated version (the *posterior*) of $p(\underline{\Theta})$ through the use of Bayes' formula. In essence this computes $P[\underline{\Theta}|\text{data}] \propto \int l[\text{data}|\underline{\Theta}]\, P(\underline{\Theta})\, d\underline{\Theta}$ where $l[\text{data}|\underline{\Theta}]$ is the likelihood of the data given a parameter vector $\underline{\Theta}$. The specification of prior distributions affords the analyst with an opportunity to inject formally elicited expert opinion about each of the model parameters or as a mechanism for introducing personal belief. While the choice is arbitrary, a prior is chosen to be either informative or vague. *Informative priors* are used when our understanding or expectation about the range of likely values for a parameter is well defined. For example, the choice of a normal distribution centered on what is believed to be the most probable value and having a small variance would constitute an informative prior. Conversely, a *non-informative* or *vague* prior is typically one that assigns equal weight to all values (such as a uniform distribution) or one that has a large variance. Other possibilities exist, such as Jeffrey's and improper priors but will not be considered here.

The computation of the integral associated with the derivation of the posterior distribution is invariably high dimensional and complex. It is generally easier (and often-times the only option available) to use numerical methods such as Markov chain Monte Carlo (MCMC) methods that are suited to implementation on a desk-top computer. The WinBUGS program is a free software tool that uses a particular implementation of MCMC estimation known as Gibbs sampling (BUGS is an acronym for Bayesian inference Using Gibbs Sampling). WinBUGS was developed at the MRC Biostatistics Unit, Cambridge University and can be downloaded from http://www.mrc-bsu.cam.ac.uk/bugs/. We have used the 'open source' version of WinBUGS available at the University of Helsinki's website http://mathstat.helsinki.fi/openbugs/. WinBUGS has its own, reasonably intuitive programming language although novices may prefer (at least initially) to use the so-called 'doodle bugs' editor which allows the user to represent his or her model in a graphical format. The WinBUGS code is then automatically generated from the graphical representation. The Bayesian model development and parameter estimation procedure is illustrated in the following section.

3. Example – estimation of a NEC for *Daphnia magna*

Biesinger et al. (1982) reported on a study into the chronic toxicity of mercury (Hg) to *D. magna*. The compounds of mercury tested were mercuric chloride ($HgCl_2$), methyl mercuric chloride (MMC), and phenyl mercuric acetate (PMA). A range of Hg concentrations was prepared and for each toxicant the number, y_i out of an initial sample of n_i individuals surviving after 21 days was recorded against the ith concentration. A summary of the data is shown in Table 1.

The response variable here is discrete and while the Poisson distribution used by Pires et al. (2002) could be adopted, this is perhaps not the most sensible choice given that some of the n_i were small and y_i is constrained to lie in the range $[0,n_i]$. A more natural candidate for $g_y(\cdot)$ in this case is the binomial distribution (Eq. (3)):

$$P[Y_i = y_i] = \binom{n_i}{y_i} \theta_i^{y_i} (1-\theta_i)^{n_i - y_i}, \quad 0 < \theta_i < 1, 0 \le y_i \le n_i \tag{3}$$

Now, $E[Y_i] = \mu_i = n_i\theta_i$ and so Eq. (2) relates the quantity $(n_i\theta_i)$ to x_i. We can also consider *proportion* $P_i = Y_i/n_i$ surviving at concentration x_i. The expectation of P_i is θ_i the *true* proportion, and thus Eq. (2) becomes

$$E[P_i] = \theta_i = \alpha \exp[-\beta(x_i - \gamma) I(x_i - \gamma)] \tag{4}$$

As noted above, there is flexibility in the choice of prior distributions. To reflect prior ignorance about the likely parameter values, we have chosen gamma priors with a large variance for each of the parameters α, β, γ in Eq. (4). The graphical representation for this model is shown in Fig. 2 and the corresponding WinBUGS code in Appendix A.

Table 1
Daphnia magna data taken from Biesinger et al. (1982) showing numbers surviving y_i after 21 days out of an initial n_i at various concentrations of three compounds of mercury.

Toxicant	Hg concentration (μg/L)	n_i	y_i
Mercuric chloride	0.05*	180	171
	0.43	80	65
	0.91	80	73
	1.82	180	160
	3.53	180	108
	5.31	20	0
Methyl mercuric chloride	0.05*	80	77
	0.17	80	79
	0.28	80	76
	0.52	80	76
	0.87	80	71
	1.14	80	0
Phenyl mercuric acetate	0.05*	35	33
	0.35	15	15
	0.54	35	30
	1.12	35	33
	1.90	35	26
	3.00	35	1

* Original concentration reported as < 0.1.

The model was run 100,000 times (after an initial 'burn-in' period of 10,000 iterations) for each toxicant with model outputs recorded at every tenth iteration (a process known as 'thinning' which reduces autocorrelation in the recorded data). Summary statistics for the resulting 10,000 MCMC iterations were gathered and are presented in Table 2. Additional diagnostics tools available in the WinBUGS software were used to check for convergence.

An example of the posterior distributions obtained from WinBUGS for phenyl mercuric acetate is shown in Fig. 3 and the fitted model in Fig. 4.

In a Bayesian analysis, it is customary to use the median of the posterior distribution as a point estimate of a parameter. We thus obtain the following point estimates for the NEC: 3.27 μg Hg/L for mercuric chloride; 0.863 μg Hg/L for methyl mercuric chloride; and 1.792 μg Hg/L for phenyl mercuric acetate. The Bayesian analog of the Frequentists' confidence interval is the Bayesian credibility interval. A ξ100% credibility interval $(0 < \xi < 1)$ is an interval containing ξ100% of the posterior distribution. Although not uniquely defined, we use the 2.5 and 97.5 percentiles of the posterior distribution to define a 95% credibility interval (Table 2 and Fig. 4).

The attractive feature of the posterior distribution is that it summarizes the combined uncertainty in the estimated NEC arising from both stochastic variation and epistemic uncertainty (i.e. incomplete knowledge). This uncertainty can be incorporated into a more comprehensive Bayesian framework for multiple species from which an ecosystem HC_x (and its uncertainty) can be determined. The approach is explained in the following section with reference to recently published data on the toxicity of pond waters at a uranium mine in Northern Australia.

4. Example – estimation of the HC_x for pond water at a uranium mine

The Ranger mine located 230 km east of Darwin, Australia is one of the world's largest open-pit uranium mines. The mine site is located within the environmentally sensitive Kakadu National

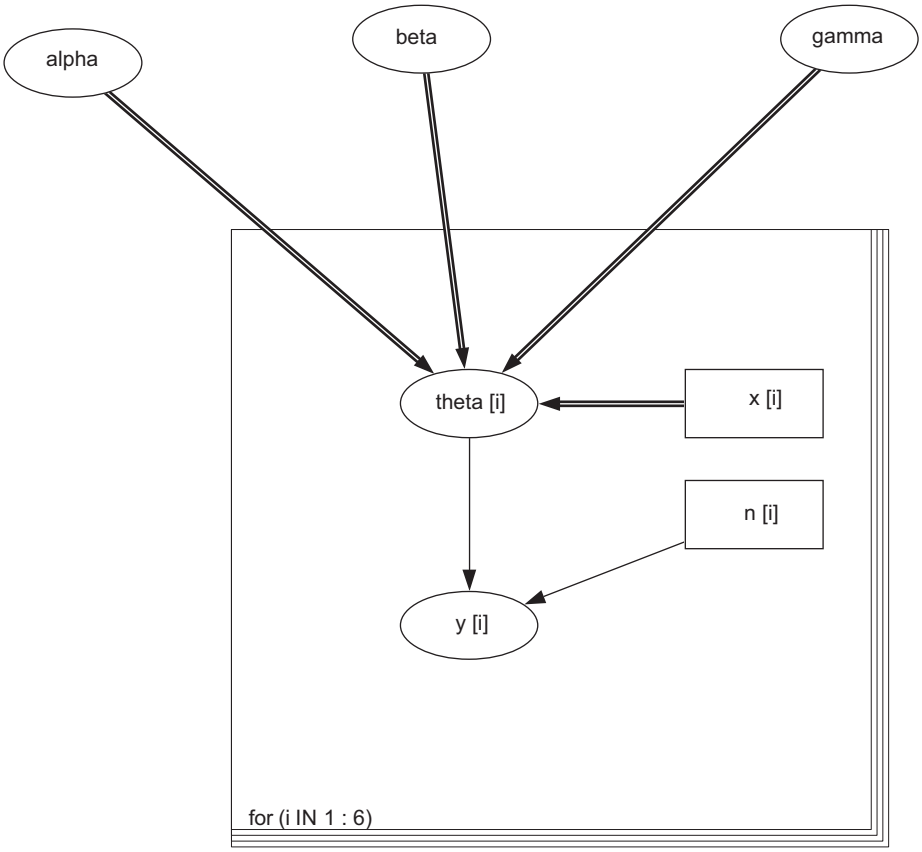

Fig. 2. Graphical representation of dose–response model for *Daphnia* example.

Table 2
Parameter estimates and selected percentiles of the posterior distribution for model given by Eq. (4).

Toxicant	Parameter	Mean	Standard deviation	Percentiles		
				$P_{2.5}$	P_{50}	$P_{97.5}$
Mercuric chloride	Alpha	0.90	0.01	0.87	0.90	0.92
	Beta	1.94	1.13	0.52	1.56	4.54
	Gamma	3.23	0.23	2.66	3.27	3.45
Methyl mercuric chloride	Alpha	0.95	0.01	0.92	0.95	0.97
	Beta	1188	3139	17	107	12,350
	Gamma	0.95	0.08	0.01	0.86	1.11
Phenyl mercuric acetate	Alpha	0.93	0.03	0.87	0.93	0.97
	Beta	2.58	1.33	1.20	2.26	5.97
	Gamma	1.82	0.20	1.57	1.79	2.59

Park. On March 9, 2007 tropical cyclone 'George' impacted the Pilbara mining region in the far north-west of Western Australia. The Ranger mine was also affected by this system with nearly 850 mm of rain falling in the 7 days to March 4, including 750 mm in one 72-h period causing flooding of the mine and the release of water into the Magela Creek. In response to this event, the Australian government undertook toxicity studies as part of an examination of options for reducing the volume of pond water stored at the mine site. One of these options involved the direct release of untreated pond water from Retention Pond 2 (RP2) to Magela Creek (Hogan et al., 2008). The toxicity studies reported by Hogan et al. (2008) were based on an analysis of the SSD obtained from five test species' IC_{10} (i.e. the concentration that results in a 10% inhibition of response relative to the control response) or equivalent data using the BurrliOZ software (www.cmis.csiro.au/envir/burrlioz/). With respect to RP2 water, Hogan et al. (2008) concluded that a dilution of 0.33% (approximately 1:300) would protect 99% of all species in Magela Creek. We have undertaken a Bayesian analysis of the data given in Appendix 5 of Hogan et al. (2008) using the methods outlined in this paper. Our results suggest that the Hogan et al. (2008) concentration is overly liberal (i.e. too high) and as such the *actual* fraction of species protected may be significantly *less* than the claimed 99%. Details of our analysis are provided below.

Hogan et al. (2008) used the following five test species: *Chlorella* (an alga); *Lemna aequinoctialis* (duckweed); *Hydra viridissima* (green hydra); *Moinodaphnia macleayi* (water flea); and *Mogurnda mogurnda* (purple-spotted gudgeon). The respective endpoints were as follows: 72 h cell division rate; 96 h plant growth; 96 h population growth; brood reproduction; and 96-h

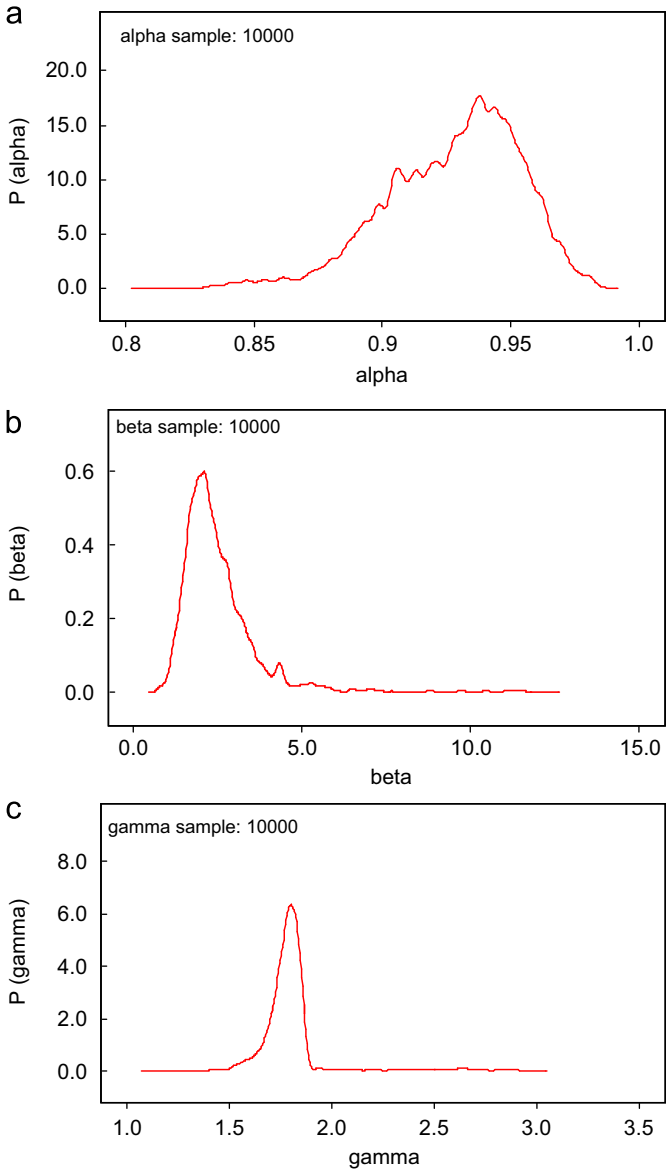

Fig. 3. Sample posterior distributions for model parameters for phenyl mercuric acetate: (a) distribution of α parameter; (b) distribution of β parameter; and (c) distribution of γ parameter.

survival. Three replicates of each species were exposed to either 6 or 7 dilutions of RP2 water except for the *M. macleayi* test which was based on 10 replicates. The complete set of data is available in Appendices 5.1–5.5 of Hogan et al. (2008) which can be downloaded from http://www.environment.gov.au/ssd/publications/ssr/197.html. An important feature of this data set is the results for *M. mogurnda* in which none of the animals showed any response at any of the 6 dilutions in the range 0–100%. The zero variance for these data is problematic for statistical distribution-fitting methods. To overcome this difficulty, Hogan et al. (2008) assumed a dilution of 50% as 'a conservative toxicity estimate for *M. mogurnda*'. The lack of any response for *M. mogurnda* data also means that our model will result in an estimate of γ (the NEC) close to zero which is unrealistic. To maintain consistency with Hogan et al. (2008) we have also adopted a 50% value for the *M. mogurnda* NEC. Furthermore, and as pointed out by one of the reviewers of an earlier draft of this paper, the response variable for *Moinodaphnia* is discrete (number of neonates) thus calling into question the appropriateness of the normal probability model. In response, we note that: (i) the normal

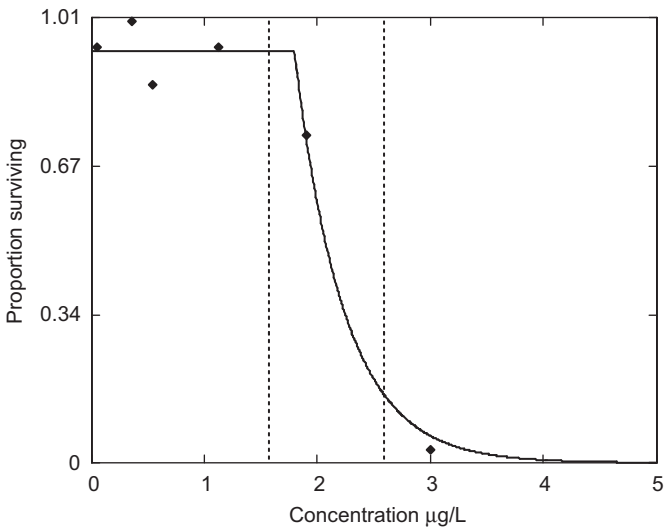

Fig. 4. Sample data (solid diamonds) and fitted model (solid line) for phenyl mercuric acetate. Vertical dashed lines indicate 95% credibility interval for NEC.

distribution was implicitly used by Hogan et al. (2008) and (ii) re-running the model assuming a Poisson response made little difference (in this case) to the estimated NEC.

5. Estimation of the no effect concentration

The response variables (Y_i) for each of the four species used by Hogan et al. (2008) are assumed to be normally distributed as $Y_i \sim N(\mu_i, \sigma_i^2)$ with μ_I given by Eq. (2) and non-informative gamma priors for the precision terms $1/\sigma_i^2$ The prior distributions assigned to the model terms in Eq. (2) are listed in Table 3. We have chosen informative priors that reflect an 'educated' guess (informed by an inspection of a plot of the data) as to the likely position of the NEC. In essence, this corresponds to what is known as 'empirical Bayes'.

The WinBUGS code for this model is given in Appendix A. The results of 50,000 MCMC simulations collated after an initial 'burn-in' of 10,000 runs are shown in Table 4. The posterior cumulative distribution functions for the fitted NECs are shown in Fig. 5. These distributions reflect the variation in estimated NECs that arise from the variation in responses from individual test organisms.

It is evident from Fig. 5 that the probability of an effect at a concentration of 22.5(%) is 50% for *Lemna* and 100% for the other three species.

A comparison of our estimated NECs with the NOECs and IC_{10} values reported by Hogan et al. (2008) is given in Table 5.

We see from Table 5 that there is reasonable agreement between all three estimates for *Chlorella* and *Hydra* however there are significant discrepancies among the estimates for *Lemna* and *Moinodaphnia*. In particular, our estimated NEC for *Lemna* is an order of magnitude greater than the IC_{10} reported by Hogan et al. (2008). The difference is an artifact of the two very different methods used to obtain the respective estimates. Our NEC is an estimate of a parameter in a dose–response model. According to Hogan et al. (2008), the IC_{10} "involves fitting straight lines between each successive concentration … then interpolating the relevant 'effect' or 'inhibition' size of interest". Thus the IC_{10} is an empirical estimate that neither accommodates data variability nor is predicated on any plausible model of the response-generating mechanism. The impact of these differences on the fitted SSD and estimated HC_1 is examined next.

Table 3
Prior distributions used for the Ranger mine RP 2 waters example.

Species				
Parameter	*Chlorella*	*Lemna*	*Hydra*	*Moinodaphnia*
α	$N(0,10^6)$	$N(0,10^6)$	$N(0,10^6)$	$N(0,10^6)$
β	$N(0,1/200)\|(0,\infty)$	$N(0,1/200)\|(0,\infty)$	$N(0,1/200)\|(0,\infty)$	$N(0,1/200)\|(0,\infty)$
γ	LN(2.124,1.708)	LN(3.201,1.582)	LN(1.221,0.437)	LN(−1.278,0.404)

Parameters relate to Eq. (2) in the text. $N(\cdot,\cdot)$ denotes a normal distribution; $N(\cdot,\cdot)|(0,\infty)$ denotes the half-normal distribution; and LN$(\cdot,\cdot)$ denotes the lognormal distribution.

Table 4
Parameter estimates and selected percentiles of the posterior distribution based on 50,000 MCMC simulations for Ranger mine RP 2 waters data.

Parameter	Mean	Standard deviation	Percentiles		
			$P_{2.5}$	P_{50}	$P_{97.5}$
γ_1	8.603	2.466	5.685	8.12	16.66
γ_2	24.95	13.08	4.775	22.61	53.29
γ_3	2.778	0.6317	1.459	2.752	4.15
γ_4	0.193	0.1565	0.01122	0.1498	0.581

Individual NECs identified as gamma parameter for each species. Values represent concentrations of pit-waters (%).

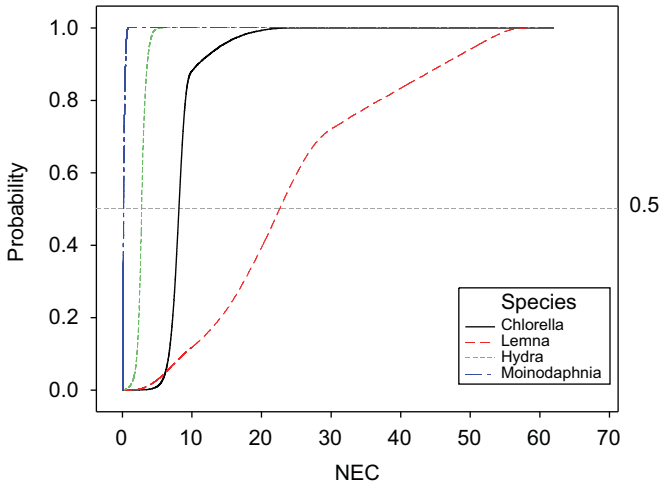

Fig. 5. Posterior distributions for gamma parameters (NEC) for Ranger mine pit water example based on 50,000 MCMC iterations.

Table 5
'No effect estimates' for Ranger mine RP 2 waters data using the following: (a) model-based NEC of this paper; (b) NOECs and IC10 estimates as given by Hogan et al. (2008).

Species	NEC(a)	NOEC(b)	IC10(b)
Chlorella	8.12 (5.68,16.66)	10	7.5 (0,16)
Lemna	22.61 (4.78,53.29)	3	2.1 (0.4,13)
Hydra	2.752 (1.46,4.15)	3	3.5 (0,4.3)
Moinodaphni	0.15 (0.01,0.58)	0.3	0.6 (0.2,0.8)

Figures in parentheses are as follows: (a) 95% credibility interval; and (b) 95% confidence interval.

6. Estimation of the HC_1

In Australia, the recommended procedure for estimating the HC_x is embedded in the software tool known as BurrliOz which is distributed with the national water quality guidelines document (ANZECC/ARMCANZ, 2000). The procedure is a generalization of the method described by Aldenberg and Slob (1993). Using the BurrliOz software with the data in Table 5 we examined the differences in the estimated HC_1. Following Hogan et al. (2008) we used a concentration of 50 (%) for the fifth species (*Mogurnda*) for the NEC, NOEC, and IC_{10}. The resulting HC_1 values are, respectively: 0.0015%; 0.04%; and 0.35% corresponding to dilutions of: 1:66,666; 1:2,500; and 1:285. Hogan et al. (2008) rounded up the last ratio and recommended that a dilution of 1:300 'would be expected to ensure the appropriate level of protection for the downstream aquatic ecosystem'. The large discrepancy among the three estimates of the HC_1 suggests that a 1:300 dilution possibly underestimates the actual dilution required to achieve the desired level of protection. Interestingly, if we follow Aldenberg and Slob (1993) and fit two-parameter log-logistic distributions to the data of Table 5 we get HC_1 estimates of 0.031% using NEC values; 0.048% using NOEC data; and 0.083% using IC_{10} values, corresponding, respectively, to dilutions of: 1:3225; 1:2082; and 1:1204. While no general conclusions can be drawn from this result, the Aldenberg and Slob (1993) method has resulted in a more consistent set of estimates for HC_1. In any event, the conclusions are the same – the IC_{10} data have resulted in a significantly larger 'safe' concentration and the dilution required is greater than 1:300.

Finally, and to conclude this analysis, we interrogate the posterior distributions for the estimated NECs to explore the range of uncertainty that is possible when the HC_1 is estimated from the fitted SSD. To this end, we randomly generated approximately 10,000 sequences of NEC quadruples from the posterior distributions of $\{\hat{\gamma}_1,\hat{\gamma}_2,\hat{\gamma}_3,\hat{\gamma}_4\}$ and, for reasons outlined above, fixed the fifth NEC at 50 (%). A two-parameter logistic distribution was fitted to each realization of NEC parameters from which the HC_1 was obtained. The histogram of the resulting collection of HC_1 estimates is shown in Fig. 6 together with a fitted log-logistic probability model. The range of HC_1 values is from 0.0018% to 0.0716% with a median value of 0.0172%. The discrepancy between this median and the value of 0.031% given above is due to the fact that the latter was the result of fitting a *single* distribution to $n=5$ median NECs and estimating the HC_1 whereas the present result is obtained by fitting distributions to individual realizations of the NECs and calculating the median of the resulting collection of $n=10{,}000$ estimated HC_1 values. The latter approach is considered to be superior since it utilizes information about the distributional properties of the individual NECs and, provided sufficient simulations are performed, will comprehensively explore the joint parameter space for $\{\gamma_1,\gamma_2,\gamma_3,\gamma_4\}$. The median concentration of 0.0172% for this method corresponds to a 1:5813 dilution and, conceptually, is equivalent to a 99:50 trigger value (ANZECC/

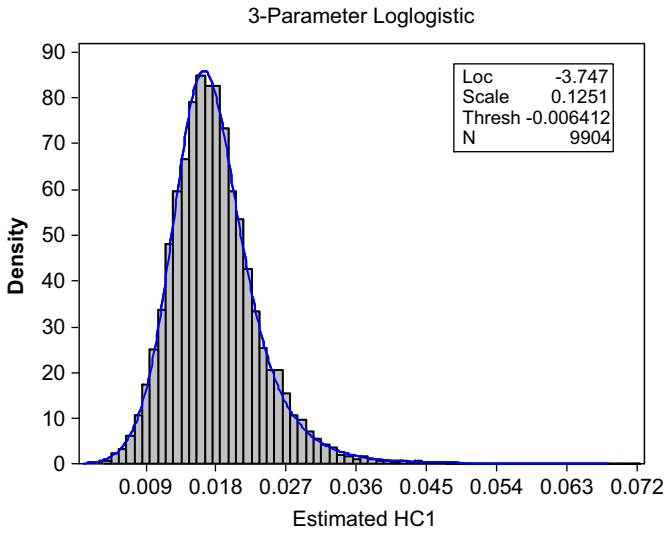

Fig. 6. Histogram of 10,000 estimated HC_1 values for Ranger mine pit water example (rectangles) and fitted 3-parameter lognormal distribution (solid line).

ARMCANZ, 2000). On the basis of this analysis and the previous results we are reasonably confident in asserting that a 1:300 dilution as recommended by Hogan et al. (2008) is too liberal by a factor of about 20.

7. Discussion

Over the past decade deterministic methods of risk assessment such as the *hazard quotient* (HQ) approach have given way to probabilistic methods. Thus for example, in Australia and elsewhere, the use of species sensitivity distributions (SSDs) is the preferred method for establishing concentration thresholds (or 'trigger values') for chemical contaminants in water bodies (ANZECC/ARMCANZ, 2000). While probabilistic ecological risk assessments (PERAs) are underpinned by a more comprehensive treatment of uncertainty than their deterministic counterparts, they are not without limitations and a number of concerns have been identified (Fox, 1999, 2006; Newman et al., 2000; Isnard et al., 2001; Pires et al., 2002; Verdonck et al., 2003). As noted by Fox (2001), conventional methods of inference via ANOVA, *t*-tests, regression and related techniques do a superb job when the attendant assumptions have been reasonably met. The problem is that environmental data are notoriously 'messy'. Data collection tends to be opportunistic resulting in samples that are anything but random. The ubiquitous normal distribution assumption underpinning conventional statistical analyses is important, but not as critical as some often overlooked requirements such as homogeneous error variance and independence. Environmental data often exhibit many, if not all of the attributes that render them particularly unsuited to conventional modes of analysis. Among these are as follows: small sample sizes; discreteness; over-dispersion; non-stationarity in space and time; heterogeneous error structures; lack of independence; and extreme skewness/kurtosis. It is not surprising therefore that the NOEC – a quantity determined as the result of a significance test applied to a small sample of non-randomly selected non-normal data, is potentially fatally flawed as an estimator of a no effect concentration.

In a recent and timely article, Newman (2008) critically reviewed the role of classical statistical inference in environmental toxicology and chemistry. His survey of 10 randomly selected articles published between 1996 and 2006 from 10 'representative journals with good impact factors' found none had made any use of Bayesian methods. This is an astonishing result and highlights the need for greater education and awareness of this important statistical paradigm. Indeed, Newman's second recommendation is that 'the teaching of statistics to environmental science students should shift away from a traditional emphasis on hypothesis testing to a more flexible approach embracing other valuable vantages, especially the Bayesian and information theory-based vantages'.

8. Conclusions

In this paper we have described a general Bayesian framework for identifying critical threshold concentrations for ecosystem protection. We have developed the models and provided sufficient mathematical and computational detail with the use of realistic examples to help facilitate the integration of Bayesian methods into the environmental toxicologists' toolkit of statistical techniques. We believe that our approach for estimating a NEC is superior to current NOEC-based methods by virtue of the following: (i) the NEC is estimated as a parameter of a general dose–response model as distinct from the NOEC which is constrained to be one of the test concentrations; (ii) statements of *precision* can be attached to the estimated NEC whereas they are inadmissible for NOECs; (iii) the Bayesian framework provides an opportunity for the researcher to inject personal belief in the form of a *prior* probability distribution for the NEC and other model parameters; (iii) both discrete and continuous probability models for the response-generating mechanism are readily handled thus removing the constraint of assumed normality as is the case with the procedure used to derive NOECs. To this end, it is hoped that this alternative paradigm may alleviate, if not remove some of the long-standing problems associated with the use of traditional modes of statistical inference for ecosystem protection.

Finally, it should be appreciated that whatever *inputs* (NEC, NOEC, IC_x, EC_x, etc.) are used in the development of an SSD, some fundamental issues concerning probabilistic ecological risk assessment (PERA) remain. As has been demonstrated in this paper, it is still possible to obtain an unrealistically low HC_x from an SSD fitted to a collection of NECs. That our analysis of the Hogan et al. (2008) data resulted in a dilution that would yield a toxicant concentration that was below background levels should not be viewed as a failure of our Bayesian NEC-based approach. This situation is not uncommon with PERAs and is the reason the Australian government attached very modest confidence levels (typically 50%) to trigger values obtained using its probabilistic method (ANZECC/ARMCANZ, 2000). The problem lies with the invariably small data sets, the non-random selection of test species, and an unfounded belief that the limited class of probability models used in such exercises is capable of adequately describing behavior in the extreme tails of an SSD. This remains an area for on-going research and development.

Acknowledgments

The author is most grateful to the following individuals for their encouragement and helpful comments on an earlier draft of this paper: Graeme Batley (CSIRO, Land and Water), Chris Humphrey, Rick van Dam, Alicia Hogan, and Andrew Harford (Environmental Research Institute of the Supervising Scientist, Supervising Scientist Division, Department of the Environment, Water, Heritage and Arts). The feedback and constructive comments from three anonymous reviewers significantly improved the final version of this paper.

Appendix A. Winbugs code for examples in text

A.1. Daphnia

```
model
   {
alpha~dgamma(0.0001,0.0001)
beta~dgamma(0.0001,0.0001)
gamma~dgamma(0.0001,0.0001)
for (i in 1:6)
   {
   theta[i] < -alpha*exp(-beta*(x[i]-gamma)*step((x[i]-
   gamma)))
      r[i]~dbin(theta[i],N[i])
   }
}
x[] r[] N[]
0.05 77 80
0.17 79 80
0.28 76 80
0.52 76 80
0.87 71 80
1.14 0 80
END
# Initial values
list(alpha=1,beta=1,gamma=0.1)
```

A.2. Mine pit-waters: single NEC

```
model
   {
      gamma~dunif(0,100)
for (j in 1:4)
   {
      alpha[j]~dnorm(0,0.000001)
      beta[j]~dnorm(0,200)C(0,)
      tau[j]~dgamma(0.0001,0.0001)
      sigma[j] < -sqrt(1/tau[j])
}
   for (i in 1:117)
   {
   mu[i] < -alpha[species[i]]*exp(-beta[species[i]]*(x[i]-
   gamma)*step((x[i]-gamma)))
      Y[i]~dnorm(mu[i],tau[species[i]])
   }
}
   # Input data
   x[] Y[] species[]
   0 1.4383 1
   0 1.3913 1
   0 1.4779 1
   0.3 1.3379 1
   .
   .
   .
   30 0 4
   END
# Initial values
list(gamma=10)
alpha[] beta[] tau[]
1 0.065 0.1
1 0.065 0.1
1 0.065 0.1
1 0.065 0.1
END
```

A.3. Mine pit-waters: multiple NECs

```
model
   {
      gamma[1]~dlnorm(2.124,1.708)
      gamma[2]~dlnorm(3.201,1.582)
      gamma[3]~dlnorm(1.221,0.437)
      gamma[4]~dlnorm(-1.278,0.404)
for (j in 1:4)
   {
      alpha[j]~dnorm(0,0.000001)
      beta[j]~dnorm(0,200)C(0,)
      tau[j]~dgamma(0.0001,0.0001)
      sigma[j] < -sqrt(1/tau[j])
}
   for (i in 1:117)
   {
   mu[i] < -alpha[species[i]]*exp(-beta[species[i]]*(x[i]-
   gamma[species[i]])*step((x[i]-gamma[species[i]])))
      Y[i]~dnorm(mu[i],tau[species[i]])
   }
}
# Initial values
alpha[] beta[] tau[] gamma[]
1 0.065 0.1 10
1 0.065 0.1 3
1 0.065 0.1 3
1 0.065 0.1 0.3
END
   # Input data
   x[] Y[] species[]
   0 1.4383 1
   0 1.3913 1
   0 1.4779 1
   0.3 1.3379 1
   .
   .
   .
   30 0 4
   END
   a[]
   0.05
   0.125
   0.025
   0.3
   4.5
   END
```

References

Aldenberg, T., Jaworska, J.S., 2000. Estimation of the hazardous concentration and fraction affected for normally distributed species sensitivity distributions. Ecotoxicol. Environ. Saf. 46, 1–18.

Aldenberg, T., Slob, W., 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. Ecotoxicol. Environ. Saf. 25, 48–63.

ANZECC/ARMCANZ, 2000. Australian and New Zealand guidelines for fresh and marine water quality. National Water Quality Management Strategy Paper No. 4, Australian and New Zealand Environment and Conservation Council/ Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia.

Biesinger, K.E., Anderson, L.E., Eaton, J.G., 1982. Chronic effects of inorganic and organic mercury on *Daphnia magna*: toxicity, accumulation and loss. Arch. Environ. Contam. Toxicol. 11, 769–774.

Billoir, E., Delignette-Muller, M.L., Péry, A.R.R., Charles, S., 2008. A Bayesian approach to analyzing ecotoxicological data. Environ. Sci. Technol. 42, 8978–8984.

Crane, M., Newman, M.C.N., 2000. What level of effect is a no observed effect?. Environ. Toxicol. Chem. 19, 516–519.

Fox, D.R., 1999. Setting water quality guidelines – a statistician's perspective. SETAC News, 17–18 May.

Fox, D.R., 2001. Environmental power analysis – a new perspective. Environmetrics 12, 437–449.
Fox, D.R., 2006. Statistical issues in ecological risk assessment. Hum. Ecol. Risk Assess. 12, 120–129.
Fox, D.R., 2008. Natural resource management. In: Melnick, E., Everitt, B. (Eds.), Encyclopedia of Quantitative Risk Assessment and Analysis. John Wiley and Sons Ltd., Chichester, UK, pp. 1600–1603.
Fox, D.R., 2009. Is the EC_x a legitimate surrogate for the NOEC?. Integr. Environ. Assess. Manage. 5, 351–353.
Grist, E.P.M., O'Hagan, A., Crane, M., Sorokin, N., Sims, I., Whitehouse, P., 2006. Bayesian and time-independent species sensitivity distributions for risk assessment of chemicals. Environ. Sci. Technol. 40, 395–401.
Hickey, G.L., Kerford, B., Dunlop, J.E., Craig, P.S., 2008. Making species sensitivity distributions reflective of naturally occurring communities: using rapid testing and Bayesian statistics. Environ. Toxicol. Chem. 27, 2403–2411.
Hogan, A., van Dam, R., Houston, M., Lee, N., 2008. Toxicity of Ranger mine RP2 and pit 3 waters to native freshwater species: 2007 wet season. Supervising Scientist Report 197, Supervising Scientist, Darwin NT, Australia.
Isnard, P., Flammarion, P., Roman, G., Babut, M., Bastien, Ph., Bintein, S., Esserméant, L., Férard, J.F., Gallotti-Schmitt, S., Saouter, E., Saroli, M., Thiébaud, H., Thomassone, R., Vindimian, E., 2001. Statistical analysis of regulatory ecotoxicity tests. Chemosphere 45, 659–669.
Jager, T., Heugens, E.H.W., Kooijman, S.A.L.M., 2006. Making sense of ecotoxicological test results: towards application of process-based models. Ecotoxicology 15, 305–314.
Kooijman, S.A.L., Hanstveit, A.O., Nyholm, N., 1996. No-effect concentrations in algal growth inhibition tests. Water Res. 30, 1625–1632.
Kooijman, S.A.L.M., 2006. An alternative for NOEC exists, but the standard model has to be abandoned first. Oikos 75, 310–316.
Lee, P.M., 2004. Bayesian Statistics: An Introduction. Oxford University Press.
Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10, 325–337.
McCarthy, M.A., 2007. Bayesian Methods for Ecology. Cambridge University Press.
Newman, M.C., 2008. What exactly are you inferring? A closer look at hypothesis testing. Environ. Toxicol. Chem. 27, 1013–1019.
Newman, M.C., Ownby, D.R., Mézin, L.C.A., Powell, D.C., Christensen, T.R.L., Lerberg, S.B., Anderson, B.-A., 2000. Applying species-sensitivity distributions in ecological risk assessment: assumptions of distribution type and sufficient number of species. Environ. Toxicol. Chem. 19, 508–515.
Pires, A.M., Branco, J.A., Picado, A., Mendonça, E., 2002. Models for the estimation of a 'no effect concentration'. Environmetrics 13, 15–27.
Posthuma, L., Suter, G.W., Traas, T.P., 2002. Species Sensitivity Distributions in Ecotoxicology. CRC Press.
Shao, Q., 2000. Estimation for hazardous concentrations based on NOEC toxicity data: an alternative approach. Environmetrics 11, 583–595.
Van der Hoeven, N., 1997. How to measure no effect. Part III: statistical aspects of NOEC, EC_x, and NEC estimates. Environmetrics 8, 255–261.
Van der Hoeven, N., 2004. Current issues in statistics and models for ecotoxicological risk assessment. Acta Biotheor. 52, 201–217.
Verdonck, F.A.M., Jaworska, J., Thas, O., Vanrolleghem, P.A., 2001. Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study. Anal. Chim. Acta 446, 427–436.
Verdonck, F.A., Aldenberg, T., Jaworska, J., Vanrolleghem, P.A., 2003. Limitations of current risk characterization methods in problematic environmental risk assessment. Environ. Toxicol. Chem. 22, 2209–2213.
Wagner, C., Løkke, H., 1991. Estimation of ecotoxicological protection levels from NOEC toxicity data. Water Res. 25, 1237–1242.