Ecotoxicology and Environmental Safety **46,** 1–18 (2000)
Environmental Research, Section B
doi:10.1006/eesa.1999.1869, available online at http://www.idealibrary.com on IDEAL®

REVIEW

Uncertainty of the Hazardous Concentration and Fraction Affected for Normal Species Sensitivity Distributions

Tom Aldenberg* and Joanna S. Jaworska†
** RIVM, Bilthoven, The Netherlands; and †Procter & Gamble, Eurocor, Strombeek-Bever, Belgium*

Received October 12, 1998

Table of Contents

Species in the environment vary according to their sensitivity to a toxicant. Because these differences in sensitivity are unique to the toxicant at consideration and laboratory data sets to assess this variability are very small due to cost, it is important to provide uncertainty estimates of (1) environmental quality objectives (hazardous concentrations) derived from these laboratory data and (2) fraction of species affected at given, or predicted, laboratory or environmental concentrations. This article focuses on the normal (Gaussian) distribution of species sensitivity. It examines and compares results of Problems (1) and (2) from two opposing statistical philosophies, Bayesian and Classical, leading to vastly different numerical approaches. For the normal model, both approaches lead to identical answers, numerically. Extrapolation factors for the lower, median, and upper estimates of the hazardous concentration at six levels of protection are derived. Furthermore, upper, median, and lower estimates of the fraction affected at given, standardized, logarithmic concentrations have been tabulated. This table can be used directly for risk assessment without reference to protection levels or hazardous concentrations. The confidence limits for hazardous concentration and fraction affected depend heavily on the number of species tested

0147-6513/00 $35.00

and are independent of the toxic substance involved (provided the model is right), due to correction for the mean and standard deviation of the toxicity data. The equivalence of confidence limits for hazardous concentration and fraction affected is captured in the *law of extrapolation*: the upper (median, lower) confidence limit for the fraction affected at the lower (median, upper) confidence limit of the hazardous concentration is equal to the fraction affected (e.g., 5%) used to define the hazardous concentration. The upper confidence limit for the fraction affected at the median estimate of the hazardous concentration for 5% of the species is a fixed number depending on the sample size of the toxicity data only. It amounts to 46% at $n=3$, down to 20% at $n=10$, and still 12% at n 30.

Key Words: **risk assessment; environmental quality objective; extrapolation; hazardous concentration; fraction affected; Bayesian statistics; confidence limit.**

1. INTRODUCTION

According to Van Straalen and Denneman (1989), a concentration of a certain toxic compound is considered hazardous for p% of a set of biological species if the probability of selecting a species from this set with a no-observed-effect concentration (NOEC) smaller than this concentration is equal to p%. Since the default value is 5%, this hazardous concentration is called the HC_5. HC_5s are used for setting quality objectives for the environment (reference for Dutch legislation). The reasoning is that if concentrations of this compound are below the HC_5, more than 95% of the biological species set considered will not display effects as determined by the chronic toxicity tests. Obviously, taking the fifth percentile is arbitrary to a large extent, although there have been attempts to validate this choice with more integrated studies, e.g., at the microecosystem or mesocosm level (Okkerman *et al*., 1993; Versteeg *et al*., 1999).

The statistical model of differences in species sensitivity with respect to a toxic compound can be used in essentially two ways:

1. to estimate confidence limits of the HC_5 from a possibly small set of species NOECs for a certain compound, e.g., in order to obtain an environmental quality objective; and

2. to estimate confidence limits of the fraction of species affected (FA) at one or more given concentrations of the toxicant, either above or below the HC_5.

The statistical model is based on either the logistic distribution (Kooijman, 1987; Van Straalen and Denneman, 1989; Aldenberg and Slob, 1993; Aldenberg, 1993) or the normal distribution (Wagner and Løkke, 1991; Smith and Cairns, 1993).

Van Straalen and Denneman (1989) already applied the method in an inverse manner, that is, to estimate the fraction of species not protected at given (measured, proposed, or predicted) environmental concentrations. They inverted their formula for the *lower* confidence limit of the hazardous concentration to obtain an estimate of the fraction affected at a given environmental concentration, but they did not discuss the nature of this estimate.

Both concepts, the hazardous concentration (HC) and the fraction of species affected (FA) at a given concentration, are fundamental to probabilistic environmental risk assessment. We may also speak of *extrapolation* methods to denote HC assessments, and *risk assessment* for assessments of FA. The relationship between quality objective setting or estimation of the HC and estimating the fraction affected (FA) is approached from two different statistical viewpoints: confidence arguments, which are in the realm of classical or sampling statistics, and credibility (or degree of belief) arguments, which belong to the domain of Bayesian statistics, or Bayesian inference (Box and Tiao, 1973; Lee, 1989; Press, 1989). We show that both approaches, although conceptually and numerically different, will nevertheless lead to the same answers. This means that the tables that summarize the uncertainty of the HC and of the FA are valid from both viewpoints.

In this paper, the focus is on the normal or Gaussian distribution only, and treatment of other distributions (e.g. the logistic) is deferred to a later stage. The reasons for doing so are fivefold. First, from the point of toxicity, there is no theoretical justification for any distribution to be the more fundamental to the subject. Second, one cannot statistically decide between different distributions at small sample size. Third, the normal distribution is very well studied, and some concepts we can recycle to our advantage, for example, the treatment of tolerance limits in quality control and the standard Bayesian treatment of the normal distribution in textbooks. Fourth, the normal model has some nice mathematical features in the form of the sample mean and standard deviation being so-called sufficient statistics. They summarize all information in the sample. Fifth, in extending the model to incorporate more structure, e.g., accounting for individual body weight or fat content or species groupings, the extrapolation problem becomes one of the area of regression analysis, or analysis of variance, for which the normal distribution of the remaining variation is the one most studied.

The paper is highly focused on the environmental risk assessment application. However, it has much broader general application of quantifying uncertainty of a fitted distribution of variability or, in other words, second-order distribution fitting (Burmaster and Wilson, 1996). This area of quantitative risk analysis is currently investigated in depth, but mainly methods from classical statistics, such as the bootstrap (Frey and Rhodes, 1998) or maximum likelihood approach (Burmaster and Thompson, 1998) have been applied so far, with an emphasis on parametric analyses. It is interesting to note that Bayesian analysis applied to species sensitivity distributions, as presented here, also leads to statements of secondary uncertainty.

In this paper, concepts of extrapolation are reviewed in Section 2. In Section 3, the statistical methodology is outlined from the point of view of classical statistics (sampling theory). Section 4 presents the approach from the point of view of Bayesian statistics. Some consequences of the results are given in Section 5 and discussed in Section 6. Mathematical and numerical details are treated in the Appendix. The core results are Table 1 (Extrapolation Factors to Estimate Hazardous Concentrations), Table 2 (Fraction Affected at Given Concentration), Table 5 (Uncertainty of Fraction Affected at Different Estimates of Hazardous Concentration), and Eq. (12) (the law of extrapolation).

2. BASIC CONCEPTS OF EXTRAPOLATION THROUGH THE SPECIES SENSITIVITY DISTRIBUTION

2.1. Hazardous Concentration and Fraction Affected

Suppose the sensitivity of different biological species for a toxicant, as expressed by their log(NOEC), or some other endpoint, is normally distributed and therefore parameterized by μ and σ. Figure 1 depicts a theoretical species sensitivity distribution (SSD) as a standard normal probability density function (PDF). The log(HC) for p% of the species population is

$$\log(\mathrm{HC}_p) = \mu - K_p \cdot \sigma \tag{1}$$

For example, $K_5 = 1.6449$. If both μ and σ were known, this calculation would suffice, provided the model is justified. In this case, $\log(\mathrm{HC}_5)$ could be calculated with deterministic

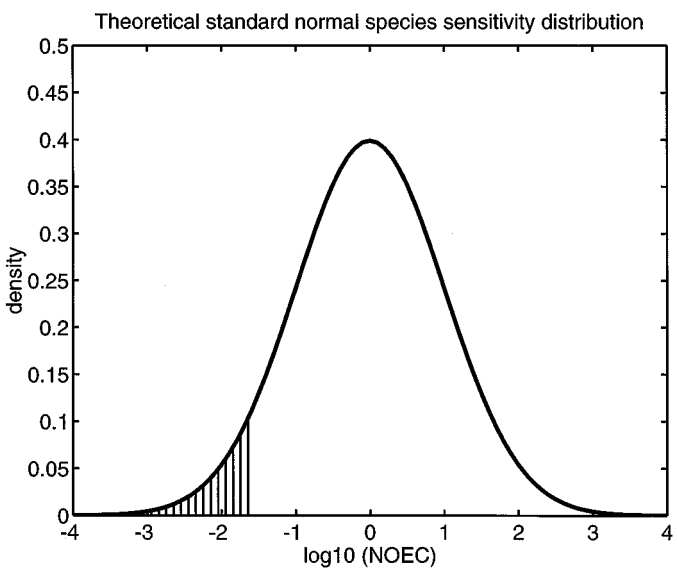

FIG. 1. Theoretical species sensitivity distribution (SSD) taken as the standard normal probability density function (PDF) over $\log_{10}$(NOEC). The fraction affected (5%) and the $\log_{10}$(hazardous concentration) = -1.6449, i.e., fifth percentile, are indicated.

precision. Figure 1 illustrates that the method essentially amounts to calculating the fifth percentile of the standard normal distribution. The FA is presented hatched, when the environmental concentration would be exactly at $\log(\mathrm{HC}_5)$, i.e.,

$$\mathrm{FA} = 5\%.$$

In this case, the NOECs of species within the hatched area would be below the environmental concentration; hence they would possibly be affected by the toxicant. At other given or predicted concentrations the FA is given by the area below the density curve to the left of the log(concentration) value, as calculated by the normal cumulative distribution function (CDF).

2.2. Extrapolation Factors

In the case of partial knowledge of the sensitivity of species for a toxicant, e.g., through a small sample of toxicity data, one calculates the mean ($\bar{x}$) and sample standard deviation (s) of the set of log-transformed toxicity data. The extrapolation method entails subtracting a factor k_s (the extrapolation factor for the sample) times s from the sample mean to estimate $\log(\mathrm{HC}_5)$:

$$\log(\hat{\mathrm{HC}}_5) = \bar{x} - k_s \cdot s. \tag{2}$$

Aldenberg and Slob (1993) present extrapolation factors for the estimation of the $\log(\mathrm{HC}_5)$ based on the assumption of logistic species sensitivity distributions for the log-transformed toxicity data. Wagner and Løkke (1991) and Smith and Cairns (1993) give analogous extrapolation factors for presumed normal distributions applied to similar data sets.

To assess the uncertainty of the HC, one must associate levels of confidence to the estimates. Aldenberg and Slob (1993) give two levels of confidence—a lower confidence limit (95% one-sided) and a median estimate—while Wagner and Løkke (1991) and Smith and Cairns (1993) focus on the lower confidence limit only. None of the three articles yields upper confidence limits, nor two-sided intervals.

There are two statistical methodologies for deriving confidence limits: classical and Bayesian. In the next two sections, confidence intervals are derived for the HC and FA using first classical statistics, or sampling theory (Section 3), and, secondly, Bayesian statistics (Section 4). Mathematical and numerical details are presented in the Appendix.

3. EXTRAPOLATION UNCERTAINTY THROUGH CLASSICAL STATISTICS

3.1. Confidence Limits of the Hazardous Concentration

In classical statistics, reasoning often refers to properties of infinitely repeated samples. In this statistical viewpoint,

the statistic of Eq. (2), which varies over different samples, is to *under*estimate the true value given by Eq. (1) in most, e.g., 95%, of the samples. This point of view is called sampling theory. The underestimation is wanted to estimate a HC at the safe side, and hence is motivated by the (eco-)toxicological application of the method. It still overpredicts in a minority (5%) of the samples. But when one applies Eq. (2) to the often single sample at hand, one feels 95% *confident* to predict on the safe side and the result is called a lower confidence limit. This reasoning will be called the confidence argument.

When the confidence to underestimate is 50%, one has a median estimate, in which case the percentage of samples that overestimate the true value is 50% as well. An upper confidence limit results, when the confidence to underestimate is only 5%, hence the majority of samples (95%) will overpredict. Lower and upper confidence limits together constitute 90% two-sided confidence intervals. In general, e.g., statistics textbooks, two-sided confidence limits are taken at the 95% level. However, in ecotoxicological extrapolation, 95% one-sided confidence levels, and hence 90% two-sided, have been used. Ultimately, the level of confidence is a risk management decision and varying the level of confidence may provide additional information to the risk manager.

The calculation of extrapolation factors for the normal distribution is mathematically analogous to the problem of tolerance limits in industrial applications (Owen, 1968, Odeh and Owen, 1980; Wagner and Løkke, 1991; Smith and Cairns, 1993). The objective is to find k_s such that the statistic in Eq. (2) underestimates the true value in Eq. (1) with probability γ:

$$\Pr(\bar{x} - k_s \cdot s \leq \mu - K_p \cdot \sigma) = \gamma. \tag{3}$$

In the Appendix, how to calculate the extrapolation factor k_s on the basis of the so-called noncentral t distribution in the case that the SSD is a normal distribution is explained.

Given the confidence level (γ), the extrapolation factor is independent of particular values of μ and σ, and hence of the substance involved, provided the model is justified. The extrapolation factor is therefore a function of the size of the sample of the toxicity data only. This means that one can take $\mu = 0$ and $\sigma = 1$, without loss of generalization. Henceforth, k_s can be tabulated for all substances.

Table 1 presents extrapolation factors for lower, median, and upper estimates of the HC at six levels of FA. Different FAs are given to relax the emphasis on the default FA of 5% somewhat. For example, 50% is used in the Netherlands for evaluating hazardous waste sites (Swartjes, 1999). The confidence levels are 95, 50, and 5%, respectively, taken as one-sided underestimates. Lower and upper confidence limits together constitute 90% two-sided confidence intervals.

3.2. *Confidence Limits for the Fraction Affected*

The methodology to estimate the FA and its uncertainty at a given (environmental) concentration is basically identical to that of estimating the HC. Recalling Eq. (3), one observes that the statistic $\bar{x} - k_s \cdot s$ varies over samples, while $\mu - K_p \cdot \sigma$ is a fixed number. Both terms are logarithmic concentrations. It will now be explained that reasoning similar to that used for the extrapolation factors can be applied to estimating the uncertainty of FA. Without loss of generalization, we again assume $\mu = 0$ and $\sigma = 1$, i.e., the standard normal distribution.

If we focus, for example, on $n = 7$ (FA $= 5\%$) in Table 1, we note that for the standard normal distribution, the log(HC) of -1.6449 is underestimated with probability 0.95 by the statistic $\bar{x} - 3.3995 \cdot s$, with mean and standard deviation calculated from samples of size 7 drawn from the standard normal distribution. Consider the sampling distribution of this statistic. Its 95th percentile equals -1.6449. One could also consider the sampling distribution of

$$\mathrm{FA}(\bar{x} - 3.3995 \cdot s), \tag{4}$$

i.e., the FA associated with the underestimating log concentrations. However, since FA is an increasing function of log concentration, that is, larger log concentrations have larger FA values, the 95th percentile of this *distribution of fractions* is exactly equal to

$$\mathrm{FA}(-1.6449) = 0.05, \tag{5}$$

i.e., the FA to define the log HC. Mathematically, Eq. (3) also holds for fractions:

$$\Pr(\mathrm{FA}(\bar{x} - k_s \cdot s) \leq \mathrm{FA}(\mu - K_p \cdot \sigma)) = \gamma \tag{6}$$

(Mood *et al.*, 1974, p. 378). Consequently, expression (4) underestimates (5) in 95% of the samples, in a similar way as (2) underestimates (1).

In practice, we have only *one* sample and one calls the particular $\bar{x} - 3.3995 \cdot s$ for the mean and standard deviation of the sample at hand a lower confidence limit of the log HC. Similarly, if a single given log concentration happens to be at $\bar{x} - 3.3995 \cdot s$, that is, if the standardized log concentration is -3.3995 after correction for mean and standard deviation of the toxicity data, then we call 0.05 (5%) an *upper* confidence limit of the FA to be estimated, in relation to this concentration. To summarize, at a standardized log concentration of -3.3995, the 95% upper confidence limit of FA is 0.05.

Table 2 presents upper, median, and lower estimates of the FA at given, or predicted (environmental), standardized log concentrations. Confidence levels (γ) are 95, 50, and 5%, respectively. Odeh and Owen (1980) present more extended tables. The use of commercially available mathematical software is described in the Appendix.

TABLE 1
Extrapolation Factors, k_s, for estimating Log Hazardous Concentrations According to Eq. (2) for Six Levels of the Fraction Affected at Three Levels of Confidence (95, 50, and 5%) as a Function of Toxicicity Data Sample Size[a]

	FA = 1%			FA = 2%			FA = 5%		
Sample size	Lower	Median	Upper	Lower	Median	Upper	Lower	Median	Upper
2	37.0936	3.3760	0.9538	32.7500	2.9624	0.7756	26.2597	2.3387	0.4748
3	10.5527	2.7645	1.1297	9.3855	2.4342	0.9416	7.6559	1.9384	0.6391
4	7.0424	2.6008	1.2462	6.2767	2.2923	1.0512	5.1439	1.8295	0.7433
5	5.7411	2.5258	1.3309	5.1206	2.2272	1.1308	4.2027	1.7793	0.8178
6	5.0620	2.4828	1.3964	4.5158	2.1899	1.1923	3.7077	1.7505	0.8748
7	4.6417	2.4551	1.4492	4.1409	2.1657	1.2417	3.3995	1.7318	0.9204
8	4.3539	2.4357	1.4931	3.8836	2.1489	1.2828	3.1873	1.7187	0.9580
9	4.1430	2.4213	1.5303	3.6950	2.1364	1.3176	3.0312	1.7091	0.9899
10	3.9811	2.4103	1.5625	3.5499	2.1268	1.3476	2.9110	1.7016	1.0173
11	3.8523	2.4016	1.5908	3.4345	2.1192	1.3740	2.8150	1.6957	1.0413
12	3.7471	2.3945	1.6158	3.3400	2.1131	1.3973	2.7363	1.6910	1.0625
13	3.6592	2.3886	1.6382	3.2611	2.1080	1.4182	2.6705	1.6870	1.0814
14	3.5845	2.3837	1.6585	3.1939	2.1037	1.4371	2.6144	1.6837	1.0985
15	3.5201	2.3795	1.6769	3.1360	2.1000	1.4542	2.5660	1.6808	1.1140
20	3.2952	2.3652	1.7492	2.9334	2.0876	1.5215	2.3960	1.6712	1.1746
30	3.0639	2.3516	1.8397	2.7246	2.0758	1.6054	2.2198	1.6620	1.2498
50	2.8624	2.3412	1.9362	2.5422	2.0667	1.6947	2.0650	1.6549	1.3294
100	2.6840	2.3337	2.0401	2.3801	2.0601	1.7907	1.9265	1.6498	1.4143
200	2.5697	2.3300	2.1182	2.2761	2.0569	1.8627	1.8372	1.6473	1.4778
500	2.4750	2.3280	2.1910	2.1910	2.0550	1.9300	1.7630	1.6458	1.5367
Inf	**2.3263**	**2.3263**	**2.3263**	**2.0537**	**2.0537**	**2.0537**	**1.6449**	**1.6449**	**1.6449**

	FA = 10%			FA = 25%			FA = 50%		
Sample size	Lower	Median	Upper	Lower	Median	Upper	Lower	Median	Upper
2	20.5815	1.7842	0.1380	11.7630	0.8874	− 0.9347	4.4645	0.0000	− 4.4645
3	6.1553	1.4985	0.3345	3.8062	0.7732	− 0.3494	1.6859	0.0000	− 1.6859
4	4.1619	1.4189	0.4439	2.6176	0.7385	− 0.1680	1.1767	0.0000	− 1.1767
5	3.4066	1.3818	0.5188	2.1497	0.7217	− 0.0660	0.9534	0.0000	− 0.9534
6	3.0063	1.3605	0.5748	1.8950	0.7119	0.0031	0.8226	0.0000	− 0.8226
7	2.7554	1.3466	0.6190	1.7322	0.7055	0.0545	0.7345	0.0000	− 0.7345
8	2.5819	1.3369	0.6552	1.6178	0.7009	0.0948	0.6698	0.0000	− 0.6698
9	2.4538	1.3296	0.6856	1.5322	0.6975	0.1277	0.6198	0.0000	− 0.6198
10	2.3546	1.3241	0.7116	1.4652	0.6949	0.1552	0.5797	0.0000	− 0.5797
11	2.2753	1.3197	0.7342	1.4111	0.6928	0.1787	0.5465	0.0000	− 0.5465
12	2.2101	1.3161	0.7541	1.3663	0.6911	0.1991	0.5184	0.0000	− 0.5184
13	2.1554	1.3132	0.7719	1.3284	0.6897	0.2170	0.4943	0.0000	− 0.4943
14	2.1088	1.3107	0.7878	1.2959	0.6885	0.2330	0.4733	0.0000	− 0.4733
15	2.0684	1.3085	0.8023	1.2676	0.6875	0.2473	0.4548	0.0000	− 0.4548
20	1.9260	1.3013	0.8585	1.1665	0.6840	0.3019	0.3866	0.0000	− 0.3866
30	1.7773	1.2944	0.9276	1.0584	0.6807	0.3670	0.3102	0.0000	− 0.3102
50	1.6456	1.2891	1.0000	0.9603	0.6782	0.4331	0.2371	0.0000	− 0.2371
100	1.5267	1.2853	1.0767	0.8696	0.6763	0.5011	0.1660	0.0000	− 0.1660
200	1.4496	1.2834	1.1335	0.8094	0.6754	0.5503	0.1169	0.0000	− 0.1169
500	1.3851	1.2823	1.1860	0.7583	0.6748	0.5949	0.0737	0.0000	− 0.0737
Inf	**1.2816**	**1.2816**	**1.2816**	**0.6745**	**0.6745**	**0.6745**	**0.0000**	**0.0000**	**0.0000**

[a] Upper and lower estimates constitute a 90% two-sided confidence interval.

Input concentrations must be standardized according to

$$-k_s = \frac{\log(\text{Conc}) - \bar{x}}{s}, \tag{7}$$

with $\bar{x}$ the mean and s the sample standard deviation of the set of log-transformed toxicity data.

The confidence levels (95, 50, and 5%) are one-sided overestimates. In Table 2, linear interpolation for noninteger

TABLE 2
Upper, Median, and Lower Estimates (95, 50, 5% Confidence) of the Fraction Affected (%) at Eleven Given Standardized Logarithmic Concentrations (Columns, $-k_s$) and Toxicity Data Sample Size (Rows, n)[a]

	− 5	− 4	− 3	− 2	− 1	0	1	2	3	4	5
	47.428	52.391	58.172	65.172	74.288	87.760	99.180	99.998	100.000	100.000	100.000
2	0.033	0.307	1.883	7.734	22.552	50.000	77.448	92.266	98.117	99.693	99.967
	0.000	0.000	0.000	0.002	0.820	12.240	25.712	34.828	41.828	47.609	52.572
	16.068	23.339	32.913	45.360	61.655	82.886	98.371	99.989	100.000	100.000	100.000
3	0.001	0.041	0.585	4.498	19.317	50.000	80.683	95.502	99.415	99.959	99.999
	0.000	0.000	0.000	0.011	1.629	17.114	38.345	54.640	67.087	76.661	83.932
	5.566	11.115	20.294	34.219	53.952	79.458	97.636	99.971	100.000	100.000	100.000
4	0.000	0.018	0.369	3.629	18.165	50.000	81.835	96.371	99.631	99.982	100.000
	0.000	0.000	0.000	0.029	2.364	20.542	46.048	65.781	79.706	88.885	94.434
	2.273	6.017	13.763	27.511	48.812	76.901	96.994	99.948	100.000	100.000	100.000
5	0.000	0.012	0.290	3.242	17.585	50.000	82.415	96.758	99.710	99.988	100.000
	0.000	0.000	0.000	0.052	3.006	23.099	51.188	72.489	86.237	93.983	97.727
	1.085	3.643	10.058	23.134	45.131	74.905	96.435	99.921	100.000	100.000	100.000
6	0.000	0.009	0.250	3.025	17.237	50.000	82.763	96.975	99.750	99.991	100.000
	0.000	0.000	0.000	0.079	3.565	25.095	54.869	76.866	89.942	96.357	98.915
	0.587	2.404	7.771	20.087	42.352	73.293	95.945	99.893	100.000	100.000	100.000
7	0.000	0.008	0.226	2.887	17.006	50.000	82.994	97.113	99.774	99.992	100.000
	0.000	0.000	0.000	0.107	4.055	26.707	57.648	79.913	92.229	97.596	99.413
	0.349	1.694	6.259	17.854	40.171	71.956	95.512	99.865	99.999	100.000	100.000
8	0.000	0.007	0.210	2.792	16.841	50.000	83.159	97.208	99.790	99.993	100.000
	0.000	0.000	0.001	0.135	4.488	28.044	59.829	82.146	93.741	98.306	99.651
	0.224	1.256	5.204	16.151	38.408	70.825	95.126	99.837	99.999	100.000	100.000
9	0.000	0.006	0.199	2.722	16.718	50.000	83.282	97.278	99.801	99.994	100.000
	0.000	0.000	0.001	0.163	4.874	29.175	61.592	83.849	94.796	98.744	99.776
	0.152	0.969	4.435	14.813	36.948	69.852	94.780	99.809	99.999	100.000	100.000
10	0.000	0.006	0.191	2.669	16.622	50.000	83.378	97.331	99.809	99.994	100.000
	0.000	0.000	0.001	0.191	5.220	30.148	63.052	85.187	95.565	99.031	99.848
	0.108	0.771	3.855	13.732	35.716	69.003	94.466	99.782	99.998	100.000	100.000
11	0.000	0.005	0.184	2.627	16.546	50.000	83.454	97.373	99.816	99.995	100.000
	0.000	0.000	0.002	0.218	5.534	30.997	64.284	86.268	96.145	99.229	99.892
	0.080	0.630	3.406	12.843	34.661	68.255	94.180	99.755	99.998	100.000	100.000
12	0.000	0.005	0.179	2.593	16.483	50.000	83.517	97.407	99.821	99.995	100.000
	0.000	0.000	0.002	0.245	5.820	31.745	65.339	87.157	96.594	99.370	99.920
	0.061	0.526	3.048	12.097	33.744	67.588	93.918	99.729	99.998	100.000	100.000
13	0.000	0.005	0.175	2.565	16.431	50.000	83.569	97.435	99.825	99.995	100.000
	0.000	0.000	0.002	0.271	6.082	32.412	66.256	87.903	96.952	99.474	99.939
	0.048	0.447	2.758	11.463	32.940	66.989	93.678	99.704	99.997	100.000	100.000
14	0.000	0.005	0.172	2.542	16.387	50.000	83.613	97.458	99.828	99.995	100.000
	0.000	0.000	0.003	0.296	6.322	33.011	67.060	88.537	97.242	99.553	99.952
	0.038	0.385	2.519	10.916	32.227	66.447	93.455	99.680	99.997	100.000	100.000
15	0.000	0.005	0.169	2.522	16.350	50.000	83.650	97.478	99.831	99.995	100.000
	0.000	0.000	0.003	0.320	6.545	33.553	67.773	89.084	97.481	99.615	99.962
	0.016	0.216	1.769	9.025	29.590	64.349	92.548	99.572	99.994	100.000	100.000
20	0.000	0.004	0.159	2.455	16.222	50.000	83.778	97.545	99.841	99.996	100.000
	0.000	0.000	0.006	0.428	7.452	35.651	70.410	90.975	98.231	99.784	99.984
	0.006	0.106	1.145	7.145	26.639	61.803	91.347	99.403	99.989	100.000	100.000
30	0.000	0.004	0.150	2.391	16.098	50.000	83.902	97.609	99.850	99.996	100.000
	0.000	0.000	0.011	0.597	8.653	38.197	73.361	92.855	98.855	99.894	99.994
	0.002	0.051	0.728	5.605	23.876	59.197	89.999	99.177	99.980	100.000	100.000
50	0.000	0.004	0.144	2.343	16.003	50.000	83.997	97.657	99.856	99.996	100.000
	0.000	0.000	0.020	0.823	10.001	40.803	76.124	94.395	99.272	99.949	99.998
Inf	**0.000**	**0.003**	**0.135**	**2.275**	**15.866**	**50.000**	**84.134**	**97.725**	**99.865**	**99.997**	**100.000**

[a] Upper and lower estimates constitute a 90% two-sided confidence interval.

standardized log-concentrations will yield only approximate results. Appendix Section 8.2.3 gives an example of smooth interpolation with the aid of repeated linear interpolation (so-called Aitken interpolation), which can be done by hand or in a spreadsheet.

3.3. Example: Cadmium

Table 3 reproduces data on cadmium toxicity to soil organisms (Van Straalen and Denneman, 1989). This data set will be our running example.

Hazardous concentration. For 95% protection (FA = 5%) and a sample size of 7 for the cadmium data set, we find from Table 1, assuming a normal distribution, the extrapolation factors k_s: 3.3995 (lower estimate); 1.7318 (median estimate), and 0.9204 (upper estimate). By applying (2), with mean and standard deviation from Table 1, and taking the anti-log, the lower estimate of the HC_5 equals $10^{(0.97124-3.3995*0.70276)} = 0.038$ μg Cd/g, the median estimate yields 0.568 μgCd/g, and the upper estimate is 2.112 μg Cd/g. Accordingly, the two-sided 90% confidence interval for the HC_5 is (0.038, 2.112), which spans a factor of almost 56.

Fraction affected. As an example environmental concentration, a value of 0.8 μgCd/g was proposed for cadmium in soils (Table 3). The standardized log concentration is

$$-k_s = \frac{\log(\text{Conc}) - \bar{x}}{s} = \frac{-0.09691 - 0.97124}{0.70276} = -1.5194.$$

Linear interpolation in Table 2 yields a median estimate of 9.665% for the FA, with an upper limit of 30.776% and a lower limit of 2.002%. Repeated linear interpolation (Appendix Section 8.2.3) on the basis of 4 points yields a 7.778% median estimate with 29.612% upper limit and 0.608% lower limit. On the basis of available mathematical software (Appendix Section 8.2.1), we arrive, with $k_s = 1.51994$, at a median estimate of 7.421% for the FA. The 90% two-sided confidence interval of FA at 0.8 μgCd/g equals (0.775%, 29.438%) on the basis of this software, which are the exact values.

Note the skewness of the confidence interval with respect to 7.421% as the median estimate. With 95% confidence the fraction of species affected may be up to almost 30% at this proposed concentration.

TABLE 3
NOEC Values for Toxicity of Cadmium (μg Cd/g) of Seven Soil Organisms[a] and Proposed Quality Objective

Species	NOEC(μg Cd/g)	$\log_{10}$(NOEC)
1	0.97	−0.01323
2	3.33	0.52244
3	3.63	0.55991
4	13.50	1.13033
5	13.80	1.13988
6	18.70	1.27184
7	154.00	2.18752
Mean:		0.97124
Standard deviation:		0.70276
Objective	0.80	−0.09691

[a] Van Straalen and Denneman (1989).

4. EXTRAPOLATION UNCERTAINTY THROUGH BAYESIAN STATISTICS

The Bayesian statistical method considers the role of sample and model reversed: the sample is fixed and unique, and the model itself is uncertain. This statistical viewpoint corresponds better to the practical situation the individual researcher is facing: there is only one sample and there are doubts what model to use, or, if the model is chosen, what parameter values to take. The uncertainty of the model is modeled by assuming that the parameters of the model are themselves distributed.

In the present case of a normally distributed species sensitivity distribution for a toxicant, the Bayesian treatment of the uncertainty of μ and σ of a presupposed normal distribution is well studied (e.g., Box and Tiao, 1973, Chap. 2, Section 2.4). If one assumes parameter values to be distributed, one must presuppose a so-called *prior* distribution for the parameters, to specify the initial state of knowledge about them, before the data are used. When the data are used, the prior is transformed into the so-called *posterior* distribution by multiplication with the classic likelihood function. This is Bayes' theorem essentially. The posterior distribution summarizes our increase in knowledge about the parameters due to the data. A Bayesian simulation centers about the evaluation of the joint posterior distribution of the parameters, in this case μ and σ. For technical details the reader is referred to the Appendix.

4.1. Probability Density Function and Its Confidence Limits: Secondary Probability Distributions

As stated above, in Bayesian statistics the roles of data and model are reversed in the sense that the data set is fixed and the model that possibly gave rise to the sample is uncertain. To illustrate this, we apply the method to the cadmium toxicity data from Table 3. Figure 2 is a plot of different normal PDFs that might have generated the data. The seven data points are displayed as a dot diagram. It is

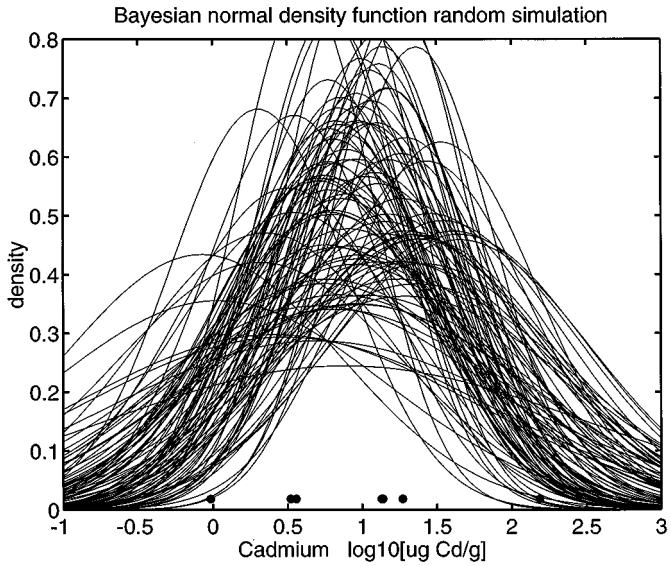

FIG. 2. Bayesian normal density spaghetti plot: random sample of 100 normal probability density functions (PDFs) drawn from the posterior distribution of μ and σ, given 7 cadmium NOEC toxicity data (dots) from Table 3.

intuitively obvious that normal density curves with a mean far below, or above, the set of points have a low probability to be the model behind the data. Similarly, normal PDFs with a mean within the reach of the data, but with a very narrow standard deviation, or extremely wide standard deviation, are unreasonable too. The Bayesian calculation generates a joint probability distribution for μ and σ that gives quantitative expression to this reasonability. It labels each possible μ and σ pair, or *model*, with a weight (probability). The plot in Fig. 2 represents a random sample of possible normal density curves. We call this type of plot a "spaghetti plot" of PDFs.

Figure 2 indicates where the normal PDF values are to be found: in fact, when reading off the possible normal density values at, e.g., -1, 0, and 1, we observe that the density values at one particular point in the vertical direction are *themselves distributed*. Outside the reach of the data points, high density values are unlikely. In the middle of the data, low density values are unlikely, but higher density values than the highest curves are unlikely too, since these would imply too small a standard deviation. Apparently, Bayesian inference generates *secondary probability distributions*: distributions of distributions (cf. Burmaster and Wilson, 1996, Burmaster and Thompson, 1998). This is implicit in the Bayesian philosophy, where in principle the parameters of the distributions employed are themselves distributed.

A more technical way to acknowledge the distributed nature of the spaghetti bundle goes as follows. The well-known formula of the normal density curve as a function of μ and σ reads

$$f(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\{-(x-\mu)^2/(2\cdot\sigma^2)\}. \qquad (8)$$

For given x, the density value f is obviously a function of μ and σ, in that the formula assigns a *single* PDF value to f for each μ and σ, albeit a density value qua interpretation. But, if μ and σ are distributed, the formula defines a *transformation* of two stochastic variables μ and σ. Hence, f, at the given x, is distributed. The distribution of f is fully determined, as soon as the (joint) distribution of μ and σ is specified. This is the case with a prior or posterior probability.

The various methods for characterizing a statistical distribution (mean, mode, percentiles, etc.) do also apply to secondary distributions. To sharpen terminology a bit, we call the distribution at given x (logarithmic concentration) a secondary distribution, because this latter distribution refers to the uncertainty with respect to a probability density value. By connecting, for example, the medians, or modes, or 95th percentiles, of the secondary distributions for each x, one obtains curves characterizing the uncertainty of the primary probability density functions. These new curves illustrate the secondary uncertainty about the model of the toxicity data. Figure 3 represents the 5th and 95th percentile curves, as well as the median curve of the secondary "normal" distribution, for the cadmium data of Table 3. The outer curves constitute a 90% Bayesian confidence interval for (primary) normal density values. The outer curves are not

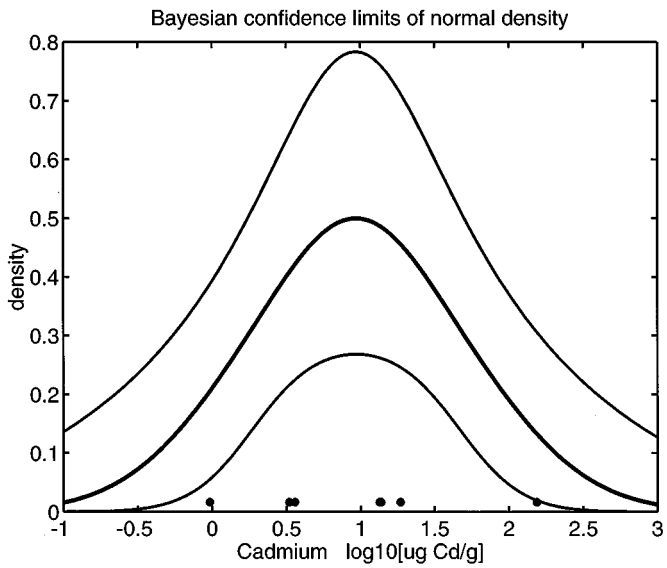

FIG. 3. Bayesian confidence limits of normal PDFs: percentiles (5th, 50th, and 95th) of posterior normal probability density function values for cadmium (Table 3).

PDFs, since they do not integrate to one. The median curve resembles a normal distribution, but probably is not exactly so. This needs further numerical analysis.

4.2. Cumulative Distribution Function and Its Confidence Limits

Operationally, looking at Fig. 2, a secondary distribution can be represented by a random sample of individual normal density curves. Each density curve defines a cumulative distribution function or CDF (Fig. 4). CDFs are useful for two things: first, for reading off how much probability is located below a certain x value; and second, for reading off percentiles. The HC for 5% of the species for a certain toxicant is defined as the fifth percentile of the species sensitivity distribution. The FA is the CDF value at a given logarithmic (environmental) concentration.

Analogous to Fig. 3, one can calculate 5th, 50th, and 95th percentiles of the vertical distribution of the cumulative curves, and connect them (Fig. 5). The individual cadmium data points from Table 3 are now plotted cumulatively. There is confusion in the literature on how to do this. Some plot at $i/(n+1)$, others at $(i-0.5)/n$. The authors find the latter one more convincing: since the empirical CDF for the sample is a staircase shaped function, with discontinuities at the sample data points, one can resolve the discontinuities by plotting the dots at the mean of the left and right empirical cumulative values, which amounts to plotting at $(i-0.5)/n$. This is a compromise between i/n, when counting from the left, and $(i-1)/n$, when counting from the right. Hence, the first cadmium point (Table 3) of 7 is plotted at $(-0.013,0.5/7) = (-0.013,0.071)$. The second point is at $(0.522,1.5/7) = (0.522,0.214)$, and so on.

One may transform the vertical axis by applying the inverse normal CDF to values of FA. This is the rationale behind normal probability graphing paper, on which each normal CDF becomes a straight line. If we do so (Fig. 6), the median curve of the cumulative values indeed seems

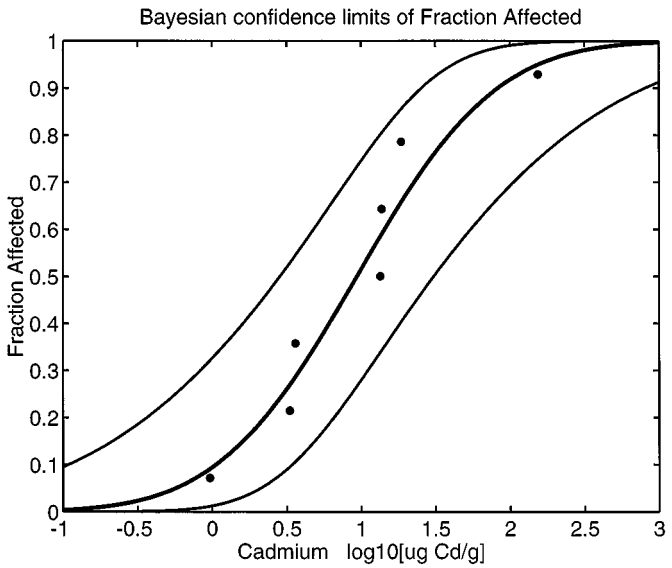

FIG. 5. Bayesian confidence limits of the fraction affected: percentiles (5th, 50th, and 95th) of posterior normal CDFs for cadmium (Table 3). Data plotted cumulatively at $(i-0.5)/n$, with i rank order, and n the number of species tested.

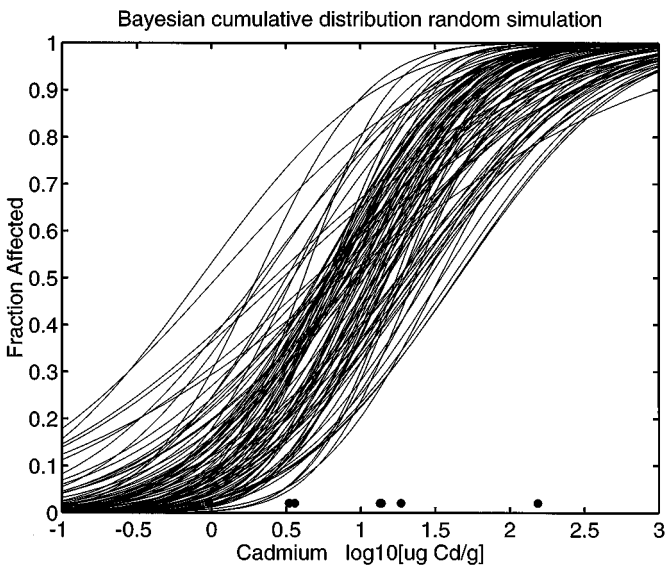

FIG. 4. Bayesian cumulative distribution spaghetti plot: cumulative distribution functions (CDFs) corresponding to the posterior normal PDFs of Fig. 2, plotted over the same cadmium data (Table 3).

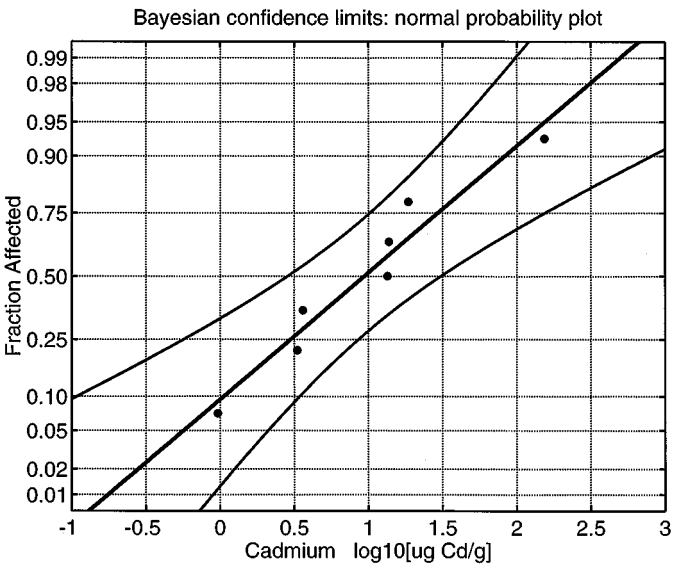

FIG. 6. Bayesian confidence limits of the fraction affected: normal probability plot of Fig. 5.

indistinguishable from a straight line, which would be surprising given the way it is calculated. Numerical analysis indeed reveals that the line is *not* an exact straight line, and, henceforth, the median curve is not the CDF of a (primary) normal distribution, but approximately so. The Bayesian confidence limits now get a familiar-looking hollow shape. The linear measure on the vertical axis is just $-K_p$.

The confidence limits obtained this way have nothing to do with the confidence of the least-squares line fit on probability paper, minimizing the "error" around the curve. The Bayesian confidence limits refer to the variation of the points themselves, and hence are broader.

4.2.1. Slicing: distribution of HC and FA and their percentiles. Reading off HCs from the simulated normal CDFs (one for each random curve) in Fig. 4 amounts to a *horizontal slice* of the cumulative spaghetti bundle, e.g., at FA = 0.05 for the HC_5. This generates a point set of log(HC)s that expresses our uncertainty with respect to its value. By increasing the number of curves indefinitely, a continuity of density curves and cumulatives result, and the point set of log(HC)s melts together to form a continuous distribution of the log(HC_5).

Figure 7 presents a continuous version, i.e. density function, of the log(HC_5). It is again a primary distribution, and, in the logarithm, it is skewed to the left. It has percentiles, a mode, moments, and so on. After taking anti-logarithms, the 5th percentile of the HC_5 distribution, the median value, and the 95th percentile of the HC_5 distribution for cadmium (Table 3) are 0.038, 0.568, and 2.111 μg Cd/g, respectively. Except for the third decimal in the upper limit, these values are identical to the classical confidence limits calculated earlier on the basis of Table 1. The Bayesian values, however, are simulated by evaluating the posterior probability of μ and σ over a fine grid of μ and σ combinations. *The two opposing statistical philosophies, Bayesian and classical, and vastly different numerical approaches, for that matter, seem to lead to the same answer for the normal (Gaussian) model.* The mathematical equivalence is proved in Appendix Sections 8.1.1 and 8.1.2.

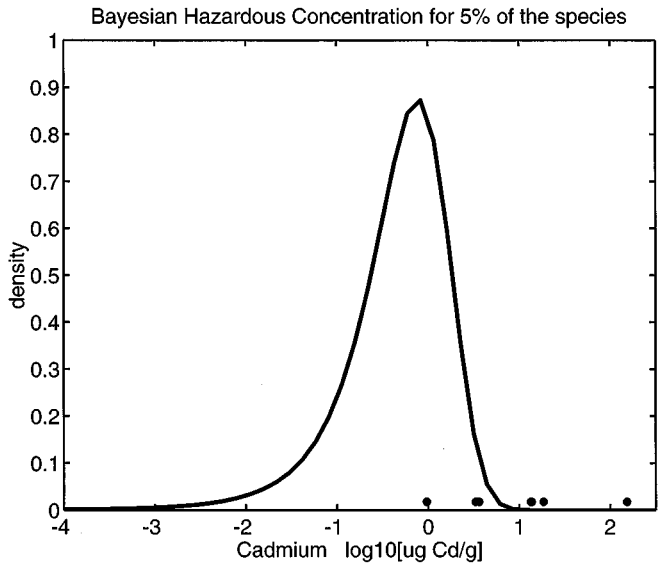

FIG. 7. Bayesian posterior probability density of the log hazardous concentration for 5% of the species for cadmium (Table 3).

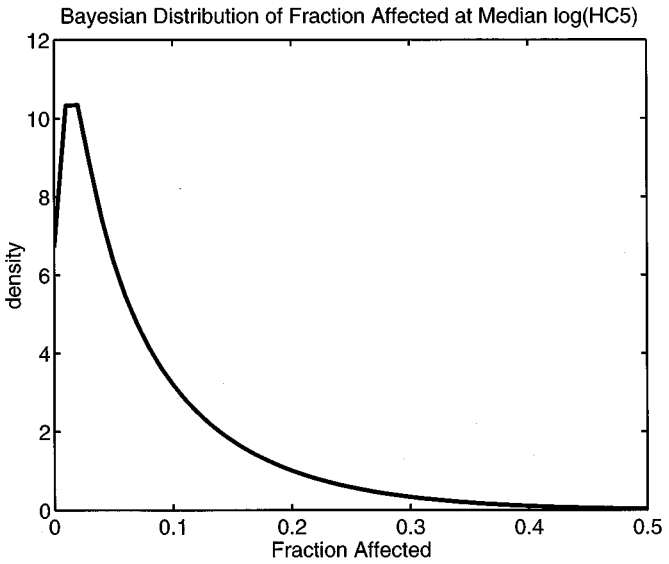

FIG. 8. Bayesian posterior probability density of the fraction affected at median log(HC_5) for cadmium (Table 3).

One can also make *vertical slices* of the simulated normal CDFs (Fig. 4). This is the uncertainty assessment of the FA at given (environmental) logarithmic concentrations. One might expect FA to be symmetrically distributed at median concentration values, while at low values, e.g., around the HC, the FA seems skewed to the right (i.e., to high values). Similarly, at highly polluted sites, FA values will be skewed to the left, which means that not all of the species may be affected. The FA distributions are again primary probability density functions and quantify our uncertainty of FA on the basis of a sample of toxicity data.

As an example, the proposed quality objective of 0.800 μg Cd/g (Table 3) is near the median estimate of the HC_5 (0.568 μg Cd/g). We first examine the posterior distribution of the FA at this median HC_5, since the intention is to protect 95% of the species, and we want to know how certain we are about the FA, especially at this estimate of the HC. At $\log_{10}$ 0.568 μg Cd/g, the vertical slice of the CDF is converted to a continuous density function and plotted in Fig. 8. The x axis of Fig. 8 is the y axis of Figs. 4 and 5. One observes how skewed to the right the distribution of FA is at this given concentration, with a tail extending to 40% and higher. The percentiles (5th, 50th, and 95th) of the FA are 0.3, 5.0, and 25.0%. We observe that the median FA indeed

is equal to 5.0%, but the 90% Bayesian confidence interval for the FA runs from 0.3% up to 25.0%.

The results for the proposed quality objective of 0.800 μg Cd/g (Table 3) are obviously similar to those in Fig. 8. The median FA is 7.4%, i.e., 2.4% higher than 5%. The 90% Bayesian two-sided confidence interval is (0.8%, 29.4%). These values are again identical to those calculated for the confidence approach (Section 3.3), as they should be on mathematical grounds. However, the interpretation differs from that in the confidence approach: in Bayesian parlance, one can speak of the (posterior) probability distribution of FA at the given concentration, without reference to an infinity of possible samples, which is the reasoning behind the confidence limits.

4.2.2. Equivalence of HC and FA limits: the inverse van Straalen Method. One might ask whether the percentile bounds on the CDF result into separate plots for the extrapolation or quality objective setting (horizontal slicing) and for the risk assessment at given environmental concentrations (vertical slicing). The answer is that one needs only one set of curves, and that each percentile curve serves both purposes. This is illustrated in Fig. 9.

Figure 9 is similar to Fig. 5 with two modifications. The first is that we have standardized the log concentration on the mean and standard deviation of the toxicity data. The median FA curve passes through the point (0, 0.5). Second, a "cross-hair," as on display devices, is added at FA $= 0.05$ (horizontal) and the median log(HC_5) (vertical). The proof of the equivalence of limits follows from the fact that CDFs are increasing functions. This means that simulated curves, as in Fig. 4, that cut the horizontal slice to the left of the median HC_5 (they are 50% in number) also cut the vertical slice above FA $= 0.05$, albeit not in the same order, i.e., intermingled. This holds for any cross-hair at any percentile point. Hence, one set of limits suffices for both horizontal and vertical assessments for any percentile or confidence level. Since, the Bayesian and confidence treatments are numerically equivalent, one may view Fig. 9 as a graphical illustration of the confidence argument for fractions affected (Section 3.2) as well.

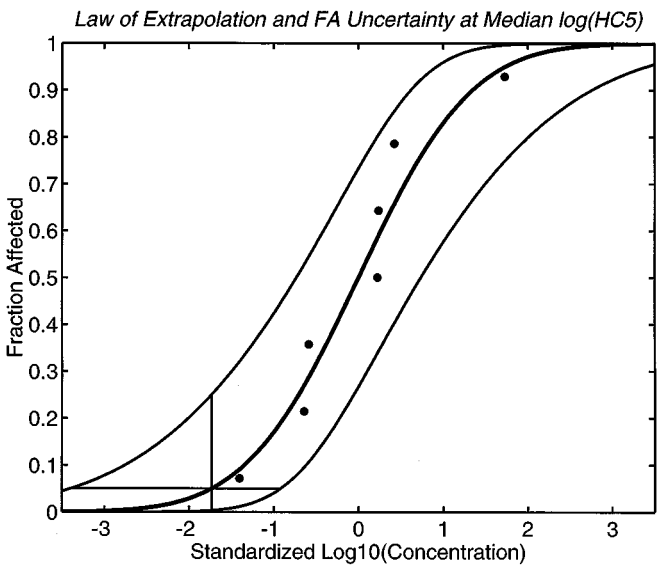

FIG. 9. Bayesian confidence limits of the fraction affected as in Fig. 5, as a function of standardized logarithmic concentration, scaled to the mean and standard deviation of the toxicity data set. Cross-hair shows the equivalence of confidence limits for fraction affected and hazardous concentration, as captured in the *law of extrapolation* (12) and Table 5. Vertical line is the uncertainty of the fraction affected at median log(HC_5). Horizontal straight line cuts CDF curves at minus the extrapolation factor values. Confidence limits and cross-hair hold for all $n = 7$ data sets.

The equivalence of HC and FA confidence limits has an important consequence, which is implicit in the Inverse Van Straalen method (Van Straalen and Denneman, 1989). We have seen that the FA percentiles at the median HC_5 for cadmium were 0.3, 5.0, and 25.0% for cadmium (Table 3). The median FA at the median HC_5 is exactly 5.0%, which it should be according to the cross-hair argument of equivalence of limits, just given. However, if one would evaluate the percentiles of FA at the *lower* limit of the HC_5, then one would obtain for the three percentiles 0.0, 0.1, and 5.0%. So, now the *upper* limit of the FA equals exactly 5.0%, which is once again the cross-hair argument. Analogously, at the upper limit of the HC_5, the lower estimate of the FA is exactly 5.0%, the median FA now is 19.0%, and the upper FA is 44.6%.

Table 4a summarizes FA uncertainties at lower, median, and upper HC_5 for cadmium (Table 3, $n = 7$). The main diagonal is filled with the 5.0% values printed in boldface which means that, at the lower HC_5, 5% of the species is affected *at the most* (upper confidence estimate). The median HC_5 yields a median 5% FA. The upper HC_5 *at least* affects 5%. Hence, the implicit conjecture of Van Straalen and Denneman (1989) to use a lower extrapolation constant to estimate an upper FA is indeed justified and can be made rigorous. In Section 5.1, this result is formulated as the *law of extrapolation*, and Fig. 9, horizontal slice, is a graphical illustration of it.

TABLE 4a

Confidence Limits (95, 50, 5%) of the Fraction Affected (%) at Similar Confidence Limits of the Log Hazardous Concentration for Cadmium from Table 3, n=7

Upper FA	**5.000**	25.009	44.576
Median FA	0.065	**5.000**	18.983
Lower FA	0.000	0.341	**5.000**
Log_{10} (μg Cd/g)	− 1.41779	− 0.24580	0.32442
Log hazardous concentation	Lower	Median	Upper

TABLE 4b
Confidence Limits (95, 50, 5%) of the Fraction Affected (%) at Similar Confidence Limits of the Log Hazardous Concentration in Standardized Logarithmic Concentrations, $n=7$

Upper FA	**5.000**	25.009	44.576
Median FA	0.065	**5.000**	18.983
Lower FA	0.000	0.341	**5.000**
Standardized log hazardous concentration	− 3.3995 Lower	− 1.7318 Median	− 0.9204 Upper

The standardization on the mean and the standard deviation points to something more. Table 4b is just Table 4a but now for *standardized* log concentrations. From the considerations in Sections 3.1 and 3.2, it is known that these standardized log concentrations are exactly the extrapolation factors from Table 1 with a minus sign. Hence, Table 4b is *universal* in that it not only holds for cadmium, but for *all* $n = 7$ *toxicity data sets*, provided the model is valid for the substances considered. But this means that Fig. 9 (or its normal probability plot form), when plotted on a standardized scale, is equally valid for all $n = 7$ data sets as well. By removing the cadmium data points from it, one would obtain just one plot, and therefore one set of confidence limits on FA that is sufficient for all $n = 7$ cases. This is a Bayesian justification of the validity of Tables 1 and 2, independent of the substance involved.

5. FRACTION AFFECTED AT HAZARDOUS CONCENTRATIONS

5.1. The Law of Extrapolation

Table 5 tabulates the FA (%) at lower, median, upper estimates of log(HC_5). The bounds on the FA are a function of sample size only. Lower and upper limits are 5%, and 95% confidence limits (one-sided). Together they make up a 90% confidence interval for the FA. From the Bayesian viewpoint, the FA limits are percentiles of posterior FA probability distributions, and they are numerically equal to the confidence limits from sampling theory in classical statistics.

The FA limits in Table 5 are independent of mean and standard deviation of the toxicity data, and therefore *universal* for each sample size, provided the model is valid for a substance. This is because HCs are calculated by subtracting the extrapolation factor times the toxicity data standard deviation from the toxicity data mean. But the estimates of the FA were also corrected for the mean and standard deviation of the toxicity data. Hence, Table 5 is independent of the mean and standard deviation of the toxicity data.

TABLE 5
Confidence Limits (5, 50, 95%) of the Fraction Affected (%) at Standardized Lower, Median, and Upper Estimates (95, 50, 5% Confidence) of Log (HC_5) and Toxicity Data Sample Size[a]

Sample size	log (HC_5) (Lower/Median/ Upper)	FA^{Lower}	FA^{Median}	FA^{Upper}
	− 26.2597	0.000	0.000	**5.000**
2	− 2.3387	0.000	**5.000**	62.626
	− 0.4748	**5.000**	35.491	80.625
	− 7.6559	0.000	0.000	**5.000**
3	− 1.9384	0.016	**5.000**	46.239
	− 0.6391	**5.000**	28.780	68.772
	− 5.1439	0.000	0.000	**5.000**
4	− 1.8295	0.073	**5.000**	37.144
	− 0.7433	**5.000**	24.862	60.055
	− 4.2027	0.000	0.006	**5.000**
5	− 1.7793	0.154	**5.000**	31.526
	− 0.8178	**5.000**	22.264	50.546
	− 3.7077	0.000	0.026	**5.000**
6	− 1.7505	0.246	**5.000**	27.737
	− 0.8748	**5.000**	20.398	48.541
	− 3.3995	0.000	0.065	**5.000**
7	− 1.7318	0.341	**5.000**	25.009
	− 0.9204	**5.000**	18.983	44.576
	− 3.1873	0.000	0.118	**5.000**
8	− 1.7187	0.435	**5.000**	22.949
	− 0.9580	**5.000**	17.868	41.360
	− 3.0312	0.001	0.181	**5.000**
9	− 1.7091	0.526	**5.000**	21.336
	− 0.9899	**5.000**	16.962	38.696
	− 2.9110	0.002	0.249	**5.000**
10	− 1.7016	0.612	**5.000**	20.036
	− 1.0173	**5.000**	16.209	36.453
	− 2.8150	0.005	0.321	**5.000**
11	− 1.6957	0.695	**5.000**	18.964
	− 1.0413	**5.000**	15.571	34.537
	− 2.7363	0.008	0.394	**5.000**
12	− 1.6910	0.774	**5.000**	18.064
	− 1.0625	**5.000**	15.023	32.879
	− 2.6705	0.014	0.466	**5.000**
13	− 1.6870	0.849	**5.000**	17.296
	− 1.0814	**5.000**	14.545	31.430
	− 2.6144	0.021	0.537	**5.000**
14	− 1.6837	0.920	**5.000**	16.633
	− 1.0985	**5.000**	14.125	30.151
	− 2.5660	0.029	0.607	**5.000**
15	− 1.6808	0.987	**5.000**	16.053
	− 1.1140	**5.000**	13.752	29.013
	− 2.3960	0.093	0.922	**5.000**
20	− 1.6712	1.279	**5.000**	13.978
	− 1.1746	**5.000**	12.363	24.788
	− 2.2198	0.278	1.404	**5.000**
30	− 1.6620	1.703	**5.000**	11.785
	− 1.2498	**5.000**	10.796	20.077
	− 2.0650	0.671	2.008	**5.000**
50	− 1.6549	2.227	**5.000**	9.857
	− 1.3294	**5.000**	9.316	15.767
	− 1.9265	1.383	2.739	**5.000**
100	− 1.6498	2.870	**5.000**	8.170
	− 1.4143	**5.000**	7.925	11.925
	− 1.8372	2.135	3.329	**5.000**
200	− 1.6473	3.402	**5.000**	7.116
	− 1.4778	**5.000**	7.003	9.542
Inf	**− 1.6449**	**5.000**	**5.000**	**5.000**

[a] The 5.000% values illustrate the *law of extrapolation* (12). Upper and lower estimates constitute a 90% two-sided confidence interval.

We can summarize the diagonal terms (**5.000%**) of Table 5 into three formulas:

$$\mathrm{FA}^{\mathrm{Upper}}\,(\log(\mathrm{HC}_p^{\mathrm{Lower}})) = p\% \tag{9}$$

$$\mathrm{FA}^{\mathrm{Median}}\,(\log(\mathrm{HC}_p^{\mathrm{Median}})) = p\% \tag{10}$$

$$\mathrm{FA}^{\mathrm{Lower}}\,(\log(\mathrm{HC}_p^{\mathrm{Upper}})) = p\%. \tag{11}$$

Note that the confidence levels (classical) in Eqs. (9), (10), and (11) are 95, 50, and 5%, respectively.

In words: *the upper* (*median*, *lower*) *confidence limit of the fraction affected at the lower* (*median*, *upper*) *confidence limit of the log* HC_p *is equal to the fraction affected p% used to define the hazardous concentration.* These relationships hold for *any* confidence level γ, being the same for both estimates of the HC and FA, and for *all* protection levels $1 - p\%$. This is called the *law of extrapolation.*

With FA^{γ} the γ-th percentile of the posterior PDF of FA, and $\mathrm{HC}_p^{1-\gamma}$ the $(1-\gamma)$-th percentile of HC_p likewise, the Bayesian version of the *law of extrapolation* reads

$$\mathrm{FA}^{\gamma}\,(\log(\mathrm{HC}_p^{1-\gamma})) = p\%. \tag{12}$$

5.2. *Uncertainty of the Fraction Affected at the Median Log Hazardous Concentration*

Table 5 also gives the off-diagonal terms. The upper FA at median $\log(\mathrm{HC}_5)$ is particularly interesting. The latter is used in policy-making. The median FA at median $\log(\mathrm{HC}_5)$, 5.000%, reveals that the objective is met, which is just the law stated above. But, the skewness of the FA, indicated by the upper limit, reveals that a considerably *larger* fraction of species may be affected at this concentration. At a toxicity data size of $n = 3$, which is not at all the exception, it is 46%, and all one can say is that with a 19 to 1 bet the FA is not larger than 46% in that case. At $n = 10$, the upper FA is 20%, while at a toxicity data sample size of 30, it is still 12%.

6. DISCUSSION

The normal or Gaussian distribution model for the variation of species sensitivity to a toxicant is one of the simplest thinkable. It has very little structure other than describing the location and scale of the species toxicity on a logarithmic toxicity endpoint axis. However, this lack of structure has some important consequences, both advantageous and less so.

One of the conceptual problems of this so-called distributional extrapolation method on the basis of the SSD is what set of species it refers to and whether it can be considered to be a random sample of that set or of some larger group of species of ecological relevance. The method has been criticized, especially when it is used as an ecosystem approach (Smith and Cairns, 1993; Forbes and Forbes, 1993; Hopkin, 1993). Issues of criticism are the lack of any structure in the model accounting for interspecies relationships and functioning of the ecosystem, the problem of bioavailability of the toxicant, and methodological questions of a more statistical nature: distributional assumptions, the percentile estimated, the method of assessing the uncertainty of the percentiles and of the fraction affected.

It is imperative that this model be extended to deal with one or more of these issues. All extensions have to cope with the trade-off between uncertainty and complexity. That means that the present model can be considered the simplest in a hierarchy of increasingly complex ones. One way of incorporating more structure is adding explanatory variables in order to reduce variability. Therefore, it is important to acknowledge that these models have to be statistical in nature. However, there are at least two major statistical philosophies around.

The estimation of hazardous concentrations, i.e., percentiles, and of fractions affected, i.e., cumulative values, can be studied from different statistical perspectives, i.e., classical statistics and Bayesian statistics, that employ divergent philosophies with regard to the uncertainty and fixedness of both the model and the sample. There is a vast literature, both statistical and philosophical, on these differences in viewpoints. However, in the case of this distributional extrapolation model, there is no need to get involved in debating foundational matters between the two major statistical philosophies, because from a pragmatic viewpoint, the numerical answers are identical. Generally speaking, this is more the exception than the rule, but due to the simplicity of the extrapolation model there is no conflict between outcomes.

For those who feel more confident with the classical approach, and find the Bayesian approach controversial to some extent, it is advantageous that both approaches yield the same answers in this simplest case. This might aid in developing confidence in the Bayesian approach as well. The interpretation of the two methods is different, however.

The classical, or confidence, method is the one most taught at university courses, although that may change. This is because the Bayesian approach is the more easily understood statistical philosophy, both conceptually as well and numerically. Many scientists have difficulty interpreting a confidence interval in the appropriate way, and a common misinterpretation, as the probability that a parameter is between certain limits, is exactly the right one from the Bayesian standpoint.

However, apart from this pedagogical aspect (cf. Lee, 1989, preface), there is a more technical reason to emphasize the Bayesian approach more than the confidence approach. That is, the Bayesian approach is the more powerful one eventually for extending the model into directions necessary to deal with the weaknesses of the present model. These are relaxation of distributional assumptions (e.g., the questions

of sensitive and insensitive species, or target versus nontarget), the incorporation of interspecies relationships, treating aspects of availability of the toxicant, nutritious or intoxicating effects, incorporating the uncertainty or variation of the environmental concentration, and so forth. We think that the conceptual device of an infinite repetition of samples, as in the sampling statistics viewpoint, does not yield enough power to accomplish these extensions, while the Bayesian stance is very general indeed.

In the present model, based on the normal distribution, the mean and standard deviation of the toxicity data play a major role. These are so-called sufficient statistics for the model. The implication is that the method does only use these two particular aspects of information in the sample, and not the information represented by the individual points. The advantage is that the raw data themselves are not needed (only mean and standard deviation), while the disadvantage is that the method may fail to indicate departures of the model due to skewness, or bimodal or multimodal distributions, or other causes.

The role mean and standard deviation of the toxicity data play is also manifest in the results. It yields extrapolation factors to be applied to these statistics that depend only on sample size and confidence level. Hence, they are easily tabulated. For the inverse application of the method, i.e., estimating the fraction affected at given (environmental) concentration, the same is true. It is only important what the value of the log concentration standardized is with respect to the mean and the standard deviation of the toxicity data. Thus, uncertainty estimates of the fraction affected can be tabulated analogously. By coupling both estimates, one can evaluate the uncertainty of the fraction affected at uncertainty estimates of hazardous concentrations. These final uncertainties are independent of the mean and standard deviation of the toxicity data, and hence universal. We observe how much uncertainty discounts the intention to not affect the majority of species, especially in the case of small data sets.

The great power of acknowledging the possibility of inverse application of the extrapolation procedure is not that one can estimate the fraction affected at given concentrations next to, and in concordance with, quality objective setting, but *instead of* quality objective setting. As Van Straalen and Denneman (1989, p. 250, third paragraph) state (italics ours): "The problem of choosing a level of protection for advisory values for environmental quality can be *circumvented* by presenting the results of risk analysis in the form of [the inverse method]." Hence, instead of setting a quality objective and asking what effect to expect at concentrations around or above the standard, for example through the methods of risk quotients, one may assess the risk of harming a fraction of the set of species considered directly through the inverse application from the toxicity data themselves.

A fixed environmental quality objective may mask the inherent uncertainties of the estimate behind it. If one decouples extrapolation and risk assessment in this way, it may vastly underestimate the risk of affecting more than a small fraction of the species set considered.

7. CONCLUSIONS

The uncertainty of the normal (Gaussian) distribution of variability is calculated using both Bayesian and classical statistics. It has been applied to SSDs for a toxicant, through calculation of the FA and the HC and their confidence intervals. This method allows for assessing both FA and HC in one mathematical framework. The Bayesian and classical results turn out to be identical in this case. The upper (median, lower) confidence limit of the FA at the lower (median, upper) confidence limit of the log HC is equal to the FA used to define the HC. The extrapolation factors that allow calculation of HC through mean and standard deviation are tabulated as a function of sample size for several levels of FA. Confidence limits of the FA at given standardized environmental concentrations were also tabulated by correcting for the mean and standard deviation of the original toxicity data.

8. APPENDIX

8.1. Theory: The Noncentral t Distribution

8.1.1. Classical statistics. For the normal distribution extrapolation factors can be calculated from the noncentral t distribution (Owen, 1968, Wagner and Løkke, 1991). From the viewpoint of sampling theory in classical statistics, one wants estimator (A2) to underestimate the true value (A1) with probability γ:

$$\log(\mathrm{HC}_p) = \mu - K_p \cdot \sigma, \tag{A1}$$

$$\log(\hat{\mathrm{HC}}_p) = \bar{x} - k_s \cdot s, \tag{A2}$$

hence,

$$\Pr(\bar{x} - k_s \cdot s \le \mu - K_p \cdot \sigma) = \gamma. \tag{A3}$$

The probability refers to repeated samples and γ is the confidence level. After reordering, one obtains

$$\Pr\left(\left[\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + K_p \cdot \sqrt{n}\right] \Big/ [s/\sigma] \le k_s \cdot \sqrt{n}\right) = \gamma. \tag{A4}$$

The left-hand side of the inequality is a random variable with a noncentral t distribution (Owen, 1968). This follows

from the fact that the random variable

$$W = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \tag{A5}$$

has a standard normal distribution, $\delta = K_p \cdot \sqrt{n}$ is a constant, and the random variable $V = (n-1)s^2/\sigma^2$ has a χ^2 distribution with $(n-1)$ degrees of freedom. Note that $\sqrt{n-1} \cdot s/\sigma$ has a χ distribution with $(n-1)$ degrees of freedom (cf. Box and Tiao, 1973, p. 88), and s/σ, the denominator of (A4), is distributed as $\sqrt{V/(n-1)}$, i.e., the square root of a χ^2 distribution with $(n-1)$ degrees of freedom divided by the degrees of freedom.

In fact, any random variable of the form

$$\frac{W + \delta}{\sqrt{V/r}}, \tag{A6}$$

with W standard normal, δ a constant (called the noncentrality parameter), and V a χ^2 distribution with r degrees of freedom, has a noncentral t distribution. For $\delta = 0$, the noncentral t-distribution becomes Student's t distribution.

Since (A4) is a cumulative probability statement with regard to the noncentral t distribution, one needs the inverse cumulative noncentral t distribution to solve for k_s, given K_p, which follows from the protection level, at a given probability (level of confidence) γ. One may also compute k_s from a given concentration and solve for K_p.

8.1.2. Bayesian statistics. The authors will prove that in the Bayesian case the noncentral t distribution shows up in a way similar to the confidence approach. The starting point is the equation

$$\Pr(\bar{x} - k_s \cdot s \le \mu - K_p \cdot \sigma) = \gamma. \tag{A3*}$$

From the Bayesian viewpoint, the probability refers to the joint posterior distribution for μ and σ on the basis of a noninformative prior for μ and σ. Upon rewriting, one obtains

$$\Pr\left(\left[\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + K_p \cdot \sqrt{n}\right] \Big/ [s/\sigma] \le k_s \cdot \sqrt{n}\right) = \gamma, \tag{A4*}$$

which is identical to (A4) qua notation, but the probability refers to the random variables μ and σ, while $\bar{x}$ and s are fixed numbers calculated from the sample. The random variable on the left of the inequality will again be noncentral t distributed, as will now be shown.

Since

$$\frac{\sigma}{\sqrt{n-1} \cdot s}$$

has an inverted χ distribution with $(n-1)$ degrees of freedom (Box and Tiao, 1973, p. 96).

$$\frac{\sqrt{n-1} \cdot s}{\sigma}$$

has a χ distribution with $(n-1)$ degrees of freedom. Henceforth, the denominator s/σ in (A4*) is distributed as $\sqrt{V/(n-1)}$, with V a χ^2 distribution with $(n-1)$ degrees of freedom.

The random variable

$$W = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \tag{A5*}$$

in the numerator of (A4*) can be written as

$$W = \frac{\mu - \bar{x}}{s/\sqrt{n}} \cdot \frac{s}{\sigma}.$$

The first factor in W

$$T = \frac{\mu - \bar{x}}{s/\sqrt{n}}$$

has a Student t distribution with $(n-1)$ degrees of freedom (Box and Tiao, 1973, p. 97), while s/σ, the second factor in W, is distributed as derived above for the same term in the denominator. Thus,

$$W = T \cdot \sqrt{V/r},$$

with V a χ^2 distribution with $r = (n-1)$ degrees of freedom. The textbook definition of Student's t distribution (e.g., Hogg and Craig, 1978), however, is

$$T = \frac{W}{\sqrt{V/r}},$$

with W the standard normal distribution and V χ^2. Thus, W in (A5*) for the Bayesian case is standard normal.

This means that the random variable at the left side of the inequality in (A4*) is of the form (A6) and therefore has a noncentral t distribution. Henceforth, the relationships between triples of K_p, k_s, γ, as defined by the cumulative noncentral t distribution in either (A4) or (A4*) are identical from both the Bayesian and the sampling points of view.

8.2. Numerical Aspects

Two numerical approaches are followed. One is by using functions for evaluating the cumulative noncentral

t distribution in commercial mathematical software packages (Section 8.2.1). The other is by simulating the posterior probability distribution of μ and σ (Section 8.2.2). From the previous section (8.1), it follows that the results should match numerically. Section 8.2.3 outlines Aitken's repeated linear interpolation scheme for estimating results at intermediate table entries.

8.2.1. The noncentral t distribution in MATLAB and Mathematica. From Section 8.1 it follows that we need values of the inverse cumulative noncentral t distribution to solve for k_s given K_p and confidence level γ. We used the MATLAB Statistics Toolbox 2.0 (Jones, 1996), as well as the Continuous Distributions package in the Standard Add-on packages (Martin, 1996) including with *Mathematica* 3.0. Unfortunately, neither is infallible. The MATLAB routine `nctcdf( )` may give erroneous answers for some small n, and `nctinv( )` has convergence problems for large n. The *Mathematica* routines based on the **NoncentralStudentTDistribution**[ ] function may not converge for large n. The tables were all calculated with MATLAB. All entries were checked with Odeh and Owen (1980, Tables 7.1 and 7.4). The figures were drawn on the basis of Bayesian simulation implemented in MATLAB as well.
The numerical approach in MATLAB was:

A function to solve for k_s:

```
function zero = fzksnct (ks, n, conf, kp)
x = ks*sqrt (n);
c = kp*sqrt(n);
df = n - 1;
zero = nctcdf (x,df,c) - conf;
```

Set the protection level and invoke the root finder:

```
kp = - norminv (0.05)
ks = fzero ('fzksnct', [kslo,kshi], accuracy, output,
      n, conf, kp)
```

A function to solve for K_p:

```
function zero = fzkpnct (kp, n, conf, ks)
x = ks*sqrt(n);
c = kp*sqrt(n);
df = n - 1;
zero = nctcdf (x, df, c) - conf;
```

Invoke the root finder:

```
kp = fzero ('fzkpnct', [kplo, kphi] , accuracy, output,
      n, conf, ks)
```

and calculate the fraction affected:

```
fa = 100*normcdf ( - kp)
```

The numerical approach in *Mathematica* was

```
kp = - Quantile [NormalDistribution [0, 1], 0.05]
```

Define the noncentral t distribution as a function:

```
nct[n_, kp_] := NoncentralStudentTDistribution
                [n - 1, kp Sqrt [n]]
```

Solve for k_s through

```
ks = Quantile [nct [n, kp], conf]/Sqrt [n]
```

Solve for K_p by invoking the root finder and calculate the fraction affected:

```
kprule = FindRoot [CDF [nct [n, kp], ks Sqrt[n]]
      = = conf, {kp, {kplo, kphi}}]
fa = 100 CDF [NormalDistribution [0,1], - kp/.kprule]
```

8.2.2. Evaluating the posterior distribution of μ and σ. In handling the Bayesian approach numerically, the authors also simulated the posterior distribution for μ and σ directly, without any reference to the noncentral t distribution.

Suppose we have a random sample of n independent observations from the normal distribution $N(\mu, \sigma^2)$, with both μ and σ unknown. The posterior probability distribution for μ and σ (Box and Tiao, 1973, p. 93), given the data vector $\mathbf{x}$ is

$$p(\mu, \sigma \mid \mathbf{x}) = c \cdot \sigma^{-(n+1)} \cdot \exp\{ - [n \cdot (\mu - \bar{x})^2 + (n-1) \cdot s^2] / [2 \cdot \sigma^2] \}.$$

This is just ordinary likelihood with one extra σ for the prior distribution of μ and σ on the basis of a noninformative prior argument. Coefficient c is a constant of proportionality. The data only manifest themselves through the mean, $\bar{x}$, and the sample standard deviation ($n - 1$ version), s. These are sufficient statistics. This means that two different toxicity data sets that have the same mean and standard deviation have the same posterior probability distribution for μ and σ. But there is an even more striking consequence of the invariance of the posterior formula for different mean and standard deviation.

When evaluating the posterior for different μ and σ the question is: what values to take? Of course, mean and standard deviation provide some guidance, but how far should one deviate from these middle values? The natural

choice seems to evaluate μ at the mean $\bar{x}$ plus and minus multiples of the standard deviation s. Similarly, multiples of s and fractions of s appear to be the natural choice for σ. Hence, we consider

$$\mu = \bar{x} + k \cdot s$$

$$\sigma = l \cdot s,$$

where k is either positive, zero, or negative, and l is positive, below 1, equal to 1, or greater than 1. By substituting these into the posterior, the result is

$$p(k, l) = c \cdot s^{-(n+1)} \cdot l^{-(n+1)} \cdot \exp\{ -[n \cdot k^2 + n - 1]/[2 \cdot l^2]\}.$$

If this expression is evaluated over a regular grid of (μ, σ) combinations, and hence (k, l) combinations, the probabilities have to be normalized, that is, added and divided each by the sum. Then, the term

$$c \cdot s^{-(n+1)}$$

cancels, and the posterior is only a function of k, l, and n. However, this has the important consequence that, given the sample size n, all posterior distributions have identical values at similar (k, l) positions in the (μ, σ) plane. That is, all size n extrapolation problems are isomorphic to the single standardized sample problem of size n, with mean zero and standard deviation one. This in turn means that one can tabulate hazardous concentrations and fractions affected as a function of sample size, independent of the actual means and standard deviations, and compare them to the corresponding results of using the noncentral t distribution.

Numerically, the authors have been experimenting with grids and random samples over μ and σ. However, it turns out that the marginal posterior distribution of σ is quite skewed to the right. Hence, the natural choice for gridding or Monte Carlo-ing over μ and σ is to reparameterize to μ and $\ln(\sigma)$. Taking

$$\tau = \ln(\sigma),$$

which implies

$$d\tau = \frac{1}{\sigma}\, d\sigma,$$

the posterior, up to proportionality, is

$$p(\mu, \tau \mid \mathbf{x}) = \exp\{ -n \cdot \tau - [n \cdot (\mu - \bar{x})^2 + (n-1) \cdot s^2]/ [2 \cdot \exp(2 \cdot \tau)]\}.$$

Note that the extra σ due to the prior in the original formula is removed again;

$$\exp\{ -n \cdot \tau\} = \sigma^{-n},$$

and essentially accounted for into the surface increment

$$d\mu \cdot d\tau = d\mu \cdot d \ln(\sigma).$$

Accordingly, in the actual implementation, the authors take the original likelihood formula with σ^{-n}, hence without the extra prior term $1/\sigma$, and grid over equally spaced $(\mu, \ln(\sigma))$ combinations; or, alternatively, a random sample is taken within a finite range of the $(\mu, \ln(\sigma))$ plane, to evaluate the posterior. The regular grid method proved to be superior in improving the precision of answers.

Another, third, alternative to evaluate the posterior is the acceptance–rejection method (Numerical Recipes). This is, numerically speaking, a very inefficient method, since it rejects $(\mu, \ln(\sigma))$ combinations, after evaluation, the more the posterior falls below its majorant, here the maximum value. However, this method yields a random sample from the posterior distribution, leading to the intuitively appealing spaghetti plots.

More sophisticated algorithms, e.g., Metropolis algorithms and the like, were not tried.

In the first two methods used (grid and random combinations), there is a large $(\mu, \ln(\sigma))$ table and its associated posterior value column vector, that after normalization, can be interpreted as probabilities. Such three-column tables are treated as a computer representation of a bivariate discrete probability density function. By dilateing the $\ln(\sigma)$-column values again to σ values with the exponential function, the probabilities carry over without change (for continuous distributions one needs corrections with the Jacobian). Now the σ values will not be regularly, or evenly, distributed any more.

The third method (acceptance–rejection/spaghetti plots) is different in that the distribution of the points themselves mimics the posterior density. The authors turn the selected points into a three-column "discrete" density function, by associating with each entry, $(\mu, \ln(\sigma))$ pair, a probability of 1/(number of points accepted); that is, all points have equal weight. In this way, all three methods lead to similar tables representing discrete probability density functions that can be treated in an analogous manner, e.g., by calculating moments, percentiles, histograms, density estimates, etc.

8.2.3. Aitken's interpolation scheme. The so-called Aitken's method of iterative linear interpolation (Abramowitz and Stegun, 1965, p. xi) can be calculated by hand or in a spreadsheet. The following example illustrates the calculation in Section 3.3 for cadmium.

A		B	C	D	E	F	G	H
1	−1.51994							
2	−2	2.887						−0.48006
3	−1	17.006	9.665					0.51994
4	−3	0.226	4.164	8.235				−1.48006
5	0	50.000	14.196	7.309	7.778			1.51994
6	−4	0.008	3.578	8.610	7.680	7.741		−2.48006
7	1	82.994	15.706	8.095	8.183	7.163	7.454	2.51994

Here, column A refers to the standardized log(*Concentration*). A1 is the input log concentration (in this case for cadmium). A2, A3, …, are the nearest Table 2 column entries in that order. Column B is the median fraction affected at $n = 7$, also from Table 2, corresponding to the entries in column A. C3 is the linear interpolant and D4 is a three-point interpolant. G7 is the six-point interpolant, 7.454, which is near the true answer, 7.421.

This is how it works: column H contains differences of A2:A7 with respect to A1. So, H2 = A2–A1, H3 = A3–A1, etc. The interpolating determinants are C3 = (B2*H3–B3*H2)/(A3–A2), C4 = (B2*H4–B4*H2)/(A4–A2), etc. D4 = (C3*H4–C4*H3)/(A4-A3), D5 = (C3*H5–C5*H3)/(A5–A3), etc.

REFERENCES

Abramowitz, M., and Stegun, I. A. (1965). *Handbook of Mathematical Functions*. Dover Publications, New York.

Aldenberg, T. (1993). *$E_T X$ 1.3a, A Program to Calculate Confidence Limits for Hazardous Concentrations Based on Small Samples of Toxicity Data*. RIVM Report, 719102015, RIVM, Bilthoven, The Netherlands.

Aldenberg, T., and Slob, W. (1993). Confidence Limits for Hazardous Concentrations Based on Logistically Distributed NOEC Toxicity Data. *Ecotoxicol. Environ. Saf.* **25,** 48–63.

Box, G. E. P., and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Addison–Wesley (1992, Wiley Classics, New York).

Burmaster D. E., and Thompson, K. (1998). Fitting second order parametric distributions to data using maximum likelihood estimation. *Human Ecol. Risk Assess.* **4,** 319–340.

Burmaster D. E., and Wilson, A. M. (1996). An introduction to second order random variables in human health risk assessments. *Human Ecol. Risk Assess.* **2,** 892–919.

Forbes, T. L., and Forbes, V. E. (1993). A critique of the use of distribution-based extrapolation models in ecotoxicology. *Funct. Ecol.* **7,** 249–254.

Frey, C., and Rhodes, D. (1998). Characterization and simulation of uncertain frequency distributions: Effects of distribution choice, variability, uncertainty, and parameter dependence. *Human Ecol. Risk Assess.* **4,** 423–469.

Hogg, R. V., Craig, A. T. (1978). *Introduction to Mathematical Statistics*. Macmillan, New York.

Hopkin, S. P. (1993). Ecological implications of '95% protection levels' for metals in soil. *Oikos* **66,** 137–141.

Jones, B. (1996). *Statistics TOOLBOX for use with MATLAB. User's Guide*. The MathWorks, Natick, MA.

Kooijman, S. A. L. M. (1987) A safety factor for LC50 values allowing for differences in sensitivity among species. *Water Res.* **21,** 269–276.

Lee, P. M. (1989). *Bayesian Statistics*: *An Introduction*. Oxford Univ. Press, Oxford. (Second ed., 1997).

Martin, E. (1996). *Mathematica 3.0 Standard Add-on Packages*. Wolfram Research, Champaign, IL.

Mood, A. M., Graybill, F. A., and Boes, D. C. (1974) *Introduction to the Theory of Statistics*. McGraw–Hill, Singapore.

Okkerman, P. C., Van de Plassche, E. J., Emans, H. J. B., and Canton, J. H. (1993). Validation of some extrapolation methods with toxicity data derived from multiple species experiments. *Ecotoxicol. Environ. Saf.* **25,** 341–359.

Odeh, R. E., and Owen, D. B. (1980). *Tables for Normal Tolerance Limits, Sampling Plans, and Screening*. Dekker, New York.

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

Press, S. J. (1989). *Bayesian Statistics*: *Principles, Models, and Applications*. Wiley Interscience, New York.

Smith, E. P., and Cairns J., Jr. (1993). Extrapolation methods for setting ecological standards for water quality: Statistical and ecological concerns. *Ecotoxicology* **2,** 203–219.

Swartjes, F. A. (1999). Risk-based assessment of soil and groundwater quality in The Netherlands: Standards and remediation urgency. *Risk Anal.* **19,** 1235–1249.

Van Straalen, N. M., and Denneman, C. A. J. (1989). Ecotoxicological evaluation of soil quality criteria. *Ecotoxicol. Environ. Saf.* **18,** 241–251.

Versteeg, D. J., Belanger, S. E., and Carr, G. J. (1999). Understanding single-species and model ecosystem sensitivity: Data-based comparison. *Environ. Toxicol. Chem.* **18,** 1329–1346.

Wagner, C., and Løkke, H. (1991). Estimation of ecotoxicological protection levels from NOEC toxicity data. *Wat. Res.* **25,** 1237–1242.