Fact Sheet on
Environmental Risk Assessment

3

Published by the International Council on Metals and the Environment (ICME)

APRIL
2001

Distribution-Based Extrapolation Approaches in the Risk Assessment of Metals in the Environment

by Nico M. Van Straalen
Institute of Ecological Science, Vrije Universiteit, The Netherlands

Discharge limits and clean-up values for environmental protection are based on a determination of how much of a pollutant can be tolerated by highly sensitive species. The goal often has been stated as protection of 95% of the species 95% of the time. This is achieved by looking at the distribution of sensitivities of all tested species and selecting a value that includes 95% of the variability. However, since its introduction at the end of the 1980s, the toxicity threshold distribution approach to environmental risk assessment has been a subject of intensive debate. Nevertheless, several European countries, the U.S.A., Canada, Australia and New Zealand have adopted some form of distribution-based extrapolation approach as the basis for refined risk assessment, in line with the recommendations given by an expert group of the OECD in 1992.

Distribution-Based Sensitivity Extrapolation:

A method for setting environmental protection levels based on scientific understanding of differing sensitivities among species and test conditions.

An important consideration that supported the implementation of these methods was the fact that their outcomes usually are in good accordance with field data and results from multispecies ecotoxicity experiments. While many of the comments and concerns are generic in nature, some are specific to metals and other naturally occurring elements. This Fact Sheet summarizes the most important arguments for and against the use of distribution-based extrapolation methods and explores the particular concerns related to metals.

Definition

The toxicity threshold distribution approach assumes that the measured variability in sensitivities among species is a function of inherent differences in species sensitivity and in environmental (exposure) conditions that can be described by some appropriate statistical distribution. The parameters of the distribution (e.g., mean and standard deviation) are estimated from a set of laboratory toxicity data (e.g., LD_{50}/LC_{50}s, No Observed Effect Concentrations [NOECs] or the geometric mean of the NOEC and LOEC [termed the Chronic Value]). Then, a certain percentile in the distribution (i.e., level of protection) is chosen. In most applications the 5^{th} percentile is used; however, the theory allows for any percentile. Finally, a confidence interval for the percentile is estimated and the environmental standard (i.e., Predicted No Effect Concentration [PNEC]) may be established at the lowest 5^{th} percentile of this confidence interval (Figure 1).

This type of distribution often is shown as a cumulative distribution frequency (Figure 2).

The final concentration derived in this way can be seen as a concentration below which the probability of finding a species that is affected by that concentration equals a specified small value (usually less than 5%). The approach fits into a risk assessment philosophy that acknowledges that environmental protection cannot eliminate risks completely, but is aimed at reducing risks to acceptable

ICME

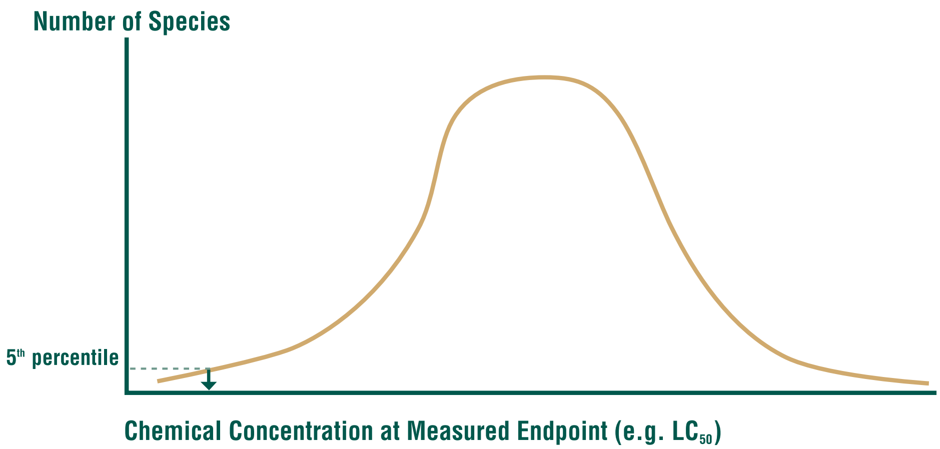

Figure 1

Toxicity Threshold Distribution

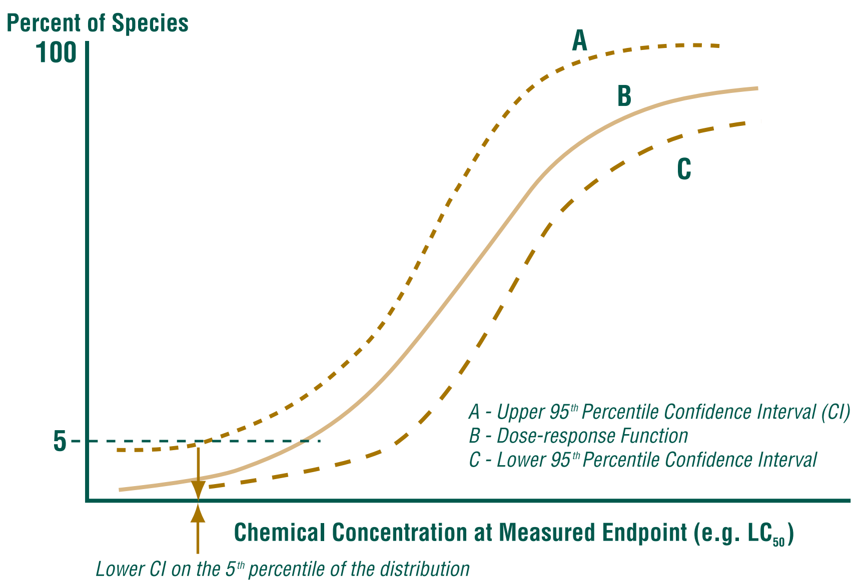

Figure 2

Cumulative Frequency Distribution

low levels. The distribution-based approach therefore also may be used in a "probabilistic risk assessment" framework to estimate the probability of affecting a given number of species at a specified environmental concentration, or distribution of concentrations.

Background

The development of distribution-based extrapolation methods (also known as the "species sensitivity distributions approach") in Europe originated with the paper by Kooijman (1987). Soon thereafter an extensive body of theory developed in Europe (summarized by Van Straalen and Van Leeuwen 2001). It is interesting that in the United States a similar path was taken (Stephan et al. 1985, Niederlehner et al. 1986), and up to around 1990, the European and American approaches developed more or less independently from each other. Nevertheless, there are many similarities, including such details as the 5% cut-off level. This likely was driven by similar discussions between science and policy. On both continents, there was a need to replace the fixed extrapolation factors (also known as "safety factors" or "uncertainty factors") that previously had been used when deriving environmental quality criteria from laboratory toxicity data. Objections to the fixed extrapolation factor approach are:

- There is no scientific basis for selection of the magnitude of fixed factors, other than that they represent whole numbers.
- The fixed factor approach does not allow flexibility, in that the factor is independent of the nature of the data.
- The degree of protection afforded by the use of a given value is generally unknown.
- Since the factor is applied to the lowest value of a set of experimentally determined NOECs or LC_{50}s, it often acts as a punishment for large data sets. The larger the data set, the greater the probability that there will be extreme values in it. This creates the very awkward situation that the derived PNEC can decrease with increasing sample size. Logically, it should be the other way around: if there is more information, the concern for uncertainty is lower and the derived PNEC should be higher.

> Extrapolation factors, Safety Factors, and Uncertainty Factors have no scientific basis, lack flexibility, and often return unrealistically low protection values.

These considerations urged scientists and policy makers to develop a more solid base for risk assessment. The aim was to increase transparency and objectivity by developing an algorithm that was based on assumptions that can be verified. It was also argued that such an algorithm could easily be updated in the light of new scientific knowledge.

> Environmental protection cannot eliminate risks completely, but is aimed at reducing risks to acceptable low levels.

It should be emphasized that the primary benefit of the distribution approach is that it provides a quantitative description of how changes in exposure relate to changes in the number or percentage of species at risk. Chapman et al. (1998) argued that each item of extrapolation should be considered separately and, if a safety factor is considered necessary, it should be based on scientific arguments relating to that item. There may be other elements in the extrapolation that require other safety factors. The authors considered it unwise to use a single (large) safety factor for many issues at the same time, because this will make the basis of these safety factors untractable. The distribution-based extrapolation approach is in line with this argument because it considers only one element of the extrapolation process.

Issues

The various points of discussion in the above debates can be classified into two categories: 1) operational issues, which concern parameterisations and numerical details of the methods, and 2) fundamental issues, which concern principal assumptions in the methodology. These two types of issues will be discussed here separately.

Operational Issues

> **Operational Issues**
> - Quality of input data
> - Choice of distribution
> - Choice of protection level
> - Choice of confidence interval

The quality of input data. When no-effect concentrations (NOECs) are used as the input data, a limitation arises because NOECs are generally estimated from step-wise dose–response studies (using analysis of variance approaches) and thus the result sometimes can be considerably higher than the true NOEC. Another issue with data quality is that many ecotoxicity experiments are conducted with freshly prepared media in which the toxicant has not established an equilibrium with binding components. As a consequence, laboratory ecotoxicity experiments tend to overestimate toxicity in the field. However, both of these data issues apply whether the fixed safety-factor or distribution approach is used.

The choice of a distribution. There are a number of statistical approaches that can be used to represent the distribution of toxicity data. It sometimes is argued that the commonly used log-normal distribution is inappropriate for ecotoxicity data, because the real data are asymmetrical, bimodal or truncated. However, there are other statistical distributions that may provide a better fit to ecotoxicity data. Numerical calculations have shown that the outcome of the extrapolation method is not very sensitive to the shape of the distribution. In fact, the normal, logistic and triangular distributions can hardly be discerned from one another if they have the same mean and standard deviation (Van Straalen and Van Leeuwen 2001). Another issue is that a taxonomically broad array of species may fall into two or more groups due to inherently different mechanistic responses to toxicants and, therefore, are not amenable to description by a single distribution. This can be rectified by grouping the data into logical taxonomic units, such as plants, animals, and microorganisms in the case of herbicides.

The choice of a protection level. As mentioned above, the distribution approach leaves the choice of the cut-off point up to the policy maker. In practice, however, only the 5% level is used. The choice of 5% is based on statistical conventions in which a 5% error is considered small enough to be acceptable in deciding whether there are “significant” effects. Additionally, it was recognized early on that use of values less than 5% resulted in values below background concentrations for metals and considerably less than the lowest chronic value in the data set. Hence the 5% level became adopted by convention.

The choice of a confidence interval. In the original paper by Van Straalen and Denneman (1989), it was proposed to set environmental quality criteria at the lower 95% confidence limit of the selected percentile. Later papers (Wagner and Løkke 1991, Aldenberg and Slob 1993) have pointed out that the algorithm proposed by Van Straalen and Denneman does not lead to the true 95% confidence interval. Better statistical procedures were developed to provide the true estimates. At the same time, it was proposed that the median of the confidence interval be used instead of the 95% lower limit of the 5th percentile. This was done to reduce the influence of data uncertainty. This influence is particularly large when sample sizes are small. In the case of large sample sizes, the difference between the median of the confidence interval and the 95% confidence limit becomes smaller and smaller. The most logical conclusion is that distribution-based models should be applied only in the case of sample sizes that are large enough to reduce the uncertainty in the estimation below an acceptable limit. In practice, it appears that sample sizes above 10 will suffice to accomplish this goal. The use of smaller data sets increases the uncertainty in the model outputs. Many of the metals have toxicity data on substantially more than 10 species, and investigations of their toxicity threshold distributions support this conclusion.

Fundamental Issues

- Protection of structure implies protection of function
- Critical species may be unprotected
- Essentiality not considered
- Representativeness of test species

Fundamental Issues

Protection of structure implies protection of function. Toxicity threshold distributions consider only the biological building blocks (ecological structure) of an ecosystem, not its functions (processes such as production, nutrient cycling, decomposition, etc.). It is assumed that the functions in an ecosystem will be safe if none of the species is exposed above its NOEC. Currently, there is a large research effort to find out how biodiversity (species richness) relates to ecosystem functioning. No clear answer can be given at the moment; however, it turns out that many ecosystems have a certain degree of functional redundancy. That is, not every loss of a species will affect ecosystem functioning. Within the same trophic group, the loss of one species can be compensated for by the growth of another species, which may take over its function. Conversely, a situation of *structural* redundancy, in which functions would be lost without effects on species, has never been demonstrated. The most problematic application of the distribution approach in this regard is with soil microbial function tests. There is extreme functional redundancy in the soil microbial community, and it is difficult to determine with accuracy the species richness. Therefore, functional assays (e.g., nitrogen fixation) generally are used to characterize ecotoxicological effects. However, it is not appropriate to place such functional endpoints on a sensitivity distribution and attempt to protect 95% of the functions; it is well known, for example, that soil ecosystems cannot function appropriately without means of decomposition, nitrogen fixation, or carbon cycling.

Critical species in the unprotected group. If the cut-off level in the distribution of sensitivities is fixed at 5%, the obvious conclusion would be that there are species that will not be protected. For the following reasons, it is hard to say whether this actually is 5% of the total number of species present at a particular location:

- Not all the tested species that are used to develop the distribution will be expected to be present at all sites. Some areas may not normally support the sensitive species, so none of the local species would be affected even at a higher percentile cut-off value.
- There generally is little information about the tail of the distribution. It is not logical to assume that the sensitivities extend down to the zero level. Even the most sensitive species will have a NOEC above zero (this is the “threshold theory” of toxicity; likewise, ecotoxicology assumes that a NOEC greater than zero exists for all endpoints). In fact, this is one of the reasons why selection of a cut-off level above zero was originally proposed.
- It is possible that some species of great ecological importance, commercial value, or with a great value for conservation will fall in the highly sensitive, unprotected group. Most authorities that have adopted the distribution-based approach have recognized this and have formulated special regulations that will prevent this from happening.

Ecologically or commerically important species may be given special protection.

For example, it may be required that at least some commercially important species, or species that are known to be in a vulnerable position, are tested and included as part of the data set. From an ecological point of view, the concern should be with the rare or highly specialized and habitat-specific species, more than with the species that take a crucial position in system productivity or nutrient cycles. "Ecological engineers" usually are abundant and robust species, with populations that are not easily affected by pollution.

Essential substances. For substances that have a biochemical role in organisms, such as many of the metals, the classical concentration–response relationship does not hold because the low concentration range induces a metabolic deficiency. There may be concentrations in the environment that are toxic for one species but, at the same time, deficient for another. The toxicity threshold distribution approach does not take these effects into account. Furthermore, since the tails of the distribution generally extend beyond any of the measured data, PNEC estimates may be produced that are below natural background concentrations. It is not yet clear how this issue of the essentiality of substances should be included in risk assessment. One line of thought is that ecotoxicity data for essential metals should not be all lumped together, but should be grouped per ecosystem type, where each ecosystem type has its own characteristic background concentration (i.e., the "metallo-region" concept). If the experimental data are then reported as measured concentrations, rather than as nominally added concentrations, the background concentration is part of the effect level and extrapolation below the background is avoided. Another approach is to develop distributions of essentiality, and verify that the PNEC does not fall below the upper 95th percentile of the essentiality distribution. Yet another approach is to divide species into "sensitive" and "tolerant" groups prior to development of the distributions.

Species sensitivity distributions may provide protection values below the required concentrations of essential elements.

Representativeness of species. In the distribution-based approach it is assumed that the species for which test data are available can be considered a random sample from the ecosystem whose protection is intended. This may be the most significant limitation with the method. It is evident that test species are *not* selected at random. If this were true, the vast majority of test species would be arthropods. In most of the larger data sets, there is more taxonomic diversity than would be obtained by random sampling, resulting in an overestimation of the standard deviation and longer tails to the distributions. Thus, the lower 5th percentile is a lower concentration than would be developed from a more representative data set. It also is known that test species are selected because of their sensitivity, or that inherently insensitive species are underrepresented in the distribution because it is not interesting to test them. On the other hand, selected test species could be among the insensitive ones because they require a certain robustness to be cultured and manipulated in the laboratory. Differences in species-specific metabolic responses to toxicants also may skew the data if the only species tested are those known to have activating enzymes, for example. A better understanding about patterns in the ecotoxicity data and the reasons why some species are more sensitive than others is needed.

Conclusion

Toxicity threshold (or species sensitivity) distributions have been developed as an approach to include more scientific reasoning into the derivation of environmental quality criteria. This approach is very flexible and can accommodate large variations in policy needs. Some limitations in the models are merely operational and can be avoided easily by adjusting the parameters. Other problems are of a more fundamental nature and require more research before an answer can be given. Regardless, the distribution-based sensitivity threshold approach is more scientifically rigorous and defensible than the fixed safety factor approach. The distribution-based approach is based on testable assumptions and holds better promise for the incorporation of scientific arguments into environmental quality criteria development and risk assessment.

References and Additional Reading

Aldenberg, T. and Slob, W. 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicology and Environmental Safety*, 25, 48-63.

Chapman, P. M., Fairbrother, A. and Brown, D. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. *Environmental Toxicology and Chemistry*, 17, 99-108.

Kooijman, S. A. L. M. 1987. A safety factor for LC_{50} values allowing for differences in sensitivity among species. *Water Research*, 21, 269-276.

Løkke, H., Christensen, B. and Møller, J. 1995. Extrapolation of the effects of glyphosate from the laboratory to the field. *Archiwum Ochrony Srodowiska*, 1, 109-120.

Niederlehner, B. R., Pratt, J. R., Buikema, A. L. J. and Cairns, J. J. 1986. Comparison of estimates of hazard derived at three levels of complexity. In *Community Toxicity Testing* (J. Cairns, Ed.). ASTM STP 920, Philadelphia, pp. 30-48.

Stephan, C. E., Mount, D. I., Hansen, D. J., Gentile, J. H., Chapman, G. A. and Brungs, W. A. 1985. Guidelines for deriving numerical water quality criteria for the protection of aquatic organisms and their uses. US EPA - Office of Research & Development, National Technical Information Service, Springfield.

Van Straalen, N. M. and Denneman, C. A. J. 1989. Ecotoxicological evaluation of soil quality criteria. *Ecotoxicology and Environmental Safety*, 18, 241-251.

Van Straalen, N. M. and Van Leeuwen, C. J. 2001. European history of species sensitivity distributions. In *Species Sensitivity Distributions in Ecotoxicology* (L. Posthuma and G.W. Suter II, Eds.). SETAC Press, Boca Raton.

Wagner, C. and Løkke, H. 1991. Estimation of ecotoxicological protection levels from NOEC toxicity data. *Water Research*, 25, 1237-1242.

About the Author:

Nico M. Van Straalen has been a professor of Animal Ecology at the Free University in Amsterdam since 1993. He teaches Evolutionary Biology, Ecology and Ecotoxicology to students of Biology, Medical Biology and Environmental Science. He is the leader of a research program on ecology and ecotoxicology of soil invertebrates, focusing on population structure, responses to soil pollution and ecological function of animals such as springtails, isopods and mites living in soil. His special interest lies in the physiology and molecular genetics of metal adaptation in the model species Orchesella cincta. He has also contributed to the risk assessment of soil pollutants by developing methods for distribution-based extrapolation. Dr. Van Straalen studied biology with a major in biophysics and minors in theoretical biology and endocrinology. He defended his PhD thesis in 1983 on "Comparative Demography of Collembola." He then developed teaching material (including a textbook) and a research program on ecotoxicology of soil invertebrates. After he became a professor, he supervised 25 PhD students in various areas of ecology. Dr. Van Straalen was one of the founding members of the Society of Environmental Toxicology and Chemistry (SETAC) in Europe. He has authored or co-authored 180 papers and book chapters since 1982.

Fact Sheet on Environmental Risk Assessment

This is the third in an occasional series of *Fact Sheets* to be produced by ICME on metal-specific issues in environmental risk assessment. Authorship selection and editorial review are coordinated by Dr. Anne Fairbrother of Parametrix, Inc. Each *Fact Sheet* is reviewed for technical merit by Dr. Erik Smolders of Katholieke Universiteit (Catholic University) Leuven, Belgium, and by a panel of experts in metal-related regulatory issues. While the *Fact Sheets* reflect the views of the authors, they are intended to provide an objective review of each topic. ICME hopes these publications provide insights into complex issues in regulatory science, and welcomes questions and comments.

These *Fact Sheets* are also available on the ICME Web site. For additional information, please contact:

The International Council on Metals and the Environment

294 Albert Street, Suite 506
Ottawa, Ontario, Canada K1P 6E6
Tel.: (+1 613) 235-4263 Fax: (+1 613) 235-2865
E-mail: info@icme.com
Web site: www.icme.com

INTERNATIONAL COUNCIL ON METALS AND THE ENVIRONMENT
Promoting policies and practices towards sustainable development
ICME