Methodology To Determine Risk Of Chemicals In Rivers Under Data Uncertainty

Frederik A.M. Verdonck[1], Joanna Jaworksa[2], Colin R. Janssen[3], Peter A. Vanrolleghem[1]

[1] Ghent University, BIOMATH, Coupure Links 653, B-9000 Gent, Belgium, frederik.verdonck@biomath.rug.ac.be

[2] Procter & Gamble, ETC, Temselaan 100,B-1853 Strombeek-Bever, Belgium

[3] Ghent University, Laboratory for Environmental Toxicology and Aquatic Ecology, Jozef Plateaustraat 22, B-9000 Gent, Belgium

ABSTRACT

The goal of probabilistic ecological risk assessment (PERA) is to estimate the likelihood and the extent of adverse effects occurring to humans and ecological systems due to exposure(s) to substances. It is based on the comparison of an exposure concentration distribution (ECD) with a species sensitivity distribution (SSD) derived from chronic toxicity data. This PERA framework was completed by also incorporating the uncertainty inherent to risk assessment. A case study on the pesticide atrazine in the surface waters of Flanders illustrates the completion. The availability of confidence intervals on the calculated risks is important for the decision-maker since these express how reliable the risk assessment is.

KEY WORDS

Species sensitivity distribution, probabilistic ecological risk assessment, water quality criteria, atrazine

INTRODUCTION

Today, over 60,000 chemicals are manufactured all over the world. Every year another 1,000 new chemicals are synthesised by researchers and scientists, and these chemicals are partly, directly or indirectly, released in the environment e.g. into the surface waters. Regulation puts constraints on these chemical emissions and these are based on environmental risk assessment.

The goal of a risk assessment is to estimate the likelihood and the extent of adverse effects occurring to humans and ecological systems due to exposure(s) to substances. Environmental risk assessment is based on the comparison of a predicted or measured exposure concentration (EC) with a ‘no effect concentration’ based on (acute or chronic) toxicity test results. In a deterministic framework, inputs to the exposure and effect prediction models are single values and the risk is calculated as simple ratios of EC and

effects. In a probabilistic framework, the environmental concentration (EC) and species sensitivity (SS) are treated as random variables taken from probability distributions (respectively ECD and SSD) which are combined to give a risk distribution.

In these types of risk assessments, the distinction between data uncertainty and variability must be made. Variability represents inherent heterogeneity or diversity in a well-characterised population. Fundamentally a property of nature, variability is usually not reducible through further measurement or study. Temporal and spatial variations of chemical concentrations can be captured in a variability distribution, called Exposure Concentration Distribution (ECD). Various species sensitivities towards a chemical can also be captured in a variability distribution called Species Sensitivity Distribution (SSD). These distributions are also used in water quality standard setting (e.g. in EU environmental risk assessment practices). In Figure 1, an example of a cumulative variability distribution is visualised by the black line.

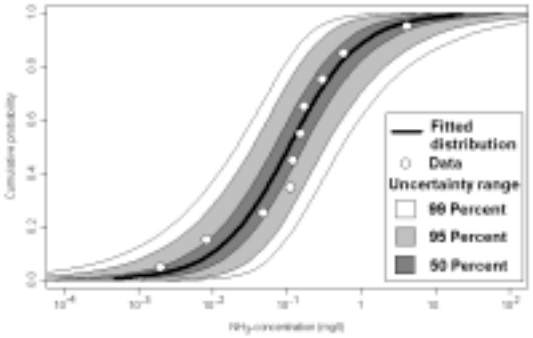

Figure 1: Variability and uncertainty of a data set

Uncertainty represents partial ignorance or lack of perfect information about poorly characterised phenomena or models (e.g. sampling or measurement error), and can partly be reduced through further research (Cullen and Frey 1999). In Figure 1, the uncertainty is visualised as a grey band around the cumulative variability distribution function. For each percentile of the variability distribution, an uncertainty or confidence interval can be calculated (i.e. the uncertainty distribution).

The characterisation of the risk of toxicants to species, when both EC and SS are variable and uncertain, is the central issue in Probabilistic Ecological Risk Assessment (PERA). The methodology is well developed in literature (Aldenberg et al. 2001). Among all risk calculation techniques available, one method was selected in this study: the risk quotient method.

The goal of this paper is to complete this PERA framework. Until now, only variability is considered. However, both EC and SS are characterised by uncertainty and variability, so risk should also be characterised by uncertainty and variability. Therefore, in this paper, the uncertainty will also be considered. Suppose, as an example, that the risk is 30%. A risk manager will feel more confident if he knows that the 90% uncertainty interval of that risk is between 25 and 35% rather than between 10 and 50%. A case study on the risk of atrazine in the surface waters of Flanders will be presented to illustrate the framework.

METHODS

Given are an EC and an SS data set. Variability and uncertainty for the EC and SS can be characterised using parametric or nonparametric procedures (Verdonck et al. 2001). For the case study, the lognormal (i.e. a parametric) model was selected. Several methods exist to determine the uncertainty band. Here, the numerical bootstrap/Monte Carlo technique was preferred since the technique is easy to understand and implement.

Modelling The Risk Distribution When Only Variability Is Considered

The probability of exceeding some randomly selected EC exceeding some randomly selected SS can be regarded as a measure of risk (Aldenberg et al. 2001). This can be written in formulae as: $\mathbf{Risk = P(EC > SS)}$.

The quotient method is well described in literature (Burmaster and Bloomfield 1996) (Rai et al. 1996). The ecological quotient estimates are used to define risks to potential ecological receptors. In environmental risk assessment, this risk quotient is an index of risk calculated by dividing an exposure estimate (EC) by a toxicity value (SS). The nominator and denominator values are in the same exposure units (e.g. mg/l) so that the ratio is dimensionless. A critical value of the risk quotient may form the basis for some regulatory action, including possible collection of more information or performing a more refined analysis (Warren-Hicks and Moore 1995).
In a probabilistic framework, the EC and SS are regarded as probability distributions rather than point estimates. As a result, the quotient will also be a probability distribution (see Figure 2). The probability of EC exceeding SS is equal to the probability that the quotient EC/SS becomes larger than 1. This probability can be considered as a measure of risk of adverse effects. This percentage can also be visualised as a column chart (see Figure 2).

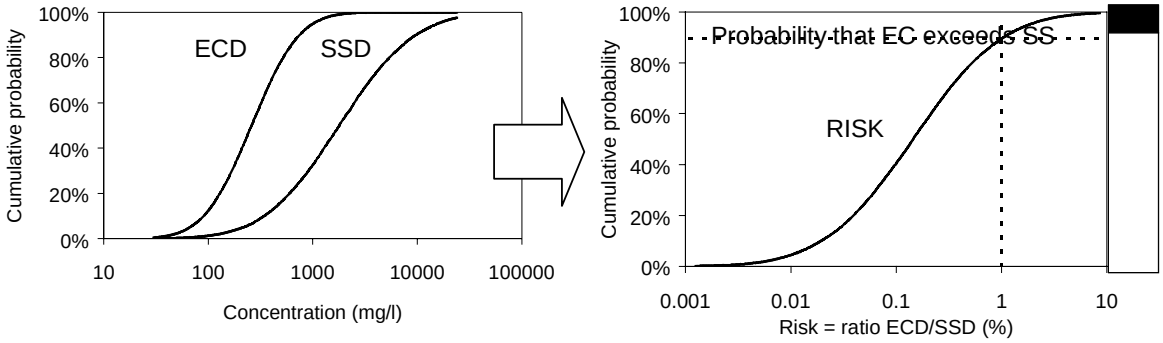

Figure 2: Calculation of risk based on ECD and SSD

When lognormal distributions are assumed for the ECD and the SSD, the risk can be calculated analytically. The result of a quotient of two lognormal distributions is again a lognormal distribution and its parameters can easily be calculated using the following equations (based on Burmaster and Bloomfield (1996) and also found by Aldenberg et al. (2001)):

$$\mu_{Risk} = \mu_{EC} - \mu_{SS}$$

$$\sigma_{Risk} = \sqrt{\sigma_{EC}^2 + \sigma_{SS}^2}$$

with μ and σ respectively the mean and standard deviation of the log-transformed data

Two important remarks have to be made. First, an important condition for using these formulas is that EC and SS are independent variables, which is the case. Second, in order to assess overlap of ECD and SSD, both sets of values have to be compatible (Aldenberg et al. 2001). One cannot compare 96h toxicity tests to hourly fluctuating concentrations at a discharge point.

Modelling The Risk Distribution When Variability And Uncertainty Are Considered

Until now, only the variability of the ECD and SSD were considered. This resulted in a risk variability distribution. The ECD and SSD are also uncertain because of sampling error. Adding a Monte Carlo sampling loop to the risk calculation can capture the uncertainty. In each run, an ECD and SSD will be selected from their respective uncertainty bands and the risk distribution will be calculated. After many runs, the risk distribution will also have an uncertainty band (Figure 3).
Remark that other uncertainty issues still need to be solved such as lab to field extrapolation uncertainties of the SS, the representativeness of the species in an SSD, model uncertainty (for choosing the lognormal distribution)… but these are not characteristic for PERA.

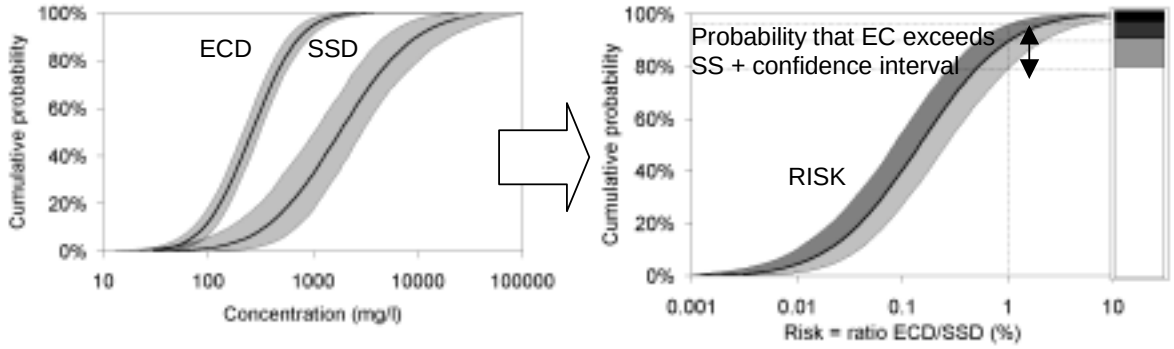

Figure 3: Calculation of the mean risk and its uncertainty interval based on ECD and SSD

The risk distribution visualised as a column chart in Figure 3 can also be visualised as a pie chart as in Figure 4. The entire pie represents 100%. The grey shades indicate how large the risk is with a pre-defined certainty. The larger the white piece, the lower the risk is. The more black, the larger the risk is. The larger the grey parts are, the more uncertainty there is on the estimated risk. The example shows that the median risk is 23% (50% certainty) and there is 95% certainty that the risk is smaller than 45%.

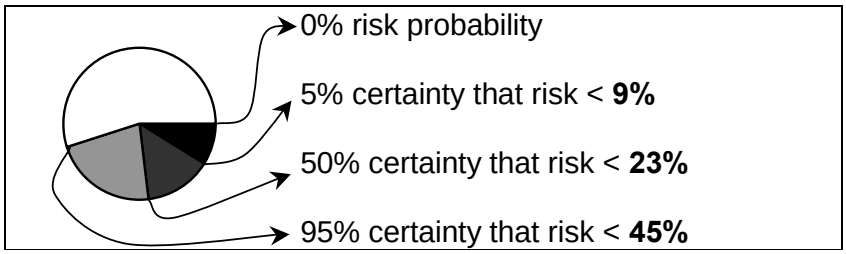

Figure 4: Visualisation of the risk of 23% and its 90%-uncertainty interval

CASE STUDY

As an illustrative case study, probabilistic risks and their 90%-uncertainty intervals were predicted for the pesticide atrazine in the river catchments of Flanders in Belgium. Since atrazine is such a widely used herbicide and the chemical nature of this compound is persistent, it is considered a great potential for groundwater and surface water contamination. Therefore, it is frequently detected. The greatest risk of atrazine runoff occurs shortly after the application because it hasn't had time to adhere to the soil particles and is still at the surface of the soil (Lipishan and Lee 1996).

The data set for the SSD consists of chronic toxicity values (NOEC: No Observed Effect Concentrations) and can be found in Versteeg et al. (1999). A lognormal distribution was assumed and fitted to the data. The fit was satisfactory.

The exposure concentrations were obtained from the Flemish environmental agency (VMM 2001). Atrazine was (mostly monthly) measured at 134 locations from 1991 till 2000. Only the reliable data from the years 1997 till 2000 were considered. It is assumed that this monitoring network is representative for all rivers in Flanders.
The cumulative empirical distribution function of all the EC is shown in Figure 5. A lognormal distribution was assumed and fitted to the data but the model does not fit very well to the data (see grey curve in Figure 5) because the data are left censored i.e. the value 50 ng/l is frequently observed.
This value corresponds with the detection limit of atrazine. Censored data can be handled in different ways. Govaerts et al. (2001) recommends to replace every value below the detection limit with a random number between zero and the detection limit (here 50). After the correction for censoring, the lognormal distribution now fits very well to the data (see black curve in Figure 5). So, for every monitoring station, a lognormal distribution was fitted. The resulting ECD represents the variation (such as temporal) of the concentration for that station.

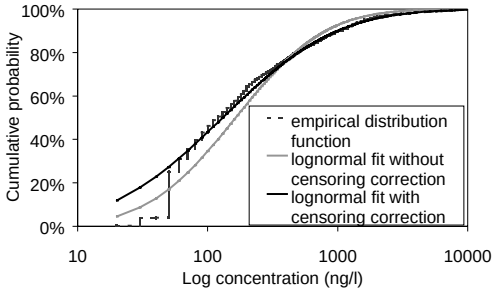

Figure 5: Cumulative probability distribution of atrazine measurements in surface waters in Flanders

RESULTS FOR THE CASE STUDY

The results of the local PERA of atrazine for all monitoring stations in the river networks of Flanders are visualised in Figure 6. Two monitoring stations in Alveringem and Aalter were selected for more detailed study. Their PERA is visualised in Figure 7.

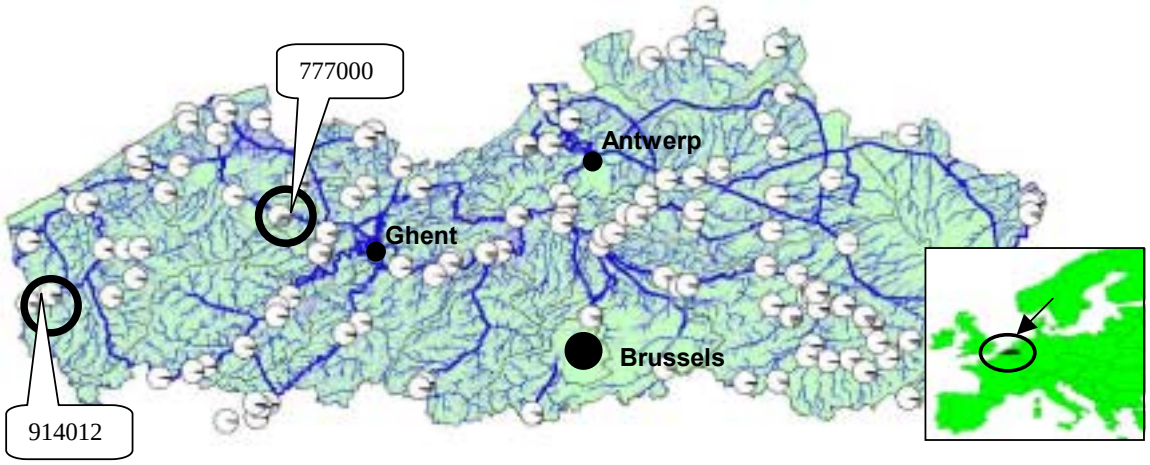

Figure 6: Atrazine risk in the catchments of Flanders (Belgium)

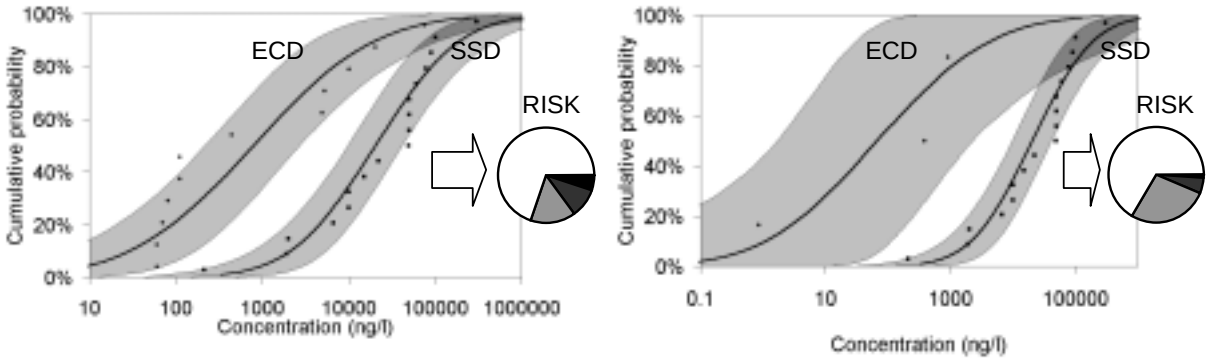

Figure 7: ECD, SSD and the risk uncertainty distribution for the monitoring points in (a) the "Westsluisbeek" in Alveringem (VMM NR. 914012) and (b) the "kanaal van Gent naar Oostende" in Aalter (VMM NR. 777000)

DISCUSSION

The results of the local PERA of atrazine in the basins of Flanders indicate that the predicted atrazine risks are around 1,6% for many monitoring locations. Most of the median risks are smaller than 5% and most of the upper risk 90% uncertainty intervals are smaller than 15%.

The first selected monitoring station (Alveringem) shows that the calculated risk is only a comparative measure. Despite the large overlap of the ECD and the SSD, the predicted mean risk is 15%. A better measure would be obtained when more attention is paid on the upper tail of the ECD and the lower tail of the SSD. The second selected monitoring station (Aalter) shows that the width of the uncertainty interval on the risk heavily depends on the uncertainty of the ECD (or SSD). The lack of more EC information results in a larger uncertainty interval for the risk.

The PERA framework improves transparency, credibility, it focuses data collection, it avoids worst-case assumptions, it improves decision support and, above all, it is more realistic compared to the current deterministic risk assessment approaches. As a result, this approach enables risk managers to evaluate the full range of variability and uncertainty instead of just using point estimates of exposure, effects and eventually risk.

CONCLUSION

A framework for uncertainty analysis in probabilistic environmental risk assessment (PERA) was proposed and illustrated with a case study. The uncertainty or confidence intervals are important for the decision-maker since these express how reliable the risk assessment is. A probabilistic approach results in a more realistic environmental risk assessment and therefore improves decision support of individual chemicals.

ACKNOWLEDGEMENTS

This research was funded by a scholarship from the Flemish Institute for the Improvement of Scientific-Technological Research in the Industry (IWT).

REFERENCES

Aldenberg, T., J. S. Jaworska, and T. P. Traas. (2001) *Normal species sensitivity distributions and probabilistic ecological risk assessment*. *in* L. Posthuma, G. W. Suter, and T. P. Traas, editors. *Species sensitivity distributions in ecotoxicology*. Lewis publishers.

Burmaster, D. E., and L. R. Bloomfield. (1996) *Mathematical properties of the risk equation when variability is present.* Human and Ecological Risk Assessment **2**:348-355.

Cullen, A. C., and H. C. Frey. (1999) *Probabilistic techniques in exposure assessment. A handbook for dealing with variability and uncertainty in models and inputs*. Plenum, New York.

Govaerts, B., B. Beck, E. Lecoutre, C. le Bailly, and P. Vanden Eeckhaut. (2001) *From monitoring data to regional distributions: a practical methodology applied to water risk assessment*. 0127, Institut de statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

Lipishan, B., and S. Lee. (1996) The Atrazine Herbicide. http://www.inweh.unu.edu/biology447/modules/intro/.

Rai, S. N., D. Krewski, and S. Bartlett. (1996) *A general framework for the analysis of uncertainty and variability in risk assessment*. Human and Ecological Risk Assessment **2**:972-989.

Verdonck, F. A. M., J. Jaworska, O. Thas, and P. A. Vanrolleghem. (2001) *Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study*. Analytica Chimica Acta **446**:429-438.

Versteeg, D. J., S. E. Belanger, and G. J. Carr. (1999) *Understanding single-species and model ecosystem sensitivity: data-based comparison*. Environmental Toxicology Chemistry **18**:1329-1346.

VMM. (2001) personal communication. *in*. Flemisch Environmental Agency, Aalst, Belgium.

Warren-Hicks, W. J., and D. R. J. Moore. (1995) *Uncertainty analysis in ecological risk assessment*. SETAC, Pensacola.

CONTACT

Frederik Verdonck, Ghent University, BIOMATH Department, Coupure Links 653, B-9000 Gent, Belgium, Phone: +32 (0)9 2645937, Fax: +32 (0)9 2646220, frederik.verdonck@biomath.rug.ac.be