

Geo-referencing of probabilistic risk of new chemicals in rivers

F.A.M. Verdonck*, C.R. Janssen**, J. Jaworska*** and P.A. Vanrolleghem*

* Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), Ghent University, Coupure Links 653, B-9000 Gent, Belgium (E-mail: frederik.verdonck@biomath.rug.ac.be)

** Laboratory for Environmental Toxicology and Aquatic Ecology, Ghent University, Gent, Belgium

*** Procter & Gamble, ETC, Strombeek-Bever, Belgium

Abstract The characterisation of the risk of (new) chemicals to species/communities, when both the exposure/environmental concentration and effects (species sensitivity) are variable and uncertain, is the central issue in Probabilistic Environmental Risk Assessment. The spatial variability is one of the largest components of the total variability. This paper tries to explicitly account for this spatial variability by geo-referencing the exposure, effect and finally probabilistic risk. Geo-referencing makes the risk assessment more refined and realistic. In addition, it is also highlighted that geo-referencing the effects of chemicals (species sensitivity distribution) is still a large unexplored area but has large potential to improve probabilistic ecological risk assessments.

Keywords Effects; exposure; risk assessment

Introduction

Yearly, thousands and thousands of existing and new chemicals are released in the environment. 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/Environmental Concentration (EC) with a “no effect concentration” based on a set of (acute or chronic) toxicity test results (i.e. testing Species Sensitivity (SS)). In the deterministic framework, inputs to the exposure and effect prediction models are single values and the risk is calculated as a simple ratio of EC and effects (SS).

This approach does insufficiently account for the inherent variability and uncertainty of the Environmental Concentration (EC) and the Species Sensitivity (SS). It is important to separate variability and uncertainty. 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. The two most important sources of variability for the EC are spatial and temporal variability. Spatial and temporal variations of chemical concentrations can be captured in a variability distribution, called Exposure Concentration Distribution (ECD). Various SS towards a chemical (i.e. inter-species sensitivity/variability) can also be captured in a variability distribution called Species Sensitivity Distribution (SSD). Note that these distributions are also used in environmental quality standard setting. In Figure 1, the variability distributions are visualised as probability distribution functions by the black line. 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 variability distribution function. For each percentile of the variability distribution, an uncertainty or confidence interval can be calculated (Verdonck *et al.*, 2002).

The characterisation of the risk of toxicants to species, when both EC and SS are variable and uncertain, is the central issue in probabilistic environmental risk assessment. As a result, the risk is no longer a simple ratio of crisp exposure and effects measures but rather a probability. In addition, the risk probability can be accompanied with a confidence or uncertainty interval (see Figure 1). In summary, probabilistic environmental risk assessment improves transparency and 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 (Warren-Hicks and Moore, 1995; Verdonck *et al.*, 2002).

Currently, risk assessments especially those for regulatory decisions are done for generic situations determined by a set of default values. However, the exposure spatial variability, for instance, can be quite high. For example in Belgium alone, atrazine concentrations in surface water range from 50 ng/l (detection limit) to more than 1 mg/l (Vandenbroele *et al.*, 2000). This is a range of five orders of magnitude. Consequently, incorporating spatial characteristics of the receiving environment could further increase realism.

The goal of this paper is first to specifically account for the spatial information in the probabilistic environmental risk assessment and second, to show the usefulness of geo-referenced probabilistic risk for new individual chemicals. For this, a hypothetical, but realistic case study will be performed. In addition, some potential application issues will be addressed.

Methodology

General methodology of probabilistic environmental risk assessment

The probabilistic environmental risk assessment methodology introduced above is well developed in the literature (Aldenberg *et al.*, 2002; Verdonck *et al.*, 2002). Evidently, it requires that variability distributions (and their uncertainty) are determined. Two different approaches can be used to determine the Exposure Concentration Distribution (ECD) and the Species Sensitivity Distribution (SSD). Data from either measurements in the environment or toxicity tests can be used directly. The alternative is to use prediction or extrapolation models (especially in case of new chemicals). However, these models also need (other) data, which can again be characterised by uncertainty and variability. As a consequence, a distinction should be made between statistical methods for estimating data uncertainty and variability (e.g. bootstrapping (Verdonck *et al.*, 2001)), and methods for propagating uncertainty and variability through mathematical models (such as Monte Carlo analysis).

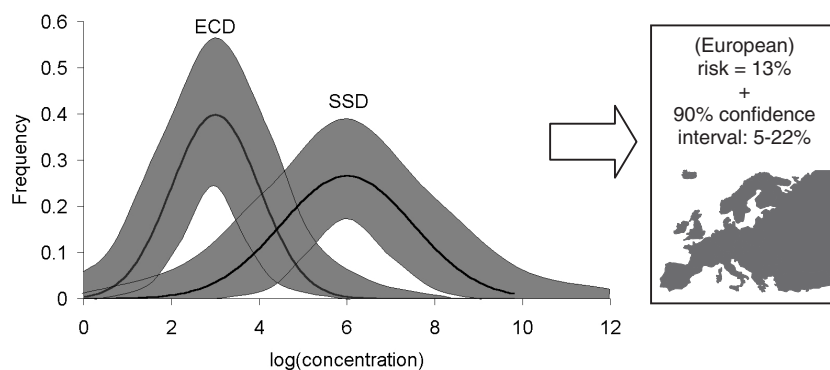


Figure 1 Probabilistic Environmental Risk Assessment is based on the comparison of an Exposure Concentration Distribution (ECD) and a Species Sensitivity Distribution (SSD), the grey bands represent 90%-uncertainty bands

Next, it will be discussed in short how to calculate a probabilistic risk and its uncertainty interval. Among all risk calculation techniques available, one method was selected in this study: the probabilistic risk quotient method. The probability of some randomly selected EC exceeding some randomly selected SS can be regarded as a measure of risk (Aldenberg *et al.*, 2002). This can be written as:

$$\text{Risk} = P(\text{EC} > \text{SS}) = P\left(\frac{\text{EC}}{\text{SS}} > 1\right) = P\left(\log_{10}\left(\frac{\text{EC}}{\text{SS}}\right) > 0\right) = P(\log_{10}(\text{EC}) - \log_{10}(\text{SS}) > 0)$$

The risk quotient is an index of risk calculated by dividing an exposure estimate (EC) by a toxicity value (SS). Its properties have been well described in the literature (Burmester and Bloomfield, 1996; Rai *et al.*, 1996). The risk quotient estimates are used to define risks to selected species representing an ecosystem. A critical value of the risk quotient may form the basis for regulatory action, including possible collection of more information or performing a more refined analysis (Warren-Hicks and Moore, 1995). In a probabilistic framework, however, EC and SS are regarded as random variables having probability distributions rather than point estimates. As a result, the quotient will also have a probability distribution. The probability of EC exceeding SS (this probability can be considered as a measure of risk of adverse effects) is equal to the probability that the quotient EC/SS becomes larger than one or that $\log_{10}(\text{EC}/\text{SS})$ becomes larger than zero (see equation above). This probability can be considered as a measure of risk of adverse effects. When lognormal distributions are assumed for the ECD and the SSD, the risk can be calculated analytically (Verdonck *et al.*, 2002). In case the ECD or SSD have a different probability distribution, the risk can always be calculated numerically by means of a Monte Carlo analysis.

Geo-referenced exposure concentration distribution (geo-ECD)

The ECD can be determined either through monitoring or through modelling. When new, unreleased individual chemicals are assessed, prediction models are the only possibility to determine an ECD. GREAT-ER is such an (aquatic) chemical exposure prediction tool for use within environmental risk assessment schemes. GREAT-ER 1.0 (Geo-referenced Regional Exposure Assessment Tool for European Rivers) calculates the distribution of the ECD of consumer “down-the-drain-chemicals” in surface waters, for individual river stretches as well as for entire catchments. The system uses a geographical information system for data storage and visualisation, combined with simple mathematical models for prediction of chemical fate. An extensive monitoring programme for boron and for Linear Alkylbenzenesulfonate in six European pilot study areas has been performed in order to validate the system. The results illustrate that GREAT-ER can deliver very accurate predictions of chemical concentrations in a river basin, provided reliable input datasets and accurate hydrological and chemical fate models are used (Schowanek *et al.*, 2001). GREAT-ER has already built in the idea of refining the exposure assessment by explicitly accounting for the spatial variability (geo-referencing the ECD). Instead of having one lumped ECD for an entire catchment (representing spatial and temporal variability), each river stretch has its own ECD (only representing temporal variability). In other words, the ECD for an entire catchment (ECD_{tot} in the example of Figure 2) has been unlumped into several ECDs (ECD 1, 2 and 3 in the example of Figure 2). A Monte Carlo analysis propagates the temporal variability of the input parameters (such as the river flow). Note that the input uncertainty is not considered here. For this, a second order or two dimensional Monte Carlo would be needed.

Geo-referenced effects: species sensitivity distribution (geo-SSD)

Geo-referenced risk is only useful when both exposure (ECD) and effects (SSD) are geo-referenced (Verdonck *et al.*, 2002). Instead of having one SSD for an entire catchment (representing spatial and other types of variability, SSD_{tot} in the example of Figure 2), each river stretch can have its own SSD (SSD A and B in the example of Figure 2).

Geo-referencing effects/SSD is still a large, unexplored area. Many factors influence a geo-SSD. The species presence is one of them. Several physico-chemical environmental circumstances create ecological niches. These result in different biodiverse communities. For example in a river environment, the community and as a result the SSD, will be different in rivers with large flow and sandy beds compared to rivers with low flow and a bed composed of rocks. Physico-chemical characteristics are also determining the bioavailability and toxicity of, for instance, metals (Janssen *et al.*, 2000) and other chemicals. Depending on the metal background concentration, biological communities in these different systems may have differentially acclimated/adapted to the natural presence of metal concentrations resulting in varying community sensitivities (Janssen *et al.*, 2000).

Depending on the chemical, one factor will be more important than the other will. It is not the goal of this paper to determine and quantify the most important factors nor to propose a framework of how to determine a geo-SSD. Geo-referenced effects are only used here to illustrate the usefulness of a geo-referenced risk. The usefulness of geo-risk may lead to more research in this area.

In the case study, an empirical approach based on considerable and arguable assumptions is used to determine a geo-SSD to study the environmental effect of a new chemical. The underlying idea of this approach is that a difference should be made between heavily polluted rivers (with low biodiversity) and rivers with a good water quality and high biodiversity. A heavily polluted river does not contain sensitive species. As a result, the SSD of a heavily polluted river will only contain more resistant species. Consequently, the risk of a new chemical affecting the species present in a heavily polluted river will be lower compared to a river with high biodiversity and sensitive species. This factor was only selected to illustrate geo-risk. Naturally, in the long term, the philosophy may be adhered to that all species in all rivers (also the current polluted ones) should be protected and a high biodiversity should be achieved and maintained.

Geo-referenced risk (Geo-risk)

Finally, once both a geo-ECD and a geo-SSD are determined, a geo-referenced risk can be calculated for every river stretch in a river basin. Figure 2 shows an example. ECD_{tot} and SSD_{tot} are the lumped, non-geo-referenced exposure and effect distributions (reflecting spatial and other types of variability). By geo-referencing the exposure (ECD 1, 2 and 3 in Figure 2), effects (SSD A and B in Figure 2) and the resulting risk assessment, the spatial variability is explicitly accounted for in each local risk assessment and as a result the risk assessment will be more realistic. The combinations of ECD 1 and SSD A and ECD 2 and SSD B give smaller risks (in comparison with the risk from ECD_{tot} and SSD_{tot}) whereas ECD 3 and SSD A will give a higher risk (in comparison with the risk from ECD_{tot} and SSD_{tot}). These local risks can afterwards again be aggregated to a lumped probabilistic risk. For this, an average of only the polluted stretches and weighted by river stretch length was used to resolve any scale-dependencies. More information on aggregation techniques in river basins can be found in Boeije *et al.* (2000).

Realistic hypothetical case study

The Rupel basin in Belgium was selected as the case study area. The Rupel is a tributary to the river Scheldt. The area is 6,700 km² large and contains the capital city of Belgium,

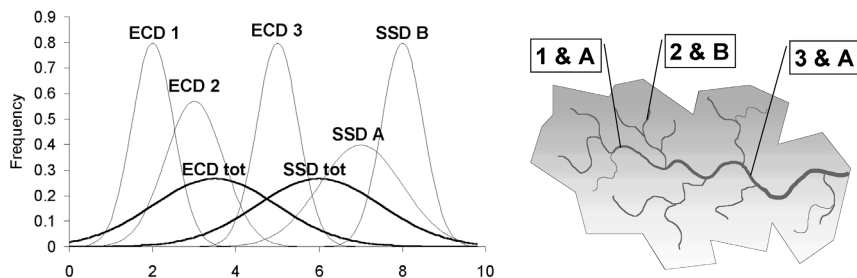


Figure 2 (left) Split up of the lumped ECD and SSD into local ECDs (1, 2, 3) and respectively SSDs (A, B) (right) by geo-referencing the risk analysis, the spatial variability is explicitly accounted for in each river stretch in a river basin

Brussels (Figure 3). The chemical under study is a new, hypothetical anionic surfactant (to be used in detergents) for widespread use, once regulated. Degradation rates, chronic toxicity values and other parameters were chosen as realistic as possible (based on existing literature information of similar chemicals).

Geo-ECDs were predicted using GREAT-ER 1.0 (situation wastewater treatment plant infrastructure 1999) (Verdonck *et al.*, 2000). Geo-SSDs were determined based on the empirical biodiversity approach described above. The Belgian biotic index, expressing species sensitivity and biodiversity of macro-invertebrates and ranging from 0 to 10 (De Pauw *et al.*, 1986), was used to get an idea about the overall biodiversity in the rivers. The biological monitoring network is rather dense (around 300 monitoring points). Only three categories of SS and biodiversity were determined: (1) moderate-good biodiversity (all species were selected in the SSD), (2) moderate-poor biodiversity (the 30% most sensitive species were removed from the SSD, see also graphically SSD A in Figure 2) and (3) poor-very poor biodiversity (the 50% most sensitive species were removed from the SSD, see SSD B in Figure 2). A log-normal distribution was fitted to the EC simulations and the three SS data sets (ECD and SSD).

Three tiers were simulated to assess the usefulness of a geo-referenced analysis. In the first tier, a non-geo-referenced probabilistic environmental risk assessment was performed (ECDtot and SSDtot). The second tier, the ECD was geo-referenced but the SSD was not (ECD 1, 2, 3, ... and SSDtot). And in the third tier, both the ECD and SSD were geo-referenced (ECD 1, 2, 3, ... and SSD A, B, ...).

Results and discussion

In the first tier, a non-geo-referenced probabilistic environmental risk assessment of the new chemical was performed on the entire Rupel basin. The resulting risk is 27%. This can be interpreted as the probability that an Environmental Concentration (EC) from a randomly selected river stretch and a randomly selected day in the year will be larger than a randomly selected Species Sensitivity (SS) from a randomly selected river stretch (with high or low biodiversity) (Aldenberg *et al.*, 2002). There was no confidence/uncertainty interval calculated for this risk probability.

The results of the geo-referenced probabilistic environmental risk assessment of the new, hypothetical but realistic, chemical on the Rupel basin (tier 3) are visualised in Figure 3 (without uncertainty intervals). The grey pattern indicates how large the risk is (the darker the river, the more the risk). It can be derived from Figure 3 that spatial risks or geo-risks can vary from 0% to larger than 90%. However, the interpretation of these geo-risks is slightly different. They can be interpreted as the probability (for a particular location) that an Environmental Concentration (EC) from a randomly selected day in the year will be

larger than a randomly selected Species Sensitivity (SS). For a particular river stretch, the spatial component is no longer needed in the interpretation because the risks are geo-referenced. Consequently, the risk assessment is more refined compared to the single number (27% in this case study) from a non-geo-referenced approach. Some rivers have higher risk probabilities of affecting the species, others have lower probabilities. These local risks could again be aggregated to a lumped probabilistic risk. Here, the aggregated risk was weighted by the length of each river stretch and resulted in risk probability of 9%. Geo-referencing reduced the risk by a factor 3.

The rivers under risk can now be studied in more detail to find the underlying causes of higher risks: is wastewater treatment plant infrastructure insufficient, under construction or to be improved, are the rivers under risk of high ecological importance, etc...? Such an analysis may help the decision-maker to avoid approving unsafe chemicals or rejecting safe chemicals. For example, the Nete basin (i.e. the Northern subbasin) has in general a very good wastewater treatment plant infrastructure, good water quality (high biodiversity) and low risks although there are some hot spots with large risks. In the Zenne river (i.e. downstream from Brussels, with bad water quality and low biodiversity), large risks were predicted even when the most sensitive species were already eliminated.

This can also be seen in risk river profile plots (see Figure 4) where probabilistic risk is plotted against the distance in the river. The three tiers are shown in Figure 4. In the first tier, a probabilistic environmental risk assessment is performed on catchment level. The calculated risk is 27% (see above) and can therefore be visualised as a straight line (independent of the distance in the river). In the second tier, the geo-risk is calculated based on a geo-ECD and a non-geo-SSD (as in GREAT-ER 1.0). In the last tier, a geo-risk is calculated based on both a geo-ECD and a geo-SSD.

In the most upstream 35 km part of the river, water quality objectives aim for sustainable fish populations. For this, all sensitive and insensitive species were used in the SSD. Consequently, tier 2 and 3 give the same risk profile. However, in the most downstream part of the river (last 40 km, almost near the mouth), water quality is so bad that the biodiversity is very low. Because of that, adding an additional, new chemical will have less/no effect to the sensitive species since they are not present. As a result, the risks for tier 3 are much smaller than for tier 2. Of course, the new chemical will impose additional stress on the existing insensitive species. It must be clear that the factor that was selected in this contribution to determine a geo-SSD should be regarded as an illustration of the

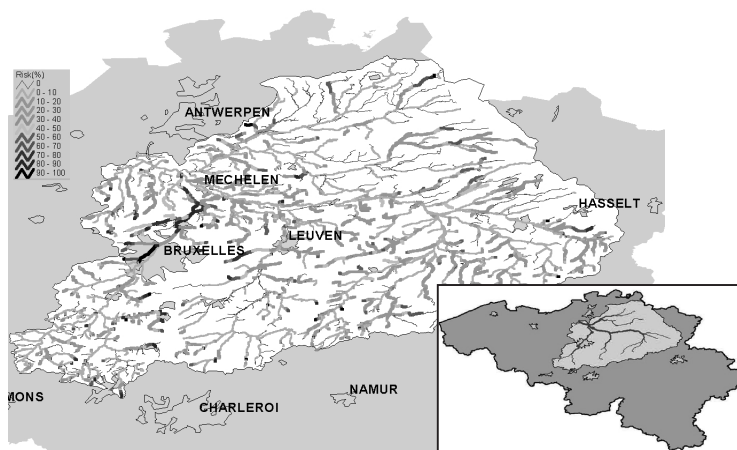


Figure 3 Geo-referenced probabilistic risk (no uncertainty intervals visualised) for a new, hypothetical but realistic, chemical on the Rupel basin in Belgium

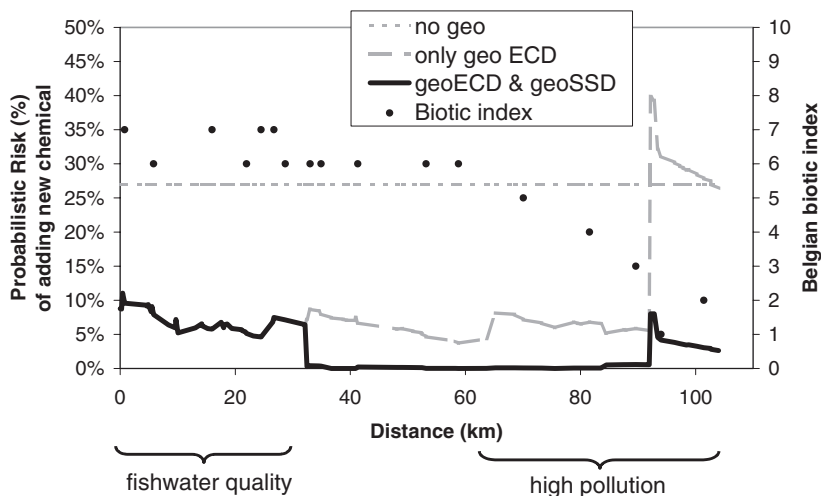


Figure 4 Probabilistic risk (of adding a new chemical) river length profile plot starting in the “Grote Gete”, “Demer”, “Dijle” and ending in the “Rupel” in the Rupel basin (Belgium)

methodology and points to the need of further development of a geo-referenced effect analysis (SSD).

An additional bottleneck for geo-referenced probabilistic environmental risk assessment may be data availability. The more information one wants to incorporate (to make the probabilistic environmental risk assessment more realistic), the more data will be needed. Data acquisition and management is however improving rapidly and will warrant the future application of the proposed methodology.

Conclusions

A geo-referenced probabilistic environmental risk assessment framework was developed and illustrated with a hypothetical, but realistic case study. Geo-referencing makes the risk assessment more realistic as spatial information is explicitly accounted for i.e. less spatial variability is lumped in the probabilistic environmental risk assessment and therefore, this is useful for assessing risk of new individual chemicals. In this case study, the probabilistic risk was reduced by a factor 3 by accounting for more geographical information. In addition, it was also highlighted that geo-referencing effects (species sensitivity distribution) is now still a largely unexplored area but has important potential to improve probabilistic environmental risk assessments.

Acknowledgements

This research was funded by a scholarship from the Flemish Institute for the Improvement of Scientific-Technological Research in the Industry (IWT). I would also like to thank B-IWA (Belgian branch of IWA) for financial support to attend the IWA Young Researchers Conference.

References

- Aldenbergh, T., Jaworska, J.S. and Traas, T.P. (2002). Normal species sensitivity distributions and probabilistic ecological risk assessment. Pages 49–102 in L. Posthuma, G.W. Suter II, and T.P. Traas, editors. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers.
- Boeije, G.M., Wagner, J.-O., Koormann, F., Vanrolleghem, P.A., Schowanek, D. and Feijtel, T.C.J. (2000). New PEC definitions for river basins applicable to GIS-based environmental exposure assessment. Contribution to GREAT-ER #8. *Chemosphere*, **40**, 255–265.

- Burmaster, D.E. and Bloomfield, L.R. (1996). Mathematical properties of the risk equation when variability is present. *Human and Ecological Risk Assessment*, **2**, 348–355.
- Cullen, A.C. and Frey, H.C. (1999). *Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Plenum, New York.
- De Pauw, N., Roels, D. and Fontoura, A.P. (1986). Use of artificial substrates for standardized sampling of macroinvertebrates in the assessment of water quality by the Belgian Biotic Index. *Hydrobiologia*, **100**, 153–168.
- Janssen, C.R., De Schampelaere, K., Heijerick, D., Muysen, B., Lock, K., Bossuyt, B., Vangheluwe, M. and Van Sprang, P. (2000). Uncertainties in the environmental risk assessment of metals. *Human and Ecological Risk Assessment*, **6**, 1003–1018.
- Rai, S.N., Krewski, D. and Bartlett, S. (1996). A general framework for the analysis of uncertainty and variability in risk assessment. *Human and Ecological Risk Assessment*, **2**, 972–989.
- Schowaneck, D., Fox, K., Holt, M., Schröder, F.R., Koch, V., Cassani, G., Matthies, M., Boeije, G., Vanrolleghem, P.A., Young, A., Morris, G., Gandolfi, C. and Feijtel, T.C.J. (2001). GREAT-ER: a new tool for management and risk assessment of chemicals in river basins. Contribution to GREAT-ER #10. *Water Science & Technology*, **43**(2), 179–185.
- Vandenbroele, M., Janssen, C.R., Bogaert, S. and Le Roy, D. (2000). Hefbomen voor een beleid gericht op duurzame ontwikkeling: uitwerking van evaluatiecriteria voor pesticiden in relatie tot de milieufiscaliteit. HL/DD/011, Laboratorium voor Milieutoxicologie en Aquatische Ecologie, Ghent University (in dutch), Gent, Belgium.
- Verdonck, F.A.M., Boeije, G.M., Schowaneck, D.R., Vanrolleghem, P.A. and ECETOC Task Force. (2000). Geography-referenced regional exposure tool for European rivers (GREAT-ER): A case study for the Rupel basin. in *Proceedings 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4)*, Banff, Alberta, Canada.
- Verdonck, F.A.M., Jaworska, J., Thas, O. and Vanrolleghem, P.A. (2001). Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study. *Analytica Chimica Acta*, **446**, 429–438.
- Verdonck, F.A.M., Jaworska, J., Janssen, C.R. and Vanrolleghem, P.A. (2002). Geography referenced probabilistic environmental risk assessment for chemical substances. in *Proceedings iEMSs 2002, Integrated Assessment and Decision Support*, Lugano, Switzerland.
- Warren-Hicks, W.J. and Moore, D.R.J. (1995). Uncertainty analysis in ecological risk assessment. SETAC, Pensacola, US.