Pergamon

Wat. Sci. Tech. Vol. 36, No. 5, pp. 251–258, 1997.

Printed in Great Britain.
0273–1223/97 $17·00 + 0·00

PII: S0273-1223(97)00481-2

A GEO-REFERENCED AQUATIC EXPOSURE PREDICTION METHODOLOGY FOR 'DOWN-THE-DRAIN' CHEMICALS[1]

Geert Boeije*, Peter Vanrolleghem* and Michael Matthies**

** Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), University of Gent, Coupre 653, B-9000 Gent, Belgium*
*** Institute of Environmental Systems Research, University of Osnabrück, Artilleriestrasse 34, D-49069 Osnabrück, Germany*

ABSTRACT

A geo-referenced simulation methodology for the prediction of aquatic exposure to individual 'down-the-drain' chemicals (consumer chemicals which mainly enter the environment via the domestic waste water route, e.g. detergents) is presented. This method uses real-world data, including their spatial and temporal variability and uncertainty. It results in statistical frequency distributions of predicted environmental concentrations (PEC).
A hybrid stochastic / deterministic simulation approach is used. Steady-state deterministic models, which describe chemical fate, form the system's core. A stochastic (Monte Carlo) simulation is applied on top of this. In the (deterministic) waste water pathway model, all processes which occur in the waste water drainage area of a discharge point are considered: emission, transport (sewers, small surface waters), and treatment (on-site treatment, waste water treatment plants). In the river model, chemical transport and conversion in main rivers is simulated. © 1997 IAWQ. Published by Elsevier Science Ltd

KEYWORDS

consumer chemicals, environmental exposure assessment, fate models, Monte Carlo simulation, risk assessment, river models, uncertainty

INTRODUCTION

In this paper, the general exposure simulation concepts applied in the GREAT-ER project (Geography-referenced Regional Exposure Assessment Tool for European Rivers) are described. The main objective of this project (Matthies et al., 1996; Feijtel et al., in press) is the development (and verification) of an aquatic exposure prediction tool for 'down-the-drain' chemicals, in a (pan-European) geo-referenced framework. This research is carried out on behalf of the European Center for Ecotoxicology and Toxicology of Chemicals (ECETOC).

For the environmental risk assessment of chemicals, the calculation of regional predicted environmental concentrations (PEC) is required by the EU legislation (EEC, 1993a, b, 1994a, b; ECETOC, 1993, 1994a). Currently, generic multimedia 'unit world' approaches (Mackay, 1991) are used for this purpose (RIVM VROM & WVC, 1994; ECETOC, 1994b). In these methods, average or default values are applied (e.g. 70% connection to waste water treatment). They do not take into account any spatial or temporal variability in environmental characteristics, river flows, or chemical emissions. The obtained results are not realistic, and are therefore only applicable at screening level (European Science Foundation, 1995).

[1] Contribution to GREAT-ER #3

The GREAT-ER concept was developed to obtain more reliable predictions, which are applicable at a higher risk assessment tier than the current methods. Geo-referenced 'real' datasets are applied instead of generic or average values. To account for temporal variability and uncertainty, a new type of PEC is introduced. These new PECs are defined as statistical time-distributions of predicted chemical concentrations in specific environments. PECs will be visualized and spatially analyzed by means of a Geographic Information System (GIS).

GENERAL SIMULATION APPROACH

From chemical market data, combined with information on the location of consumers and their emission habits, geo-referenced domestic chemical emissions are predicted. These emissions are further processed in sewer (and/or small surface water) and treatment models, to obtain predicted chemical fluxes to rivers. The discharge fluxes are entered into a river model, resulting in (geo-referenced) predictions of chemical concentrations in the considered river systems.

To deal with statistically distributed inputs and outputs, a hybrid simulation approach is used, involving both stochastic and deterministic techniques. The model core is deterministic. By means of Monte Carlo simulation, a stochastic layer (e.g. NRA, 1990) is added on top of this core. A large number of 'shots', which are discrete samples from the distributed data set, are generated. For each of these 'shots', the deterministic model is called, which contains a mechanistic description of the considered processes in the rivers and in the waste water drainage areas. Process rates are derived from knowledge about chemical properties and process specifics. Finally, the (discrete) results from each 'shot' are statistically analyzed, to obtain distributed results as simulation output.

For reasons of data set simplicity and performance, only steady-state model formulations are applied. Hence, a number of fundamental assumptions are made: constant chemical consumption, constant flows (within each individual steady-state model run), and constant environmental properties (e.g. waste water treatment infrastructure, river properties). If needed, the absence of true steady-state conditions (e.g. due to diurnal patterns in product consumption) and the influence of long-term variations can be implicitly incorporated into the data set through the parameters' statistical properties.

To allow a straightforward mass-balancing approach, all determinands (chemical levels, water flows) are expressed as fluxes. Chemical mass fluxes Φ *(mass / time)* are applied to describe chemical loads. Water flows are expressed as volumetric fluxes Q *(volume / time)*. In the models, chemical concentrations C *(mass / volume)* are not used, because they are not independent of water flows, and they do not describe chemical transport. When concentrations are explicitly required, they are calculated as: $C = \Phi / Q$.

Simulations can be performed for different scenarios (i.e., evaluations of different chemicals and chemical consumption patterns). The simulation input consists of a scenario-independent and a scenario-dependent data set. These data are expressed as statistical frequency distributions, incorporating seasonality and parameter uncertainty. Environmental characteristics are constant, and hence non-scenario-dependent. These include the river network structure, flow and flow velocity distributions, discharge point locations, treatment plant information, emission data, properties of sewers and small surface waters, etc. Chemical-specific information is scenario-dependent: chemical properties (i.e., biological, chemical and physical properties, specific process rates,...), and chemical market data (i.e., per capita product consumption rates). Market data are geo-referenced in the same way as the waste water information (i.e., related to waste water discharge points).

The simulation results (PECs) are frequency distributions of chemical concentrations, encapsulating temporal variability. For risk assessment purposes, these can be expressed as mean and 95th percentile values. Two PEC types are produced: concentrations in each stretch of the river network, and concentrations in the small surface waters of the waste water drainage areas. They are geo-referenced in the same way as the input data set: river PECs are associated with a river network structure, and waste water drainage area PECs are associated with discharge points. Within one water body, a differentiation is made between the maximal PEC (i.e., upon discharge), the minimal PEC (i.e., after degradation processes), and an average value.

SEGMENTATION

A strict segmentation is applied at all levels (Figure 1). Geographies are divided into geographical segments, which are connected in parallel or in series. In each geographical segment, one or more processes (connected in parallel or in series) occur. Each process further consists of one or more sub-processes (connected in parallel). Each sub-process has an input and an output terminal. Input and output operations, as well as model calculations, are performed at the sub-process level.

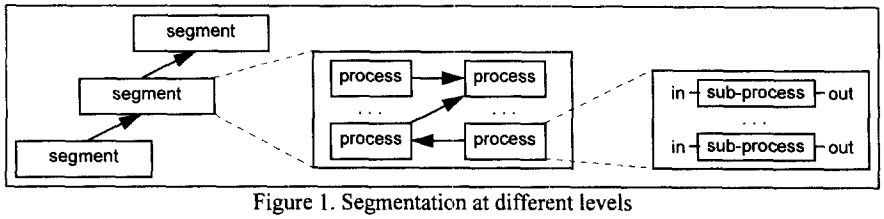

Figure 1. Segmentation at different levels

A river system is represented by means of a digital river network, which consists of an interconnected set of 'main river' stretches. The segmentation is determined by the occurrence of homogeneous environmental conditions and constant flow distributions within each stretch, and by the location of discharge points. The selection of 'main rivers' depends on the applied geographical scale (see below). With each waste water discharge point into the river, a waste water drainage area is associated, where chemicals are emitted, transported and possibly removed. The total drainage area of a 'main river' is segmented into Geographic Units (GU), each of which represents the drainage area of one discharge point, and is associated with one 'main river' stretch (Figure 2).

A geographical segment consists of a 'main river' stretch, and (if a discharge occurs) a Geographic Unit (Figure 2). Interconnection between segments is through the 'main river' stretches. With one discharge point into the 'main river', one GU is associated. Hence, all waste water drainage area processes upstream of the discharge point, are to be lumped into this single GU. These processes can be real (e.g. a single treatment plant) or hypothetical (e.g. multiple domestic discharges grouped into a single 'aggregated discharge').

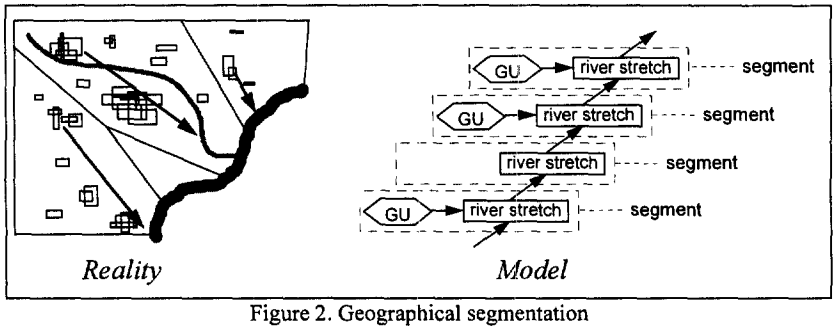

Figure 2. Geographical segmentation

The modeling and simulation methodologies, as well as the GIS data methodology, are scale-independent. In the upscaling process (i.e., moving from a smaller, detailed scale to a larger, less detailed scale), multiple discharge points can be aggregated into single (hypothetical) discharges; several smaller rivers are no longer considered as 'main rivers', but are transferred to the waste water discharge model and aggregated into a single (hypothetical) 'small surface water'. In the large scale approach, the mouth of a small rivers' catchment into a large river is represented by a discharge point. Hence, for different scales, only the geo-referenced data set is different; the applied models are identical. In Figure 3, a system is modeled at a small (left) and at a large scale (right). In this example, a complex system of 'main rivers' is used for the small scale, each with individual waste water discharges. In the large scale approach, the system is reduced to a single 'main river' with a single discharge point.

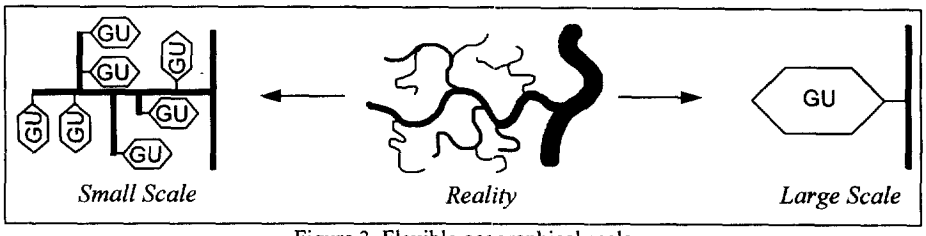

Figure 3. Flexible geographical scale

DETERMINISTIC MODEL

In environmental exposure assessment, several applications relate to 'new' chemicals, of which the safe use is to be assessed. Environmental concentration measurements of these chemicals, which are required for statistical modeling (e.g. Helsel and Hirsch, 1992), can not be obtained. Hence, knowledge-derived deterministic models should be applied.

In the deterministic model, all geographical segments are sequentially simulated (upstream to downstream). For each segment, the influent (from upstream segments) is calculated; if required the waste water pathway simulation is performed; the 'main river' processes are simulated; and finally the effluent (flowing to downstream segments) is calculated.

Segment Selection

A recursive tree-walking algorithm is used to find the correct segment sequence. The algorithm climbs up into the river network from the most downstream segment (i.e., the root), until the most upstream segments are detected. At confluences, both upstream directions are climbed, one after the other. At bifurcations, the upstream climbing is ended if the upstream part has already been climbed before, i.e. when coming from the other side of the bifurcation. Next, the network is again descended. Each segment which is encountered during the descent is selected and simulated. This way, influent data (from upstream) are always available to downstream segments. An example of the selection methodology is shown in Figure 4 (the numbers indicate the sequence in which the segments are called for simulation).

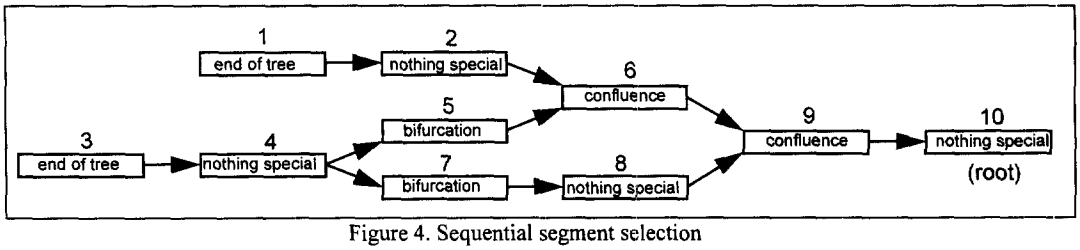

Figure 4. Sequential segment selection

Influent and Effluent Calculation

Segment influents and effluents are discrete values. For the influent, only chemical fluxes are calculated. Each segment's influent flow is set to the segment's 'main river' flow, which is taken as such from the (hydraulically consistent) geo-referenced data set. At a confluence, complete and instantaneous mixing is assumed, hence the influent is equal to the sum of the upstream effluents. At a bifurcation, the upstream segment's effluent is split into two fractions, proportional to the (known) river flows in each stretch downstream of the bifurcation. The effluent of a segment is identical to the effluent of the segment's 'main river' stretch. Hence, no effluent calculations are to be performed as such.

Individual Process Modeling Concept

At the conceptual level (described in this paper), models are considered as 'open boxes', each of which applies to one sub-process. The general model expression for each sub-process is:

$$x_{out} = a \cdot x_{in} + b \quad (1)$$

where x_{in} and x_{out} are the sub-process input and output values for the considered determinand, b is the emission value, and a is the conversion factor. Emissions and conversion factors are obtained at the detailed model level (which is not described here). At this detailed level, several model formulations (e.g. Monod or first-order kinetics describing biodegradation) and solution algorithms (analytical or numerical) can be applied.

Emission models are described by one emission value for each determinand. Their output is this set of emission values. Obviously, no input is required. Chemical fate is simulated by means of transport / conversion models. These are described by the chemical removal fraction R, which is calculated at the detailed model level. A ground water leakage fraction L can also be applied. L is part of the sub-process outflow fractionation (see below), taken as such from the geo-referenced data set. The input vector of a transport / conversion model, as well as the output vector, is a set containing a value for each simulated determinand.

Waste water pathway model

The waste water pathway model is used to predict the properties of discharges into 'main river' stretches. It consists of both emission and transport / conversion processes. Due to the nature of 'down-the-drain' chemicals, only point discharges are applicable (diffuse emissions, e.g. via runoff, do not occur).

Three types of chemical emission into the waste water pathway are considered:
- domestic emission (blackwater and greywater);
- non-domestic emission (industrial and agricultural); and
- land runoff.

Five transport / conversion processes are considered along the pathway:
- domestic on-site treatment (e.g. septic tanks);
- sewers (combined or separate);
- waste water treatment;
- ground water leakage; and
- small surface waters.

The outflow of each sub-process is split into a number of fractions, each of which is connected to one downstream sub-process. One or more downstream steps can be 'skipped', e.g. a domestic emission outflow fraction can be connected directly to the 'main river' discharge point. Emission processes are at the most upstream level - they do not receive any inflow. The sequence of the transport / conversion processes is: on-site treatment → sewer → treatment plant → ground water → surface water. The architecture of the interconnection between processes is identical for each segment. The (simplified) interconnection for domestic waste water pathway processes is shown in Figure 5.

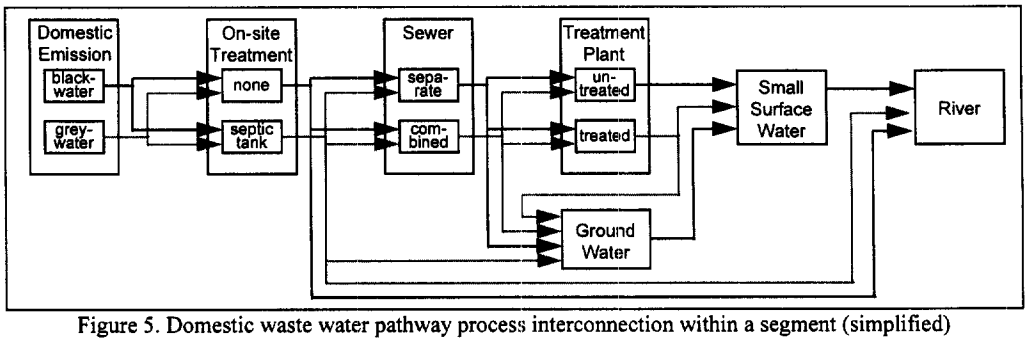

Figure 5. Domestic waste water pathway process interconnection within a segment (simplified)

River model

The river model is used for the 'main river' process of all segments. The model's inflow is the sum of the segment influent (from upstream segments) and the segment's waste water discharge (if applicable). The river model is a transport / conversion model, without leakage step. Water flows are not simulated but taken as such from the geo-referenced data set. Hence, the process is completely determined by the chemical removal fraction R, which is obtained at the detailed model level (e.g. Cowan et al., 1993; Trapp & Matthies, 1996).

Calculation Approach

The system of steady-state model equations for each determinand can be expressed as: (see above, equation 1)

$$X_i = A_i \cdot X_i + B_i \tag{2}$$

with X_i = state variables vector for determinand i, with elements x^i_k
A_i = (square) transport/conversion matrix for determinand i, with elements $a^i_{k,l}$
B_i = emission vector for determinand i, with elements b^i_k

The vectors X_i and B_i are partitioned per process, within a process per terminal, and finally within a terminal per sub-process (e.g. sewer - outputs - combined sewer). The process sequence in the arrays is from upstream to downstream; the terminal sequence is first input, then output. Hence, determinand values at a specific situation are independent of values at any subsequent situation (nor of themselves). Consequently A_i is a lower triangular matrix with zero-diagonal. In A_i the element $a^i_{k,l}$ is the conversion factor of x^i_l to x^i_k. In an input to output conversion, this represents the non-removed fraction. In a transport step, this is the fraction of x^i_l (upstream output) which is sent to x^i_k (downstream input).

Emission values are taken as such from the B_i vectors, hence the A_i rows referring to emissions are all-zero.

The A_i rows referring to transport/conversion inputs are partitioned into:

- a first set of columns (relating to upstream processes), which contain the upstream outflow fractions directed to the current process; and
- a second set of columns (relating to the current and to downstream processes), which is all-zero.

Outputs only depend on the (sub-)process inputs. The A_i rows referring to these outputs are partitioned into:

- a first set of columns (relating to upstream processes) which is all-zero;
- a second set (relating to the inputs of the current process) which is a diagonal matrix, with the input to output conversion factors for each sub-process on the diagonal; and
- a third set of columns (relating to the outputs of the current processes and to further downstream situations) which is again all-zero.

The equations can not be solved using the matrix calculation

$$X_i = (I - A_i)^{-1} \cdot B_i \qquad \text{with } I \text{ the unity matrix} \tag{3}$$

because the elements of the A_i matrices are only obtained as the calculations proceed: values in A_i at row k may be derived from (previously calculated) values of any determinand in any upstream process:

$$a^i_{k,l} = f\left(x^1_{1\cdots k-1}, \cdots, x^m_{1\cdots k-1}\right) \qquad \text{with } m = \text{number of determinands} \tag{4}$$

Hence, a sequential solution (from 'top' to 'bottom') is required:

$$x^i_\lambda = \sum_{k=1}^{\lambda-1} a^i_{\lambda,k} \cdot x^i_k + b^i_\lambda \qquad \text{with } \lambda \text{ varying from 1 to the total number of rows} \tag{5}$$

STOCHASTIC SIMULATION

The stochastic simulation takes into account seasonality of the determinands and parameter uncertainty. Seasonality deals with major environmental variation throughout the year(s). Parameter uncertainty deals with the difficulties to estimate model parameters, and with the inherent variability in specific processes.
By means of Monte Carlo simulation, discrete 'shots' of the data and parameter set (e.g. flows, process parameters, market data,...) are generated. With these discrete values, the entire geography is simulated, using the deterministic model (Figure 6). In particular, discrete segment inputs and outputs are applied. The main advantage of this approach, is that no correlation analysis between flows and chemical fluxes is needed. In an alternative approach, individual Monte Carlo simulations are performed for each segment. Hence, segment inputs and outputs are distributions. As flow and chemical flux are not necessarily independent, the correlation between them must be known to allow correct Monte Carlo sampling. However, for 'new' chemicals, this information may not be obtainable.

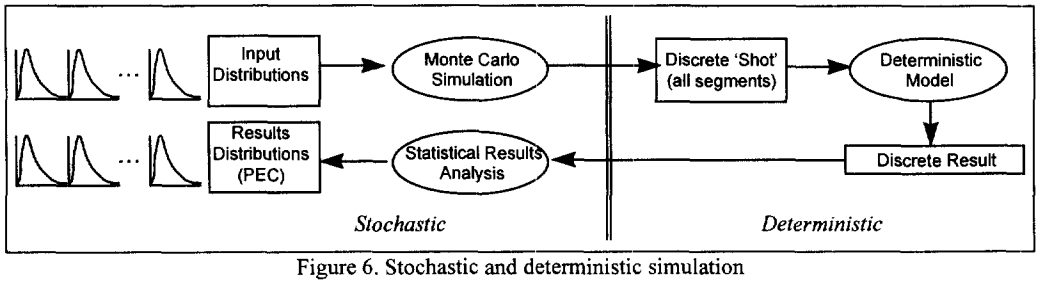

Figure 6. Stochastic and deterministic simulation

In each river stretch, a range of flows can occur. The probability of each individual flow within this range is given by the flow distribution, which is approximated by a log-normal distribution (NRA, 1990) and which is typically described by its mean flow and 95th percentile low flow (Q95) (Gustard et al., 1992). A flow scenario is defined as a percentile from a flow distribution.

For one 'shot', flow scenarios between segments are assumed to be fully correlated. For each 'shot' a single flow percentile (which is sampled from a uniform distribution between 0 and 1) is applied to all river stretches. This assumption implies that flow scenarios 'follow the water flow' from upstream to downstream. The time-lag between the occurrence of a specific flow percentile in one segment and in the immediate downstream segment, is assumed to be equal to the first segment's hydraulic residence time. For very large systems, the validity of this approach may be limited, due to major climatological differences in different parts of the catchment.

Within segments, the correlation between the waste water discharge flow and the 'main river' flow is to be specified in the data set. For large rivers, it will generally be assumed that these flows are uncorrelated. Dry weather waste water flow is obviously independent of the 'main river' flow. Wet weather flow in combined sewers depends on short-term rain events, rather than longer-term climatological conditions. One can expect the highest chemical fluxes to occur at high waste water flows (e.g. due to treatment plant bypassing and combined sewer overflows). Hence, for chemical risk assessment, the uncorrelated approach is probably the most appropriate, as it does not overlook this worst-case scenario (high chemical loads combined with low river flows). For small rivers, on the other hand, correlation of river flows with waste water discharge flows may become significant.

Parameter uncertainty is incorporated into the data set distributions. All parameters which are not flow-scenario dependent are assumed to be uncorrelated. Hence, the sampling from the distributions is performed independently.

CONCLUDING REMARKS & FUTURE PLANS

A geo-referenced simulation method for the prediction of aquatic exposure to individual 'down-the-drain' chemicals was developed. The fate of these chemicals, taking into account emission, sewage transport, treatment and river processes, is simulated by means of deterministic models. The natural variability, and the uncertainty in the required information, is dealt with via a Monte Carlo simulation. This finally results in statistical distributions of predicted environmental concentrations (PEC), which can be used for risk assessment.

A prototype of the presented simulation system was tested by means of a large sample data set, containing 16,000 geographical segments. A 1,000 'shot' Monte Carlo simulation took less than 1.5 hours on a Windows NT workstation (Pentium 150 MHz). This indicates the feasibility of the approach for detailed regional or large-scale pan-European simulations. Appropriate chemical fate models for each sub-process will be implemented into the simulator. The system will be applied to two pilot study areas (Yorkshire, UK, and Lambro, Italy) for the surfactant LAS (linear alkylbenzene sulphonate). An environmental monitoring study will provide measured data, required for model calibration and for results verification.

ACKNOWLEDGEMENT

The authors would like to thank the members of the ECETOC GREAT-ER task force, A. Young (IH, Wallingford), and F. Koormann and J.-O. Wagner (University of Osnabrück) for their valuable comments and suggestions. In addition, the authors would like to thank the Environmental Risk Assessment Steering Committee (ERASM) of the Association Internationale de la Savonnerie, de la Détergence et des Produits d'Entretien (AIS) and the Comité Européen des Agents de Surface et Intermédiaires Organiques (CESIO), and the UK Environment Agency for their financial support, and ECETOC for management support.

REFERENCES

Cowan, C.E., Caprara, R.J., White, C.E., Merves, M.L. & Gulloti, M.J. (1993). *A model for predicting the fate of 'down-the-drain' consumer product ingredients in United States rivers.* Water Environment Federation - 66th annual conference & exposition, Anaheim, Ca., Oct. 3-7, 1993, 351-358.

ECETOC (1993). *Environmental Hazard Assessment of Substances*. Technical Report No. 51.

ECETOC (1994a). *Environmental Exposure Assessment*. Technical Report No. 61.

ECETOC (1994b). *HAZCHEM, a mathematical model for use in risk assessment of substances*. Special Report, No. 8.

EEC (1993a). *Commission Directive 93/67/EEC of 20 July 1993 laying down the principles for assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/458/EEC.* Off. J. of the European Communities L 227/9.

EEC (1993b). *Council Regulation (EEC) No 793/93/EEC of 23 March 1993 on the evaluation and control of the risks of existing substances*. Off. J. of the European Communities, L 84/1.

EEC (1994a). *Commission Regulation of laying down the principles for the assessment of risks to man and the environment of existing substances in accordance with Council Regulation (EEC) No 793/93/EEC.*

EEC (1994b). *Technical guidance documents in support of the risk assessment directive (93/67/EEC) for substances notified in accordance with the requirements of Council Directive 67/548/EEC*, Brussels.

European Science Foundation (1995). *Problems and perspectives of regional fate modeling of xenobiotics. A strategy for enhancing the role of multimedia mass balance models for the stewardship of chemicals.* Proceedings of a workshop held at Hillerroed, Denmark, Jan. 30 -Feb. 1, 1995.

Feijtel, T., Boeije, G., Matthies, M., Young, A., Morris, G., Gandolfi, C., Hansen, B., Fox, K., Holt, M., Koch, V., Schröder, R., Cassani, G., Schowanek, D., Rosenblom, J. and Niessen, H. (in press). *Development of a Geography-referenced Regional Exposure Assessment Tool for European Rivers - GREAT-ER*. Chemosphere.

Gustard, A., Bullock, A. & Dixon, J.M. (1992). *Low flow estimation in the United Kingdom.* Institute of Hydrology report 108, Wallingford.

Helsel, D.R. & Hirsch, R.M. (1992). *Statistical Methods in Water Resources.* Studies in Environmental Science 49. Elsevier, Amsterdam.

Mackay, D. (1991). *Multimedia environmental models, the fugacity approach.* Lewis publishers. ISBN 0-87371-242-0.

Matthies, M., Koormann, F., Boeije, G. and Feijtel, T. (1996). *The Identification of Thresholds of Acceptability and Danger: The Chemical Presence Route*. EUROTOX '96 (Alicante), Proceedings.

NRA (1990). *SIMCAT consent setting model - User guide.* The National Rivers Authority's R&D Programme. Project No. 022, Output ref. P-61, NRA, Bristol.

RIVM, VROM & WVC (1994). *Uniform system for the evaluation of substances (USES), version 1.0.* National Institute of Public Health and Environmental Protection (RIVM), Ministry of Spatial Planning and the Environment (VROM), Ministry of Welfare, Health and Cultural Affairs (WVC). The Hague. Distribution No. 11144/150.

Trapp, S. and Matthies, M. (1996). *Dynamik von Schadstoffen - Umweltmodellierung mit CemoS. Eine Einführung*. Springer-Verlag, Heidelberg.