EUFRAM

Concerted action to develop a European Framework for probabilistic risk assessment of the environmental impacts of pesticides[1]

Work Package 10

SOFTWARE AND DATABASES FOR PROBABILISTIC ASSESSMENT[2]

December 2004

Leo Posthuma (RIVM)[3], Udo Hommen (IME Fraunhofer)[3], Tom Aldenberg (RIVM), Emilio Benfenati (Mario Negri Institute), Dick De Zwart (RIVM), Auteri Domenica (ICPS), Scott Ferson (Applied Biomathematics), Hector Galicia (Springborn), Andy Hart (CSL), , Robert Luttik (RIVM), Tharacad Ramanarayanan (BayerCrop Science), Donna Randall (US-EPA), Keith Solomon (University of Guelph), Theo Traas (RIVM), Paul Van Den Brink (Alterra)

Contents

[1] EUFRAM is supported by European Commission's 5th Framework Programme (www.cordis.lu), contract number QLK5 - CT 2002 01346. Further information and news about EUFRAM is provided at www.eufram.com.
[2] This document is the responsibility of its publishers and in no way represents the views of the European Commission or its services.
[3] Work Package (co)leaders.

1 OBJECTIVES

EUFRAM Work Package 10 has two objectives, viz.:

"…. (to) establish a common methodology for selecting software and databases for probabilistic pesticide risk assessments, *and* evaluate examples currently in use by partners for possible adoption as harmonised standards in the EU" (EUFRAM contract, 2002).

2 BACKGROUND

2.1 General

A workshop on "Probabilistic Risk Assessment for Pesticides in Europe: Implementation & Research Needs" (**EUPRA**) was held in the Netherlands, in June 2001 (Hart 2001). Broadly speaking, the added notion of "probabilistic" (**P**) to the concept of "risk assessment" (**RA**) refers to methods that attempt to quantify variability and/or uncertainty in factors that influence risk, and express risk in terms of the probability and magnitude of adverse effects. The methods currently used for these assessments are predominantly deterministic rather than probabilistic. They use fixed values for exposure, toxicity and risk, and attempt to allow for variability and uncertainty by using worst-case assumptions and safety factors. A probabilistic approach in risk assessment, **PRA**, would allow for variation and uncertainty by using distributions, instead of fixed values, for exposure, toxicity and risk.

Probabilistic methods are common or are being increasingly adopted for various purposes in human society, like insurance industry, safety design of e.g. nuclear power plants, and the design of dams and other constructions. PRA might also be explicitly useful in assessing the impact of plant protection products (pesticides) on the environment. Directive 91/414/EEC and its Annexes require EU member states to analyse these risks before authorising pesticides for sale.

2.2 EUPRA recommendations on Software and Databases for PRA

The EUPRA workshop considered the potential of probabilistic methods for assessment of ecological risks of plant protection products in relation to the above mentioned Directive. The EUPRA workshop reviewed the state of the art, and made detailed recommendations. Together, these recommendations were shaped into the EUFRAM-project, the "Concerted Action to develop a European Framework for probabilistic risk assessment of the environmental impacts of pesticides".

Amongst others, the EUPRA recommendations specifically addressed the availability and selection of **Software and Databases for probabilistic risk assessment**, since PRA requires software and data sets when the approach is to be practically implemented as assessment approach in the European practice.

This report specifically focuses on the EUPRA recommendations on Software and Databases. The detailed recommendations for this work package, according to Hart (2001), are:

1. that standard software tools for probabilistic assessment should be adopted, at a level of complexity appropriate for users in all parts of the EU regulatory arena.
2. that, in order to avoid duplication of effort, consideration should be given to whether tools developed elsewhere could be appropriate for use in the EU, either in whole or in part.
3. that there is an urgent need to catalogue existing data that would assist the development and operation of probabilistic approaches, and to collate these data in publicly-available, quality-controlled databases. The catalogue of data should include pesticide-specific data (e.g. toxicity, for use in developing SSD methods) and general data (e.g. geographical and ecological data).
4. that databases and software should be made easily accessible, to promote harmonised approaches

2.3 Current situation and problems in Software and Databases for PRA

Regarding the availability of **Software**, the EUPRA report already mentioned the existence of a relevant body of work of US-EPA (http://www.epa.gov/scipoly/sap/index.htm), but it is unclear whether the approach can be applied as is in the EU-context.

Regarding the availability of **Data**, it was noted that some data-sharing initiatives are already underway (the RED-project, the SEEM-project, the US-EPA Ecotox database), but that data sharing may be hindered by issues related to commercial interest, confidentiality and data ownership.

Regarding **Application** of Software and Databases, an array of technical limitations hampers the application of PRA-techniques in practice. As yet, no single, easily accessible database covering all data sources that are needed has been designed, and it is questionable whether combining a set of existing databases would solve the issue. Similarly, software has mostly been developed for in-company use and for various separate steps in the PRA-process. However, the existing software programs do not systematically cover all steps that can be identified in a systematic framework for PRA, while arrangements to guarantee broader availability are lacking. Finally, the requirements that usually accompany systematic application within a legal framework are not met.

Solutions for these problems require co-operation between the owners of the data and software developers. Eventually, sharing data and software would provide benefits to all stakeholders. Regarding the need for extra data, the collaboration could indirectly lead to a reduction in the amount of animal testing required in future.

2.4 Steps taken so far to address prevailing problems

To address the EUPRA recommendations, **EUFRAM Work Package 10** was designed, to focus specifically on Software and Databases. This report provides a compilation of criteria for evaluating databases and software, and, by using those criteria, it provides a first evaluation of some of the operational databases and software items currently available.

The work in Work Package 10 has been formatted according to the following steps:

1) The making of an inventory of existing Software and Databases, resulting in a Preliminary paper composed from contributions written by experts in the field (Posthuma *et al.* 2003),

2) The definition of preliminary criteria to select Software and Databases that could be appropriate for any relevant step in the process of PRA (Posthuma *et al.* 2003), and

3) The targeted evaluation of the inventory of Software, Databases and preliminary criteria at an expert workshop on Software and Databases for PRA (held in June/July 2003), in which both the preliminary criteria were discussed and refined, and in which the existing Software and Databases were confronted with the refined criteria, to identify the currently most promising examples for implementation (De Zwart *et al.* 2003).

This report marks the phase of compiling the work done so far and serves as basis for discussions for the subsequent steps, i.e., comparison with results from the other Work Packages. This implies a broadening of the discussion and a move towards practical applications in case studies (within EUFRAM, next step) and final implementation in regulations (after EUFRAM).

After the discussion and comparison to the results of the other Work Packages (October, 2004), the following phases will thus follow:

4) testing and use in selected case studies (2004/5)

5) selection of the most promising items and training of end-users at the EUFRAM end-user workshops 1 – 3 in 2005 to 2006.

2.5 Contents of this report

This report presents the Work Package results as obtained until present (September 2004) on the steps 1 – 3, esp. on:

- the common methodology for selecting Software and Databases for probabilistic pesticide risk assessments, that is: to list the criteria to evaluate Software and Database items within the EUFRAM framework
- the inventory of existing software and database items
- the evaluation of the examples that have been presented at the Software and Database workshop of 2003, and that are currently in use by partners, using the evaluation criteria

These results are presented after the description of the possible processes in PRA for pesticides, to provide some background information on the issues being discussed and their conceptual linkages.

Eventually, the steps 4 and 5 could lead to the possible adoption of Software and Database items as harmonised standards in the EU.

3 IDENTIFICATION OF SOFTWARE AND DATA NEEDS IN PRA

3.1 Schematic overview

Software and databases for PRA of plant protection products will eventually serve within the framework being developed and tested within EUFRAM. Thus, the precise requirements imposed by the EUFRAM-framework will only be clear *after* publication of the Framework to be produced by EUFRAM. Nonetheless, insight in the types of assessments that can be the subject within a PRA is crucial to focus the attention to the right issues. That is: those issues that are likely part of the eventual PRA-process designed by EUFRAM. This Chapter provides the schematic background for the analyses that are made.

Risk of a compound is commonly considered to be present when exposure exceeds a (conceptual) no-response level, or (more general) a relevant sensitivity parameter. This general definition shows that risk assessments consist of both an exposure and an effect assessment.

A general scheme showing the association between the different aspects (potentially) relevant for PRA is given in Figure 3-1. Exposure and sensitivity and associated effects are shown in the central part of the Figure, and are addressed by modelling (either or not validated or verified by measurements in the exposure medium or on effects observed in exposed organisms).

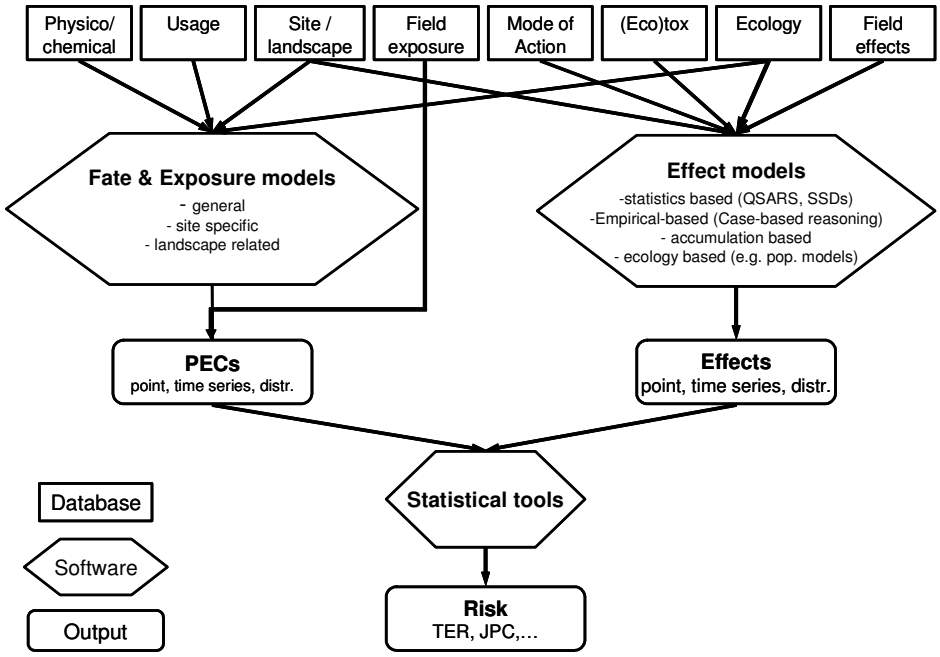

Figure 3-1. Schematic representation of possible steps in risk assessments, identifying the 'position' of Software, Databases and Output. For explanation: see text.

Both exposure and effects are influenced by an array of factors. Conceptually, these range from the physico-chemical properties of the compound, usage pattern, site/landscape characteristics to the ecological niches of the exposed species when

considering exposure. When considering effects, the array is from site and landscape data, via the toxic mode of action, ecotoxicological sensitivity, ecological characteristics of the species to field effects (upper part).

For EUFRAM, it is assumed that field data of exposure are usually not available (i.e. for registration of new pesticides), but that exposure is predicted by modelling. Equivalently, the (conceptual) no-response level (or sensitivity of the exposed or target organism) is not known. It can e.g. be predicted from a statistical description of sensitivity data (a Species Sensitivity Distribution, **SSD**), in the derivation of which one may or may not include incorporation of (aut)ecological information to select the species data pertinent to the assessment target.

Both the (statistical) distribution of (predicted) exposure concentrations (**PEC**s, in space or time) and the distribution of effects are obtained by modelling (middle part). In both cases, the model may be using simple assumptions (use only few parameters of the top selection) and approaches, or more complex ones, the latter according to a tiering principle. The PEC data are 'overlaid' with the (statistical) distribution of effects, in most cases an SSD, e.g. of No Observed Effect Concentrations, **NOEC**s) to produce a so-called Joint Probability Curve (**JPC**) and to calculate the final risk parameters, of which various (equivalent) forms exist, e.g. the Expected Total Risk (**ETR**) and the Ecological Risk (Aldenberg *et al.* 2002), (Solomon & Takacs 2002), (Van Straalen 2002) and others). As an example see Figure 3-2.

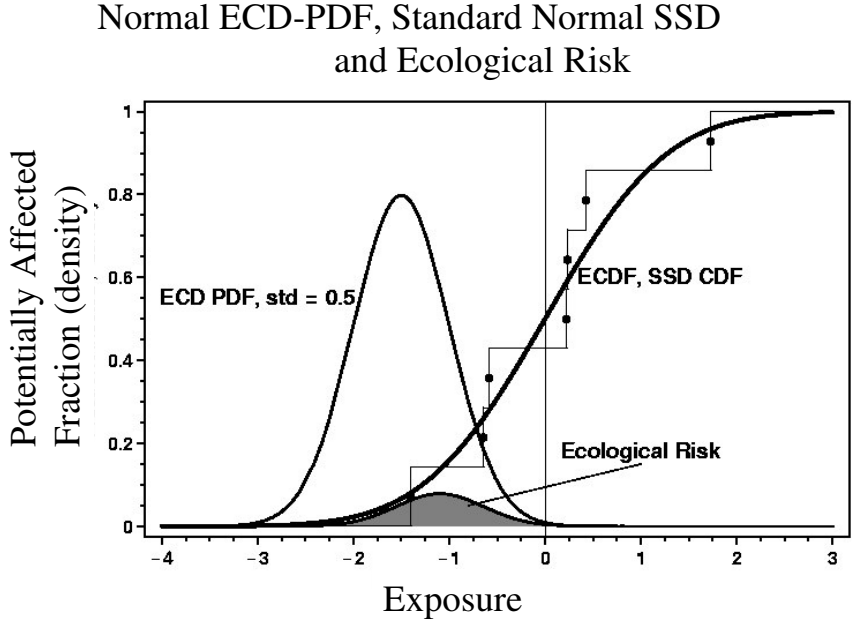

Figure 3-2. An example of a probabilistic approach in ecological risk assessment. Sensitivity data (a Species Sensitivity Distribution-Cumulative Distribution Function, SSD-CDF) are overlaid with exposure data (an Exposure Concentrations Distribution-Probability Density Function, ECD-PDF). The ecological risk is calculated from both distributions according to (Aldenberg *et al.* 2002).

Aldenberg (pers. comm.) has shown that all models in which this type of overlays are produced, though described in different terms, boil down to one, universal statistical approach. Moreover, as linkage to the deterministic approach in risk assessment, the

well-known PEC/PNEC ratio approach, or the equivalent Toxicity Exposure Ratio (**TER**) approach used for pesticides in Europe, can statistically be seen as the simplest form of a JPC-approach. That is: one that consists of the ratio of two point estimates, of PEC and PNEC, rather than of broader distributions.

While the JPC-approach is probabilistic in kind, the first principles of this approach do not necessarily reach farther than the use of statistical distributions of PEC and PNEC *per se*. That is: the first principles of the methods do not necessarily use (aspects of) ecological theory to predict risk. In the right part of the Figure, 'ecology-based models', such as population models, are identified. In this case, the risk assessment approach is (at least in part) building forth on ecological first principles. For example, the risk for extinction (or *vice versa*: the probability of maintaining a viable population) can be calculated using e.g. information on birth rate, survival time, juvenile period and death rate (and distributions thereof) under pesticide exposure.

The scheme is worked out below in more detail.

3.2 Aspects of Fate and Exposure

The issue of 'Exposure' (left side of Figure 3-1) is commonly addressed by Fate and Exposure Modelling. Both terms can be used (often interchangeable) in the context of risk assessment. These models (when interpreted in their narrow-sense meanings) capture two processes:

1. Fate modelling: Modelling used to predict the place where compounds used in agricultural practice 'end up' in the environment. That is: to predict the local concentrations in environmental compartments
2. Exposure modelling: Modelling used to predict the level of exposure experienced by an ecological receptor (e.g., an individual, a species, a community), given a certain concentration in an environmental compartment. That is: to predict the biologically relevant concentration of a compound. A compound needs not be completely available for uptake due to partitioning processes (e.g., sorption to the matrix), or due to specific ways in which organisms explore their environment (niche choice, behaviour, food choice).

Fate and exposure models may require the following types of data:

1. Compound-specific data: Data on the inherent physical and chemical properties of the compound, used as key parameters in Fate and Exposure models
2. Usage data: Data on the use of the compound (amounts, crops, periods of use, application methods), used in Fate and Exposure modelling
3. Site-characteristics: Data on the landscape in which the compound is used, including data on matrix properties (e.g., soil type), for Fate and Exposure modelling.
4. Characteristics of exposed organisms: Data on the exposed species, in terms of routes of exposure (food intake rate, et cetera)

Despite the interchangeable use of the terms, the emphasis in modelling is often on fate modelling.

The data (3 – 6) can be obtained only in part from generic databases. Generic databases pertain to the 'universal' (or 'generic') properties of the system, such as pertaining to compound properties like the octanol-water partition coefficient, the

fugacity, et cetera. The other part contains modelling parameters that are of interest for site- or species specific assessments, such as in the case of using soil and/or landscape properties of an identified site to predict fate and exposure levels at such a site. Further, 'intermediate' data sets can be envisaged, such as databases on measured crop residues, which may be used in risk assessment next to- or instead of Predicted Environmental Concentrations.

The combination of data from databases with the Fate and Exposure models eventually results in the *predicted fate or exposure*, either in scenario conditions (hypothesised conditions as stand-in for real world possibilities) or in real conditions (an identified site for which an assessment is needed). This can take the format of a single-value estimate, or of a distribution of fate/exposure data in space and/or time.

3.3 Aspects of Effects data

The issue of 'effects' (middle-right part of Figure 3-1) can potentially be addressed in various ways. The more 'refined' a risk assessment should be (higher tiers), the more specific the effects data are needed, and the more realistic the assessment approach is in comparison to the system being assessed.

In lower-tier, generic approaches, statistical distributions of sensitivity data that were generated in laboratory conditions might be considered sufficient as first principles. For example, one can use lab-derived NOECs to construct the relevant Species Sensitivity Distribution (Posthuma *et al.* 2002a) using all available data. In more specific approaches, one might focus on certain species or species types typical for a given landscape, or to the exposure level of interest (e.g., use EC50s to construct the SSD rather than using NOECs, see e.g. (Posthuma *et al.* 2002b). This implies species selection or endpoint selection prior to SSD-construction, which would require a database containing different sensitivity endpoints (NOECs, EC50s, ECx), a database on (aut)ecological preferences of species and a database on biogeographical distribution ranges of species.

In assessments where specific protection endpoints are defined, e.g., in terms of risk for extinction of identified species, attention would be focused on the (aut)ecological characteristics of the species, to fill out how the parameters of the population model respond to exposure.

As final issue, pesticides are often developed to elicit specific toxic action. This relates to the concept of (Toxic) Mode of Action (**TMoA)** - a concept pertaining to the molecular-level interactions between a compound molecule and a receptor molecule inside an organisms' tissues. Risks for organisms possessing the target molecule are expectedly larger than for organisms that lack the receptor site. For the latter organisms, the presence of only baseline toxicity would mark the 'ideal pesticide' in this respect. However, also apparent (secondary) toxicity is possible, whereby the compound molecule interacts with non-target molecules, to cause toxicity to an extent that is often expected to be smaller than for target toxicity. This shows up as different slopes and positions of SSDs constructed of data sets of target organisms as compared to those of non-target organisms (Figure 3-3).

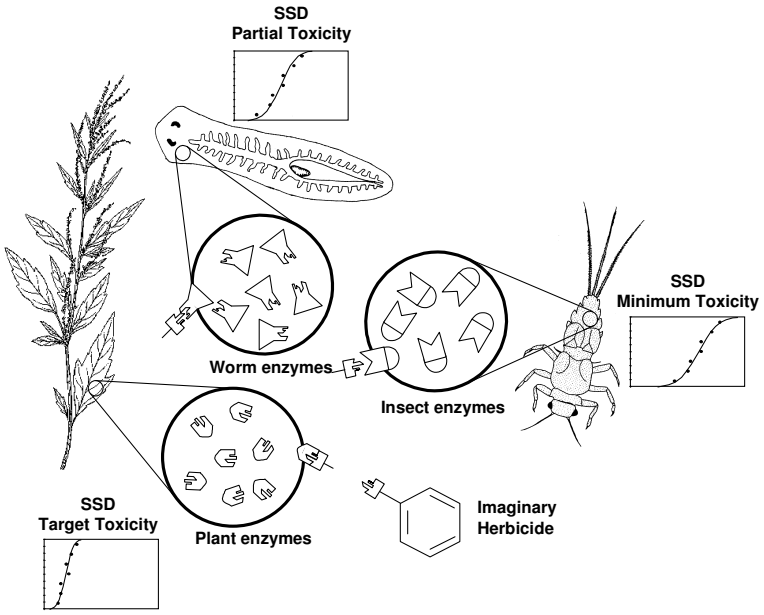

Figure 3-3. Conceptual representation of the influence of TMoA on the shape of the SSD curve when considering classical, toxicological target toxicity, and other effects.

Currently, although the TMoA of a pesticide is often well-defined for the target (pest) organism, it is a matter of further scientific development as to assign primary and secondary (or even tertiary) TMoAs to address non-target organisms. This would be an addition to the concepts of baseline toxicity and primary target toxicity as accepted theories (Posthuma *et al.* 2002b).

The issue of TMoA, and its 'capture' in databases and theory and models has recently gained in importance. In the evaluation of new pesticides, it is useful to consider if data are available for substances with similar TMoA as the compound of concern. Such data are useful to obtain a preliminary view on the risks of a compound, for example by the Case-Based Reasoning approach followed in PERPEST, see www.perpest.alterra.nl (Van Den Brink *et al.* 2002).

3.4 Software and databases potentially needed for PRA

Given the schematic steps that can be part of PRA (Figure 3-1) one can easily deduct the complete set of databases and types of software needed for an assessment of pesticide risks. From the databases, one might apply widely defined data queries to select a large set of data assumed relevant for the assessment. For example, one may select one or all NOEC-data to construct a single sensitivity distribution. From the models, one may choose models founded in statistics, e.g., the distribution of sensitivities (an SSD). Alternatively, one may select models founded on conceptual descriptions of the phenomenon (e.g., population models), on truly mechanistic phenomena (molecular interaction of the pesticide molecule with receptor molecules), or one may use purely empirical models (e.g. Case-based reasoning).

Key to EUFRAM is the notion of 'probabilistic'. This means that distributions of data are modelled rather than point estimates, which asks for specific features of the model, and of the 'software environment' in which the model is implemented. Since

there are 'standard' software environments ('shells') that are specifically designed for probabilistic approaches, those shells can be considered as relevant issue for evaluation within the EUFRAM development process.

Probabilistic approaches also ask for specific attention for options to allow for the required ways of data querying. Usually, an array of data is needed to grasp variability. Further, for the registration of new pesticides, it could be profitable to have options to select data from (toxicologically) related compounds, and highly profitable to have access to (as yet) non-public data.

EUFRAM may eventually develop a probabilistic risk assessment approach based on tiering principles, or not. Decisions on this aspect are taken outside of the scope of this report.

4 SOFTWARE AND DATABASE EVALUATION CRITERIA

4.1 Approaches to define evaluation criteria

Evaluation criteria for software and databases can be taken from different open sources. Gross criteria lists can be found easily, focusing on the general sets of criteria that Software and Database developers apply to optimise their products. The sites http://www.so.cc.va.us/vccsit/Archive/SoftEval.htm and http://www.so.cc.va.us/vccsit/softchek.htm give examples of general criteria and general checklists. These general inventories show that software and database developers focus on an array of checkpoints, under various main headings. The criteria headings range from purely technical criteria for the IT-infrastructure up till criteria focusing on the population of users (e.g., are users trained or not; can training be easily arranged) and legal criteria (administrative requirements, e.g., under EU-regulations).

Since general checklists like those cited above are not tailored to the EUFRAM targets, a tailored inventory of criteria was made amongst current experts in (Probabilistic) Risk Assessment. This process has led to a draft list of potential objective criteria applicable to EUFRAM, as shown in Table 4-1 (for software for PRA) and Table 4-2 (for databases for PRA).

The criteria are grouped under various main 'headers', focusing mainly on science and application in PRA-practice (both from a users' point of view as well as the point of view of the legal context).

Table 4-1. Inventory list of objective criteria that can be used to evaluate the appropriateness of an available software item for the eventual EUFRAM targets.

Category	Criterion	Subcriterion
Science	Model coverage	All PRA aspects
		Part of PRA aspects: fate
		Part of PRA aspects: exposure
		Part of PRA aspects: effects
		Part of PRA aspects: risk
	Model underpinning	Published
		Validation available
	Model specifications	Option to account for variability and uncertainty
		Option to choose optimal model for problem
		Options to shape PRA approach to problem
		Custom-made tools (e.g., in Excel) appropriately execute desired function
		Built-in tools (e.g., in Excel) appropriately execute desired function
Application	Ease of use	'Feeding' of input data
		Exporting of output
		Printing of output
		Screen interface
		Shape of output format: numerical
		Shape of output format: graphical
	Usefulness of output	Option to show confidence bands
	Shell availability	Broadly dispersed and known
		Free availability
		Current dispersal over potential users
		Windows-based
		Web-enabled use possible
	Format	Dependency of executable of shell version
	Technical capabilities	No limitations on required data entry (max. n of rows/columns)
	Future use / legal adoption	Suitability for EUFRAM targets

Merging of the overall ideas from (general) software and database developers and experts in the field provide the list of key criteria on which specific focus is to be put for developing the EUFRAM framework. Such a merging of ideas took place during the Software and Databases workshop organised in the EUFRAM context, June 2003.

The development process (including the workshop discussions) implies that the criteria have been *listed* as a long-list before existing software and database items were exhibited and before some hands-on experience was gained, and that they have been *refined* and *specified for PRA-practice* in the EUFRAM-project *after* gaining hands-on experience.

In the subsequent paragraphs, the criteria for software and databases are addressed. The criteria provided below can be not only interpreted as selection criteria (i.e., to choose amongst existing software and database items), but also as guidance to software and database development. The experts' views provide the general criteria of the ideal software and databases that should be developed in case such matters lack.

Table 4-2. Inventory list of objective criteria that can be used to evaluate the appropriateness of an available database for the eventual EUFRAM targets.

Category	Criterion	Sub-criterion
Science	Data validity	Raw data = data entry
		Quality earmark
	Data aspect	Fate
		Exposure
		Effect
	Data context	Values for modifying factors (contextual data, covariates)
	Data quality	Quality of the study (e.g, GLP, compliance with regulation x)
		Identification of 'doubles'
		Indication of uncertainty on estimated values (e.g., EC50 with C.I.)
		Information on limit of detection
	Data coverage	Allowing for versatile data query / sub-selections
		Wide coverage of existing data
	Database	Published
Application	Ease of use	'Feeding' of new data
		Exporting of output
		Exporting of output selections only
		Printing of output
		Screen interface
		Linkage to software for PRA
	Shell availability	Broadly dispersed and known
		Free availability
		Windows-based
		Web-enabled use possible
	Format	Dependency of executable of shell version
	Technical capabilities	No limitations on required data entry (max. n of rows/columns)
General	Future use / legal adoption	Suitability for EUFRAM targets

4.2 Definition of EUFRAM evaluation criteria

During the workshop, the draft general and 'expert'-biased criteria lists were discussed for software and databases in view of the targets of EUFRAM. Criteria were assigned to the categories 'highly desirable', 'desirable', or 'undesirable' for EUFRAM. In case there were no good 'candidate software or database items' for adoption as EUFRAM standard, these criteria can be of help in the further development of software. The results (refined criteria) are summarised in Table 4-3 for software and in Table 4-4 for database items. It should be noted, that the criteria that were mentioned will not uniformly apply to all models. The scope of the model will eventually determine the criteria that will apply most. For example, if one wants to model a specific habitat, then it is desirable that the model can be recalibrated to obtain the right level of detail. For other modelling problems, recalibration might add too much degrees of freedom for an assessment, e.g., for standard regulatory assessments. The Tables show that the workshop focused on positive rather than negative criteria: many of the EUFRAM-specific criteria were highly desirable or desirable. Only few criteria for software choice and development were characterised as 'undesirable'. The characterisation was, in those cases, addressing the fact that the characteristics would be undesirable in the case of lower-tier, regulatory assessments. In such cases, it is, e.g. in modelling, undesirable to have a wide choice of distributions, since this would lead diversification of results obtained, e.g., dependent on the preferences of the user.

Table 4-3. EUFRAM expert workshop suitability criteria for Software choice and development.

	Highly desirable	Desirable	Undesirable
Scientific characteristics			
Ability to handle different groups of compounds	X		
Multiple Spatial scale (global/regional/local)		X	
Ability to handle (pesticide) mixtures		X	
Ability to handle metabolites		X	
One stop shop (incl exposure, effects and risk		X	
Option to account for uncertainty	X		
Option to account for variability	X		
Option to account for dependencies/correlation	X		
Empirically validated (models only)	X		
Flexible model structure	For research	For higher tier	For lower tier regulatory
Wide choice of distributions	X		For lower tier
Published in the peer-reviewed literature		X	
Expert panel peer review	X		
Option to consider temporal variability/uncertainty	For higher tier/research	X	
Option to consider spatial variability/uncertainty	for higher tier/research	X	
Requires more than typically available data			for lower tier regulatory
Requires recalibration for new regions/countries			X
Contains non-data-derived adjustment factors			X
Technical			
Quality assured software	X		
Version control protocol available	X		
Model fully documented	X		
Input data error checking (incl. units)	X		
Backwards/forward compatible	X		
PC based	X		
Broadly dispersed and known shell	X		
Potential to link to other software/databases		X	
Dependence on other proprietary software			X
Ease of use			
Exporting of output	X		
Reproduce/save past run	X		
Help screens	X		
Intuitive interface	X		
Free of use restrictions	X		
Any windows-based application	X		
Ability to cut/paste/link input data		X	
Roll-over (mouse-over) help		X	
Flexibility in data input (e.g. units)		X	
Reference fields for input data		X	
Default data and worked examples		X	
Differential User-modes (default and advanced		X	
Execution in minutes Vs. hours on Std. PC		X	
Free of costs to user		X	
Web-enabled use		X	
Support and maintenance			
Feedback loops to improve software	X		
Related to regulatory context	X		
Long-term support, updating, and maintenance	X		
User support		X	
User groups		X	

Table 4-4. EUFRAM expert workshop suitability criteria for Database choice and development.

	Highly desirable	Desirable	Undesirable
Scientific characteristics			
Modifying factors (contextual data, covariates)	X		
References for data documented	X		
Quality criteria documented	X		
Identification and handling of 'duplicates'	X		
Information on limit of detection for non detects	X		
Quality measure (e.g. measured / nominal data)	X		
Statistical method described		X	
Indication of uncertainty on estimated values (EC50 with CI)		X	
Quality of the study (GLP, compliance with regulation x)		X	
Reported with original number of digits		X	
Slope of the response lines used to estimate the EC50, LOEC or NOEC		X	
Information to judge the power of tests (e.g. sample size, statistical power)		X	
Raw data included		X	
Links to related compounds		X	
Comprehensiveness of data per endpoint		X	
Comprehensiveness of endpoints		X	
Ease of use			
Exporting of output selections	X		
Free of use restrictions		X	
User friendliness of search terms (common names)		X	
Ability to reproduce past query		X	
Flexibility in data input (e.g. units)		X	
Free of costs		X	
Includes summary tools to qualify sub-selections		X	
Software characteristics – technical			
Broadly dispersed and known	X		
Format remains in exported data	X		
Error checking routines for input	X		
Database documentation	X		
Extensible (by database manager)	X		
Allows for versatile data query / sub-selections		X	
Keeps track of changes to records		X	
No aggregate (text) fields (e.g. min & max instead of range)		X	
Web-enabled use possible		X	
Windows-based		X	
Open design for software access		X	
Support and maintenance			
Feedback loops to improve context of database	X		
Related to regulatory context		X	

Overall, the Tables show that a practice-weighted view on the scientific state-of-art in risk modelling and the data needed for that has been adopted. For the models, this means that the criteria have not been set at a currently and conceptually ideal level, but at a level that can be reached in the near future. The latter also depends on the availability and quality of the available data to 'feed' the models. For example, the idea to have all the raw data of an ecotoxicity study being reported in a database (one of the criteria) could have received universal support as being highly desirable from a conceptual and flexibility point of view. However, this criterion would lead to rejection of the current databases (e.g., for ecotoxicity data currently > 150.000 records), in which commonly only *derived* data are given, such as EC50s or NOECs. Adapting such a database to store the original raw test data would require a major

effort, generally viewed as unrealistic. Evidently, the situation described here in the example is already data rich. This means that the ideal model, if any, should also be considered with view on the available data, with reference to both quality and amount.

4.3 Summary properties of ideal software and databases for PRA

The Tables provided in the previous paragraphs during the workshop contain subjective evaluations of criteria being highly desirable or desirable, reflecting the balance of workshop participants' views of their relative importance, weighted by their views of reachable goals. The lists are not intended as absolute criteria for acceptance or rejection of particular tools. Rather, they are intended as criteria that should be considered by prospective users when choosing between existing tools, and as a guide for the development of new or improved tools.

With a view upon the end-users of the software and databases for PRA, the workshop defined a more comprehensive view, superimposed on the separate tables for software and databases. Based on the criteria in the Tables, such an ideal system might be characterised as follows. It would:

1. provide a comprehensive set of models for assessing the ecological risks of pesticides and their metabolites, including both exposure and effects.
2. provide comprehensive, appropriate, referenced, quality-controlled data for all model inputs including pesticide use, physico-chemical properties, toxicity, ecological and landscape factors, and field data.
3. include dynamic links between data and models.
4. include appropriate methods for taking account of uncertainty, variability and dependencies, and for incorporating spatial and temporal variation.
5. offer different modes for different users and purposes, including standardised models and data for lower tier assessments and more flexibility for higher tier assessments.
6. be fully tested and documented, approved by appropriate expert panels, published in peer-reviewed literature and, as far as possible, empirically validated.
7. run efficiently on computer systems typically used by regulatory assessors, or be usable remotely via the internet.
8. have an intuitive user interface including automatic error-checking, extensive help facilities, and convenient methods for input and output of data and results.
9. be provided with long-term maintenance and user support services.
10. be suitable for regulatory use at both national and European level.

Researchers are likely to require more flexibility and specialist tools, that would be difficult to incorporate in a single system, but they would still benefit from the other characteristics mentioned.

It is unlikely that such an ideal system will be achieved quickly, but it is useful to indicate the direction in which developments might be most useful. In the short term, available tools may be judged by the extent to which they meet this ideal, or more specifically the criteria in the Tables (see next Chapter).

5 INVENTORY OF SOFTWARE ITEMS

5.1 Overview

At present (September 2004), 26 software items are considered within the EUFRAM project (Table 5-1.).

The software items can be differentiated into

- general tools (not especially developed fort the risk assessment of pesticides but for analysis of uncertainty and variability),
- models to estimate fate and/or exposure,
- models to predict effects, and
- tools to combine probabilistic fate and effects estimations to calculate risks and calculate Joint Probability Curves.

This differentiation must not seen as an exclusive characterisation, because some tools integrate fate and effect estimation. Fate models are further differentiated with respect to their scale (general, site specific or landscape related), while effect models were divided into statistics based models (e.g. tools to calculate SSDs), ecological models (e.g. population models) and models which consider bioaccumulation.

The inventory shows that for each of the options needed in a PRA-scheme, as conceptualised in Figure 3-1, one or more software items have been developed:

- Crystal Ball (MS-Excel add-in) and Risk Calc form general shells for e.g. uncertainty analysis,
- an array of fate and exposure situations is addressed,
- various ways of effect prediction can be chosen, and
- 9 software items can be chosen to calculate risk from the knowledge of exposure and sensitivity.

All this does, however, *not* mean that adopting (few) models from the set solves the problem, that is: there is no set of current models that, even when combined, fits the description of the ideal set of EUFRAM modelling tools (see previous Chapter). The software items have been developed in different ‘languages’, which may or may not correspond to each other. Software may, in its implementation, be very specific for particular targets, excluding direct use in the EUFRAM framework.

In order to be able to investigate whether the set of models can be used as basis for further developments, the next paragraphs provide:

1) information on the software items addressed in the inventories and discussed so far
2) the confrontation of the software items to the criteria as defined earlier, showing results of the workshop

Evidently, an array of models has not been evaluated. These models were considered of too limited importance for further development for EUFRAM.

Table 5-1. Software reviewed for the use in probabilistic risk assessment within EUFRAM (status September 2004)

Acronym	General tool (e.g. uncertainty analysis)	Fate and exposure prediction			Effect prediction			Risk prediction
		General model	Site / situation specific	Landscape related	Statistics based	Ecology based	Bioaccumulation based	(e.g. JPC calculation)
Crystal Ball	X							
Risk Calc	X							
FOCUS MACRO			X					
FOCUS PEARL			X					
FOCUS PELMO			X					
FOCUS PRZM			X					
MCPELMO			X	X				
SimpleBox		X	X					
AQUATOX			X			X	X	
Daphnia						X		
DEMETRA					X			
PERPEST					X	X		
RAMAS ecotox						X	X	
Secondary Poisoning						X	X	
SimpleChain						X	X	
SSWD					X			
Zebrafish						X		
Busy	X				X			X
ETX					X			X
GPS1						X		X
PRAT2					X			X
ARRA			X		X			X
TIM			X		X			X
USES		X			X		X	X
WEBFRAM			X		X	X		X
IQ-Tox			X		X			X

Description of Software items

The following short descriptions were extracted from the summary and fact sheets of a meta-database developed for EUFRAM-WP10 in which all model and programming information is stored for detailed reference. For more information this meta-database with more specific information can be downloaded from the EUFRAM web site (www.eufram.com in the directory of this report (deliverable database)). Specific attention is given to the aspect “probabilistic”. This aspect is judged as “yes” when the software has an explicit probabilistic design, as a whole, or in major sub-models.

5.1.1 General tool: Crystal Ball

Purpose:

Forecasting and risk analysis taking uncertainty and variability into account

Characteristics:

Crystal Ball is a simulation program that allows analysis of risks. CB considers uncertainty (amount of data or lack of knowledge) and variability (natural/intrinsic) by means of a two-dimensional simulation for characterising risk. The ‘Two-dimensional Simulation’ tool runs an outer loop to simulate the uncertainty values, and then keeps constant the uncertainty values while it runs an inner loop of the whole model to simulate the variability. The primary output of this process is a chart depicting a series of cumulative frequency distributions.

CB is implemented as an add-in to Microsoft EXCEL and utilises Monte Carlo and Latin Hypercube sampling. The program allows creating several alternative scenarios for each model. Forecast, trend, and overlay charts can be generated into reports. A sensitivity analysis option is available.

Availability:

Commercial software of Decisioneering Inc., Standard edition 882 € in September 2004. Website: http://www.crystalball.com

Probabilistic:

Yes.

5.1.2 General tool: Risk Calc

Purpose:

General tool to support probability bounds analysis, standard fuzzy arithmetic, and classical interval analysis.

Characteristics:

Risk Calc does not require specification of precise details of statistical distributions and their dependency relationships when empirical data are lacking. Risk Calc uses traditional methods such as probability theory and interval analysis and the newest techniques such as probability bounds analysis and fuzzy arithmetic to quantify uncertainty in risk assessments.

Standalone Windows program, includes a book that explains how to use the software and supplies examples from reliability engineering, human health assessment, and environmental and ecological risk analysis

Availability:

Commercial software of Applied Biomathematics, available from Lewis Publishers (149.95 $). Website: http://www.ramas.com/riskcalc.htm

Probabilistic:

Yes.

5.1.3 Fate and exposure: SimpleBox

Purpose:

Developed originally in the 1980‘s as a scientific tool for researchers of environmental fate processes, SimpleBox has formed the regional distribution module in the EUropean System for Evaluation of Substances (EUSES) since 1996.

Characteristics:

SimpleBox is a multimedia (air, water, soil) compartment model of the so-called Mackay type and performs mass balance calculations for the environmental compartments considered (40 in all). The mass balance terms are: emissions into each compartment, advection with air and water, intermedia exchange (atmospheric deposition, volatilisation from water, soil and vegetation, soil run off and erosion, sedimentation and re-suspension), and removal (leaching from soil to groundwater, burial in deep sediments, degradation in all compartments). SimpleBox solves the set of mass balance equations in two different modes: steady-state (Mackay level 3)

612 mode, or quasi-dynamic (Mackay level 4) mode. SimpleBox takes the nesting
613 approach to accounting for spatial variability: it models a small area of high emission
614 intensity (hot spot) as a local scale, nested inside a regional scale, which is nested
615 inside a continental scale, and so on.

616 SimpleBox is written in MS-Excel spreadsheet format, allowing the user full access to
617 system parameters and process formulations. Probabilistic inputs can be handled via
618 Monte-Carlo simulation software (e.g. Crystal Ball).

619 **Availability**:

620 Report: http://www.rivm.nl/bibliotheek/rapporten/719101029.html

621 Free available from ha.den.hollander@rivm.nl or d.van.de.meent@rivm.nl

622 **Probabilistic:** Can be, but needs additional software.

623 ### *5.1.4 Fate: FOCUS Models*

624 **Purpose**:

625 FOCUS is FOrum for the Co-ordination of pesticide fate models and their USe. The
626 organisation is an initiative of the European Commission to harmonise the calculation
627 of predicted environmental concentrations (PEC) of active substances of plant
628 protection products (PPP) in the framework of the EU Directive 91/414/EEC. FOCUS
629 is based on co-operation between scientists of regulatory agencies, academia and
630 industry.

631 **Characteristics**:

632 The FOCUS **groundwater** scenarios are a set of nine standard combinations of
633 weather, soil and cropping data which collectively represent agriculture in the EU for
634 the purposes of a Tier 1 EU-level assessment of leaching potential. The scenarios
635 and their derivation are described in detail in a published report. The scenarios have
636 been implemented as sets of input files for four simulation models:

637 – MACRO is a physically-based one-dimensional numerical model of water flow
638 and reactive solute transport in field soils.

639 – PEARL is an acronym of Pesticide Emission Assessment at Regional and
640 Local scales. It is a one-dimensional numerical model of pesticide behaviour in
641 the soil-plant system.

642 – PELMO is a one dimensional simulation model simulating the vertical
643 movement of pesticides in soil by chromatographic leaching.

644 – PRZM (Pesticide Root Zone Model) is a one-dimensional non-deterministic
645 compartment model for the prediction of chemical movement in unsaturated
646 soils by vertical chromatographic leaching.

647 The FOCUS **surface water** scenarios are used to assess the potential contamination
648 of active substances and metabolites of plant protection products to surface water.
649 They form a part of the review process for active substances in the EU.

650 The FOCUS concentration estimation methodology was developed as a tiered
651 approach with four levels of assessment. The Step 1 has been defined as a relatively
652 simple calculation based on a maximal loading and a fixed scenario, while the Step 2
653 allowed multiple applications and regional variation across Europe.

The program STEPS1-2 in FOCUS is a stand-alone Surface water Tool for Exposure Predictions -Steps 1 & 2 for the derivation of PEC values in water and sediment based upon the chosen scenario. The tool requires a minimum of input values (molecular weight, water solubility, DT50soil, Koc, DT50sediment/water, number of applications, application interval and application rate) and is designed to evaluate both active substances and metabolites.

Step 3 of the approach consists of the scenarios developed, while Step 4 allows a detailed site-specific approach in case all Steps fail. The FOCUS surface water scenarios are a set of ten standard combinations of weather, soil and cropping data and water bodies, which collectively represent agriculture in the EU for the purposes of a Step 3 EU-level assessment of concentration estimation. The scenarios and their derivation are described in detail in a published report (FOCUS 2001, see Documentation in FOCUS Surface water).

The scenarios have been implemented as sets of input files for three simulation models estimating the influence of drainage, run-off and fate on the final concentration estimations:

- The model MACRO is used to estimate the drainage as a sub-surface loading to surface waters,
- the model PRZM accounts for run-off as a superficial loading to surface water and finally
- the model TOXSWA, which takes into account the dissipation processes in surface waters itself.

The results of the MACRO-model or the PRZM-model are used as input into the model TOXSWA in addition to the drift input. The resulting concentrations from TOXSWA are used in the risk assessment process to calculate the Toxicity Exposure Ratios (TER) for aquatic organisms.

To minimise the influence of the user on the outcome of the PEC estimation as many as possible of the input variables have been fixed, leaving only the dossier data as main input data. To take care of this input process and to guide the user through the correct scenarios to run depending on the use of the PPP a computerised shell is developed as well. This shell, called SWASH, helps the user through the exposure assessment.

FOCUS models are not developed for probabilistic risk assessment. However, in PRZM, series of daily weather data over a period of 20 (40, 60) years can be used as inputs to calculate 80th percentile of concentrations.

Availability:

All models and documentation can be downloaded at http://viso.ei.jrc.it/focus/index.html.

Probabilistic:

Some modelling aspects have probabilistic elements.

5.1.5 Fate: MCPELMO

Purpose: Stochastic prediction of the vertical movement of pesticides in soil by chromatographic leaching

Characteristics:

The temporal variability of weather conditions and the spatial variability of soil properties in Germany served as inputs into Monte-Carlo simulations using stochastic version of PELMO 2.01 respectively FOCUS PELMO. For example, MCPELMO allows calculation of the probability to exceed threshold levels of pesticides in groundwater for different scenarios in Germany.

Availability:

Email contact: michael.klein@ime.fraunhofer.de

Probabilistic:

Yes.

5.1.6 Fate and Effects: AQUATOX

Purpose:

Analysis and prediction predict of the effects of chemical pollutants and other stressors on aquatic ecosystems.

Characteristics:

AQUATOX is an ecosystem model that simulates the transfer of biomass and chemicals from one compartment of the ecosystem to another. It does this by simultaneously computing important chemical and biological processes over time. AQUATOX can predict not only the fate of chemicals in aquatic ecosystems, but also their direct and indirect effects on the resident organisms. Therefore it has the potential to help establish the cause and effect relationships between chemical water quality, the physical environment, and aquatic life.

Required input data are loading to the water body, general site characteristics, chemical characteristics of any organic toxicant, and biological characteristics of the plants and animals. AQUATOX comes bundled with data libraries that provide default data. This is of particular importance for the biological data, which are probably the most difficult for a user to obtain.

Latin hypercube facilitates uncertainty analysis with choice of 4 distributions of any or all model parameters and loading.

Availability:

Free download of the program and detailed information at http://www.epa.gov/ost/models/aquatox/

Probabilistic:

Yes.

5.1.7 Effect: Daphnia

Purpose:

Extrapolation from effects on life table data of *D. magna* to the population level considering variability between individuals

Characteristics:

Daphnia describes the population dynamics of the water flea *Daphnia magna* by means of a detailed model of the individual life histories of the animals driven by available food levels and toxicant concentrations. The model focuses on the demographic stochasticity due to the different life table properties of the simulated individuals. The model was validated based on population experiments under flow-trough conditions at different constant concentration of a toxicant (3.4-dichloroaniline).

Daphnia is a stand-alone Windows program. Documentation of the model is only available in German.

Availability:

Free on request from udo.hommen@ime.fraunhofer.de

Probabilistic:

Yes.

5.1.8 Effects: DEMETRA

Purpose:

Prediction of toxicity of pesticides and related compounds (such as metabolites) from chemical structure of the compound.

Characteristics:

DEMETRA is an on-going project (Start 2003-01-01, end 2006-06-30), clustered with EUFRAM. This project aims to develop software that will give a quantitative prediction of the toxicity of a molecule, in particular molecules of pesticides, candidate pesticides, and their derivatives. The input will be the chemical structure of the compound, and the software algorithms will use “Quantitative Structure-Activity Relationships” (QSARs).

Five endpoints have been selected for prediction: Trout, Daphnia, Quail (acute and dietary toxicity), Honeybee. The data set to build up and validate the model includes 420 pesticides (the number is increasing for inclusion of further compounds for testing. Basics of the SW will be a collection of hybrid systems addressing a single endpoint. Five different hybrid systems will be developed for the five endpoints above indicated. Each hybrid system will integrate several QSAR models.

Availability:

Web based tool, will be available at http://www.demetra-tox.net

Probabilistic:

Probabilistic elements present.

5.1.9 Effects: PERPEST

Purpose:

Prediction of the Ecological Risks of PESTicides in freshwater ecosystems

Characteristics:

PERPEST predicts the effects of a particular concentration of a pesticide on various (community) endpoints, based on empirical data extracted from the literature. The method that it uses is called Case-Based Reasoning (CBR), a technique that solves new problems (e.g., what is the effect of pesticide A?) by using past experience (e.g., published microcosm experiments). The database containing the "past experience" has been constructed by performing a review of freshwater model ecosystem studies evaluating the effects of pesticides. This review assessed the effects on various endpoints (e.g. community metabolism, phytoplankton, macro-invertebrates) and classified them according to their magnitude and duration.

The PERPEST model searches for situations in the database which resemble the question case, based on relevant (toxicity) characteristics of the compound. This allows the model to predict effects of pesticides for which no evaluation on a semi-field scale have been published. PERPEST results in a prediction showing the probability of classes of effects (no, slight or clear effects, plus an optional indication of recovery) on the various grouped endpoints.

Availability:

Free download of the Windows program as well as the documentation at http://www.perpest.alterra.nl

Probabilistic:

Yes.

5.1.10 Effects: RAMAS ecotox

Purpose:

Shell to develop (structured) population and food web model including simple fate models

Characteristics:

Models of population dynamics and toxicant kinetics are constructed using a simple Windows interface, and linked to bioassay data. Parameters can be specified as scalars, intervals or distributions, to take account of environmental variability and ignorance. Monte Carlo simulations are then used to predict future population trajectories, and calculate the risk of adverse events such as extinction's or algal blooms.

Availability:

Commercial software 395 $ (academic), 595 $ (non-profit) or 995 $ (regular). Website: http://www.ramas.com/ecotox.htm

Probabilistic:

Yes.

5.1.11 Exposure and effects: Secondary Poisoning

Purpose:

Probabilistic calculation of bioaccumulation risks for terrestrial foodwebs of predators

Characteristics

The model uses a species-specific approach to estimate the bioaccumulation potential in terrestrial food webs. Four birds of prey (Sparrow Hawk, Kestrel, Barn owl and Little owl) and two mammalian predators (Badger and Weasel) were modelled. Bioaccumulation factors (BAFs) for major food items of predators were collected and it was assumed that the available data were a sample of a log-logistic distribution. Critical intake levels are collected from NOECs. These were corrected for differences between laboratory animals and animals in the wild: metabolic rate, caloric content of food, and food assimilation efficiency. The model compares food intake concentrations with critical food concentrations to estimate risk to top predators. The model was applied to DDT and Cadmium.

The model is written in Mathcad.

Availability:

Free on request from tp.traas@rivm.nl

Probabilistic:

Yes.

5.1.12 Exposure and effects: SimpleChain

Purpose:

Calculation of effect of persistent compounds on competition within trophic levels and effects on overall ecosystem function

Characteristics:

Generic trophic levels were modelled: primary production (as input), microbial degradation, detritivore invertebrates and carnivorous invertebrates. The model was applied to Copper and Zinc. Model structure can be easily adapted for the number of competing species within trophic levels (limited by computer memory). The model uses ecological parameters (max uptake rate, half saturation constants, assimilation efficiency, production efficiency, death rate, and for detritus: decay rate) to construct parallel food chains of the most efficient producers. Efficient is defined as a more efficient resource utilisation (higher production efficiency). These are combined randomly with species sensitivity, sampled from species sensitivity distributions. The SSD parameters (mean, stdev) can be calculated from toxicity databases.

The software is implemented in Matlab.

Availability:

Not yet decided, in development. Contact: tp.traas@rivm.nl

Probabilistic:

Yes.

5.1.13 Effects: SSWD (Weighted SSDs)

Purpose:

Calculation of Species Sensitivity Weighted Distributions (SSWD) and hazardous concentration (HCx)

Characteristics:

SSWD offers a procedure in which ecotoxicity concentration data can be weighted to account for redundant data for each species or genus and for the disproportion in the data number between the taxonomic groups or trophic levels was developed.

Three distributions (log-empirical, log-normal and log-triangular) can be selected at the same time. The preceding weights are accounting for the calculation of the distribution's parameters and the hazardous concentration. Confidence limits are estimated by bootstrap and the number of bootstrap samples can be defined by the user. The goodness of fit (for log-normal and log-triangular distribution) is tested by a Kolmogorov-Smirnov (KS) test with Dallal-Wilkinson approximation. The SSWD procedure is a Macro ".xla" of Excel.

Availability:

SSWD.xla can be free loaded from the http://chimie.ineris.fr . Email contact: philippe.ciffroy@edf.fr

Probabilistic:

Yes.

5.1.14 Effects: Zebrafish

Purpose:

Extrapolation from effects on life table data of *Danio rerio* to the population level considering variability between individuals

Characteristics:

Zebrafish is an Individual Based Model, which simulates a small laboratory population of the zebrafish *Danio rerio*. The model includes stochastic descriptions of reproduction and foraging behaviour and allows analysis of the consequences of effects on the individual level to the level of the population.

Zebrafish is a stand-alone Windows program. Documentation of the model is only available in German.

Availability:

Free on request from udo.hommen@ime.fraunhofer.de

Probabilistic:

Yes.

5.1.15 Risk calculation: Busy (Bayesian Uncertainty System)

Purpose:

Uncertainty analysis in Ecological Risk Assessment using Bayesian statistics

Characteristics:

Busy provides a flexible and easy to use set of probability distributions to carry out Probabilistic ERA. This includes kernel density estimation to account for non-symmetric and/or multi-modal data sets. The SSD part of Busy focuses on the (species) toxicity data only, while extending the analysis to secondary (2-D) distributions. This is done through Bayesian statistics. Bayesian Logistic Dose-Response modelling implements a small-sample approach to uncertainty analysis of dose-response data sets.

Busy is implemented in the computer programming system Mathematica. In order to run Busy, Mathematica (4.2) must have been installed on your computer. However, there is no need to learn Mathematica. Busy is run through so-called Notebooks that can be coupled to MS-Excel for exchange of data.

Availability:

Developed in the framework of the CEFIC/ACC Long Range Research Initiative Project: Uncertainty Analysis in Ecological Risk Assessment. It will become freely available to CEFIC members, as well as others interested. No web site, email contact: tom.aldenberg@rivm.nl.

Probabilistic:

Yes.

5.1.16 Risk calculation: ETX (various versions)

Purpose:

Derivation of risk limits and/ or estimate % species affected (PAF) by means of SSD-theory.

Characteristics:

The software fits standard distribution functions (normal SSDs) to a set of data. From the SSD, either percentile values on the concentration axis (e.g. the 5th percentile, the HC5), or the fraction of species exposed above the relevant toxicity endpoint can be estimated. This fraction is called the 'potentially affected fraction', PAF. Each quantity estimated with the software is given with a confidence interval, based on normal statistics (Aldenberg & Jaworska, 2000). Different test for the goodness of the SSD fit to a log-normal distribution are implemented, results are given as tables and diagrams. If a set of PECs is entered, Joint Probability Curves are calculated.

Availability:

ETX-2.0 is available as executable and associated RIVM-report, to replace older versions (Van Vlaardingen, 2004). Email contact: tp.traas@rivm.nl or tom.aldenberg@rivm.nl, and/or (autumn 2004) http://www.rivm.nl/bibliotheek/rapporten/*.*.

Probabilistic:

Yes.

5.1.17 Effects: GPS1 (General Population Simulator)

Purpose:

Estimation of effects and recovery using a simple population model

Characteristics:

The General Population Simulator estimates effects of toxicants on populations and the possible recovery by means of the logistic population growth model. Exposure is modelled as exponential decay after single or multiple application of a toxicant, while population dynamics is simply described by growth rate and carrying capacity. Toxicity data are given as LC50 or inhibition of (population) growth rate. Uncertainty about the input parameters can be handled via Monte-Carlo-Simulation while

temporal variability can be considered by stochastic modelling of the carrying capacity.

GPS is a stand-alone Windows program.

Availability:

Free on request from udo.hommen@ime.fraunhofer.de

Probabilistic:

Yes.

5.1.18 Risk calculation: PRAT2

Purpose: Calculation of SSDs, % of species affected (PAF), and Exceedence Profiles

Characteristics:

PRAT2 is a simple Excel sheet to fit log-normal distributions to toxicity data and/or environmental concentrations and to calculate the % of species affected for a given PEC or Exceedence Profiles (Joint Probability Curves) showing the probability that the toxicity threshold of a certain proportion of species is exceeded.

PRAT2 uses only the implemented functions of Excel without any macro programming and it is restricted to log-normal models.

Availability:

Free as email attachment from ksolomon@uoguelph.ca

Probabilistic:

Yes.

5.1.19 Exposure, Effects, Risks: Aquatic Level II Refined Risk Assessment (ARRA 2.0)

Purpose:

Estimation of the likelihood and magnitude of effects on aquatic species that are vulnerable to pesticide exposure in edge-of-field situations

Characteristics:

US-EPA is developing a software tool (**A**quativ Level II **R**efined **R**isk **A**ssessment) that facilitates execution of the exposure simulation. It uses an exposure module based on the Agency's PRZM model (used here as a run-off model) linked with the VVWM (Varying Volume Water body Model) .A toxicity data analysis module calculates measurement endpoints and species sensitivity distributions for fish and invertebrates. The effects and exposure data are integrated using a two-dimensional Monte Carlo analysis in the probabilistic risk module, yielding estimates of the probability and magnitude of effects to aquatic organisms, as well as estimates of uncertainty associated with those predictions.

Availability:

Latest information can be found at http://www.epa.gov/oppefed1/ecorisk

Probabilistic:

Yes

5.1.20 Exposure, Effects, Risks: Terrestrial Integration Model (TIM 2.0)

Purpose:

Estimation of the risk of acute effects on birds due to exposure to pesticides in-crop

Characteristics:

The revised Level II **T**errestrial **I**ntegration **M**odel is a multimedia exposure/effects model that can be used to address acute mortality levels in generic or specific species over a user-defined exposure window. The spatial scale is at the field level, such that the field and surrounding area are assumed to meet habitat requirements for each species. As an overall simplifying assumption, contamination of edge or adjacent habitat from drift is assumed to be zero.

The major parameters addressed in the model are multimedia estimates for different routes of exposure; food habits of defined generic or selected specific species, hourly ingestion/inhalation rates of food, water, and air and dermal residue transfer rates from contaminated vegetation as a function of body weight, frequency of feeding and drinking on the sprayed field, distribution of residues on/in vegetation, water (dew and puddles), and air, degradation/dissipation rates of pesticide residues in each environmental media considered and acute toxicity dose-response relationships.

Availability:

Free download from http://www.epa.gov/scipoly/sap/index.htm#march.

Probabilistic:

No.

5.1.21 Exposure, Effect, Risks: USES 4.0

Purpose:

Initial and refined risk assessment of industrial chemicals and pesticides.

Characteristics:

USES is a decision-supporting instrument enabling rapid and efficient assessment of the risks attached to new and existing substances, and agricultural and non-agricultural pesticides. The PC program, manual and background document, are mainly intended for use in the Netherlands. USES is intended mainly for initial and refined risk assessment rather than comprehensive analysis. Interpreting estimated risks for risk management is also beyond its scope. Risk assessment with USES takes account of international policy on substances and the associated regulations. The USES 4.0 risk assessment system for new and existing substances is fully equivalent to EUSES. USES 4.0 also incorporates, as much as possible, the EU Uniform Principles referred to in Directive 91/414/EEC concerning the placing of plant protection products on the market and Directive 98/8/EC concerning the placing of biocidal products on the market.

For probabilistic assessments, additional tools like Crystal Ball are needed.

Availability:

See http://arch.rivm.nl/csr/risk.html

Probabilistic:

No.

5.1.22 Exposure, Effects, Risks: WEBFRAM

Purpose:

Development of a web-integrated model (WEBFRAM), for the assessment of ecological risks from pesticides.

Characteristics:

Ongoing project, sponsored by DEFRA (PS2304). Report available in December 2006

Availability:

No own web site yet, information from:

http://www.silsoe.cranfield.ac.uk/ecochemistry/research/project/ps2304.htm

Probabilistic:

Not clear yet

5.1.23 Exposure and Effects: Instrument for Quantification of toxic pressure (IQ-tox)

Purpose:

Multi-criteria evaluation of multi-substance exposure data, with corrections for bioavailability of compounds (hence: an exposure element), using functions of matrix/media characteristics. No fate modelling.

Characteristics:

IQ-Tox is a dedicated MS Access application with VBA procedures and custom made graphics procedures for evaluating site-specific exposure or monitoring results. Very flexible and versatile program for quantitative multi-criteria analysis. Unlimited sets of environmental and response variables can be defined and grouped to form different environmental quality indices related to predetermined human activities. Ecotoxicity related indices can be calculated according to the mixed model ms-PAF (multi-substance Potentially Affected Fraction) approach for mixture toxicity. Any variable that can be quantified and rated for environmental acceptability can be incorporated. Several options and weighting procedures are open for grouping.

IQ-Tox has no probabilistic components yet.

Availability:

An Access *.mbe file can be obtained from d.de.zwart@rivm.nl

Probabilistic:

No.

5.2 Confronting Software items with criteria

The following results when the software items are confronted with the criteria developed during the workshop Table 5-2. The Table shows, regarding the scientific aspects of the software for PRA:

- that various software items were not confronted to the full list of criteria developed during the Software and Database workshop. Therefore, contact persons were asked after the workshop to check the criteria for "their" software. This is an ongoing process (see Recommendations chapter). Other items were not available at the workshop (these items do not have evaluation entries).[4]
- that software items that address the *same* type of issues (e.g., calculating SSDs) differ with regards to the apparent suitability for further development in the EUFRAM framework.
- that software items can undergo substantial evolution. Experience over the last year shows that major changes in development status may occur within a year. An example is the evaluation of EXT(2000), a dedicated Excel program, at the workshop in 2003, while a versatile stand-alone executable is available in autumn 2004.
- that the evolution of software appears to take place in line with the criteria as established for part of the studied items

and that, regarding the operational features:

- some items obtain positive remarks on many aspects, whereas others obtain mainly negative marks. Note that this suggests large differences in the purpose of the designer (own use versus planned multi-usage by third parties), and not scientific weakness of the latter.

Given these observations, and the large variability in judgement on the individual criteria for each software item, the workshop was crucial to gain overview. To that end, a SWOT analyses were made. The results are presented in Table 5-3. Note that the SWOT provides a preliminary overview on the proposed and evaluated items, and that the Recommendations (see Chapters 8 and 9) suggest to continuously update the views that are presented. Note further that the Tables are the results obtained in the June 2003 workshop. The ideas shown are to be seen as a wish list.It is likely that various wishes are not fulfilled for the existing software, or that they are difficult to be fulfilled in the near future.

[4] Owners of software and databases are asked to provide further facts as items develop..

Table 5-2. Overview of the confrontation of the software items with the criteria for software development.

	SimpleBox	AQUATOX	Daphnia	DEMETRA	PERPEST	SSWD	Zebrafish	GPS1	IQ-Tox
Scientific characteristics									
Ability to handle different groups of compounds	+	+	+	+	+	+	+	+	+
Multiple Spatial scale (global/regional/local)	+	-	-	-	-		-	-	-
Ability to handle (pesticide) mixtures	+	+	-	-	-	-	-	-	+
Ability to handle metabolites	-	+	-	+	-		-	-	+
One stop shop (including exposure, effects and risk characterisation)		-	-	-	-		-	-	+
Option to account for uncertainty	+	+	-	+	+	+		+	-
Option to account for variability	+	+	+	+	+	+	+	+	-
Option to account for dependencies/correlation	+		-		-		-	-	-
Empirically validated	+	+	+		+		+	-	-
Flexible model structure	+	+	-	+	-		-	-	-
Wide choice of distributions	+	+	-	-	-	-	-	-	+
Published in peer-reviewed literature	+		-	+	+	+	+		-
Expert panel peer review	+	+	-	+	+		-	-	-
Temporal variability/uncertainty	+	+	-	-	+		-	+	-
Spatial variability/uncertainty	+	-	-	-	+		-	-	-
Requires more than typically available data	-	-	-		-		-	-	-
Requires recalibration for new regions/countries	-	-	-		-		-	-	-
Contains non-data-derived adjustment factors	-		-		-		-	-	-
Technical									
Web-enabled use possible	-		-	+	+	-	-	-	-
Quality assured software		+	-	+	+		-	-	-
Version control protocol available		+	-		+		-	-	-
Model FULLY documented	+	+	-	+	+	+	+	-	-
Input data error checking (incl. units)	-	-	-		+		-	-	+
Backwards/forward compatible		-			+		-	-	-
PC Based	+	+	+	+	+		+	+	+
Broadly dispersed and known shell	+	+	+	+	-	+	+	+	+
Potential to link to other software/databases	-	+	-	+	-		-	-	-
Dependence on other proprietary software	-	-	-	+	+	+	-	-	+
Ease of use									
Flexibility in data input (e.g. units)	+	-	+		+	+	+	+	-
Free of costs	+	+	+	+	+	+	+	+	+
Exporting of output	+	+	+		+	+	+	+	+
Reproduce/save past run	+	+	+		+	+	+	+	+
Help screens	-	+	+		+		+	-	-
Intuitive interface	+	+	+	+	+	+	+	+	+
Free of use restrictions	+	+	+	-	+	+	+	+	+
Any windows-based	+	+	+		+	+	+	+	-
Ability to cut/paste/link input data	+	+	+		+		+	+	-
Roll-over (mouse-over) help	-	+	-		-		-	-	+
Reference fields for input data	-	+	-		-		-	-	-
Default data and worked examples	-	+	+	+	-		+	+	-
Differential User-modes (default and advanced user modes)	-	+	-		+		-	-	-
Execution in minutes Vs. hours on Std. PC	+	+	+		+	+	+	+	+
Shape of output format: numerical		+	+	+	+	+	+	+	+
Shape of output format: graphical		+	+		+	+	+	+	+
Option to show confidence bands	-	+	+		-	+	+	+	-
Support and maintenance									
Related to regulatory context	+	+	-	+	+	+	-	+	-
Feedback loops to improve software		+	-		+		-	-	+
Long-term support/updating/maintanence	-	+	+		+		-	-	+
User support	-	+	-		-		+	-	+
User groups	-	-	-		-		-	-	-

Table 5-3. Results of the SWOT analyses for the software items (restricted to items presented at the workshop). Suitability 1 = directly usable in EUFRAM, 2 = with modification, 3 = not suitable, 0 = not ranked during the workshop).

Acronym	Probilistic module included	Research tool	Potential for regulatory use	SWOT summary / Remarks	Suitability for EUFRAM
Crystal Ball	X	X	X	+ user friendly and flexible tool - expensive Excel add-in, limited sample size, lack of Bayesian methods, limited sensitivity analysis	0
Risk Calc	X	X	X	+ flexibility: statistical freedom in the treatment of distributions, automatic check of units - people not (yet) familiar with p-bounds, complex tool, not free available	0
FOCUS PELMO	(X)		X	+ developed especially for EU registration, quality checked, user friendly - fixed scenarios, not really probabilistic yet	2
SimpleBox	(X)		X	Included in EUSES + simple, generic model, well studied, free available, flexible excel sheet - global or regional scale only, steady state, not dynamic, formulars not protected, requires expert user, uncertainty not included (but possible with Crystal Ball)	2
AQUATOX	X	X	X	Higher tier tool + state of the art dynamic ecosystem model(s), flexible, very user friendly, calibration and validation continues, used by US EPA, free on web - model equations fixed, complex	2
DEMETRA	(X)			Clustered with EUFRAM, under development, will allow to predict toxicity values with uncertainty reported	0
PERPEST	X	X	X	Database based, only regulary update necessary + new approach to use knowledge from exisiting micro-/mesocosm studies to estimate ecological effects of new compounds, user friendly interface, peer reviewed, free on the web - limited flexibility (input of new data), restricted to substance with MoA covered in the database	1
RAMAS ecotox	X	X		Shell, higher tier tool + focus on ecological endpoints, user friendly model builder interface, flexible - not free available (commercial), parameters dificult to assess	2
Secondary Poisoning	X	X	X	Potentially useful + simple trackable algorithm, clear structure, free on request, - depends on Crystal Ball	2
SimpleChain	X	X	X	+ simple mechanistic model - depends on Matlab, not user friendly	2
SSWD	X			not available at the workshop	0
Busy	X	X	X	Under develeopment + powerful statistical tool - complex, not clearly user friendly, needs Mathematica	0
ETX	X		X	Simple Excel based tool + complete SSD and JPC tool, user friendly, free - diagrams might be improved	1
GPS1	X		X	+ simple population model to analyse effects and recover , user friendly interface - restricted to single population, software not quality checked and documented	2
PRAT2	X	X	X	Teaching tool + simple SSD and PEC distr. tool, easy to use, free - no detailed statistics (e.g. confidence limits)	2
ARRA	X		X	+ all-in-one: includes different tools (e.g. fate/ exposure models, dose-response models, SSDs,…); especially developed for regulation of pesticides , user friendly interface, straightforward package for basic PRA - specific to US-EPA, use of not S.I. u	2
TIM	X		X	+ integrates fate and effects, user friendly interface (Excel based), dose-response orientated, free on the web - restricted to birds & mammals, no insects, no beneficial arthropods, no chronic data, US-EPA scenarios only, little flexibility, adaptation	2
USES	(X)		X	+ focusses on many aspects, linked to regulation, accepted in EU (EUSES), tested over the years, can be adapted for pesticides - not yet probabilistic (needs Crystal Ball)	2
WEBFRAM	(X)		X	Under development + comprehensive models, linked to current legislation, strong links to EUFRAM, free, web based - not necessarily probabilisitic, UK rather than European focus (only UK scenarios)	0
IQ-Tox			X	+ user friendly interface, AMOEBA presentation, creates simpe to interprete output, für site specific RA - not realy probabilistic	3

From the latter Table, it can be concluded that most software items obtain the rating "2", implying that the item has to be adapted to the general design criteria to a smaller or larger extent. Only two programs were considered sufficiently adapted to the EUFRAM criteria that they gained the rating "1", which implies that those items only need undergo minor adaptation when they would be adopted in the EUFRAM framework. Various programs obtained the remark that they were especially suited to a lower or higher tier of an eventual risk assessment scheme. A small number of items has already been adopted in an official, sometimes EU-level, risk assessment framework (e.g., EUSES).

6 INVENTORY OF DATABASES

6.1 Characteristics / types of Database items

At present (September 2004), 12 database items are considered within the EUFRAM project (Table 6-1). The databases address different parts of a risk assessment framework, focusing on:

- physicochemical properties (e.g., in RIVM e-toxBase)
- ecotoxicological sensitivities (e.g., in US-EPA Ecotox and RIVM e-toxBase)
- ecological properties (e.g., Assiwerk/Enerwerk)
- field conditions (e.g., POND-FX)

Like in the software item classification, this differentiation must not be seen as exclusive. Some databases store data of different kinds, for example, the RIVM e-toxBase can store test data as variable as physico-chemical properties of a compound, ecotoxicity test data (NOECs, EC50s, BSAFs), Mode-of-Action data, metabolite data, and taxonomic information.

Table 6-1. Databases reviewed for the use in probabilistic risk assessment within EUFRAM (status September 2004)

	Compound			Biota		Field		
Acronym	Physico-chem. properties	MoA	Usage	(Eco-) toxicity	Ecology	Site charac-teristics	Fied exposure	Field effects
EPA EFD	X						X	
RED				X			X	
RIVM e-toxBase	X	(X)		X				
ECOTOX-EPA				X				X
COMET				X				
NTP	X							
ECOTOX CD				X	X			
MINAMB	X			X			X	
SEEM			X	X				
Assiwerk / Enerwerk					X			
Pond-FX					X	X		
SEISMIC						X		

6.2 Descriptions of Databases

The following short descriptions were extracted from the summary and fact sheets of a meta-database developed for EUFRAM-WPR10 in which all model and programming information is stored for detailed reference. For more information this meta-database with more specific information can be downloaded from the EUFRAM web site (www.eufram.com/xxx).

6.2.1 Fate: EPA EFD (Pesticide Environmental Fate Database)

Purpose:

This is an environmental fate database containing studies describing what happens to pesticides in the environment after they are applied. The database contains fate and transport data.

Characteristics:

EPA's Office of Pesticide Programs (OPP) collects and reviews a variety of environmental fate studies submitted by pesticide manufacturers in support of the registration of pesticide products. Environmental fate studies describe what happens to a pesticide in soil, water, and air after it has been applied and include the following types of studies: product chemistry, metabolism, hydrolysis, photolysis, field dissipation, bioaccumulation, adsorption/desorption and leaching. After reviewing the data in these studies, OPP scientists summarise the information in Data Evaluation Reports (DERs), Reregistration Eligibility Decision Documents (REDs), science chapters, Emergency Use Exemptions, and other environmental fate reports. In 2000, OPP initiated the development of a pesticide environmental fate database which will allow the user to search and view the data, query the fate database, and print reports that are found in these summary reports. The database contains environmental fate and transport data for about 250 pesticide active ingredients. The Pesticide Program plans to complete the initial version of this database by the end of 2002 and will be adding additional active ingredients during the next two years.

Availability:

The Pesticides Environmental Fate Database will be put in service by the US EPA in late spring or early summer 2003, as Microsoft Access database. For more information: Liu.Larry@epa.gov. Website (only descriptive as yet: http://www.epa.gov/oppefed1/general/databasesdescription.htm

6.2.2 Ecotoxicity and metabolites: RED

Purpose:

Compilation of a European Reference Database for Ecotoxicology Data on Plant Protection Products and their Metabolites

Characteristics:

Council Directive 91/414/EEC requires extensive ecotoxicological testing of active substances and plant protection products. For an overall ecotoxicological assessment it is insufficient to evaluate a limited number of endpoints in an isolated way. Comparative assessments of different substances and/or probabilistic assessments of the variability of data sets or analyses of sensitivity distributions of different species become increasingly important not only in the scientific community but also in the context of regulatory decision making. In order to perform such comparative or probabilistic assessments a body of validated data is required. The goal of the RED project therefore is to initiate the development of a database containing this type of data which is freely accessible and which will allow the statistical distributions of the ecotoxicological effects of active parent compounds and their metabolites to be explored.

Availability:

Under development. See further:

http://www.silsoe.cranfield.ac.uk/ecochemistry/research/project/red.htm

6.2.3 Ecotoxicity, metabolites and physico-chemical data: RIVM e-toxBase

Purpose:

Development of the RIVM e-toxBase started in 2000, in order to provide an array of RIVM ecotoxicity-related projects with a tool to store, handle and retrieve data in a reproducible way. The database now covers more than 166,000 entries, and is ready for internal use.

Characteristics:

The RIVM e-toxBase is a server-based relational database with strict integrity rules, ensuring a reliable and unambiguous means of storing and retrieving ecotoxicity data. Access to the database is user-restricted. Depending on their specific rights, users can read, qualify, store and export data in the RIVM e-toxBase.

Both the database itself and the interface were developed in MS Access. By now, the database has been transferred to a robust, server-based environment (a stand-alone application in Visual Basic). The interface has three main functions: selecting and viewing data, easy export of a selection of data to a MS Excel sheet and entering new data. A range of selection criteria like test type (EC50, NOEC), substance, species name and literature references are available for querying the data in the RIVM e-toxBase. After a selection is made, a user is immediately presented with the remaining number of tests. Advanced search options are available, like the taxonomy browser, which allows a user to select test results for entire taxonomic groups. Using the flexible export function to Excel, any field from the database can be selected for export. The entry of new tests is restricted to certain users within RIVM.

Currently, data coverage concerns mostly aquatic ecotoxicity data, pesticides included. Compound characteristics (e.g., Kow), and breakdown products (parent/metabolite) can however also be stored. The database can store scientific entries as well as meta-data (which include literature references but also an array of quality statements, which may differ between the scopes of the different RIVM-projects). The largest source of data is the US EPA ECOTOX database. Data from the ECOTOX database have first been ‘cleaned’ in order to fit the strict integrity rules of the RIVM e-toxBase. New data are added each day, as RIVM projects generate or store ecotoxicity data. Some of the RIVM data sets are (at least in part) confidential. In the latter case, access to the data is restricted to certain users.

The data model design of the RIVM e-toxBase is highly flexible, so that new types of data can be stored without having to make major structural modifications to the database. The user-friendly interface (and in future web-interface) offers access to the database to a wide audience, without requiring any special knowledge of ecotoxicology or databases.

Availability:

To be decided, currently only RIVM-internal use. Development of a web-enabled searching technique is underway.

6.2.4 Ecotoxicity: Ecotox-EPA

Purpose:

The ECOTOXicology database (ECOTOX) is a source for locating single chemical toxicity data for aquatic life, terrestrial plants and wildlife. ECOTOX is a useful tool for examining impacts of chemicals on the environment.

Characteristics:

Peer-reviewed literature is the primary source of information encoded in the database. Pertinent information on the species, chemical, test methods, and results presented by the author(s) are abstracted and entered into the database. Another source of test results is independently compiled data files provided by various United States and International government agencies. ECOTOX was created and is maintained by the U.S.EPA, Office of Research and Development (ORD), and the National Health and Environmental Effects Research Laboratory's (NHEERL's) Mid-Continent Ecology Division. The development of the ECOTOX was started in 1995, and in March of 1996, it was released to governmental users through talnet access procedures. In February, 2000 ECOTOX was released as a web based interface system. The ECOTOXicology database (ECOTOX) is a source for locating single chemical toxicity data from three U.S. Environmental Protection Agency (U.S. EPA) ecological effects databases; AQUIRE, TERRETOX, and PHYTOTOX. Aquatic data in AQUIRE are limited to test organisms that are exclusively aquatic (saltwater and freshwater). Species that are associated with the water but do not have gills, such as ducks and geese, are included in the terrestrial database. Amphibians are included in both AQUIRE and TERRETOX databases, with the life stages that exist exclusively in the water (e.g., tadpole) located in AQUIRE and the terrestrial life-stage (e.g., adult) in TERRETOX. Bacteria and virus are not included in the ECOTOX database. TERRETOX is the terrestrial animal database. It's primary focus is wildlife species but when data gaps exist for a particular chemical, data for domestic species are included. PHYTOTOX is a terrestrial plant database.

Availability:

The ECOTOX database can be accessed using your web browser software via the Internet at http://www.epa/gov/ecotox. For more detailed information regarding field data definitions, refer to the pertinent coding guidelines.

6.2.5 Ecotoxicity: COmputerized Molecular Evaluation of Toxicity (COMET)

Purpose:

The data set has been prepared within the EC funded project COMET (Computerized Molecular Evaluation Of Toxicity), ENV4-CT97-0508. It includes 235 pesticides. On the basis of these pesticides predictive models for toxicity have been evaluated, for trout, daphnia, quail, and rat.

Characteristics:

COMET includes a data set with toxicity values for trout, daphnia, quail, and rat. Values have been taken from the Pesticide Manual, RTECS, ECDIN, AQUIRE, HSBD. Min, max and suggested values are indicated. Suggested value is the minimum value, unless it is an outlier. COMET also developed QSAR models, calculating chemical descriptors with various software types, such as CODESSA and Pallas. Data coverage pertains to several chemical classes of pesticides; trout, daphnia, quail, and rat. Mathematical algorithms are used for QSAR models: linear multiregression, neural networks, fuzzy logic, B-spline and genetic algorithms.

Availability:

Excel for toxicity data and chemical descriptors. Available at request: Benfenati@marionegri.it.

6.2.6 (Eco)toxicity: NTP (National Toxicology Program database)

Purpose:

Federal and State Regulatory Agencies use the NTP study data in considering the need for regulation of specific chemicals to protect human health. The National Toxicology Program (NTP), amongst other tasks, provides information about potentially toxic chemicals to health regulatory and research agencies, the scientific and medical communities, and the public.

Characteristics:

The database presents written summary information on test reports. Currently, data coverage concerns mostly: compounds identification (CAS number, chemical formula, ect..), physical-chemical properties (solubility, volatility, stability, etc..), toxicity data (LD50 on several species mainly mammals, carcinogenicity, mutation data, teratogenicity, etc..) and other data. Some information is reported in the text form; some data are not numerical ones.

Availability:

The NTP database can be accessed using your web browser software via the Internet at http://ntp-server.niehs.nih.gov/

6.2.7 Ecotoxicity: ECOTOX CD (CD Ecological Modelling and Ecotoxicology)

Purpose:

The ECOTOX-CD provides extensive tables, data and parameters needed by modellers, theoretical scientists, environmental managers, ecologists and toxicologists to carry out estimations and calculations. Information on the environmental effects of chemical substances is also included.

Characteristics:

The original print edition - the Handbook of Ecological Parameters and Ecotoxicology - contains data and references from the scientific literature in a web-based search environment. Examples of data include: growth parameters, lethal concentrations (LC50), modelling equations, emissions, degradation of chemical substances, background concentrations, concentration factors, biological effects, octanol/water partition coefficients, excretion and uptake rates and compositions of living organisms. Examples of the models include: eutrophication models, models for dispersion of chemical compounds, models for growth and competition of different organisms, and models describing the global environmental cycles.

Availability:

Commercial, www.elsevier-ecotox.com

6.2.8 MINAMB

Purpose:

The MINAMB database stores physico-chemical characteristics, ecotoxicological and mammal toxicological data as well as environmental fate data and information about metabolites for around 450 substances.

Characteristics:

The data are organized in an Access database.

Availability:

Not yet decided, contact: domenica.auteri@icps.it

6.2.9 Effects: Database for pesticides and their metabolites (SEEM)

Purpose:

The SEEM database was commissioned by de European Commission, DG Health and Consumer Protection and contains ecotoxicity data for pesticides and pesticide metabolites. It is currently ready for use by ICPS and DG-SANCO. The main objectives for its establishment were the relationship between acute and chronic toxicity, the relationships between sensitivities of individual endpoints of parent substances and their metabolites, and the relationship between toxicity for Daphnids and sediment dwellers.

Characteristics:

The SEEM-database consists of data on 200 active ingredients and 130 metabolites in and Excel spreadsheet.

Availability:

Restricted (EU Commission), see also http://www.icps.it/

6.2.10 Ecology: Assiwerk/Enerwerk

Purpose:

Database of assimilation efficiency, water content and energy content of food for birds and mammals. The RIVM database is extended by information from the Central Science Laboratory (York, UK) and from Franz Bailain (Germany).

Characteristics:

The database contains data in Excel format and can be used to express dietary information into exposure data for birds and mammals (e.g., a diet of 100% insects for a 15 g passerine bird is *x* gram food per day).

Availability:

Contact person is C.E.Smit@RIVM.NL. Website: http://www.rivm.nl/csr, new site under construction.

6.2.11 Ecology (Field): Pond-FX

Purpose:

In order to better predict the likely effects of a pollutant (pesticide, or otherwise) on the "health" of natural resources such as farmland ponds, the original research project focused on the need to understand how the ecology of the pond life can alter, or mediate, the effects that might be expected purely from the results of toxicity tests. A prerequisite for such investigations is to collect information about the various life history characteristics thought to play a role in mediating the risk that a population of organisms will be both: (a) exposed to the pollutant, (b) unable to recover from any adverse effects within an acceptable period of time.

Characteristics:

The Pond-FX web site provides ways to access the data. Most of the information is contained in a relational database which, although somewhat complex, means that there is great flexibility in how the information may be explored. The links offer selected starting points for exploring the database but do not restrict the information that can be accessed. You will find that most of the information on a page will be linked to the other areas of the database. The data coverage is: 1. A classification of the types of pond found on farmland around Britain and the invertebrate pond life found in them. 2. Information about the life history of pond organisms (including dispersal ability, preferred water body types, feeding habits, etc.). 3. Miscellaneous information such as rarity status, common names and photographs of some of the pond animals.

Availability:

http://www.ent.orst.edu/PondFX

6.2.12 Spatial Environmental Information System for Modelling the Impact of Chemicals (SEISMIC)

Purpose:

SEISMIC is an environmental information system designed to provide detailed soil and weather data necessary to parameterise a wide range of chemical fate models for any environmental scenario in England and Wales.

Characteristics:

SEISMIC is programmed in MS Visual Basic and utilises MS-Access databases and ESRI MapObjects technology. It does not require additional GIS software to run. SEISMIC contains both spatial and parameter data, on climate, agricultural crops, soil types and aquifer types. Further, it contains weather data.

Availability:

Under licence, Annual lease fee or one-off payment, http://www.silsoe.cranfield.ac.uk

6.3 Confronting Database items with EUFRAM Criteria

The following Table results when the Database items are confronted with the criteria developed during the workshop (Table 6-2). The Table shows, regarding the scientific aspects of the Databases for PRA:

- that the database items were not all tested / presented at the Software and Database workshop (some items do not have evaluation entries) and that the criteria list developed during the workshop could not be applied to the items there (input from developers / contact persons is necessary)
- that there are various databases available for a single purpose, e.g. collecting ecotoxicity data in US-EPA Ecotox and RIVM e-toxBase
- that databases have highly different purposes, such as an embedded target of predicting toxicity based on stored toxicity data (e.g. QSAR, embedded formulae) *versus* pure storage and retrieval functions

and that, regarding the operational features:

- some items obtain positive remarks on many aspects, whereas others obtain mainly negative marks. This suggests large differences in the purpose of the

designer (own use versus planned multi-usage by third parties), and not scientific weakness of the latter.

Given these observations and the large variability in judgement on the individual criteria for each software item, the workshop was crucial to gain overview. To that end, a SWOT analyses were made. The results are presented in Table 6-3. Note again that the SWOT provides a preliminary overview on the proposed and evaluated items, and that the Recommendations (see Chapters 8 and 9) suggest to continuously update the views that are presented.

Table 6-2. Overview of the confrontation of the database items with the criteria for database development.

	EPA EFD	RED	RIVM e-toxBase	ECOTOX-EPA	COMET	NTP	ECOTOX CD	MINAMB	SEEM	Assiwerk / Enerwerk	Pond-FX	SEISMIC
Data characterististics												
Modifying factors (contextual data, covariates)			+									
References for data documented			+									
Quality critera documented			+		-							
Identification and handling of 'duplicates'			+									
Info on limit of detection for non detects			-									
Quality measure (e.g. measured / nominal data)			+									
Statistical method mentioned			-									
Indication of uncertainty on estimated values (EC50 with CI)			-		-							
Quality of the study (GLP, compliance with regulation x)			+		-							
Reported with original number of digits			+									
Slope for EC50, LOEC for NOEC			-									
Info to judge tests (e.g. sample size, power)			+									
Raw data included			-									
Link to related compounds			-									
Comprehensiveness of data per endpoint					+							
Comprehensiveness over endpoints												
Technical characteristics												
Format remains in export			+									
Error checking routines for input			+									
Database documentation			+									
Extensible (by db manager)			+									
Allowing for versatile data query / sub-selections			+									
Keep track if records have changed			+									
No aggregrate (text) fields (e.g. min & max instead of range)			+		+							
Web-enabled use possible			-		+							
Windows-based			+		+							
Open design for software access			-									
Ease of use												
Exporting of output selections			+									
User friendliness of search terms (common names)			+									
Reproduce past query			+									
Flexibility in data input (e.g. units)			+									
Free of costs					+							
Includes summary tools to qualify subselections												
Feedback loops to improve context of database												
Related to regulatory context												

Table 6-3. Results of the SWOT analyses of the databases presented at the Software and Database workshop (suitability 1 = directly usable in EUFRAM, 2 = with modification, 3 = not suitable, 0 = not ranked during the workshop)

Acronym	SWOT summary / reamarks	Suitability for EUFRAM
EPA EFD	+ many data on fate, metabolites included, goodqueries, continued entry - restricted to chemicals registered by US-EPA, not yet freely available	1
RED	not available at workshop	0
RIVM e-toxBase	+ integrates different data sets, large number of endpoints, influenced by the EU registration procedure, user friendly interface, flexible,fate and physicochemical data can be added - not yet available for public, dependece on US-EPA for update	1
ECOTOX-EPA	+ large data base of ecotox data, flexible queries, web based, free - quality check not done properly, different units, no slopes given for EC-values	1
COMET	QSAR models and input data + toxicity values given with min and max, metabolites included, free available - includes all tests, regardless of use of standardised methods	2
NTP	not available at workshop	0
ECOTOX CD	not checked at workshop	0
MINAMB	not available at workshop	0
SEEM	+ brings many different high quality data together, includes metabolites, easy to use - not yet availble for public, only species used in EU dossiers considered, no real database with user interface (Excel file)	2
Assiwerk/Enerwerk	+ large DB for bioaccumulation related parameters, raw data included - no relational database (Excel files), no user-interface, restricted to bird and mammals, limited availability	2
Pond-FX	+ user friendly web application - no output/export device, no info about quality and source of data, no updates	2
SEISMIC	not available at workshop	0

From the latter Table, it can be concluded that most databases obtain the rating “2”, implying that the item has to be adapted to the general design criteria to a smaller or larger extent. Three databases were considered sufficiently adapted to the EUFRAM criteria that they gained the rating “1”, which implies that those items only need undergo minor adaptation when they would be adopted in the EUFRAM framework. A specific difference between the two items, being both ecotoxicity databases, is that the US-EPA Ecotox database currently allows for web-enabled searching and data collection by any thrid party, whereas the RIVM e-toxBase is currently for in-house use only. Various databases were not evaluated after obtaining hands-on experience during the workshop.

7 PRELIMINARY CONCLUSIONS REGARDING THE EUPRA RECOMMENDATIONS

The detailed recommendations for this work package, according to Hart (2001), are, as repeated from the introductory pages of this report:

1. that standard software tools for probabilistic assessment should be adopted, at a level of complexity appropriate for users in all parts of the EU regulatory arena.
2. that, in order to avoid duplication of effort, consideration should be given to whether tools developed elsewhere could be appropriate for use in the EU, either in whole or in part.
3. that there is an urgent need to catalogue existing data that would assist the development and operation of probabilistic approaches, and to collate these data in publicly-available, quality-controlled databases. The catalogue of data should include pesticide-specific data (e.g. toxicity, for use in developing SSD methods) and general data (e.g. geographical and ecological data).
4. that databases and software should be made easily accessible, to promote harmonised approaches.

As to these recommendations, in this stage of the EUFRAM project, it can be concluded that (regarding the various issues) the following remarks apply:

1. Availability of suitable software and databases is as yet not sufficient:
   - various software tools are available for the various necessary steps that can be envisaged for a tiered protocol for risk assessment;
   - the set of tools do not show the required “inter-linkage” as needed for a ‘protocolised’ risk assessment scheme in neither the technical (IT-technique) nor the logical (scientific rationale) way;
   - not all possible assessment steps are covered; and
   - most items require further development, according to the set of defined criteria, to move into the direction of optional regulatory adoption.
2. Adoption of tools requires close fit to the (eventual) regulation; developments seem unavoidable from this perspective.
3. There is as yet no open data source for all relevant types of data, this is except for the databases that can as yet be queried via a web interface (e.g., the US-EPA Ecotox database). This implies that part of the data is hidden or unavailable for use, either due to technical reasons or due to use limitations imposed by data(base) ownership. In addition, the databases that are available, at best provide data which are of limited use for the risk assessment of new pesticides (i. e. ecotoxicity data), and when data are available they were mostly collected for lower tier assessments. Especially, data that might be useful for risk assessment of different pesticides are difficult to obtain, e.g. databases about landscape characteristics and species ecology.

4. Efforts are needed to adapt the existing set of software and database items to the probabilistic risk assessment framework being developed in EUFRAM. These efforts can be focused by:

   (1) seeking complete coverage on the *types* of software and database items needed (e.g., do we cover the complete cause-effect chain), and

   (2) seeking the best implementation format per assessment step, by e.g. adapting the best candidates towards functioning in a formalised scheme as targeted in EUFRAM.

With regards to the latter, note that most software programs are relatively simplistic in the ways that they can deal with uncertainty, and that many practical issues need to be addressed when adopting any program for implementation (e.g., handling data, account for uncertainty, other distributions that the log-normal distribution, et cetera).

8 RECOMMENDATIONS FOR INTEGRATED FRAMEWORK

The EUFRAM framework for probabilistic risk assessment will be formatted, in the next stage, according to a line of further theoretical/conceptual developments resulting from the Work Packages and according to end-user testing experiences. This is to merge sound science to practical approaches.

It is recommended to:

1. Identify the tiers, steps and approaches that are chosen as candidates for adoption in the EUFRAM framework, as being currently developed in the other Work Packages of the EUFRAM development program
2. To describe the software and databases that are needed to fill out the conceptual approaches of EUFRAM as identified under (1)
3. To arrange that information on software and databases as collected in the meta-database is updated, concomitant with the further development of software and databases
4. Given the conceptual framework and the most recent descriptions of required software and databases: to re-consider the findings of the software and database evaluation made so far, so as to
   - re-confirm current findings and / or
   - to identify changes in software and databases that would influence suitability for EUFRAM
5. To make a preliminary collection of the most suitable software and databases, in order to allow for practicality tests, i.e., in the earlier end-user workshops, but also in a focused testing program by experienced risk assessors
6. To request, from the end-user workshops and the focused testers, the
   - final selection of candidate software and databases for regulatory adoption
   - the description of a development plan for EUFRAM software and databases, encompassing both the existing items themselves as well as non-existing items that are needed, and the options to link / develop all

items in such a way that evolution towards the 'ideal situation' is done according to a structured approach

7. To be prepared for making an education / instruction plan (ready when software and databases are adopted) for use in the end-user testing phase (testing the instructions as a relevant issue next to testing the software and databases)

The items as identified below should be considered in the work package on Case studies (WP8), and those passing this basic proof-of-practicality should be forwarded to the end-user test phase.

Based on the current stage of evaluation, software considered relevant for further study in the EUFRAM process is:

- *Fate / Exposure:*

The FOCUS group has provided different models that are or will likely be accepted for the registration of pesticides in Europe. It seems to be a logical next step to check if the existing models can be adapted to consider variability and uncertainty.

- *Effects:*

There are two user-friendly tools available for the construction of SSDs that have been used in peer-reviewed publications (ETX and SSWD). Both tools have some minor shortcomings, but with minor refinements they can be used both for probabilistic risk assessment, with or without weighting of species. ETX has undergone major change, in line with the criteria as established, and will be available as executable in Autumn 2004.

Ecology based models are rarely used in the risk assessment of pesticides, yet. However, AQUATOX and RAMAS provide user-friendly shells for complex models that might be used in higher tiers to support the assessment. However, simple population models to estimate recovery rates have also become available recently (e.g. HERBEST (http://www.alterra-research.nl/pls/portal30/docs/FOLDER/PERPEST/HERBEST/INDEX.HTM HERBEST), or GPS1).

Another type of effect model was evaluated to be directly suitable for EUFRAM: PERPEST, an expert system that uses results of micro-/mesocosm studies to estimate effects on aquatic communities. At the moment the database contains insecticide and herbicide studies; data from fungicide studies are currently being entered (van den Brink, pers. comm.)

- *Risk characterisation and integrated tools:*

There are different commercial software tools available for uncertainty analysis that can be used to combine (probabilistic) outcomes of fate and effects assessment. However, due to their flexibility these tools are more useful for research or special application in higher tier risk assessments. For lower tiers specifically developed tools seem to be preferable.

The current software of US-EPA provides such solutions for lower tier aquatic and terrestrial risk assessment, with one software package for each purpose. Of course, the assumptions made, the models used and the scenarios analysed are related to the regulatory context in the US. Thus, the tools might show the direction of the solution for the EU but can not be used directly. However, the tools should be available at the end user workshop to check if this kind of tools is accepted and desired by representative users in Europe.

Databases considered relevant for further study in the EUFRAM process are:

Physico-chemical properties

Data for “old” pesticides are stored in a variety of databases, most of which not dedicated to pesticides only. Further investigations in this field are warranted.

Fate and Exposure

Consider the US-environmental fate database as example of type of database.

Ecotoxicity:

- US-EPA Ecotox database
- RIVM e-toxBase
- SEEM database
- RED database

Ecology / Field

The *type* of database can be defined in comparison with existing databases (e.g., POND-FX); dedicated databases for the area under investigation need be made.

9 RECOMMENDATIONS FOR FURTHER WORK

For further work within EUFRAM, it is recommended:

1. to make arrangements, within Work Package 10 or linked to it, for the monitoring of the progress that is being made in software and database evolution, eventually yielding continuously updated insight in the qualifications of the available items as well as in software and database evaluation (against the criteria) and in the results of (end-user) testing
2. to prepare the meta-database as instrument for capturing not only qualifications, but also testing experiences
3. to establish a workgroup that arranges and stimulates the monitoring, the capture of test results and the evaluation thereof, including the capture of newly emerging results from other (EU-)projects, eventually yielding software and database that fits to the requisites of EUFRAM as close as possible, or the description of actions that should lead to that result.

10 REFERENCES

Aldenberg, T., Jaworska, J. S. & Traas, T. P. (2002) Normal Species Sensitivity Distributions and Probabilistic Ecological Risk Assessment. Species Sensitivity Distributions in Ecotoxicology (eds L. Posthuma, Suter G.W.II & T. P. Traas), pp. 49-102. CRC Press, Boca Raton, FL, U.S.A.

De Zwart, D., Hommen, U. & Posthuma, L. EUFRAM WP10 - Expert Workshop on Software and Databases for Probabilistic Risk Assessment of plant protection products, June 30th - July 2nd 2003, Bilthoven, The Netherlands. Workshop report. EUFRAM -report. 2003. www.eufram.com.

1575 Hart, A. Probabilistic risk assessment for pesticides in Europe: Implementation and
1576 Research needs. Report of the European workshop on Probabilistic Risk
1577 Assessment for the Environmental Impacts of Plant Protection Products
1578 (EUPRA). Hart A. 2001. York, UK, Central Science Laboratory.

1579 Posthuma, L., Aldenberg, A., Benfenati, E., De Zwart, D., Domenica, A., Ferson, S.,
1580 Galicia, H., Hart, A., Hommen, U., Luttik, R., Ramanarayanan, T., Randell, D.,
1581 Solomon, K. R., Traas, T. P. & Van Den Brink, P. J. Preliminary Paper on
1582 Databases and software for Probabilistic Risk Assessment of the environmental
1583 impacts of pesticides. EUFRAM (internal draft report). 2003.

1584 Posthuma, L., Suter, G. W., II & Traas, T. P. (2002a) Species Sensitivity Distributions
1585 in Ecotoxicology. pp. 587 pp. CRC Press, Boca Raton, FL, U.S.A.

1586 Posthuma, L., Traas T.P., De Zwart, D. & Suter G.W.II (2002b) Conceptual and
1587 technical outlook on Species Sensitivity Distributions. Species Sensitivity
1588 Distributions in Ecotoxicology (eds L. Posthuma, Suter G.W.II & T. P. Traas),
1589 pp. 475-510. CRC Press, Boca Raton, FL, U.S.A.

1590 Solomon, K. R. & Takacs, P. (2002) Probabilistic risk assessment using Species
1591 Sensitivity Distributions. Species Sensitivity Distributions in Ecotoxicology (eds
1592 L. Posthuma, Suter G.W.II & T. P. Traas), pp. 285-314. CRC Press, Boca
1593 Raton, FL, U.S.A.

1594 Van Den Brink, P. J., Roelsma J, Van Nes EH, Scheffer M & Brock, T. C. M. (2002)
1595 PERPEST model, a case-based reasoning approach to predict ecological risks
1596 of pesticides. Environ.Toxicol Chem., 21, 2500-2506.

1597 Van Straalen, N. M. (2002) Theory of Ecological Risk Assessment based on Species
1598 Sensitivity Distributions. Species Sensitivity Distributions in Ecotoxicology (eds
1599 L. Posthuma, Suter G.W.II & T. P. Traas), pp. 37-48. CRC Press, Boca Raton,
1600 FL, U.S.A.

1601 Van Vlaardingen, P.L.A., T.P. Traas, Wintersen, A., Aldenberg, T. (2004) ETX2.0. A
1602 program to calculate Hazardous Concentrations and Fraction Affected, based
1603 on normally distributed toxicity data. RIVM-report no. 601501028. National
1604 Institute for Public Health and the Environment, Bilthoven, The Netherlands.