

# EUFRAM

## Concerted action to develop a European Framework for probabilistic risk assessment of the environmental impacts of pesticides<sup>1</sup>

### Work Package 10

## PRELIMINARY PAPER ON DATABASES AND SOFTWARE FOR PROBABILISTIC RISK ASSESSMENT OF THE ENVIRONMENTAL IMPACTS OF PESTICIDES<sup>2</sup>

May 2003

Leo Posthuma (RIVM)<sup>3</sup>, 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), Udo Hommen (IME Fraunhofer), Robert Luttik (RIVM), Tharacad Ramanarayanan (BayerCrop Science), Donna Randall (US-EPA), Keith Solomon (University of Guelph), Theo Traas (RIVM), Paul Van Den Brink (Alterra)

## 1 CONTENTS

1	Contents.....	1
2	Executive Summary.....	2
3	Introduction.....	5
4	General approaches .....	7
5	Listing and ranking of judgement criteria for databases and software .....	10
6	Compilation of existing databases and software .....	17
7	Evaluation: databases and software confronted with criteria.....	47
8	Preliminary conclusions and recommendations .....	48

---

<sup>1</sup> EUFRAM is coordinated by CSL ([www.csl.gov.uk](http://www.csl.gov.uk)) and supported by European Commission's 5<sup>th</sup> Framework Programme ([www.cordis.lu](http://www.cordis.lu)), contract number QLK5 - CT 2002 01346. Further information and news about EUFRAM is provided at [www.eufram.com](http://www.eufram.com).

<sup>2</sup> EUFRAM partners welcome public comment and discussion on this document. Comments may be contributed via the EUFRAM website at [www.eufram.com](http://www.eufram.com). This document is the responsibility of its publishers and in no way represents the views of the European Commission or its services.

<sup>3</sup> Work Package leader.

## 2 EXECUTIVE SUMMARY

Probabilistic methods can be used to estimate the frequency and magnitude of impacts in risk assessment, and to take account of variability and uncertainty. The EU Scientific Steering Committee's Working Group on the Harmonisation of Risk Assessment has recommended that these methods should be considered for adoption in the general area of human and environmental health.

The EUPRA workshop (2001) considered the potential of probabilistic methods for assessment of ecological risks of plant protection products, under European Directive 91/414/EEC. The workshop reviewed the state of the art, and made detailed recommendations, amongst others on databases and software to be used for probabilistic risk assessment.

This *Preliminary Paper on Databases and software* appears as response to the recommendations on databases and software of the EUPRA workshop. It has been prepared within the framework of the *EUFRAM project*, the "Concerted Action to develop a European Framework for probabilistic risk assessment of the environmental impacts of pesticides", within Work Package 10 "Software and databases for probabilistic assessments". This report, in a preliminary fashion, 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 software and database items currently available.

Three further versions of the report will appear:

- First, the preliminary report will be discussed by the *EUFRAM Steering Committee* in May, 2003. The report, when extended by the amendments suggested by the *Steering Committee*, will yield the *Preliminary Paper on Databases and software* (Internal project paper)
- Second, that version will be discussed in an *Expert Workshop on Databases and software* to be held in June, 2003. This will yield *Version 1* of the report (to be posted on public website).
- Third, after receiving feedback from the other work packages of EUFRAM, the final *Version 2* of the report will be prepared (to be posted on public website).

The last report version will be the basis for the second half of the *EUFRAM* project, which amongst others focuses on end-user testing.

## **Preface**

This report is compiled as result of a short compilation process (January – March 2003).

A first attempt was made to collect the preliminary contributions of the expert members of *EUFRAM* work package 10 and other experts on three subjects in this preliminary paper, as basis for targeted discussions, especially at the workshop.

The contributions are shaped in:

1. a draft list of criteria for the evaluation of databases and software for probabilistic risk assessment
2. a draft list of available databases and software items
3. a draft program for the expert workshop on databases and software to be held in June, 2003.

Contributions have been received from contractors of work package 10 of the *EUFRAM* project.

## **Acknowledgements**

The work package leader is very grateful to the colleagues and organisations that contributed to this preliminary paper for the *EUFRAM* work package on Software and Databases, to the European Commission and RIVM for funding, the Steering Committee of *EUFRAM* and Andy Hart for general coordination of the *EUFRAM* project, Klazien Groeneveld for starting preparations for the workshop, and various colleagues at RIVM within and outside the context of *EUFRAM* for sharing their expertise on software and databases.

## 3 INTRODUCTION

### 3.1 Problem definition

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” to the concept of “risk assessment” 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 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 requires EU member states to analyse these risks before authorising pesticides for sale. The EUPRA workshop considered the potential of probabilistic methods for assessment of ecological risks of plant protection products in relation to 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 databases and software to be used 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 databases and software. These recommendations are (see Hart, 2001):

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.

Regarding the availability of software, the EUPRA report already mentioned the existence of a relevant body of work of US-EPA.

Regarding the availability of data, it was noted that some data-sharing initiatives were already underway, but that data sharing may be hindered by issues related to commercial interest, confidentiality and data ownership. Further, technical limitations hamper the application of PRA-techniques in practice: as yet, no single, easily accessible database covering all data sources needed has been designated or designed. Solutions for these problems require co-operation with the owners of the data. Eventually, sharing data would provide benefits to all stakeholders, and would reduce the amount of animal testing required in future. To underscore the need and urgency for solving these problems, it can be mentioned that the EU recently issued a restricted Call for Tender, to promote the design of a freely accessible Reference Database (the RED-project) for compiling ecotoxicology data on plant protection products and their metabolites (EEC, 2002). The work for that project was suggested to built forth on a nucleus that was earlier established from an earlier contract, that is: the *SEEM* project (EEC, 2001).

To address the *EUPRA* recommendations, *EUFRAM* work package 10 focuses specifically on databases and software. This report, in a preliminary fashion, 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.

### 3.2 Objectives

According to the *EUFRAM* project description, the objectives of the work package on software and databases are:

1. to establish a common methodology for selecting databases and software for probabilistic pesticide risk assessments
2. to compile examples of available databases and software
3. to evaluate examples currently in use by partners for possible adoption as harmonised standards in the EU

The current report subsequently addresses which objective criteria could apply for selecting databases and software (Chapter 5), it gives a compilation of existing databases and software (Chapter 6), and it confronts the software and the databases with the objective criteria (Chapter 7).

### 3.3 Method overview and limitations of this report version

The current version of the report is to be considered a *preliminary paper* only, that is: various versions will follow. Subsequent versions will be produced after discussing the paper in the broader *EUFRAM* context of progress and

implementation activities. The latter will be part of the *EUFRAM* phase of end-user testing and implementation. The subsequent papers are thus:

1. Version 1 - September 2003 (Public)
2. Version 2 - September 2004
3. *EUFRAM* end report - December 2006

The current version will thus consist of a first gross attempt to compile evaluation criteria and existing databases and software. The contributions obtained so far stem from selections from the experience of the work package members and interested others, extended with some specific experiences with software and database implementation from ongoing works in the institute of the work package leader (RIVM).

The compiled material will especially be used as discussion paper in the spring 2003-meeting of the *Steering Committee* and in the *Expert Workshop*. Hence, the current report is by no means to be considered either complete or final:

- the list of criteria can be extended or changed on the basis of amendments suggested by work package participants or others
- the list of available databases and software can be extended by adding new items, new items being optionally proposed by work package participants or others
- the confrontation of the available software- and database items with the criteria, in the current report, only takes the format of finger exercises, since neither the criteria nor the list of items that are to be evaluated have been discussed yet.

In accordance with the approach followed in the *EUPRA* workshop, however, the format for the confrontation between criteria and items takes the format of a SWOT-analysis, that is: Strengths, Weaknesses, Opportunities and Threats will be listed to end up with condensed evaluation results.

## 4 GENERAL APPROACHES

### 4.1 Questions to be asked

The currently available databases and software items (see Chapter 6 for draft shortlists) have not been designed for the specific targets of the *EUFRAM* project. For example, a software item may be addressing only parts of the potential framework for PRA of plant protection products. A software item may further be user-friendly only for the specific tasks implemented by its developer. A database may not be freely accessible. In addition, a compilation of databases may be needed to execute the PRA process for a compound.

Each database and software item can, in principle, be considered in the *EUFRAM* project for eventual implementation as PRA-tools in risk

assessment practice in the EU. However, other *EUFRAM* work packages likely yield views on database fields that are to be preferred (since they are crucial to the preferred PRA-models), and on software (since software may contain preferred characteristics, like an explicit way of handling confidence limits throughout the process of PRA).

By acknowledging the role of the other work packages, and by accepting that the results of those work packages will co-evolve with the current work package, the problems for this work package are:

- what exactly are the subjects of evaluation (Section 4.2), and
- how to choose among alternatives (Section 4.3), especially when candidate items do have both profitable and unprofitable characteristics when confronted with the *EUFRAM* targets.

## 4.2 The subjects of evaluation

Regarding the question on ‘what exactly needs to be evaluated’, the *EUFRAM* software and database work package explicitly looks at functional products that execute a task in the PRA process, being databases and software for PRA. In this work package, the evaluation efforts are *not* focused on which of two alternative theoretical concepts that execute the same type of task should be preferred. For example, alternative methods for uncertainty analysis are treated in work package 4, and probabilistic approaches for use with typical data sets are treated in work package 5. The other work packages should deliver the general conceptual approaches and indicate preferred ones, or test them later on (the work package on case studies, and the end-user workshops), while the current work package only looks at the technical implementation formats of those general approaches.

### 4.2.1 Database evaluation

The meaning of ‘functional product’ for the item ‘database’ is conceptually summarised in Figure 1.

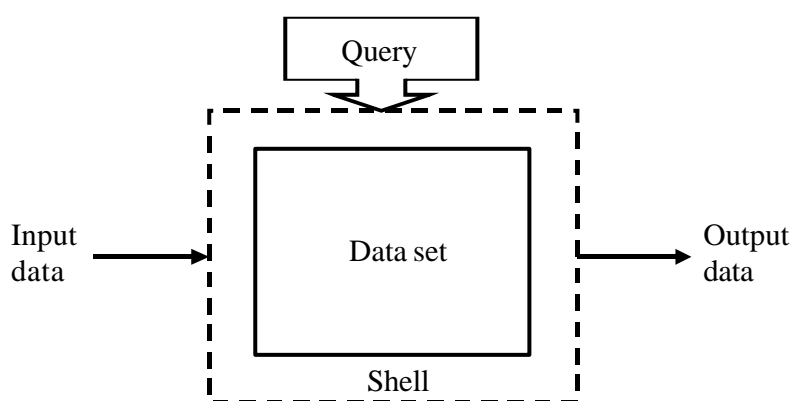


Figure 1. The item that is evaluated regarding databases is any functional product that consists of the components shown in the graph. For further explanation: see text.

All data together compiled in electronic format form the data set. This data set is captured in a shell, like a spreadsheet or a database program. The shell itself is empty, like a spreadsheet tab in MS-Excel, or an empty MS-Access

file. An operational database should allow for adding new input data (e.g., for newly marketed compounds) according to structured sets of data fields that define the characteristics of the data to be entered (e.g., concentration fields are defined in either  $\mu\text{mol/L}$  or as  $\mu\text{g/L}$ , but no use of both within a single database). Further, it should be able to deliver output data for PRA, either in the format of reproductions of selections of raw data, or as processed data (e.g., averages), based upon the possibility to query for the required selection of data. Only together, these matters make up a functional database. Any product of this kind (being either a set of data in a spreadsheet format or in a true database format) can be considered as a database worthwhile evaluation for *EUFRAM*.

#### 4.2.2 Software evaluation

The meaning of 'operational product' for the item 'software' is conceptually summarised in Figure 2.

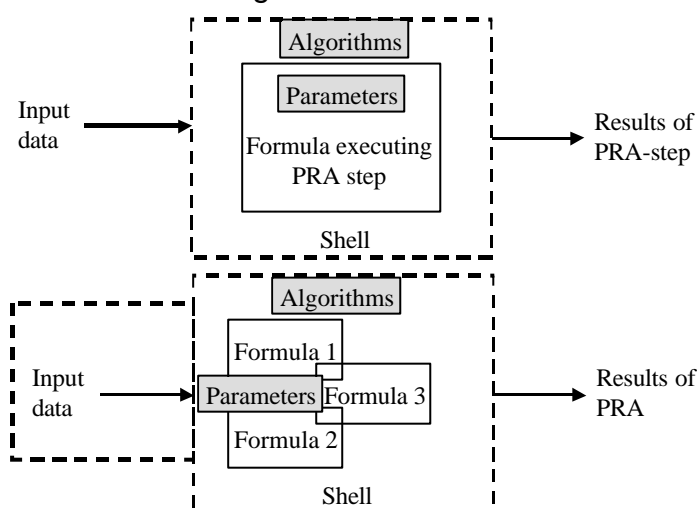


Figure 2. The item that is evaluated regarding software is any operational product that consists of the components shown in the graph. It can have various formats regarding the ways components are built-in into the product (compare top and bottom as two examples). For further explanation: see text.

The formula and algorithms executing or supporting a PRA step (top) or a set of formulae and algorithms that execute or support various PRA-steps (bottom), are the mathematical and representation(s) and technical approaches of the PRA-concept that is being addressed. For example: the log-normal model for a Species Sensitivity Distribution approach is a formula for one step in PRA. For example: the generation of random numbers, or the way of constructing a graphical output, or of solving formulae that have no analytical solution are executed by programmed algorithms. The set of formulae and algorithms is captured in a shell, like a general spreadsheet shell (like MS-Excel), or specific statistical or modelling shells (like MatLab), which themselves are 'empty'. Functional software should allow for adding the relevant input data on the compound to be assessed for the PRA processes. Parameters to execute the PRA-step, e.g. on the species that is modelled, may be contained in the model formulae, or they may be required as additional input (next to the input data pertaining to the compound under study).. Regarding the addition of input data on the characteristics of the

compound being investigated there are two options. The input data can originate from a data base query, whereby the database is not necessarily linked to the software product. Data are then transferred 'manually' to the software shell. As an alternative, the functional software may contain part of the input data (depicted by the extended shell in the bottom graph only). The software should further be able to deliver meaningful output for PRA, either as Results of (a single) PRA-step, or as (final) Results of PRA. Together, these matters make up a functional software unit. Any product of this kind, even when the software addresses only a small (but relevant) part of a PRA, can be considered as a software unit worthwhile evaluation for *EUFRAM*.

#### **4.2.3 Additional items for evaluation: linkages between operational products**

Evidently, the possibilities of linkages between the required database(s) and the software, and between separate software units within the different steps of PRA when an overall-programme is not available, are important to consider in addition to the databases and software items themselves. This relates to the general *EUFRAM* criteria that the chosen databases and software items should preferably easily applied and freely accessible.

### **4.3 Outline of the evaluation process**

Regarding the question 'how to choose among alternatives' for adoption as practical tool, it is evident that each existing item has profitable characteristics for the aims for which it was designed – but do these profitable characteristics also hold for to the general European implementation of PRA to assess plant protection products?

To enable a conscious choice between optional candidates, or to suggest merging items in the case of databases that are complimentary, each item should be confronted with a set of (preferably) objective criteria. In the next Chapter, such criteria are listed in draft, and sorted according to different evaluation targets of this work package, such as user-friendliness, accessibility and fit to eventual implementation as *EUFRAM*-tool for PRA. In Chapter 6 the available databases and software items are listed. In Chapter 7 the criteria are used to evaluate the databases and software items.

## **5 LISTING AND RANKING OF JUDGEMENT CRITERIA FOR DATABASES AND SOFTWARE**

### **5.1 Approach**

A first draft of a list of objective criteria was obtained from collecting experiences of work package members. These experiences resulted from the works on the development of databases and software at the members' institutes, or the use thereof. Note that the list can, in later report versions, be extended with views collected from reviews of the open and grey literature (e.g., US experiences), and changes are to be expected based on interactive discussions in the *Steering Committee* and at the workshop.

The collected criteria are categorised as to various points of attention. Relevant subdivisions are possible regarding, e.g.:

- aspects related to application characteristics (e.g., free accessibility, user friendliness, data coverage of databases)
- aspects related to scientific characteristics (e.g., validation of software, documentation, linkage to adopted theoretical concepts for PRA, et cetera)

Other categorisations are possible, e.g., in view of the SWOT analyses of Chapter 7. The list of criteria for Strengths and Weaknesses might cover the above items, while Opportunities and Threats might suggest taking broader issues into account. Examples of the latter are, that the selected items, despite them being “optimal” per item, together need not yield a complete approach.

## 5.2 General sources for evaluation criteria

In general, there are many considerations to be taken into account to evaluate software and databases. A restricted web search on ‘software+evaluation+criteria’ learned that various resources can be considered to get raw criteria lists for generic database and software evaluation (that is: not specific for the targets of PRA as technique or *EUFRAM* as project). For example, <http://www.so.cc.va.us/vccsit/Archive/SoftEval.htm> lists the criteria types and items for basic database and software evaluation as summarised in Table 1.

---

### General evaluation criteria

---

1. Learner population(s) – *describe all elements of the population; profile the learners*
    - Demographics
    - Learning styles
    - Common characteristics
    - Unique characteristics
  2. Learning environment – *in what context is the "learning" to occur*
    - Format
    - Time frame
    - Outcomes – content goals
    - Instructor/facilitator – expertise & expectations
  3. Technology infrastructure
    - Capabilities & limitations
    - "Product" requirements
    - Administrative requirements
    - Implementation & maintenance
    - Licensing parameters
  4. Instructional goals
    - What needs to be accomplished
    - How can this best be done
    - Which technologies/approaches are most appropriate
    - Philosophical or theoretical basis
- 

Table 1. General software criteria as published on <http://www.so.cc.va.us/vccsit/Archive/SoftEval.htm>

The authors of this source have also provided a criteria list that provides a broadened structured checklist, as an extension to the draft Fact Sheet filled out by work package members for *EUFRAM*. The extended checklist is provided in Table 2.

<b>General information</b>			
Product name			
Producer			
Price options			
Content covered			
Intended audience(s)			
Documentation (provided/useful)			
Program free of bugs & errors			
Entry level skills required (describe)			
Computer skills required (describe)			
Other technical skill requirements (describe)			
<b>General Features</b>			
	Yes	No	Comments
Glossary			
Index			
Menus			
Writing style/technique			
Grammar			
Contextual examples			
Maintains user interest			
User records/activities are secure			
<b>Content</b>			
	Yes	No	Comments
Content error-free			
Appropriate for intended audience			
Sufficient depth of coverage			
Level of difficulty options			
Appropriate use of technology			
Content can be modified/updated			
Content can be added			
Objectives clearly stated			
Outlines			
Summaries			
Reviews/tests			
Test types			
Pre-tests			
Post-tests			
Multiple choice			
Essay			
Short answer			
True/false			
Projects			
Port folios			
<b>Technical aspects</b>			
	Yes	No	Comments
Installation directions			
Technical support			
Self-paced access			
Screens easy to read			
Navigation -- back and forward home			
Instructions can be bypassed			
Instructions are accessible throughout program			
Help option			
Exit options (stop & restart at a point)			
Appropriate hyperlinks			
Number			
Relevance			
Easily added			
Easily deleted			
Font type			
Icons			
Location			
Logic			
Graphics -- stills			
Appropriate use			
Appropriate for content			
Clearly labeled			
Video			
Appropriate for content			
Quality			

Table 2. Formalised checklist for software and database evaluation (from <http://www.so.cc.va.us/vccsit/softchek.htm>).

### 5.3 EUFRAM criteria used so far

The collected criteria from the work package members, which have served in the information compilation process so far, have been generated mainly from the specialists' perspectives and are sorted. The resulting sorted lists of potential objective criteria are summarised in the subsequent paragraphs, starting from the definition of 'the ideal case'. The collected criteria are ranging from ones that may eventually appear as highly relevant for *EUFRAM* targets to ones that can eventually be considered of only weak relevance. The *Expert Workshop on Software and Databases* will need to compile the general and the *EUFRAM*-specific criteria, and evaluate them, as to determine their relevance for the end product of the *EUFRAM* project. The workshop produces a workshop report, presenting the program, work methods and results. The Strengths, Weaknesses, Opportunities and Threats analysis (SWOT-analysis) methodology will be used to structure the feedback to the draft criteria list, both from the *Steering Committee* and from the workshop participants.

### 5.4 Databases

Scientifically, an ideal database for PRA of plant protection products should encompass all necessary data fields that serve as input data or parameters in the appropriate PRA models. In view of the use in PRA, the fields that contain data used in PRA-assessment steps should originate from various tests (e.g., NOECs for various species). The necessary data fields are defined by the regulatory requirements set by the Commission Directive 91/414 for Annex III (formulated plant protection products) and where applicable supported by Annex II (Active ingredients), or the specifications thereof defined in the *EUFRAM*-project. Thereby, all fields are unequivocally defined and given in standard units throughout the data set. Further, the data compiled in the data fields should be scientifically valid (no errors between raw data and database entry), of sufficient (known) and traceable quality, and traceable to the original data sources (studies). Further, it should allow for easy entering of new data, efficient querying of the data set regarding various query-definitions, and easy export of the data for modelling purposes (i.e., output formatted in the way that the software can be 'fed' directly). Finally, it should be documented well, both scientifically and practically, and it should be available for broad use.

When this is considered an appropriate verbal translation of 'the ideal database for PRA', then a derived list of science-related criteria can be produced to test databases (by the SWOT analyses of each of the available software items as made in draft in Chapter 7). Such a list is (in draft) summarised in Table 3 according to the categorisation into scientific and application criteria. No evaluation of relative importance has been given yet. That is a task for the workshop.

Category	Criterion	Subcriterion	Evaluation
Science	Data validity	Raw data = data entry	
Science	Data validity	Quality earmark	
Science	Data aspect	Fate	

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

Table 3. Draft list of objective criteria that can be used to evaluate the appropriateness of an available database for the eventual EUFRAM targets

## 5.5 Software

Scientifically, an ideal software package for the PRA of plant protection products should be able to execute all scientifically required assessment steps, that are (like for the ideal database) defined according to the Commission Directive and the EUFRAM project. That is: there should be assessment models for exposure characterisation, for effect characterisation, and for risk quantification. The steps executed in data handling should be scientifically valid and traceable. Further, an ideal package should enable easy 'feeding' on the input data on exposure and effect, and it should be able to provide output in one or more meaningful ways. Furthermore, it should be documented well, both scientifically and practically, and it should be available for broad use.

When this is considered an appropriate verbal translation of 'the ideal software for PRA', then a derived list of criteria can be produced to test software items (by the SWOT analyses of each of the available software items as made in draft in Chapter 7). Such a list is (in draft) summarised in Table 4,

according to the categorisation into scientific and application criteria. No evaluation of relative importance has been given yet. That is a task for the workshop.

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

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

In addition to this list, a project partially linked to the *EUFRAM* project (the DEFRA-projects, see Summary Sheets for information) generated the following criteria list for software evaluation:

## **WEBFRAM-Indicative features of web software**

### *User interface*

- Home page
- Information pages, help pages, links to relevant websites (material provided by module projects)
- Scenario selection screens
- Data entry screens
- Output/report option selection screens
- Option to do model calculations online or in batch
- Results screens including a range of graphical and tabular outputs, suitable for use in reports/documents, with option to print/download/save online
- Option to save data/scenarios for later re-use
- Consistent look and feel for all modules
- Default options and data for every module (so user can just run or modify a standard example)
- Information screens on the standard examples (generated by case studies in module projects)
- List-server to announce changes/additions to registered users
- Option to limit functionality available to different types of users

### *Model inputs*

- Option to vary analysis options (e.g. number of iterations) – but decide during project how much to limit these choices
- Default distributions for all parameters (from module project case studies)
- Option to specify own distribution (type and parameters) for each parameter
- Option to use default or user's raw data for bootstrapping?

### *Uncertainty analysis*

- Type(s) of uncertainty propagation to be decided during project and likely to be limited to one or two, but may include: first-order error analysis, 1D and 2D Monte Carlo, Bayesian updating, P-bounds, interval analysis.
- Need to be able to do uncertainty propagation within the software generated by the project, or by running existing packages using macros.
- Sensitivity analysis (methods to be decided during project)
- Need to consider what can be done online vs. in batch mode

### *Testing*

- Alpha and beta testing (see above for details)

### *Maintenance*

- Technical manual and code to be provided, to enable customer/others to maintain/upgrade/expand system after delivery
- Technical spec. for server requirements
- Documentation within program code
- Functionality for customer to modify default data, including alteration of constants and substitution of altered distribution parameters or raw data.
- Functionality for customer to modify default settings of other options (analysis, reporting etc)

## 6 COMPILATION OF EXISTING DATABASES AND SOFTWARE

### 6.1 Overview of available databases

The databases covered so far in this project phase are summarised in Table 5. No detailed attempts have been made in this phase to extend the items suggested so far by a web-search for 'new' items. Preliminary activities, searching for the combination of the word database (or similar) and targets items such as fate, exposure, ecotoxicity, ecology and even taxonomy revealed a large number of possible items of very diverse origins, coverage, quality and scopes. In later phases, extended and better targeted web-searches into areas that are crucial but as yet uncovered might provide the identities of these items. As yet, the crucial databases are assumed to be within reach of the work package members, and those are listed here.

In the column "EUFRAM-need", the items are to be categorised as to the question whether the *EUFRAM Steering Committee* and the other work packages considers an item crucial, relevant, or irrelevant (++, +, -) for detailed evaluation.

Coverage	Owner	Name	EUFRAM-need (++, + or -)	Code *#
<i>Items covered in the Summary Sheets and Fact Sheets</i>				
Chem. properties	IME	Propertbase		IME-D1
Chem., properties	ICP	MINAMB		ICP-D1#
Chem. properties	Other	Various items		
Fate	US-EPA	Pesticides Environm. Fate Database		EPA-D1*
Exposure (assim. eff. food)	RIVM	Assiwerk		RIVM-D1*
Exposure (energy content food)	RIVM	Enerwerk		RIVM-D2*
Effects	RIVM	RIVM-e-toxBase		RIVM-D3*
Effects (ecotox)	US-EPA	US-EPA Ecotox database		EPA-D2*
Effects (ecotox)	IME	Propertbase		IME-D1

Effects (ecotox)	ICP	MINAMB	ICP-D1#
Effects (human)	US-EPA	US-EPA NTP database	EPA-D3*
Effects and metabolites	EU/Cranfield	RED-database	EU/Cranf-D1
Effects and metabolites	ICP	SEEM-database	ICP-D1#
Ecology (not submitted)		e.g., POND-FX (?)	
Ecology (not submitted)		e.g., EcolMAT initiative	
Meta-database		LIMOD	

Table 5. Compilation of available databases, categorised according to concepts that are covered. For detailed information: see overview descriptions in the Summary Sheets (subsequent paragraphs) and concise specification information in the Fact Sheets. \* marks that written Summary Sheet is available for gaining overview on the general purposes and designs of items (see Chapter 6.3), # marks that tabulated fact sheet information is available for easy comparisons among items.

The compiled list of databases shows that database items on compound properties, fate, exposure, effects, ecology and metabolism have been suggested, whereby for part of the suggested items detailed information is provided, whereas for other items it is suggested that have them considered. Note that, as yet, suggestions for specific databases on ecological characteristics of the species under consideration are lacking, although there have been recent initiatives in the framework of the European Science Foundation to fill this gap by developing such a database (see the EcolMAT web pages:

[http://www.esf.org/esf\\_domain\\_applications.php?language=0&activity=5and6&domain=3](http://www.esf.org/esf_domain_applications.php?language=0&activity=5and6&domain=3)).

Further, for many assessment activities, the basic physico-chemical properties of the compounds under investigation are needed, and no specific suggestions on this have been provided, while various 'general' sources do exist.

As preliminary conclusion for this stage of assessing the availability and usefulness of databases for the context of *EUFRAM*, it appears that the data sources seem relatively scattered (availability), incomplete (coverage per item) and not fully covering the field (coverage of PRA-assessment steps). The software and database workshop needs to consider the actions to be taken to complete on (1) the coverage of PRA-assessment steps and (2) the coverage of data in one or a lower number of sources for each of the steps. It seems *a priori* undesired to apply PRA methods on the basis of data sets on e.g. ecotoxic effects when those data should be compiled from different sources with incomplete coverage each.

## 6.2 Overview of available software items

The software items covered so far in this project phase are summarised in Table 6. No detailed attempts have been made in this phase to extend the items suggested so far by a web-search for 'new' items. Preliminary activities, searching for the combination of the word software or models (or similar) and targets items such as fate, exposure, ecotoxicity, ecology and even taxonomy revealed a number of possible items of very diverse origins, coverage, quality and scopes. In later phases, extended and better targeted web-searches into

areas that are crucial but as yet uncovered might provide the identities of these items. As yet, the major required software items are assumed to be within reach of the work package members, and those are listed here.

In the column “EUFRAM-need”, the items are to be categorised as to the question whether the *EUFRAM Steering Committee* and the other work packages considers an item crucial, relevant, or irrelevant (++, +, -) for detailed evaluation.

Coverage	Owner	Name	EUFRAM-need Code * # (++, + or -)
Fate			
Exposure	IME	MCPHELMO	IME-M1*#
Exposure	IME	FOCUS??	
Exposure			
Effects (SSD)	IME	SSD.xls	IME-M2*#
Effects (SSD)	IME	SSD.delphi	IME-M3*#
Effects (pop.)	IME	GPS (General Population Simulator)	IME-M4*#
Effects (pop.)	IME	Daphnia	IME-M5*#
Effects (pop.)	IME	Zebra	IME-M6*#
Effects (cosm)	IME	SAM3 (Sim. Of Aquatic Microcosm)	IME-M7*#
Effects (web)	IME	FLEXSWAC	IME-M8*#
Effects		RAMAS	
Effects	US-EPA	AQUATOX	EPA-M1
Effects	US-EPA	Ecofram works	EPA-M2
Effects		Ecotox Cd (Jorgensen)	
Effects	Alterra	PERPEST	Alt-M1#
Effects	RIVM	ETX	RIVM-M1*#
Effects	RIVM	Busy	RIVM-M2*#
Effects	RIVM	Sec. Poisoning	RIVM-M3*#
Effects	Alterra	PERPEST	Alt-M1 *
Meta / project	CSL e.a.	WEBFRAM-program '03-'07	CSL-M1*#
Effects		LIMES	

Table 6. Compilation of available software items, categorised according to concepts that are covered. For detailed information: see overview descriptions in the Summary Sheets (subsequent paragraphs) and concise specification information in the Fact Sheets. \* marks that written Summary Sheet is available for gaining overview on the general purposes and designs of items (see Chapter6.2), # marks that tabulated fact sheet information is available for easy comparisons among items.

The compiled list of software items shows that an array of models has been suggested, whereby for part of the suggested items detailed information is provided, whereas for other items it is suggested to have them considered. Emphasis in the submitted software items is on effects assessments. The

collected information is absent or small for exposure and fate models. However, suggestions have been made to screen various web sites, such as:

- <http://www.riskworld.com/SOFTWARE/SW5SW001.HTM>. This website provides a list of risk assessment related software. It seems to include only US programs and companies, it is not specifically oriented on pesticides or environmental risk assessment, but some links might be worth to check. For example:
  - 1) The environmental software of the Cadmus Group, such as the Software and User's Manual for "Aquatic Ecological Risk Assessment: A Multi-Tiered Approach", a program operationalising SSDs.
  - 2) The environmental software of "Lakes Environmental". The software items of this company for air dispersion modeling, risk assessment (human health and ecological), and emergency release modeling are widely used by consultants, industries, governmental agencies and academia. All programs are Windows based and provide full graphical visualisation, including import of base maps from favorite formats. The software incorporates GIS tools by ESRI, the makers of Arc View and Arc Info, to provide users with the capability of GIS-presentation of PRA-results. Their advanced risk assessment software for human health (IRAP-h View) and ecological risk (EcoRisk View) is based on the most recent guidance from the U.S EPA. Comprehensive risk assessments can now be completed in a fraction of the time traditionally required. ISC-AERMOD View is useful in the field of air dispersion modelling by providing three U.S. EPA models, ISCST3, AERMOD, and ISC-PRIME under an integrated interface. For denser-than-air modelling and accidental releases, SLAB View provides an interface to the U.S EPA SLAB model. RMP View and SEVEX View are used to prepare risk management plans or comprehensive emergency response plans.

or

- <http://www.wiz.uni-kassel.de/ecobas.html>. On this web site, the Register of Ecological Models (REM) is provided. The REM is a meta-database for existing mathematical models in ecology.

Furthermore, a recently published volume on ecological modeling in risk assessment (Pastorok et al., 2001) may be a valuable reference for extending the list of optional modeling tools. The volume provides a description of the model features beyond the short Summary Sheets of this report, and provides general evaluations that are made in the context of that volume.

As preliminary conclusion for this stage of assessing the availability and usefulness of software items for the context of *EUFRAM*, it appears that the data sources seem relatively scattered (availability), incomplete (coverage per item) and showing an emphasis on items for effects assessment, which is unlikely to provide sufficient coverage of PRA-assessment steps. The software and database workshop needs to consider the actions to be taken to complete on (1) the coverage of PRA-assessment steps and (2) the choice or combination of items to cover separate assessment steps.

### **6.3 Available operational databases: Summary Sheets**

The available operational databases are summarised on the subsequent pages, each item has approximately one page of Summary Sheet in writing.

### **6.3.1.1 Summary Sheet: Cranfield – RED -database**

(Description by A. Boxall)

#### **Name of the database, owner, identification**

The RED database is currently being developed by Cranfield Centre for EcoChemistry for the European Commission. Contact persons are a.boxall@cranfield.ac.uk and c.sinclair@cranfield.ac.uk.

#### **Background information**

The development of the RED database was started in January 2003 and the final database will be available in October 2003. The database will contain validated data on the ecotoxicity of pesticides and their transformation products which will allow the statistical distributions of the ecotoxicological effects of active parent compounds and their metabolites to be explored.

#### **Characteristics regarding software**

The database is currently under development. The final version will be web-enabled and the exact characteristics of the software that will be used are currently under discussion.

#### **Characteristics regarding data coverage**

Currently, data coverage concerns mostly aquatic ecotoxicity data for pesticides and their transformation products. Information on selected physico-chemical properties, chemical structure and toxic mode of action are also contained in the database.

#### **Basics of the database**

The data base is designed to explore relationships between parent and metabolite ecotoxicity and to provide insights into the sensitivity of different species to both parent compounds and their metabolites. All data in the database will be public domain and it is envisaged that the database will provide a valuable resource for the risk assessment of pesticides.

#### **Characteristics regarding output**

The RED database will be web enabled which will mean that users will be able to extract data online.

Planned or expected further development (future versions)

It is envisaged that once the database is completed, it will be hosted by the European Commission and will be regularly updated to include newly available data.

### 6.3.1.2 Summary Sheet: Cranfield - SEISMIC

(Description by J. Hollis)

#### **Name of the database, owner, identification**

The SEISMIC software system has been developed by The National Soil Resources Institute of Cranfield University on behalf of the UK Department of Environment Food and Rural Affairs, the Pesticides Safety Directorate and the UK Crop Protection Association. It is available on CD-ROM for use to support the prediction of pesticide environmental exposure in England and Wales and is subject to an annual lease for periods of up to 5 years. Contact persons are [j.hollis@cranfield.ac.uk](mailto:j.hollis@cranfield.ac.uk) and [m.j.d.dufour@cranfield.ac.uk](mailto:m.j.d.dufour@cranfield.ac.uk).

#### **Background information**

Complex, mechanistic environmental fate models play an increasing role in environmental risk assessment for diffuse-source pesticides. SEISMIC is a unique and easy to use environmental information system designed to provide the detailed soil and weather data necessary to parameterise a wide range of chemical fate models for any environmental scenario in England and Wales. Launched in 1994, SEISMIC has now become the regulatory-preferred system for supporting the estimation of diffuse-source pesticide exposure.

#### **Characteristics regarding software**

The latest version of SEISMIC (v. 2.0) is programmed in MS Visual Basic and utilises MS Access databases and ESRI MapObjects technology. The software is compatible with MS Windows 9\*, 2000, NT, and XP. It does not require any additional GIS software to run.

#### **Characteristics regarding data coverage**

SEISMIC holds both spatial and parameter data derived from NSRI's comprehensive Land Information System, LandIS. The spatial distribution of climate characteristics, agricultural crops, soil types and aquifer types are all held as 5 x 5 km grid data giving the percentage cover of each category in each grid cell. These can be overlaid to map the distribution of user-specified combinations. Cropping patterns based on agricultural statistics for the years 1988, 1995 or 2000 can be selected. A national summary of the proportion of land that is represented by the user-selected combination is provided.

Soil parameter data representing all 412 soil types in the spatial database can be accessed to parameterise models. For each soil type, data is available for all significantly different soil layers to a depth of 1.5m or rock and for the following properties: layer upper depth, layer lower depth, particle-size fractions, organic carbon content, pH, bulk density, hydraulic characteristics including water retention at different suctions and saturated conductivity. For each soil type, data is available for the following land uses: Arable, short-term rotational grassland, permanent grassland and semi-natural vegetation. Computed 30 year daily weather datasets for representative sites across the two countries are also supplied to drive the models. Weather parameters included are: Precipitation, Maximum, minimum & mean temperature, Sunshine hours, Wind speed and Relative humidity. Probability distributions of the variable for each station show that the computed datasets match the measured data up to the 99th percentile extreme values.

#### **Basics of the database**

Interaction of the spatial datasets in the system is driven by drop-down menus. Generated maps can be exported as image files and easily compared using adjacent windows. Data files of soil property characteristics and daily weather can be easily downloaded for model input.

**Characteristics regarding output**

Version 2.0 of SEISMIC has been both alpha and beta-tested. Soil property and daily weather data files are exported in both .dbf and .txt format, whereas map images are exported as .bmp files or in enhanced metafile format. All the spatial data are protected within the system and cannot be downloaded by users. Records of each user session are generated automatically and kept as a session.txt file that identifies all queries and actions carried out within a user session.

**Planned or expected further development (future versions)**

Plans exist to link the SEISMIC system to an Aquatic Ecosystems Database for the UK to enable the derivation of data on both the physical environment and biological characteristics of aquatic ecosystems relevant to user-specified crops.

**6.3.1.3**

### 6.3.1.4 Summary Sheet: WEBFRAM

(Description by A Hart)

#### **Name of the software, owner, identification**

WEBFRAM is the collective name for a suite of research projects funded by UK Department for Environment, Food and Rural Affairs for the benefit of the UK Pesticides Safety Directorate. The key output of the projects will be a suite of software modules to perform probabilistic assessment of pesticide exposure and effects for aquatic and terrestrial organisms. The overall contact persons are [a.hart@csl.gov.uk](mailto:a.hart@csl.gov.uk) and [h.thompson@csl.gov.uk](mailto:h.thompson@csl.gov.uk).

#### **Background information**

The individual projects are:

WEBFRAM 1 – will establish common principles for the whole suite of projects including approaches to uncertainty and variability, receive probabilistic models and supporting data from the other WEBFRAM projects, and implement them as user-friendly web-enabled software.

WEBFRAM 2 – will develop probabilistic models for aquatic organisms

WEBFRAM 3 – will develop probabilistic models for terrestrial vertebrates

WEBFRAM 4 – will develop probabilistic models for non-target arthropods

WEBFRAM 5 – will develop probabilistic models for soil organisms

WEBFRAM 6 – will develop approaches for evaluating the acceptability of risk (unlikely to be implemented as software)

WEBFRAM 7 – will develop new approaches for assessing long-term and population-level risks to birds and mammals (these may be taken up in WEBFRAM 3)

The project runs for 4 years from January 2003; no software is available yet. Early outputs will include:

June 2003 (EUFRAM software workshop). Prototype webpages to demonstrate the proposed structure of the user interface.

October 2003 (EUFRAM case study meeting 1). Model description and case study results for acute risks to birds and mammals.

March 2004 (EUFRAM case study meeting 2). First prototypes of web-enabled modules available for demonstration.

December 2004 (EUFRAM end-user workshop). First web-enabled modules ready for beta-testing.

#### **Characteristics regarding software**

WEBFRAM 2-6 and 7 are expected to develop their models using various types of software, but the final versions of their model algorithms will all be re-programmed and web-enabled within WEBFRAM 1. This will be done by Cadmus Group (Seattle office, USA), using a common user-interface developed with ASP.Net, a subset of the new Microsoft .Net programming platform (in effect, an improved extension of Java).

#### **Characteristics regarding data coverage**

The software will include generic data relating to UK assessment scenarios (e.g. characteristics of environment and focal species), which will be used as default but may be replaced by the user's own data for specific assessments. The software will also contain examples of pesticide-specific data (e.g. toxicity) for tutorial and demonstration purposes.

### **Characteristics of models and output**

The customers' requirement is for a web-enabled framework of models that incorporate variability and uncertainty into the assessment of risk to non-target species. The models will not necessarily be probabilistic if other approaches can be shown to be satisfactory. The models are required to be closely compatible with current regulatory approaches; this probably implies that outputs of all modules will include current assessment endpoints such as TERs, accompanied in some cases by novel endpoints. The customers specified that the models should address issues such as species sensitivity distributions, spatial, temporal and regional scales, population and community effects, the impact of food intake on exposure, and pesticide residues in food of non-targets. The precise approach may vary between sub-projects according to the issues for each type of organism, but are expected to be compatible with general principles established in WEBFRAM 1. It is also intended to take account of approaches considered or recommended in EUFRAM, and to seek EUFRAM feedback on WEBFRAM prototypes.

### 6.3.1.5 Summary Sheet – ICPS - SEEM-database

(Summary by D. Auteri)

#### **Name of the database, owner, identification**

The SEEM-db is developed by International Centre for Pesticide and Health Risk Prevention (ICPS) and was Sponsored by the European Commission, DG Health and Consumer Protection (Contract N° B1-3330/2001216). It is currently ready for internal use for ICPS and DG SANCO analysis in the field of ecotoxicity assessments of pesticides and their metabolites. Contact persons are [domenica.auteri@icps.it](mailto:domenica.auteri@icps.it) and [Wolf-Martin.Maier@cec.eu.int](mailto:Wolf-Martin.Maier@cec.eu.int).

#### **Background information**

The SEEM project, developed in 2002, was aimed at establish an ecotoxicological database of pesticides and their metabolites useful for the investigation of the statistical distribution and inter-association of certain end points of interest. The main objective of interest was to investigate the following relationships:

The relationships between acute and chronic toxicity of a parent substances or a metabolites to fish, daphnia, algae and earthworms.

The relationships among individual end points (acute vs acute, chronic vs chronic) of parent substances and their metabolites by pair-wise comparison, and the cases where a metabolite is of 10-fold higher toxicity than the parent substance.

The relationships between toxicity of a parent substance or a metabolite of daphnia and sediment dwellers (*Chironomous*).

#### **Characteristics regarding software**

The dBase was created using MS EXCEL spreadsheet. MS EXCEL was chosen because, according to the call for tender, records can be easily added afterwards, and data retrieval by a filter allowing multiple selection criteria is easy for any user. The extraction of sub-sets of data for further analyses is guaranteed. The dBase was splitted into three spreadsheets: one for the active ingredients, one for metabolites and one for the legend. The spreadsheets were organised in a matrix structure (rows corresponding to substances and columns to their attributes). The active ingredient (a.i.) and metabolite spreadsheets are linked each other by an id-code number that unequivocally allows retrieving all the metabolites generated by an active substance.

#### **Characteristics regarding data coverage**

A total number of 200 a.i. and 130 metabolites was included in the dBase. Data coverage concerns mostly aquatic ecotoxicity data. In particular, for each a.i and metabolite the following end point were included: Fish LC<sub>50</sub>, NOEC (acute), NOEC (chronic); Daphnids (and aquatic invertebrates) EC(LC)<sub>50</sub>, NOEC (acute), NOEC (chronic); Sediment dwellers NOEC (chronic); Algae and Aquatic plants EC<sub>50</sub>, NOEC (acute); Earthworms LC<sub>50</sub>, NOEC (chronic); Soil microorganisms (Effect % of N and C mineralisation); Octanol–water partitioning Log P; Dissociation (acidic) pK<sub>a</sub>; Chemical class; Use.

#### **Basics of the database**

The data model is flexible, so as to allow the selection of sub-sets of data of interest (i.e. by group of organisms, species, end point) for the statistical analysis. The lists of end points from the ECCO peer review process (downloaded from CIRCA, or provided by the ECCO team) and from the national review process (provided by Italian Ministry of Health) were selected for the project, as the two sources of available validated data. Literature was searched as a complementary source.

**Planned or expected further development (future versions)**

See RED project.

### **6.3.1.6 Summary Sheet: RIVM – e-toxBase**

(Summary by L. Posthuma)

#### **Name of the database, owner, identification**

The RIVM EcoTox-db is developed by RIVM, and is currently ready for internal use for RIVM projects in the field of ecotoxicity assessments of any type of toxic compounds. Contact persons are [d.de.zwart@rivm.nl](mailto:d.de.zwart@rivm.nl) and [l.posthuma@rivm.nl](mailto:l.posthuma@rivm.nl).

#### **Background information**

The development of the RIVM EcoTox-db was started in 2000, to provide an array of RIVM ecotoxicity-related projects with a tool to store, handle and retrieve data in a reproducible way. The project types at RIVM range from projects on the derivation of Environmental Quality Criteria to projects on site-specific risk assessments. The database now covers more than 30,000 entries, and is ready for internal use. Management decisions and technical developments being awaited, the EcoTox db may develop into a tool allowing for (limited?) web-enabled searching and retrieval of data.

#### **Characteristics regarding software**

In the phase of internal use, the RIVM EcoTox-db operates on the internal RIVM network. Its data model was developed in MS-Access, but transfer to Oracle has been recently started to fit into RIVM-standards. The user interface is now programmed in MS-Access. The user interface was designed to allow for easy and flexible querying and/or export of data, allowing for making appropriate sub-selections from the data set pertaining to the specific problem. The export function creates Excel spreadsheets, of the desired data types only. Each project can thus retrieve only the types of data needed.

#### **Characteristics regarding data coverage**

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 data cover both 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).

#### **Basics of the database**

The data model is highly flexible, so as to enable capturing developing insights in the data needed to execute Ecological Risk Assessments. The resources of the data are other open sources of compiled data, and sources within RIVM. Part of these resources are confidential, and the database allows for keeping these confidential.

#### **Characteristics regarding output**

Queried data are exported to Excel. The RIVM-EcoTox db will soon be formally adopted for internal use in RIVM-projects and collaborative projects.

#### **Planned or expected further development (future versions)**

Plans exist to extend on data coverage (e.g., extend on PPPs, extend on terrestrial and sediment data, et cetera), on web-enabled search facilities, and other activities needed to use the RIVM EcoTox-db as standard tool in collaborative projects and beyond.

### **6.3.1.7 Summary Sheet: RIVM – Assiwerk and RIVM - Enerwer2 databases**

(description by R. Luttik)

#### **NAME OF THE DATABASE, OWNER, IDENTIFICATION**

Assiwerk and Enerwer2

Robert Luttik

Two excel files one with data on assimilation efficiency for avian and mammalian species of certain types of food (260 values) and the other on the caloric content of different food types (700 values) and in some cases information about the water content of these food types (150 values).

For more information on the Assiwerk and Enerwer2 database contact:

Robert Luttik, PO Box 1, 3720 BA Bilthoven, The Netherlands, Telephone: + 31 30 2742795, Fax: + 31 30 2744401, E-mail: Robert.Luttik@RIVM.NL

#### **BACKGROUND INFORMATION**

The two excel sheets were prepared in 1994 and 1995 for the research project that resulted in the publication given below.

T. P. Traas, R. H. Jongbloed, and R. Luttik (1996) A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. I. Model description and uncertainty analysis for dichlorodiphenyltrichloroethane (DDT) and cadmium. *Ecotoxicology and Environmental Safety* 34, 264-278.

R. H. Jongbloed, T. P. Traas, and R. Luttik (1996) A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. II. Calculations for dichlorodiphenyltrichloroethane (DDT) and cadmium. *Ecotoxicology and Environmental Safety* 34, 279-306.

#### **CHARACTERISTICS REGARDING SOFTWARE**

It are just plain excel files

#### **CHARACTERISTICS REGARDING DATA COVERAGE**

Together with the research carried out by Joe Crocker of CSL (York) this will provide the state of the art for these important parameters for risk assessment for birds and mammals.

#### **BASICS OF THE DATABASE**

Excel

#### **CHARACTERISTICS REGARDING OUTPUT**

Not applicable

#### **PLANNED OR EXPECTED FURTHER DEVELOPMENT**

See above: data set of Joe Crocker should be combined with this data set.

### 6.3.1.8 Summary Sheet: US-EPA – Ecotox database

(<http://www.epa.gov/ecotox/>)

(Summarised kindly by Emilio Benfenati)

#### NAME OF THE DATABASE, OWNER, IDENTIFICATION

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. 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.

For more information on the ECOTOX database contact:

ECOTOX Support Staff

U.S. Environmental Protection Agency

Office of Research and Development

National Health and Environmental Effects Research Laboratory

Mid-Continent Ecology Division (MED)

6201 Congdon Boulevard

Duluth, Minnesota 55804

Telephone: 218-529-5225

Fax: 218-529-5003

E-mail: [ecotox.support@epa.gov](mailto:ecotox.support@epa.gov)

#### BACKGROUND INFORMATION

The development of the ECOTOX was started in 1995, and in March of 1996, it was released to governmental users through a 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. In the development of PHYTOTOX, the tests represent mostly agricultural chemicals and predominantly agricultural species.

#### CHARACTERISTICS REGARDING SOFTWARE

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 coding guidelines for the aquatic database (Aquatic Coding Guidelines) and the terrestrial database (Terrestrial Coding Guidelines). ECOTOX (ECOTOXicology Database System) is a comprehensive computer-based system that provides single chemical toxic effect data for aquatic life, terrestrial plants, and terrestrial wildlife.

## **CHARACTERISTICS REGARDING DATA COVERAGE**

Database searches can be conducted using either a Quick Query or an Advanced Query menu. The Quick Query supports searches on habitat, taxonomic kingdom, species common or Scientific name, Chemical Abstract Service Registry number, chemical name, observed effect group and publication year. The Advanced Query menu includes all options under Quick Query, and enables you to focus on more specific criteria such as study site type (e.g., laboratory, field), exposure media (e.g., freshwater, soil), route of chemical exposure (e.g., oral, diet), and statistically-derived endpoints (e.g., LD50, NOEL).

The bioconcentration factor (BCF) is a unitless value describing the degree to which a chemical can be concentrated in the tissues of an organism in the aquatic environment (View Endpoint Code List for full definition). A bioconcentration endpoint is coded as either wet (or unknown) or as dry weight (BCF and BCFD, respectively). An accumulation effect is coded. If the author does not calculate a BCF, the test is recorded as a residue measurement effect with a blank Endpoint and BCF field.

## **BASICS OF THE DATABASE**

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; ACQUIRE, TERRETOX, and PHYTOTOX. The ACQUIRE and TERRETOX databases contain information on lethal, sublethal and residue effects. The ACQUIRE database includes toxic effects data on all aquatic species including plants and animals and freshwater and saltwater species. TERRETOX is the terrestrial animal database. It's primary focus is wildlife species but the database does include information on domestic species. PHYTOTOX is a terrestrial plant database that includes lethal and sublethal toxic effects data.

The resources of the data come from several database:

ECOTOX (ACQUIRE, TERRETOX, PHYTOTOX)

Some independently compiled data sets have been transferred into ECOTOX from external sources. The data sets must meet the ECOTOX data parameter and quality assurance guidelines. Data sets available in ECOTOX are from Dutch dataset, EPA: Fathead Minnow Acute Toxicity database (MED-Duluth), French dataset, German dataset, EPA: Office of Pesticides Program database, Russian dataset, USGS Acute Toxicity database.

## **CHARACTERISTICS REGARDING OUTPUT**

Data obtained from compiled data files must meet the minimum data requirements and quality assurance guidelines defined for each ECOTOX database component. The key data fields that must be included are: test chemical name, test organism, test duration, effect, and effect concentration or application rate. Documentation describing the test methods must be provided within the publication. If tests are missing key parameters, the data are rejected. No effort is made to locate unreported data (e.g., authors are not contacted, citations referring to methods used are not obtained). During the incorporation of an electronic data file, a quality assurance check of the CAS number, species scientific name, and reference citation is completed. Search results can be downloaded either as an ASCII delimited file format, which can be transferred into a database or spreadsheet, or in a browser-viewable report format.

## **PLANNED OR EXPECTED FURTHER DEVELOPMENT**

Plans exist to extend on data coverage.

### **6.3.1.9 Summary Sheet: US-EPA – NTP database**

(Summarised by Emilio Benfenati)

#### **NAME OF THE DATABASE, OWNER, IDENTIFICATION**

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) coordinates toxicological testing programs within the Department of Health and Human Services (DHHS), strengthen the science base in toxicology; It develops and validates improved testing methods; and provides information about potentially toxic chemicals to health regulatory and research agencies, the scientific and medical communities, and the public.

For more information on the NTP database contact:

NTP Webmaster E-mail: [ntpwm@niehs.nih.gov](mailto:ntpwm@niehs.nih.gov)

Central Data Management (CDM)

P.O. Box 12233, MD EC-03

Research Triangle Park, NC 27709

Telephone: (919) 541-3419

E-mail: [cdm@niehs.nih.gov](mailto:cdm@niehs.nih.gov)

NTP Liaison and Scientific Review Office

P.O. Box 12233, MD A3-01

Research Triangle Park, NC 27709

Telephone: (919) 541-0530

E-mail: [liaison@starbase.niehs.nih.gov](mailto:liaison@starbase.niehs.nih.gov)

#### **BACKGROUND INFORMATION**

More than 80,000 chemicals are registered for use in commerce in the United States, and an estimated 2,000 new ones are introduced annually for use in everyday items such as foods, personal care products, prescription drugs, household cleaners, and lawn care products. The effects of many of these chemicals on human health are unknown, yet people and our environment may be exposed to them during the manufacture, distribution, use, and disposal or as pollutants in our air, water, or soil. Although relatively few chemicals are thought to pose a significant risk to human health, safeguarding public health depends on identifying the effects of these chemicals and the levels of exposure at which they may become hazardous to humans. The National Toxicology Program (NTP) was established in 1978 by the Department of Health and Human Services (DHHS) to coordinate toxicological testing programs within the Department. The NTP is an interagency program consisting of relevant toxicology activities of the National Institutes of Health's National Institute of Environmental Health Sciences (NIH/NIEHS), the Centers for Disease Control and Prevention's National Institute for Occupational Safety and Health (CDC/NIOSH), and the Food and Drug Administration's National Center for Toxicological Research (FDA/NCTR). The NIH's National Cancer Institute (NIH/NCI) was a charter agency; however, the NCI Carcinogenesis Bioassay Program was transferred to the NIEHS in 1981. The NCI remains active in the Program through membership on the NTP Executive Committee. NTP's mission is to evaluate agents of public health concern by developing and applying tools of modern toxicology and molecular biology.

#### **CHARACTERISTICS REGARDING SOFTWARE**

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

### **BASICS OF THE DATABASE**

The National Institute of Environmental Health (NIEHS) Sciences Division of Extramural Research and Training (DERT) in collaboration with the National Toxicology Program (NTP) has initiated a new and unique grants program which funds investigator-initiated research to provide data to aid in defining the mechanism of action of agents under study by the NTP. This new program uses the NIH R03 Small Grant mechanism to encourage investigator-initiated hypothesis-driven investigations that utilize animals/tissues/cells from animals undergoing the NTP 2-year cancer bioassay or shorter toxicological characterizations. The National Toxicology Program (NTP) conducts toxicity/carcinogenesis studies on agents suspected of posing hazards to human health. Chemical-related study information is submitted to NIEHS and is archived and maintained in a central location (Central Files) so that all study information can be monitored and tracked efficiently. Currently, more than 800 chemical studies are on file. NTP Information is routinely provided to industry and the public on an as requested basis.

### **CHARACTERISTICS REGARDING DATA COVERAGE**

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

### **PLANNED OR EXPECTED FURTHER DEVELOPMENT**

Plans exist to extend on data coverage and on web-enabled source facilities, moreover several annual meeting on specific subjects (Endocrine Disruptors, Phthalates, etc..) will be organized.

### **6.3.1.10 Summary Sheet: US-EPA – Pesticides Environmental Fate Database**

(Provided by H. Galicia)

#### **Name of the Database, owner, identification**

Pesticides Environmental Fate Database will be put in service by the US EPA in late spring or early summer 2003.

#### **Background information**

This is a federal effort. Namely by the Information Products Bulletin (IPB) which is a joint effort between the U.S. Environmental Protection Agency (EPA) and the Environmental Council of the States (ECOS) to inform stakeholders and the public about upcoming significant information products. There is a caution remark since the IPB is not intended to be a compilation of the information products that have already been completed and released to the public by EPA or the States. So this may mean that it has a less official character.

This is an environmental fate database containing studies describing what happens to pesticides in the environment after they are applied. Developed in FY 2000, this database contains fate and transport properties for 250 pesticides registered in the U.S.

I am afraid that there is no further details to convey, only that the contact information is the AAship: Office of Prevention, Pesticides, and Toxic Substances/Office of Pesticide Programs

- EPA Office/Organization: Office of Prevention, Pesticides, and Toxics/Office of Pesticide Programs

Harry Liu, 703-305-5372, liu.larry@epa.gov

#### **6.4 Available operational software: Summary Sheets**

The available operational software is summarised on the subsequent pages, each item one page of Summary Sheet in writing.

### **6.4.1.1 Summary Sheet: Alterra - PERPEST**

(Prepared by P.J. van den Brink)

#### **Name of the software, owner, identification**

The software package PERPEST (Prediction Ecological Risks PESTicides) is developed in the framework of research programme 359 of the Dutch Ministry of Agriculture, Nature Management and Fisheries. It will be released in 2003, and becomes freely available on the web ([www.perpest.alterra.nl](http://www.perpest.alterra.nl)). Contact person is [paul.vandenbrink@wur.nl](mailto:paul.vandenbrink@wur.nl).

#### **Background information**

Effects of pesticides on aquatic communities and ecosystems can be predicted using large simulation models like food-web models. Ecological models, however, are often incomplete and usually have many uncertain parameters, making their predictions uncertain too. In fact, experts are often better able to predict effects of toxicants compared to these models. PERPEST is a model that uses Case-Based Reasoning to predict the effects of a particular concentration of a pesticide on a defined aquatic ecosystem. It mimics the expert by basing its prediction on published information about the effects of pesticides on the structure and function of aquatic ecosystems as observed in semi-field experiments. The PERPEST model can be used in ecological risk assessment when uncertainties are large and data availability is poor. The PERPEST model can then provide some indication of the types of uncertainties that are likely to be large, and thus of the type of data that must be gathered for a refined risk assessment (e.g. endpoints and exposure concentrations of interest).

#### **Characteristics regarding software**

The PERPEST model 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. For each class, the bootstrapped 95% confidence limits are presented in a table. The model can be used in both predictive (e.g. effects of a new pesticide) as retrospective (e.g. measured concentrations of a pesticide in a river) risk assessment. The PERPEST output can be used to translate spatially and temporally distributed measured or modelled concentration data into effect concentrations, i.e. it can be used as a risk indicator.

#### **Characteristics regarding data coverage**

The database (called the case base) consist of two data sets, one containing the results of the review of the effects of pesticides observed in semi-field experiments and one containing chemical properties and ecotoxicological profiles, based on the results of standard laboratory tests of insecticides and herbicides. The first data set comprises case studies in which the effect of a particular concentration of a particular pesticide is evaluated using freshwater microcosms or mesocosms. The endpoints evaluated were classified into 8 different ecological endpoint groups, which were different for insecticides and herbicides. Within each of the eight ecological endpoint groups, the most sensitive endpoint was selected for assignment to an effect class. The responses observed for these groups were assigned to 0 or to the five effect classes, ranging from no effects demonstrated to clear long-lasting effects

A total of 90 experiments were evaluated (113 references published between 1980 and 1998). This evaluation resulted in 333 cases (substance and concentration combination), with 171 evaluating the effects of herbicides and 162 those of insecticides.

#### **Basics of the software**

PERPEST is programmed in Delphi but can be installed without Delphi.

### **Characteristics regarding output**

Output is both presented in Figures (Pie charts) as tables which can be exported to other programmes like Excel.

### **Planned or expected further development (future versions)**

An extension of PERPEST to the ecological risks of Fungicides is planned for the coming two years.

## 6.4.1.2 Summary Sheet: IME - Simulation models

(Description by U. Hommen)

### Name of the software, owner, identification

The programs IME GPS, Daphnia, Zebra, SAM, and FLEXSWAC provide ecological models developed for the analysis and prediction of the effects of toxicants on populations and communities. Because program structure and user interface are very similar for all these models, they are described in one summary sheet here. Contact: [hommen@ime.fraunhofer.de](mailto:hommen@ime.fraunhofer.de).

### Background information

The five programs cover ecological models of different complexity:

**GPS** (General Population Simulator) uses the simple logistic growth model to estimate effects of and recovery after application of a toxicant.

**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 original model was developed by Dülmer (1998) and later extended by U. Hommen.

Another Individual Based Model is **Zebra**, which simulates a small laboratory population of the zebrafish *Danio rerio*. The model includes stochastic descriptions of reproduction and foraging behaviour. As for the Daphnia model, the objective is to extrapolate from effects in individuals (e. g. reduced egg production) to the level of the population (e. g. extinction risk).

**SAM** (Simulation of an Aquatic Microcosm) models a simple planktonic food web, comprising of phytoplankton, cladocerans, copepods and rotifers. It has been developed to support interpretation of mode ecosystem studies, but can also be used to estimate potential indirect effects of a toxicant. Uncertainty about model parameters can be taken into account by Monte-Carlo simulations.

**FLEXSWAC** is a Windows version of the SWACOM, the Standard Water Column Model, developed by O'Neill et al. (1982) and described in detail in the book of Bartell et al. (1992). It is a generic model of a complex food web in a lake (19 populations in 4 trophic levels).

### Characteristics regarding software

All models have been implemented in a Delphi shell offering different (context specific) types of simulation experiments: single (deterministic) runs, response analysis, calibration, Monte-Carlo-simulations

### Characteristics regarding data coverage

Due to the different complexity of the models, different amounts and types of data are needed. For the population dynamics, default parameter sets are given, except for the GPS where the intrinsic growth rate and the capacity of a population has to be entered. Effects are described by LC50 and/or EC50 for reproduction, while the exposure module needs magnitude and timing of inputs and half times for the assumed exponential decrease. Input data are read from Excel sheets and can be modified with the program.

### Basics of the software

The implemented models cover a range of population and food web models linked to a simple exposure model. Exposure is modelled as an exponential decrease of the toxicant concentration in the water body after one or several applications respectively assumed inputs (e.g. run-off events). Uncertainty about magnitude and timing of the inputs as well as the half-life can be considered via Monte-Carlo simulations. Alternatively, concentration time series can be read from input files. Population dynamics is linked to exposure via dose response functions for mortality or reproduction (described by LC50 or EC50 and slopes). Population dynamics is described using the logistic growth model (GPS), detailed Individual Based Models (Daphnia, Zebra), or differential equation systems using parameters such as maximum consumption rate, respiration rate, assimilation efficient, temperature optima etc. (SAM, FLEXSWAC).

### **Characteristics regarding output**

The programs provide time series of forcing functions (e. g. the toxicant concentration) and state variables (e.g. population densities), and summary statistics for Monte-Carlo-Simulations (e. g. mean population densities of the MC-runs over time, minimum abundance per simulation, extinction risk for different treatment levels). . The tables are provided as Excel compatible sheets. Diagrams can also be copied and pasted, as well as exported into bitmaps.

### **Planned or expected further development (future versions)**

Documentation and help files have to be improved. Some documentation is only available in German yet.

### 6.4.1.3 Summary Sheet: IME SSD calculators

(Description by U. Hommen)

#### **Name of the software, owner, identification**

IME SSD-xls and SSD-delphi were developed by U. Hommen and are used for IME projects in the field of risk assessments of pesticides or other toxic compounds. Contact: [hommen@ime.fraunhofer.de](mailto:hommen@ime.fraunhofer.de).

#### **Background information**

The software can fit species sensitivity distributions to a collection of similar toxicity data (e.g. NOECs, LC50s, or EC50 from different species). The Delphi program includes a variety of SSD approaches: the different 'factor models' (e.g. van Straalen & Denneman, Lokke & Wagner, Aldenberg & Jaworska), linear regression models using logit and probit transformations, and Bootstrap methods (Newman). The Excel spreadsheet focusses on the recent method of Aldenberg & Jaworska (2000) and simple linear regression fits of cumulative probit and logit distributions. It is mostly used to create different types of SSD based diagrams.

#### **Characteristics regarding software**

Both versions are limited to a single set of toxicity data ( $n \leq 30$  in the moment), which can be directly entered or pasted from other Windows spreadsheets.

#### **Characteristics regarding data coverage**

In addition to the toxicity data, the programs include factors given by different authors to estimate SSDs from the mean and the standard deviation of the log-transformed toxicity data.

#### **Basics of the software**

The software fits different distribution functions (SSDs) to a set of data. From the SSD it estimates the so called 'hazardous concentrations' for a given % of species (e.g. the HC5) or the fraction of species affected at a given concentration.

#### **Characteristics regarding output**

The software calculates (depending on the model used) goodness of fit ( $R^2$  or  $\chi^2$ ), basic statistics, the 5th and the 50th percentile (resp. HC5 and HC50) and plots of cumulative distribution functions, linear fits of the transformed data, % of protected species versus PEC, or threshold concentration versus % protected species.

#### **Planned or expected further development (future versions)**

The program and the Excel sheet have only used by U. Hommen to calculate HC5 and create SSD diagrams. Therefore, documentation is poor. It seems questionable, if it is worth to spend time into improvements if RIVM ETX is more refined and will be freely.

### **6.4.1.4 Summary Sheet: RIVM – ETX**

(Description by T.P. Traas)

#### **Name of the software, owner, identification**

ETX was developed by RIVM, and is used for RIVM projects in the field of risk assessment and standard setting of toxic compounds. Contact persons are [tp.traas@rivm.nl](mailto:tp.traas@rivm.nl) and [tom.aldenberg@rivm.nl](mailto:tom.aldenberg@rivm.nl).

#### **Background information**

The ETX software can fit species sensitivity distributions to a collection of similar toxicity data. Usually, NOECs, LC50s or LD50s from different species are combined in such a distribution, called a species sensitivity distribution (SSD). The software was developed in 1993 (ETX 1.3a) to fit normal, logistic and triangular distributions to a set of log-transformed toxicity data. The normal distribution theory was updated in 2000 (ETX-2000).

#### **Characteristics regarding software**

The software is implemented in a custom format (ETX1.3a) or in EXCEL (ETX-2000) with a few added menu items. Data input is limited to single sets of toxicity data ( $n < 200$ ) and environmental exposure (point estimates). Statistical tables and goodness of fit procedures are implemented but hidden from view and protected against accidental editing.

#### **Characteristics regarding data coverage**

The software does not incorporate data except statistical tables. Toxicity data can be typed or pasted in from any data source with normal ms-windows copying modes.

#### **Basics of the software**

The software fits different distribution functions (SSDs) to a set of data. From the SSD, we can either estimate percentile values on the concentration axis (e.g. the 5th percentile, the HC5), or estimate the fraction of species exposed above the relevant toxicity endpoint. This fraction is called the 'potentially affected fraction', PAF. Each quantity estimated with the software is given with a confidence interval, based on either logistic (Aldenberg & Slob, 1993) or normal statistics (Aldenberg & Jaworska, 2000).

#### **Characteristics regarding output**

The software calculates Goodness- of fit, basic statistics, the 5th and the 50th percentile (resp. HC5 and HC50) and plots a histogram and a cumulative distribution function of the SSD and the data (figures only in ETX-2000).

#### **Planned or expected further development (future versions)**

Plans exist to implement SSDs in the new EUSES software (<http://ecb.jrc.it/new-chemicals/>), to develop a full stand-alone version of ETX and to make a web-enable version in the WEBFRAM program, funded by DEFRA (UK).

### **6.4.1.5 Summary Sheet: RIVM – Busy (Bayesian Uncertainty System)**

(Prepared by T. Aldenberg)

#### **Name of the software, owner, identification**

The software package Busy (Bayesian Uncertainty System) is developed in the framework of the CEFIC/ACC Long Range Research Initiative Project: Uncertainty Analysis in Ecological Risk Assessment. It will be released in 2003, and becomes freely available to CEFIC members, as well as others interested. Contact person is <mailto:Tom.Aldenberg@rivm.nl>.

#### **Background information**

The research objectives are the development of software tools to make probabilistic uncertainty analysis available to both beginning and expert practitioners of environmental risk assessment of chemicals through easy to use computer programs; to conduct a detailed and critical review of uncertainty concepts and methods, as well as current software for probabilistic uncertainty analysis; to document guiding principles and instructions for practical usage, and to work out applications in the framework of EUSES, a system for the evaluation of new and existing chemicals in the environment in the EU.

#### **Characteristics regarding software**

The first release (early 2003) will focus on three applications:

Probabilistic Ecological Risk Assessment

Bayesian Normal Species Sensitivity Distributions (SSDs) and two-D Monte Carlo analysis

Bayesian Logistic Dose-Response models.

The second release (late 2003) will implement a number of uncertainty analysis techniques as in current usage. 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.

#### **Characteristics regarding data coverage**

Probabilistic Ecological Risk Assessment deals mainly with the comparison of exposure and toxicity data through fitting statistical distributions and calculating Ecological Risk. The exposure data may come from the field, and the toxicity data from the lab. Busy provides a flexible and easy to use set of probability distributions to carry out the assessment. 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.

#### **Basics of the software**

Busy is designed from the outset as an object-oriented system. This means that models and data structures are arranged in a hierarchical tree allowing for the inheritance of methods down the tree. Functions (procedures) having the same name, adapt to the object receiving the call. This provides a very clear and intuitive textual user interface. On top of this Templates (MS-Excel sheets) are pre-programmed Busy scripts that accomplish some task, e.g. SSD fitting, or Risk Assessment.

**Characteristics regarding output**

Busy runs in Mathematica Notebooks, either independently or coupled to Excel and provides both numerical and graphical output (distributions, uncertainty plots), which can be saved to files.

**Planned or expected further development (future versions)**

Busy aims at an easy to use system for carrying out Bayesian statistics, by non-statisticians. It should be able to run the basic models as outlined in recent Bayesian textbook: Normal inference, ANOVA, regression, Logistic inference, and so on.

### **6.4.1.6 Summary Sheet: RIVM – PRA Secondary Poisoning**

(prepared by T.P. Traas)

#### **Name of the software, owner, identification**

The PRA model for secondary poisoning was developed by RIVM, and was used for RIVM projects in the field of risk assessment of metals. Contact person is tp.traas@rivm.nl.

#### **Background information**

In 1996 Traas et al. described a probabilistic model for calculating bioaccumulation risks for terrestrial foodwebs of predators. A species-specific approach was followed: four birds of prey (Sparrow Hawk, Kestrel, Barn owl and Little owl) and two mammalian predators (Badger and Weasel) were modelled. The model was applied to DDT and Cadmium. Model structure can be easily adapted to accommodate different types of food webs.

#### **Characteristics regarding software**

The software is implemented in Mathcad.

#### **Characteristics regarding data coverage**

The software does not incorporate data. (Parameters of) Distributions need to be specified for bioaccumulation factors (BAFs) for different food items. Different toxicants have different accumulation factors and uncertainty of the factors is calculated from empirical BAFs.

#### **Basics of the software**

Data input are distributions of bioaccumulation factors for food items of specific predators.

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.

#### **Characteristics regarding output**

Data output is the expected distribution of food concentrations in predators, and the probability that critical intake levels are exceeded. By inverting the model, the distribution of corresponding maximum permissible soil concentrations can be calculated.

#### **Planned or expected further development (future versions)**

Plans exist to rewrite the model in EXCEL with an add-on for Monte Carlo analysis such as Crystal Ball.

## 6.5 Overview of missing databases or software items for PRA

This section of the final report will summarise items that are not summarised in Summary Sheets or Fact Sheets, but that may remain valuable for EUFRAM during the project period, e.g., when new insight develop in other work packages. Table 7 and Table 8 list these items. In the current report version the tables are empty.

Coverage	Owner	Name	Code
Fate			
Exposure			
Effects			
Effects and metabolites			

Table 7. Potential databases that have not been summarised in Summary Sheets or Fact Sheets, but that may be relevant to be evaluated in the context of the *EUFRAM* project.

Coverage	Owner	Name	Code
Fate			
Exposure			
Effects			
Effects and metabolites			

Table 8. Potential software items that have not been summarised in Summary Sheets or Fact Sheets, but that may be relevant to be evaluated in the context of the *EUFRAM* project.

## 7 EVALUATION: DATABASES AND SOFTWARE CONFRONTED WITH CRITERIA

### 7.1 Approach for databases and software

The available software items and databases that are eventually compiled in Chapter 6 are eventually evaluated by the criteria proposed in Chapter 5, according to a two-step approach. The first approach is a quick-scan. By compiling the eventual series of fact sheets into a single table (horizontal: products, vertical: criteria) it will be possible to make a raw + and 0 and – score on presence/absence of features, and this allows to discriminate between the items with most deficiencies or strengths. In the second step, a SWOT (Strengths, Weaknesses, Opportunities, Threats) approach can be executed using the criteria in more detail (not: *is* a criteria fulfilled, but *how*). Evidently, for some of the items that wouldn't pass the quick scan due to technical weaknesses, the scientific content may be worthwhile to keep. For example: a set of unique data in a 'weak technical shell' should not be

discarded. These exercises will be undertaken at the workshop. Results of this process will be summarised in this Chapter.

## 7.2 Evaluation of Databases

As a preliminary example, and assuming that the quick-scan step has been passed, the RIVM – e-toxBase is confronted with the criteria of Table 3, and the preliminary SWOT-results are shown in Table 9.

### Strengths

Raw data = data entry  
Quality earmark  
Effect  
Allowing for versatile data query / sub-selections  
'Feeding' of new data  
Exporting of output  
Exporting of output selections only  
Printing of output  
Screen interface  
Windows-based

### Weaknesses

Fate: data not yet entered  
Exposure: data not yet entered  
Wide coverage of existing data is as yet absent  
Published: only internal document  
Linkage to software for PRA: not needed: via Excel  
Broadly dispersed and known: no  
Free availability: not arranged, might become partly possible only  
Web-enabled use possible: no, might become possible  
Suitability for EUFRAM targets

### Opportunities

Define EUFRAM scope and quality scale  
EUFRAM-queries can likely be met  
Mass import from other data sets possible  
Now from Access to Excel, develop technique for web  
Now from Access to Excel, develop technique for web  
Develop web-querying facilities  
Fate data can be adopted  
Exposure data can be adopted  
Data model does not pose restrictions to data types

### Threads

Sources of mass data have use limitations  
Data coverage is unbalanced (more aquatic than terrestrial)  
Web-querying pending (procedural /technical)

Table 9. Preliminary SWOT analysis (illustration of a possible outcome) of one the databases to be evaluated for *EUFRAM*.

As an overall conclusion on this example, this database has advantages and disadvantages according to the *list* of criteria of Chapter 5. After that the criteria have been weighted, the databases can be sorted as to potential profitability to *EUFRAM* targets.

Results of the evaluation process will be added to this report after the workshop.

## 7.3 Evaluation of Software

The text in this section will be generated according to the outline of the previous section, and will be presented after the workshop.

# 8 PRELIMINARY CONCLUSIONS AND RECOMMENDATIONS

## 8.1 Preliminary conclusions

Preliminary conclusions will be formulated on the separate issues that have been treated in this report, that is: on criteria, on their ranking, on the available

software and database items and on their evaluation, including an analysis of what is lacking. The text for this section will be generated after the workshop.

### **8.1.1 Detailed conclusions on databases**

- what *do* we have
- what is *missing*
- what could especially be developed to reach the *EUFRAM* targets
- which databases do we recommend for end-user testing phase of *EUFRAM*

### **8.1.2 Detailed conclusions on software**

- what *do* we have
- what is *missing*
- what could especially be developed to reach the *EUFRAM* targets
- which software items do we recommend for end-user testing phase of *EUFRAM*

### **8.1.3 Overall conclusions**

Addressing the eventual target of WP10 of EUFRAM: in how far does this report contribute to the overall EUFRAM targets? In this section, an answer will be provided to the main targets, which are (copied from the *EUFRAM* task lists):

*“Detailed recommendations for a common methodology for selecting databases and software for probabilistic pesticide risk assessments, and evaluate examples currently in use by partners for possible adoption as harmonised standards in the EU.”*

## **8.2 Issues to be discussed at Project Meeting 1**

Issues to be discussed at Project Meeting 1 are:

- does the general approach suit to the needs of *EUFRAM*
- are there major issues of concern generated by other work packages that should be taken into account
- if so, which issues of concern are there specifically?
- which software and database items are lacking and which need to be handled *per se*, e.g., in consideration of the case study workshops and the end-user testing phase?

## **8.3 Issues to be discussed at Software and Database evaluation workshop**

Issues to be discussed at the workshop are, or in the further shaping of the workshop documentation:

- is the list of *a priori* generated criteria correctly worded with respect to the idea of them being 'objective'
- is the list complete?
- is the list operational for using in the quick-scan and the SWOT-analysis, that is: would each section of the SWOT profit from the criteria list to generate remarks on Strength, Weakness, Opportunity, Threat?
- are the lists of databases and software items sufficiently complete?
- do the general conclusions generated for the workshop documentation hold, or should they be amended?
- Can we recommend specific items for end-user testing?

These issues are captured in the draft programme.

#### **8.4 Recommendations**

This section will address the specific actions to be undertaken or considered regarding implementation of the findings of this evaluation process within the framework of the *EUFRAM* project, especially with respect to implementation possibilities.

## References

- Hart, A. (ed.) 2001. Probabilistic risk assessment for pesticides in Europe: implementation and research needs. Report of the EUPRA workshop. Central Science Laboratory, UK. Report can be downloaded free of charge, from: <http://www.eupra.com>.
- EEC (2001) DG SANCO Contract N° B1-3330/2001216, Statistical Evaluation of available Ecotoxicology data on plant protection products and their Metabolites.
- EEC (2002) DG SANCO, Restricted Call for Tender SANCO/2002/E1/RED.
- Pastorok, R.A., Bartell, S.M., Ferson, S., and Ginzburg L.R. (2001) Ecological modeling in risk assessment. Chemical effects on populations, ecosystems and landscapes. CRC-Press, Boca Raton, FL, USA.