

# EUFRAM

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

Work Package 5

## PRELIMINARY PAPER ON PROBABILISTIC APPROACHES FOR TYPICAL DATASETS<sup>2</sup>

June 2003

Andy Hart (CSL)<sup>3</sup>, Colin Brown, CU), Pamela Byrne (PCS-DAFRD), Niels Elmegaard (DMU), Scott Ferson (AB), Hector Galicia (SL-SRC), Joanna Jaworska (P&G), Robert Luttik (RIVM), Keith Solomon (CTUoG)

### 1 CONTENTS

1	Contents .....	1
2	Objectives .....	2
3	Background.....	2
4	What are typical data? .....	3
5	What challenges do typical data present? .....	6
6	What approaches should we consider? .....	8
7	Actions for next phase of EUFRAM.....	12
8	References .....	13

---

<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 OBJECTIVES

This work package aims to progress one of the main recommendations of the EUPRA workshop, that research should be conducted to develop and evaluate probabilistic approaches that can be used for typical datasets<sup>4</sup>. The specific objectives of the work package are:

1. To identify the challenges presented by the types and quantities of data typically available for routine pesticide assessments in Europe
2. To identify probabilistic approaches that may be applicable to those datasets
3. To identify the potential advantages and disadvantages of each approach, and any other obstacles to their adoption
4. To identify possible ways to overcome the potential disadvantages and obstacles,
5. To recommend research actions required to evaluate and improve the most promising approaches
6. To facilitate coordinated implementation of those research actions by EUFRAM partners and others<sup>5</sup>,
7. To collate research results and integrate them into guidance for end-users.

Objectives 1-5 are addressed in this paper, which is based on a draft version developed by the WP members together with discussions with other project members at the EUFRAM project meeting on 12-14 May 2003. Objectives 6 and 7 will be pursued throughout the remainder of the project, and especially between May 2003 and June 2004, when the first draft of guidance is due.

## 3 BACKGROUND

There is no generally-accepted definition of the term “probabilistic risk assessment”, and it has been used to describe a variety of approaches. Broadly speaking, it refers to assessments that use probability distributions to quantify uncertainty and/or variability in the estimation of risk.

The EUPRA workshop (Hart, 2001) recognised a widespread perception that probabilistic methods are “data-hungry”. However, it also recognised that it is difficult to make a fair comparison between probabilistic and deterministic methods. Most existing examples of probabilistic approaches have increased the complexity of the assessment model (e.g. by incorporating landscape features or non-standard routes of exposure), or used non-standard types or quantities of data (e.g. toxicity studies for large numbers of species, or radio-tracking studies of birds), as well as treating it

---

<sup>4</sup> EUPRA recommendation 8: “research should be conducted to develop and evaluate probabilistic approaches that can be used when data are limited”. The word “limited” refers to the common perception that typical datasets provide insufficient information for probabilistic assessments.

<sup>5</sup> EUFRAM does not provide funds for research, although it does provide some funds for “enrichment” of case studies and evaluation of software. Rather the intent is that ongoing activities of partners will be “concerted”. This could include forming links between existing activities, adapting those activities if possible to address the needs more closely, and collaborating to initiate new activities.

probabilistically. This has created the impression that complex models and large datasets are an intrinsic requirement of probabilistic approaches.

Simple (1-dimensional) Monte Carlo, one of the more widely-known methods for probabilistic modelling, does require a lot of information, including precise knowledge of the shape, mean and variance of the distribution for each input variable, and the dependencies between input variables. However, some probabilistic methods are designed specifically for situations where data are very limited, and/or to enable the incorporation of expert judgement (Ferson, 2002). Advocates for these methods argue that it is precisely when data are limited, and uncertainty is highest, that these methods are most necessary. Methods for limited datasets have the potential to provide practical improvements more rapidly and for more pesticides than methods that require extensive or precise supporting information. However, the appropriateness of these methods for regulatory assessment of pesticides has not yet been evaluated. The EUPRA workshop therefore concluded that research should be conducted to develop and evaluate probabilistic approaches that can be used with the quantities of data typically available for pesticide assessments.

#### **4 WHAT ARE TYPICAL DATA?**

EUFRAM is primarily concerned with pesticide assessments conducted in the EU under Directive 91/414/EEC.

The types of data required for pesticide risk assessments are defined in the annexes to 91/414/EEC, and in guidance documents published by the European Commission. From the perspective of companies and regulators, the quantity of data that are generated and evaluated for a typical new pesticide could be regarded as substantial, and also costly. However, for the purpose of estimating distributions for probabilistic assessments, they are very limited. To estimate a distribution, a statistician would usually like to have tens of individual data points. But for a typical pesticide, only 1-2 studies are available for each type of data (or for some types of data, 3 or 4).

The numbers of ecotoxicity studies typically submitted in the EU for first tier assessments is indicated in Table 1. In most cases, multiple studies of the same type are for different species. Table 1 includes mainly studies with technical-grade active substance, although studies with formulations are also included in cases where they were submitted as part of the first tier data package. Table 1 does not include additional studies submitted for refined assessments, nor does it include literature studies that sometimes are cited in submissions for “existing” (previously registered) pesticides. Overall, the numbers in the table slightly over-estimate the numbers of studies typically submitted for new active substances. For existing pesticides, the total number of studies available will often be significantly greater, but only for those endpoints that required refined assessment.

**Table 1:** Number of ecotoxicity studies submitted per pesticide for first tier assessments, averaged over 7 new and 9 existing active substances<sup>6</sup>. Does not include additional studies submitted for refined assessments.

Type of data	Average number of studies submitted
<b>Birds</b>	
Acute oral toxicity	1.8
Short-term dietary toxicity	2.4
Sub-chronic toxicity and reproduction	1.9
<b>Aquatic life</b>	
Acute toxicity to fish	4.8
Chronic toxicity test on juvenile fish	0.5
Fish early life stage toxicity test	0.9
Fish life cycle test	0.06
Bioconcentration in fish	0.8
Acute toxicity to aquatic invertebrates	3.8
Chronic toxicity to aquatic invertebrates	1.4
Effects on algal growth	3.2
Effects on sediment dwelling organisms	0.9
Aquatic plants	0.4
<b>Non-target arthropods</b>	
<b>Bees</b>	
Acute toxicity (oral/contact)	1.8
Other arthropods (glass plate studies, excluding extended lab studies)	1.6
<b>Earthworms</b>	
Acute toxicity	2.6
Sublethal effects	1.1
<b>Other</b>	
Effects on soil non-target micro-organisms	2.8
Effects on biological methods for sewage treatments	0.9

**Table 2:** Types of studies typically submitted for environmental fate assessments<sup>7</sup>.

<b>Route of degradation:</b>
Aerobic: Mineralisation, non-extractable residues, relevant metabolites, etc.
Supplemental: Anaerobic degradation and soil photolysis.
<b>Rate of degradation:</b>
DT50 in laboratory and under field circumstances (sometimes information on soil accumulation and plateau concentrations),
DT50 for anaerobic conditions and soil photolysis.
<b>Soil adsorption/desorption:</b>
Information on the $K_f$ , $K_{oc}$ and $K_{s/l}$ .
<b>Mobility in soil:</b>
Column leaching, aged residues leaching, lysimeter, field studies.
<b>Route and rate of degradation in water:</b>
Hydrolysis (DT50), Photolytic degradation (DT50), Degradation in water/sediment (DT50).
<b>Information on residues in crops:</b>
Can be used for calculating a DT50. This study is very often available but hardly ever used in the environmental risk assessment.

<sup>6</sup> These active substances are currently awaiting or under evaluation in the UK. There were two herbicides, one plant growth regulator, one nematicide, two molluscicides, seven insecticides and three fungicides. Uses included two indoor-only products, one granule, two pellets, two seed treatments and twelve sprays. Data provided by Mark Clook, PSD (pers. comm.).

<sup>7</sup> List provided by Robert Luttkik, RIVM (pers. comm.).

In addition to the pesticide-specific data listed in Tables 1 and 2, assessments of exposure use a number of general parameters. For example, the first-tier assessment of bird exposure requires estimates of body weight, daily food consumption, and the concentration of pesticide in food. An official guidance document provides standard point estimates for these parameters (European Commission, 2002b). Body weight is an arbitrary figure that is considered broadly representative of a relevant type of bird (e.g. 10g for a small insectivorous species). Daily food consumption is an average value estimated from allometric equations relating energy requirement to body weight, together with general estimates of the energy content of relevant food types and its assimilation efficiency in birds. Generic estimates are provided for initial concentrations of pesticides on various types of food. For plant materials and large insects, these estimates are approximate 90<sup>th</sup> percentiles of empirical data pooled for a number of different pesticides. Due to lack of direct measurements, concentrations on small insects are estimated by using, as a surrogate, data for plant material with a similar surface area-to-volume ratio.

Similarly, the initial exposure assessment for aquatic organisms (European Commission, 2002a) estimates worst-case concentrations in water for a standard scenario comprising a static ditch of 30 cm water depth, and a 5 cm deep sediment layer with organic carbon content of 5% and bulk density 0.8 kg/l. The application rate is assumed to be worst case, e.g. the maximum season's usage applied as a single dose. Spray drift into the ditch is estimated by the 90<sup>th</sup> percentile of measurements made 1m from the edge of the crop in a generic set of field experiments. Run-off/erosion and/or drainflow are included as a single fixed loading of 15% of the application rate that occurs on the day of application.

Other examples of limitations in existing data mentioned by Work Package members include the lack of measured concentrations in most types of toxicity studies, lack of uniformity in toxicity study designs for non-target arthropods, the uncertainties affecting determination of no-effect levels, and variation, uncertainty and bias in fate parameters. Perhaps greatest are the uncertainties that arise in extrapolating from assessments based on laboratory studies to the processes of impact and recovery in real ecosystems, as these involve major biological/ecological processes that are completely omitted from the basic assessment model.

It can be seen from these examples and Tables 1 and 2 that the data available for typical assessments are limited in several senses:

- some parameters use specific measurements for the pesticide under assessment, but the sample sizes are typically small (1-5),
- some parameters use generic estimates based on pooled data for other pesticides and/or contexts – a form of extrapolation or surrogacy,
- both the specific and generic estimates are derived from particular experimental conditions (e.g. laboratory toxicity studies), which are assumed to be relevant to field conditions – again a form of extrapolation
- some parameters use generic assumptions based on expert knowledge
- many factors and processes that may influence risk are ignored.

In addition, measurements of toxicity and other parameters are subject to limitations in their accuracy and precision (measurement uncertainty).

## 5 WHAT CHALLENGES DO TYPICAL DATA PRESENT?

### 5.1 Difficulties quantifying variability and uncertainty

Probabilistic approaches involve quantifying variability and uncertainty by replacing fixed point estimates with distributions. This is made difficult by the data limitations of typical datasets, as identified in the preceding section.

- *Small sample sizes* imply increased uncertainty in the choice of distribution and in estimating distribution parameters such as the mean and variance. For example, it makes it difficult to decide whether a species sensitivity distribution is lognormal. Even if the distribution shape for an SSD can be assumed with confidence, a small sample implies large confidence limits on the mean and variance.
- *Extrapolation from surrogate data* is likely to involve more uncertainty than using direct measurements. Uncertainty may be quantifiable for extrapolations that have been calibrated empirically (e.g. the relationship between body weight and energy expenditure). Expert judgement may be the only basis available for quantifying uncertainty in uncalibrated extrapolations (e.g. the relationship between toxicity measured in laboratory tests and toxicity to the same species in the field).
- *Measurement uncertainty* is well characterised for some types of study (e.g. analytical chemistry) but less so for others (e.g. between-laboratory variation in toxicity studies).
- *Parameter estimates based on expert judgement* introduce subjective uncertainty about the relationship between the chosen values and the distributions of values that occur in the field. In some cases (e.g. body weights of birds), ample data are available to replace arbitrary estimates with distributions.
- *Omission of factors or processes that influence risk* imply model uncertainty – uncertainty arising from structural deficiencies in the model. Even those factors and processes that are included in the model are unlikely to be represented perfectly. It may be possible to quantify model uncertainty to some extent by developing alternative models, which include and omit the uncertain elements. However, there is always likely to be some remaining uncertainty about the model structure, which could only be characterised subjectively.

The severity of these technical difficulties is illustrated by the controversies surrounding the derivation and use of species sensitivity distributions (e.g. Forbes & Calow, 2002), e.g. the number of species required for an SSD, bias in the selection of test species, and uncertainty introduced by using SSDs as an index of impacts at community level.

It was suggested at the EUFRAM project meeting (May 2003) that it may not be necessary to account explicitly for all the uncertainties mentioned above, because they are in effect “covered” by conservatism in other parts of the assessment (e.g. extreme worst case scenarios). However, it is difficult to judge whether the level of “cover” is sufficient without quantifying the uncertainties. Furthermore, a common aim of probabilistic assessments is to make assessment scenarios less extreme, so to maintain an appropriate degree of “cover” the other uncertainties should also be quantified.

## 5.2 Wide confidence bounds on risk estimates

If the uncertainties associated with typical datasets were quantified, the resulting risk estimates may have very wide confidence bounds.

It has been suggested that risk estimates with wide confidence bounds would be useless for decision-making. However, this contradicts the general principle that risk assessors should be as explicit as possible about scientific uncertainty, so that decision-makers can take account of it (e.g. Codex 2002, European Commission 2003). In other words, decision-makers need to know how uncertain the risk assessment is, even if it is very uncertain.

It has also been suggested that, if probabilistic risk estimates would be very uncertain, then it would be preferable to resort to deterministic estimates with safety factors. The difficulty with this is that the level of protection provided by existing safety factors (e.g. TER thresholds) is uncertain. In some cases, the degree of protection may be assessed by comparing predicted risks with effects observed in semi-field or field studies (e.g. aquatic mesocosms, see preliminary paper from EUFRAM Work Package 9). In other cases, it would be hard to determine how large safety factors should be without doing a probabilistic analysis (SCP, 2002). If existing safety factors are conservative, as is widely thought, then probabilistic analysis should confirm this.

If a probabilistic assessment does produce bounds that appear unreasonably wide, this suggests that we have additional knowledge that we are using implicitly to make judgements about the bounds. If so, we should try to incorporate that additional knowledge in the probabilistic assessment so that its contribution to reducing uncertainty is properly accounted for.

In summary, if it is accepted that risk assessors should be explicit about uncertainty, then the key issue for EUFRAM is not whether wide confidence bounds are useful, rather it is whether probabilistic methods can estimate them reliably.

## 5.3 Credibility of assessments based on typical data

It is important to emphasise that the challenges presented by typical datasets are not purely technical. They also present a major challenge to the *credibility* of a probabilistic risk assessment with scientists, decision-makers and stakeholders.

People place more confidence in a deterministic assessment that uses lots of data than one that uses few data. In effect, they are using quantity of data as a simple measure of the uncertainty of the outcome. This is very reasonable, provided the data have been properly used. It is also the only practical option, because a deterministic assessment does not provide any other measure of uncertainty.

There is a natural tendency to react to a probabilistic assessment in the same way. However, a complete probabilistic assessment should quantify uncertainty due to data limitations, for example by placing confidence bounds on risk estimates. Those bounds should then be used as the measure of uncertainty. If the bounds are robust, then the quantity of data becomes irrelevant, because it is already accounted for in the bounds. Obviously this only holds if the bounds properly reflect the uncertainty due to lack of data. Therefore, a high priority should be attached to determining whether probabilistic methods can account for uncertainty due to limited data. This applies to all the types of data limitations listed above (small sample sizes; use of

surrogate data, extrapolation or expert judgement; and omission of key factors or processes), although they may require different approaches. No method can account for all uncertainty (e.g. that due to the omission of unknown factors), so the key question is whether probabilistic methods can account for uncertainty better than alternative approaches.

People are comfortable with deterministic assessments based on few data, if they include an appropriate safety or uncertainty factor: many regulatory decisions are made on this basis. This suggests that probabilistic assessments using few data would also be acceptable if people were satisfied that they account adequately for uncertainty, as argued above. Interestingly, although founded on general experience most of the uncertainty factors currently used for deterministic assessments are somewhat arbitrary, and in most cases the level of protection they provide is unknown (SCP, 2002; EU, 2002a, b). Furthermore, there is no clear guidance on how much uncertainty factors should be reduced when additional toxicity data are provided. The question is whether probabilistic methods for limited datasets can improve on this.

#### **5.4 General concerns that may be exacerbated by small datasets**

Several of the concerns that stakeholders have about probabilistic approaches in general are likely to be enhanced when small datasets are considered.

- The lack of worked examples is more severe for small datasets, because almost all existing examples of probabilistic assessments for pesticides involve large datasets.
- Parametric methods are sensitive to assumptions about distribution shape, which are more uncertain for small datasets.
- Small datasets may require different approaches (e.g. probability bounds). Users are currently less familiar with these methods and their software than the more popular approaches such as Monte Carlo simulation.
- Accounting for increased uncertainty due to the limitations of typical datasets may increase the complexity of the assessment model (e.g. to incorporate sampling uncertainty).
- Probabilistic methods would be used more frequently, and by more people, if they were accepted for typical datasets, than if they were used only for large datasets. This would increase the need for training and user-friendly software, for effective means of communicating methods and results, and for efficient processes for peer review.

## **6 WHAT APPROACHES SHOULD WE CONSIDER?**

### **6.1 Probabilistic methods that can incorporate uncertainty due to data limitations**

The previous sections suggest that the fundamental requirement for assessments based on typical data is that they take appropriate account of the uncertainties caused by the data limitations.

Methods for quantifying uncertainty are described in Work Package 4, which also considers the general advantages and disadvantages of each approach. Work Package 5 will focus on their suitability for use with typical datasets.

A preliminary list of potential advantages and disadvantages is given in Table 3. It is emphasised that this is intended purely as a starting point for discussion. Three types of activity are needed to refine this assessment:

1. Consultation with experts in each methodology to clarify how serious the disadvantages are and whether they can be overcome (e.g. confounding of subjectivist and frequentist interpretations of probability in 2D Monte Carlo, or the lack of a sampling theory for probability-bounds).
2. Development of worked examples for pesticide assessments with typical datasets, comparing the different approaches side-by-side.
3. Discussion amongst EUFRAM partners to evaluate the findings from 1 and 2 and seek a consensus on which approaches are suitable for use in regulatory assessments of pesticides.

Consideration should be given to using multiple approaches in combination, where appropriate. One of the limitations of Monte Carlo is that it requires strong assumptions about distribution shape and dependencies, which P-bounds avoids. One of the general limitations of P-bounds is that it provides no information about probability density within the bounds, whereas 2D Monte Carlo does. The solution may be to use both methods and overlay one on the other (e.g. Regan et al. 2002). The 2D Monte Carlo shows you where the answer is most likely to be, and the P-bounds show you how much worse it could be if your assumptions about distribution shape and dependencies are wrong – both valuable bits of information for the risk assessor and decision-maker. Doing this properly requires incorporation of sampling uncertainty in the P-bounds. Methods for doing this exist, but theory to support them is currently lacking (Ferson et al. 2003). It would be useful to establish which of the other methods can also be used in combination.

Unfortunately it is unlikely that there will be complete consensus amongst the experts on which methods to use, as there are fundamental theoretical issues about which they disagree (e.g. between frequentist and Bayesian viewpoints). For some purposes, however, the differing approaches give similar or identical results (e.g. Aldenberg & Jaworska, 2000). In such cases, at least, ease of use may determine the method of choice.

**Table 3.** Preliminary assessment of selected approaches for dealing with uncertainty in risk assessments with limited data (heavily adapted from Ferson, 2002 and Ferson, pers. comm.). The advantages and disadvantages are offered only as a starting point for discussion and investigation (e.g. in case studies).

	<b>Advantages</b>	<b>Disadvantages</b>
<b>Worst case analysis</b>	<ul style="list-style-type: none"> <li>account for uncertainty by being conservative</li> </ul>	<ul style="list-style-type: none"> <li>difficult to decide how conservative each parameter should be without doing an uncertainty analysis</li> <li>difficult to quantify overall conservatism when worst cases are combined</li> </ul>
<b>Interval analysis</b>	<ul style="list-style-type: none"> <li>bounds account for uncertainty by being conservative</li> </ul>	<ul style="list-style-type: none"> <li>same disadvantages as worst case analysis</li> </ul>
<b>Monte Carlo simulation</b>	<ul style="list-style-type: none"> <li>can incorporate sampling uncertainty about parameter estimates for some distribution types including Normal &amp; Lognormal (Vose, 2000)</li> </ul>	<ul style="list-style-type: none"> <li>does not take account of uncertainty about distribution shape, nor other types of model uncertainty</li> </ul>
<b>Second-order Monte Carlo</b>	<ul style="list-style-type: none"> <li>separates variability and uncertainty</li> <li>can handle model uncertainty in a limited way</li> </ul>	<ul style="list-style-type: none"> <li>does not take account of uncertainty about distribution shape</li> <li>confounds frequentist and subjectivist interpretations of probability</li> </ul>
<b>Probability bounds analysis</b>	<ul style="list-style-type: none"> <li>can account for uncertainty about parameter values, distribution shapes, dependencies, and model form</li> </ul>	<ul style="list-style-type: none"> <li>must truncate infinite tails on distributions for every parameter with sampling uncertainty</li> <li>lacks theoretical basis for interpreting its treatment of sampling uncertainty</li> </ul>
<b>Range-moment propagation</b>	<ul style="list-style-type: none"> <li>same as probability bounds but simpler to use</li> </ul>	<ul style="list-style-type: none"> <li>cannot use further information about distribution and correlations (must resort to probability bounds)</li> </ul>
<b>Bayesian methods</b>	<ul style="list-style-type: none"> <li>accounts in a robust and logical way for subjective uncertainty</li> <li>Bayesian updating provides a natural method for refining the assessment as new data is added</li> <li>can be used to combine different types of information</li> </ul>	<ul style="list-style-type: none"> <li>lack of consensus about use of subjective uncertainty and non-informative priors</li> <li>seem to be more difficult than other methods for non-specialists to understand and use</li> </ul>

## 6.2 Make more use of existing data

Large amounts of data already exist for some generic parameters that are treated as fixed in current deterministic assessments. For example, this is true of several parameters used in terrestrial vertebrate exposure assessment such as body weight and the parameters used to estimate daily food intake. These data could be collated and used to estimate standard distributions.

Large amounts of data already exist for pesticide-specific parameters such as toxicity. If difficulties gaining access to these data can be resolved (which is the objective of EUFRAM Work Package 11), they could be used to develop generic (pooled) estimates and/or uncertainty factors. For example, Luttik & Aldenberg (1997) calculated pooled estimates of between-species variation in acute toxicity for mammals and birds, and used them to derive extrapolation factors. This type of approach could be refined and applied to more sorts of data. In addition, consideration should be given to using Bayesian methods to combine generic estimates (“prior” distributions) with specific data for the pesticide under assessment.

Consideration should be given to using formal methods such as Bayesian updating for combining different types of information, e.g. combining laboratory-based risk estimates and information from field studies and monitoring.

### **6.3 Develop a suite of standard assessment scenarios**

Generic parameters affecting exposure vary widely between habitats and regions, and between species (e.g. proximity of crops to surface water, soil types, foraging behaviour of birds, etc). It may not be necessary to quantify all this variability if a limited number of standard scenarios can be defined for use in risk assessment, as has been done by the Focus workgroups. Standardisation of scenarios is likely to increase the acceptance of new approaches by regulators, and should contribute to promoting consistency between different assessors and reduce the potential for misleading results. However, it will be important to understand how the standard scenarios relate to the wider range of conditions in the field, to help decision-makers interpret their significance (FOCUS, 2002).

If a limited number of standard scenarios are defined, it becomes more feasible to develop reliable distributions for the key generic parameters. These data need only be generated once, provided that arrangements can be made for sharing them.

### **6.4 Refine each assessment only as much as is necessary, and target data collection on key uncertainties**

The use of tiered assessment approaches is a familiar method of adjusting the level of effort to the needs of the assessment. The ECOFRAM project considered ways of applying this to probabilistic assessments. The ECOFRAM aquatic work group proposed four levels of refinement, with increasingly sophisticated assessment tools at each level (ECOFRAM, 1999a). The ECOFRAM terrestrial work group suggested that a probabilistic assessment might start with worst case assumptions for most parameters and the progressively replace them with distributions, one parameter at a time (ECOFRAM, 1999b). This process would continue until either (a) the risk was shown to be acceptable, or (b) it became apparent that further refinements were unlikely to reduce the risk estimate significantly. Moore et al. (1999) argued that it is inappropriate to mix worst case assumptions and distributions within the same model, and that one should start instead with distributions for all (important) parameters. In both this and the terrestrial ECOFRAM approach, sensitivity analysis would be used to identify which parameters contribute most to uncertainty in the overall output (assessment endpoint). Data collection could then be targeted on those parameters that were likely to be the most cost-effective options for uncertainty reduction. Further investigation is required to evaluate these and other possible

approaches to optimise the assessment process and minimise the data needed to reach regulatory decisions.

## 7 ACTIONS FOR NEXT PHASE OF EUFRAM

The project meeting on 12-14 May 2003 agreed on the following activities to explore and evaluate the issues and options discussed in this paper.

Action	Who <sup>8</sup>	When
1. Investigate more thoroughly the advantages and disadvantages of alternative approaches to accounting for uncertainty in assessments with typical (small) datasets, by reviewing the literature and by consulting relevant experts. Attempt to identify ways to avoid or mitigate disadvantages, including the lack of a sampling theory for P-bounds. Find out which methods are complementary and/or can be used in combination. Summarise findings in draft section for WP4/5 report(s).	CSL	30/9/03
2. Collaborate with the case studies (in WP8 and also outside EUFRAM) to test alternative approaches for uncertainty analysis separately and in combination. Examine their ability to account for different types of uncertainty arising from limited datasets (e.g. sampling uncertainty, extrapolation uncertainty when using surrogate data, subjective uncertainty when using expert judgement, model uncertainty). Report findings in WP5 report.	CSL	30/6/04
3. Review research on approaches that may reduce uncertainty when data are limited, e.g. using extrapolations between species, between chemicals, between endpoints, QSARs, QSPRs. Report findings in WP5 report.	CSL	30/6/04
4. Consider the feasibility and usefulness of establishing a register or archive of assessment scenarios and distributions for generic parameters. Draft design for database for discussion at PM2. Report findings in WP5 report.	CSL	30/6/04
5. Determine support for holding a small meeting of SSD researchers to share ideas for estimating species sensitivity distributions from small datasets and promote collaboration. Report outcome in WP5 report.	CSL	30/6/04
6. Cranfield University to report on their work on general guidance on selecting approaches for attributing and estimating distributions, according to the amount of data available. Include in WP5 report.	CU	30/4/04

In addition, some issues relevant to WP5 will be addressed by activities in other work packages: WP4 will seek to develop effective ways of communicating the principles of different methods of uncertainty analysis, and WP3 will consult with end-users and decision-makers to explore how they would use information on uncertainty in risk assessment.

<sup>8</sup> EUFRAM partner acronyms. CSL = Central Science Laboratory, CU = Cranfield University.

## 8 REFERENCES

- Aldenberg T, Jaworska J. 2000. Uncertainty of the hazardous concentration and fraction affected for Normal species sensitivity distributions. *Ecotoxicology and Environmental Safety*, 46: 1-18.
- Codex, 2002. Proposed draft working principles for risk analysis for application in the framework of the Codex Alimentarius. Codex Alimentarius Commission, ALINORM 03/33 Appendix II, April 2002.
- Dubus IG, Brown CD, Beulke S, Turner NL. 2002. Uncertainty and probabilistic approaches to pesticide fate modelling. Cranfield Centre for EcoChemistry, report to Defra. Project PL0598.
- ECOFRAM 1999a. ECOFRAM Aquatic Draft Report. Ecological Committee on FIFRA Risk Assessment Methods.
- ECOFRAM 1999b. ECOFRAM Terrestrial Draft Report. Ecological Committee on FIFRA Risk Assessment Methods.
- European Commission, 2002a. Guidance document on aquatic ecotoxicology in the context of Directive 91/414/EEC. Sanco/3268/2001 rev.4 (final), 17 October 2002, Brussels.
- European Commission, 2002b. Guidance document on risk assessment for birds and mammals under Council Directive 91/414/EEC. Sanco/4145/2000, 25 September 2002, Brussels.
- European Commission, 2003. Risk assessment of food borne bacterial pathogens: quantitative methodology relevant for human exposure assessment. Final report. Adopted by the Scientific Steering Committee, 16-17 January 2003. Brussels: European Commission.
- Ferson S, 2002. RAMAS RiskCalc 4.0 Software. Risk assessment with uncertain numbers. Lewis Publishers.
- Ferson S, Kreinovich V, Ginzburg L, Myers DS, Sentz K. 2003. Constructing probability boxes and Dempster-Shafer Structures. Report number SAND2002-4015. Sandia National Laboratories.
- Focus Surface Waters Report. (full reference?).
- Forbes V & Calow P. 2002. Species sensitivity distributions revisited: a critical appraisal. *Human and Ecological Risk Assessment*, 8: 473-492.
- Hart, A. 2001. Probabilistic risk assessment for pesticides in Europe: implementation and research needs. Report of the European workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products (EUPRA). Central Science Laboratory, Sand Hutton, UK. 109pp.
- Luttik, R. and Aldenberg, T. 1997. Extrapolation factors for small samples of pesticide toxicity data: special focus on LD50 values for birds and mammals. *Environmental Toxicology and Chemistry*, 16: 1785-1788.
- Moore DRJ, Sample BE, Suter GW, Parkhurst BR and Teed RS. 1999. Risk-based decision making: The East Fork Poplar Creek case study. *Env Toxicol Chem*: 18: 2954-2958.

Regan HM, Hope BK, Ferson S. 2002. Analysis and portrayal of uncertainty in a food-web exposure model. *Human and Ecological Risk Assessment*, 8: 1757-1777.

SCP, 2002. Opinion of the Scientific Committee on Plants on the Draft Guidance Document on Risk Assessment for Birds and Mammals under Directive 91/414/EEC. DG SANCO, European Commission, Brussels.

Vose, D. 2000. *Risk analysis: a quantitative guide*. 2<sup>nd</sup> edition. John Wiley & Sons Ltd.