Report to the Department for Environment, Food & Rural Affairs

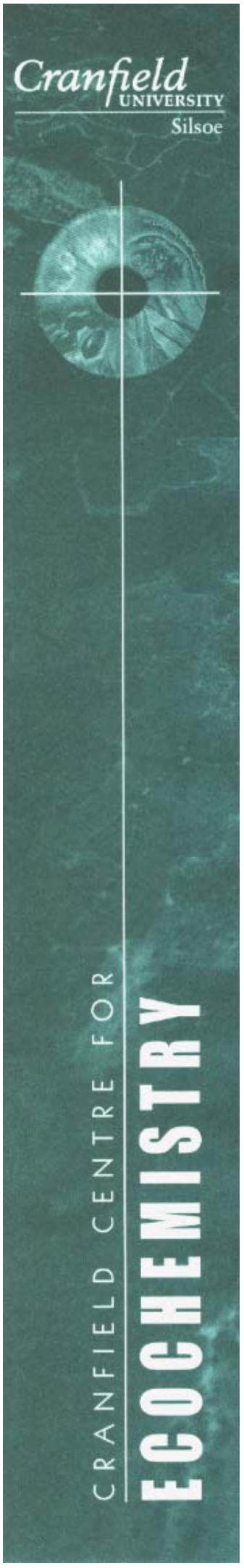
DEFRA project PL0548
Study number: JA3755E

Uncertainty and probabilistic approaches to pesticide fate modelling

I.G. Dubus, C.D. Brown, S. Beulke & N.L. Turner

December 2002

Cranfield Centre for EcoChemistry
Cranfield University, Silsoe, Beds MK45 4DT, UK
www.cranfield.ac.uk/ecochemistry



Cranfield UNIVERSITY

Uncertainty and probabilistic approaches to pesticide fate modelling

by

I.G. Dubus, C.D. Brown, S. Beulke & N.L. Turner

December 2002

DEFRA Project PL0548

Cranfield Centre for EcoChemistry research contract No. JA 3755E

Cranfield Centre for EcoChemistry
Cranfield University, Silsoe, Beds MK45 4DT, United Kingdom
Tel: 01525 863000 Fax 01525 863533
ecochemistry@cranfield.ac.uk www.cranfield.ac.uk/ecochemistry

Foreword

In 2000, the Department for Environment, Food & Rural Affairs commissioned Cranfield Centre for EcoChemistry to conduct research to identify the merits and shortcomings of the use of probabilistic approaches to assess pesticide exposure. The present report provides a review of the sources of uncertainty in pesticide fate modelling, investigates the technical robustness of Monte Carlo based probabilistic approaches, presents a case study to demonstrate the application of a number of approaches and discusses the pros and cons of introducing a probabilistic element in pesticide fate modelling from a conceptual, technical and regulatory perspective.

The preferred reference to the present report is as follows:
Dubus I.G., Brown C.D., Beulke S. & Turner N.L. (2002). Uncertainty and probabilistic approaches to pesticide fate modelling. Cranfield Centre for EcoChemistry research report for DEFRA PL0548, Cranfield University, Silsoe, UK, 138p.

Acknowledgements

The authors are grateful to the UK Department for Environment, Food & Rural Affairs for the funding of this research within the scope of the programme PL05: Fate and Behaviour of Pesticides in Soil and Water. Thanks are due to Syngenta for the provision of data on environmental properties of atrazine. The discussion presented in Chapter 3 is the result of an informal collaboration between Igor Dubus and Peter Janssen (RIVM, The Netherlands). Gerard Heuvelink (University of Amsterdam, The Netherlands) kindly commented on the section dedicated to spatially referenced data in Chapter 2. The opinions expressed and conclusions drawn in this report are those of the authors, not necessarily of the project's sponsor.

Contact details

Any enquiries regarding this report should be addressed to:
Dr Igor Dubus
Cranfield Centre for EcoChemistry
Cranfield University, Silsoe
Beds MK45 4DT, United Kingdom
Tel: +44 (0)1525 863378
Fax: +44 (0)1525 863533
E-mail: i.dubus@cranfield.ac.uk

Executive summary

Although there is currently strong interest in Europe in introducing a probabilistic component into environmental risk assessment procedures to support pesticide registration, very little research has been conducted on probabilistic aspects of pesticide fate modelling. The overall aim of the project was to investigate the merits and shortcomings of probabilistic modelling for predicting pesticide exposure from a conceptual, technical and registration perspective. Particular attention was paid to the robustness of Monte Carlo modelling, the most popular probabilistic technique, and to estimation of the confidence that should be attributed to results of probabilistic modelling.

A comprehensive review of the different sources of uncertainty that can potentially affect predictions from pesticide fate models was undertaken. The review demonstrated that pesticide fate modelling is laced with uncertainty, thereby supporting the use of probabilistic approaches to estimate the uncertainty associated with model predictions. The work suggests that probabilistic approaches that are typically being deployed to account for uncertainty in pesticide fate modelling, in particular Monte Carlo modelling, ignore a number of key sources of uncertainty (including model error and modeller subjectivity).

Monte Carlo modelling exercises were repeated for various numbers of model runs and for a number of replicated random samples (the random samples were generated by Latin Hypercube Sampling using different seed numbers). The probability of exceeding a specific concentration, an indicator of major interest within the context of environmental risk assessment for pesticides, was found to be significantly influenced by the seed number used in the sampling, even when a large number of model runs (5000) were undertaken. Within the constraints imposed by running times of pesticide fate models, this suggests that Monte Carlo results should be considered inherently variable and that the levels of variability reported have the potential to affect decision-making. Similar issues can be anticipated when percentile concentrations (95^{th}-, 99^{th}-percentile concentrations) are derived.

The influence on Monte Carlo results of subjective choices made during the implementation of the technique was investigated by considering multiple options for parameterising lognormal distributions and repeating the modelling for different levels of parameter correlation. Probabilities of exceeding a concentration of 0.1 μg/l were found to be significantly influenced by these subjective choices. Levels of variation reported could be of relevance to decision-making if exceedance probabilities were close to threshold levels.

A case study was undertaken to illustrate the application of a range of probabilistic approaches to pesticide fate modelling. A range of probabilistic techniques of varying complexity were applied to predict potential leaching of atrazine within the context of a groundwater risk assessment. Calculations included two simple approaches based on FOCUS groundwater scenarios and two higher-tier approaches (Monte Carlo based modelling and scenario-based modelling). The two refined FOCUS approaches and Monte Carlo modelling built directly on FOCUS modelling and expressed to a lesser or greater extent the uncertainty arising from that associated with timing of application or properties of the pesticide. The research suggests that these approaches are unlikely to further inform a regulatory decision where a standard FOCUS prediction exceeds 0.1 µg/l and that the generic nature of the FOCUS scenarios makes them inappropriate to support higher tier modelling. Scenario-based modelling was implemented by selecting 16 environmental scenarios covering the range of soil and climatic conditions under which atrazine might be applied in England and Wales and by running the MACRO model for each of the scenarios selected using long-term (20 years) weather sequences. An overall probability of impact at the larger scale was obtained by weighting results for each scenario by their abundance in the broader landscape. In contrast to other approaches considered which were based on one or a number of worst-case scenarios, scenario-based modelling permits an assessment that better reflects the variability in soil and climatic conditions across the proposed area of use. However, the typical application of deterministic models in scenario-based modelling ignores the uncertainty resulting from that in sorption and degradation parameters, which is known to be of importance for the prediction of pesticide loss. The Monte Carlo and scenario-based modelling are complementary to some extent and the two approaches should ideally be combined.

The research raised a number of conceptual and technical issues with regard to the typical application of Monte Carlo techniques to pesticide fate models. Some of these were related to the specific nature of the modelling undertaken to assess pesticide exposure within the registration context whereas some others were more universal. The research suggests that extreme care should be exercised when considering results originating from Monte Carlo modelling and limitations of the technique in terms of robustness should not be overlooked. Nevertheless, the consideration of uncertainty in the modelling, and especially of that originating with sorption and degradation parameters, remains a priority as this constitutes a weak point in exposure assessment. Further research into Monte Carlo modelling as well as other probabilistic techniques for environmental exposure is required to try to address the issues outlined in the project and to increase the robustness of probabilistic estimates to support decision-making in the context of pesticide registration.

Table of contents

CHAPTER 3: ASPECTS OF REPEATABILITY IN MONTE CARLO MODELLING

CHAPTER 4: DEPENDENCE OF MONTE CARLO RESULTS ON PARAMETERISATION ASSUMPTIONS

CHAPTER 5: A CASE STUDY ILLUSTRATING THE USE OF PROBABILISTIC APPROACHES FOR EXPOSURE ASSESSMENT

CHAPTER 6: OVERALL CONCLUSIONS

Chapter 1:

Introduction

1 THE RATIONALE FOR PROBABILISTIC RISK ASSESSMENT

Current environmental risk assessment procedures for pesticide registration in the European Union rely on the comparison between exposure and ecotoxicological endpoints (surface waters) or a legal threshold concentration (groundwater). A tiered approach is implemented to concentrate on those compounds which are most likely to be harmful to the environment and not penalise those which pose little threat. Relevant ecotoxicological endpoints are typically derived through laboratory tests using a few representative organisms whereas the estimation of predicted environmental concentrations for exposure (PECs) relies heavily on the use of predictive models, especially at higher tiers. For surface waters, the ratio between a relevant PEC and ecotoxicological endpoint is calculated (termed TER for Toxicity:Exposure Ratio[*]) and compared to threshold values which are dependent on the organism considered. These arbitrarily-set thresholds (typically factors of 10) act as 'safety' (also know as 'uncertainty') factors in the risk assessment. A compound is considered to pose little threat to surface water organisms if the TER exceeds the relevant threshold. For groundwater, PECs for leaching (generally estimated at 1-m depth) are compared to a threshold concentration of 0.1 μg/l, irrespective of the toxicity or ecotoxicity of the compound. The current approach, often referred to as point estimate or deterministic risk assessment, is designed to provide upper-bound estimates of risk through the integration of a range of worst-case assumptions (in the assessment of exposure, in the assessment of effects, through the use of safety factors). Such worst case analyses are useful at lower tiers of the assessment because they are relatively quick to deploy and act as a screening step. Compounds which have shown a potential for environmental impact undergo additional work to better quantify the likelihood and magnitude of the risk involved.

Although point estimate methods are useful in the context of pesticide registration, the degree of environmental protection provided by the approach is unknown to a large extent. Under-conservatism would mean environmental impact while over-conservatism would mean that

[*] Note that the reciprocal of the TER quotient is used in the US (Hazard quotient=Exposure/Toxicity)

some compounds that could benefit agriculture would not meet environmental criteria for registration despite being relatively innocuous to the environment. Attempts to get an understanding of the actual risk to the environment are hampered by the integration of worst-case assumptions and aspects of uncertainty in the approach. This would require a detailed characterisation of the uncertainty associated with exposure and the receptors under consideration. For these reasons, there has been a strong interest both in the US (ECOFRAM, 1999) and in Europe (EUPRA, 2001) in probabilistic risk assessment techniques which are expected to address many of the limitations of point estimate risk assessment.

Probabilistic risk assessment (PRA) is an approach to risk assessment which integrates uncertainty considerations and uses probabilities or probability distributions to characterise risk. Uncertainty considerations can be integrated into the assessment of exposure, effects, or both. In contrast to point estimate risk assessment, the overall objective of the method is to avoid worst-case assumptions and come up with a more realistic assessment of risk. Table 1 shows the strengths and weaknesses of probabilistic risk assessment as proposed by the EUPRA workshop (EUPRA, 2001).

Potential strengths and opportunities of PRA	**Potential weaknesses and threats of PRA**
PRA can produce outputs which are more meaningful ecologically	PRA techniques are more complex
PRA techniques can quantify variability and uncertainty to some extent	Some PRA techniques require more data
PRA makes better use of the available data	PRA may be difficult to communicate
PRA techniques enable the identification of factors which most influence risk assessment results	PRA may lead to misleading results
PRA may provide an alternative to field testing and help to focus the testing where required	There is no agreement on outputs to look at and on decision-making procedures
PRA promotes better science	PRA is difficult to validate

Table 1. Potential strengths and weaknesses of probabilistic risk assessment (PRA) when compared to the more traditional point estimate risk assessment currently in place (after EUPRA, 2001).
Entries in the table assume that probabilistic methods are implemented and used in appropriate ways.

2 PROBABILISTIC APPROACHES TO EXPOSURE ASSESSMENT

The present project specifically investigated aspects of probabilistic approaches for estimating exposure. Probabilistic approaches for effects and the combination of exposure and effects

within a probabilistic framework were outside the scope of the present investigations. These are covered to some extent in other research projects commissioned by DEFRA (PN 0932; PN0933).

A wide range of methods are available to propagate uncertainty in numerical assessments (Saltelli, 2000), but only a few have been applied in the context of pesticide fate modelling. These are differential analysis (Loague, 1991; Loague *et al.*, 1996; Freissinet *et al.*, 1999; Li *et al.*, 1998; Diaz-Diaz *et al.*, 1999), fuzzy logic (Freissinet *et al.*, 1998; 1999) and Monte Carlo analysis (Carsel *et al.*, 1988; Zhang *et al.*, 1993; Nofziger *et al.*, 1994; Soutter and Pannatier, 1996; Soutter and Musy, 1998, 1999; Zacharias *et al.*, 1999; Ma *et al.*, 2000; Trevisan *et al.*, 2001; Dubus and Brown, 2002; Carbone *et al.*, 2002). *Differential analysis* (of which First-Order Uncertainty Analysis is part) relies on the approximation of equations using Taylor series and allows the uncertainty to be expressed analytically. The difficulty in the application of the technique comes from the need to determine partial derivatives in the Taylor series approximation. This limits the application of the technique to simple models based on one or a few equations (*e.g.* the GUS index equation, the Attenuation Factor / Retardation factor approach) or to a few equations from more complex codes. The *fuzzy logic* approach to uncertainty assessment relies on the assignment of membership functions to model input parameters and their combination using fuzzy logic theory. This allows the derivation of an imprecision range for model predictions and a degree of confidence for this range. Again, the method can realistically only be applied to a small number of equations. In contrast, there are few restrictions to the application of *Monte Carlo* methods to complex computer models. The Monte Carlo approach is discussed in more detail in the next paragraph.

3 MONTE CARLO MODELLING

Monte Carlo procedures for assessing model sensitivity and uncertainty in the modelling are generic and have been applied to numerous fields of science. The wide adoption of Monte Carlo techniques results from i) the conceptual simplicity of the approach which can be easily communicated by scientists and risk assessors and understood by decision-makers; ii) the extensive literature reporting on the application of Monte Carlo approaches; iii) the availability of off-the-shelf software packages which facilitate the application of these techniques by non-experts; and, iv) the intrinsic benefits of the approach in terms of uncertainty assessment (*e.g.* derivation of cumulative distribution functions exploring the full range of possibilities). A Monte Carlo uncertainty analysis typically involves five stages

(Figure 1): i) the selection of model inputs and outputs on which the analysis will be performed; ii) the attribution of variation ranges and distributions to each parameter selected and the specification of their dependence/correlation; iii) the generation of a random sample from the distributions assigned to parameters; iv) the running of the model for each of the sample elements; and, v) the examination of model predictions in statistical terms (*e.g.* estimation of the mean and variance, construction of cumulative distribution charts). Monte Carlo analysis can also be used to assess model sensitivity. In this case, the fifth step consists of examining the relationship between changes in model predictions resulting from changes in model input parameters using scatterplots, regression analysis or correlation measures.

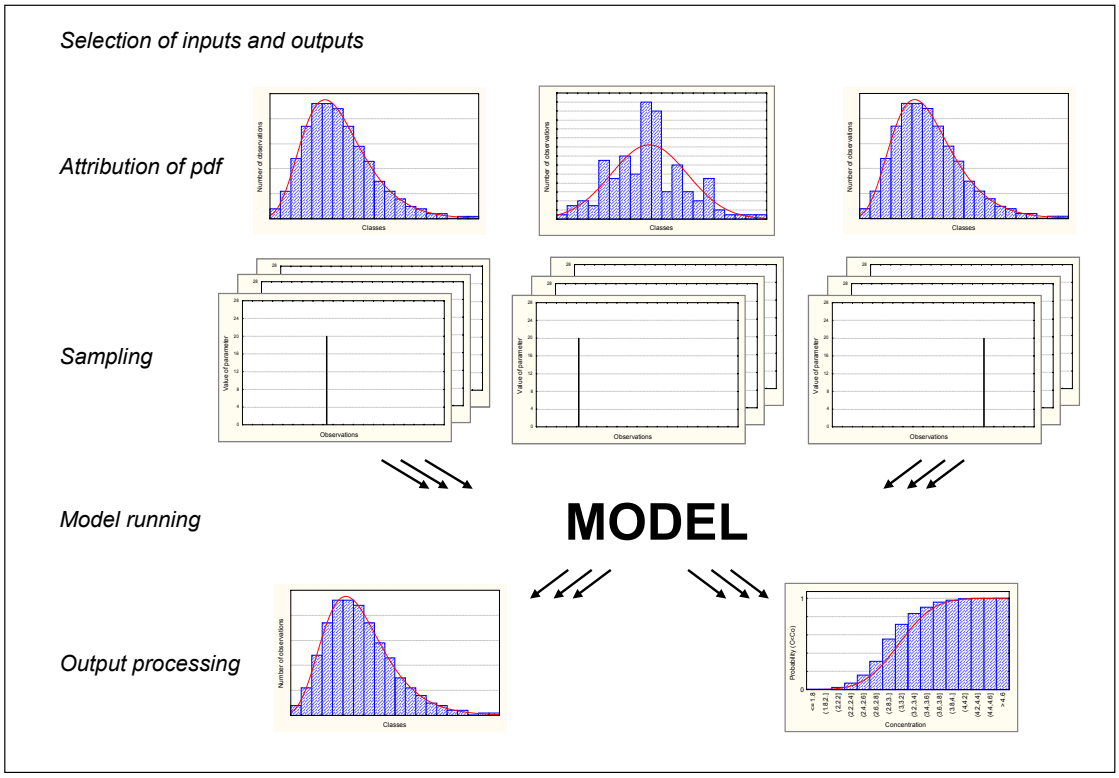


Figure 1. Schematisation of the Monte Carlo approach

The selection of the inputs to be included in the Monte Carlo analysis (step 1) should be based on a combination of information on the uncertainty in model input parameters and on the sensitivity of the model to these inputs (Dubus & Brown, 2002) although in practice it usually results from a subjective choice made by the modeller.

The attribution of probability density functions (also known as pdf's; step 2) is the key step in the deployment of Monte Carlo analysis as this will influence the final results of the analysis. Distributions which are typically available in dedicated Monte Carlo software packages

include: normal, lognormal, uniform, triangular, Poisson, binomial, exponential, geometric, Weibull, beta, Chi^2, gamma, logistic, Pareto and Pearson. A range of visual assessment, statistics and goodness-of-fit tests can be used to select the most appropriate distributions when data are available (Table 2). Where data are not available or not numerous enough, the parameterisation of distribution functions is usually based on literature information or expert judgement. In this case, uniform or triangular distributions tend to be used.

Basic statistics	Mean, median, mode, variance, coefficient of variation, skewness, kurtosis, percentiles
Visual assessments	Histograms, box and whisker plots, quantile plots, normal probability plots
Statistical tests	Chi^2, Kolmogorov-Smirnov, Lilliefors, Anderson-Darling, Filliben, Shapiro-Wilk, d'Agostino

Table 2. Range of methods which can be used to support the attribution of distribution functions from experimental data

One of the disadvantages of the Monte Carlo approach is that it is computationally intensive. For this reason, the sampling stage (step 3) is usually designed to minimise the number of model runs to be undertaken. The Latin Hypercube Sampling (LHS) scheme (McKay *et al.*, 1979) is widely used. LHS is considered to be an extremely efficient sampling scheme when compared to purely random MC sampling because it provides a better coverage of the parameter space and allows a significant reduction in the sample size to achieve a given level of confidence without compromising the overall quality of the analysis (Helton, 1993). In LHS, the probability distributions are divided into ranges of equal probability and one sample is taken from each equi-probability interval. Restricted pairing techniques are available for inducing correlations between variables in LHS (Iman & Conover, 1982; Iman and Shortencarier, 1984). A number of software packages are available to generate Monte Carlo random samples. Packages which have been extensively evaluated by the authors and which proved suitable for pesticide fate modelling are Crystal Ball 2000 (Decisioneering, 2000), @RISK (Palisade, 2002) and UNCSAM (Janssen *et al.*, 1992). Crystal Ball 2000 and @RISK are popular add-ons to Microsoft Excel. UNCSAM was developed at RIVM in the Netherlands to support environmental risk assessments. All these packages will cover all stages of uncertainty and sensitivity analyses, from the selection of probability density functions to the analysis of the Monte Carlo results.

Step 4 in Monte Carlo modelling consists of the automatic running of the model for hundreds to thousands of iterations. Each run will integrate a set of random values for input parameters selected for inclusion in the analysis. The replacement of input values in input files, the running of the pesticide fate model and the extraction of the relevant model ouputs can be

automated using dedicated packages such as SENSAN, a utility provided as part of the PEST package (Doherty, 2002). Alternatively, these automation procedures can be coded in any programming language. The implementation of a full Monte Carlo analysis from Excel (*i.e.* random sampling, replacements in input files, running of the model, extractions in output files, processing of the results) is theoretically feasible, but would prove more difficult than managing the whole modelling process from DOS, the approach which is typically favored.

The final step in Monte Carlo modelling is to examine the results from a statistical perspective. The statistics that are likely to be of most interest to risk assessors within the context of pesticide fate modelling and that can be derived from a cumulative distribution chart (Figure 2) are probabilities of exceedance of a particular concentration (*e.g.* 0.1 µg/l or a ecotoxicological endpoint) or specific percentile concentrations (*e.g.* median, 95th-, 99th-percentile concentrations).

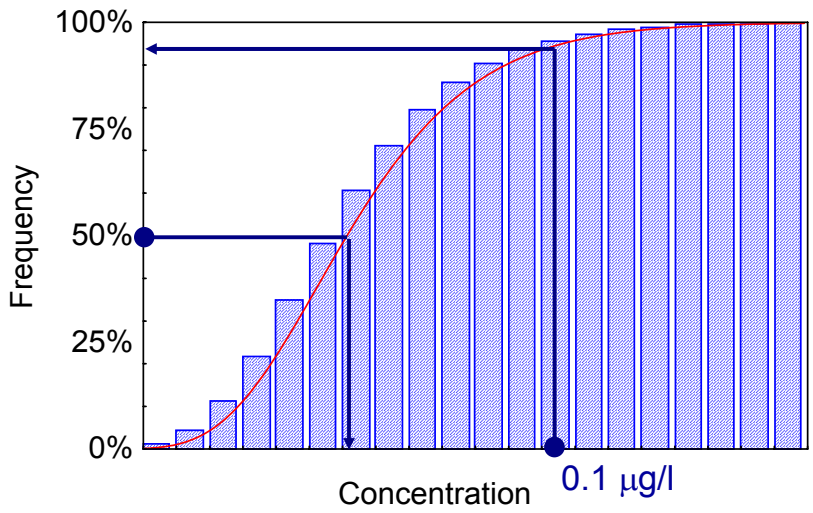


Figure 2. Example of cumulative distribution chart obtained through a Monte Carlo analysis. In this example, the 50th-percentile (median) concentration is 0.05 µg/l while the probability of predicting a concentration below 0.1 µg/l is 95% (the probability of exceeding 0.1 µg/l is 5%).

The default implementation of the Monte Carlo approach has been presented. It should be noted that a popular refinement to the standard Monte Carlo modelling is to differentiate the contribution of stochastic variability and uncertainty (in the sense of lack of knowledge; also known as incertitude) in a two-dimensional Monte Carlo analysis. Investigations of applications of 2D Monte Carlo analysis was considered outside the scope of the present project.

4 PRESENTATION OF THE RESEARCH

Research investigations undertaken within the scope of the present project were aimed at *estimating the confidence that should be attributed to results of probabilistic approaches within the context of pesticide fate modelling*. The research concentrated on Monte Carlo modelling as this is by far the most popular probabilistic approach to dealing with uncertainty and the one that currently receives the most attention within the pesticide fate modelling community. The deployment of scenario-based modelling which is also considered of interest in Europe and elsewhere was also demonstrated. The robustness of the Monte Carlo approach and its potential usefulness in pesticide fate modelling was considered from a conceptual, scientific and regulatory perspective.

The document is presented in the form of a collection of five individual chapters. Each chapter benefits from a synopsis located at the front and can be considered on its own. The *present chapter* is an introductory chapter aimed at presenting the rationale for the introduction of probabilistic approaches in pesticide fate modelling, the range of probabilistic approaches available to pesticide fate modellers, the concept of Monte Carlo modelling and the research undertaken. *Chapter 2* is a comprehensive review on the sources of uncertainty in pesticide fate modelling. An attempt is made to assess the magnitude of the different uncertainties inherent in the prediction of pesticide fate. Implications with regard to current approaches to pesticide fate modelling are discussed. *Chapter 3* presents investigations on the issue of replicability in sensitivity and uncertainty (Monte Carlo modelling) analyses. Sensitivity and uncertainty analyses were repeated for a number of replicated random samples and the variability of probabilistic estimates obtained was compared to that predicted by the statistical theory. Implications for the confidence that should be assigned to results of probabilistic exercises are discussed. *Chapter 4* reports on further tests aimed at assessing the robustness of probabilistic modelling. Investigations concentrated on the impact on results of Monte Carlo modelling of i) the strategy used to parameterise probability density functions; and, ii) the specification of correlation coefficients between model input parameters. *Chapter 5* reports on a case study aimed at presenting results of various probabilistic approaches within the scope of a groundwater leaching assessment for atrazine in England & Wales. Approaches deployed differed in their complexity and included Monte Carlo and scenario-based modelling. Approaches are discussed individually and their potential benefits within the context of environmental risk assessment for pesticide registration are reviewed. *Chapter 6* is a concluding chapter. Implications with regard to the usefulness of probabilistic approaches for pesticide fate modelling are discussed.

References

Carbone J.P., Havens P.L. & Warren-Hicks W. (2002). Validation of Pesticide Root Zone Model 3.12: employing uncertainty analysis. Environmental Toxicology and Chemistry, 21:1578-1590.

Carsel R.F., Parrish R.S., Jones R.L., Hansen J.L. & Lamb R.L. (1988). Characterizing the uncertainty of pesticide leaching in agricultural soils. Journal of Contaminant Hydrology, 2:111-124.

Decisioneering (2000). Crystal Ball 2000 – User manual. Decisioneering Inc., Denver, Colorado, USA.

Diaz-Diaz R., Loague K. & Notario J.S. (1999). An assessment of agrochemical leaching potentials for Tenerife. Journal of Contaminant Hydrology, 36:1-30.

Doherty J. (2002). PEST – model independent parameter estimation. Watermark Computing, Corinda, Australia.

Dubus I.G. & Brown C.D. (2002). Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. Journal of Environmental Quality, 31:227-240.

ECOFRAM (1999). http://www.epa.gov/oppefed1/ecorisk/

EUPRA (2001). Probabilistic risk assessment for pesticides in Europe. Implementation and research needs. Report from the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products, The Netherlands, June 2001.

Freissinet C., Erlich M. & Vauclin M. (1998). A fuzzy logic-based approach to assess imprecisions of soil water contamination modelling. Soil & Tillage Research, 47:11-17.

Freissinet C., Vauclin M. & Erlich M. (1999). Comparison of first-order analysis and fuzzy set approach for the evaluation of imprecision in a pesticide groundwater pollution screening model. Journal of Contaminant Hydrology, 37:21-43.

Helton J.C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliability Engineering and System Safety, 42:327-367.

Iman R.L. & Conover W.J. (1982). A distribution-free approach to inducing rank correlation among input variables. Communications in Statistics - Simulation and Computation, B11:311-324.

Iman R.L. & Shortencarier M.J. (1984). A Fortran 77 program and user's guide for the generation of Latin hypercube and random samples for use with computer models. SAND83-2365. Sandia National Laboratories, Albuquerque, New Mexico, USA.

Janssen P. M. H., Heuberger P. S. C. & Sander R. (1992) UNCSAM 1.1: a Software Package for Sensitivity and Uncertainty Analysis, Manual, RIVM-report nr. 959101004, Bilthoven, The Netherlands.

Li Z.C., Yost R.S. & Green R.E. (1998). Incorporating uncertainty in a chemical leaching assessment. Journal of Contaminant Hydrology, 29:285-299.

Loague K. (1991). The impact of land use on estimates of pesticide leaching potential: assessments and uncertainties. Journal of Contaminant Hydrology, 8:157-175.

Loague K., Bernknopf R.L., Green R.E. & Giambelluca T.W. (1996). Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii: review. Journal of Environmental Quality, 25:475-490.

Ma L., Ascough J.C., Ahuja L.R., Shaffer M.J., Hanson J.D. & Rojas K.W. (2000). Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. Transactions of the ASAE, 43:883-895.

McKay M.D., Conover W.J. & Beckman R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21:239-245.

Nofziger D.L., Jin-Song Chen & Haan C.T. (1994). Evaluating the Chemical Movement in Layered Soil Model as a tool for assessing risk of pesticide leaching to groundwater. Journal of Environmental Science and Health, A29:1133-1155.

Palisade (2002). Guide to using @RISK. Palisade Corporation, Newfield, New York, USA.

Saltelli A., Chan K. & Scott E.M. (2000). Sensitivity analysis. John Wiley & Sons, Chichester, UK.

Soutter M. & Musy A. (1998). Coupling 1D Monte-Carlo simulations and geostatistics to assess groundwater vulnerability to pesticide contamination on a regional scale. Journal of Contaminant Hydrology, 32:25-39.

Soutter M. & Musy A. (1999). Global sensitivity analyses of three pesticide leaching models using a Monte Carlo approach. Journal of Environmental Quality, 28:1290-1297.

Soutter M. & Pannatier Y. (1996). Groundwater vulnerability to pesticide contamination on a regional scale. Journal of Environmental Quality, 25:439-444.

Trevisan M., Calandra R., Vischetti C., Esposito A. & Padovani L. (2001). Pesticide leaching potential in the Trasimeno Lake area. Assessment of uncertainty associated with the simulation process. In: A. Walker (Ed.), Pesticide behaviour in soils and water, BCPC Symposium Proceedings No. 78, Farnham, UK, pp. 245-250.

Zacharias S., Heatwole C.D., Persaud N., Bruggeman A.C., Kumar D. & Smith C.N. (1999). Stochastic simulation of field-scale pesticide transport using Opus and GLEAMS. Journal of Environmental Quality, 28:411-423.

Zhang H., Haan C.T. & Nofziger D.L. (1993). An approach to estimating uncertainties in modeling transport of solutes through soils. Journal of Contaminant Hydrology, 12:35-50.

Chapter 2

Sources of uncertainty in pesticide fate modelling (synopsis)

The estimation of environmental exposure mainly relies on predictions of deterministic mathematical models and uncertainty in the prediction of exposure is typically largely ignored. There is therefore worldwide interest in the use of probabilistic approaches to assess pesticide exposure, as a means to account for uncertainty.

The first steps in conducting a probabilistic analysis of a system are: i) to identify where the uncertainties comes from; and, ii) to pinpoint those uncertainties that are likely to most affect the predictions of the system. The present chapter aims at addressing those two points within the context of exposure assessment for pesticides through a review of the different sources of uncertainty in pesticide fate modelling and a discussion of the likely magnitude of these uncertainties. The extensive listing of sources of uncertainty clearly demonstrates that pesticide fate modelling is laced with uncertainty. More importantly, the work suggests that probabilistic approaches that are typically being deployed to account for uncertainty in pesticide fate modelling ignore a number of key sources of uncertainty which are likely to have a significant effect on the prediction of environmental concentrations for pesticides (*e.g.* model error, modeller subjectivity).

Future research should concentrate on i) quantifying the impact these ignored uncertainties have on exposure assessments; and, ii) developing procedures that enable their integration within probabilistic assessments.

The research is presented in the form of a paper which was submitted to The Science of the Total Environment in November 2002.

Chapter 2

Sources of uncertainty in pesticide fate modelling

1 INTRODUCTION

Probabilistic approaches to environmental risk assessment for pesticides and other chemicals are currently receiving a vast amount of interest, especially in the US (ECOFRAM, 1999; Solomon *et al.*, 2000) and in Europe (Jager *et al.*, 2001; EUPRA, 2001). The application of probabilistic techniques to the assessment of exposure is traditionally expected to yield a number of benefits including: i) the quantification of the uncertainty associated with model predictions; ii) the identification of the factors which most influence model predictions; iii) the promotion of critical examination of the data and exposure models; and, iv) the generation of more meaningful outputs for subsequent decision-making (EUPRA, 2001). The overall aim of any probabilistic exercise is to account for 'uncertainty'. Uncertainty is a capacious term used to encompass a multiplicity of concepts (Morgan & Henrion, 1990) and has different meaning in various disciplines (Mowrer, 2000). Although numerous contributions in the literature deal with uncertainty, very few authors have actually defined what they mean by this term within the context of their research. Terminology which has been related to uncertainty within the context of contaminant modelling include: variation, variability, ambiguity, heterogeneity, approximation, inexactness, vagueness, inaccuracy, subjectivity, imprecision, misclassification, misinterpretation, error, faults, mistakes, artefacts. Within the context of this paper, the term is used in its widest sense and is meant to represent the combination of factors of various origins leading to a lack of confidence with regard to the description of the system under study. The terminology used encompasses both stochastic variability and incertitude.

The very first step in any assessment of uncertainty should be to compile a list of the different sources of uncertainty (Warren-Hicks and Moore, 1998). Listing of uncertainties is useful in that it provides a means to i) explicitly acknowledge the different uncertainties which will either be integrated in the modelling or not; ii) identify those uncertainties which are difficult to integrate into uncertainty analyses carried out using traditional methods (*e.g.* Monte Carlo or first order uncertainty analyses); and, iii) draw attention to the confidence that should be

assigned to the results of probabilistic modelling. Although uncertainty analyses and probabilistic modelling exercises within the context of pesticide fate modelling have been reported in the literature (*e.g.* Carsel *et al.*, 1988; Fontaine *et al.*, 1992; Nofzinger *et al.*, 1994; Maund *et al.*, 2001), sources of uncertainty have never been discussed in detail.

This paper reviews the sources of uncertainty associated with pesticide fate modelling. For illustration, emphasis is put on simulation of pesticide leaching through soil, but the analysis and discussions are relevant to other pesticide fate models and the broader field of environmental modelling. Contributions to the overall uncertainty are presented in a practical classification which follows the different steps in the modelling process: i) the acquisition of basic (primary) data in the field or in the laboratory, ii) the derivation of model input parameters from basic data or by other means, and iii) the modelling itself. Sources of uncertainty which are rarely acknowledged, but which may have a significant influence on estimations of environmental concentrations, are also discussed.

2 UNCERTAINTY IN THE PRIMARY DATA

The primary data are defined here as the basic physical, chemical and environmental properties determined either in the field or in the laboratory which are either directly fed into a model or used to derive input parameters for the model. Depending on the amount of information available to the modeller, primary data might include site characteristics (*e.g.* longitude, latitude, elevation), soil properties (*e.g.* organic carbon content, particle size distribution, bulk density, water retention data), weather conditions (*e.g.* rainfall, minimum and maximum temperatures, wind, humidity and radiation data), pesticide properties (*e.g.* solubility, Henry's constant, laboratory data on sorption and degradation) or results of field experiments (*e.g.* soil residues at different depths, pesticide loss by run-off, pesticide concentrations in water sampled using suction cups, pesticide concentrations in leaching or drainage waters). Although some primary data can be measured directly in the field (*e.g.* meteorological data, measurement of soil hydraulic conductivity), a common approach is to take samples in the field and to transfer them to the laboratory for analytical determination. Uncertainty in the primary data will arise from the spatial and temporal variability of environmental variables, from sampling procedures in the field and from analyses in the laboratory. These are discussed below.

2.1 Spatial and temporal variability of environmental variables

A large number of papers have reported on the spatial variability of pesticide residues or leaching in the field (Rao & Wagenet, 1985; Vischetti *et al.*, 1997) or in lysimeters (Flury *et al.*, 1998; Simmonds & Nortcliff, 1998). These have been attributed to some extent to the variability in space of environmental properties which in turn influence predictions of pesticide leaching models, such as physical (Bosch & West, 1998), chemical (Wood *et al.*, 1987) and biological (Parkin, 1993; Soulas & Lagacherie, 2001) data. Temporal variability has been the subject of much less research. Causes of spatial variability are traditionally classified into intrinsic or extrinsic factors. Taking the soil system as an example, intrinsic variability is due to natural conditions in soils whilst extrinsic variability is that imposed on a field as part of the crop production practices. Examples of soil characteristics which exhibit intrinsic variations are texture and mineralogy whilst tillage, fertilizer and pesticide applications, harvesting, and removal of crop residues all contribute to the development of an extrinsic variability. Some variables might be intrinsic in the first place but become extrinsic (Cambardella *et al.*, 1994). Extrinsic variability may either increase or decrease as farmers have created variation by enclosing, reclaiming, clearing and fertilising the land, though within fields they have removed some by cultivation and drainage (Webster, 2001). Extrinsic variability has often been considered a major reason for variability of pesticide residues in the field (Taylor *et al.*, 1971; Wauchope *et al.*, 1977).

Spatial dependency of measurements in the field occurs when samples collected close to one another are more similar than samples collected at greater distances (Cambardella *et al.*, 1994). Although the application of classical statistics to instances where spatial dependence is present is valid (Brus & de Gruijter, 1997), the use of geostatistics designed to exploit non-random and correlated data (Yost *et al.*, 1993) is often preferred (Hamlett *et al.*, 1986). Geostatistical analyses have been performed to study the spatial variability of pesticide sorption (Cambardella *et al.*, 1994; Novak *et al.*, 1997; Jacques *et al.*, 1999) and degradation (Price *et al.*, 2001) in the field. The reader is referred to Webster and Oliver (2001) for a description of geostatistical techniques. Describing uncertainty using geostatistics is not an activity exempt from uncertainty itself as variogram uncertainty may be large (Jansen, 1998) and spatial interpolation may be undertaken using different techniques.

Sorption and degradation properties of soils play a particularly important role within the context of pesticide fate modelling (Boesten & van der Linden, 1991; Dubus & Brown, 2002). Spatial variability in pesticide sorption has been reported for a range of geographical entities,

including individual fields showing some homogeneity in terms of soil series (Lennartz, 1999; Elabd *et al.*, 1986; Jacques *et al.*, 1999), fields or contiguous fields made of different soil series (Jaynes *et al.*, 1995; Novak *et al.*, 1997; Oliviera *et al.*, 1999), catchments (Coquet & Barriuso, 2002) and a large number of soils contrasting in their geographical origin (Barriuso & Calvet, 1992; Dubus *et al.*, 2001). Variability in the sorption distribution coefficient Kd can generally be reduced by normalising it to the organic carbon content (Hamaker & Thompson, 1972), but the variability of the resulting Koc often remains considerable (Figure 3). Some authors consider that the use of Koc reduces the variability in Kd and that variability in Koc is generally due to experimental errors and variation in the nature of the organic matter (Gerstl, 2000) although others have observed that normalisation of Kd to organic carbon failed to reduce variability compared to that of Kd (Elabd *et al.*, 1986; Beck *et al.*, 1996). In terms of shape of distribution to describe spatial variability, both normality (Coquet & Barriuso, 2002; Wood *et al.*, 1987; Nordmeyer, 1994; Elabd *et al.*, 1986; Lennartz, 1999) and log-normality (Novak *et al.*, 1997; Jaynes *et al.*, 1995) of Kd have been reported. In most studies accounting for the variability in the sorption distribution coefficient Kd, a normal distribution was hypothetised (Nofziger *et al.*, 1994; Franke & Teutsch, 1994; van der Zee and Boesten, 1993; Di & Aylmore, 1997) although uniform (Souttier & Panatier, 1996) and log-normal (Dubus & Brown, 2002) distributions have also been used. The practical implementation of Monte Carlo modelling for sorption has been handled differently by various authors. Nofziger *et al.* (1994) considered that Koc was normally distributed and then derived Kd by multiplying Koc by fixed values of organic carbon. In contrast, some studies derive Kd on the basis of fixed values of Koc and a distribution of organic carbon values (Lafrance & Banton, 1995; Zacharias *et al.*, 1999). Disregarding one of the two components of the variability in Kd is likely to have noticeable effects on probabilistic modelling (Coquet & Barriuso, 2002).

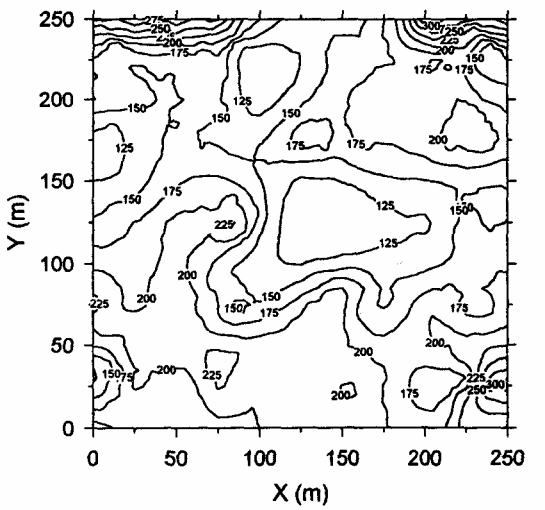


Figure 3. Distribution of Koc for atrazine within a 6.25-ha field.
The contour map was generated from 2601 block-kriged estimates of Koc for atrazine based on a 5 by 5 m grid pattern using block estimates (reproduced from Novak *et al.*, 1997)

The spatial variability of degradation has received less attention than that for sorption. Walker and coauthors (2002) investigated the variation in degradation of isoproturon at the field level and reported DT50 values varying from 6 to 30 days. Spatial differences in degradation were found to occur over a range of a few metres or even less. Parkin & Shelton (1992) reported CVs for DT50s of carbofuran of 9.4 to 17.5% for conventional tillage systems and 25.5 to 49.2% for no tillage systems (within field variability and row effect). Vischetti *et al.* (1997) reported CVs of 11 to 110% for DT50 for metamitron and chloridazon in three Italian field soils (one silt loam, two clay loams). Walker & Brown (1983) investigated the variability of DT50 for metribuzin and simazine for 10 different plots in three fields (sandy loam; CV of 7 and 21% for simazine and metribuzin, respectively) and the within-field variability (sandy loam; 20 samples per plot; CV 23 and 25% for simazine and metribuzin, respectively). Walker *et al.* (2001) suggested that the variability in the degradation of isoproturon was related to microbial biomass and diversity, which appeared to be influenced by pH conditions at the field scale. Wolt *et al.* (2001) consider that there is a case supporting the use of a lognormal distribution for describing the variability in degradation data because physicochemical properties of the soil environment that may impact half-life, such as aggregate and particle size distribution, exchangeable cations and diffusion coefficients are also lognormally distributed. Rao & Davidson (1982) reviewed data on degradation from a range of sources for 31 pesticides. The coefficient of variation for topsoil metabolism or dissipation was usually greater than 40% and averaged 73% (Cohen *et al.*, 1995). Wolt *et al.* (2001) considered that the quantification of the variability amongst soils should be a priority.

This can be achieved through replicated experiments with a large number of soils which are undertaken in such a way as to minimise experimental error.

Relatively large variabilities have been reported for pesticide residues in the field. Uncertainty in pesticide residues observed in the field shortly after application might be due to some extent to the variability associated with the spraying of the pesticide (Vischetti *et al.*, 2001). These authors reported CVs of 9.5 to 31% for concentrations of four herbicides sampled using spray traps and CVs of 30 to 61% for concentrations in soil 1-2 h after spraying. Larger CVs of 17-74% for application rates have been reported elsewhere (Vischetti *et al.*, 1998; Walker & Brown, 1983; Smith & Parrish, 1993). Residues of various contaminants in the environment have often been found to be log-normally distributed and Ott (1990) provided a theoretical justification the ubiquity of this skewed distribution in the description of environmental concentrations of pollutants.

Issues of spatial variability in the context of pesticide fate modelling are not limited to pesticide and soil properties as rainfall data may also be significantly variable (Krajewski *et al.*, 1998). Variability in rainfall will directly affect the water balance in pesticide fate models. Goodrich *et al.* (1995) assessed the uncertainty in rainfall data due to sampling equipment and spatial variability and concluded that the assumption usually made of spatial uniformity in rainfall at the small watershed scale did not hold for a 4.4-ha catchment characterised by convective storms. Chaubey and co-authors (1999) investigated the influence of the use of rainfall data from different weather stations on the uncertainty associated with predictions of a runoff model and concluded that a large uncertainty in the model outputs can be expected if this rainfall property is not accounted.

2.2 Sampling procedures and measurement error

The term ‘measurement error’ refers to the uncertainty originating from the sampling in the field and the determination of physical or chemical properties of samples. Differences in sampling procedures for pesticide residues between organisations and individuals is likely to contribute to the overall uncertainty in the determination of persistence of pesticides in the field. The reader is referred to the review paper by Ramsey (1998) for detailed discussion of sampling as a source of measurement uncertainty. It is not clear how the uncertainty introduced by sampling procedures compares with that associated with for example spatial or analytical variability. Walker and Brown (1981) compared the degradation of metamitron,

atrazine and propyzamide in fresh soil with that in soil stored for seven months either at 5°C, frozen, or sieved and air-dried. They found considerable differences in degradation rates both between the pretreatments and between stored and fresh soil for metamitron and atrazine. Differences in storage conditions and duration may contribute under some specific circumstances to the variability in the determination of degradation rates and pesticide concentrations in soil samples. Procedures for the preparation of sediment and soil samples (storage, subsampling, sieving, drying/rewetting, homogenisation) and the subsequent addition of organic compounds (addition method, carrier solvents, mixing) have been demonstrated to be largely variable between scientists and this may introduce unintentional bias in numerous experimental studies (Northcott & Jones, 2000). As well as a "random error", laboratory measurements may be subject to a “systematic error” which results from bias in the measuring apparatus and experimental procedure. Systematic errors arise, for example, because a measurement device is imperfectly calibrated, consistently used incorrectly or an incorrect multiplier or scaling factor is used in computations (Millstein, 1994). Uncertainty may also arise from the use of laboratory or field materials which have been demonstrated to interact with pesticides such as ceramic porous cups (Beltran *et al.*, 1995), walls of PVC lysimeters (Koskinen *et al.*, 1999), flasks and tubes (Topp & Smith, 1992) or filtering devices (Mouvet & Jücker, 1997; Clausen, 2000; Ahmad *et al.*, 2001).

With regard to rainfall, Krajewski *et al.* (1998) compared rainfall measurement from a tipping bucket gauge and an optical gauge at the same location and reported significant differences between the two measuring devices, especially at the beginning of rainfall events. Armstrong *et al.* (1996) reported a maximum of 10% variability in rainfall recorded on a small field using nine rain gauges. The impact of potential errors in the measurement of rainfall volumes on predictions of a cereal model was demonstrated by Heinemann *et al.* (2002). They simulated plant development and water balance for a range of crops under different precipitation scenarios. Scenarios were generated synthetically by assuming random over- or under-estimation of actual rainfall by measurement using rain gauges. Accounting for these potential errors in rainfall measurement introduced significant uncertainty in model predictions, especially in dry years. Specific factors which may influence the reliability of rainfall data include differences in measurement height, the presence of a building or natural feature which may influence the distribution of wind at the point of measurement (Verschoor *et al.*, 2001), losses by evaporation, the absence of heating systems to enable the measurement of snowfall, differences in collector shapes, the non-instantaneous character of buckets when being emptied, the presence of impurities deposited in buckets and inadequate calibration (Viton, 1970; Krajewski *et al.*, 1998).

3 UNCERTAINTY IN THE DERIVATION OF MODEL INPUT PARAMETERS

Process simulation models are by their very nature complex and it is not easy to make them simple to use (Garen *et al.*, 1999). A "good" model should strike a balance between complexity (which can be grossly estimated by the number of model parameters) and accuracy (goodness of fit *vs.* experimental data) (Beck *et al.*, 1997). Pesticide leaching models require values to be attributed to a large number of parameters (FOCUS, 2000). Although the use of experimental data for a direct parameterisation of the model is recommended, these data are not always available and most input parameters will require manipulation of the data by the modeller. For instance, units of the model input may be different from that of the field or laboratory measurement or the model might need surrogate parameters which reduce primary data to a few variables. Examples include the use of DT50 values (surrogate for the degradation data), sorption distribution coefficients (surrogate for the sorption isotherm data) and parameters of the van Genuchten or Brooks and Corey equations (surrogate for the water release and hydraulic conductivity curves). The difficulty in using experimental data directly in the modelling holds particularly for the more complex models which may include conceptual theories for which parameters are particularly difficult to derive. In these instances, the modeller might decide to i) leave the parameters at their default values, ii) make an educated guess using expert judgement, iii) extract values from existing databases, or iv) derive the values from empirical functions presented in the literature. The adoption of one procedure or the other will mostly depend on the experience of the individual with the model or similar programs and his/her knowledge of the processes the parameters refer to. Each procedure will introduce uncertainty into the modelling. The uncertainty in model input parameters may not directly be transposed into uncertainty in modelling predictions as the latter will result from a combination of uncertainty in model inputs and sensitivity of the model.

3.1 Derivation of parameters from primary data

As seen before, uncertainty in degradation can originate from spatial and temporal variability of degradation in the field, from storage and preparation of the soil samples and from slight differences in experimental conditions in the laboratory. Additional uncertainties in the modelling originate from the derivation of values for input parameters. A typical example is the variability in DT50 values arising from their derivation from laboratory data. Leake *et al.* (1995) calculated DT50 values for four different pesticides by applying six decay equations to laboratory degradation data using six curve fitting programs. Calculated DT50 values for the

four compounds were found to be strongly affected by the equation and package used (Table 3).

	Number of fittings	Range (days)	Coefficient of variation (%)
DT50 values			
Dataset 1	8	16-38	35.1
Dataset 2	7	95-110	5.4
Dataset 3	9	95-180	24.0
Dataset 4	9	64-180	41.3
DT90 values			
Dataset 1	8	32-120	42.3
Dataset 2	6	336-380	4.2
Dataset 3	9	317-411	8.1
Dataset 4	9	241-698	33.9

Table 3. Variability in DT50 and DT90 values introduced by the fitting of different equations to four degradation datasets (different compounds).
Only those fittings with r2>0.7 were considered. Adapted from Leake *et al.* (1995).

Other sources of uncertainty in the derivation of DT50 values from degradation data include i) the application of first-order kinetics to data which do not follow this decay pattern, ii) the use of linear transformation of the data as opposed to a direct non-linear estimation (Beulke & Brown, 2001); and, iii) modeller subjectivity (Boesten, 2000). An example of introduction of variability in results of pesticide leaching models introduced by procedures for deriving input parameters was presented by Beulke *et al.* (2001). They simulated the leaching of two pesticides using sorption data for four soils and investigated the variation in 80th-percentile concentrations resulting from two different procedures: a) taking mean values for sorption properties and running the model for a single iteration; and, b) running the model four times, one for each set of sorption data, and taking the mean of the four predicted concentrations. The use of the former approach was demonstrated to reduce the estimated mean pesticide concentration by factors of *ca.* 2 and 30 for the two pesticides as opposed to the use of the latter approach. Such a difference originates from the non-linear character of the Freundlich representation of sorption and of pesticide leaching models. Aspects of interaction between model non-linearity and error are discussed in Addiscott & Tuck (2001) using examples based on a solute leaching model. The difference between “evaluate first” and “interpolate first” also raises concern with regard to the uncertainty in large-scale spatial simulations, particularly where there is an implicit assumption that using spatially averaged parameters is equivalent to applying spatial averaging to the model output (Heuvelink & Pebesma, 1999; Addiscott & Tuck, 2001). Discrepancies can also be introduced in the modelling if different statistics to summarise sorption and degradation are used. As degradation data are generally

log-normally distributed (Wolt *et al.*, 2001), the median of a population of data points may be very different from its mean.

3.2 Procedures to derive input parameters using limited information

3.2.1 Use of pedotransfer functions

Soil hydrological properties are subject to a large spatial variability and strongly dependent on the measurement technique (Vereecken, 1992) and are thus seldom available to the modeller. In order to help with the model parameterisation, empirical equations, known as pedotransfer functions (PTFs; Bouma, 1989), have been established which express relationships between basic soil properties (*i.e.* those which are measured routinely or can be derived from a soil map) and these parameters which are difficult to measure. Typical examples are the derivation of retention data or parameters of retention curve equations (Tietje & Hennings, 1993) and the derivation of saturated hydraulic conductivities (Pachepsky *et al.*, 1999). PTFs suffer from a number of shortcomings which introduce additional uncertainty in the model parameterisation. First, the uncertainty in the databases which are used to derive relationships is likely to transfer to the derivation of values using these relationships. Uncertainty in databases arises from the aggregation of data from a wide variety of sources and the typically uneven sample size of datasets (Warren-Hicks and Moore, 1998). Secondly, these functions usually have a domain of validity which is often unknown to the end-user and thus ignored. Few pedotransfer functions are applicable to a wide range of texture or soil types. Also, the differences between values estimated with different PTFs results in variability in estimated parameters (Tietje & Tapkenhinrichs, 1993). Tiktak *et al.* (1999) performed an uncertainty analysis of a model describing the deposition of cadmium in the Netherlands and found that the largest source of uncertainty in coarse-textured soils was that introduced by an empirical relationship to estimate cadmium sorption in soils. Satisfaction with the accuracy of PTFs is dependent on the predicted use of the relationships (Pachepsky *et al.*, 1999). For instance, PTFs were used to derive hydraulic properties for the MACRO model and it was proposed that the use of these PTFs may be sufficiently accurate for registration purposes (Jarvis *et al.*, 2000). Inskeep *et al.* (1996) investigated the influence of the resolution of input parameters on predictions of a simple and a more complex pesticide fate model against mean travel times estimated from experimental data. Soil input values were derived from detailed soil profile characterisation or from PTFs applied to textural data obtained from a USDA database whilst weather data were taken from detailed local measurements or were estimated from a climate database and a weather generator. Model predictions compared well with the data when high resolution inputs were used whereas the use of inputs derived from coarse resolution data led

to poorer results. The use of PTFs for estimating soil hydraulic conductivity was identified as a key problem for obtaining reliable LEACHM predictions (Inskeep *et al.*, 1996). An alternative to estimating hydrological parameters using PTFs and databases with basic soil data (*e.g.* texture, organic matter content) is to use soil databases which provide measured values for water retention and hydraulic conductivity curves for specific classes of soil (*e.g.* Wösten *et al.*, 1994; Hollis *et al.*, 1993). Again, the use of such generic data is likely to introduce some uncertainty in the modelling, although the procedure is useful for performing regional assessments and for parameterisation of a model under conditions of limited data.

3.2.2 Use of spatially referenced data

The inaccuracies and uncertainties of maps have interested scholars for decades, if not centuries, but the advent of computers, and geographical information systems (GIS) in particular, have given the topic new impetus and relevance (Zhang & Goodchild, 2002). Although the end products of GIS applications (typically large, attractive multicoloured maps or tables) give an impression of reliability and accuracy on which clear cut decisions can be robustly based (Mowrer, 2000), they may be subject to significant uncertainties. Applications in pesticide fate modelling may draw on spatial data for i) deriving local environmental properties for a given site through the use of spatial databases (*e.g.* Hallet *et al.*, 1995); or, ii) performing spatially distributed pesticide exposure assessments at different scales (see Cryer *et al.* (2001), Shukla *et al.* (2000), Verro *et al.* (2002) and Bach *et al.* (2001) for recent applications at the scales of catchment, county, region and country, respectively). Uncertainty in spatial applications may arise from numerous sources. A detailed discussion of all these sources is beyond the scope of this paper and the reader is invited to refer to dedicated books on the subject for additional information (*e.g.* Goodchild & Gopal, 1989; Heuvelink, 1998a; Zhang & Goodchild, 2002). The major sources of uncertainty are nevertheless briefly discussed below.

Error propagation takes on an extra dimension when models are used in conjunction with GIS (Addiscott & Tuck, 2001). Data stored in GIS may be acquired through digitisation or scanning of existing maps or by interpolation from point observations. The acquisition of data from existing maps through digitisation means that the variability / uncertainty in the maps is being propagated to the digitised version. The digitisation process itself will introduce additional uncertainty in the data (Burrough & McDonnell, 1998). Although maps usually display data with a number of homogeneous zones separated with crisp boundaries, these boundaries, such as the limit between two soil series, are often gradual and mapping units are rarely truly homogeneous in the real world (Heuvelink, 1998a). This may lead to

misclassification of location in map units or positional errors of class boundaries (Tarantola *et al.*, 2002). The spatial misrepresentation of classes using sharp boundaries and uniform units can be accounted for using fuzzy logic (Burrough, 1989). Also, the representation of spatial attributes in the form of uniform grid cells ('mixed pixels'; Burrough & McDonnell, 1998) will carry some uncertainty. Each grid cell typically contains a single value of an attribute, so the information presented (typically a mean value) does not reflect the variability in the attribute displayed. One possible answer to the problem of mixed pixels is to store multiple map layers with *e.g.* mean and variance or percentiles of the attribute. In vector-raster conversion, the mixed pixel problem leads to the dilemma of whether to classify a cell according to the class covering the geographic reference point (*e.g.* the centre of the cell) or according to the dominant class occurring in the cell.

The acquisition of data through interpolation from point observations will have involved extrapolations and generalisation assumptions and there is therefore always a risk of a particular point in space being misclassified. Some of this uncertainty can be accounted for using geostatistics (Webster & Oliver, 2001). The point data are themselves uncertain since they are subject to positional and measurement error (see paragraph above on uncertainty in primary data). Data will carry additional uncertainty if their sampling location is assessed using a Global Positioning System (Mowrer, 2000) or if they have been acquired through remote sensing techniques (Friedl *et al.*, 2001). Although some of the physical, chemical and biological variables measured in the field will vary in time over different scales (*e.g.* days, season, year, decade), datapoints will traditionally be represented as a static value in the GIS and the geographically referenced data will therefore only represent a single occurrence or state in time.

As for a numerical model, uncertainty in the data stored in the GIS may be propagated to the spatial outputs produced through manipulation of the data within the GIS, depending on the sensitivity of the calculations applied to the data. Additional uncertainty may be introduced when data from sources of different quality are combined, when multiple classification classes are combined (Stine & Hunsaker, 2001) or when data are transformed from vector to raster (Burrough, 1986). It is important to realise that the outcome of a GIS exercise at a given scale may be influenced by the spatial resolution of the data (Tarantola *et al.*, 2002). Finke *et al.* (1999) compared results of a soil acidification model parameterised using the 1:1,000,000 soil map and the CORINE land cover database to those obtained when using smaller scale high-resolution independent data for soils (1:50,000 scale) and vegetation for the Netherlands. A total of 69% of the surface area of the Netherlands was misclassified in the large scale map when compared to the highly detailed Dutch data. This was reflected in the modelling, with

40-50% of the variance for predicted nitrate concentrations and exceedance probabilities being attributed to errors in categorical data. A set of recommendations on the estimation of the risk of pesticide leaching at the regional scale using GIS has been provided by Loague *et al.* (1996). Although uncertainty in GIS applications is recognised from a research point of view, hardly any GIS in use can present the user with information about the confidence limits that should be associated with the end product (Heuvelink, 1998a).

3.3 Other sources of uncertainty in the derivation of input parameters

Additional uncertainty factors associated with the derivation of input parameters from primary data include the treatment of replicates and outliers in a dataset, the selection of a representative variable (*i.e.* arithmetic or geometric mean, median), the use of inadequate units or the rounding of values. It is interesting to note that the typical rule used for rounding (round up if the last figure is five or more, round down if it is four or less) that will introduce a bias in the process as for a given random number, the probability of rounding up will be 5/9 while that of rounding down will be 4/9. Measurement errors will perpetuate and propagate during arithmetic operations to obtain the required input parameter. Millstein (1994) recommends that a simultaneous and proportional increase in the precision of all variables will help to control error arising from loss of significant digits in a computation with one numerical division. Particular attention has to be paid to units of parameters since discrepancies in units between pesticide leaching models are common. Another source of uncertainty arises from the lack of detailed information provided in scientific reports used to support model parameterisation. Examples include the reporting of pH values of soil samples without the specification of the background electrolyte and the reporting of the particle size distribution as sand, silt, clay without any reference to particle-size limits. Whereas FAO and some national systems (*e.g.* USA, France) define the boundary between sand and silt fractions at 50 μm, others (*e.g.* UK, Germany) set the division at 60 or 63 μm (Wösten *et al.*, 1999). Procedures for harmonising particle-size distributions from different sources adhering to different classification systems have been reported by Nemes *et al.* (1999).

4 OTHER UNCERTAINTIES RELATED TO THE MODELLING

Modelling the fate of pesticides represents a challenge to any modeller since the behaviour of the compound in soil is influenced by a large number of physical, chemical and biological processes, some of them probably unknown to the scientist. Uncertainty may therefore originate from model error (also referred to as structural error or model inadequacy), the inability of the model to represent given reality accurately even when correct model inputs are

being used. Other sources of uncertainty inherent in the use of pesticide leaching models which are often overlooked are the influence of the choice of a particular model, subjectivity introduced by the modeller, linguistic imprecision and the inappropriate use of concepts implemented in the models. These are discussed below. Sources of uncertainty which were considered to be outside the scope of the present review include: human error (through unstable or biased experimental procedures, interpretation, typing error or the simple variation between people; Stine & Hunsaker, 2001) and uncertainty resulting from the upscaling of models to a larger scale than that they have been developed for; Gaunt *et al.*, 1997).

4.1 Inability of the model to describe experimental observations

The fact that a model is unable to simulate experimental observations even when an appropriate set of model inputs is used has been referred to as structural error, conceptual errors, uncertainties in the conceptual model and model error (Beck *et al.*, 1997). Although model error can be due to errors in translation (the process of converting a scientific concept into a set of equations or computer code; Addiscott & Tuck, 2001), these "errors" will most often originate from the non-inclusion or inappropriate representation of significant processes in the model. Unfortunately, it is seldom easy to quantify the model error (Heuvelink, 1998b). The error is typically revealed through evaluation exercises where the model is tested against experimental data. It is common to say that models are only a simplification of reality, but this adage is particularly true within the context of pesticide fate modelling. Excellent fit for simulations of field leaching data have rarely been reported in the literature and this is not surprising given the complexity of the system to be described. Few existing pesticide fate models if any include satisfactory descriptions of processes such as volatilisation, sorption kinetics and desorption, degradation in the subsoil (Vanclooster *et al.*, 2000) or the extent of sorption and degradation in macropore structures in relation to those in the soil matrix (Jarvis, 2003).

4.2 Selection of a pesticide fate model

Choosing an appropriate model can be a difficult and confusing task, especially for the non-expert (Del Re & Trevisan, 1995; Garen *et al.*, 1999). PRZM is the primary model used in the United States to assess the fate of a pesticide in the unsaturated zone whereas four models (PRZM, PELMO, PESTLA and MACRO) are mainly used in the European Union (FOCUS, 2000). These models have a great deal in common and the current tendency appears to be towards the lessening of their differences (Travis, 2000). However, they also have their own

specificities and comparison studies based on simulations carried out with similar datasets have demonstrated that large differences between models can sometimes be obtained. Vink *et al.* (1997) tested the capability of five leaching models to simulate the water flow and leaching of bromide, simazine and aldicarb from the bottom of an undisturbed clay lysimeter. Soil, hydrological and pesticide properties were derived from the soil used, but no calibrations were carried out. A large variability in the predictions of the five models for pesticide leaching was found, especially with respect to leaching patterns. This variability in model predictions demonstrates that the choice of the model to be used in the modelling is a significant source of uncertainty in pesticide fate modelling. In theory, the selection of a particular model from those recommended for pesticide registration should result from a balance between considerations regarding the model (*i.e.* model capabilities, data required, extent of validation, ease of use and documentation; Russell *et al.*, 1994), the objectives of the modelling exercise (*e.g.* predictive simulation, simulation of field data, calibration) and the availability of data for model parameterisation (Di & Aylmore, 1997), but external factors such as known preferences of national registration authorities also come into play (Travis, 2000). A comparison of process descriptions to guide selection of a leaching model has been presented for calibration applications (Dubus *et al.*, 2002).

4.3 Modeller subjectivity

Brown *et al.* (1996) reported the simulation of pesticide leaching in a sandy loam soil by five modellers using three pesticide fate models (LEACHP, PRZM-2 and VARLEACH). Although modellers were provided with the same information (site, weather, soil, crop, pesticide properties and application data), no two simulations with any of the three models were identical. Predictions for total residues in soil 7.5 months after application varied between the five modellers with CVs of 8 to 58% depending on the model used. When predicting the maximum pesticide concentration at 1 m depth, the series of figures obtained for the five individuals were: 0, 0, 0.01, 0.02 and 0.11µg/l. User-dependency of the results was attributed to a number of key input parameters which could not be derived from the experimental information provided and were thus open to considerable subjectivity according to the experience and knowledge of the individual modeller (Brown *et al.*, 1996). Other recent studies with a larger number of models and datasets have confirmed the user-dependent variability of model outputs (*e.g.* Tiktak, 2000; Boesten and Gottesbüren, 2000). Boesten (2000) reported a ring test exercise in which 18 pesticide fate modellers received raw data from laboratory degradation studies and were asked to derive DT50 values (at 10°C) and sorption distribution coefficients for ethoprophos and bentazone, respectively. Calculated

DT50 values ranged from 92 to 346 days (mean 191 days, CV 29%) and from 33 to 204 days (mean 83 days, CV 46%) for the two compounds (Figure 4). Variation in sorption distribution coefficients was smaller with coefficients of variation of 21 and 13% for the two pesticides. Variability in the derivation of environmental properties for these two pesticides was attributed, in decreasing importance, to the expert judgement introduced in establishing the relationship between transformation rate and soil temperature, the inclusion/exclusion of data for incubation times larger than 100-200 days and the use of different fitting procedures. Boesten (2000) suggested the subjective influence of the model user was much greater than the variability introduced by the use of different models. The uncertainty is likely to be magnified for users of the more complex simulation models, who may be specialists in some aspects of the subject, but relatively inexperienced in others (Jarvis *et al.*, 2000). The large majority of ring test / validation studies carried out with pesticide leaching models involved individuals who had significant experience in the use of leaching models. It is expected that simulations carried out by people who are less experienced with specific models will carry an even larger uncertainty. The uncertainty in model predictions resulting from the influence of the modeller can be reduced by providing more detailed guidance on model parameterisation (Brown *et al.*, 1996) and by discussing existing procedures for the derivation of input parameters between modellers (Boesten, 2000).

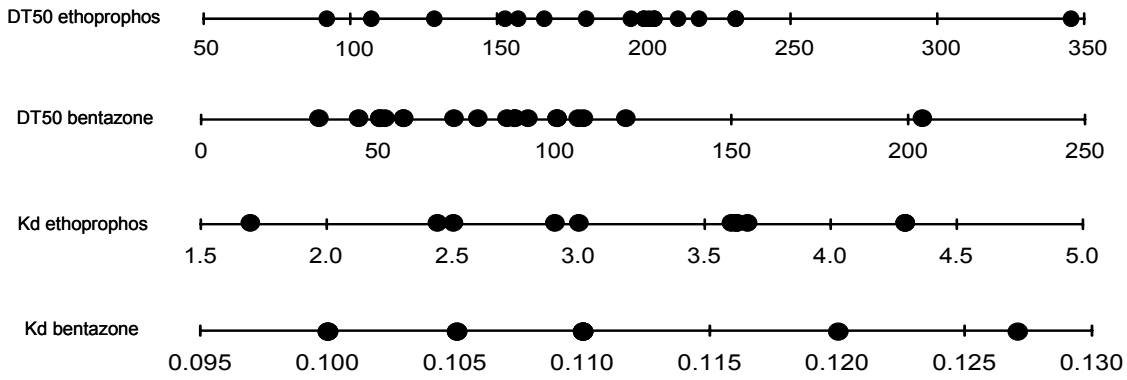


Figure 4. Variability in the selection of DT50 and Kd values for ethoprophos and bentazone by a range of modellers.
Selection was based on the same degradation and sorption datasets for all individuals (adapted from Boesten, 2000)

4.4 Inadequate use of concepts implemented in models

Perhaps one of the largest sources of error in modelling is the use of concepts which do not apply to a particular modelling situation. Examples of pitfalls are the inadequate use of the Koc (or Kom) approach and the parameterisation of pesticide decay. Sorption is one of the

key processes regulating the transport, transformation and biological effects of organic chemicals in soils and aquatic systems (Calvet, 1989). Sorption is often characterised by the Koc, the sorption distribution coefficient normalised to organic carbon, which is defined as:

$$Koc = \frac{Kd}{OC} * 100 \qquad [1]$$

where Kd is the linear sorption distribution coefficient determined in batch equilibrium experiments, usually expressed in l/kg,

OC is the organic carbon content, expressed in %.

The Koc was introduced to reflect the linear relationship between Kd and OC which was demonstrated for non-ionic compounds (Hamaker and Thompson, 1972) because this variable has the virtue of being broadly independent of any particular soil (Hamaker, 1975). Nevertheless, the Koc concept tends to be used by modellers for all molecules regardless of their ionisation status. The Koc approach is usually invalid for ionisable compounds (Wauchope *et al.*, 2002) since there is often no clear influence of organic matter on sorption (Shimizu *et al.*, 1992; Ukrainczyk and Ajwa, 1996). Another pitfall relates to the parameterisation of degradation. Most pesticide leaching models use first-order kinetics to describe the degradation of pesticides in soil. If a first-order model is selected in the model, then it is a necessity that DT50 values (or degradation rates) derived by first-order kinetics are used, even if other equations fit the data better. Inputing DT50 values in a model that was derived with another equation is a common mistake since the modeller can be provided with DT50 values without further information on the way these were derived. The inadequate application of model concepts is likely to contribute significantly to the overall uncertainty of the modelling and should obviously be avoided. It is essential that people who use models are aware of the underlying concepts and data requirements. Adequate documentation, help screens, version control and the provision of training are all possible ways to address this.

4.5 Linguistic imprecision

Linguistic imprecision was identified as a source of uncertainty by Morgan & Henrion (1990), as we often refer to events or quantities with imprecise language. For instance, the “water flow rate in the Thames” is an ill-specified quantity, which may be better qualified as “the net water flow rate of the Thames at the Westminster bridge in London at noon on 12 September 1998” (Morgan & Henrion, 1990). Another side of linguistic imprecision which might lead to uncertainty is the description of two different concepts by the same term. Sorption data originating from batch equilibrium studies are often described using the non-linear Freundlich equation:

$$X = Kf\, C^{1/n} \qquad [2]$$

where X is the amount adsorbed,

C is the concentration of pesticide in solution,

and Kf and 1/n are coefficients (referred to as the Freundlich distribution coefficient and exponent) .

It is common practice to normalise the Kf coefficient to organic matter and to report this new variable as Koc. As pointed out earlier, the initial definition of the Koc refers to the linear distribution coefficient. There is therefore possible confusion on the term Koc which could either represent the linear or Freundlich distribution coefficient normalised to organic carbon. Differences in the values for the two different Koc will be maximum for compounds which exhibit strong non-linear sorption. Model developers do not always help to clarify the situation. Freundlich sorption was introduced in MACRO when the model was updated from version 4.0 to 4.1 in 1998, but ZKD, the name of the parameter for distribution coefficients remained the same. The fact that a user has to input Kf coefficients in a parameter called ZKD can be somewhat confusing and might lead to inadequate parameterisation of the model.

A second examples relates to the use of the term "field capacity" to refer to a specific point on the water retention curve to which a wetted, freely-draining profile is considered to drain within a defined period. Uncertainty in the modelling might arise because the reference tension varies between countries: -5 kPa in the UK, -10 kPa in the Netherlands, -33 kPa in the US, for instance. Similarly, different leaching models implement different definitions (-5 kPa in MACRO, -33 kPa in PRZM, -10 kPa in PESTLA) and the model user can introduce significant uncertainty if he does not select the appropriate tension.

4.6 Model calibration

Model calibration is a cornerstone in pesticide fate modelling (Dubus *et al.*, 2002). It is sometimes considered that model calibration can help reduce the uncertainty in the modelling through the derivation of values for input parameters that help to improve the simulation of experimental data, but this might not always be the case. The calibration might be ill-posed (Carrera & Neuman, 1986) and uncertainty will originate from the fact that multiple combinations of input parameters will provide a similar fit to the experimental data (concept of equifinality; Beven and Binley, 1992). In other words, there is no clear optimum parameter set that calibrates the model. The predictive uncertainty arising from the lack of a unique solution to the calibration problem can be accounted for using dedicated procedures such as the Generalized Likelihood Uncertainty Estimation (Beven, 2001), the Pareto Optimal Set

procedure (Yapo *et al.*, 1998) or the non-linear predictive analyser of the PEST inverse modelling package (Doherty, 2002). Although these approaches will provide a confidence interval for each optimised parameter and for subsequent extrapolations on the basis of these parameters, the uncertainty estimates provided will be dependent on subjective choices, such as the selection of an objective function or the limit from which you consider that the model is not calibrated anymore (Beven, 2001). Also, some authors argue that the error measures determined for one set of conditions cannot be extrapolated to other applications (Heuvelink, 1998b). Another reason which may prevent a decrease in uncertainty resulting from calibration is 'parameter lumping' (Dubus *et al.*, 2002). Parameter lumping occurs when there is enough flexibility in the modelling system to allow for changes in specific parameters to compensate for inaccuracies, uncertainties and limitations associated with experimental data, modelling and calibration procedures. Lumped parameters can only be obtained by calibration and will have lost their physical, chemical or biological definition (Jansen, 1998). As noted by Melching (1995), the fit to the data using erroneous and biased parameter values obtained through calibration may not necessarily be greatly different from that obtained with true (*i.e.* assumed hypothetical values) data and parameter values because of the curve-fitting nature of the calibration process, which will provide, in effect, an adjustment factor and bias. In cases where lumping occurs, calibration will result in an increase in the parameter uncertainty, although this may not be transparent to the modeller.

There is continuing debate in the literature about the relationship between calibration and extrapolation. For some, successfully calibrating a model demonstrates its ability to simulate a specific set of conditions and allows one to extrapolate to other points in space and time (*e.g.* Matsui *et al.*, 2002). For others, the calibration of complex deterministic models tends to be specific to the conditions at the site for which experimental data were collected and no extrapolation should be carried out with a calibrated parameter set. For instance, Heuvelink (1998b) considers that when model development involves calibration of model parameters, then model error will almost certainly increase when the model is extrapolated to another area or to another time period. Calibration may therefore decrease or increase uncertainty depending on the problem at hand. Calibration issues in the context of the analysis of uncertainty of mathematical models of water quality have been discussed by Beck (1987).

5 DISCUSSION AND IMPLICATIONS

5.1 Contribution of the individual sources of uncertainty to the overall uncertainty in pesticide fate modelling

The listing of sources of uncertainty clearly shows that pesticide fate modelling is laced with uncertainty. Most of these uncertainties will significantly affect the outcome of environmental exposure assessments, although their influence will vary. A priority should therefore be to estimate the magnitude of the uncertainties listed above and whether they transfer into uncertainty in the estimation of exposure. This task is made difficult because only a few uncertainties are easily quantified, many or most are quantified with difficulty, and several may not be quantifiable at all (Oreskes, 1998). Also, parameter uncertainty may be transferred differently through the modelling, from suppression to large exaggeration (Addiscott & Tuck, 2001) depending on the sensitivity of the model. For these reasons, very few attempts to differentiate between the contributions of the different sources of uncertainty to the overall uncertainty in exposure modelling have been reported (Zhang *et al.*, 1993; Nofziger *et al.*, 1994; Loague *et al.*, 1996). Zhang *et al.* (1993) investigated the influence on CMLS predictions of the uncertainty in model input parameters and year-to-year rainfall variability using first-order second moment analysis and Monte Carlo simulation. Both sources of uncertainty were found to have comparable effects on the magnitude of the uncertainty of model predictions and the combined variabilities appeared to magnify the overall uncertainty. Nofziger *et al.* (1994) considered the influence of the uncertainty in weather, soil and pesticide properties on predictions for bentazone leaching. Uncertainty contributions were found to vary between the two years simulated and the overall uncertainty in predictions when combining the three classes of uncertainty was 28 orders of magnitude. Loague (1994) investigated the influence of reductions in data uncertainties on the overall uncertainty levels predicted in a number of uncertainty analyses. The approach is similar in its concept to a cost/benefit analysis which would identify where additional resources could be allocated to allow a reduction of the uncertainty in the assessment of pesticide leaching. As noted by Loague *et al.* (1996), more information does not necessarily translate into improved model performance and one needs to consider 'when enough is enough' (James & Gorelick, 1994). A useful step in gaining an understanding of the uncertainty in pesticide fate modelling resulting from the uncertainty in model input parameters is to assess the extent of the latter. Such an approach was recently undertaken by an IUPAC working group which derived information on the uncertainty in sorption in the form of rules of thumb (Wauchope *et al.*, 2002). The authors consider that i) the batch experiment probably varies from the true

average Kd in the field of the same soil by a factor of two; ii) the variability in Kd in the field is to be attributed to variation of the organic matter content in the field and of the organic matter itself and typically has a CV of about 50%; iii) a Kd determined for different soils will vary by approximately one order of magnitude; iv) a CV of 30-60% is common in multi-soil studies and reflects the variability in the sorption capacity of the organic matter and in the measurement of the organic carbon content; and, v) Koc values reported for different studies with multiple soils are expected to vary by an order of magnitude. These kinds of variability will render modelling predictions largely uncertain. The application of a similar approach for other key model input parameters would be useful. A number of sensitivity analyses have demonstrated that predictions of pesticide fate models for leaching will mainly be influenced by sorption and degradation parameters (Boesten, 1991; Soutter & Musy, 1998), although a number of models have also show large sensitivities to hydrological parameters (Dubus & Brown, 2002; Wolt *et al.*, 2002).

Model error, also known as structural error or model inadequacy, is the most basic form of error in modelling (Cohen *et al.*, 1995; Addiscott & Tuck, 2001), but is notoriously difficult to estimate. In chemistry, the analytical bias can be estimated independently by the use of certified reference materials (Ramsey, 1998) or by comparing a given method to others (Heber *et al.*, 1998). An attempt to transpose this general evaluation approach to environmental fate modelling has occasionally been made through the comparison of results of simple models against those obtained with more complex codes to assess model error. For instance, Loague *et al.* (1996) compared the results of the simple 'Attenuation Factor' index method with the more conceptually rigorous PRZM model. However, the procedure is limited since relative agreement with a complex model does not guarantee agreement with behaviour in the field. The more complex model, although considered a reference, could be broadly unable to simulate what happens in the field. For these reasons, model error is traditionally assessed using a comparison of model predictions and experimental data (model evaluation, model testing or model 'validation'). The inherent capability of models to describe field data can be assessed by simulating the experimental data using information on the study site but without knowledge of the pesticide concentrations found at the site ('predictive' or 'blind' simulation). It is increasingly recognised that pesticide leaching models cannot accurately simulate field data in predictive mode, partly because of variability and uncertainty aspects (Vanclooster *et al.*, 2000). The difficulty in assessing model error for blind simulations is that inadequate fits to the data obtained have understandably not been reported in the literature. Any estimate of model error derived through review activities is probably biased and it may be better to rely on expert judgement. The most popular technique for evaluating model error has therefore been to calibrate the model against the data to be

simulated (Dubus *et al.*, 2002). This approach is not flawless as the calibration problem in pesticide fate modelling tends to be ill-posed (Neumann & Carrera, 1986) and multiple combinations of model input parameters will provide a similar fit to the data (Beven & Binley, 1992). Uncertainty propagation arising from the estimation of parameters by calibration has been the subject of much research, especially within the context of water quality modelling (Beck, 1987). Whether calibration is used or not, the testing of the model is likely to be as much an evaluation of the capabilities of the modeller as of those of the model (Brown *et al.*, 1996; Boesten, 2000). The variability in modelling results introduced by the modeller was found to be of the same order of magnitude as that associated with the field results (Brown *et al.*, 1996). This source of uncertainty has been shown to contribute to the modelling uncertainty to such an extent that it may prevent the evaluation of model concepts (Tiktak, 2000).

The contribution of sources of uncertainty other those in model input parameters is largely unknown and further research is clearly required in this area. It can be strongly suspected that, as noted by Bailar (1988; reported by Costanza *et al.*, 1992), random variability is simply swamped by other kinds of uncertainties. The implications of focusing on a limited number of sources of uncertainty are discussed below.

5.2 The uncertainty iceberg

Techniques which have been applied to propagate uncertainty in pesticide fate modelling include differential analysis (Loague, 1991; Loague *et al.*, 1996; Freissinet *et al.*, 1999; Li *et al.*, 1998; Diaz-Diaz *et al.*, 1999), Fourier amplitude sensitivity test (Fontaine *et al.*, 1992), Monte Carlo analysis (Carsel *et al.*, 1988; Zhang *et al.*, 1993; Nofziger *et al.*, 1994; Soutter and Pannatier, 1996; Soutter and Musy, 1998, 1999; Zacharias *et al.*, 1999; Ma *et al.*, 2000; Trevisan *et al.*, 2001; Dubus & Brown, 2002; Carbone *et al.*, 2002) and fuzzy logic (Freissinet *et al.*, 1998; 1999). These uncertainty analyses have investigated the effects of input uncertainty on model predictions and therefore implicitly assume that i) the major sources of uncertainty are those associated with model input parameters; ii) the model is structurally correct; iii) modeller subjectivity is minimal; and, iv) an adequate parameterisation of a model is possible. Jansen (1998) noted that in the best case, these uncertainty analyses can only provide an optimistic estimate of prediction error, since model errors cannot be fully captured by such analyses and because of restrictions in the sources of uncertainty considered. As noted earlier, model error is notoriously difficult to estimate and cannot be easily implemented in traditional uncertainty assessment techniques. Modeller subjectivity can be

assessed by asking different individuals to simulate the same modelling situation (Brown *et al.*, 1996; Boesten, 2000; Tiktak, 2000), but the magnitude of the differences in model predictions will remain to some extent specific to the particular modelling exercise considered. Somewhat perversely, the selection and implementation of techniques designed to account for uncertainty are themselves subject to significant uncertainty. For instance, in the specific case of Monte Carlo based probabilistic assessments, aspects which are likely to influence overall results include the selection of input parameters to be included in the analysis (Nofziger *et al.*,1994), the type (Mosleh, 1986; Brattin *et al.*, 1996) and parameterisation (Brattin *et al.*, 1996) of probability distribution functions attributed to input parameters, the absence or presence of correlation between variables (Iman & Conover, 1982; Smith *et al.*, 1992), the extent of the correlations considered, the sampling scheme used (Saltelli *et al.*, 2000) and the seed number used in the sampling (Dubus & Janssen, 2003).

It is generally considered important to differentiate in quantitative risk analyses between stochastic variability (or 'Type A' uncertainty) and uncertainty due to the lack of knowledge (or incertitude or 'Type B' uncertainty) (IAEA, 1989; Hoffman & Hammonds, 1994; US EPA, 1996). This is because variability is an inherent and irreducible property of the scenario being evaluated, while incertitude is not an inherent property and can (at least in theory) be reduced by collecting additional data or information or performing additional analysis (Brattin *et al.*, 1996). Real variability will not change as a result of better or more extensive measurements although it may be more accurately known (Hattis & Burmaster, 1994). Although in theory every component of a risk assessment model could be represented as having both variability and incertitude, this duality is not taken into account and the two uncertainty classes are considered to be exclusive (Kelly & Campbell, 2000). The approach typically results in a two-dimensional (also known as 'nested' or 'double loop') Monte Carlo analysis being carried out where the contribution of each of these broad classes is assessed. Thompson & Graham (1996; reported by Cullen & Frey, 1999) presented a classification of the different types of approaches for differentiating between variability and incertitude in Monte Carlo uncertainty analysis. Mathematical operations on probabilities can be utilised without distinction between whether the distribution arises from variability or incertitude or whether the basis for the distribution is subjective or data driven (Hoffman *et al.*, 1999). Two-dimensional Monte Carlo analysis should be conducted with care, as typical applications which have been reported in the literature have only focused on a limited number of sources of uncertainty and may have ignored significant others. Kelly & Campbell (2000) reported that the uncertainty they considered was generally restricted to those types of uncertainty that can be modelled with probability distributions, but also noted that model error is typically the largest source of uncertainty in ecological risk assessments. Results from two-dimensional

uncertainty analyses might therefore grossly underestimate uncertainty and convey a false sense of accuracy.

A simple but useful representation of the challenges facing the pesticide fate modelling community and those who base part of their decision-making on the results of pesticide leaching models is presented in Figure 5. In the past (iceberg [a]), the uncertainty associated with pesticide fate modelling was largely ignored by modellers and decision-makers and the uncertainty iceberg was largely concealed. The last 10 years have seen an increasing recognition of variability in the natural environment, in particular in model sorption and degradation parameters, and the influence this may have on the uncertainty in modelling predictions and on the subsequent decision-making. Consequently, procedures to account for uncertainties in model input, such as Monte Carlo analysis, have been recommended and applied (iceberg [b]). A listing of the sources of uncertainty in pesticide fate modelling reveals that many are not being accounted for by traditional approaches to uncertainty analysis and that the underwater section of the uncertainty iceberg may be much larger than anticipated (icebergs [c] and [d]). The significant presence of submerged uncertainty (iceberg [d]) would weaken results of probabilistic modelling if it only dealt with the emerged part of uncertainty. More research is urgently required to estimate the ratio emerged / submerged matter currently implemented in probabilistic modelling exercises. As noted by Cornell (1972; reported by Melching, 1995), it is important that we avoid the tendency to model only those probabilistic aspects that we think we know how to analyse. It is far better to have an approximate model of the whole problem than an exact model of only a portion of it.

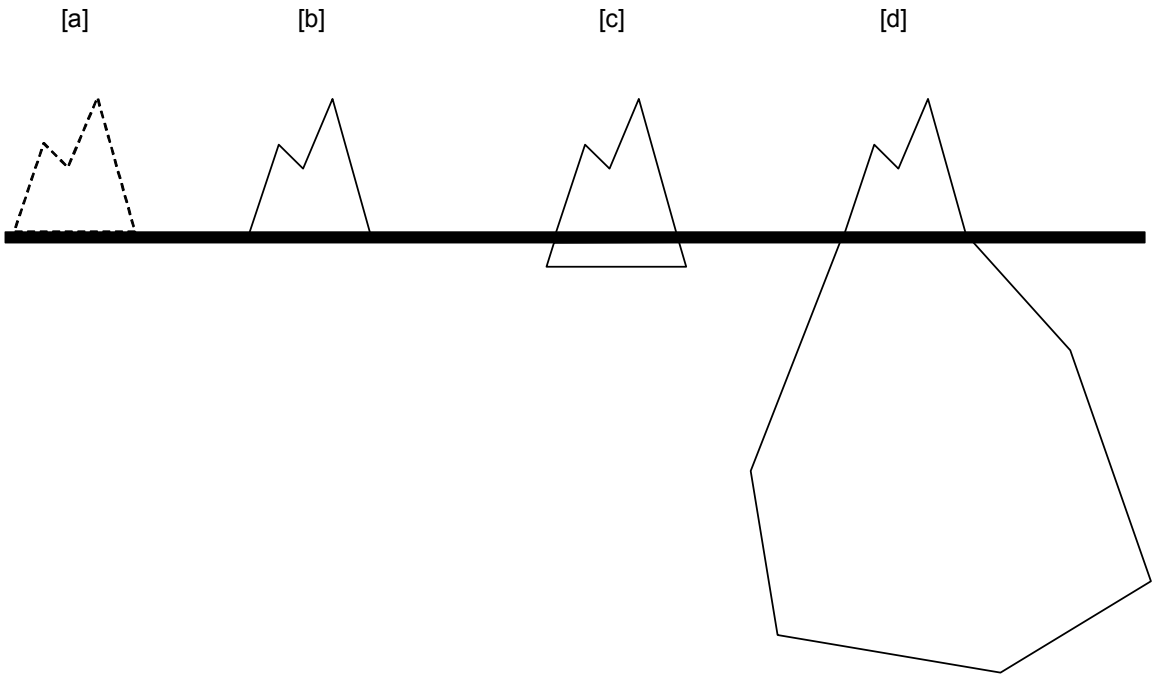


Figure 5. The uncertainty iceberg

Although uncertainty in pesticide fate modelling has been ignored in the past [a], there has been a number of attempts to quantify uncertainty over the last 10 years [b]. The challenge is now to ascertain whether the uncertainty which is accounted for represents a large [c] or small [d] proportion of the overall uncertainty in pesticide fate modelling.

6 CONCLUSIONS

Research on uncertainty is concerned with the identification of uncertainty, its description and with predicting its effects on the outputs of the analysis. The present paper has demonstrated that errors and uncertainties accumulate in pesticide fate modelling in various forms and disguises. Further research is clearly required to i) assess the magnitude of the different sources of uncertainties affecting pesticide fate modelling; and, ii) integrate into probabilistic modelling exercises those significant sources of uncertainty that are not currently considered (*e.g.* model error, modeller subjectivity). There is no doubt that assessing the uncertainty in inputs of soil and water models (Jansen, 1998) and model error (Heuvelink, 1998b) is a difficult task. Particular complications arise from the facts that: i) the most common approach in numerous other fields of science has been to only consider the uncertainty in model input parameters; and, ii) the effect of structural errors cannot be quantified when the 'structurally correct' model is unknown, and, strictly speaking, this is always the case with any natural system (Jansen, 1998). Dealing with uncertainty is not an individual matter, but rather a team effort. Working towards a better quantification of uncertainties and their reduction should be

a common effort across the whole community of people who have a direct or indirect input into the study of the fate of pesticides, from individuals involved in designing soil sampling strategies in the field right through to the modellers and eventually to decision-makers.

Deterministic models implicitly assume that data on which the parameterisation is based are error-free and parameters and model structures are considered as completely known (Tarantola *et al.*, 2002). Although there have been some attempts at putting variability at the heart of the modelling itself (Jury & Gruber, 1989; Gustafson & Holden, 1990; van der Zee & Boesten, 1991; Wu *et al.*, 1997), deterministic programmes are likely to remain the main tools for simulating the fate of pesticides for the foreseeable future. Accounting for uncertainty in pesticide fate modelling is still in its infancy, but progress in this domain is crucial if robust estimates of uncertainty in model predictions are to be obtained. As pointed out by a number of authors (Jury *et al.*, 1987; Greig-Smith, 1992; Costanza *et al.*, 1992), uncertainty may limit the use of a model for making regulatory decisions unless the uncertainty is figured into the decision making process. Disclosure of uncertainty and characteristics of the uncertainty such as the sources and likely magnitudes is needed for the decision maker to understand how confident he or she can be about the decisions that are being made.

Acknowledgements

The authors gratefully acknowledge the funding of this work by the UK Department for Environment, Food and Rural Affairs on behalf of the Pesticides Safety Directorate (project PL0548). The opinions expressed in this paper are those of the authors and may not reflect those of the study sponsors. Thanks are due to G.B.M. Heuvelink (University of Amsterdam, The Netherlands) for commenting on the section on uncertainty in spatially referenced data.

References

Adams J. (1998). A 'Richter scale for risk'? Scientific management of uncertainty versus management of scientific uncertainty. Interdisciplinary science reviews, 23:146-155.

Addiscott T.M. & Wagenet R.J. (1985). Concepts of solute leaching in soils – a review of modelling approaches. Journal of Soil Science, 36:411-424.

Addiscott T.M. & Tuck G. (2001). Non-linearity in modelling soil processes. European Journal of Soil Science, 52:129-138.

Ahmad R., Kookana R.S. & Alston A.M. (2001). Syringe filtration as a source of error in pesticide residue analysis in environmental samples. Bulletin of Environmental Contamination and Toxicology, 66:313-318.

Armstrong A.C., Portwood A.M., Leeds-Harrison P.B., Harris G.L. & Catt J.A. (1996). The validation of pesticide leaching models. Pesticide Science, 48:47-55.

Bach M., Huber A. & Frede H.G. (2001). Modeling pesticide losses from diffuse sources in Germany. Water Science & Technology, 44:189-196.

Bailar III J.C. (1988). Scientific inferences and environmental problems. The uses of statistical thinking. Institute for Environmental Studies, University of North Carolina at Chapel Hill.

Barriuso E. & Calvet R. (1992). Soil type and herbicides adsorption. International Journal of Environmental and Analytical Chemistry, 46:117-128.

Beck M.B. (1987). Water quality modelling: a review of the analysis of uncertainty. Water Resources Research, 23:1393-1442.

Beck A.J., Harris G.L.,Howse K.R., Johnston A.E. & Jones K.C. (1996). Spatial and temporal variation of isoproturon residues and associated sorption/desorption parameters at the field scale. Chemosphere, 33:1283-1295.

Beck M.B., Ravetz J.R., Mulkey L.A. & Barnwell T.O. (1997). On the problem of model validation for predictive exposure assessments. Stochasic Hydrology and Hydraulics, 11:229-254.

Beltran J., Hernandez F., Lopez F.J. & Morell I. (1995). Study of sorption processes of selected pesticides on soils and ceramic porous cups used for soil solution sampling. International Journal of Environmental and Analytical Chemistry, 58:287-303.

Beulke S. & Brown C.D. (2001). Evaluation of methods to derive pesticide degradation parameters for regulatory modelling. Biology and Fertility of Soils, 33:558-564.

Beulke S., Brown C.D. & Dubus I.G. (2001). Characterisation of sorption for the modelling of pesticide fate. In: A. Walker (Ed.), Pesticide behaviour in soils and water, BCPC Symposium Proceedings No. 78, Farnham, UK, pp. 51-56.

Beven K. (2001). Rainfall-runoff modelling: the primer. John Wiley & Sons Ltd., Chichester, UK.

Beven K. & Binley A. (1992). The future of distributed models – model calibration and uncertainty prediction. Hydrological processes, 6:279-298.

Bosch D.D. & West L.T. (1998). Hydraulic conductivity variability for two sandy soils. Soil Science Society of America Journal, 62:90-98.

Boesten J.J.T.I. (1991). Sensitivity analysis of a mathematical model for pesticide leaching to groundwater. Pesticide Science, 31:375-388.

Boesten J.J.T.I. (2000). Modeller subjectivity in estimating pesticide parameters for leaching models using the same laboratory dataset. Agricultural Water Management, 44:389-409.

Boesten J.J.T.I. & van der Linden A.M.A. (1991). Modeling the influence of sorption and transformation on pesticide leaching and persistence. Journal of Environmental Quality, 20:425-435.

Boesten J.J.T.I. & Gottesbüren B. (2000). Testing PESTLA using two modellers for bentazone and ethoprophos in a sandy soil. Agricultural Water Management, 44:283-305.

Bouma J. (1989). Using soil survey data for quantitative land evaluation. Advances in Soil Science, 9:177-213.

Brattin W.J., Barry T.M. & Chiu N. (1996). Monte Carlo modeling with uncertain probability density functions. Human and Ecological Risk Assessment, 2:820-840.

Brown C.D., Baer U., Günther P., Trevisan M. & Walker A. (1996). Ring test with the models LEACHP, PRZM-2 and VARLEACH: variability between model users in prediction of pesticide leaching using a standard data set. Pesticide Science, 47:249-258.

Brown C.D., Bellamy P.H. & Dubus I.G. (2002). Prediction of pesticide concentrations found in rivers in the UK. Pest Management Science, 58:363-373.

Brus D.J. & de Gruijter J.J. (1997). Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma, 80:1-59.

Burrough P.A. (1986). Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford, UK.

Burrough P.A. (1989). Fuzzy mathematical methods for soil survey and land evaluation. Journal of Soil Science, 40:477-492.

Burrough P.A. & McDonnell R.A. (1998). Principles of geographical information systems. Oxford University Press, Oxford, UK.

Calvet R. (1989). Adsorption of organic chemicals. Environmental Health Perspectives, 82:145-177.

Cambardella C.A., Moorman T.B., Novak J.M., Parkin T.B., Karlen D.N., Turco R.F. & Konopka A.E. (1994). Field-scale variability of soil properties in Central Iowa soils. Soil Science Society of America Journal, 58:1501-1511.

Carbone J.P., Havens P.L. & Warren-Hicks W. (2002). Validation of Pesticide Root Zone Model 3.12: employing uncertainty analysis. Environmental Toxicology and Chemistry, 21:1578-1590.

Carrera J. & Neuman S.P. (1986). Estimation of aquifer parameters under transient and steady state conditions. 2. Uniqueness, stability and solution algorithms. Water Resources Research, 22:211-227.

Carsel R.F., Parrish R.S., Jones R.L., Hansen J.L. & Lamb R.L. (1988). Characterizing the uncertainty of pesticide leaching in agricultural soils. Journal of Contaminant Hydrology, 2:111-124.

Chaubey I., Haan C.T., Salisbury J.M. & Grunwald S. (1999). Quantifying model output uncertainty due to spatial variability of rainfall. Journal of the American Water Resources Association, 35:1113-1123.

Clausen L. (2000). Retention of pesticides in filter membranes. Journal of Environmental Quality, 29:654-657.

Cohen S.Z., Wauchope R.D., Klein A.W., Eadsforth C.V. & Graney R. (1995). Offsite transport of pesticides in water: mathematical models of pesticide leaching and runoff. Pure & Applied Chemistry, 67:2109-2148.

Coquet Y. & Barriuso E. (2002). Spatial variability of pesticide adsorption within the topsoil of a small agricultural catchment. Agronomie, 22:389-398.

Cornell C.A. (1972). First-order analysis of model and parameter uncertainty. In: Proceedings of the International Symposium on Uncertainties in Hydrology and Water Resources Systems. University of Arizona, Tucson, Arizona, Vol. 3, 1245-1272.

Costanza R., Funtowicz S.O. & Ravetz J.R. (1992). Assessing and communicating data quality in policy relevant research. Environmental Management, 16:121-131.

Cryer S.A., Fouch M.A., Peacock A.L. & Havens P.L. (2001). Characterizing agrochemical patterns and effective BMPs for surface waters using mechanistic modeling and GIS. Environmental Modelling & Assessment, 6:195-208.

Cullen A.C. & Frey H.C. (1999). Probabilistic techniques in exposure assessment. Plenum Press, New York.

Del Re A.A.M. & Trevisan M. (1995). Selection criteria of xenobiotic leaching models in soil. European Journal of Agronomy, 4:465-472.

Di H.J. & Aylmore L.A.G. (1997). Modeling the probabilities of groundwater contamination by pesticides. Soil Science Society of America Journal, 61:17-23.

Diaz-Diaz R., Loague K. & Notario J.S. (1999). An assessment of agrochemical leaching potentials for Tenerife. Journal of Contaminant Hydrology, 36:1-30.

Doherty J. (2002). PEST – model independent parameter estimation. Watermark Computing, Corinda, Australia.

Dubus I.G., Barriuso E. & Calvet R (2001). Sorption of weak organic acids in soils: clofencet, 2,4-D and salicylic acid. Chemosphere, 45:767-774.

Dubus I.G. & Brown C.D. (2002). Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. Journal of Environmental Quality, 31:227-240.

Dubus I.G., Beulke S. & Brown C.D. (2002). Calibration of pesticide leaching models: critical review and guidance for reporting. Pest Management Science, 58:745-758.

Dubus I.G. & Janssen P.H.M. (2003). Issues of replicability in Monte Carlo modelling: a case study with a pesticide leaching model. Submitted to Environmental Toxicology and Chemistry in September 2002.

ECOFRAM (1999). http://www.epa.gov/oppefed1/ecorisk/

Elabd H., Jury W.A. & Cliath M.M. (1986). Spatial variability of pesticide adsorption parameters. Environmental Science & Technology, 20:256-260.

EUPRA (2001). Probabilistic risk assessment for pesticides in Europe. Implementation and research needs. Report from the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products, The Netherlands, June 2001.

Finke P.A., Wladis D., Kros J., Pebesma E.J. & Reinds G.J. (1999). Quantification and simulation of errors in categorical data for uncertainty analysis of soil acidification modelling. Geoderma, 93:177-194.

FOCUS (2000). FOCUS groundwater scenarios in the EU plant protection review process. EC document reference Sanco/321/2000 re. 2, 202 pp.

Fontaine D.D., Havens P.L., Blau G.E. & Tillotson P.M. (1992). The role of sensitivity analysis in groundwater risk modelling for pesticides. Weed Technology, 6:716-724.

Flury M., Yates M.V., Jury W.A. & Anderson D.L. (1998). Variability of solute transport in field lysimeters. In: Führ F., Hance R.J., Plimmer J.R. & Nelson J.O. (eds). The lysimeter concept. ACS Symposium series 699, pp. 65-75.

Franke H.J. & Teutsch (1994). Stochastic simulation of the regional pesticide transport including the unsaturated and the saturated zone. Ecological Modelling, 75/76:529-539.

Freissinet C., Erlich M. & Vauclin M. (1998). A fuzzy logic-based approach to assess imprecisions of soil water contamination modelling. Soil & Tillage Research, 47:11-17.

Freissinet C., Vauclin M. & Erlich M. (1999). Comparison of first-order analysis and fuzzy set approach for the evaluation of imprecision in a pesticide groundwater pollution screening model. Journal of Contaminant Hydrology, 37:21-43.

Friedl M.A., McGwire K.C. & McIver D.K. (2001). An overview of uncertainty in optical remotely sensed data for ecological applications. In: Spatial uncertainty in ecology, Hunsaker C.T., Goodchild M.F., Friedl M.A. & Case T.J. (Eds). Springer-Verlag, New York, USA. pp. 258-283.

Garen D., Woodward D. & Geter F. (1999). A user agency's view of hydrologic, soil erosion and water quality modelling. Catena, 37:277-289.

Gaunt J.L., Riley J., Stein A & Penning de Vries F.W.T (1997). Requirements for effective modelling strategy. Agricultural Systems, 54:153-168.

Gerstl Z. (2000). An update on the Koc concept in regard to regional scale management. Crop Protection, 19:643-648.

Goodchild M. & Gopal S. (1989). Accuracy of spatial databases. Taylor & Francis, , London, UK.

Goodrich D.C., Faures J.M., Woolhiser D.A., Lane L.J. & Sorooshian S. (1995). Measurement and analysis of small-scale convective storm rainfall variability. Journal of Hydrology, 173:283-308.

Greig-Smith P.W. (1992). A European perspective on ecological risk assessment, illustrated by pesticide registration procedures in the United Kingdom. Environmental Toxicology & Chemistry, 11:1673-1689.

Gustafson D.I. & Holden L.R. (1990). Nonlinear pesticide dissipation in soil: a new model based on spatial variability. Environmental Science & Technology, 24:1032-1038.

Hallet S.H., Thanigasalam P. & Hollis J.M. (1995). SEISMIC: a desktop information system for assessing the fate and behaviour of pesticides in the environment. Computers and Electronics in Agriculture, 13: 227-242.

Hamaker J.W. & Thompson J.M. (1972). Adsorption. In: Organic chemicals in the soil environment, vol.1. C.A. Goring & J.W. Hamaker (Eds), Marcel Dekker, New York, 49-143 pp.

Hamaker J. W. (1975). The Interpretation of Soil Leaching Experiments. In: Environmental Dynamics of Pesticides. Haque R. and Freed V. H. (Eds). Plenum Press, New York, pp. 115 - 133.

Hamlett J.M., Horton R. & Cressie N.A.C. (1986). Resistant and exploratory techniques for use in semivariogram analyze. Soil Science Society of America Journal, 50:868-875.

Hattis D. & Burmaster D.E. (1994). Assessment of variability and uncertainty distributions for practical risk analyses. Risk Analysis, 14:713-730.

Heber V., Siebers J., Nolting H.G., Vetten H.J., Kreuzig R. & Bahadir M. (1998). Investigations on the applicability of two ELISA types for the determination of chlorpyrifos in soil samples compared with a gas chromatographic method. Fresenius Journal of Analytical Chemistry, 360:739-741.

Heinemann A.B., Hoogenboom G., Chojnicki B. (2002). The impact of potential errors in rainfall observation on the simulation of crop growth, development and yield. Ecological Modelling, 157:1-21.

Heuvelink G.B.M. (1998a). Error propagation in environmental modelling with GIS. Taylor & Francis, London, 127 p.

Heuvelink G.B.M. (1998b). Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems, 50:255-264.

Heuvelink G.B.M. & Pebesma E.J. (1999). Spatial aggregation and soil process modelling. Geoderma, 89:47-65.

Hoffman F.O. & Hammonds J.S. (1994). Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Analysis, 14:707-712.

Hoffman F.O., Chambers D.B. & Stager R.H. (1999). Uncertainty is part of making decisions. Human and Ecological Risk Assessment, 5:255-261.

Hollis J.M., Hallett S.H. & Keay C.A. (1993). The development and application of an integrated database for modelling the environmental fate of pesticides. Proceedings Brighton Crop Protection Conference – Weeds, 3:1355-1364.

IAEA (1989). Evaluating the reliability of predictions made using environmental transfer models. Safety series No. 100, Vienna, Austria.

Iman R.L. & Conover W.J. (1982). A distribution-free approach to inducing rank correlation among input variables. Communications in Statistics – Part B- Simulation and Computation, 11:311-334.

Inskeep W.P., Wraith J.M., Wilson J.P., Snyder R.D., Macur R.E. & Gaber H.M. (1996). Input parameter and model resolution effects on predictions of solute transport. Journal of Environmental Quality, 25:462-462.

Jacques D., Mouvet C., Mohanty B., Vereecken H. & Feyen J. (1999). Spatial variability of atrazine sorption parameters and other soil properties in a podzoluvisol. Journal of Contaminant Hydrology, 36:31-52.

Jager T., Vermeire T.G., Rikken M.G.J. & van der Poel P. (2001). Opportunities for a probabilistic risk assessment of chemicals in the European Union. Chemosphere, 43:257-264.

James B.R. & Gorelick S.M. (1994). When enough is enough: the worth of monitoring data in aquifer remediation design. Water Resources Research, 30:3499-3513.

Jansen M.J.W. (1998). Prediction error through modelling concepts and uncertainty from basic data. Nutrient Cycling in Agroecosystems, 50:247-253.

Jarvis N.J. (2003). Macropore flow and preferential flow. In: The Encyclopaedia of Agrochemicals. John Wiley and Sons, Chichester, UK (in press).

Jarvis N.J., Brown C.D. & Granitza E. (2000). Sources of error in model predictions of pesticide leaching: a case study using the MACRO model. Agricultural Water Management, 44:247-262.

Jaynes D.B., Novak J.M., Moorman T.B. & Cambardella C.A. (1995). Estimating herbicide partition coefficients from electromagnetic induction measurements. Journal of Environmental Quality, 24:36-41.

Jury W.A. (1983). Chemical transport modelling. current approaches and unresolved problems. Chemical mobility and reactivity in soil systems. ASA Special Publication 42, 49-64.

Jury W.A., Focht D.D. & Farmer W.J. (1987). Evaluation of pesticide groundwater pollution potential from standard indices of soil-chemical adsorption and biodegradation. Journal of Environmental Quality, 16:422-428.

Jury W.A. & Gruber J. (1989). A stochastic analysis of the influence of soil and climatic variability on the estimate of pesticide groundwater pollution potential. Water Resources Research, 25:2465-2474.

Kelly E.J. & Campbell K. (2000). Separating variability and uncertainty in environmental risk assessment – making choices. Human and Ecological Risk Assessment, 6:1-13.

Koskinen W.C., Cecchi A.M., Dowdy R.H. & Norberg K.A. (1999). Adsorption of selected pesticides on a rigid PVC lysimeter. Journal of Environmental Quality, 28:732-734.

Krajewski W.F., Kruger A. & Nespor V. (1998). Experimental and numerical studies of small-scale rainfall measurements and variability. Water Science & Technology, 37:131-138.

Lafrance P. & Banton O. (1995). Implication of spatial variability of organic carbon on predicting pesticide mobility in soil. Geoderma, 65:331-338.

Leake C.R., Humphreys S.P. & Austin D.J. (1995). Factors influencing the modelling of pesticide degradation in soil and the estimation of half-life (DT50) and DT90 values. In: Pesticide Movement to Water, BCPC monograph No. 62, pp. 195-200.

Lennartz B. (1999). Variation of herbicide transport parameters within a single field and its relation to water flux and soil properties. Geoderma, 91:327-345.

Li Z.C., Yost R.S. & Green R.E. (1998). Incorporating uncertainty in a chemical leaching assessment. Journal of Contaminant Hydrology, 29:285-299.

Loague K. (1991). The impact of land use on estimates of pesticide leaching potential: assessments and uncertainties. Journal of Contaminant Hydrology, 8:157-175.

Loague K. (1994). Regional scale groundwater vulnerability estimates: impact of reducing data uncertainties for assessments in Hawaii. Ground Water, 32:605-616.

Loague K., Giambelluca T.W., Green R.E., Liu C.C.K., Liang T.C. & Oki D.S. (1989). Simulation of organic chemical movement in Hawaii soils with PRZM: 2. Predicting deep penetration of DBCP, EDB and TCP. Pacific Science, 43:362-383.

Loague K., Bernknopf R.L., Green R.E. & Giambelluca T.W. (1996). Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii: review. Journal of Environmental Quality, 25:475-490.

Ma L., Ascough J.C., Ahuja L.R., Shaffer M.J., Hanson J.D. & Rojas K.W. (2000). Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. Transactions of the ASAE, 43:883-895.

Matsui Y., Itoshiro S., Buma M., Matsushita T., Hosogoe K., Yuasa A., Shinoda S. & Inoue T. (2002). Predicting pesticide concentrations in river water with a hydrologically calibrated basin-scale runoff model. Water Science & Technology, 45:141-148.

Maund S.J., Travis K.Z., Hendley P., Giddings J.M. & Solomon K. (2001). Probabilistic risk assessment of cooton pyrethroids: V. Combining landscape-level exposures and ecotoxicological effects data to characterize risks. Environmental Toxicology & Chemistry, 20:687-692.

Melching C.S. Reliability estimation. In: Computer models of watershed hydrology. Singh V.P. (Ed.), Water Resources Publication, CO, pp. 69-118.

Millstein J.A. (1994). Propagation of measurement errors in pesticide application computations. International Journal of Pest Management, 40:159-165.

Moir W.H. & Mowrer H.T. (1995). Unsustainability. Forest Ecology and Management, 73:239-248.

Morgan M.G. & Henrion M. (1990). Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York.

Mosleh A. (1986). Hidden sources of uncertainty: judgement in the collection and analysis of data. Nuclear Engineering and Design, 93:187-198.

Mouvet C. & Jücker C. (1997). Influence of various filters on the concentration of pesticides dissolved in water. Environmental Science & Technology, 31:2434-2437.

Mowrer H.T. (2000). Uncertainty in natural resource decision support systems: sources, interpretation and importance. Computers and Electronics in Agriculture, 27:139-154.

Nemes A., Wösten J.H.M., Lilly A. & Oude Voshaar J.H. (1999). Evaluation of different procedures to interpolate particle-size distributions to achieve compatibility within soil databases. Geoderma, 90:187-202.

Nofziger D.L., Jin-Song Chen & Haan C.T. (1994). Evaluating the Chemical Movement in Layered Soil Model as a tool for assessing risk of pesticide leaching to groundwater. Journal of Environmental Science and Health, A29:1133-1155.

Northcott G.L. & Jones K.C. (2000). Spiking hydrophobic organic compounds into soil and sediment: a review and critique of adopted procedures. Environmental Toxicology and Chemistry, 19:2418-2430.

Novak J.M., Moorman T.B. & Cambardella C.A. (1997). Atrazine sorption at the field scale in relation to soils and landscape position. Journal of Environmnental Quality, 26:1271-1277.

Oliveira R.S., Koskinen W.C., Ferreira F.A., Khakural B.R., Mulla D.J. & Robert P. (1999). Spatial variability of imazethapyr sorption in soil. Weed Science, 47:243-248.

Oreskes N. (1998). Evaluation (not validation) of quantitative models. Environmental Health Perspectives, 106:1453-1460.

Ott W.R. (1990). A physical explanation of the lognormality of pollutant concentrations. JAPCA – The Journal of the Air & Waste Management Association, 40:1378-1383.

Pachepsky Y.A., Rawls W.J. & Timlin D.J. (1999). The current status of pedotransfer functions: their accuracy, reliability, and utility in field- and regional-scale modelling. In: Assessment of Non-Point Source Pollution in the Vadose Zone, Corwin D.L., Loague K., Ellsworth T. (eds). AGU Geophysical Monograph 108, pp. 223-234.

Parkin T.B. (1993). Spatial variability of microbial processes in soil – a review. Journal of Environmental Quality, 22:409-417.

Parkin T.B. & Shelton D.R. (1992). Spatial and temporal variability of carbofuran degradation in soil. Journal of Environmental Quality, 21:672-678.

Price O.R., Walker A., Wood M. & Oliver M.A. (2001). Using geostatistics to evaluate spatial variation in pesticide/soil interactions. In: A. Walker (Ed.), Pesticide behaviour in soils and water, BCPC Symposium Proceedings No. 78, Farnham, UK, pp. 233-238.

Ramsey M.H. (1998). Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources. Journal of Analytical Atomic Spectrometry, 13:97-104.

Rao P.S.C. & Davidson J.M. (1982).Retention and Transformation of selected pesticides and phosphorus in soil-water systems: a critical review. EPA 600/3-82-060.

Rao P.S.C. & Wagenet R.J. (1985). Spatial variability of pesticides in field soils: methods for data analysis and consequences. Weed Science, 33:18-24.

Russell M.H., Layton R.J. & Tillotson P.M. (1994). The use of pesticide leaching models in a regulatory setting – an industrial perspective. Journal of Environmental Science and Health, 29A:1105-1116.

Saltelli A., Chan K. & Scott E.M. (2000). Sensitivity analysis. John Wiley & Sons, Chichester, UK.

Seiler F.A. & Alvarez J.L. (1996). On the selection of distribution for stochastic variables. Risk Analysis, 16:5-18.

Sevruk B. (1982). Methods of correction for systematic error in point precipitation measurement. World Meteorological Organization, Operation Hydrology report No. 21, WMO-No. 589, Geneva, Switzerland.

Shimizu Y., Yamazaki S. & Tersahima Y. (1992). Sorption of anionic pentachlorophenol in aquatic environments: the effect of pH. Water Science Technology, 25:41-48.

Shukla S., Mostaghimi S., Shanholt V.O., Collins M.C. & Ross B.B. (2000). A county-level assessment of ground water contamination by pesticides. Ground Water Monitoring and Remediation, 20:104-119.

Simmonds L.P. & Nortcliff S. (1998). Small scale variability in the flow of water and solutes, and implications for lysimeter studies of solute leaching. Nutrient Cycling in Agroecosystems, 50:65-75.

Small M.J. & Mular J.R. (1987). Long term pollutant degradation in the unsaturated zone with stochastic rainfall infiltration. Water Resources Research, 23:2246-2256.

Smith C.N. & Parrish R.S. (1993). A field study to evaluate leaching of aldicarb, metolachlor and bromide in a sandy loam soil. Journal of Environmental Quality, 22:562-577.

Smith A.E., Ryan P.B. & Evans J.S. (1992). The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. Risk Analysis, 12:467-474.

Solomon K., Giesy J. & Jones P. (2000). Probabilistic risk assessment of agrochemicals in the environment. Crop Protection, 19:649-655.

Soulas G. & Lagacherie B. (2001). Modelling of microbial degradation of pesticides in soils. Biology and Fertility of Soils, 33:551-557.

Soutter M. & Pannatier Y. (1996). Groundwater vulnerability to pesticide contamination on a regional scale. Journal of Environmental Quality, 25:439-444.

Soutter M. & Musy A. (1998). Coupling 1D Monte-Carlo simulations and geostatistics to assess groundwater vulnerability to pesticide contamination on a regional scale. Journal of Contaminant Hydrology, 32:25-39.

Soutter M. & Musy A. (1999). Global sensitivity analyses of three pesticide leaching models using a Monte Carlo approach. Journal of Environmental Quality, 28:1290-1297.

Stine P.A. & Hunsaker C.T. (2001). An introduction to uncertainty issues for spatial data used in ecological applications. In: Spatial uncertainty in ecology. C.T. Hunsaker, M.F. Goodchild, Friedl M.A. & Case T.J., Springer, New York, pp. 91-107.

Tarantola S., Giglioli N., Jesinghaus J. & Saltelli A. (2002). Can global sensitivity analysis steer the implementation of models for environmental assessments and decision-making? Stochastic Environmental Research and Risk Assessment, 16:63-76.

Taylor A.W., Freeman W.H.P. & Edwards W.M. (1971). Sample variability and the measurement of dieldrin content of a soil in the field. Journal of Agricultural and Food Chemistry, 19:832-836.

Thompson K.M., Burmaster D.E. & Crouch E.A.C. (1992). Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments. Risk Analysis, 12:53-63.

Tietje O. & Hennings V. (1993). Bewertung von Pedotransferfunktionen zur Schätzung der Wasserspannungskurve. Z. Pflantzenernähr. Bodenk., 156:447-455.

Tietje O. & Tapkenhinrichs M. (1993). Evaluation of pedo-transfer functions. Soil Science Society of America Journal, 57:1088-1095.

Tiktak A., Leijnse A. & Viddenberg H. (1999). Uncertainty in a regional-scale assessment of cadmium accumulation in the Netherlands. Journal of Environmental Quality, 28:461-470.

Tiktak A. (2000). Application of pesticide leaching models to the Vredepeel dataset. II. Pesticide fate. Agricultural Water Management, 44:119-134.

Topp E. & Smith W. (1992). Sorption of herbicides atrazine and metolachlor to selected plastics and silicone rubber. Journal of Environmental Quality, 21:316-317.

Travis K.Z. (2000). Pesticide leaching models – past, present and future. In: Proceedings of the Brighton Crop Protection Conference – Pests and Diseases, BCPC, Farnham, Surrey, UK, pp. 69-76.

Trevisan M., Calandra R., Vischetti C., Esposito A. & Padovani L. (2001). Pesticide leaching potential in the Trasimeno Lake area. Assessment of uncertainty associated with the simulation process. In: A. Walker (Ed.), Pesticide behaviour in soils and water, BCPC Symposium Proceedings No. 78, Farnham, UK, pp. 245-250.

Ukrainczyk L. & Ajwa H.A. (1996). Primisulfuron sorption on minerals and soils. Soil Science Society of America Journal, 60:460-467.

US EPA (1996). Summary report for the workshop on Monte Carlo analysis, EPA/630/R-96/010, Risk assessment forum, Washington, DC.

van der Zee S.E.A.T.M. & Boesten J.J.T.I. (1991). Effects of soil heterogeneity on pesticide leaching to groundwater. Water Resources Research, 27:3051-3063.

van der Zee S.E.A.T.M. & Boesten J.J.T.I. (1993). Pesticide leaching in heterogeneous soils with oscillating flow: approximations. The Science of the Total Environment, 132:167-179.

Vanclooster M., Boesten J.J.T.I., Trevisan M., Brown C.D., Capri E., Eklo O.M., Gottesbüren B., Gouy V. & van der Linden A.M.A. (2000). A European test of pesticide leaching models: methodology and major recommendations. Agricultural Water Management, 1-19.

Vereecken H. (1992). Derivation and validation of pedotransfer functions for soil hydraulic properties. In: Indirect methods for estimating the hydraulic properties of unsaturated soils. van Genuchten M.T. & Leij F.J. (Eds.). Riverside, California: University of California, Riverside. Pp. 473-488.

Verschoor A.J., Boesten J.J.T.I., Leistra M., van der Linden A.M.A., Linders J.B.H.J. & Pol J.W.W. (2001). Evaluation of pesticide leaching in lysimeter and field studies. Parent substances. RIVM report 601506 007/2001.

Verro R., Calliera M., Maffioli G., Auteri D., Sala S., Finizio A. & Vighi M. (2002). GIS-based system for surface water risk assessment of agricultural chemicals. 1. Methodological approach. Environmental Science & Technology, 36:1532-1538.

Vink J.P.M., Gottesbüren B., Diekkrüger B. & van der Zee S.E.A.T.M. (1997). Simulation and model comparison of unsaturated movement of pesticides from a large clay lysimeter. Ecological Modelling, 105:113-127.

Vischetti C., Businelli m., Marini M., Capri E., Trevisan M., DelRe A.A.M., Donnaruma L., Conte E. & Imbroglini G. (1997). Characterisation of spatial variability structure in three separate field trials on pesticide dissipation. Pesticide Science, 50:175-182.

Vischetti C., Perniola M., Scarponi L. & Tarantino E. (1998). Field and lysimeter study on the leaching of bromide ion and the herbicides imazethapyr and bentazone in a clay loam soil in southern Italy. Fresenius Environmental Bulletin, 7:641-648.

Vischetti C., Scarponi L., Esposito A., Vianello M., Zanin G. (2001). Characterisation of the structure of spatial variability for initial pesticide concentration in a field trial. In: A. Walker (Ed.), Pesticide behaviour in soils and water, BCPC Symposium Proceedings No. 78, Farnham, UK, pp. 295-300.

Viton P. (1970). Pluviomètres et pluviographes utilises par la Météorologie Nationale. In: Techniques d'étude des facteurs physiques de la biosphere, INRA, Paris. pp. 245-253.

Walker A. & Brown P.A. (1981). Effects of soil storage on degradation rates of metamitron, atrazine and propyzamide. Proceedings of the EWRS Symposium: Theory and practice of the use of soil applied herbicides, 63-71.

Walker A. & Brown P.A. (1983). Spatial variability in herbicide degradation rates and residues in soil. Crop Protection, 2:17-25.

Walker A., Jurado-Exposito M., Bending G.D. & Smith V.J.R. (2001). Spatial variability in the degradation rate of isoproturon in soil. Environmental Pollution, 111:407-415.

Walker A., Bromilow R.H., Nicholls P.H., Evans A.A. & Smith V.J.R. (2002). Spatial variability in the degradation rates of isoproturon and chlorotoluron in a clay soil. Weed Research, 42:39-44.

Warren-Hicks W.J. & Moore D.R.J. (1998). Uncertainty analysis in ecological risk assessment. Proceedings from the Pellston workshop on uncertainty analysis in ecological risk assessment, 23-28 August 1995. Pellston, Michigan. Pensacola FL: Society of Environmental Toxicology and Chemistry, 341 pp.

Wauchope R.D., Chandler J.M. & Savage K.E. (1977). Soil sample variation and herbicide incorporation uniformity. Weed Science, 25:193-196.

Wauchope R.D., Yeh S., Linders J.B.H.J., Kloskowski R., Tanaka K., Rubin B., Katayama A., Kordel W., Gerstl Z., Lane M. & Unsworth J.B. (2002). Pesticide soil sorption parameters: theory, measurement, uses, limitations and reliability. Pest Management Science, 58:419-445.

Webster R. (2001). Statistics to support soil research and their presentation. European Journal of Soil Science, 52:331-340.

Webster R. & Oliver M.A. (2000). Geostatistics for Environmental Scientists. John Wiley and Sons Ltd, 286 p.

Wolt J.D., Nelson H.P., Cleveland C.B. & van Wesenbeeck I.J. (2001). Biodegradation kinetics for pesticide exposure assessment. Reviews in Environmental Contaminant Toxicology, 169:123-164.

Wolt J., Singh P., Cryer S. & Lin J. (2002). Sensitivity analysis for validating expert opinion as to ideal data set criteria for transport modelling. Environmental Toxicology & Chemistry, 21:1558-1565.

Wood L.S., Scott H.D., Marx D.B. & Lavy T.L. (1987). Variability in sorption coefficients of metolachlor on a Captina silt loam. Journal of Environmental Quality, 16:251-256.

Wösten J.H.M., Veerman G.J. & Stolte J. (1994). Water retention and hydraulic conductivity characteristics of topsoils and subsoils in the Netherlands: the Staring series. Technical document 18, DLO-Winand Staring Centre, Wageningen, The Netherlands.

Wösten J.H.M., Lilly A., Nemes A. & Le Bas C. (1999). Development and use of a database of hydraulic properties of European soils. Geoderma, 90:169-185.

Wu Q.J., Ward A.D., Workman S.R. & Salchow E.M. (1997). Applyting stochastic simulation techniques to a deterministic vadose zone solute transport model. Journal of Hydrology, 197:88-110.

Yapo P.O., Gupta H.V. & Sorooshian S. (1998). Multi-objective global optimization for hydrologic models. Journal of Hydrology, 204:83-97.

Yost R., Loague K. & Green R.E. (1993). Reducing variance in soil organic carbon estimates: soil classification and geostatistical approaches. Geoderma, 57:247-262.

Zacharias S., Heatwole C.D., Persaud N., Bruggeman A.C., Kumar D. & Smith C.N. (1999). Stochastic simulation of field-scale pesticide transport using Opus and GLEAMS. Journal of Environmental Quality, 28:411-423.

Zhang J. & Goodchild M.F. (2002). Uncertainty in geographical information. Taylor & Francis, London, 266 p.

Zhang H., Haan C.T. & Nofziger D.L. (1993). An approach to estimating uncertainties in modeling transport of solutes through soils. Journal of Contaminant Hydrology, 12:35-50.

Chapter 3

Aspects of repeatability in Monte Carlo modelling (synopsis)

This chapter reports on investigations carried out within the scope of the present project and of an earlier DEFRA project (PL0532, Sensitivity analyses for the pesticide leaching models used for pesticide registration in Europe). The research is presented in the form of a paper which was submitted to Environmental Toxicology and Chemistry in October 2002. The paper was co-authored by Igor Dubus and Peter Janssen (RIVM, The Netherlands).

The research was aimed at investigating repeatability issues in Monte Carlo based approaches. Monte Carlo techniques are extensively used to account for the uncertainty in model predictions resulting from that in input parameters and to assess model sensitivity. Although termed 'random', the generation of series of numbers by Monte Carlo sampling is not random and identical samples of numbers can be obtained if the same seed numbers are used (the seed number is used to initiate the sequence of 'random' samples). Sensitivity analyses and probabilistic modelling were undertaken for various number of runs of the pesticide leaching model PELMO and these were repeated 10 times with different seed numbers.

The ranking of PELMO input parameters according to their influence on predictions for leaching (sensitivity analysis) was found to be stable for the most influential parameters. For those less influential parameters, the sensitivity ranking was found to be greatly influenced by the seed number used.

For Monte Carlo modelling, probabilities of exceeding 0.1 μg/l were found to be significantly influenced by the seed number used in the random sampling of values for the two parameters considered (Koc and DT50). The coefficients of variation associated with the probability estimates decreased with increasing number of runs, but were still in the order of 5-9% when 5000 runs of the model were undertaken. These levels of uncertainty will affect the confidence that should be attributed to probabilistic estimates of exposure and will limit the usefulness of Monte Carlo approaches for environmental risk assessments. It is recommended that attention is paid to replicability aspects when modellers devise their approach to probabilistic modelling and when decision makers examine results of probabilistic exercises.

Chapter 3

Aspects of repeatability in Monte Carlo modelling

1 INTRODUCTION

Modelling techniques based on Monte Carlo (MC) sampling are increasingly being used in environmental disciplines (Beven & Binley, 1992). The two main applications are typically: i) the study of the relationship between model input and output (sensitivity analysis; Helton, 1993); and, ii) the assessment of the variability / uncertainty in modelling predictions resulting from the variability / uncertainty in model input parameters (uncertainty analysis, also referred to as probabilistic modelling or MC modelling; Morgan & Henrion, 1990). A MC uncertainty analysis typically involves six stages: i) the selection of outputs of interest on which the analysis will be performed; ii) the selection of model input parameters to be included in the analysis; iii) the attribution of ranges and distributions to each parameter selected and specification of their dependence/correlation; iv) the generation of a random sample from the distributions assigned to parameters; v) the running of the model for each of the sample elements; and, vi) the examination of model predictions in statistical terms (*e.g.* estimation of the mean and variance, construction of cumulative distribution charts). In a sensitivity analysis, the sixth step consists of examining the relationship between changes in model predictions resulting from changes in model input parameters using scatterplots, regression analysis or correlation measures (Saltelli *et al.*, 2000). MC sensitivity and uncertainty analyses have been reported for pesticide leaching models (*e.g.* Carsel *et al.*, 1988; Soutter & Musy, 1999; Ma *et al.*, 2000; Dubus & Brown, 2002). Reasons for the widespread adoption of MC techniques in the study of model sensitivity and uncertainty in modelling include: i) they are transparent, conceptually simple, easy to explain and can therefore be readily understood by decision makers; ii) they are generic and can be applied to virtually any model; iii) they have been applied on numerous occasions in the past; iv) they are relatively easy to implement when using appropriate software; and, v) they normally do not require modifications to the model code.

Various sampling procedures are used in MC studies and these include random sampling, stratified sampling and quasi-random sampling (Saltelli *et al.*, 2000). Random sampling offers the advantage of providing unbiased estimates of the mean and variance of the variable

considered, but a large sample size is required to cover the sample space. Making several thousands of model runs may not be practical, especially for models which are expensive to run in computational terms. These practical shortcomings of traditional MC random sampling have led to the wide adoption of a stratified random scheme known as Latin Hypercube Sampling (LHS) (McKay *et al.*, 1979). LHS is considered to be an extremely efficient sampling scheme when compared to purely random MC sampling because it provides a better coverage of the parameter space and allows a significant reduction in the sample size to achieve a given level of confidence without compromising the overall quality of the analysis (Helton, 1993). In LHS, the probability distributions are divided into ranges of equal probability and one sample is taken from each equi-probability interval. In median LHS, one sample is taken from the median of each equal-probability interval, while in random LHS one sample is taken at random within each interval (Morgan & Henrion, 1990). Restricted pairing techniques are available for inducing correlations between variables in LHS (Iman & Conover, 1982; Iman and Shortencarier, 1984). Although sampling schemes are termed random, the numerical processes involved are deterministic and therefore not random in the strictest sense. 'Random' number generation typically requires the specification of a seed number which will be used to initiate the random sequence. The specification of different seed numbers will result in different random samples. Although MC simulation is widely used in many fields of science, little research has been conducted on what influence the generation of a random sample with a particular seed number may have on the overall outcome of probabilistic modelling (replicability issue).

The present paper reports on investigations carried out to assess the robustness of MC applications in pesticide fate modelling (MC based sensitivity analyses and probabilistic modelling) and, hence, the confidence that should be attributed to results obtained through these techniques. The research looked at the influence on results of sensitivity and uncertainty analyses of: i) the seed number used to generate random values for model input parameters; and, ii) the size of the random sample (equivalent to the number of model runs undertaken). Sensitivity and uncertainty analyses were undertaken for the pesticide leaching model PELMO with varying numbers of model runs and all analyses were repeated 10 times with different seed numbers.

2 MODELLING METHODS

2.1 Model selection

The pesticide leaching model used was PELMO (PEsticide Leaching MOdel) in its version 3.00 SP2, the latest release available at the time the work was started. The model has been widely used in the assessment of the potential for pesticides to reach groundwater resources within the context of their European registration and has benefited from a number of evaluation studies (Klein, 1994; Klein *et al.*, 1997). Investigations reported in this paper involved a large number of model runs and PELMO was selected from a range of models suited for the simulation at hand as the model with the fastest running time.

PELMO is a one-dimensional leaching model which integrates descriptions of water movement and pesticide transfer through the soil column (Klein, 1995). The model was adapted from an early version of the US-EPA Pesticide Root Zone Model (Carsel *et al.*, 1984) and the current versions of the two models are very similar. Hydrology is described using a "tipping-bucket" approach where water will only flow to the next soil layer if field capacity is exceeded. This means that upwards movement of water and associated solutes resulting from drying at the soil surface in response to losses of water by evapotranspiration cannot be simulated. Solute transport is simulated using the convection-dispersion equation. Subroutines describing the fate of reactive solutes include the description of i) sorption by the Freundlich equation; ii) degradation using first-order kinetics; iii) pesticide losses in run-off using the modified Soil Conservation Service curve number technique; iv) pesticide losses via soil erosion using the modified soil loss equation; v) volatilisation using a simple approach combining diffusion and partitioning between the liquid and gaseous phases in the soil; and, vi) complex metabolic schemes. Outputs of the model are provided on a daily basis. The model was initially developed to simulate average behaviour at the field scale although it has been used to simulate smaller systems (*e.g.* lysimeters; Klein *et al.*, 1997).

2.2 Scenario simulated

The scenario considered consisted in the simulation of the leaching of a pesticide in the German soil 'Borstel' with replicated climatic data from Hamburg. The selection of soil and climatic information was not meant to represent a realistic leaching situation, but rather to represent common practice in the assessment of the leaching potential within the context of pesticide registration in Europe. The Borstel soil is a sandy loam over sand (USDA textural

classification) which exhibits relatively low organic carbon contents throughout the profile (1.5, 1.0 and 0.2% in the 0-30, 30-60 and 60-75 cm layers, respectively; no organic carbon from 75 cm downwards). A detailed description of the profile can be found elsewhere (Klein, 1995). The weather dataset used was that known as the 'Hamburg normal' which consists of a number of replicated years of weather data for the year 1978 for Hamburg (annual rainfall 777 mm). This dataset is designed to represent average conditions with regard to rainfall in northern Germany.

Properties of the pesticide were selected from within the range of those of registered compounds to give a reasonable likelihood of leaching to groundwater (US EPA, 1992). A laboratory DT50 value of 58.5 days (assumed to have been determined at 20ºC and at 40% field capacity) was initially considered in the modelling while the sorption distribution coefficient normalised to organic carbon (Koc) and the Freundlich exponent were set to 91.45 ml/g and 0.895, respectively.

There are no guidelines available yet from pesticide regulatory authorities on what an acceptable level of exceedance of a particular threshold concentration is, but it is anticipated that within the context of probabilistic modelling, the acceptance level will be below 5% exceedance. Accordingly, an application rate of 15 g active substance/ha was selected to provide on average a probability of just below 5% that the annual average concentration in the 20th year simulated exceeds the European Union threshold for drinking water (0.1 μg/l). The pesticide was assumed to be applied to the soil on 15 May in each of the 20 years simulated.

2.3 Automation of modelling tasks

The output variable of interest in the present exercise was the annual average pesticide concentration in leachate in the 20th year simulated. This endpoint is traditionally compared to a threshold concentration of 0.1 μg/l within the context of pesticide registration in Europe. Annual average concentrations tend to increase over the years for relatively persistent compounds and a relatively long simulation period (20 years) was selected to make sure a plateau was reached by the end of the run period. Annual average concentrations were computed from the PELMO raw output files wasser.plm (leaching volumes) and chem.plm (pesticide loads) using a Perl script (Perl, 2002). The precision in the reporting by PELMO of the leaching volumes and pesticide loads is four and five significant figures, respectively. The automatic modification of values for selected parameters in the PELMO input files, the

running of the PELMO model and the post-processing script, and the extraction of values from the script output were automated using the SENSAN utility provided in the PEST package (Doherty, 2000).

2.4 Replicability in sensitivity analysis

The aim of the sensitivity analyses was to establish a classification of PELMO input parameters according to their influence on PELMO predictions for leaching, *i.e.* to identify the parameters to which the model is sensitive or insensitive. The sensitivity of PELMO was studied using MC sampling combined with regression analysis. This popular methodology (Saltelli *et al.*, 2000) has recently been used to investigate the sensitivity of the preferential flow model MACRO (Dubus & Brown, 2002). In the present instance, 21 PELMO input parameters were attributed normal (18 parameters), log-normal (2 parameters) or uniform (1 parameter) distributions on the basis of expert judgement (Table 4). For the normal and log-normal probability density functions, lower and upper values for variation were arbitrarily selected (Table 4) and it was assumed that these corresponded to 2.5th and 97.5th percentiles of the distributions, respectively. This is equivalent to assuming that there is a 1-in-20 chance that a value would fall outside the range defined by the lower and upper values presented in Table 4. For organic carbon content, bulk density and water capacity, only the parameter in the top horizon was included in the analysis and values for the parameter in the deeper horizons were modified by the same variation applied to the topsoil value.

Parameter [a]	Description	Nominal value	Lower value [b]	Upper value [b]	Distribution	Variance
ANET	Depth of evapotranspiration computation (cm)	10	5	15	normal	6.51
CINT	Maximum interception storage (cm)	0.28	0.21	0.35	normal	1.28×10-3
AMXD	Maximum active rooting depth (cm)	110	82.5	137.5	normal	1.97×102
COVM	Maximum soil cover by the crop (%)	100	80	100	uniform	-
BUD*	Bulk density (g/cm3)	1.5	1.35	1.65	normal	5.86×10-3
WC*	Initial soil water content (%)	0.2	0.15	0.25	normal	6.51×10-4
SA1	Sand content first horizon (%)	68.3	64.9	71.7	normal	3.04
CL1	Clay content first horizon (%)	7.2	6.8	7.6	normal	3.37×10-2
OC*	Organic carbon (%)	1.5	1.35	1.65	normal	5.86×10-3
SA2	Sand content second horizon (%)	67	63.65	70.35	normal	2.92
CL2	Clay content second horizon (%)	6.7	6.365	7.035	normal	2.92×10-2
SA3	Sand content third horizon (%)	96.2	92.4	100.0c	normal	3.66
CL3	Clay content third horizon (%)	0.9	0.855	0.945	normal	5.27×10-4
APPL	Application rate (kg/ha)	0.015	0.01125	0.01875	normal	3.66×10-6
DT50	Laboratory half-life (days)	58.5	29.25	117	log-normal	3.85×102
TEMP	Temperature of laboratory incubation (ºC)	20	19	21	normal	2.60×10-1
QTEN	Increase in degradation for a temp. increase of 10ºC (-)	2.2	1.76	2.64	normal	5.03×10-2
MOIS	Soil moisture content during incubation experiment (%)	40	30	50	normal	2.60×101
MEXP	Exponent for moisture correction (-)	0.7	0.56	0.84	normal	5.10×10-3
KOC	Sorption coefficient normalised to organic carbon (ml/g)	91.45	45.725	182.9	log-normal	9.41×102
NF	Freundlich exponent (-)	0.895	0.716	1.074	normal	8.34×10-3

Table 4. PELMO input parameters included in the sensitivity analysis and parameterisation of probability density functions

[a] The star indicates those parameters which are depth dependent. Variations of values in the subsoil were linked to that in the topsoil.

[b] Lower and upper values were assumed to correspond to the 2.5th and 97.5th percentiles of the normal and log-normal distributions.

[c] A truncation at 100% was integrated into the sampling to avoid the use of unrealistic values.

Latin Hypercube Sampling (LHS) was used to sample values for the 21 parameters from their probability density functions. The two sampling packages which were used in the present research were Crystal Ball 2000 (version 5.1; Decisioneering, 2000) and @RISK (version 4.0.5; Palisade, 2000) which both implement random LHS. Four different sample sizes were considered (250, 1000, 2500 and 5000 elements) and for each sample size, ten different replicated samples were generated by varying the seed number used in the sampling. Seed numbers were selected randomly from a uniform distribution to avoid any bias in their selection. The maximum number of elements (5000) was determined by limitations with regard to computational resources. No truncation was applied in the sampling except for the distribution of sand content in the third horizon which was truncated at a maximum of 100% to avoid the model crashing.

PELMO was run automatically for each element of each replicated sample and the annual average pesticide concentration in leachate in the 20th year simulated was automatically extracted (see above). The total number of PELMO runs was therefore 10*(250+1000+2500+5000) = 87500.

The sensitivity of PELMO to changes in input parameters was derived by regression analysis with raw (*i.e.* non ranked) and rank-transformed data. The rank transformation consisted in replacing each value for inputs and outputs by their rank in the sample. For instance, the largest organic carbon content randomly selected in the sample of size 250 received the rank 1 while its smallest value was attributed the rank 250. Rank transformation is a popular transformation in sensitivity analysis which is used to decrease non-linearity/non-monotonicity in highly non-linear systems (Iman & Conover, 1979). Both raw and ranked data were analysed in the same way. Annual average concentrations in the 20th year simulated were related to model inputs using standardised multiple linear regression using the STATISTICA package (version 6.0; Statsoft, 2002) as follows:

$$Y = \sum_{i=1}^{21} b_i \times X_i + \varepsilon \qquad [1]$$

where Y is the standardised concentration, Xi is a standardised input parameter, bi is the regression coefficient for each Xi and ε is the regression error.

The sensitivity of PELMO to each input parameter is given by the coefficient bi for this parameter, which is known as the standardised regression coefficient SRC (raw data) and the

standardised ranked regression coefficient SRRC (rank transformed data) (Hamby, 1994). S(R)RC values can be positive or negative. A positive S(R)RC means that an increase in the input parameter will result in an increase in the output whereas a decrease in the output would result for a parameter with a negative SRRC. For each replicated sample and for raw and ranked data, absolute values of S(R)RC were used to sort the 21 PELMO input parameters according to their influence on the prediction of the annual average concentration in the 20th year of simulation. SRRC indices have been shown to be one of the most robust and reliable means to assess model sensitivity (Saltelli & Marivoet, 1990).

2.5 Replicability in probabilistic modelling

The probabilistic modelling undertaken was based on MC simulations and was restricted to the variation of two parameters only: the Freundlich sorption coefficient normalised to organic carbon (Koc) and the half-life of the pesticide (DT50). PELMO has been shown previously to be most sensitive to parameters related to sorption and degradation (Dubus *et al.*, 2000) and these two parameters play an important role within environmental risk assessments for pesticides.

Data on environmental properties of pesticides are expected to be limited in most instances where an environmental risk assessment is needed and not numerous enough to support a robust attribution of probability distribution functions to parameters via distribution fitting. In the present case, log-normal distributions were attributed to Koc and DT50 through expert judgement. The focus of the present paper was on illustrating replicability issues and no distinction between the contributions of stochastic variability and incertitude (lack of knowledge) to the overall uncertainty was made in an effort to keep the modelling system as simple as possible. In our experience, a common rule of thumb is that Koc and DT50 values obtained through standard laboratory experiments carried out will vary approximately within a factor of two from the median values, although some consider that a factor of four might be more appropriate. This expert knowledge was reflected in the parameterisation of the log-normal distributions by considering that 95% of the overall probability for the two parameters was contained within a range defined by M/2 and M×2 where M is the value considered in the initial scenario. This is equivalent to saying that there is a one-in-20 chance that a given value will be outside the range [M/2; M×2]. Two cases were then made by considering the possibility of using truncation in the sampling, which would reflect the belief of the modeller that an extremely small or large value for Koc or DT50 as sampled from a log-normal distribution would not be realistic. Truncation levels were arbitrarily set to the 0.5th and

99.5th percentiles of the log-normal distributions used for Koc and DT50. A graphical summary of the parameterisation of statistical distributions for DT50 and Koc is provided in Figure 6.

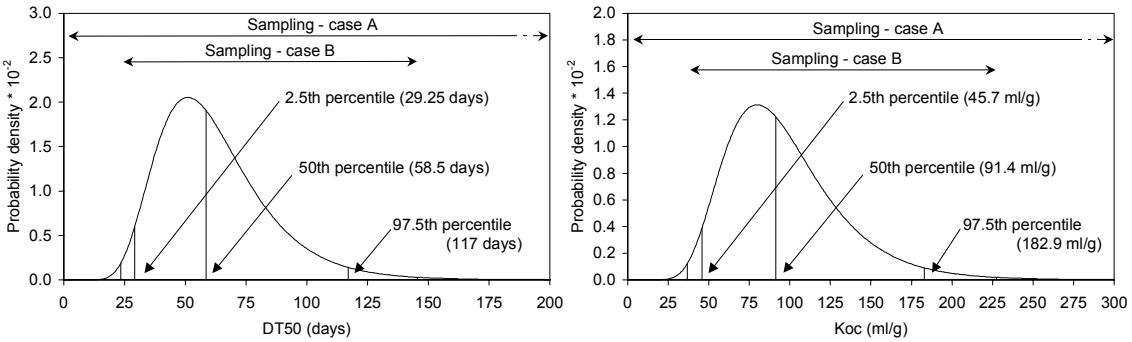


Figure 6. Parameterisation of probability density functions for DT50 and Koc
95% of the probability is contained within the range [median/2; median*2]
Case A: no truncation in the sampling; Case B: truncation at the 0.5th and 99.5th percentiles

The LHS procedure was the same as that used for sensitivity analyses. Twelve different sample sizes were considered (10, 50, 100, 150, 200, 250, 500, 750, 1000, 1500, 2500 and 5000 elements) and for each sample size, ten different samples were generated by varying the seed number used in the sampling. Again, seed numbers were selected randomly from a uniform distribution to avoid any bias in their selection and the maximum number of elements (5000) was determined by limitations with regard to computational resources.

PELMO was run for each element of each replicated sample and the annual average pesticide concentration in leachate in the 20th year simulated was automatically extracted (see above). For each sample, the probability p of exceeding a threshold concentration of 0.1 µg/l was calculated as follows:

$$p = \frac{n(C > 0.1\mu g/l)}{N} \qquad [2]$$

where n (C>0.1µg/l) is the number of model runs for which the average annual concentration in the 20th year of simulation (C) exceeded 0.1 µg/l,
N is the total number of model runs carried out in the sample.

The total number of PELMO runs undertaken to investigate replicability issues in probabilistic modelling was
10*(10+50+100+150+200+250+500+750+1000+1500+2500+5000) = 315,200. Adding the number of PELMO runs for the sensitivity analysis, PELMO was run for a total of 402,700

times. Each PELMO run took between 17 to 35 seconds to complete depending on the computer used.

3 RESULTS

3.1 Replicability in sensitivity analysis

Model sensitivity was determined by regression analysis relating the annual average pesticide concentration in leachate in the 20th year simulated to model inputs. In those instances where regression analyses were carried out on the raw untransformed data, relatively small r2 values were obtained (0.31<r2<0.49; Table 5). The poor quality of the regressions reflected the inherent non-linear character of the relationship between inputs and outputs in pesticide fate models and may question the validity of the sensitivity results obtained. The application of a rank transformation to the input and output data resulted in a large increase in r2 values (0.95<r2<0.97), confirming the usefulness of this approach.

	Raw data		Rank-transformed data	
Number of model runs	Mean	Standard deviation	Mean	Standard deviation
250	0.415 (a)	0.040	0.958 (a)	0.005
1000	0.360 (b)	0.023	0.955 (b)	0.002
2500	0.358 (b)	0.012	0.954 (b)	0.003
5000	0.347 (b)	0.009	0.954 (b)	0.002

Table 5. Coefficient of determination (r^2) obtained by regression analysis on the raw and rank-transformed data for different number of model runs

Means having the same letter are not different at the 5% probability level (post-hoc Bonferroni's test)

Results of replicability investigations are presented in Figures 7 (raw data) and 8 (ranked-transformed data) which provide a rapid and effective visual assessment of the stability of sensitivity results. The four charts in each of the two figures present results of sensitivity analyses carried out for the 10 replicated samples. Every chart corresponds to a different sample size, *i.e.* a different number of PELMO runs. Each of the 21 lines in the charts (one for each parameter) joins the 10 sensitivity ranks obtained for a particular parameter in the replicated sensitivity analyses. A chart with only horizontal lines would be obtained if the ranking of parameters would be stable whatever the random sample used. Charts reflecting poor replicability will appear disorganised.

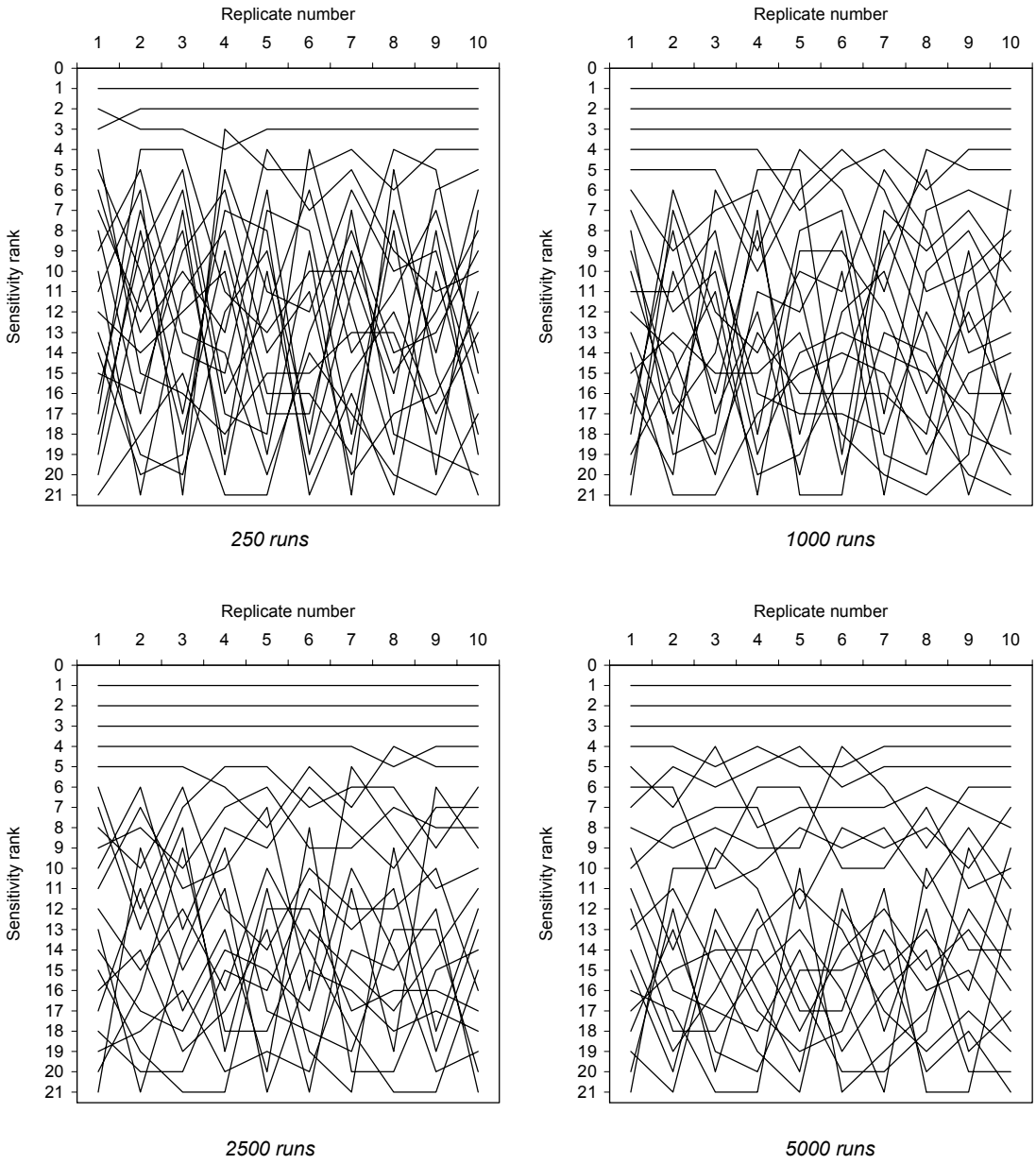


Figure 7. Stability in sensitivity ranking obtained for 10 different seed numbers and four sample sizes (untransformed data)

A larger sensitivity rank denotes a greater sensitivity of the model to this parameter.

Each line joins sensitivity rankings obtained for the 10 different seed numbers for one particular input parameter.

A horizontal line means that the same sensitivity ranking was obtained for the 10 replicated random samples

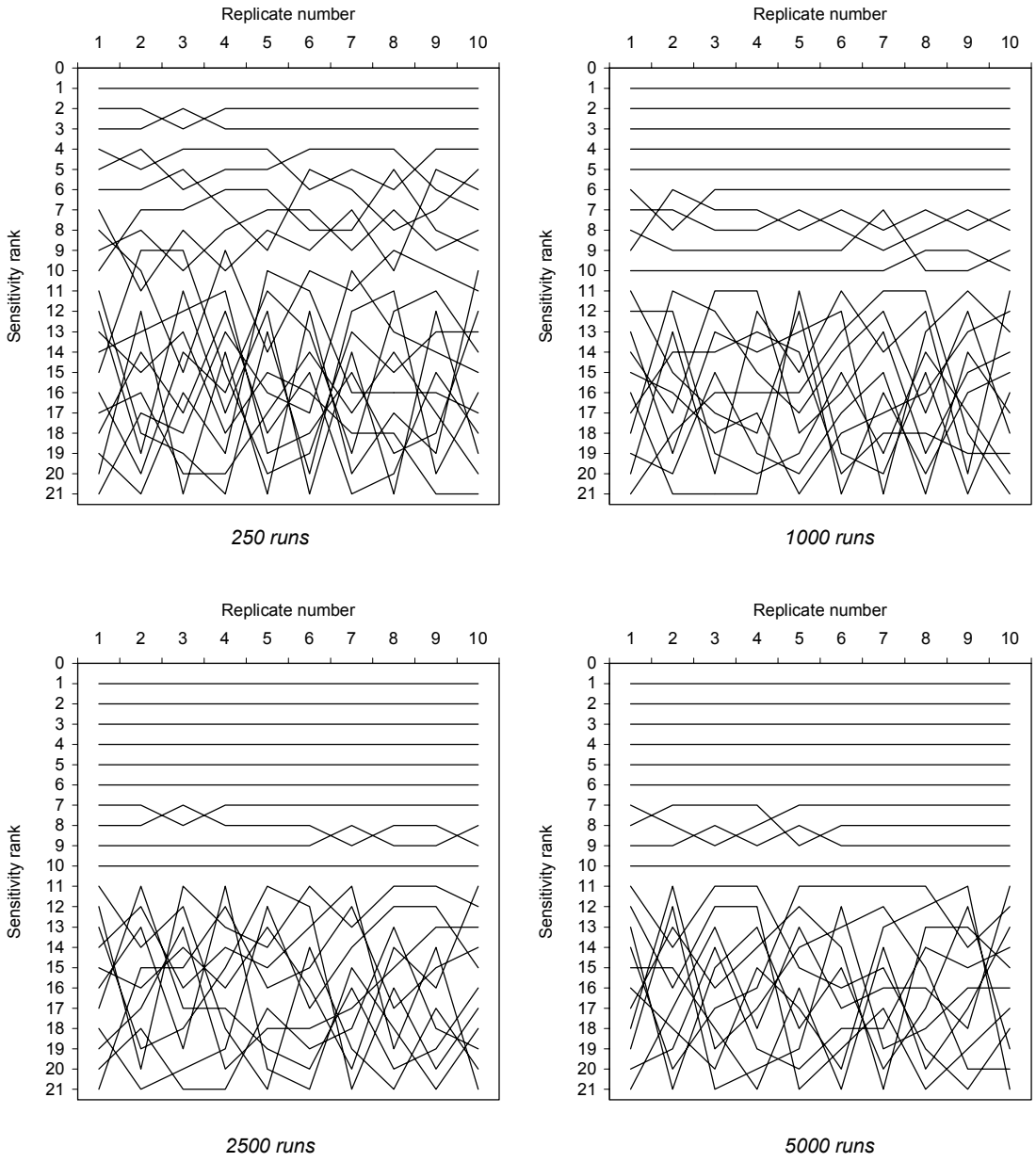


Figure 8. Stability in sensitivity ranking obtained for 10 different seed numbers and four sample sizes (rank-transformed data)

A larger sensitivity rank denotes a greater sensitivity of the model to this parameter.
Each line joins sensitivity rankings obtained for the 10 different seed numbers for one particular input parameter.
A horizontal line means that the same sensitivity ranking was obtained for the 10 replicated random samples

Figure 7 obtained for the raw data shows that replicability of sensitivity rankings was pretty poor except for the three most influential parameters. All charts were relatively disorganised with only the top three parameters displaying horizontal lines. Figure 9 presents SRC absolute values for PELMO input parameters together with their 95% confidence intervals for the case '5000 runs' and provides a heuristic explanation of the instability for the various

replicates observed for the less influential parameters. Confidence intervals for the first three most influential parameters are clearly separated whereas those for the remaining 18 parameters overlap. Although it should be noted that the strict conditions associated with the derivation of these confidence intervals in terms of residue randomness are not fulfilled when applying a regression approach to a deterministic model, the approach provides a tentative form of guidance as to whether the model output is significantly influenced by the various model inputs (Helton & Davis, 2000; 2002).

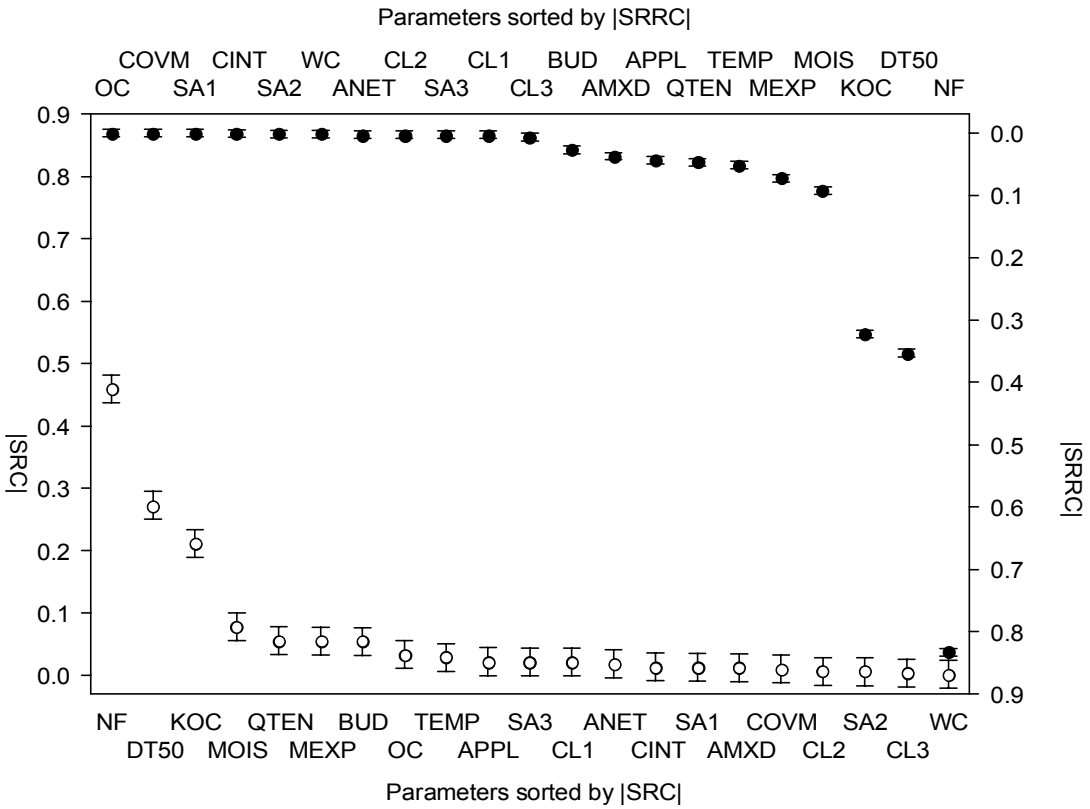


Figure 9. Classification of PELMO input parameters according to absolute values of SRC and SRRC indices (5000 runs)

Intervals for each point represent the 95% confidence intervals;
SRC (Standardised Regression Coefficient): open symbols;
SRRC (Standardised Ranked Regression Coefficient): closed symbols.

The rank transformation of the data led to an improvement in the stability of the ranking of PELMO input parameters for different replicated random samples (Figure 8). Also, the increased stability obtained by increasing the number of model runs was visible. Again, the relative stability of the most influential parameters for the case '5000 runs' was reflected to some extent by the absence of overlap of the confidence intervals for SRRC (Figure 9). Based on the SRRC results obtained for 5000 runs, the 10 input parameters which most

influenced PELMO predictions for pesticide were (parameter names are presented in Table 4):

NF > DT50 > KOC > MOIS > MEXP > TEMP ≈ QTEN ≈ APPL ≈ BUD ≈ OC

Parameters which had the largest influence on PELMO predictions for leaching were those related to pesticide sorption and degradation. These sensitivity results are in agreement with those obtained in one-at-a-time and MC-based sensitivity analyses undertaken for the PELMO model (Dubus *et al.*, 2000).

The present investigations demonstrated that the stability in the ranking of model input parameters according to their influence on model predictions will be dependent on the ranking of the parameters itself. For those parameters to which the model is more sensitive, sensitivity ranking will not be influenced by the use of different seed numbers in replicated samples. In contrast, the seed number used in the sampling will severely affect the ranking of those parameters with a lesser influence on model predictions. Replicability issues in sensitivity analysis, which have also been reported by Saltelli & Homma (1992) and Helton *et al.* (1995), will tend to decrease as the number of model runs is increased.

3.2 Replicability in probabilistic modelling

Probabilistic modelling exercises were conducted for different numbers of model runs and repeated for different seed number values used in the generation of the Latin hypercube samples. Model runs were processed to derive the probability that the average annual concentration in the 20th year of simulation will exceed 0.1 μg/l. Variability in the derived probabilities is presented in Figures 10 (no truncation in the sampling) and 11 (truncation in the sampling). The figures present the estimated minimum, average and maximum exceedance probabilities as well as the coefficient of variations obtained for the 10 different seed numbers. Results are presented for samples generated using the package Crystal Ball 2000. Similar results were obtained when the @RISK package was used (data not shown).

Significant variabilities in the exceedance probability obtained for different seed numbers were found for all sample sizes (Figures 10 and 11). Coefficients of variation for the 10 probabilities varied between 5.2 and 211% and between 9.3 and 162% for the non truncated and truncated case, respectively. There was an initially rapid then slow decrease in the coefficients of variation with increasing number of runs. Levels of variability for the largest number of PELMO runs considered (5000 model runs) were 5.2% and 9.3% for the two cases. The average exceedance probabilities for the 10 replicates was found to vary significantly

with the number of model runs undertaken. For the non-truncated case, the average exceedance probability over the 10 replicates was found to vary between 2 and 4.6% whereas that for the case where truncation in the sampling was considered ranged from 1.95 to 3.80%. The smaller range obtained in the latter case is likely to have been caused by the lack of sampling in the low and high probability tails of the log-normal distributions. The range of exceedance probabilities predicted (*i.e.* the distance between the minimum and maximum values) was found to be relatively large in both instances. For instance, the minimum and maximum exceedance probabilities when 500 PELMO runs were carried out were 2.4 and 5.2%, respectively (Figure 10). The probabilities of exceedance predicted when 5000 model runs were carried out for the two cases (3.43% when no truncation is considered and 2.84% when truncation is implemented) were significantly different at the 5% probability level. This shows that slight changes in the parameterisation of probability distributions (here the introduction of truncation at the 99% probability level) and in sampling procedures affect predicted exceedance probabilities.

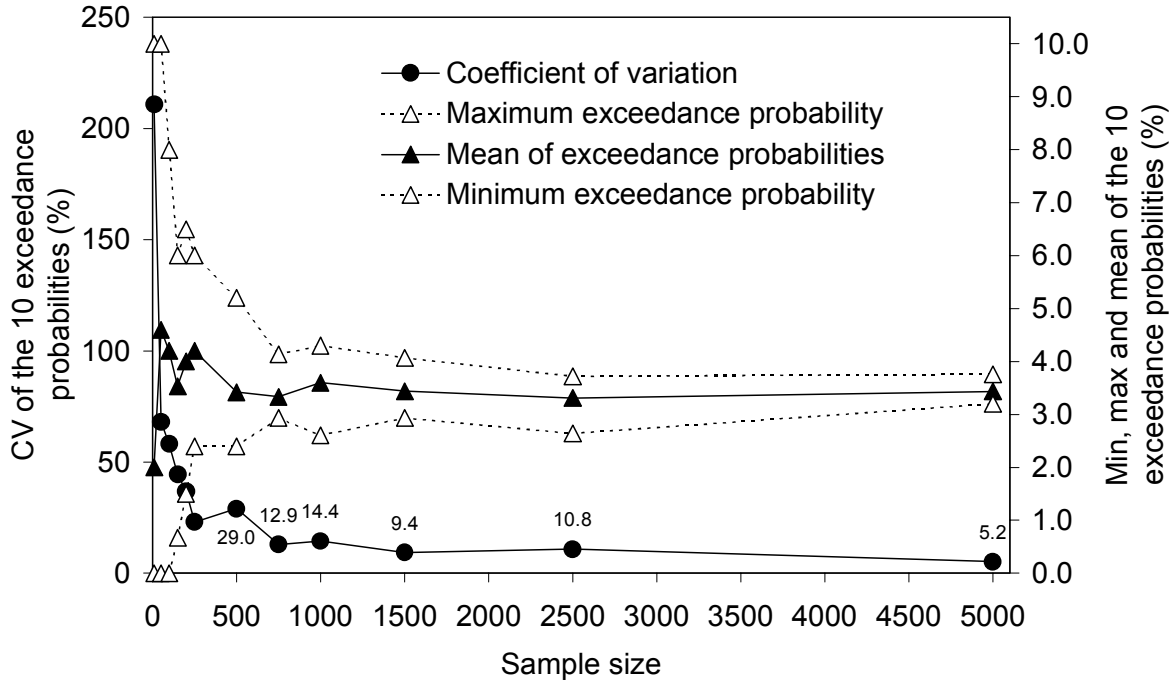


Figure 10. Variability in the probability of exceeding 0.1μg/l obtained when repeating Monte Carlo simulations 10 times with different seed numbers [no truncation in the sampling].
The sample size (equivalent to the number of PELMO runs undertaken for each of the 10 replicated samples) was varied from 10 to 5000.

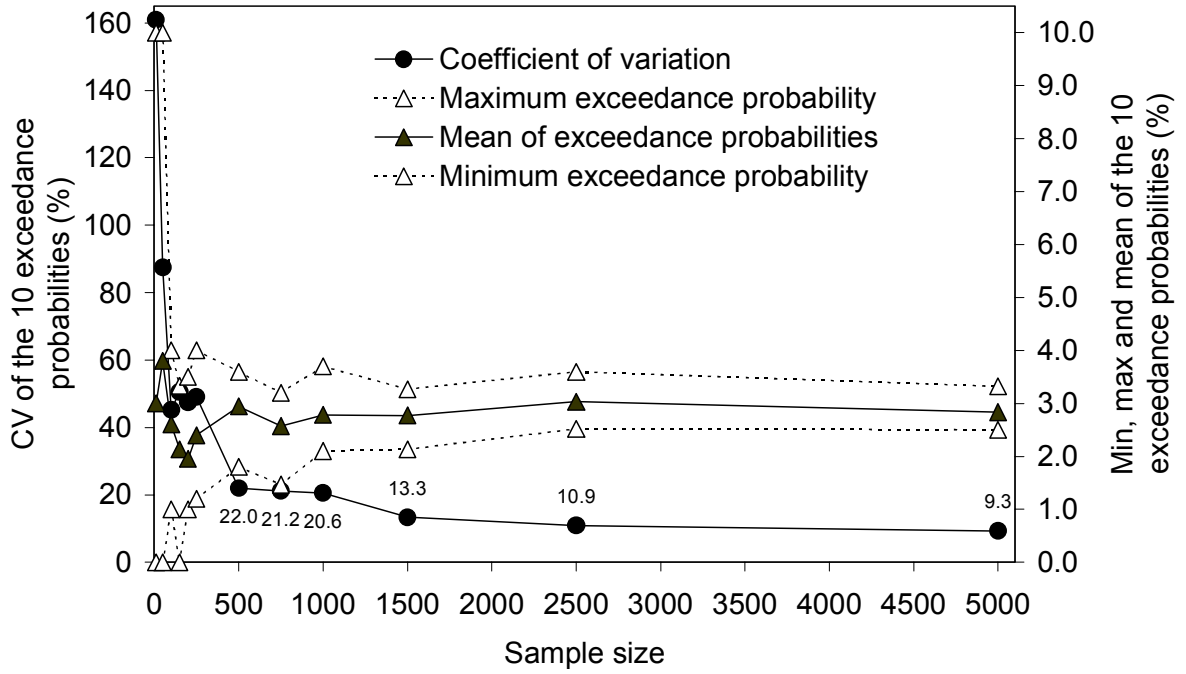


Figure 11. Variability in the probability of exceeding 0.1µg/l obtained when repeating Monte Carlo simulations 10 times with different seed numbers
[random samples were truncated at the 99% probability level; see text for details].
The sample size (equivalent to the number of PELMO runs undertaken for each of the 10 replicated samples) was varied from 10 to 5000.

The experimental results can be partially supported by theoretical considerations on the asymptotic behaviour of confidence intervals of probabilities to exceed a particular concentration, based on the fact that estimating a probability of exceedance on the basis of a random sample is equivalent to estimating the probability of success in a binomial distribution. Using the results of the binomial test in Conover (1971), the coefficient of variation of the non-parametric exceedance probability estimator (equation [2]) can be expressed as follows under conditions of random sampling and asymptotic behaviour:

$$CV = \frac{\sqrt{\frac{1}{p} - 1}}{\sqrt{N}} \qquad [3]$$

where CV is the coefficient of variation,
p is the (unknown) true exceedance probability,
N is the size of the random sample.

Applying equation [3] to the results obtained experimentally (N=5000 runs; p=0.0343 and p=0.0284) yields CV's of 7.5 and 8.3% for the non-truncation and truncation case, respectively, which agree to some extent with the empirical results (estimated CV's of 5.2 and

9.3% for the non-truncation and truncation case, respectively). Differences between empirical and theoretical results can be attributed to the facts that: i) empirical results are based on LHS while calculations of theoretical figures assume random sampling; and, ii) 'true' exceedance probabilities are unknown.

Iman and Helton (1985) consider that satisfactory results with regard to the coverage of the range of each parameter for evaluating the uncertainty in model output can be obtained with LHS if the size of the sample N follows the equation $N > 4/3 \times p$ (p being the number of parameters to be sampled). The results obtained here suggest that the equation should not be used to infer stability in results of sensitivity analysis and in estimates of exceedance probabilities derived by uncertainty analysis. A much larger sample size will be required for these purposes. From equation [3], it can be established that the number of random samples (N) should fulfil inequation [4] in order to be $100\times(1-\alpha)\%$ confident that the estimate of the exceedance probability p has no more than f% deviation (in relative sense) from its corresponding true (but unknown) value (Conover, 1971):

$$N \geq \kappa^2 \times \left(\frac{1}{p} - 1\right) \times \left(\frac{100}{f}\right)^2 \qquad [4]$$

where $\kappa = z\alpha/2$ denotes the $100\times(1-\alpha/2)$th percentile of the standard normal distribution. For $\alpha=0.05$ ($\kappa=1.96$), inequation [4] will provide the minimum number of random samples N which are required to be 95% confident that the estimated exceedance probability is within f% of its true value. Should we wish to have a high accuracy (say f=1%), then 153664, 345744 and 729904 model runs will be required for exceedance probabilities of 20%, 10% and 5%, respectively. It should be noted that inequation [4] is only valid for random sampling. Calculations based on this inequation when using LHS should be considered as upper-bounds as LHS is generally expected to lead to more efficient estimates asymptotically than random sampling (Stein, 1987).

Probabilistic modelling based on MC sampling currently receives much attention in the pesticide fate modelling community as a possible means to account for the uncertainty and variability in model inputs (*e.g.* ECOFRAM, 1999; EUPRA, 2001). MC approaches to uncertainty estimation have been applied successfully in numerous fields of science (Saltelli *et al.*, 2000). The main disadvantage of the approach is typically the large number of runs associated with the implementation of the approach (Helton, 1993). The need to keep the computational load to a minimum, in particular for models which are computationally expensive to run, has resulted in a widespread use of the stratified sampling technique Latin Hypercube Sampling, which allows a reduction in the number of runs to be undertaken. LHS

is traditionally considered a very efficient sampling scheme, but the present research demonstrates that care should be exercised when minimising the number of model runs for estimating exceedance probabilities. Uncertainties in the exceedance probability in the order of 5 to 10% (in terms of coefficients of variation) were noted when carrying out replicated probabilistic exercises for 5000 PELMO runs, while uncertainties of *ca.* 20-30% were obtained when 500 model runs were carried out. These levels of variability can be considered to be fairly large considering that only two parameters were included in the MC analyses.

Uncertainty levels such as those reported here will be of relevance to pesticide registration, especially if they are close to a threshold level used in decision making to grant authorisation for placement of pesticides on the market. Computational limitations in the present instance meant that finding a number of runs which would provide smaller levels of uncertainty (say CV 1%) by repeating analyses using larger sample sizes was not feasible. Although the present paper focused on the uncertainty in the probability of exceeding a particular threshold concentration, results suggests that similar issues of replicability will be encountered when considering specific percentiles in the distribution (data not shown). The research presented here was tailored to provide an exceedance level corresponding to a percentile of 3-5%, which represents the upper limit of exceedance levels that decision-makers will be faced with. Figure 12 shows the uncertainty in the estimated exceedance probability expressed in the form of coefficients of variation as a function of the probability level considered. The figure suggests the probability estimates will become more uncertain as they decrease. Percentiles used to support regulatory decisions are often much smaller than those considered in the present study (*e.g.* 0.1% for acute dietary assessments for pesticides; Wolt, 1999) and estimates are therefore expected to carry even larger uncertainty levels than those reported in the present study. It is expected that these uncertainty levels could be decreased by: i) increasing the sample size (running models for a larger number of times); ii) assessing the use of more efficient sampling schemes (*e.g.* importance sampling; Glasserman *et al.*, 2001); and, iii) investigating more efficient estimates of exceedance probabilities (*e.g.* (semi-)parametric methods; extreme value theory; Bassi *et al.*, 1998).

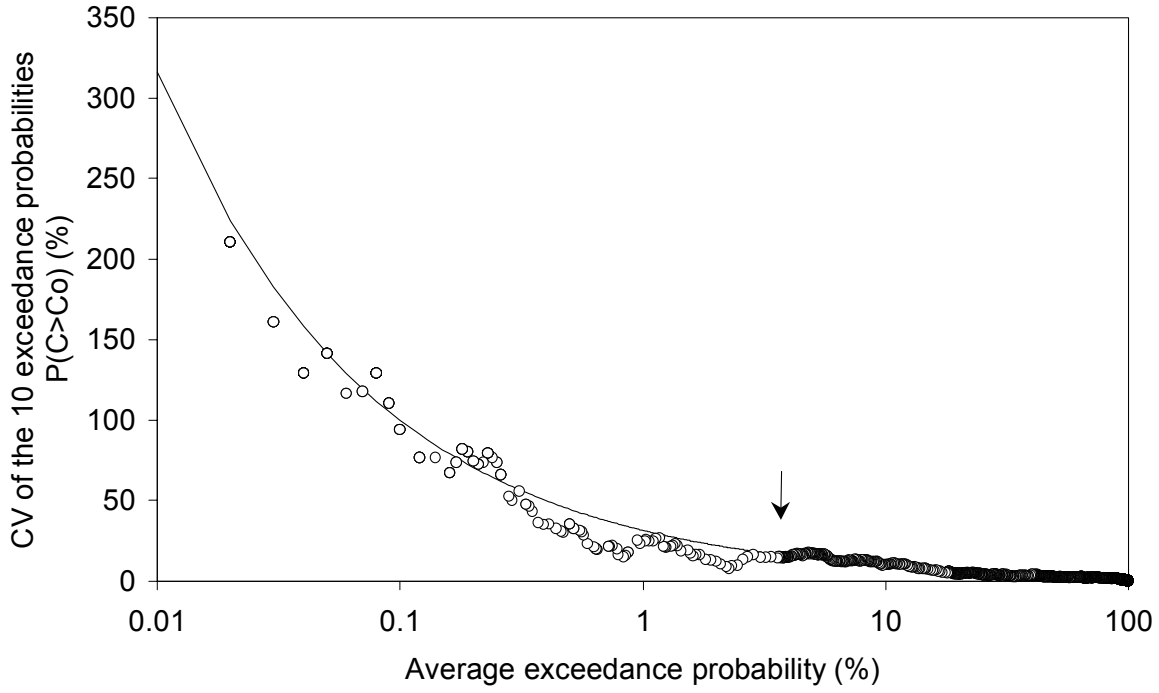


Figure 12. Variation of the coefficient of variation of exceedance probabilities for 10 replicated samples as a function of the level of exceedance probability

Dots indicate empirical results obtained in the present study while the plain line refers to theoretical estimates derived from equation [3] using an asymptotic approximation
The chart corresponds to the untruncated sample of size 1000;
the arrow indicates the average exceedance probability of 0.1 μg/l (3.6%)

4 CONCLUSIONS

The robustness of Monte Carlo approaches for assessing model sensitivity and the associated variability/uncertainty in model predictions was investigated for the pesticide leaching model PELMO. Replicated analyses for different seed numbers and for different sample sizes suggest that i) sensitivity results obtained through Latin Hypercube Sampling in combination with multiple linear regression are only stable for those parameters which most influence model predictions; ii) the robustness of the parameter ranking in sensitivity analyses can be assessed using charts plotting S(R)RC and their confidence intervals; iii) the estimation of exceedance probabilities from probabilistic modelling is inevitably subject to uncertainty which may affect subsequent decision making; and, iv) stability of results can be improved by applying a rank transformation to model input and output (sensitivity analysis) and by increasing the size of the random sample (both sensitivity and uncertainty analyses). The use of alternative, more efficient, techniques for sampling and estimating exceedance probabilities may also lead to a reduction in uncertainty levels in probabilistic modelling. Although present investigations focused on a pesticide leaching model, it is expected that the same

replicability issues will be found for environmental models which share the same general characteristics (non-linearity; large number of sensitive parameters; strong levels of model sensitivity).

The uncertainty resulting from the use of different random samples should be investigated in probabilistic modelling as it may affect subsequent decision making. The principal recommendation arising from the present research is that probabilistic analyses should be replicated a number of times (10 times were used here) or investigated theoretically to ensure that i) uncertainty/inaccuracy levels in results are low; or ii) where significant uncertainty/inaccuracy has been shown, it is appropriately accounted for in the decision-making process. Attention should be paid to replicability aspects by modellers when devising their approach to the modelling and by decision makers when examining results of probabilistic approaches.

The uncertainty introduced by the use of different seed numbers in the sampling is only one of a number of technical procedures in MC simulations which may lead to results being uncertain. Non-robustness in probabilistic results introduced by other technical aspects such as the influence of the attribution of probability density functions to model input parameters or correlation between parameters deserves investigation.

References

Bassi F., Embrechts P. & Kafetzaki M. (1998). Risk management and quantile estimation. In: A practical guide to heavy tails. Adler *et al.* (eds). Boston, Birkhaeuser, pp. 111-130.

Beven K. & Binley A. (1992). The future of distributed models – model calibration and uncertainty prediction. Hydrological Processes, 6:279-298 (1992).

Carsel R.F., Smith C.N., Mulkey L.A., Dean J.D. & Jowise P. (1984). User's manual for the Pesticide Root Zone Model (PRZM): Release 1. Report EPA-600/3-84-109. U.S. Environment Protection Agency, Athens, GA.

Carsel R.F., Parrish R.S., Jones R.L., Hansen J.L. & Lamb R.L. (1988). Characterizing the uncertainty of pesticide leaching in agricultural soils. Journal of Contaminant Hydrology, 2:111-124.

Conover W.J. (1971). Practical nonparametric statistics. John Wiley & Sons, New York, USA.

Decisioneering (2000). Crystal Ball 2000 – User manual. Decisioneering Inc., Denver, Colorado, USA.

Doherty J. (2000). PEST – model-independent parameter estimation. Watermark Computing, Corinda, Australia.

Dubus I.G. & Brown C.D. (2002). Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. Journal of Environmental Quality, 31:227-240.

Dubus I.G., Brown C.D. & Beulke S. (2000). Sensitivity analyses for leaching models used for pesticide registration in Europe. SSLRC Research Report to MAFF No. PL0532, Cranfield University, Silsoe, UK.

ECOFRAM (1999). http://www.epa.gov/oppefed1/ecorisk/

EUPRA (2001). Probabilistic risk assessment for pesticides in Europe. Implementation and research needs. Report from the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products, The Netherlands, June 2001.

Glasserman P., Heidelberger P. & Shahabuddin P. (2001). Efficient Monte Carlo methods for value-at-risk. In: Mastering Risk, vol. 2: Applications. Alexander C. (Ed.). Financial Times-Prentice Hall, London.

Hamby D.M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. Environmental Monitoring and Assessment, 32:135-154.

Helton J.C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliability Engineering and System Safety, 42:327-367.

Helton J.C., Johnson J.D., McKay M.D., Shiver A.W. & Sprung J.L. (1995). Robustness of an uncertainty and sensitivity analysis of early exposure results with the MACCS reactor accident consequence model. Reliability Engineering and System Safety, 48:129-148.

Helton J.C. & Davis F.J. (2000). Sampling based methods. In: Sensitivity Analysis, Saltelli A., Chan K. & Scott E.M. (Eds.), John Wiley & Sons, Chichester, UK. Pp. 101-153.

Helton J.C. & Davis F.J. (2002). Illustration of sampling-based methods for uncertainty and sensitivity analysis. Risk Analysis, 22:591-622

Iman R.L. & Conover W.J. (1979). The use of rank transformation in regression. Technometrics, 21:499-509.

Iman R.L. & Conover W.J. (1982). A distribution-free approach to inducing rank correlation among input variables. Communications in Statistics - Simulation and Computation, B11:311-324.

Iman R.L. & Helton J.C. (1985). A comparison of uncertainty and sensitivity analysis techniques for computer models. NUREG/CR-3904, SAND84-1461, Sandia National Laboratories, Albuquerque, New Mexico, USA.

Iman R.L. & Shortencarier M.J. (1984). A Fortran 77 program and user's guide for the generation of Latin hypercube and random samples for use with computer models. SAND83-2365. Sandia National Laboratories, Albuquerque, New Mexico, USA.

Klein M. (1994). Evaluation and comparison of pesticide leaching models for registration purposes. Results of simulations performed with the pesticide leaching model. Journal of Environmental Science and Health A, 29:1197-1209.

Klein M. (1995). User manual – PELMO, Pesticide Leaching Model, Version 2.01, Fraunhofer-Institut für Umweltchemie und Ökotoxikologie, Schmallenberg, Germany.

Klein M., Müller M., Dust M., Görlitz G., Gottesbüren B., Hassink J., Klowskowski R., Kubiak R., Resseler H., Schäfer H., Stein B. & Vereecken H. (1997). Validation of the

pesticide leaching model PELMO using lysimeter studies performed for registration. Chemosphere, 35:2563-2587.

Ma L., Ascough J.C., Ahuja L.R., Shaffer M.J., Hanson J.D. & Rojas K.W. (2000). Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. Transactions of the ASAE, 43:883-895.

McKay M.D., Conover W.J. & Beckman R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21:239-245.

Morgan M.G. & Henrion M. (1990). Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York.

Palisade (2000). Guide to using @RISK. Palisade Corporation, Newfield, New York, USA.

Perl (2002). Practical Extraction and Report Language. www.perl.com.

Saltelli A. & Marivoet J. (1990). Nonparametric statistics in sensitivity analysis for model output – a comparison of selected techniques. Reliability Engineering & System Safety, 28:229-253.

Saltelli A. & Homma T. (1992). Sensitivity analysis for model output. Performance of black box techniques on three international exercises. Computational statistics & Data Analysis, 13:73-94.

Saltelli A., Chan K. & Scott E.M. (2000). Sensitivity analysis. John Wiley & Sons, Chichester, UK.

Soutter M. & Musy A. (1999). Global sensitivity analyses of three pesticide leaching models using a Monte Carlo approach. Journal of Environmental Quality, 28:1290-1297.

Statsoft (2002). STATISTICA System Reference. Statsoft Inc., Tulsa, Oklahoma, USA.

Stein M. (1987). Large sample properties of simulations using Latin Hypercube Sampling. Technometrics, 29:143-151.

US EPA (1992). Pesticides in ground water database. A compilation of monitoring studies: 1971-1991. National Summary, United States Environmental Protection Agency, Washington, District of Columbia, USA.

Wolt J.D. (1999). Exposure endpoint selection in acute dietary risk assessment. Regulatory Toxicology and Pharmacology, 29:279-286.

Chapter 4

Dependence of Monte Carlo results on parameterisation assumptions (synopsis)

Investigations were carried out on the influence on probabilistic results of i) the parameterisation of statistical distributions attributed to model input parameters; and, ii) the attribution in the sampling of correlation between model input parameters.

A lognormal distribution was attributed to Koc and DT50, the two parameters considered for variation in the Monte Carlo modelling. The lognormal distribution was parameterised on the basis of five different realistic assumptions. Some of the random samples generated were truncated to reflect the belief of a modeller that the sampling of unrealistically small or large values for these parameters is inappropriate. The probability of exceeding a concentration of 0.1 μg/l was found to be significantly affected by the parameterisation of the probability density function. In the case where exceedance probabilities were <5%, truncation in the sampling resulted in a reduction of the exceedance probability from 4.4% to 2.0%. The range of variation of exceedance probabilities for the five realistic parameterisation assumptions were 1.2-4.4%, 44.8-49.2% and 95.6-100% depending on the level of exceedance considered. These levels of variability can be considered of significance within the context of decision-making.

A range of correlation levels between Koc and DT50 were introduced in the generation of random samples for these two parameters. Probabilities of exceedance of a particular threshold concentration were significantly influenced by the correlation considered. Attributing correlation between the two parameters either increased or decreased exceedance probabilities depending on the location of the threshold concentration with respect to the overall probability distribution. In most cases, the probability of exceeding 0.1 μg/l when no correlation was specified was larger than any of those obtained when correlations were introduced in the sampling.

These results suggest that the outcome of probabilistic modelling exercises will be significantly influenced by small changes in the parameterisation of probability density functions.

Chapter 4

Dependence of Monte Carlo results on parameterisation assumptions

1 INTRODUCTION

The previous chapter reported on investigations of replicability issues in Monte Carlo modelling and focused on the influence of the seed number used in the sampling on results of probabilistic modelling and sensitivity analyses. The use of replicated random samples (*i.e.* generated using different seed numbers) was found to significantly influence probabilistic results even when adequate sampling methods were used and numerous model runs were undertaken. In an effort to further estimate the robustness of Monte Carlo techniques when applied to pesticide fate models, investigations were carried out on the influence on probabilistic results of i) the attribution of statistical distributions to model input parameters; and, ii) the attribution of correlation between model input parameters. Investigations were based on synthetic datasets for Koc and DT50. The results were presented in the form of a poster at the 12th annual meeting of SETAC Europe in Vienna in 2000. A copy of the poster is provided at the end of the chapter.

2 MODELLING METHODS

2.1 Scenario used in the modelling

The pesticide leaching model PELMO was selected for these investigations on the basis of its relatively short running time. The scenario simulated and automation procedures were identical to those described in the previous chapter. In short, the leaching of a pesticide was simulated in the Borstel soil cultivated with maize using a 20-year climatic series for Hamburg ('Hamburg normal' dataset). The Koc and laboratory DT50 values initially selected for the modelling were 91.45 ml/g and 58.5 days, respectively. The pesticide was assumed to be applied to the soil on 15 May in each of the years simulated. The potential for leaching was assessed through the annual average pesticide concentration simulated in the 20th year of simulation.

2.2 Monte Carlo modelling

Monte Carlo modelling was undertaken for two parameters only, the Freundlich sorption coefficient normalised to organic carbon (Koc) and the laboratory half-life of the pesticide (DT50). Log-normal distributions were attributed to the two parameters. Based on computational limitations at the time of the study, 250 model runs were undertaken. The sampling was based on the Latin Hypercube Sampling method. The influence of the attribution of a particular seed number on probabilistic results (see previous chapter) was not investigated.

2.2.1 Influence of the parameterisation of statistical distributions on probabilistic results

The Koc values used to support parameterisation were assumed to range between 46 and 183 ml/g with a median of 91.45 ml/g while DT50 values were assumed to range between 29 and 117 days and to have a median of 58.5 days. Boundary values were obtained by dividing and multiplying the median values (M) by a factor of two. A total of five different assumptions were considered in the parameterisation of the probability density functions (pdf) attributed to Koc and DT50:

pdf 1 It is assumed that 95% of values are within the range defined by [M/2; M*2] and any value for Koc and DT50 can be sampled through random sampling.

pdf 2 It is assumed that 95% of values are within the range defined by [M/2; M*2] and that no value outside can be sampled between the 0.5^{th} and 99.5^{th}-percentiles of the log-normal distribution (truncation at the 99%-probability level).

pdf 3 It is assumed that 95% of values are within the range defined by [M/2; M*2] and that no value outside this range can be sampled (truncation at the 95%-probability level).

pdf 4 It is assumed that 99% of values are within the range defined by [M/2; M*2] and any value for Koc and DT50 can be sampled through random sampling.

pdf 5 It is assumed that 99% of values are within the range defined by [M/2; M*2] and no value outside this range can be sampled (truncation at the 99%-probability level).

It should be noted that all these assumptions are realistic and could be made by an individual or a group of individuals parameterising probability density functions. The introduction of truncation in the sampling would reflect the belief of the modeller that the sampling of

extremely small or large values of Koc and DT50 would not be realistic. The two levels of probability contained in the range [M/2; M*2] reflects possible discrepancies between individuals with regard to the chance of sampling an additional value outside the range [M/2; M*2]. In the case where it is assumed that 95% of values are within [M/2; M*2], it is considered that there is a 1/20 chance of getting a value outside this range if an additional experiment was carried out. In the case of the 99% probability level, this probability is reduced to 1/100. The different parameterisations for DT50 are presented graphically in Figure 13.

The level of the probability to exceed 0.1 µg/l can be expected to influence the results to some extent and the whole analysis was therefore repeated for three levels of exceedance: i) P(C>0.1 µg/l) < 5%; ii) P(C>0.1 µg/l) ≈ 50%; and, iii) P(C>0.1 µg/l) > 95%. The exceedance levels were adjusted by modifying the application rate of the substance.

2.2.2 Influence of the attribution of correlation on probabilistic results

In the vast majority of cases where Monte Carlo modelling has been carried out, model input parameters have been assumed to be independent although it can be argued that most input parameters of models are inter-related in some way or another. Not specifying correlation in the sampling means that no restriction is assumed on the combination of variables to be considered in the sampling.

The impact of specifying correlations between Koc and DT50 on results of Monte Carlo modelling was investigated for the scenario described earlier for the five types of parameterisation tested (pdf 1 to pdf 5, paragraph 2.2.1). Correlations that were studied were r = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0 and results of Monte Carlo modelling obtained for these correlations were compared to those obtained when no correlation was specified in the sampling. The range of random samples obtained is presented in Figure 14.

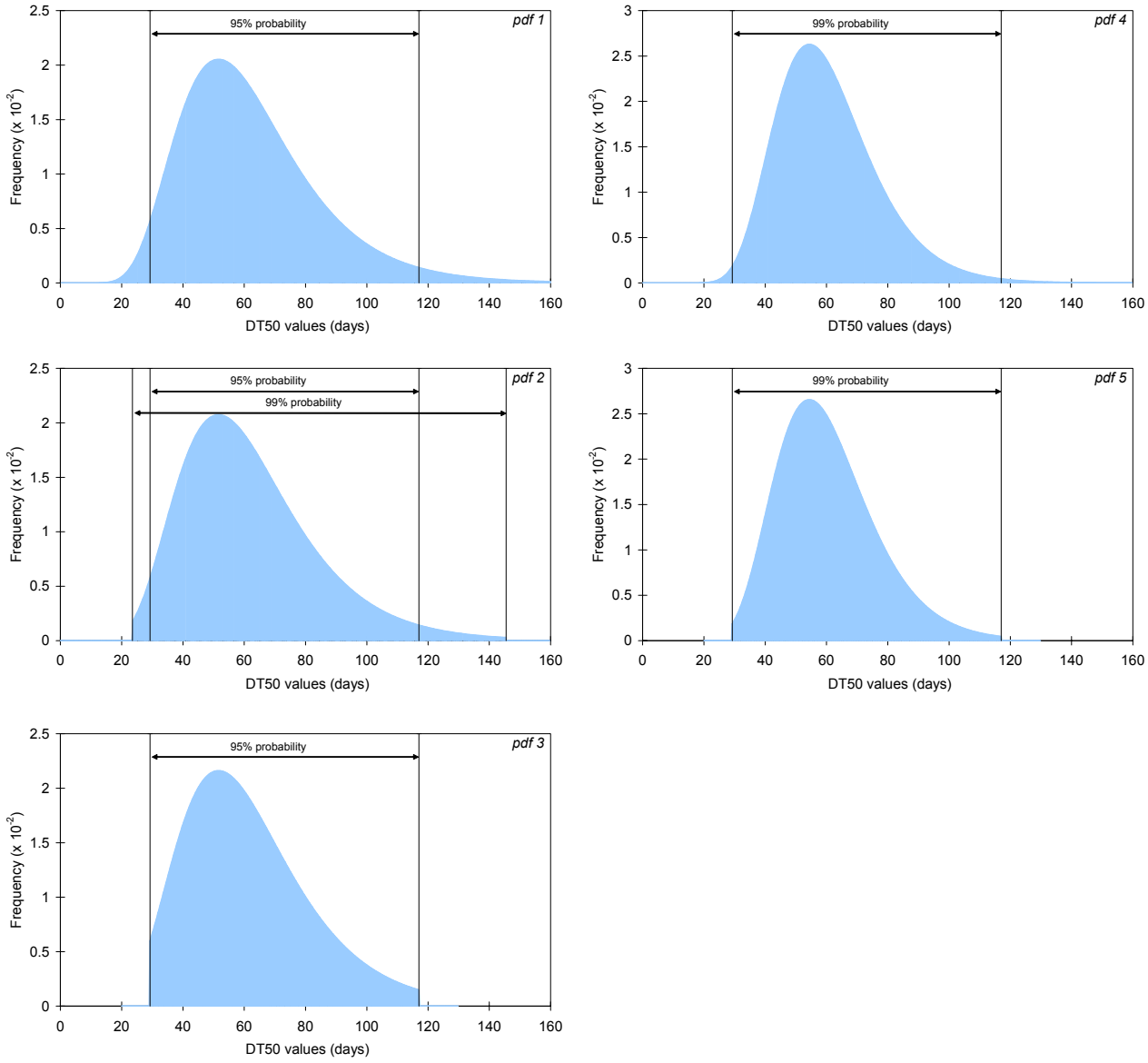


Figure 13. Graphical presentation of the five assumptions used for attributing probability density functions to DT50

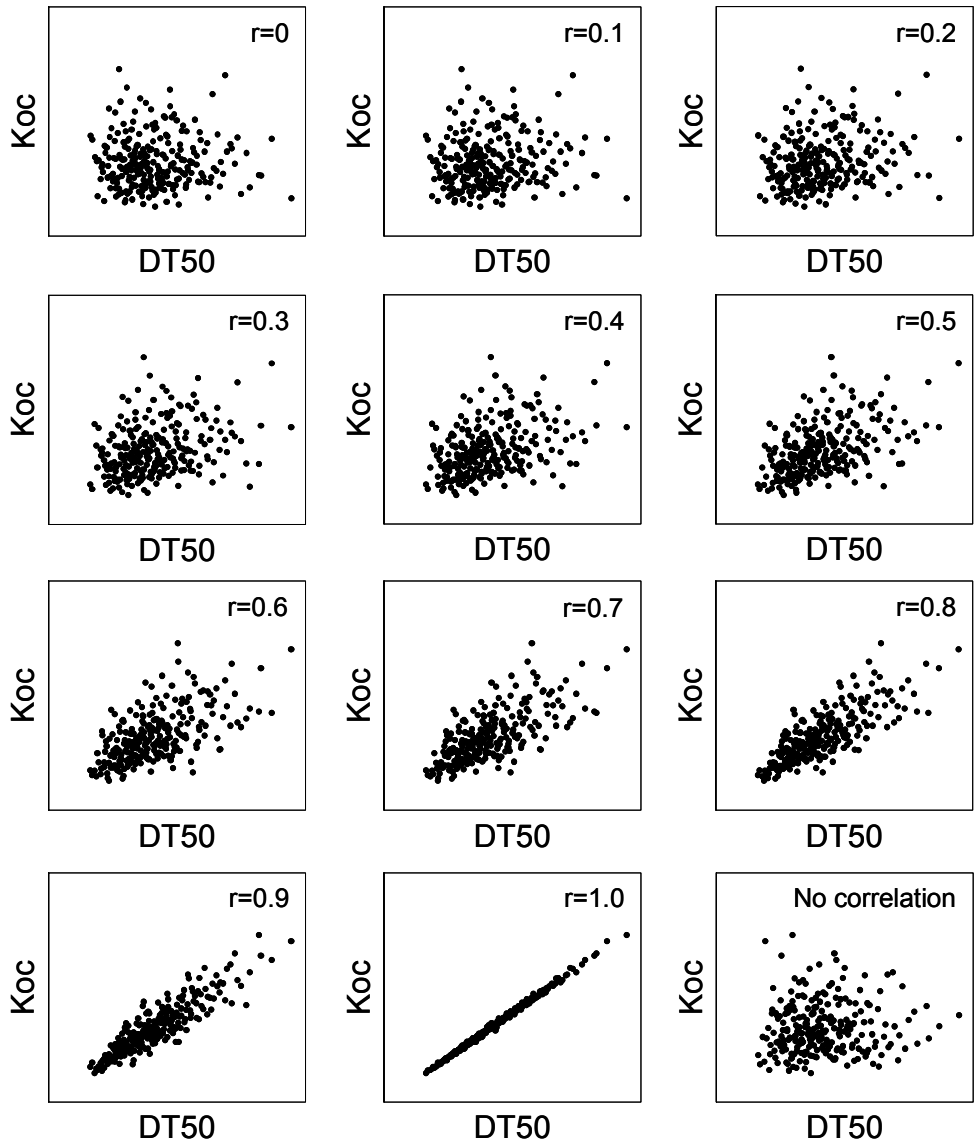


Figure 14. Random samples obtained for different correlation coefficients between Koc and DT50 (samples are presented for pdf 2 only)

3 RESULTS

3.1 Influence of the parameterisation of statistical distributions on probabilistic results

Cumulative distribution functions showing the cumulative probability of predicting a concentration less than 0.1 µg/l are presented in Figures 15 to 17 for the five different parameterisations of the log-normal distribution and for the three levels of exceedance.

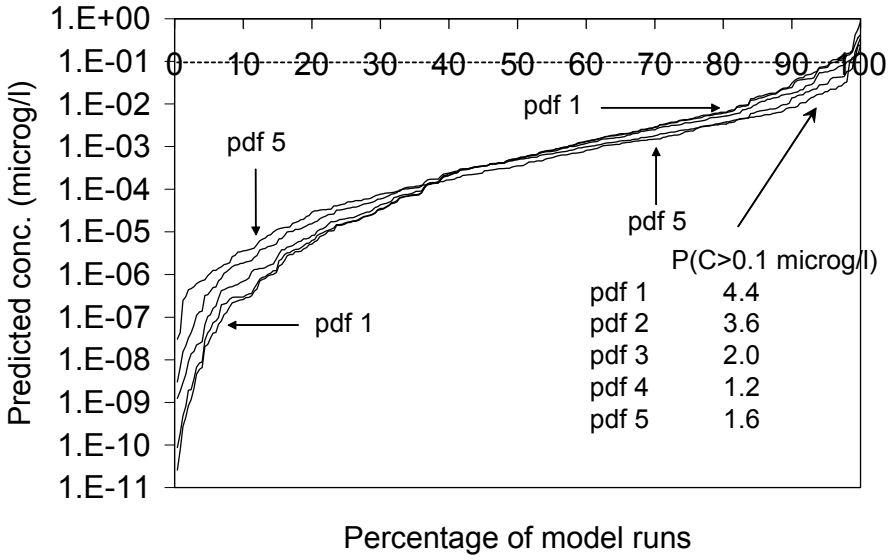


Figure 15. Cumulative distribution charts displaying the probability of simulating a concentration below a given concentration

The application rate of the substance was adjusted to provide a probability of exceeding 0.1 µg/l < 5%

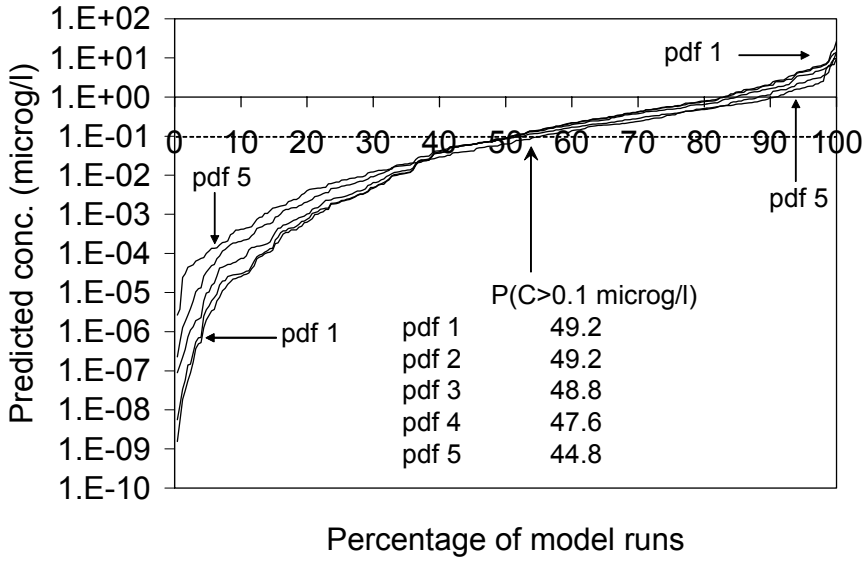


Figure 16. Cumulative distribution charts displaying the probability of simulating a concentration below a given concentration

The application rate of the substance was adjusted to provide a probability of exceeding 0.1 µg/l of about 50%

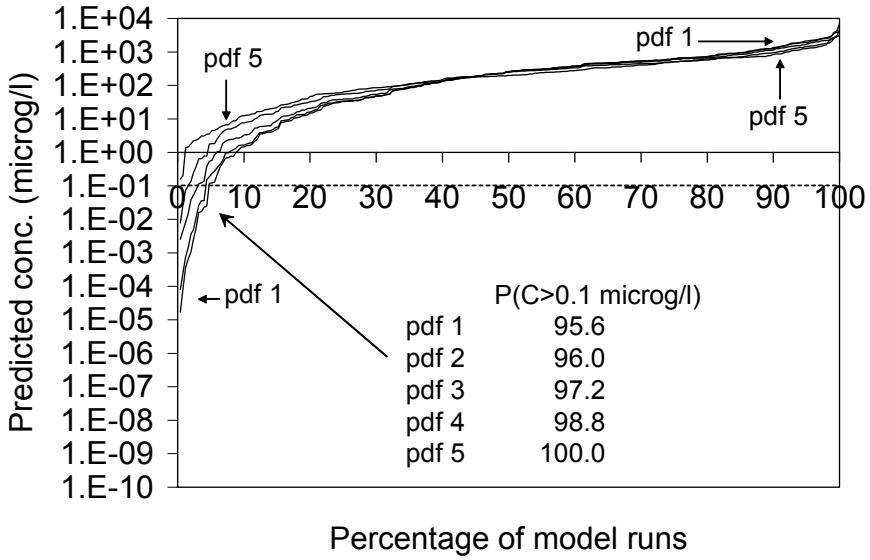


Figure 17. Cumulative distribution charts displaying the probability of simulating a concentration below a given concentration

The application rate of the substance was adjusted to provide a probability of exceeding 0.1 µg/l > 95%

Where exceedance probabilities were in the range of 5% or less (Figure 15), the different parameterisations of the log-normal distribution were found to influence the probability of exceeding 0.1 µg/l. These probabilities ranged between 1.2 and 4.4%. The use of truncation in the sampling resulted in a decrease in the probability of exceedance (pdf 1 *vs.* pdf 2; pdf 4 *vs.* pdf 5) and the tighter the truncation applied, the smaller the exceedance probability (pdf 2 *vs.* pdf 3). These levels of variability can be considered rather significant when put in the context of decision-making. The largest differences between the different cumulative distribution functions in Figure 16 were observed towards the tails of the distributions (*i.e.* for the smallest and largest percentiles).

In contrast, where exceedance probabilities were more than 95% (Figure 17), truncation resulted in an increase in the exceedance probability (pdf 1 *vs.* pdf 2; pdf 4 *vs.* pdf 5) and the tighter the truncation applied the larger the exceedance probability (pdf 2 *vs.* pdf 3). This change of influence is related to the change of positioning of the different curves in the cumulative distribution charts (Figures 15 to 17). Again, the influence of the use of different strategies for parameterising the lognormal distribution was found to influence exceedance probabilities and the level of resulting variability may be of importance to decision-makers.

Where intermediate levels of exceedance were simulated, probabilities of exceedance were more stable and less subject to variation as a result of the strategy used for parameterisation.

This is reflected in Figure 16 through the lesser variation of the different cumulative distribution curves in the horizontal direction in the middle of the chart.

3.2 Influence of the attribution of correlation on probabilistic results

The cumulative distribution charts obtained for a range of correlation coefficients between Koc and DT50 are presented in Figure 18 (assumption "pdf 1") and 19 (assumptions "pdf 2" to "pdf 5". The specification of correlation between the two parameters was found to have a strong influence on the cumulative distribution chart. Specifying a positive relationship between Koc and DT50 (r = 1.0) resulted in fairly flat cumulative distribution curves in Figures 18 and 19, which reflected the prediction of a small range of variation of pesticide concentrations by the model. Sorption and degradation processes compensated for one another in the prediction of pesticide leaching. The derivation of exceedance probabilities for fairly flat curves such as those reported for the large correlation coefficients will carry significant uncertainty.

Specifying a positive correlation between Koc and DT50 resulted in either an increase or decrease in percentile concentrations depending on the percentile considered. For percentiles less than *ca.* 50%, an increase in concentrations was obtained whilst a decrease was obtained for percentiles over *ca.* 50%. In the specific case considered here (probability of exceeding 0.1 µg/l of *ca.* 5% when no correlation is assumed), the stronger the correlation considered the smaller the exceedance probability. It should be noted that specifying correlations could result in severe influence on exceedance probabilities if the threshold concentration corresponded to relatively flat sections of cumulative distribution curves. This is reflected in Figure 20 which presents the variation of exceedance probabilities as a function of the correlation considered in the sampling for three threshold concentrations, 0.001, 0.01 and 0.1 µg/l. A large influence of correlations can be noted for a threshold concentration of 0.001 µg/l for r > 0.5.

In four of the five parameterisation scenarios considered, the probability of exceeding 0.1 µg/l when no correlation was specified was larger than any of the probabilities obtained when correlations were introduced in the sampling (Figures 18 and 19).

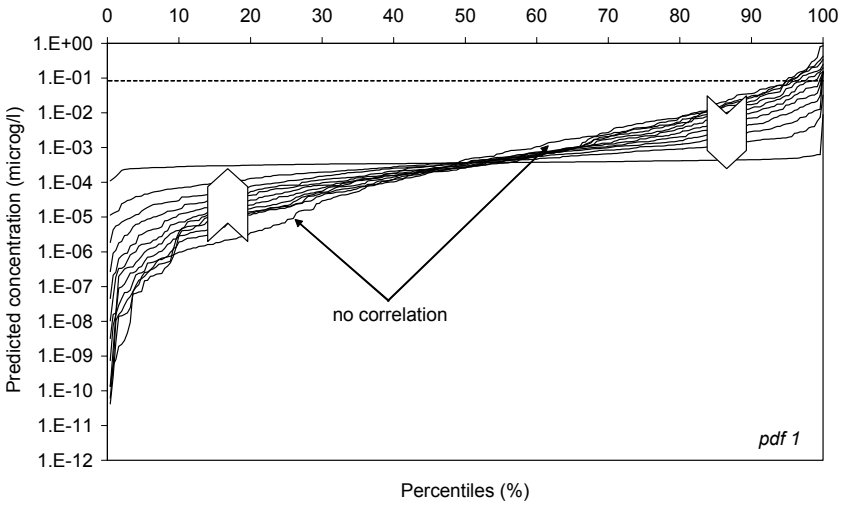


Figure 18. Cumulative distribution charts obtained for random samples generated using different correlation coefficients between Koc and DT50 (for the assumption "pdf 1")

The chevrons point towards curves obtained for the larger correlations

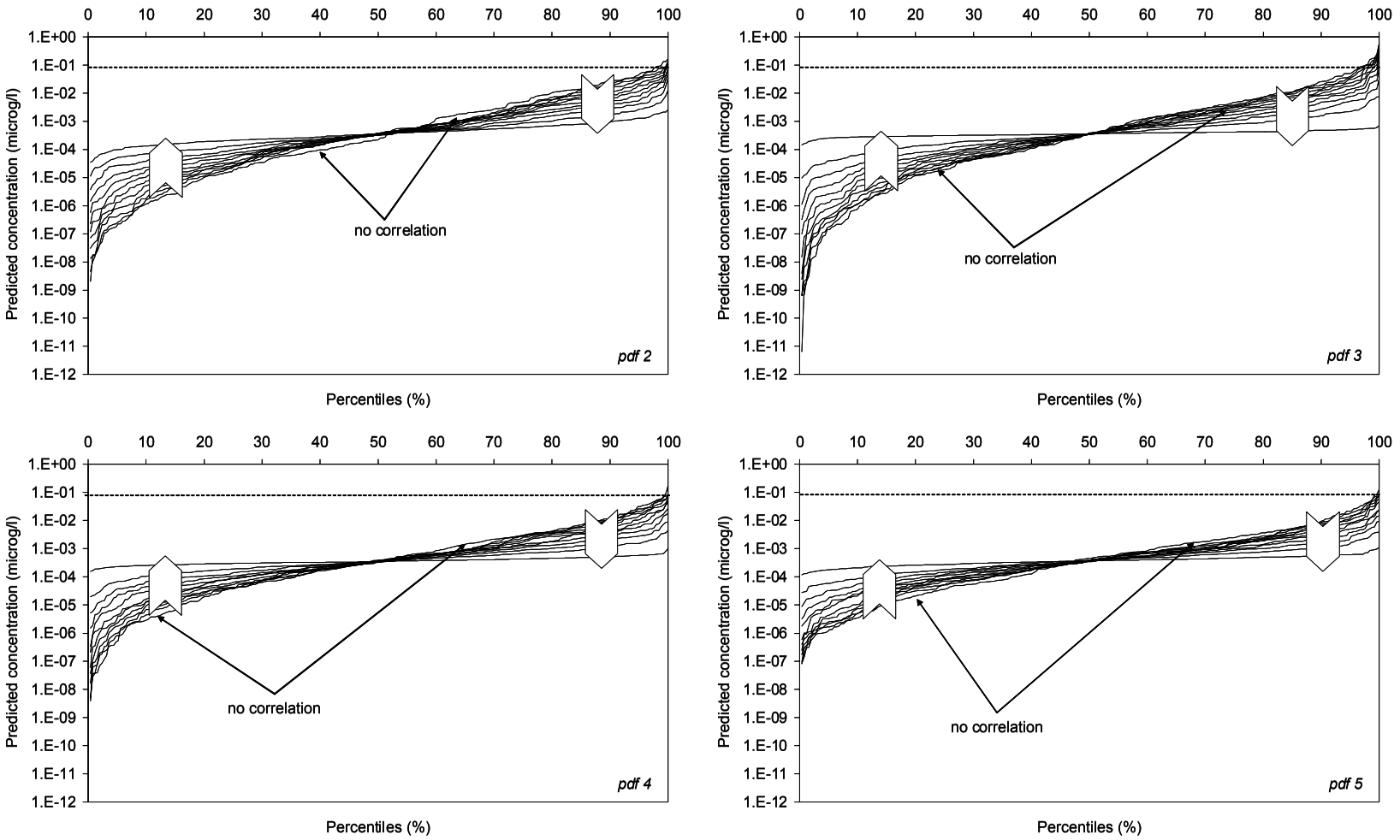


Figure 19. Cumulative distribution charts obtained for random samples generated using different correlation coefficients between Koc and DT50 (for the assumptions "pdf 2" to "pdf5")

The chevrons point towards curves obtained for the larger correlations

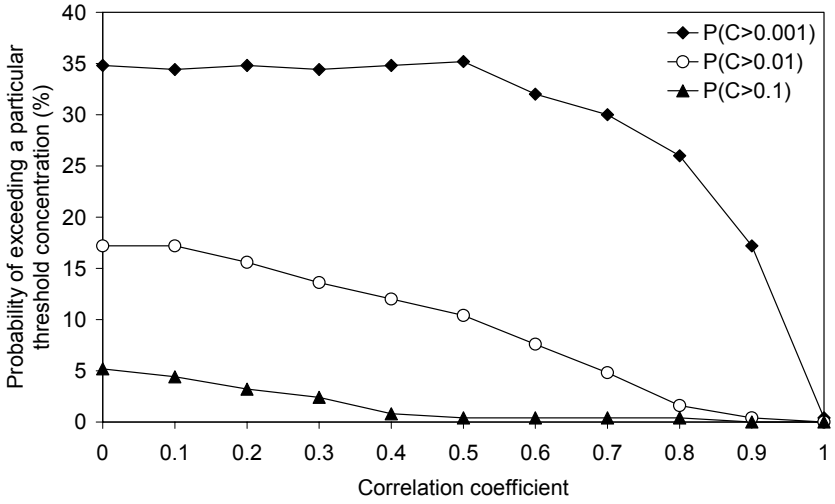


Figure 20. Influence of the correlation specified between Koc and DT50 on the probability of exceeding 0.001, 0.01 and 0.1 µg/l (assumption "pdf 1")

4 CONCLUSIONS

Investigations were carried out on the influence on probabilistic results of i) the attribution of statistical distributions to model input parameters; and, ii) the attribution of correlation between model input parameters. The results suggest that the outcome of probabilistic modelling exercises will be significantly influenced by small changes in the parameterisation of probability density functions. Levels of variability reported can be considered of significance when put in the context of decision-making.

PROBABILISTIC APPROACHES AND PESTICIDE EXPOSURE: A LONG WAY FROM HOME?

Igor Dubus*, Colin Brown & Sabine Beulke
Cranfield Centre for EcoChemistry, Cranfield University, Silsoe, Bedfordshire MK45 4DT, United Kingdom

Take-home messages

- **The appropriateness of Monte Carlo methods for providing refined assessments for pesticide exposure should not be taken for granted.**
- **Our investigations suggest that Monte Carlo results will be largely influenced by subjective choices (e.g. selection of pdf's; specification of correlation between parameters) and factors we have little control over (e.g. seed number used in the sampling scheme). This may restrict the usefulness of probabilistic approaches for assessing pesticide exposure.**
- **Monte Carlo modelling only considers the uncertainty in model predictions resulting from that in input parameters. There are numerous other sources of uncertainty in pesticide fate modelling that should not be ignored.**

Introduction

Probabilistic approaches are seen in Europe as a potential means to address the uncertainty in environmental exposure assessment for pesticides.

Monte Carlo modelling is used in numerous fields of science. One of the specificities of pesticide exposure assessment is that it relies on predictions of complex, non-linear, models. The following question should therefore be asked and addressed:

How much confidence should we place in results of Monte Carlo studies for the assessment of pesticide exposure?

The likely concentrations of a pesticide in groundwater were estimated using the leaching model PELMO and Monte Carlo simulations. Investigations were carried out to assess the influence on Monte Carlo results of: 1) the attribution of probability density functions; 2) the use of different seed numbers; and, 3) the specification of correlation between input parameters.

Approach

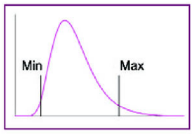
Context: assessment of the potential for leaching of atrazine to depth
Leaching model: PELMO 3.00
Leaching scenario: Borstel soil, Hamburg "normal" weather, annual application
Application rates selected to provide different ranges of P (C>0.1 µg/l): 2 scenarios
Target output variable: annual pesticide leaching in the 20th year of simulation

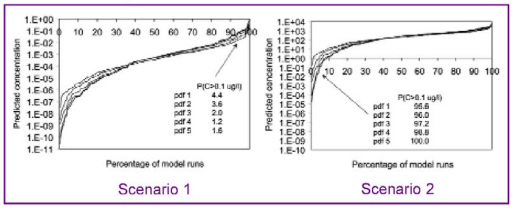
Parameters selected for MC modelling: Koc and DT50 (from sensitivity analysis)
Information available on parameter variability: assumed to be scarce
Sampling scheme: Latin Hypercube Sampling
Size of random samples: 250
Software for random sampling: Crystal Ball 2000

Attribution of probability density functions

Assumptions on variation of experimental values ("min-max"):
Koc: 46-183 ml/g; DT50: 29-117 days. Five different assumptions considered:

- pdf 1: 95% of values are within [min;max], no truncation of random samples
- pdf 2: 95% of values are within [min;max], truncation of samples at 99%
- pdf 3: 95% of values are within [min;max], truncation of samples at 95%
- pdf 4: 99% of values are within [min;max], no truncation of random samples
- pdf 5: 99% of values are within [min;max], truncation of samples at 99%

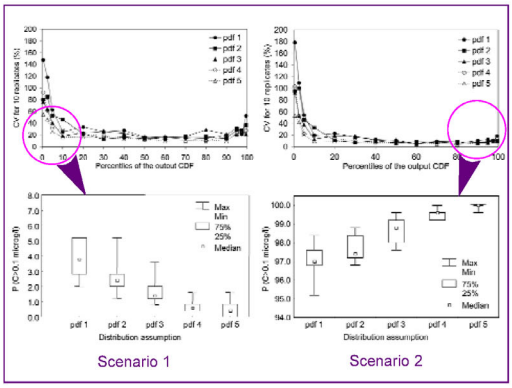




There are multiple ways to assign statistical distributions to experimental data. This will impact on probabilities of exceedance of a particular threshold. Do we need to envisage a standardised approach?

Repeatability

MC modelling was repeated 10 times using different seed numbers for the five pdf assumptions presented on the left

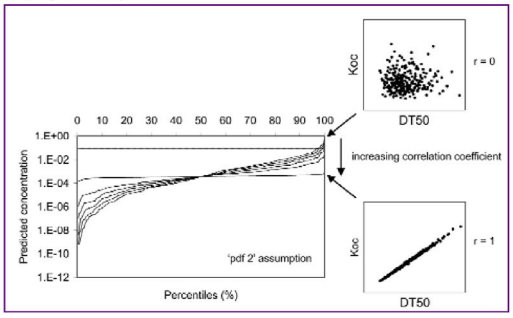


Results of Monte Carlo modelling are significantly dependent on the seed number used in the random sampling. Does this mean we have to repeat the analysis several times? Is this feasible? Desirable?

Correlation

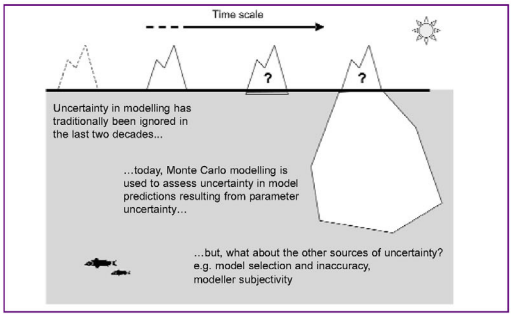
Koc and DT50 can be expected to be correlated to some extent, but attributing a correlation coefficient is difficult.

The influence of specifying different correlation coefficients between Koc and DT50 in the sampling was investigated.



Some sort of correlation between Koc and DT50 should probably be specified in the sampling. This correlation will largely influence the probability of exceedance derived. What should be the basis for selecting a correlation coefficient? Can the approach be standardised?

Sources of uncertainty



Is Monte Carlo only covering the tip of the uncertainty iceberg?

The funding of the project PL0548 "Uncertainty and probabilistic approaches to pesticide fate modelling" by DEFRA on behalf of PSD is gratefully acknowledged

To whom correspondence should be addressed:
i.dubus@cranfield.ac.uk tel: +44 (0)1525 863378
www.silsoe.cranfield.ac.uk/ecochemistry

Poster presented at the 12th annual meeting of SETAC Europe

(Vienna, 12-16 May 2002)

Chapter 5

A case study illustrating the use of probabilistic approaches for exposure assessment (synopsis)

Accounting for uncertainty in exposure assessments can take numerous forms. In the present instance, a range of probabilistic approaches of varying complexity were applied to predict potential leaching of atrazine within the context of a groundwater risk assessment. Calculations included two simple approaches based on FOCUS groundwater scenarios and two higher-tier approaches. These were: i) refined FOCUS modelling to account for the uncertainty in the application date of the compound; ii) refined FOCUS modelling to account for the uncertainty in sorption and degradation parameters in a simple way; iii) Monte Carlo modelling to account for the uncertainty in sorption and degradation parameters; and, iv) scenario-based modelling to account for the variability in environmental conditions within the use area of the compound. The PEARL model was used in the first three approaches whilst modelling with the preferential flow model MACRO was undertaken in the fourth. The approaches varied in their complexity and the number of model runs undertaken (from 12 to 10,000 runs).

The two refined FOCUS approaches and Monte Carlo modelling built directly on FOCUS modelling and expressed to a lesser or greater extent the uncertainty arising from that associated with timing of application or properties of the pesticide. Whilst it may be useful to attempt to quantify uncertainty for FOCUS estimations of exposure, each approach yields a distribution of potential outcomes spread around the original deterministic output. Whilst more scientifically defensible than a single deterministic prediction, the distributions are unlikely to further inform a regulatory decision where a standard FOCUS prediction exceeds 0.1 µg/l. The FOCUS scenarios are a consensus benchmarking system for leaching of pesticides whereby regulatory decisions can be reached based on a surrogate for concentrations in groundwater (*i.e.* annual average concentration in soil water leaching to 1 m depth). As such, the FOCUS approach rests on a framework of inter-related assumptions and it is probably inappropriate to impose higher tier modelling onto this framework. As already noted in Chapters 3 and 4, results of the Monte Carlo modelling were found to be dependent on subjective choices made during its implementation and to be affected by repeatability issues.

Scenario-based modelling was implemented by selecting 16 environmental scenarios covering the range of soil and climatic conditions to which atrazine will be applied in England and Wales and by running the MACRO model for each of the scenarios selected using long-term (20 years) weather sequences. Predicted concentrations in leaching were derived for each scenario. An overall probability of impact at the larger scale was obtained by weighting results for each scenario by their abundance in the broader landscape. In contrast to other approaches considered which were based on

one or a number of worst-case scenarios, scenario-based modelling permits an assessment that better reflects the variability in soil and climatic conditions across the proposed area of use. However, the typical application of deterministic models in scenario-based modelling ignores the uncertainty resulting from that in sorption and degradation parameters, which is known to be of importance for the prediction of pesticide loss. In the present instance, a first attempt at dealing with these sources of uncertainties was made by considering three different combinations of Koc and DT50. Both scenario-based modelling and Monte Carlo modelling are affected by issues related to the subjectivity in the implementation of the modelling approaches and do not consider aspects of uncertainty related to model error.

The Monte Carlo and scenario-based modelling are complementary to some extent. Monte Carlo modelling provides a solution to accounting for the uncertainty in model predictions resulting from that in model input parameters (and in particular to that in sorption and degradation values). In contrast, scenario-based modelling is particularly suited to the uncertainty in the modelling resulting from the consideration of a range of environmental conditions. Ideally, the two approaches should be combined. Although this integrated approach is desirable, its implementation is currently limited by the running time of pesticide fate models and the lack of knowledge about the robustness of probability estimates derived through Monte Carlo approaches. An additional constraint is the user-subjectivity associated with the selection of representative scenarios in scenario-based modelling.

Chapter 5

A case study illustrating the use of probabilistic approaches for exposure assessment

Important note

The following case study presents an assessment of the likelihood of atrazine being transferred to groundwater resources under UK conditions. Although some regulatory data were used in the exercise, it should be noted that the case study is largely theoretical. Certain aspects were selected to illustrate the potential utility of probabilistic assessments of exposure, but the assessment is far from comprehensive. Results are unlikely to be representative of the true leaching potential of atrazine.

1 INTRODUCTION

There is currently strong interest in the introduction of probabilistic approaches into risk assessment procedures for pesticides (ECOFRAM, 1999; EUPRA, 2001). The overall objective is to strengthen the risk assessment by accounting for uncertainties inherent in assessment procedures. Environmental risk assessments for pesticides are typically based on a comparison between likely concentrations to be found in the environment (exposure) and either ecotoxicological/toxicological endpoints (surface water assessment; groundwater risk assessment in the US) or threshold concentrations (0.1 μg/l for groundwater assessment in Europe). The present chapter aims at exemplifying the application of a number of probabilistic approaches to the assessment of exposure. The example is based on assessment of the likelihood of atrazine being transferred to groundwater resources in the UK. Atrazine was selected for this case study as the compound has been used in the past for assessing probabilistic approaches for exposure and effects, and numerous environmental fate data for the compound were available. These were kindly provided by Syngenta. The case study is largely hypothetical and far from exhaustive. For instance, it ignores higher tier experimental data and processes such as time-dependent sorption which may have a significant impact on the leaching potential of the compound.

Accounting for uncertainty and variability in exposure can take many different forms, from simple approaches considering the influence on exposure of considering a limited number of different environmental fate variables to more complex procedures requiring the running of a pesticide fate model for numerous iterations. This diversity was reflected in the present study through the use of five different approaches to exposure assessment (Table 6). The first approach is that recommended by the FOCUS groundwater working group (FOCUS, 2000) and does not explicitly account for any uncertainty, although it should be noted that the uncertainty associated with different climates and soils is indirectly dealt with. The second and third approaches are simple and direct extensions of the FOCUS framework to account to some extent for the uncertainty associated with pesticide properties and the application date. The fourth approach is a simple example of application of the popular Monte Carlo modelling technique and was devised to account for uncertainty in pesticide properties. Finally, the fifth approach (scenario-based modelling) aimed at representing the variability in environmental conditions at a large scale. Merits and shortcomings of the various approaches are discussed in detail later in the chapter.

	Approach	Description	Uncertainty accounted for	Model used and number of model runs	Output of the approach
1	'Standard' FOCUS approach	Twenty-six-year runs for the four FOCUS scenarios which are considered representative for UK conditions (Châteaudun, Hamburg, Kremsmünster and Okehampton). Median Koc and DT50 values were used.	*Implicitly:* that associated with different locations and different years	PEARL: One 26-year run for each of the four scenarios, hence *4 runs* MACRO: One 26-year run for the Châteaudun scenario	5 concentrations
2	Simple refinement to the FOCUS modelling approach #1	FOCUS modelling for the four representative scenarios using 3 different combinations of Koc and DT50. Median case: median Koc, median DT50 Better case: 75th-percentile Koc, 25th-percentile DT50 Worse case: 25th-percentile Koc, 75th-percentile DT50	*Explicitly:* that associated with Koc and DT50 (simple approach) *Implicitly:* that associated with different locations and different years	PEARL: One 26-year run for each of the 4 scenarios and for each of the 3 combinations of Koc and DT50, hence *12 runs*	12 concentrations
3	Simple refinement to the FOCUS modelling approach #2	FOCUS modelling for the four representative scenarios using different application dates within a two-week application window (median case used)	*Explicitly:* that in the application date *Implicitly:* that associated with different locations and different years	PEARL: 15 26-year runs for each of the 4 scenarios, hence *60 runs*	Four series of 15 concentrations
4	Monte Carlo modelling	Monte Carlo modelling for the Okehampton scenario (adopting the FOCUS approach)	*Explicitly:* that associated with Koc and DT50 and that associated with different years	PEARL: 500 26-year model runs (arbitrarily selected) were repeated 5 times with different seed numbers, hence *2500 runs*	For the particular scenario selected, the probability of exceeding 0.1 μg/l and concentrations for a given percentile
5	Scenario-based modelling	Modelling with 30-year weather series for 16 scenarios (4 soils, 4 climates) selected to represent the maize growing area overlying aquifers in England and Wales where the compound will be used. The modelling was repeated for three combinations of pesticide properties (see approach #2 for details)	*Explicitly:* that associated with soil and climatic conditions where the compound is used, that associated with different years, that associated with Koc and DT50 (simple approach)	MACRO: 1 30-year run for each of the 16 scenarios considered and for each of the 3 combinations of Koc and DT50. hence *48 runs*	For each scenario, series of 30-year concentrations and probabilities of exceeding 0.1 μg/l A probability of exceeding 0.1 μg/l and concentrations for a given percentile for the maize growing area in England and Wales

Table 6. Probabilistic approaches used in the present work (exposure assessment for atrazine within the context of a groundwater risk assessment)

Care was taken to reflect current risk assessment procedures in the EU and most of the techniques were thus implemented using the FOCUS groundwater framework as a starting point. The work was based on PEARL and MACRO, two pesticide leaching models which have been widely used in the EU within the context of pesticide registration. The preferential flow model MACRO was used for the scenario-based modelling as preferential flow can be expected to play a significant role in the transfer of pesticide through the soil profile in England and Wales.

It should be noted that the main objective of the work was to demonstrate the possibilities offered by the different approaches used for probabilistic modelling rather than undertaking in-depth assessments for each of these approaches. This work should be seen as a starting point to initiate discussions on a possible use of probabilistic approaches in exposure assessment and not as guidance for performing probabilistic assessments of exposure.

2 'STANDARD' FOCUS MODELLING

2.1 Modelling methods

FOCUS groundwater modelling (FOCUS, 2000) was undertaken for the four FOCUS scenarios which are considered representative of conditions in England and Wales (Hollis *et al.*, 2001; Renaud & Brown, 2001): Châteaudun, Hamburg, Kremsmünster and Okehampton. The model selected for the present leaching assessment was the PEARL model (Tiktak *et al.*, 2000) in its FOCUS version 1.1.1. The selection was based on the fact that PEARL is a Richards' equation model with a relatively short running time. The preferential flow model MACRO was also used to assess the potential for atrazine to leach to depth for the Châteaudun scenario. Atrazine was assumed to be applied to a maize crop at an application rate of 1 kg a.s./ha 7 days after emergence. Sorption properties used in the modelling were Koc 91.45 ml/g and nf 0.895. These were taken as the median of 20 values from company data. A significant number of field DT50 values (61) were available for atrazine and the use of these values was preferred over that of laboratory data. Consequently, subroutines for correction of degradation for variations of temperature and moisture were switched off by setting the 'molar activation energy' and the 'exponent for the effect of liquid' to zero (PESTLA) and by setting the exponents in the temperature response (TRESP) and in the moisture relation (EXPB) to zero (MACRO). The field DT50 value used in the modelling

was taken as the median of the 61 values derived from field dissipation experiments (43 days). The subroutine for volatilisation was turned off in PESTLA by setting the saturated vapour pressure to zero.

2.2 Results for the 'standard' FOCUS modelling

Results obtained through a standard FOCUS modelling approach are presented in Table 7. The table presents 80th percentile annual average concentrations at 1-m depth from a series of 20 years of simulation predicted by PEARL for each of the four scenarios which are considered relevant to conditions in England and Wales.

Scenario	80th percentile concentration (μg/l)
Châteaudun	0.084
Hamburg	0.041
Kremsmünster	0.060
Okehampton	0.206

Table 7. Results obtained through the standard FOCUS modelling approach using PEARL

Eightieth percentile annual average concentrations of atrazine were predicted to be <0.1 μg/l for three of the four scenarios considered. The largest concentration (0.21μg/l) was simulated for the Okehampton scenario. The 80th percentile concentration when using MACRO for the Châteaudun scenario was 0.31 μg/l.

3 REFINED FOCUS MODELLING

Two attempts to account for uncertainty in a simple way while remaining within the constraints of the FOCUS approach were considered. These are presented below.

3.1 Modelling methods

3.1.1 Effect of sorption and degradation properties on predicted leaching

Sorption and degradation parameters are known to be variable and uncertain (Chapter 2). Two additional sets of Koc and field DT50 values were used in the modelling to provide a simple first-step assessment of the uncertainty in predicted concentrations resulting from these two uncertain parameters (Table 8). The 'better' case was constructed by selecting the 25th

percentile DT50 and the 75th percentile from the range of values available from company data (61 field DT50 values, sorption data for 20 soils) while the 75th and 25th percentiles for DT50 and Koc, respectively, were used for the 'worse' case. Percentiles were determined in Microsoft Excel using the 'percentile' function. The selection of properties was not meant to represent best or worst case conditions for leaching, but rather to give an idea about the uncertainty associated with FOCUS predictions for atrazine (hence the use of 25th and 75th percentiles and wording 'better' and 'worse' cases). The approach may not be appropriate for those compounds (*e.g.* ionic pesticides) which show a strong relationship between sorption and degradation. The modelling was undertaken using PEARL. The same assumptions for application date and rate as those used in the standard FOCUS modelling were made, *i.e.* application at 1 kg a.s./ha 7 days after emergence. To be consistent with the previous modelling, the main output considered was the 80th percentile annual average concentration as directed by FOCUS guidelines (FOCUS, 2000).

	Koc (ml/g)	Field DT50 (days)
Median case	91.45 ml/g (50th percentile)	43 days (50th percentile)
Better case	116.5 ml/g (75th percentile)	29 days (25th percentile)
Worse case	79 ml/g (25th percentile)	60 days (75th percentile)

Table 8. Combination of Koc and DT50 values used in the FOCUS modelling
The percentiles considered are given in parentheses

3.1.2 Effect of application date on predicted leaching

An application date of 7 days after emergence was arbitrarily selected in the standard FOCUS modelling (section 2 of the present chapter). The date of pesticide application in relation to the main rainfall events is expected to have some influence on predicted concentrations in leachate. Since FOCUS GW scenarios do not integrate a procedure for selecting application dates in relation to rainfall, a range of application dates were considered in the modelling. The application date for atrazine was varied between –7 and 7 days from the target application date (15 application dates in total). The analysis was carried out for the four scenarios previously considered. Please note that the selection of boundary application dates was arbitrary in the present instance. The selection of these dates would normally be based on the Good Agricultural Practice table for the compound of interest. PEARL runs were only undertaken for the set of median values for Koc and field DT50 values (*i.e.* the modelling was not repeated for the better and worse cases). Again, the same assumption for application rate was used, *i.e.* application at 1 kg a.s./ha. To be consistent with the previous modelling, the main output considered was the 80th percentile annual average concentration as directed by FOCUS guidelines (FOCUS, 2000).

3.2 Results for the refined FOCUS modelling

3.2.1 Effect of sorption and degradation properties on predicted leaching

In total, three combinations of Koc and DT50 values were used in the modelling with PEARL (Table 8). Combined with the four environmental scenarios considered, the analysis resulted in 12 predicted concentrations (Table 9).

	80th-percentile concentration (μg/l)		
Scenario	Better case	Median case	Worse case
Châteaudun	<0.001	0.084	1.307
Hamburg	<0.001	0.041	0.597
Kremsmünster	<0.001	0.060	0.815
Okehampton	0.001	0.206	1.903

Table 9. Results of a refined approach to FOCUS modelling to account for the effect of sorption and degradation properties on PEARL predictions

PEARL predictions were found to vary by several orders of magnitude. This reflects the large sensitivity of model predictions to changes in these parameters (Dubus *et al.*, 2000). The maximum concentrations for all four scenarios in the worse case were found to exceed the 0.1 μg/l threshold, reflecting the importance of taking uncertainty into account in the decision-making process.

3.2.2 Effect of the application date on predicted leaching

The influence of the uncertainty associated with the application data on PEARL predictions for atrazine leaching was investigated by varying the date of application -7 to +7 days from the default application date initially considered (7 days post emergence). The investigations concentrated on the median values for Koc and DT50, *i.e.* no attempt was made to account simultaneously for the uncertainty in pesticide properties (see previous section) and that in the application date. Results for the four scenarios are presented graphically in Figures 21 to 24.

Predicted concentrations were only marginally affected by the limited change in the application date when compared to the results obtained for the influence on the modelling of pesticide properties (previous section). Still, the effect could be of significance for other

compounds and those situations where the leaching is close to the 0.1 μg/l level or if larger variations in the application date were considered.

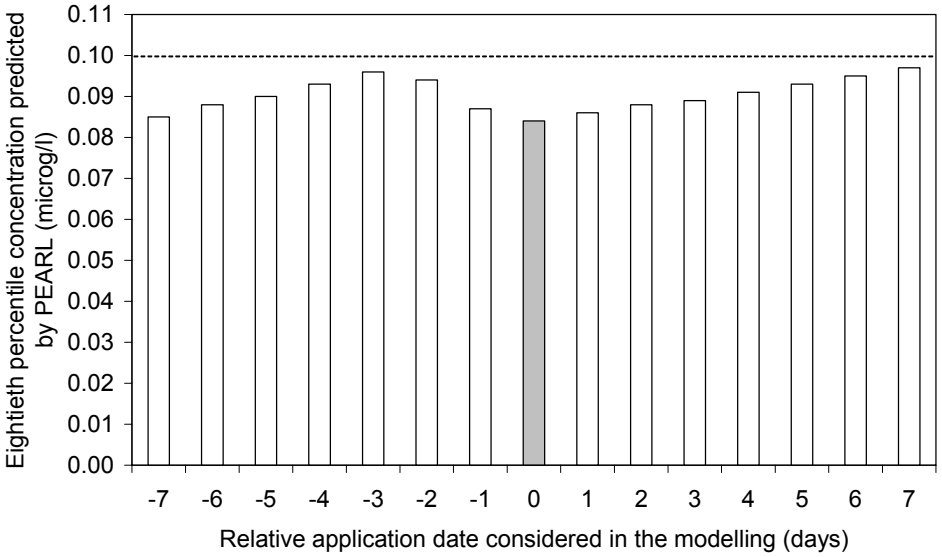


Figure 21. PEARL predictions for 80th percentile concentrations obtained when varying the application date by + / - 7 days (Châteaudun scenario).

The default application date selected in the earlier modelling (7 days post emergence) is shown in grey. The dotted line indicates the 0.1 μg/l leaching level.

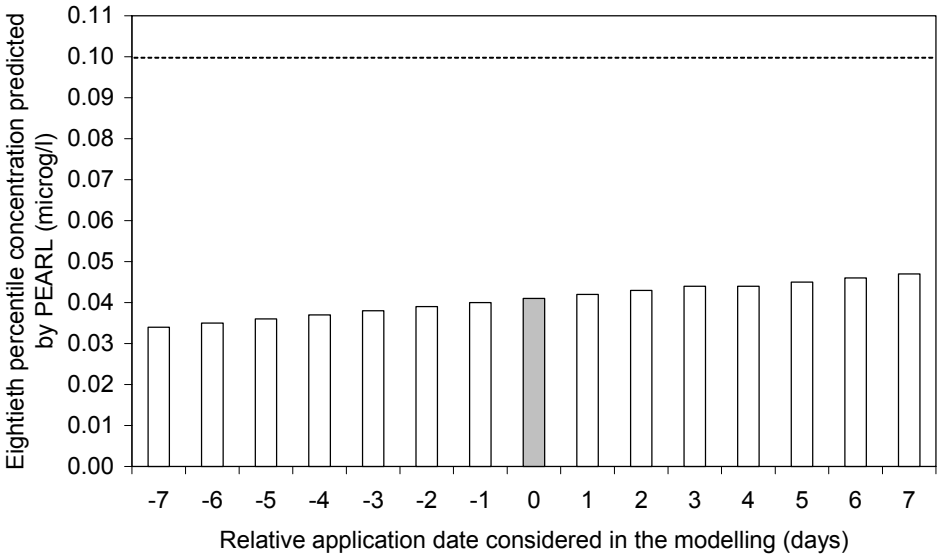


Figure 22. PEARL predictions for 80th percentile concentrations obtained when varying the application date by + / - 7 days (Hamburg scenario).

The default application date selected in the earlier modelling (7 days post emergence) is shown in grey. The dotted line indicates the 0.1 μg/l leaching level.

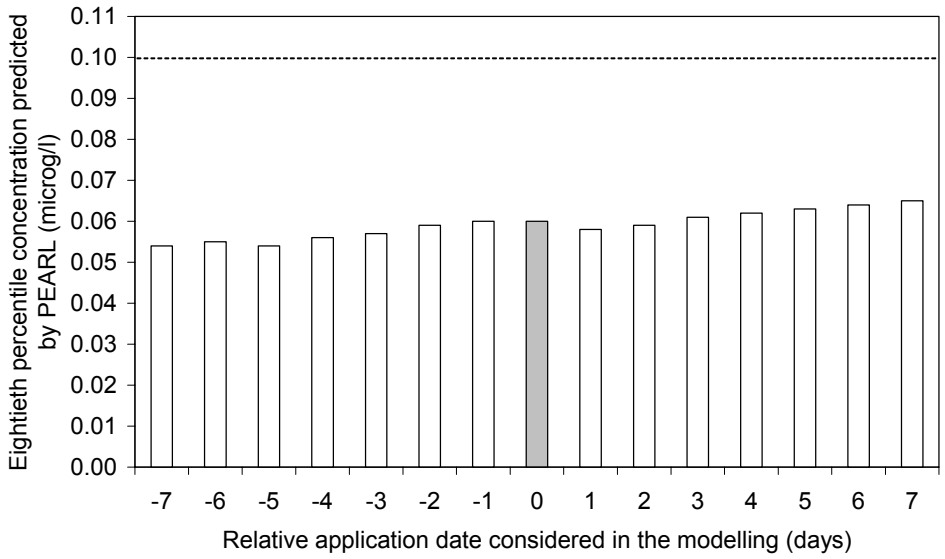


Figure 23. PEARL predictions for 80th percentile concentrations obtained when varying the application date by + / - 7 days (Kremsmünster scenario).

The default application date selected in the earlier modelling (7 days post emergence) is shown in grey. The dotted line indicates the 0.1 μg/l leaching level.

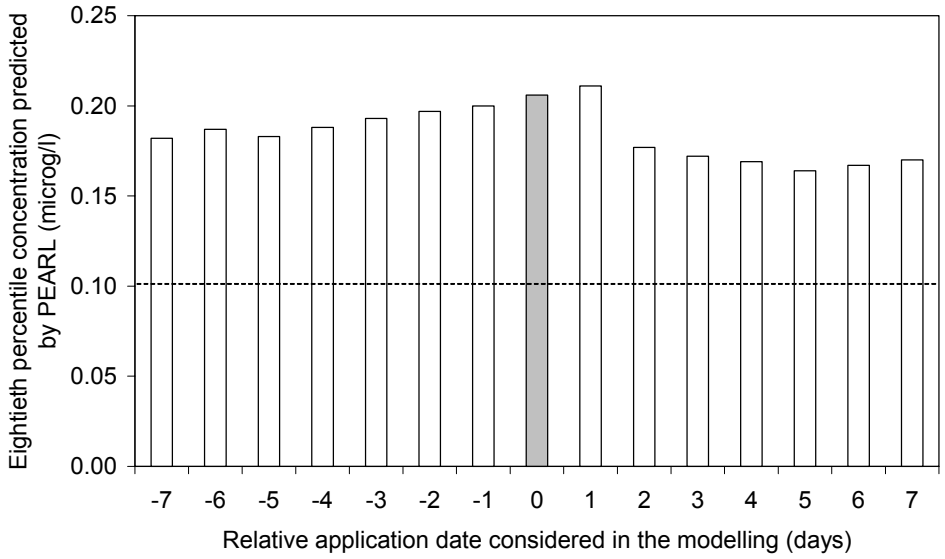


Figure 24. PEARL predictions for 80th percentile concentrations obtained when varying the application date by + / - 7 days (Okehampton scenario).

The default application date selected in the earlier modelling (7 days post emergence) is shown in grey. The dotted line indicates the 0.1 μg/l leaching level.

4 MONTE CARLO MODELLING

Monte Carlo methods are the most popular techniques to account for uncertainty in numerous scientific disciplines as i) they are easy to understand and transpose; ii) they are relatively easy to implement provided the right software tools are being used; iii) they have been used for many years and are believed to lead to more informed decision-making. A number of working groups have recently called for a greater use of the Monte Carlo framework in assessing the risk of pesticides impacting on the environment.

Monte Carlo simulations were carried out using the PEARL model for the Okehampton scenario. The present exercise was not intended to provide an in-depth analysis of the propagation of uncertainties through PEARL, but rather to exemplify the types of results that can be obtained by Monte Carlo modelling.

4.1 Modelling methods

The PEARL model was selected for Monte Carlo modelling as this model was used in investigations reported earlier in the report and its relatively short execution time allows the running of the model for numerous instances. The modelling was undertaken for the Okehampton scenario as the use of this scenario in the FOCUS modelling resulted in the largest concentrations predicted by PEARL. Parameters to be included in the Monte Carlo analysis were selected using a combination of information about their uncertainty and their influence on model predictions (model sensitivity) and were 1) the sorption distribution coefficient normalised to organic matter Kom Kfom, and 2) the field DT50 value. Sorption and degradation have been identified in earlier DEFRA work (PL0532) as the parameters to which predictions for pesticide loss are most sensitive in most instances. In contrast to the standard FOCUS approach which concentrates on the 80th-percentile concentration, the 20 years simulated were used in the assessment.

4.1.1 Examination of the measured data for Koc and DT50

A total of 20 Koc values and 61 field DT50 values were available to support the attribution of probability density functions for the two parameters. Field DT50 values were preferred over laboratory estimates as i) the number of laboratory values available was smaller; and, ii) field DT50 values are considered to provide a more realistic measure of field behaviour than laboratory values. Summary statistics about the two datasets are provided in Table 10 while distribution charts are presented in Figures 25 to 28. Only normal and lognormal distributions

were considered in the present instance as it was felt that limitations in the number of datapoints available did not allow the testing of alternative distributions in a robust way.

	Koc (ml/g)	Field DT50 values (days)
Range	39-173	13-105
Mean	99.5	42.2
Median	91.5	43
Skewness	0.49	0.62
Kurtosis	-0.14	-0.05
Standard deviation	34.6	22.7
Coefficient of variation (%)	34.8	50.3

Table 10. Statistics on the datasets available for Koc and field DT50 values (from Syngenta)

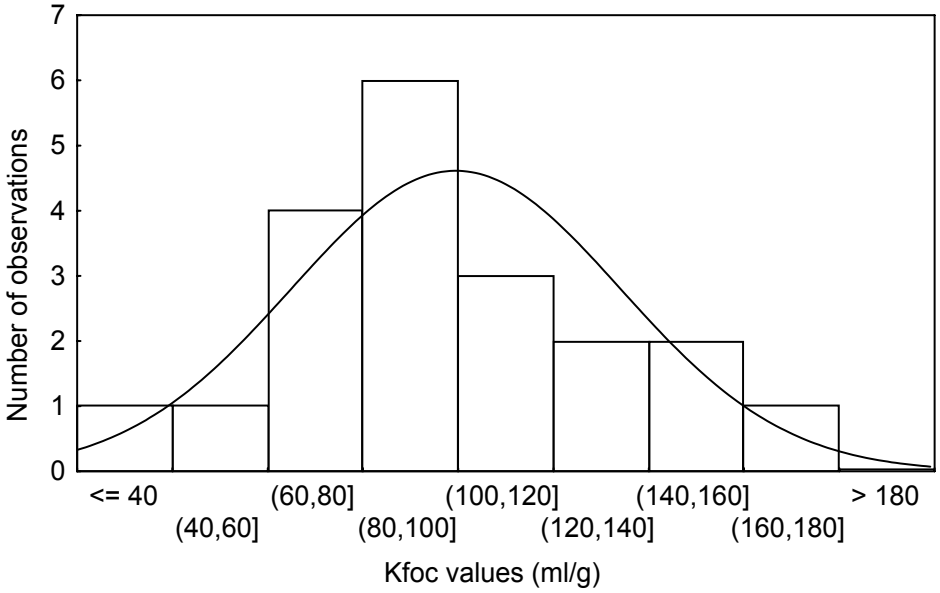


Figure 25. Probability density function of Koc values
The line indicates the fitting of a normal distribution to the data.

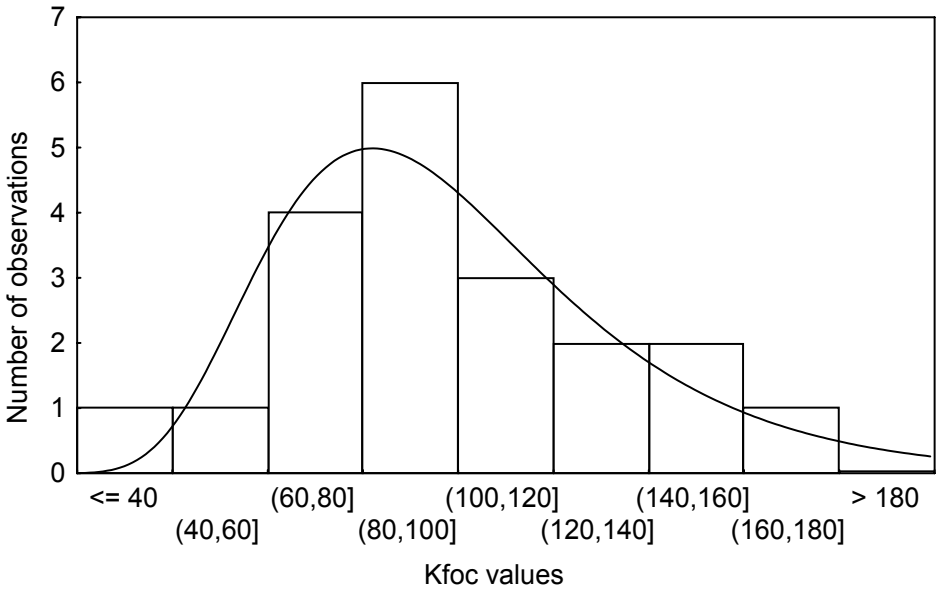


Figure 26. Probability density function of Koc values
The line indicates the fitting of a lognormal distribution to the data.

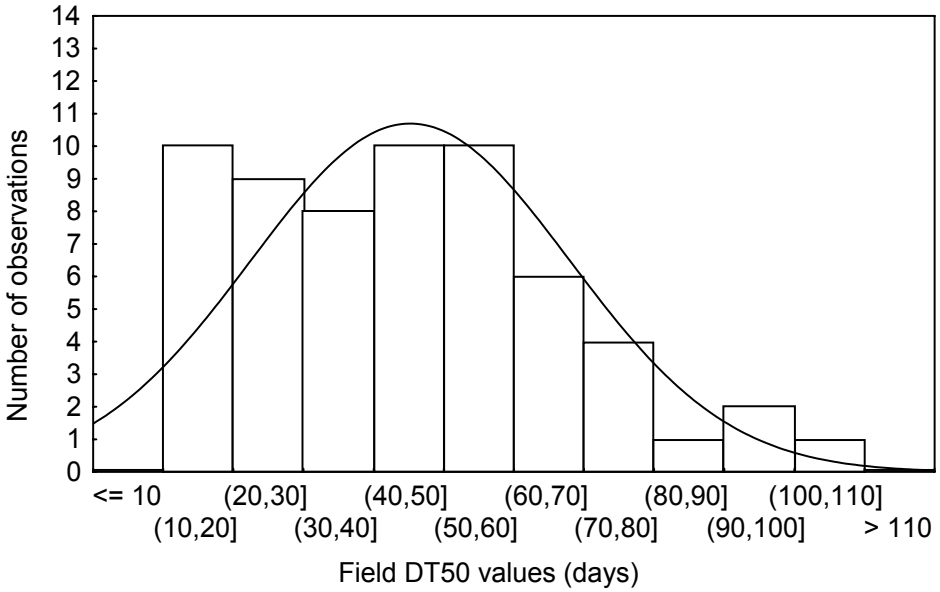


Figure 27. Probability density function of field DT50 values
The line indicates the fitting of a normal distribution to the data.

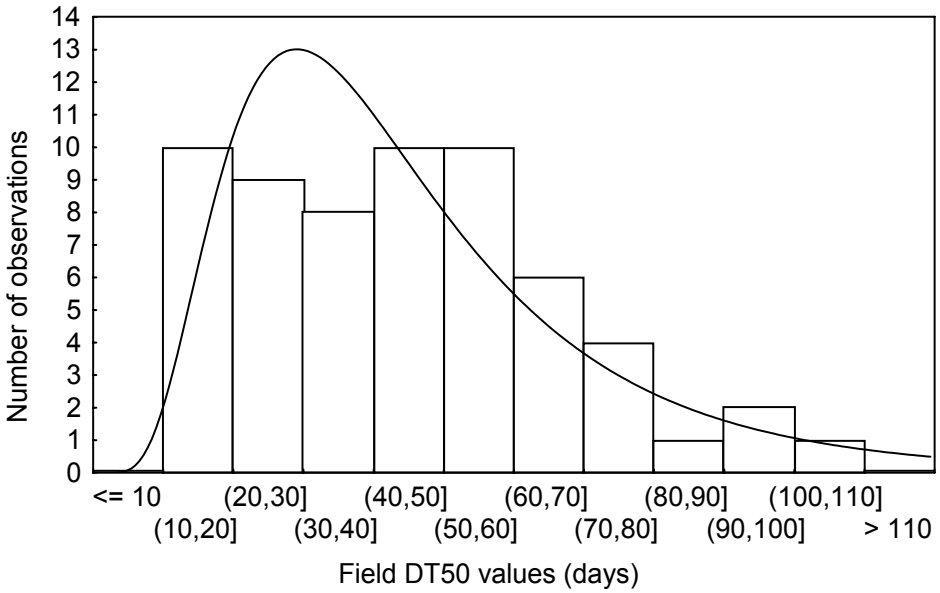


Figure 28. Probability density function of field DT50 values
The line indicates the fitting of a lognormal distribution to the data.

Making a judgement on the fit of the normal and lognormal distributions to the Koc and DT50 data on the basis of Figures 25 to 28 was difficult and complementary goodness-of-fit tests were carried out. It should be noted that the selection of distributions on the basis of statistical tests remains a subjective exercise (Saltelli *et al.*, 2000). Three goodness-of-fit tests were used: i) Chi-square (χ^2); ii) Kolmogorov-Smirnov; and, iii) Anderson-Darling. The *Chi-square* test is the oldest and most common of the goodness-of-fit tests. It gauges the general accuracy of the fit. In the implementation of the test in the most popular Monte Carlo

packages, the distribution is broken down into areas of equal probability and the test compares the data points within each area to the number of expected datapoints on the basis of a χ^2 distribution. The *Kolmogorov-Smirnov* test is computed as the largest vertical distance between the cumulative distribution function of the empirical points and that of the law tested. The *Anderson-Darling* test is similar to the Kolmogorov-Smirnov test except for that it puts more weight on the differences between the two distributions at their tails compared to the mid-ranges. The fit of the lognormal distribution to the data is tested by evaluating a normal fitting on the logtransformed data. Additionally, the Shapiro-Wilk W test was computed. This test is commonly used to assess normality.

Results of the different tests are provided in Table 11. The p value associated with each statistic is a reflection of the confidence that should be attributed to the test. Since p values were small for all tests carried out ($p<0.30$) and results of the different tests were conflicting to some extent, the derivation of any conclusion with regard to the distribution to be assigned to Koc and DT50 on the basis of Table 11 was not considered robust in the present instance.

	Number of datapoints	Shapiro-Wilk W test (and p value)	p value χ2 test	Kolmogorov-Smirnov statistics	Anderson-Darling statistics
Koc data – normal fit	20	0.968 (<0.02)	0.273	0.174	0.318
Log Koc data – normal fit [a]	20	0.976 (<0.03)	0.094	0.108	0.199
DT50 data – normal fit	61	0.952 (<0.72)	0.253	0.084	0.646
Log DT50 data – normal fit [a]	61	0.958 (<0.87)	0.140	0.087	0.732

+ The fitting of a normal law to log-transformed data is equivalent to the fitting of a lognormal law to the original data.

Table 11. Results of goodness-of-fit tests carried out on the Koc and DT50 data using Crystal Ball 2000

Closer fit to the normal law is indicated by larger Shapiro-Wilk W statistics, larger p values for the χ^2 test and smaller values for the Kolmogorov-Smirnov and Anderson-Darling statistics.

4.1.2 Attribution of probability density functions to Koc and DT50

Attempts were made to assign normal and log-normal distributions to the two parameters selected for inclusion in the Monte Carlo modelling on the basis of the examination of summary statistics (in particular, median, mean, skewness and kurtosis), distribution charts (Figures 25 to 28) and goodness-of-fit tests (Table 11). The three approaches did not allow a robust assignment of distributions to the two parameters. This could be mainly attributed to the rather limited datasets available to support distribution assignment (though it should be noted that those for atrazine will be considerably larger than those for most other compounds).

In consequence, the distribution fitting approach was discarded and an empirical approach based on a triangular distribution was used. This distribution is widely used in elicitation, the process of attributing and parameterising probability distributions on the basis of expert knowledge. Parameterisation of the triangular distribution made use of the minimum, median and maximum values of the datasets for Koc and DT50. The assumptions made were that:

- The most likely values for Koc and DT50 (modes of the triangular distributions) were the medians of the datasets of measured values;
- The probability of having a value outside the range defined by the minimum and maximum values of the datasets was inversely proportional to the number of datapoints in the dataset and is 100/N (expressed in %) where N is the number of values in the dataset. This parameterisation reflects the fact that the confidence of having a new measurement in the [minimum; maximum] range will increase with the number of values in the datasets.

The approach adopted can potentially be applied to the whole spectrum of size of datasets for pesticide sorption and degradation, from those with four or so values to those larger ones.

The detailed parameterisations of the triangular distribution functions for Koc and DT50 are presented in Table 12 and Figures 29 and 30.

	Koc (ml/g)	Field DT50 value (days)
Most likely value (mode)	91.45	43
2.5th percentile	39	-
97.5th percentile	173	-
0.82th percentile	-	13
99.18th percentile	-	105

Table 12. Parameterisation of the triangular distributions for Koc and DT50
The most likely values were taken as the median of the datasets available.
The lower percentile (expressed in %) was taken as 100/(2N) where N is the number of values in the datasets for Koc and DT50 (20 and 61 values, respectively)
The larger percentile was set to 100*(2N-1)/(2N).
The value for the lower percentile was set to the minimum value from the dataset while that for the larger percentile was set to the maximum value.

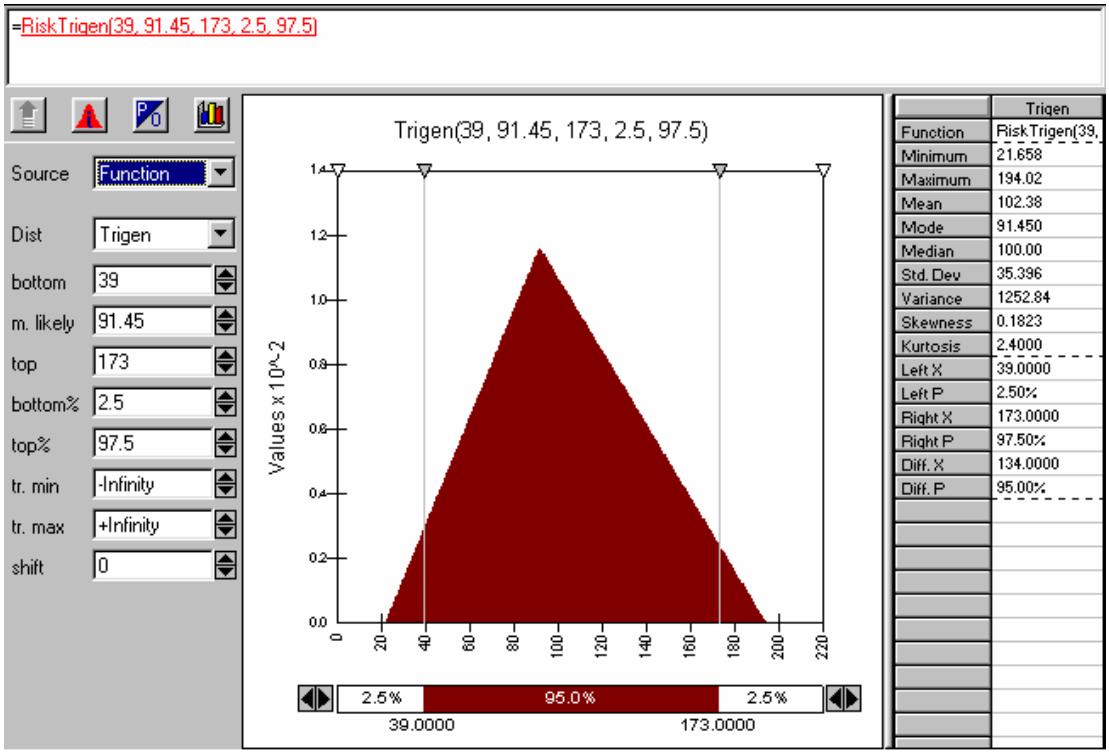


Figure 29. Screenshot from @RISK showing the parameterisation of the triangular distribution for Koc

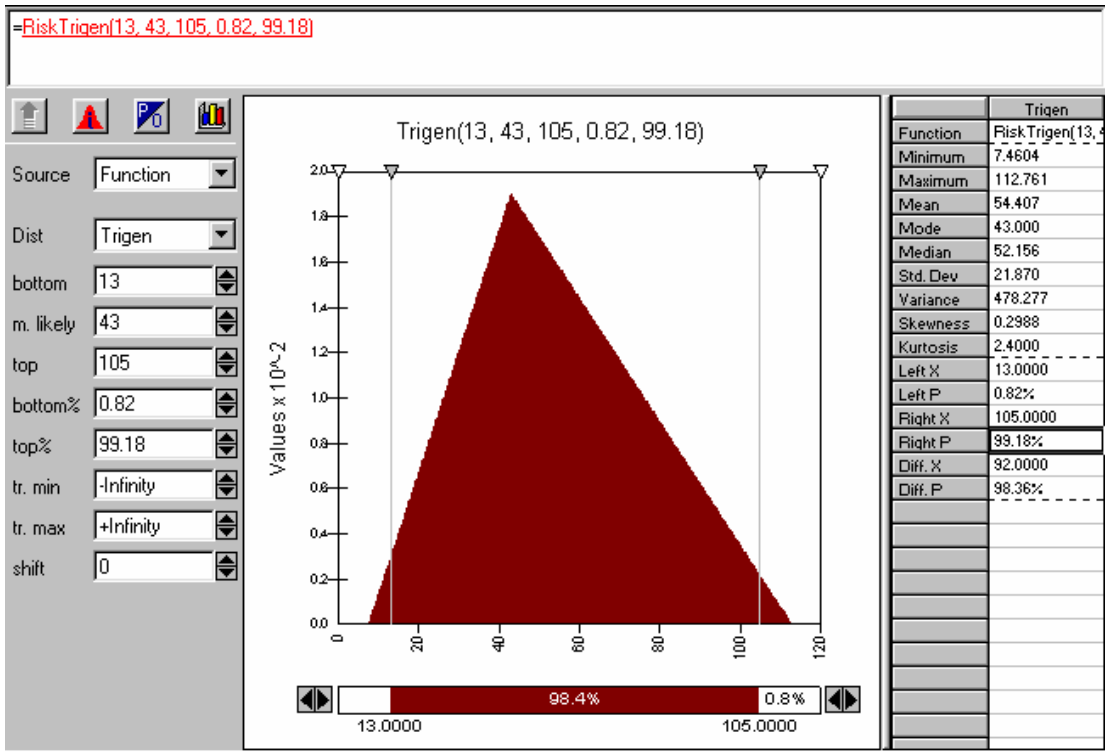


Figure 30. Screenshot from @RISK showing the parameterisation of the triangular distribution for DT50

The fitting of the normal, lognormal and triangular distributions are compared in Figure 31 (Koc) and Figure 32 (DT50).

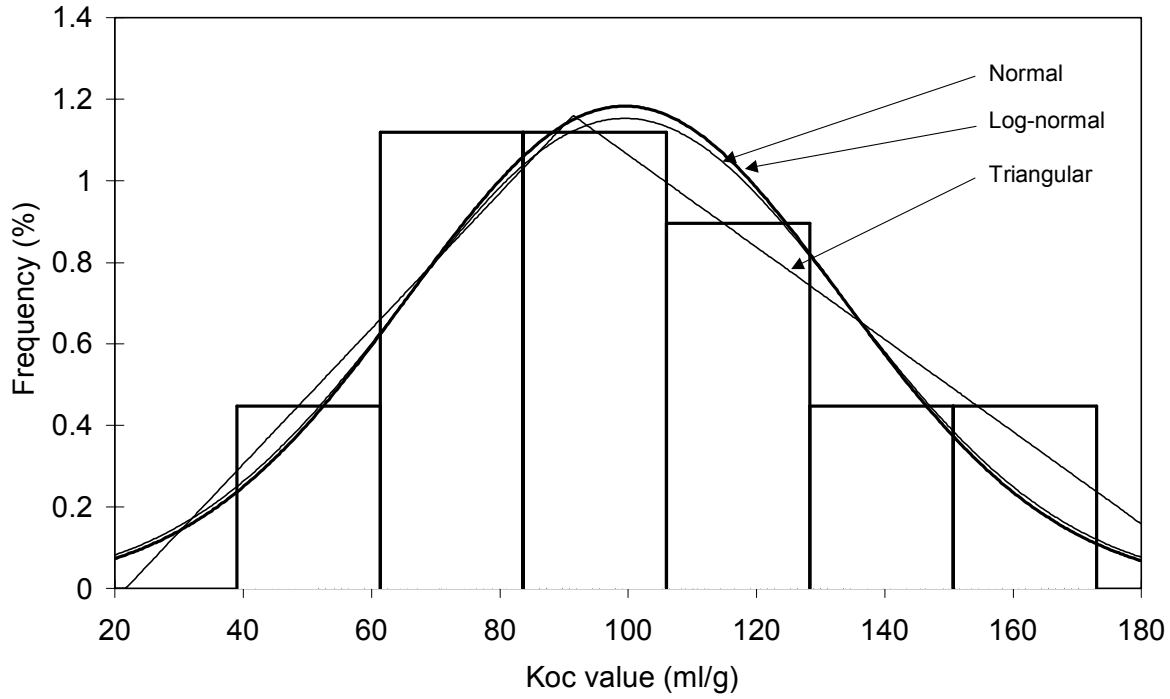


Figure 31. Fitting of the normal, lognormal and triangular distributions to Koc data for atrazine

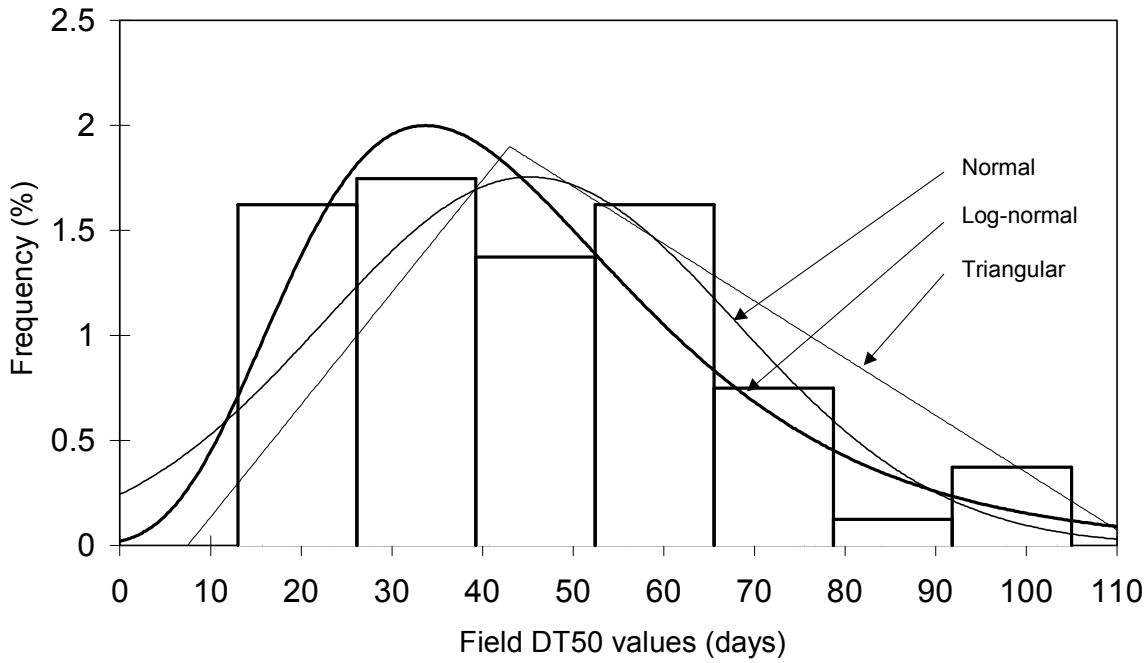


Figure 32. Fitting of the normal, lognormal and triangular distributions to field DT50 data for atrazine

4.1.3 Monte Carlo sampling

Monte Carlo sampling was implemented in the @RISK package using the Latin Hypercube method which is recognised as an efficient sampling technique to explore the parameter space

(Helton, 1993). The number of model runs undertaken was set to 500 on the basis of computational requirements. Results of the Monte Carlo modelling will be influenced by the seed number used to initiate random sequences (Chapter 3). The modelling was therefore repeated for five different sets of 500 runs resulting in a total of 2500 iterations of the model. The total number of years of running was 50000.

Koc values were sampled and these were subsequently divided by 1.72 to provide a distribution of Kom values to be fed into the PEARL parameter file. The modification of Kom and DT50 values in the PEARL input file, the running of the model and the extraction of the relevant concentration were automated using the SENSAN program provided in the PEST package (Doherty, 2000).

4.2 Results of the Monte Carlo modelling

Cumulative distribution functions for predicted concentrations of atrazine are shown in Figure 33. The different curves correspond to the five random samples generated with different seed numbers. Table 13 shows, for the five replicated sets of Monte Carlo simulations, the probabilities of exceeding 0.1μg/l and selected concentrations for a number of distribution percentiles. Percentiles were found to be significantly affected by the random sample used (CV 6-15%). The probability of exceeding 0.1 μg/l based on the simulation of concentrations at 1-m depth for 20 years for the Okehampton leaching scenario when accounting for the uncertainty in Koc and DT50 values was 53.5% on average.

Figure 33 also presents the concentrations obtained through FOCUS modelling for the 20 years of simulation. The range of concentrations predicted in Monte Carlo modelling was much larger than that obtained through FOCUS modelling because the uncertainty in model predictions arising from that in pesticide properties (which is ignored in the standard FOCUS modelling) is significantly larger than that associated with inter-annual variations in weather patterns. On Figure 33, the median of the Monte Carlo results (0.139 μg/l) is significantly larger than that predicted using the FOCUS run of 20 years (0.055 μg/l). This can be attributed to some extent to the parameterisation of the triangular distributions for Koc and DT50. The triangular distribution is typically parameterised using the minimum, maximum and most likely value (the tip of the triangular distribution). To reflect expert knowledge, the most likely value in the triangular distribution was taken here as the median of the datasets for sorption and degradation. The triangular distribution being skewed to the left, this most likely

value will be less than the actual median of the distribution for Koc and DT50, hence the difference in the median of predicted concentrations.

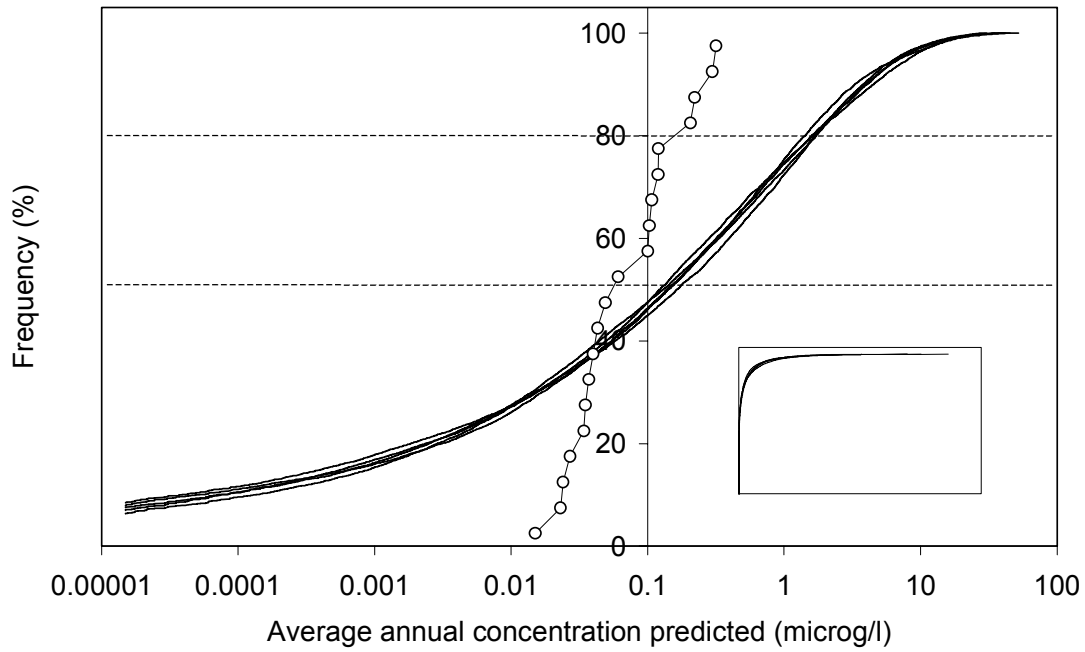


Figure 33. Cumulative distribution functions for atrazine concentrations predicted by PEARL
The five different line curves correspond to the use of replicated random samples in Monte Carlo modelling. The insert to the right shows the Monte Carlo cumulative distribution functions on a non log-transformed axis. The line with round open symbols presents concentrations for the 20 years simulated using the default FOCUS framework for the Okehampton scenario (see section 2). The dotted lines indicate the 50th- and 80th- percentiles.

	Rep. #1	Rep. #2	Rep. #3	Rep. #4	Rep. #5	Mean	CV (%)
P(C≤0.1μg/l), in %	47.4	45.0	46.4	47.5	46.1	46.49	2.3
P(C>0.1μg/l), in %	52.6	55.1	53.6	52.5	53.9	53.51	2.0
50th percentile concentration (μg/l)	0.12	0.16	0.14	0.13	0.14	0.14	11.3
75th percentile concentration (μg/l)	1.02	1.21	1.05	1.14	0.98	1.08	8.6
80th percentile concentration (μg/l)	1.62	1.72	1.59	1.68	1.42	1.61	7.2
90th percentile concentration (μg/l)	4.52	3.78	3.96	4.06	3.30	3.92	11.3
95th percentile concentration (μg/l)	8.17	6.53	7.03	6.76	6.94	7.08	9.0
99th percentile concentration (μg/l)	18.80	16.68	17.27	15.80	17.41	17.19	6.4
99.9th percentile concentration (μg/l)	38.25	30.36	31.38	34.31	25.53	31.96	14.8

Table 13. Results of the Monte Carlo modelling (probabilities of predicting concentrations below/ above 0.1 μg/l and concentrations for selected percentiles)
Results were derived from the curves presented in Figure 33

5 SCENARIO-BASED MODELLING

In scenario-based modelling, a model is used to predict the environmental fate of a pesticide in a number of soils under different sets of climatic conditions. Soils and climates used in the modelling are selected as representative of those conditions under which the compound will be used. In the present instance, four soils and four weather scenarios were used to represent the maize growing area overlying aquifer resources in England and Wales where atrazine is likely to be used.

5.1 Modelling methods

5.1.1 Model selection

Preferential flow has been recognised as an important process for the transfer of pesticides through soils in the UK (Brown *et al.*, 2000). The MACRO model was selected for scenario-based modelling because it has the potential to simulate leaching of agrochemicals through structured soils where macropore flow may be an important process. In coarser soils where macropore flow is less likely to occur, the model reverts to a classical solution of Richards' equation and the convection-dispersion equation to describe soil water flow and solute transport, respectively. The latest stand-alone version of the model (version 4.3, 30 August 2001) was used in the present study (Jarvis, 2001).

5.1.2 Derivation of modelling scenarios

The target area (here, maize-growing land overlying aquifers in England and Wales; *ca.* 1.1 x 105 ha; Figure 34) was divided into environmental scenarios comprising discrete classes of soil type and climate. The analysis was undertaken using the SEISMIC database (Hallett *et al.*, 1995). First, areas of maize cultivation were overlaid onto aquifer boundaries to generate an estimate that *ca.* 35% of the area cultivated with maize in England and Wales directly overlies major or minor aquifers. The soil series making up the area with potential for direct leaching to groundwater were then divided into five broad classes (Table 14) based upon vulnerability to leaching of pesticides. The division was made subjectively based upon soil factors including organic carbon content, texture, likelihood of macropore flow and average depth to groundwater. Representative soil series lying at the vulnerable end of each class were attributed to four of the five soil classes (Table 14). No representative series was selected to cover Class 5 as the aquifer is protected from vertical transport of water and pesticide by the presence of the slowly permeable/impermeable layer or the presence of a

peaty layer in the profile which will display very strong sorption. Thus four soils were selected for modelling to represent the range of maize soils overlying aquifers that have potential for leaching of atrazine to depth. Together, they represent 27% of the maize growing area of England and Wales. There is no risk of direct leaching of pesticide to groundwater for the remaining 73% of the maize area. For each representative series, profile information was extracted from SEISMIC (Table 15) and used to parameterise the model (see Dubus & Brown (2002) for a description of the approach to selecting soil hydraulic parameters).

Class	Description	Representative soil series	Area as % of total maize
1	Heavy soils with by-pass flow to shallow groundwater situated at *ca.* 1m	Enborne	4.0
2	Loamy soils over gravel or sandy soils with shallow groundwater (*ca.* 1 to 2 m depth)	Hall	2.9
3	Sandy and shallow soil with deep groundwater (>5 m depth)	Cuckney	3.3
4	Deep loamy and clayey soils with deep groundwater (>5 m depth)	Ludford	16.8
5	Slowly permeable and impermeable soils or soils with peaty layers	-	7.8
		Total	34.8

Table 14. Classification and relative extent of maize soils overlying aquifers in England and Wales

Information is available on depth to groundwater for each class in Table 14. Hence, leaching was simulated to groundwater depth to a maximum of 2 m below the soil surface (there is no information on substrate properties to support modelling of deeper layers). Simulations considered leaching to 1 m for the Enborne series and to 2 m for the Hall, Cuckney and Ludford series. Hall soils are very stony in subsoil layers. For modelling purposes, it was assumed that the stones play no part in leaching processes and the profile was shortened according to the relative stone content (actual profile depth simulated 1.25 m).

	Depth interval (cm)	Organic carbon (%)	Sand (%)	Silt (%)	Clay (%)	Bulk density (g/cm^3)	pH (-)
Enborne							
Horizon1	0-25	3.9	19	42	39	1.12	7.1
Horizon 2	25-50	1.6	14	46	40	1.26	7.2
Horizon 3	50-70	0.9	20	45	35	1.34	7.3
Horizon 4	70-100	0.8	23	42	35	1.38	7.3
Hall							
Horizon1	0-30	2.7	67	21	12	1.25	6.8
Horizon 2	30-50	0.6	71	20	9	1.39	6.4
Horizon 3	50-60	0.4	71	20	9	1.40	6.5
Horizon 4	60-70	0.2	76	18	6	1.45	6.5
Horizon 5 [a]	70-200	0.1	87	9	4	1.47	6.9
Cuckney							
Horizon1	0-30	1.7	72	18	10	1.40	6.9
Horizon 2	30-50	0.6	79	14	7	1.40	6.9
Horizon 3	50-70	0.4	76	14	10	1.47	6.7
Horizon 4	70-100	0.2	82	11	7	1.45	6.6
Horizon 5 [a]	100-200	0.1	82	11	7	1.45	6.6
Ludford							
Horizon1	0-25	2.2	38	40	22	1.30	6.8
Horizon 2	25-50	0.9	36	42	22	1.37	6.9
Horizon 3	50-75	0.5	34	36	28	1.40	7.1
Horizon 4	75-100	0.3	37	36	27	1.50	7.2
Horizon 5 [a]	100-200	0.1	37	36	27	1.50	7.2

[a] Properties for the 1-2 m horizon were taken as similar to the horizon which goes to 1 m except for the organic carbon content which was set to 0.1 %

Table 15. Selected soil properties for the four representative soil series
Some profile depths have been adjusted to allow for
the description of leaching to groundwater situated at >1 m depth.

Areas of maize cultivation in England and Wales were divided into four climatic classes (Figure 35) designated 'dry' (<625 mm precipitation per annum), 'medium' (625-750 mm p.a.) 'wet' (750-850 mm p.a.) and ‘very wet’ (>850 mm p.a.). Four weather datasets were then selected from the SEISMIC database as representative of the four climatic classes. The datasets comprised 30 years of daily weather (precipitation, maximum and minimum air temperature, solar radiation) simulated using the WGEN model. Average annual rainfall for the four datasets was 589 (Cambridge NIAB), 713 (Mylnefield), 815 (Keele) and 1112 mm (Rosewarne). Potential evapotranspiration was calculated using Linacre’s equation and adjusted using a fixed site-specific factor such that model simulations matched long-term averages for actual evapotranspiration. Long-term (30 years) simulations were started with two years of dummy weather data to allow the hydrological component of the model to reach equilibrium prior to the first application of pesticide.

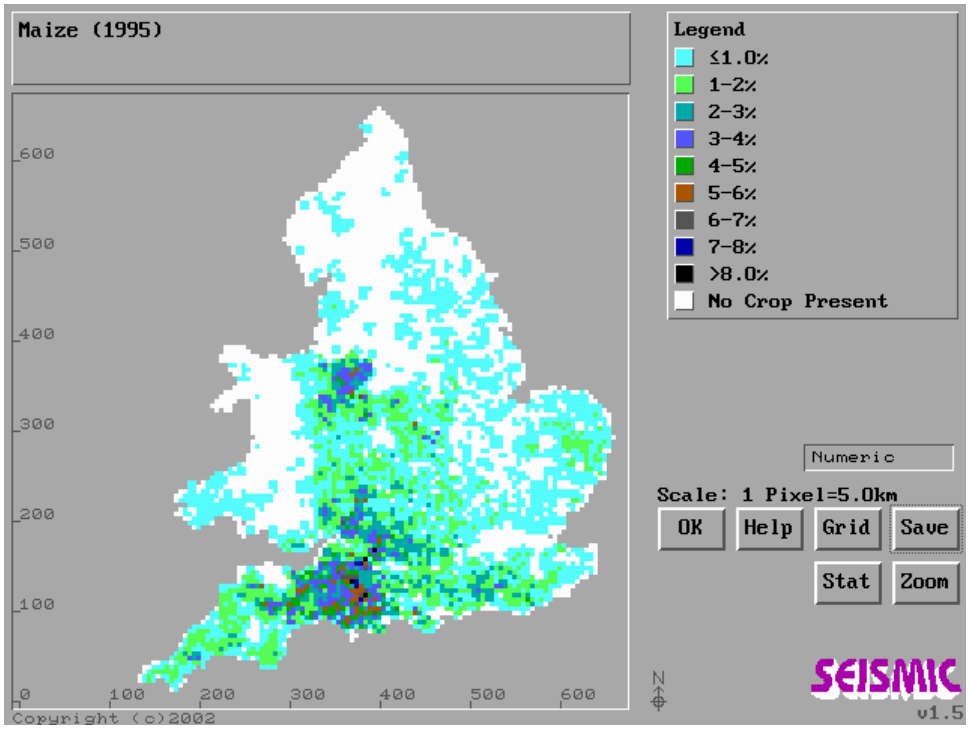


Figure 34. Distribution of maize cultivation in England and Wales (SEISMIC data, 1995 cropping)

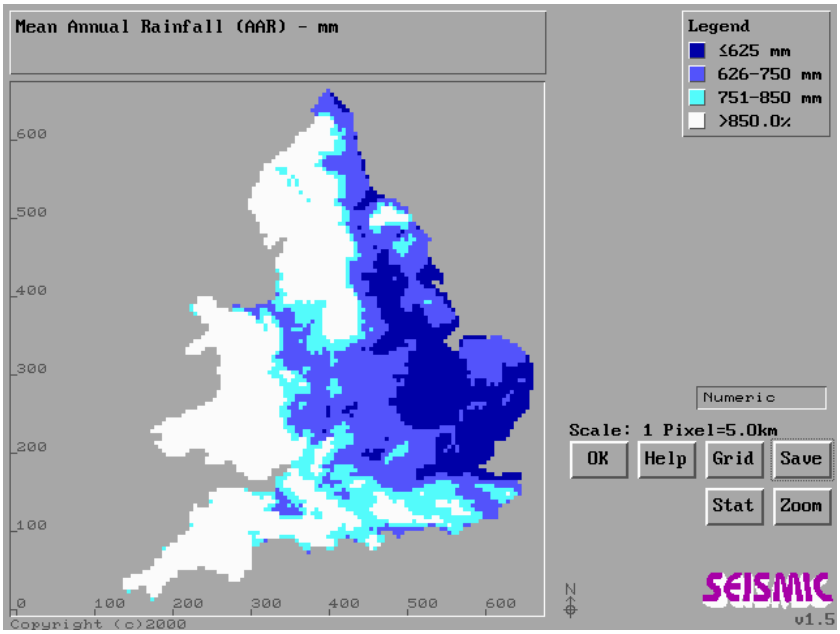


Figure 35. Distribution in England and Wales of the four climatic scenarios selected for modelling (SEISMIC data)

5.1.3 Cropping data and pesticide properties

Maize cultivation was modelled throughout the 30-year simulation period. Atrazine was assumed to be applied at the rate of 1 kg a.s./ha to a maize crop in the spring of each of the 30 successive seasons. The target application date was 1 June. Where rainfall exceeded 2 mm on the selected date or 7 mm on either of the two subsequent days, sowing and application were delayed to the first subsequent dry day using the CHAPAR (change parameter) facility within MACRO. No interception by the crop was considered (*i.e.* the entire spraying solution reached the soil surface). Crop parameters representative of agricultural practices in England and Wales were used for soils from the four series.

Degradation rates in the four soils were specified in the model using field data and corrections for the variation of soil moisture and temperature conditions were therefore turned off. In an effort to investigate the uncertainty in the results originating from that in the selection of degradation and sorption values, a range of values were considered in the modelling. Apart from the median case (median field DT50 and Koc values derived from the sorption and degradation datasets), a 'better case' and a 'worse case' were considered (Table 16). The combinations of parameter values are identical to those which were used in the FOCUS modelling (see Section 3.1.1). Degradation rates were corrected down the profile according to the organic carbon content and sorption values in each layer using the equation implemented in MACRO_DB (Jarvis *et al.*, 1997). The Freundlich exponent was taken as the median of the 20 values available (nf 0.895).

	Koc	Field DT50
Median case	91.45 ml/g (50^{th} percentile)	43 days (50^{th} percentile)
Better case	116.5 ml/g (75^{th} percentile)	29 days (25^{th} percentile)
Worse case	79 ml/g (25^{th} percentile)	60 days (75^{th} percentile)

Table 16. Combination of Koc and DT50 values used in the MACRO modelling
The percentiles considered are given in parentheses

5.1.4 Modelling strategy

MACRO 4.3 was used to simulate the potential for leaching of atrazine to depth for each of the four representative soils and the four weather scenarios for each of the three combinations of sorption and degradation parameters. This gave a total of 48 long-term (30-year) simulations. Primary outputs from MACRO were percolation volumes at the bottom of the soil profile (*i.e.* 1 m for the Enborne series, 2 m for the remaining three soil series) and concentrations of atrazine in percolating water.

Cumulative distribution charts were constructed for annual average concentrations of atrazine in soil water leaching from the bottom of the soil profile. Annual average concentrations were calculated for each of 30 years simulated per scenario, resulting in a total of 1440 annual averages. The series of 30 values were then weighted according to the prevalence of the respective scenario across England and Wales. Slowly permeable, impermeable and peaty soils which represent 7.8% of the maize-growing area in England and Wales were attributed concentrations of zero and were included in the analysis in a similar way.

5.2 Modelling results

5.2.1 Results for individual scenarios

Summary results for the Enborne series are presented in Table 17. Patterns of annual average concentrations (AAC) of atrazine in percolation water over the 30-year simulation period for the four representative soil series are shown in Figures 36 to 39. Only the results for the combinations of median pesticide properties are presented.

Annual percolation was predicted to range from 77 mm (Enborne; dry scenario) to 954 mm (Hall; very wet scenario), whilst average annual concentrations of atrazine in leaching were predicted to vary between <0.001 µg/l and 2.2 µg/l, between <0.001 µg/l and 5.7 µg/l and between <0.001 µg/l and 9.8 µg/l for the better, median and worse case, respectively. The largest concentrations were all reported for the combination of the Enborne series and the wet climate scenario.

The relative vulnerabilities of the four soil classes (based on 30-year average concentrations for atrazine) for the three combinations of pesticide properties were:

Enborne > Hall > Ludford > Cuckney

Thirty-year average annual and maximum annual concentrations for each of the 48 scenarios are presented in Table 18.

	Dry scenario		Medium scenario		Wet scenario		Very wet scenario	
Year	Percolation (mm)	AAC[a] (µg/l)	Percolation (mm)	AAC (µg/l)	Percolation (mm)	AAC (µg/l)	Percolation (mm)	AAC (µg/l)
1959	101	<0.001	225	0.007	348	0.647	684	1.053
1960	212	0.459	219	0.134	363	0.836	733	2.684
1961	87	0.307	255	0.378	199	0.390	768	2.419
1962	104	0.036	301	0.664	406	1.723	601	2.921
1963	162	0.764	191	0.233	254	0.644	761	2.270
1964	191	0.370	346	0.465	416	2.041	650	1.662
1965	77	0.127	205	0.551	157	0.731	641	2.147
1966	158	0.058	99	0.081	149	0.546	566	1.656
1967	192	0.821	304	0.368	442	1.086	761	2.458
1968	148	0.260	300	0.543	320	0.700	501	1.227
1969	111	0.198	246	0.186	254	0.569	570	2.300
1970	190	1.068	287	0.262	331	1.231	888	3.739
1971	138	0.299	369	0.690	441	5.686	884	2.263
1972	105	0.059	240	0.614	418	1.710	864	2.782
1973	184	0.268	363	0.565	364	1.041	693	2.567
1974	129	0.235	326	0.700	285	0.421	616	1.361
1975	213	1.073	278	0.455	294	0.556	674	1.929
1976	147	0.300	218	0.282	310	0.794	614	1.703
1977	124	0.067	188	0.231	281	0.364	785	3.848
1978	155	0.132	242	0.394	420	3.769	506	3.031
1979	287	0.715	258	0.650	380	1.836	588	1.732
1980	280	1.299	361	0.772	339	1.003	760	3.321
1981	205	0.799	248	0.282	214	0.389	555	1.439
1982	110	0.264	263	0.266	178	0.273	743	3.262
1983	120	0.024	171	0.102	311	0.674	435	1.286
1984	186	0.111	204	0.318	193	0.186	607	1.135
1985	98	0.443	191	0.374	281	0.875	746	3.292
1986	138	0.585	139	0.061	354	1.530	608	1.025
1987	126	0.051	224	0.240	304	1.114	613	1.226
1988	124	0.191	156	0.046	304	1.151	602	0.977
Mean[b]	153	0.379	247	0.364	310	1.151	667	2.157
Min.	77	<0.001	99	0.007	149	0.186	435	0.977
Max.	287	1.299	369	0.772	442	5.686	888	3.848

[a] annual average concentration of atrazine in percolation

[b] the long-term mean was calculated as the sum of daily loads divided by the sum of daily percolation

Table 17. Percolation volumes and concentrations of atrazine in percolation water simulated for the Enborne series with 30-year runs of "dry", "medium", "wet" and "very wet" weather (median case scenario)

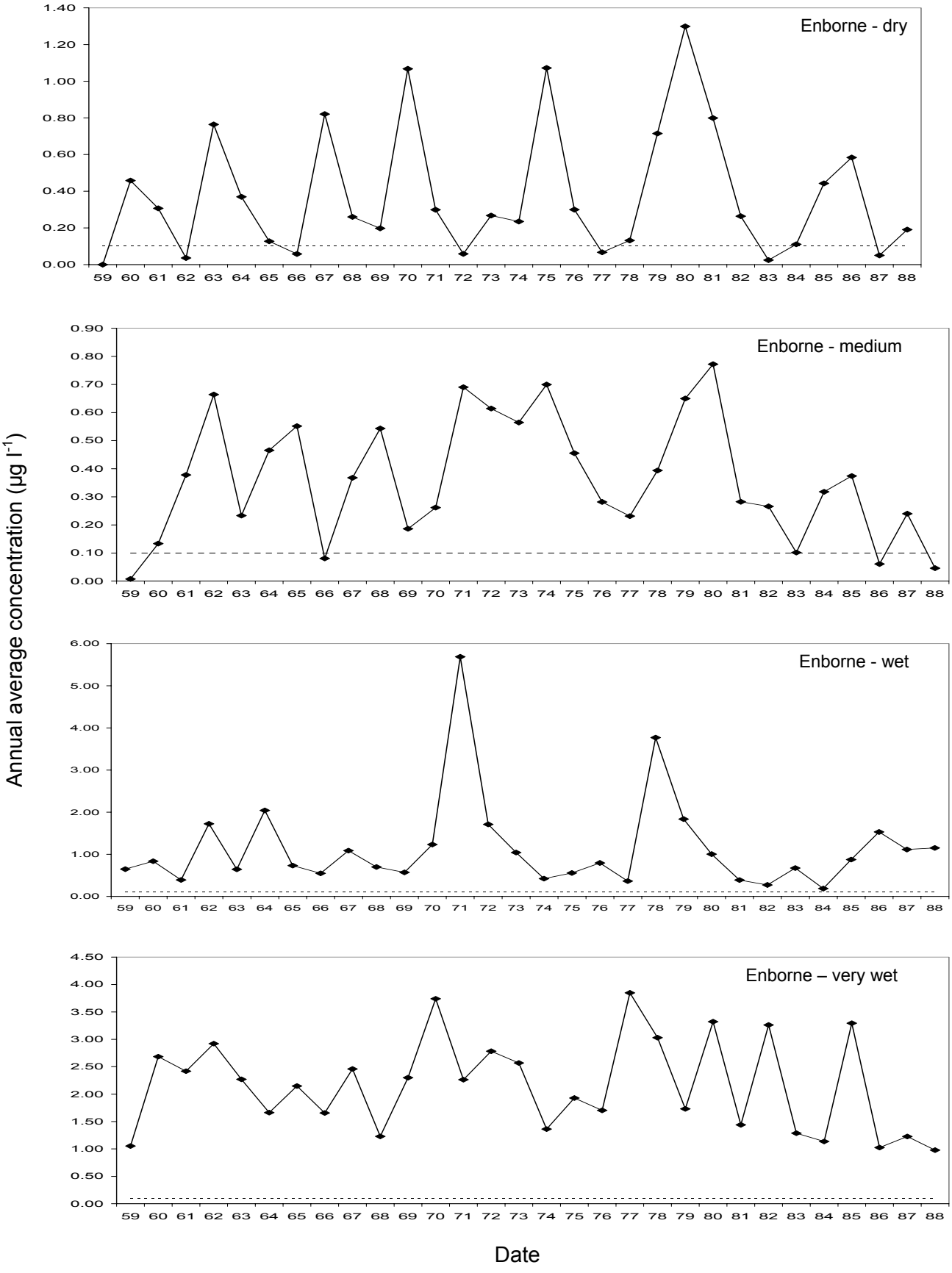


Figure 36. Annual average concentrations of atrazine in percolation water for the Enborne series for the four weather scenarios (median case scenario)

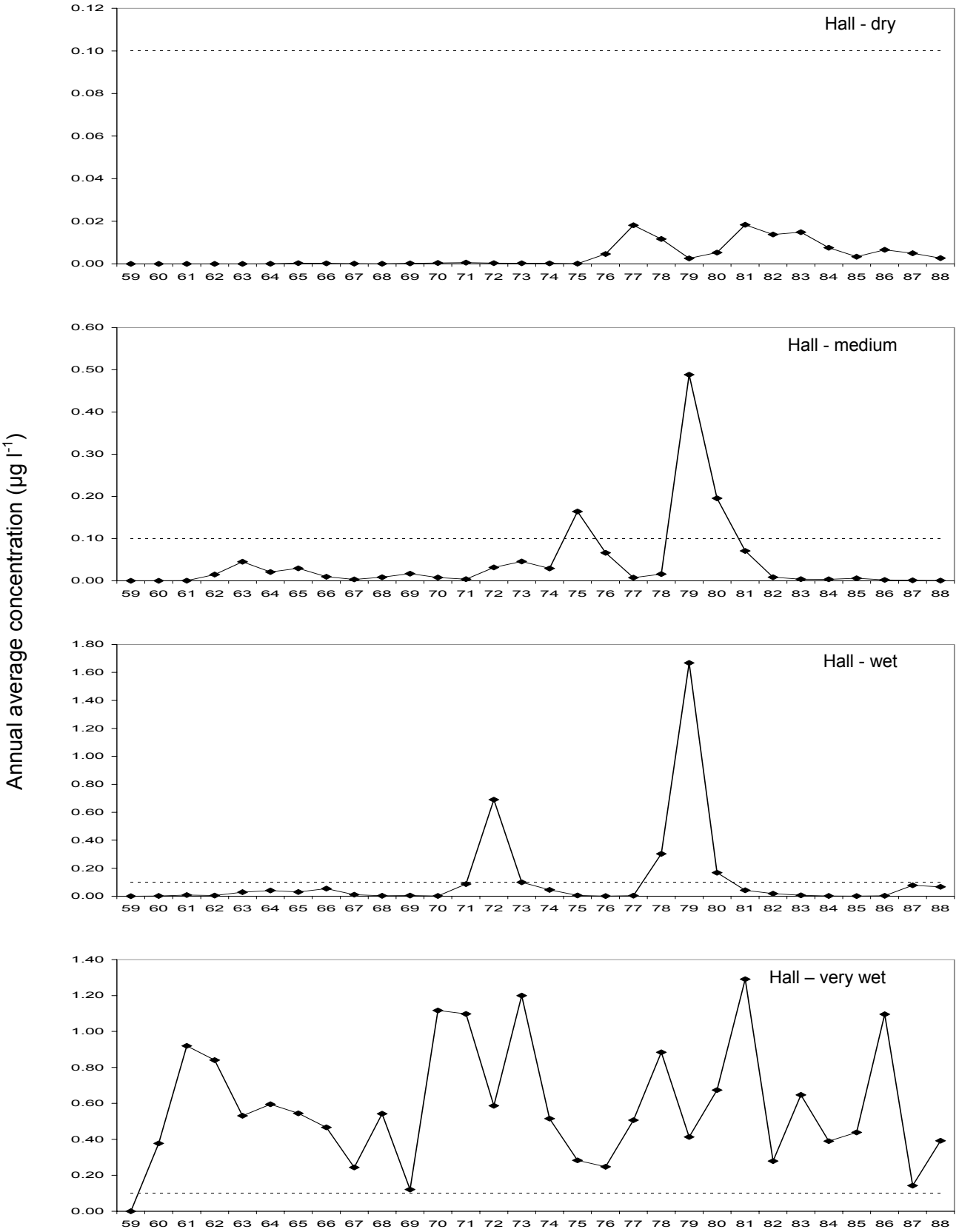


Figure 37. Annual average concentrations of atrazine in percolation water for the Hall series for the four weather scenarios (median case scenario)

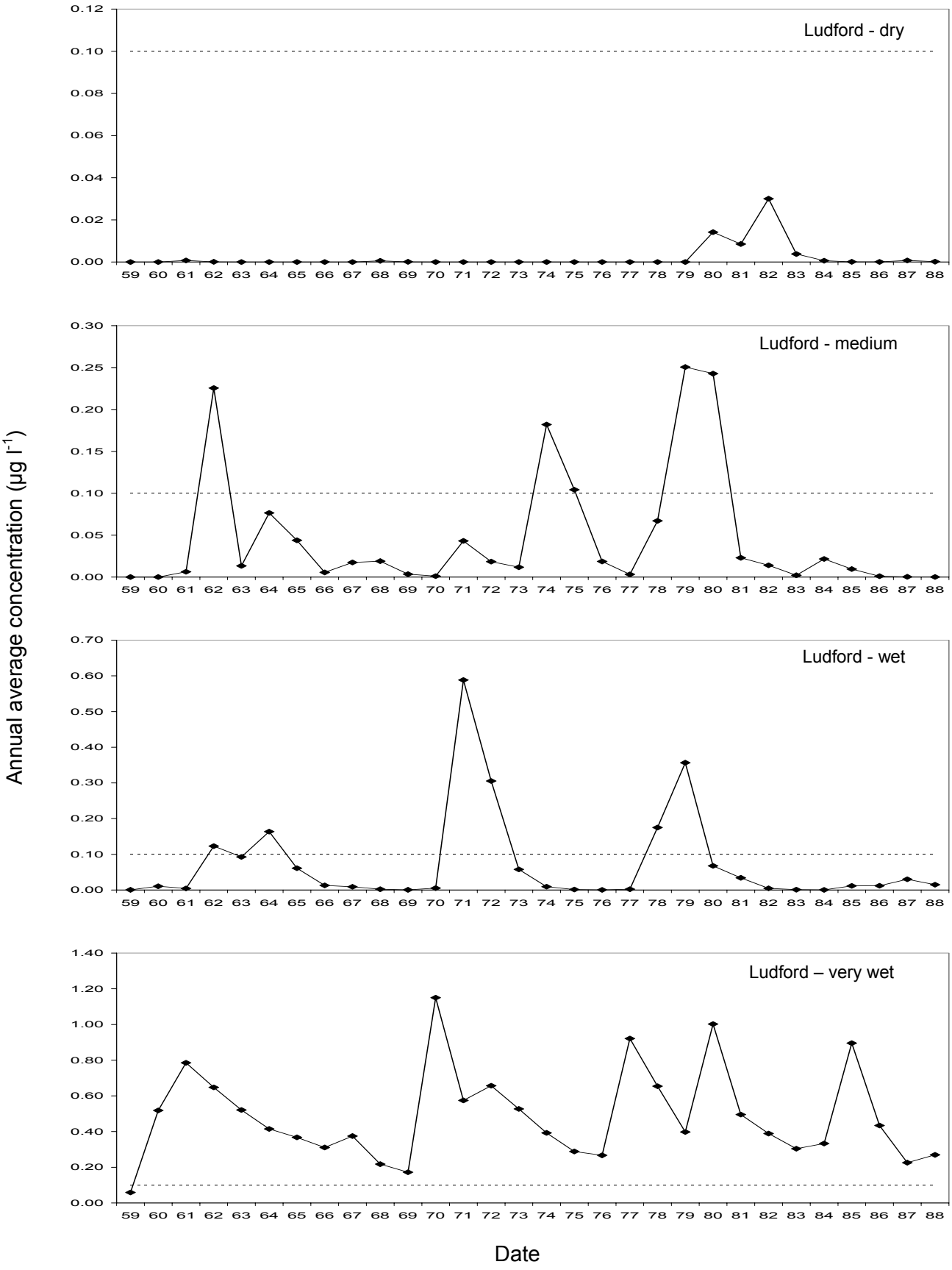


Figure 38. Annual average concentrations of atrazine in percolation water for the Ludford series for the four weather scenarios (median case scenario)

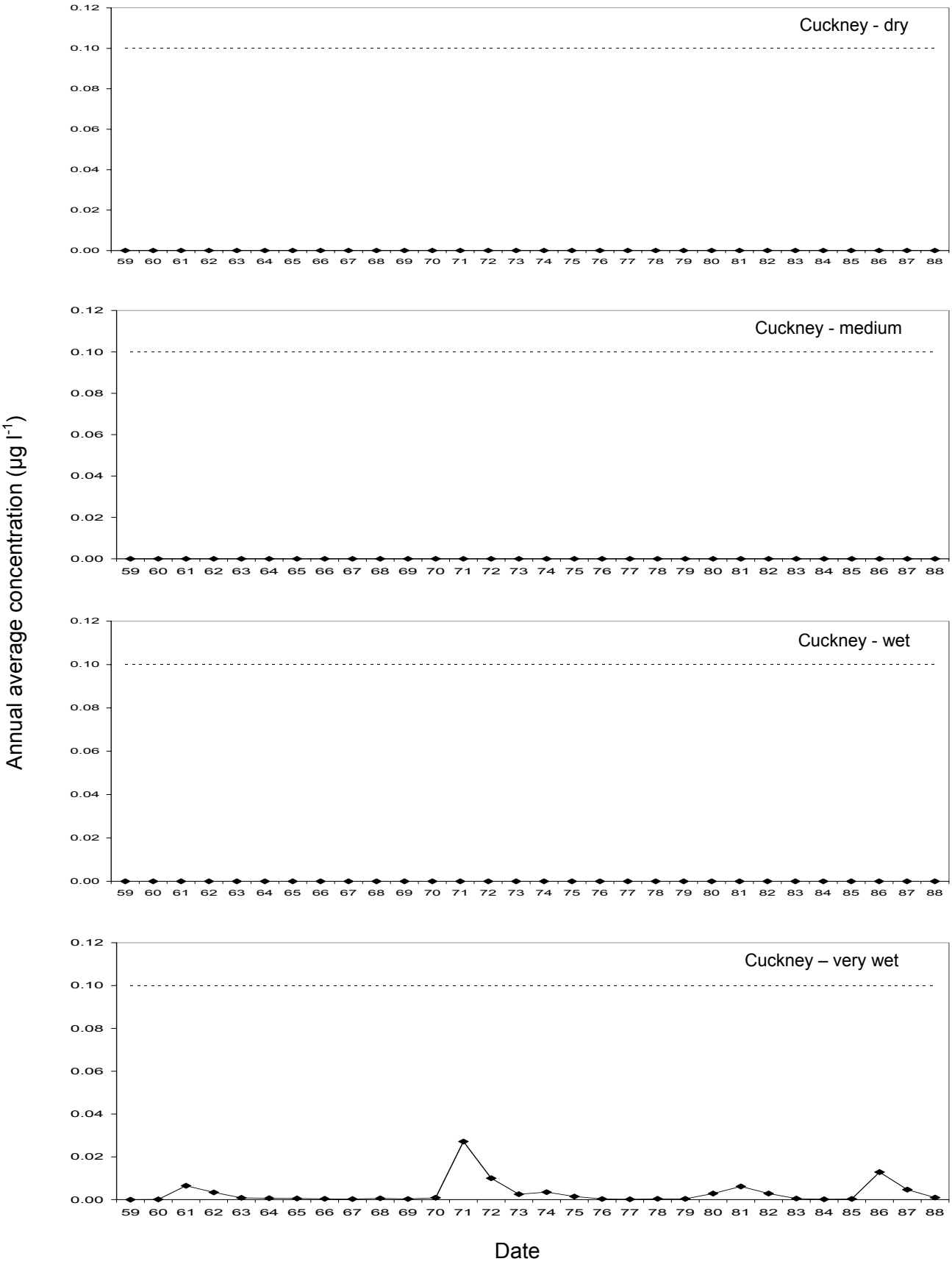


Figure 39. Annual average concentrations of atrazine in percolation water for the Cuckney series for the four weather scenarios (median case scenario)

Soil series	Climatic scenarios			
	Dry	Medium	Wet	Very wet
Better case				
Enborne	0.06 (0.28)	0.05 (0.15)	0.27 (2.16)	0.55 (1.26)
Hall	<0.01 (<0.01)	<0.01 (0.04)	0.01 (0.30)	0.06 (0.24)
Ludford	<0.01 (<0.01)	<0.01 (0.03)	0.01 (0.14)	0.08 (0.32)
Cuckney	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)
Median case				
Enborne	0.38 (1.30)	0.36 (0.77)	1.15 (5.69)	2.16 (3.85)
Hall	<0.01 (0.02)	0.04 (0.49)	0.12 (1.67)	0.58 (1.29)
Ludford	<0.01 (0.03)	0.05 (0.25)	0.07 (0.59)	0.49 (1.15)
Cuckney	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (0.03)
Worse case				
Enborne	1.09 (3.46)	1.17 (2.31)	2.76 (9.81)	4.77 (7.56)
Hall	0.07 (0.33)	0.25 (1.61)	0.44 (4.01)	1.91 (3.54)
Ludford	0.01 (0.16)	0.19 (0.87)	0.22 (1.10)	1.36 (2.43)
Cuckney	<0.01 (<0.01)	<0.01 (<0.01)	<0.01 (0.02)	0.09 (0.40)

Table 18. Thirty-year annual average concentrations of atrazine in leaching for each of the 48 scenarios considered in the modelling.

All expressions are expressed in μg/l.

The 30-year maximum annual concentrations are given in parentheses

5.2.2 Cumulative distribution charts

Cumulative distribution charts were constructed for atrazine for average annual concentrations. The 30-year average annual concentrations (an example is reported in Table 17 for the Enborne series) were extracted, resulting in 480 individual values (4 soils, 4 climates, 30 years of simulation) for each of the three combinations of pesticide properties. The series of 30 values were then weighted by their abundance across England and Wales (Table 19). Soils not overlaying aquifers and soils overlaying aquifers, but which show little vulnerability to pesticide leaching were attributed concentrations of zero and were included in the analysis in a similar way.

Soil type	Extent of soil within each climatic scenario (%)				Total extent (%)
	<625 mm AAR	625-750 mm AAR	751-850 mm AAR	>850 mm AAR	
Not overlaying aquifers	-	-	-	-	65.2
Not vulnerable [a]	0.7	2.9	3.4	0.8	7.8
Enborne	0.7	1.3	2.0	0.0	4.0
Hall	0.2	0.8	1.0	0.9	2.9
Cuckney	0.2	1.3	1.2	0.6	3.3
Ludford	0.8	4.1	4.1	7.8	16.8
Total	2.6	10.3	11.7	10.2	100.0

[a] Impermeable or peaty soils

Table 19. Proportion of the total maize growing land in England and Wales accounted for by each scenario (SEISMIC data, 1995 cropping)

Figures 40 to 42 presents cumulative distribution charts for the better, median and worse case as defined by the combination of pesticide properties, respectively. The charts can be used to derive a given percentile concentration or the probability of exceeding a particular concentration. Table 20 presents the probability of leaching concentrations exceeding 0.1 μg/l and 95th percentile concentrations. A common approach in pesticide fate modelling is to select pesticide properties as medians of datasets. In the present instance, the use of median Koc and DT50 values resulted in a probability of 13.8% for atrazine to exceed a concentration of 0.1 μg/l in leaching at 1-m depth from soils cultivated in maize in England and Wales. This probability estimation was found to be significantly influenced by the values attributed to Koc and DT50 (Table 16).

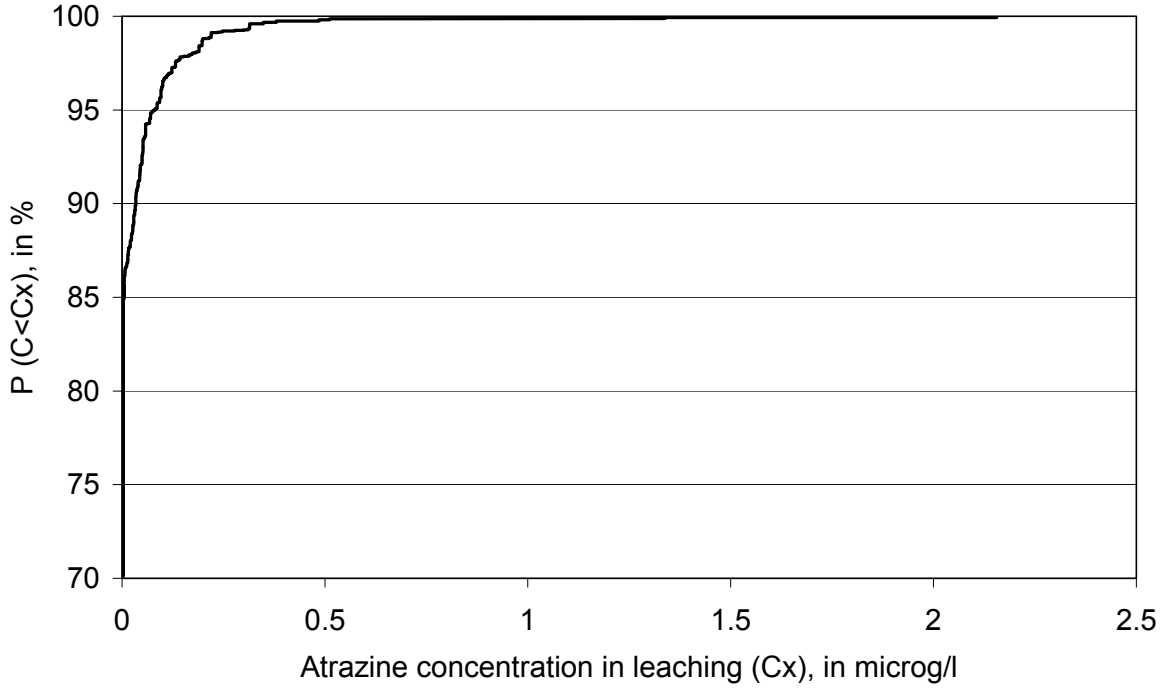


Figure 40. Cumulative distribution chart for concentrations of atrazine in leaching (better case scenario).

The y axis displays the probability that a concentration C is below the reference concentration Cx for the modelling scenarios considered.

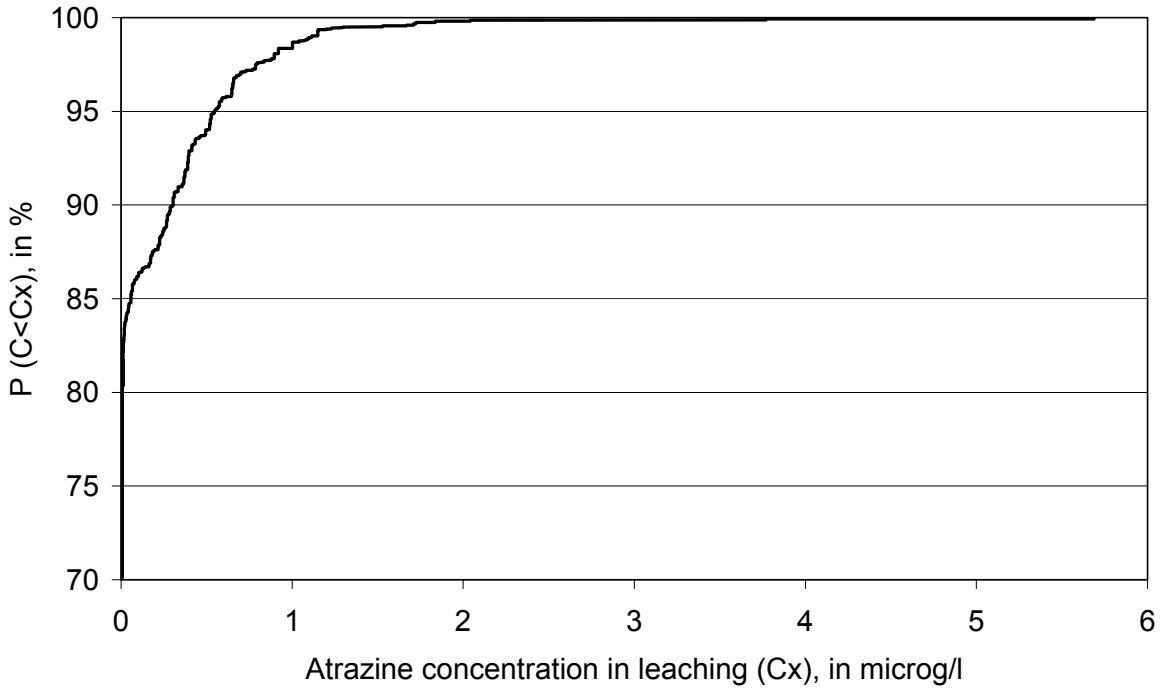


Figure 41. Cumulative distribution chart for concentrations of atrazine in leaching (median case scenario).

The y axis displays the probability that a concentration C is below the reference concentration Cx for the modelling scenarios considered.

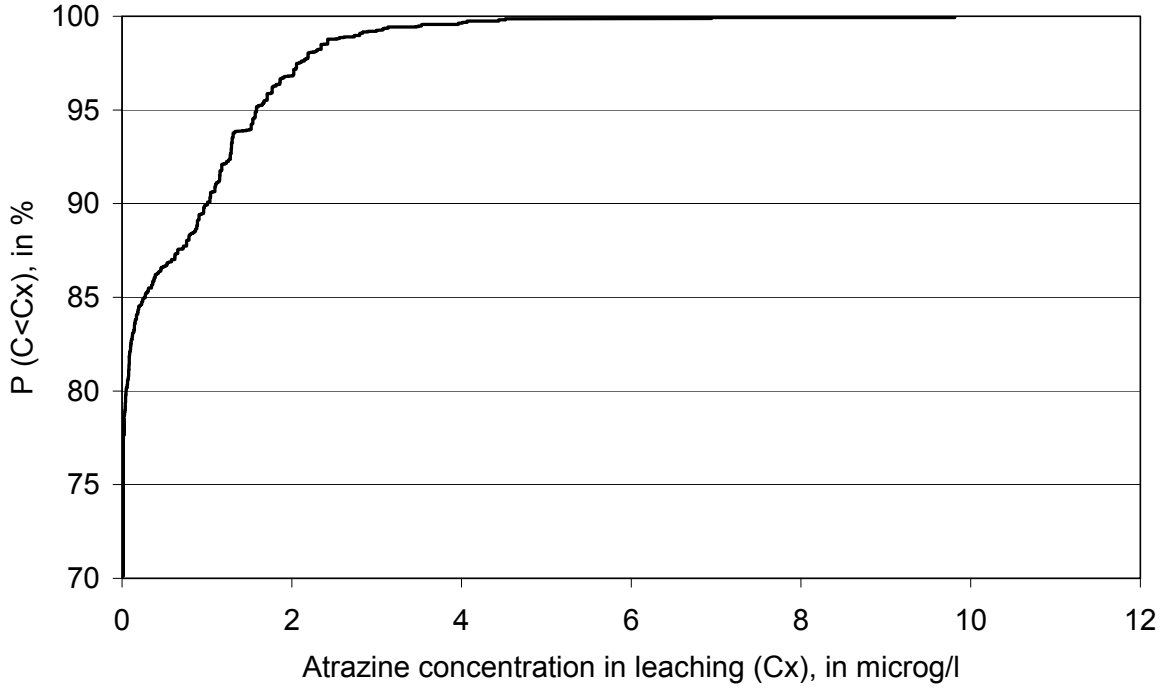


Figure 42. Cumulative distribution chart for concentrations of atrazine in leaching (worse case scenario).

The y axis displays the probability that a concentration C is below the reference concentration Cx for the modelling scenarios considered.

	Probability of exceeding 0.1 μg/l (%)	95th percentile concentration (μg/l)
Better case (75th percentile Koc, 25th percentile DT50)	3.7	0.08
Median case (50th percentile Koc, 50th percentile DT50)	13.8	0.55
Worse case (25th percentile Koc, 75th percentile DT50)	17.7	1.58

Table 20. Probabilistic results obtained through scenario-based modelling for atrazine within the context of a groundwater leaching assessment (overall estimates at the scale of the maize growing area).

Results have been derived from Figures 40 to 42.

6 MERITS AND SHORTCOMINGS OF THE INDIVIDUAL APPROACHES USED

6.1 FOCUS modelling

FOCUS modelling for atrazine for the four FOCUS scenarios representative of leaching conditions in England and Wales was initially undertaken. This simple standardised approach accounts for uncertainty to a limited extent as the models are run for different locations and for a number of different years with different weather data. The system is not designed to account for most of the uncertainty associated with leaching assessments (for example, only one set of pesticide properties is considered). Rather, the FOCUS system provides a benchmarking for the potential for a compound to leach within the European agricultural area.

6.2 Simple refinements to the FOCUS modelling

In an effort to provide a simple intermediate approach between the first-step FOCUS procedures and the more demanding investigations such as Monte Carlo modelling and scenario-based modelling, the FOCUS approach was refined by considering the influence on model predictions of i) the use of different application dates in the modelling; and, ii) the use of different combinations of Koc and DT50, the two parameters to which a pesticide fate model is normally most sensitive.

Model predictions were influenced only slightly by changes in application date, but the variability could be important in borderline cases. The variation of the application date was restricted to a narrow interval (initial application date +/- 7 days) for demonstration purposes, but this interval would usually extend to boundary dates allowed by the Good Agricultural Practice table for the product of interest. Greater influence of application date on model output can be expected for mobile compounds with a short half-life or for compounds applied very close to the beginning or end of the winter leaching period.

Model predictions were found to be largely influenced by sorption and degradation parameters emphasising the large uncertainty associated with the selection of representative values. The approach adopted provides a simple estimation of the uncertainty for relatively little additional effort. However, the calculations provide no additional information on the relative likelihood of leaching under actual conditions of use and the approach is of limited practical value in moving forward from basic FOCUS groundwater calculations.

6.3 Monte Carlo modelling

A simple implementation of Monte Carlo modelling was undertaken to demonstrate the nature of results that can be obtained through this technique. More in-depth assessments based on the technique, in particular two-dimensional Monte Carlo analysis which aims at separating between stochastic variability and uncertainty (in the sense of lack of knowledge; also known as 'incertitude'), were considered outside the scope of the present illustrative exercise. In comparison to earlier assessments which concentrated on the 80th-percentile concentrations, data for the 20 years simulated in each run were used in the leaching assessment. The probability of exceeding 0.1 µg/l based on the simulation of concentrations at 1-m depth for 20 years for the Okehampton leaching scenario when accounting for the uncertainty in Koc and DT50 values was 53.5% (average of five runs with replicated random samples). It should be noted that the unique scenario simulated does not cover the range of atrazine use conditions in England and Wales. The leaching estimate was found to be influenced by the random sample generated (CV of 6-15% for specific percentile concentrations were reported). A number of subjective choices were made in the analysis (the use of a triangular distribution as opposed to other distributions, the absence of truncation in the sampling) and it should be noted that other choices may have resulted in different estimates. Critically, it was assumed that there was no correlation between sorption and degradation of atrazine, so pairings of properties from extremes of the two distributions were allowed.

The Monte Carlo modelling undertaken here focused on the most vulnerable scenario identified from initial FOCUS modelling. The uncertainty in predictions arising from that associated with pesticide half-life and Koc was characterised. Aspects of subjectivity mean that the Monte Carlo analysis (in common with other approaches) should be seen as an interpretation of the data rather than as a definitive expression of the uncertainty. The probability distribution included data from all years and thus the assessment also included the uncertainty arising from inter-annual variation in weather data.

The primary drawback of the approach is that it is not informative about the relevance of the Okehampton scenario to areas of maize cultivation in the UK, nor about the vulnerability of the scenario relative to the broader agricultural area. Monte Carlo modelling built directly on FOCUS procedures, so key areas of uncertainty are still excluded. These might include (i) depth to groundwater and simulation of leaching below 1 m depth; (ii) intensity of maize cultivation and atrazine use and thus potential for dilution within groundwater by water

arising from non-treated areas; (iii) more robust descriptions of key processes such as degradation in the subsoil or time-dependent sorption.

As the modelling builds directly on the FOCUS groundwater framework, it would be reasonably straightforward to agree a regulatory endpoint for comparison against 0.1 µg/l. Standard FOCUS modelling nominally combines 80th percentiles for weather and soil properties with median values for pesticide properties to give an overall 90th percentile value on which to base a regulatory decision. Selecting the 80th percentile annual average concentration from the probability distribution in Figure 33 should give a similar level of protection as the soil properties remain unchanged. However, Table 13 provides an 80th percentile concentration of 1.6 µg/l (CV 7%) which is almost an order of magnitude larger than the value obtained with basic FOCUS modelling for Okehampton (0.2 µg/l). This anomaly arises because i) the uncertainty arising from that in pesticide properties (ignored in standard FOCUS modelling) far outweighs that arising from inter-annual variation in weather patterns; and, ii) statistical distributions in the Monte Carlo modelling were parameterised in such a way that medians of the distributions were larger than medians of the sorption and degradation datasets. Monte Carlo modelling has thus quantified two major sources of uncertainty (pesticide properties and weather) to provide a distribution of results for the Okehampton scenario. However, the approach has not moved the regulatory process significantly forwards. To do this, it would be necessary to generate information on the relative likelihood of a particular result arising under actual conditions of use.

6.4 Scenario-based modelling

Scenario-based modelling consists of selecting a number of environmental scenarios covering the range of soil and climatic conditions to which the compound will be applied. In contrast to the FOCUS approach, the objective is not to provide a worst-case assessment, but to cover a wide range of vulnerability to leaching. Modelling is undertaken for each of the scenarios selected using long-term (20 years) weather sequences. Predicted concentrations in leaching are available for each scenario. An overall probability of impact at the larger scale can be obtained by weighting results for each scenario by their abundance in the broader landscape. In the case presented (an assessment of the risk of atrazine to impact on groundwater resources following an application to a maize crop), situations that were integrated into the assessment included those areas grown in maize, but which do not overlay aquifers and those areas which overlay aquifers, but which are not vulnerable to leaching because of the properties of the soils. The probability of having concentrations of atrazine in the water

percolating out of the soil profile in England and Wales when considering median Koc and DT50 values was estimated to be 13.8%. Uncertainty in pesticide properties was accounted for in the analysis in a simple manner by repeating the modelling with two combinations of Koc and DT50 values. Estimates based on better (75^{th}-percentile Koc and 25^{th}-percentile DT50) and worse (25^{th}-percentile Koc and 75^{th}-percentile DT50) cases were 3.7 and 17.7% , respectively.

The approach has already been applied to groundwater and surface water risk assessments and necessitates the definition of scenarios relevant to the compound considered each time an assessment is undertaken. In contrast to other approaches considered which were based on one or a number of worst-case scenarios, scenario-based modelling permits an assessment that better reflects the variability in soil and climatic conditions within a realistic framework. However, the default implementation of scenario-based modelling ignores the uncertainty resulting from that in sorption and degradation parameters, which is known to be of importance for the prediction of pesticide loss. In the present instance, an attempt at dealing with these sources of uncertainties was made by considering three different combinations of Koc and DT50. Run times for the MACRO model precluded a more robust analysis in combination with multiple environmental scenarios. The selection of scenarios to represent the spectrum of leaching conditions throughout the country for a particular compound and the subsequent model parameterisation is subjective. In common with the Monte Carlo approach, it is likely that different probability estimates would be obtained by different individuals.

The scenario-based approach attempted to quantify aspects of uncertainty associated with soil type, climate, weather patterns and pesticide properties. Whilst the impact of weather patterns was investigated through the use of 120 different years of data, the remaining aspects included only limited combinations (three to four). Realistic depths to groundwater were considered for two soils, but simulations were restricted to 2 m depth for soils with deeper groundwater because of a lack of data on substrate properties. Uncertainty arising from model error was not considered explicitly (as in Monte Carlo modelling), but the MACRO model was selected because it incorporates the process of preferential flow which was considered important within the assessment framework.

Scenarios were selected for the compound and use of interest and summary statistics on relative abundance were used to assess the relative likelihood of a particular outcome within the wider landscape. The approach could be refined by adding information on the intensity of occurrence of vulnerable scenarios to address likely dilution of any residues within groundwater. The resolution of SEISMIC (5 x 5 km pixels) means that information on

intensity of occurrence is more useful for a groundwater assessment (where water is drawn from a large area) than for a surface water assessment.

7 CONCLUSIONS

Accounting for uncertainty in exposure assessments can take numerous forms. In the present instance, a range of probabilistic approaches of varying complexity were applied to predict the likely concentrations of atrazine leaching to depth. Applications included two simple and two higher-tier approaches (Monte Carlo modelling and scenario-based modelling).

Three of the approaches built directly on FOCUS modelling and expressed to a lesser or greater extent the uncertainty arising from that associated with timing of application or properties of the pesticide. Whilst it may be useful to attempt to quantify uncertainty for FOCUS estimations of exposure, each approach yields a distribution of potential outcomes spread around the original deterministic output. This is more scientifically defensible than a single deterministic prediction, but the distributions are unlikely to support a regulatory decision where a standard FOCUS prediction exceeds 0.1 μg/l. The FOCUS scenarios are a consensus benchmarking system for leaching of pesticides whereby regulatory decisions can be reached based on a surrogate for concentrations in groundwater (*i.e.* annual average concentration in soil water leaching to 1 m depth). As such, the FOCUS approach rests on a framework of inter-related assumptions and it is probably inappropriate to impose higher tier modelling onto this framework.

Higher tier approaches attempt to provide a more accurate analysis of likely concentrations of a pesticide in groundwater. To do this, it is necessary both to address over-simplifications and worst-case assumptions in the modelling and to provide an estimate of the reliability of any resulting prediction (*i.e.* quantify uncertainties). Refined modelling might comprise the introduction of additional processes into the description of the system on the basis of experimental data (*e.g.* over-riding FOCUS assumptions on rates of degradation in the subsoil or including a two-site sorption routine). However, the first step in obtaining a probabilistic estimate of exposure seems necessarily to be the derivation of representative environmental scenarios. These scenarios should cover the range of conditions likely under the specific conditions of use and information should be available on their relative extent.

Higher-tier approaches move away from some of the worst-case assumptions inherent in Step 1 calculations, so it becomes increasingly desirable to provide a measure of the

uncertainty associated with any predicted outcome. The Monte Carlo and scenario-based modelling demonstrated above are complementary to some extent. Monte Carlo modelling provides a solution to accounting for the uncertainty in model predictions resulting from that in model input parameters (and in particular to that in sorption and degradation values). In contrast, scenario-based modelling is particularly suited to the uncertainty in the modelling resulting from the consideration of a range of environmental conditions. Ideally, the two approaches should be combined, *i.e.* the uncertainty in the model predictions for scenario-based modelling resulting from that in Koc and DT50 values should be assessed through Monte Carlo modelling. Although this integrated approach is desirable, its implementation is currently limited by the running time of pesticide fate models and the lack of knowledge about i) the robustness of probability estimates derived through Monte Carlo approaches and ii) the influence of the subjectivity inherent in the scenario-based modelling. Numerous model runs will be required in Monte Carlo modelling to try to achieve robust probability estimates. The preferential flow model MACRO appears to be the most appropriate model for simulating a wide spectrum of leaching conditions in England and Wales, but the slow running time of the current version (version 4.3b) means that the running of the model for tens/hundreds of thousands of times cannot be envisaged at this stage. The release of a new version of the model or its emulation may address computational issues.

Although the application of probabilistic approaches to environmental risk assessments for pesticides currently receives much attention in Europe, little research has been devoted to aspects of the estimation of aquatic exposure. The present chapter has presented a range of options to assess the probability of leaching of atrazine in England and Wales. It is intended to promote discussion on the role that should be given to probabilistic approaches in risk assessment procedures.

References

Brown C.D., Hollis J.M., Bettinson R.J. & Walker A. (2000). Leaching of pesticides and bromide through lysimeters from five contrasting soils. Pesticide Science, 56:83-93.

Doherty J. (2000). PEST – model-independent parameter estimation. Watermark Computing, Corinda, Australia.

Dubus I.G. & Brown C.D. (2002). Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. Journal of Environmental Quality, 31:227-240.

Dubus I.G., Brown C.D. & Beulke S. (2000). Sensitivity analyses for leaching models used for pesticide registration in Europe. SSLRC research report for DEFRA PL0532, 407pp.

ECOFRAM (1999). http://www.epa.gov/oppefed1/ecorisk/

EUPRA (2001). Probabilistic risk assessment for pesticides in Europe. Implementation and research needs. Report from the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products, The Netherlands, June 2001.

FOCUS (2000). FOCUS groundwater scenarios in the EU plant protection product review process. Report of the FOCUS Groundwater Scenarios Workgroup, EC document reference Sanco/321/2000, 197p.

Hallett S.H., Thanigasalam P. & Hollis J.M. (1995). SEISMIC – a desktop information system for assessing the fate and behaviour of pesticides in the environment. Computers and Electronics in Agriculture, 13:227-242.

Helton J.C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliability Engineering and System Safety, 42:327-367.

Hollis J.M., Smith H.L. & Brown C.D. (2001). Relevance of the European Union FOCUS groundwater scenarios to UK agriculture. SSRLC research report for DEFRA PL0528, 37 pp.

Jarvis N.J. (2001). Web site www.mv.slu.se/BGF/MACROHTM/MACRO.HTM, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Jarvis N.J., Hollis J.M., Nicholls P.H., Mayr T. & Evans S.P. (1997). MACRO_DB: a decision support tool for assessing pesticide fate and mobility in soils. Environmental Software, 12:251-265.

Renaud F. & Brown C.D. (2001). Comparison of vulnerability to leaching of two FOCUS groundwater scenarios and soils of the UK H1b leaching potential sub-class. Cranfield Centre for EcoChemistry research report No. JA3750E for DEFRA PL0536, 23 pp.

Saltelli A., Chan K. & Scott E.M. (2000). Sensitivity analysis. John Wiley & Sons Ltd, Chichester, UK.

Tiktak A., van den Berg F., Boesten J.J.T.I., van Kraalingen D., Leistra M. & van der Linden A.M.A. (2000). Manual of FOCUS PEARL version 1.1.1. RIVM report 711401 008, Alterra report 28, 144 pp.

Chapter 6

Overall conclusions

There is currently strong interest in the assessment of probabilistic approaches to refine environmental risk assessment procedures for pesticides. The present project represents one of the first critical assessments of the application of probabilistic approaches to models for predicting pesticide exposure. The main objective of the project was to assess the robustness of probabilistic approaches and to estimate the confidence that should be attributed to their results. Most of the investigations concentrated on Monte Carlo approaches as it is the most popular technique for dealing with uncertainty in environmental modelling, and in pesticide fate modelling in particular.

The uncertainty iceberg

A comprehensive review of the different sources of uncertainty that can potentially affect predictions of pesticide fate models was undertaken. The review demonstrated that pesticide fate modelling is laced with uncertainty, thereby supporting the use of probabilistic approaches to estimate the uncertainty associated with model predictions. Uncertainties which are expected to most affect estimations of exposure are those related to variability and uncertainty in the most influential parameters in models (parameter uncertainty), but also those associated with model selection, model error and model parameterisation. Monte Carlo modelling only considers the former type of uncertainty and may therefore ignore significant contributions to the overall uncertainty (see the concept of the 'uncertainty iceberg' in Chapter 2). This suggests that results of Monte Carlo modelling should be considered *uncertain in nature*. The degree of uncertainty introduced by those sources of uncertainty which are not covered by Monte Carlo modelling should be investigated further so that the level of confidence that should be attributed to Monte Carlo results is assessed.

Limitations of Monte Carlo modelling from a technical perspective

Monte Carlo modelling exercises were repeated for various numbers of model runs and for a number of replicated random samples (the random samples were generated by Latin Hypercube Sampling using different seed numbers). The probability of exceeding a specific concentration, an indicator of major interest within the context of environmental risk assessment for pesticides, was found to be significantly influenced by the seed number used in the sampling. The variation in the probability of exceeding 0.1 μg/l was found to decrease when increasing the number of model runs, but was still significant (coefficient of variations 5-9%) when a large number of runs (5000) were undertaken. Within the constraints imposed by running times of pesticide fate models, this suggests that Monte Carlo results should be considered *inherently variable* and that the levels of variability reported have the potential to affect decision-making. Similar issues can be anticipated for percentile concentrations (95^{th}-, 99^{th}-percentile concentrations) and the closer the concentration to the tail of the distribution the more inherent variability is to be expected. Two solutions to the repeatability issue can be considered: i) dramatically increasing the number of model runs to decrease variation in the results; and, ii) repeating the Monte Carlo modelling for a number of replicated random samples to assess variation in the results. Both solutions imply the running of pesticide fate models for a large number of runs (tens, hundreds of thousands of times). Limitations associated with the running time of the complex models used in pesticide fate modelling mean that robust implementation of Monte Carlo modelling would require a very large computational effort, which may not be compatible with a routine integration of the technique in the assessment of pesticide exposure.

The influence on Monte Carlo results of subjective choices made during the implementation of the technique was investigated by considering multiple options for parameterising lognormal distributions and repeating the modelling for different levels of parameter correlation. Probabilities of exceeding a concentration of 0.1 μg/l were found to be significantly influenced by these subjective choices. Levels of variation reported could be of relevance to decision-making if exceedance probabilities were close to threshold levels. User subjectivity in the undertaking of Monte Carlo modelling could be decreased by standardising modelling procedures, but this may conflict with the broad philosophy of higher tier risk assessment.

Probabilistic approaches and environmental risk assessment for pesticide registration

The deployment of probabilistic assessments is clearly positioned towards the higher end of higher tier approaches for environmental risk assessment. As such, the main objective of probabilistic approaches should be to provide a *more accurate* and *more realistic* assessment of exposure and effects. For exposure, it appears that the deployment of modelling approaches which consider the whole spectrum of environmental conditions and vulnerability to pesticide contamination, such as the scenario-based modelling presented here, is helpful for refining knowledge on the risk involved and should therefore be encouraged. Undertaking Monte Carlo modelling for assessing the uncertainty in model predictions resulting from the uncertainty in input parameters for one or a few artificial scenarios is unlikely to bring additional useful information to support decision-making. In particular, basing the Monte Carlo modelling on the FOCUS approach was deemed to be inappropriate as the FOCUS approach rests on a framework of inter-related assumptions. In contrast, applying Monte Carlo techniques to scenarios defined within the scope of scenario-based modelling would help to assess the influence of parameter uncertainty on model predictions for the various scenarios. In any case, the limitations attached to the different approaches should be kept in mind: i) the user subjectivity involved in the selection of vulnerability scenarios (scenario-based modelling); ii) the user subjectivity attached to the implementation of the Monte Carlo approach (selection of parameters to include in the analysis, selection and parameterisation of distribution functions, specification of correlation in the sampling); iii) the inherent uncertainty associated with Monte Carlo results due to the random component in the approach; iv) the user subjectivity involved in the selection and parameterisation of the pesticide fate model (both scenario-based and Monte Carlo modelling); and, v) the fact that the model may not represent behaviour in the field (model error; both scenario-based and Monte Carlo modelling). Future research should be directed towards addressing these limitations.

A common criticism of probabilistic modelling is that it can be used to justify moving away from some of the worst-case assumptions built into schemes for deterministic risk assessment. The case study presented in Chapter 5 illustrates that this is unlikely to be the case for estimations of aquatic exposure. Whilst current lower-tier procedures rest on a number of assumptions, the systems are relatively well characterised and assumptions are generally "reasonable worst-case" rather than "absolute worst-case". In this, aquatic exposure may differ from other disciplines (e.g. human dietary exposure to pesticides) where several absolute worst-case assumptions may be combined to provide an unrealistic estimate of

exposure. For aquatic exposure, a probabilistic assessment is likely to result in a distribution of exposure values around the original deterministic estimate, provided that the conceptual model of the system and the scenario remain the same. In fact, the challenge in estimating aquatic exposure will be to consider the broader relevance of modelling scenarios to actual areas of use and to move the conceptual model away from edge-of-field to consider factors influencing exposure within the broader landscape. Probabilistic risk assessment is not the tool to support this shift in approach, but it may prove invaluable in estimating the uncertainty around any more realistic estimate of exposure.

Concluding remarks

The research raised a number of conceptual and technical issues with regard to the typical application of Monte Carlo techniques to pesticide fate models. Some of these were related to the specific nature of the modelling undertaken to assess pesticide exposure within the registration context whereas some others were more universal (e.g. the fact that Monte Carlo modelling only addresses parameter uncertainty and may ignore other significant sources of uncertainty). The research suggests that extreme care should be exercised when considering results originating from Monte Carlo modelling and limitations of the technique in terms of robustness should not be overlooked. Nevertheless, the consideration of uncertainty in the modelling, and especially of that originating from sorption and degradation parameters, remains a priority as this constitutes a weak point in exposure assessment. Further research into Monte Carlo modelling as well as other probabilistic techniques for environmental exposure is required to try to address the issues outlined in the project and to increase the robustness of probabilistic estimates to support decision-making in the context of pesticide registration.