General enquiries on this form should be made to:
Defra, Science Directorate, Management Support and Finance Team,
Telephone No. 020 7238 1612
E-mail: research.competitions@defra.gsi.gov.uk

SID 5 Research Project Final Report

- **Note**

  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible. The SID 5 (Research Project Final Report) is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website. A SID 5 must be completed for all projects.

  A SID 5A form must be completed where a project is paid on a monthly basis or against quarterly invoices. No SID 5A is required where payments are made at milestone points. When a SID 5A is required, no SID 5 form will be accepted without the accompanying SID 5A.

- This form is in Word format and the boxes may be expanded or reduced, as appropriate.

- **ACCESS TO INFORMATION**

  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: PS2305

2. Project title: Development of a web-based pesticide risk assessment module for below-ground invertebrates

3. Contractor organisation(s):
(1) University of Southampton, UK
(2) ECT Őkotoxikologie GmbH, Flörsheim, Germany
(3) Alterra, Wageningen, The Netherlands
(4) NERI, Silkeborg, Denmark

4. Total Defra project costs: £ 200,256

5. Project: start date .... 01 December 2002

end date .... 30 June 2005

6. It is Defra's intention to publish this form.
Please confirm your agreement to do so..YES ☒ NO ☐

(a) When preparing SID 5s contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.

Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the SID 5 can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.

In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

(b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

- A systematic review was carried out to investigate the effects of pesticides on soil invertebrates in lower-tier (laboratory) and higher-tier (semi-field, terrestrial model ecosystem (TME) and field) studies. Over 2000 data records for relevant endpoints (e.g., acute mortality, chronic reproduction, or field effects) were quality-checked and extracted into Excel databases. The data represents 277 pesticide substances and more than 70 invertebrate species or groups.

- In laboratory studies, the standard earthworm test species *Eisenia* spp. (*Eisenia fetida* + *E. andrei*) is one of the least sensitive soil invertebrate groups to insecticides. Examination of acute-chronic ratios shows that for pesticides overall, in the majority (85%) of cases, the acute mortality earthworm test would not provide the most sensitive risk assessment.

- In model ecosystem and field studies, the modal effects class overall (71 pesticides studied) is for no observed effect. Arachnida, Nematoda and Formicidae usually are not affected by fungicides, whilst Lumbricidae, Nematoda and Formicidae appear generally unaffected by herbicides.

- Pronounced and persistent effects in higher-tier studies have been observed particularly when Lumbricidae and Enchytraeidae were exposed to fungicides and when Arachnida, Collembola and Lumbricidae were exposed to insecticides.

- The available data in the public domain is biased towards the same soil invertebrates that are used as standard test species in the regulatory risk assessment of pesticides. Limited ecotoxicological information is available on species that might be considered more ecologically relevant. Most pesticides (88%) have fewer than 10 data sets; relatively few pesticides (16%) have both laboratory and higher-tier data.

- Based on a minimum of five input data, species sensitivity distributions (SSD) for soil invertebrates and hazardous concentration estimates (HC_5) can be derived for only eleven pesticide substances (atrazine, carbendazim, chlorpyrifos, copper compounds (grouped), diazinon, dimethoate, gamma-HCH, lambda-cyhalothrin, parathion, pentachlorophenol and propoxur). Sensitivity distributions are restricted to acute mortality (LC_{50}) due to a lack of data for other toxicity endpoints.

- Of the eleven pesticides for which HC_5 could be calculated (based on a minimum of five species data), copper compounds have no comparable higher-tier data and dimethoate only one higher-tier data set. Higher-tier NOEC and LOEC can be estimated only for atrazine. To clarify acceptability thresholds for higher-tier effects, we suggest that appropriate low concentrations should be included in TME and field studies.

- For the pesticides that had comparable higher-tier data, the range of uncertainty in the HC_5 estimates included concentrations at which pronounced higher-tier effects on soil invertebrates have been observed. Based on these (limited) data, the HC_5 estimates for these pesticides would at best be protective only of individual taxonomic groups, not the intended 95% of species.

- Arthropods and oligochaete worms clearly differ in their sensitivity to pesticides. Addition of a single Collembola species to a Lumbricidae-based SSD for chlorpyrifos resulted in decrease in both the HC_5 estimate and its certainty by approximately four orders of magnitude.

- Distribution-based risk assessment approaches such as SSD are currently incompatible with lower-tier testing of pesticides on soil invertebrates because insufficient information is generated to characterise sensitivity distributions. TME studies might support distribution-based risk assessment in future, if they enable concentration-response data to be obtained simultaneously for different species.

- So far, TME studies have only been evaluated for carbendazim. There is good agreement between the effects of carbendazim observed in TME studies and those observed in field studies, but the comparison is available for only four soil invertebrate groups.

- At present, HC_5 does not appear promising as a tool to improve the soil invertebrate pesticide risk assessment, due to (1) insufficient data to calculate sensitivity distributions, (2) insufficient higher-tier information to validate distribution-based predictions, and (3) evidence suggesting that in the few cases that HC_5 can be calculated they would not adequately protect soil invertebrate communities.

- The implications of these findings for the development of an internet-based risk assessment module for this project are discussed. We conclude that the limited availability of data does not preclude the development of a useful internet risk assessment module for soil invertebrates, although there would be less support available from empirical data than in other related WEBFRAM risk assessment modules.

The findings of this project suggest that some improvements could be made in the way pesticide effects studies are carried out and reported, to clarify interpretation of effects and prediction of risk to soil invertebrates. We recommend research that:

(1) investigates pesticide risks for species other than the standard test species;

(2) broadens the range of species sensitivity data to include more species and chemicals; and

(3) addresses the limitations of current studies, particularly with regard to:

(i) testing appropriate concentration ranges in higher-tier studies;

(ii) improving (and where possible harmonising) the reporting of studies;

(iii) evaluating the potential of TMEs for risk assessment of a range of chemicals representing different toxic modes of action.

These recommendations represent good scientific practice and would clarify the ecological implications of the current deterministic risk assessment scheme; they are not dependent on whether pesticides regulatory agencies endorse the use of distribution-based approaches to risk assessment.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also

seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:

- the scientific objectives as set out in the contract;
- the extent to which the objectives set out in the contract have been met;
- details of methods used and the results obtained, including statistical analysis (if appropriate);
- a discussion of the results and their reliability;
- the main implications of the findings;
- possible future work; and
- any action resulting from the research (e.g. IP, Knowledge Transfer).

1. Project objectives

1.1 Scientific objectives of the project

The overall objective of this project (WEBFRAM-5) is to investigate whether the application of probabilistic methods of uncertainty analysis can improve the risk assessment for effects of pesticides on below-ground invertebrates. The project addresses this objective in two ways: First, analysis of existing pesticide effects data is used to examine which methods of uncertainty analysis could be used. Second, where possible (subject to availability of data), internet-based risk assessment tools are provided to enable stakeholders to independently and impartially evaluate the current deterministic risk assessment approach against alternative methods. The project output aims to be consistent with the current EU regulatory guidance (according to the Directive 91/414/EEC and its amendments) and will complement the output from other WEBFRAM projects on non-target arthropods (WEBFRAM-4), birds and mammals (WEBFAM-3 & 7) and aquatic organisms (WEBFRAM-2). The WEBFRAM projects together (coordinated in project WEBFRAM-1), are intended to provide a comprehensive internet-based risk assessment resource by autumn 2006.

The WEBFRAM-5 project has five specific objectives: (1) acquire relevant data; (2) identify initial (year 1) pesticide use scenarios and case studies for evaluation of uncertainty analysis methods; (3) develop and refine the case studies for internet-enabling as online risk assessment tools; (4) identify further (year 2) pesticide use scenarios and case studies for evaluation of uncertainty analysis methods, and refine these for internet enabling; (5) finalise testing of the internet risk assessment tools and disseminate the findings.

1.2 Progress with the scientific objectives

Objective 1 has been completed using a systematic review approach to locate and obtain relevant pesticide effects data for soil invertebrates. In total, 2920 effects data sets have been extracted into Excel databases, comprising 1950 lower-tier (laboratory) data sets and 970 higher-tier (model ecosystem and field) data sets.

Objectives 2, 3 and 4 have been completed as far as the development of three case studies of pesticide effects on soil invertebrates: (a) effects of the herbicide atrazine on earthworms (Lumbricidae) in maize; (b) effects of the fungicide carbendazim on earthworms in winter cereals; (c) effects of winter and summer applications of the insecticide chlorpyrifos on earthworms and springtails (Collembola) in winter or spring cereals. These case studies will provide example toxicity and exposure data to support the internet-based risk assessment tools. The web-enabling of the internet tools is being undertaken primarily by DEFRA Central Science Laboratory (CSL) in collaboration with the Cadmus Group (in project WEBFRAM-1). The WEBFRAM-5 (PS2305) project team and CSL have agreed an internet format for the project output and this is currently being prepared for uploading by CSL and Cadmus.

Objective 5 has not been fully completed yet as testing of the internet models can only commence once the models have been fully uploaded to the internet. However, the species sensitivity distribution model that will be employed has been evaluated in detail and some potential limitations of the initial version identified. The revised version should provide greater flexibility for stakeholders to control graphical output.

WEBFRAM-5 is the first of the WEBFRAM projects to have its final internet risk assessment models uploaded to the internet. Accordingly, CSL are checking that the proposed format meets the requirements of both PSD and the other WEBFRAM projects. Slight delays in the internet enabling of the project output beyond the formal end data of WEBFRAM-5 mainly reflect this effort to harmonise the overall WEBFRAM output and ensure efficiency of the future uploading of the internet tools. Full completion of all project objectives is expected once these checks have been completed.

2. Methods and materials

A summary of the methods is given here; for further details please see Frampton et al. (2006), Jaensch et al. (2006) and the full project report (86pp).

2.1 Data collection

A systematic review was carried out to locate pesticide effects data for soil invertebrates. For reasons of commercial confidentiality, data owned by agrochemicals companies is excluded. Accordingly, this report is confined to information available freely in the public domain.

Pesticide active substances (excluding microbial and fungal) were considered relevant if they are currently approved, or have previously been approved, for use in European agriculture. Copper compounds (acetates, carbonates, chlorides, nitrates, oxides, oxychlorides and sulphates) are included as fungicides (grouped together) because they are the group of pesticides tested most widely on soil invertebrates. However, most of the copper data is from ad hoc studies, with only one commercial formulation (containing copper oxychloride) represented. Data was also extracted for some other pesticide-related substances that have been extensively tested on soil invertebrates. These include 4-nitrophenol (a metabolite of the insecticide parathion), 3,4-dichloroaniline (a metabolite of the herbicide diuron) and monochlorobenzene (a pesticide formulation additive). Two types of data were extracted from the literature and databases: lower-tier data (from laboratory tests) and higher-tier data (from terrestrial model ecosystem (TME), semi-field or field tests).

2.1.1 Lower-tier data

Data were extracted if a relevant measurement endpoint (e.g., acute or chronic LC_{50}, EC_{50} or NOEC) was given, together with sufficient supporting information to allow the endpoint to be interpreted (i.e., information on the species and test conditions, pesticide formulation type, application rate and the type and duration of exposure). Data was accepted for any euedaphic (soil-dwelling) invertebrate groups other than micro-organisms.

Relevant data were extracted into an Excel database. A separate data record was entered for each of the following test variables: species, pesticide type, substrate type, exposure concentration, assessment endpoint (e.g., growth, reproduction or mortality) and test duration (acute or chronic). Each data record contains all the available toxicological effects information for the given combination of test variables.

2.1.2 Higher-tier data

Data from terrestrial model ecosystem (TME), semi-field and field studies were scrutinised and identified as relevant if all of the following criteria were met:

- The data concerned euedaphic invertebrates (excluding micro-organisms), or were reported for a euedaphic life stage (for example the larvae of certain flies and beetles). If no further discrimination was possible, broad taxonomic groups were included (i.e., total Arthropoda or total Insecta).
- The measurement endpoint(s) given were relevant to field populations or communities. The assessed endpoints were abundance, biomass, mortality, development and behaviour (e.g., surface migration).
- Effects of specified individual pesticide(s) were not confounded with other variables (such as other chemical applications).
- The spatial and temporal characteristics of the reference (control) treatment were appropriate for interpreting pesticide effects.
- The study design and analysis were reported in sufficient detail to enable quantitative interpretation of pesticide effects (i.e., based on adequate replication and statistical evaluation).

Relevant data were extracted into a second Excel database. A separate data record was entered for each species (or group), pesticide type, application method and geographical locality as well as for some other variables that differed between studies (e.g., crop type and sampling date).

2.2 Data reliability

Lower-tier data were assigned to four reliability classes following a scheme proposed by Klimisch et al. (1997): (1) Reliable without restriction; (2) Reliable with restrictions; (3) Not reliable; (4) Not assignable. Due to the complexity and singularity of higher-tier studies (nearly no guidelines exist) it was not possible to assess the reliability of the study results in the same way as for laboratory tests. The only field test guideline using soil invertebrates (for earthworms; ISO 1999b) is usually performed as part of the pesticide registration process; thus these results are not available in the open literature. The reliability of higher-tier studies was therefore evaluated according to expert judgement. Studies were scrutinised in terms of their ecological relevance (the species monitored and the spatial and temporal aspects of the study design) and whether the study design and analysis were appropriate (taking into consideration the extent of replication, possible bias or confounding of variables, and the appropriate application of statistical tests). To avoid discarding relevant information, higher-tier studies were excluded only if they were considered clearly unreliable (for example if based on an inappropriate experimental design or presenting erroneous or biased data).

2.3 Data analysis

2.3.1 Lower-tier data

If the sensitivity to a pesticide is known for a range of species, parameters of the species sensitivity distribution (SSD) may be used to predict the concentration of the substance at which a certain proportion *x* (%) of the species would be affected (the hazardous concentration, HC*x*). This concept, which was conceived by Kooijman (1987) and van Straalen & Denneman (1989), and developed by Wagner & Løkke (1991) and Aldenberg & Slob (1993), is recognised as a potentially useful tool in environmental protection and risk assessment (e.g., Hart 2001). The HC_5 is usually used as a benchmark for the (ecologically arbitrary) maximum permissible environmental concentration, i.e. the concentration that would affect no more than 5% of the species, giving a '95% protection level'. The SSD concept depends on a number of assumptions, including that the chosen distribution (usually log-normal or log-logistic) is appropriate and that the species represented in the SSD are representative of those in the environment to which the HC_5 applies. Critics of the SSD approach question the validity of these and other assumptions (Forbes & Calow 2002), and also the implication that 5% of species are effectively considered expendable (Hopkin 1993). Nevertheless, the SSD concept is increasingly being used to inform decisions in regulatory (particularly aquatic) risk assessment, although with few exceptions (e.g., Badejo & van Straalen 1992) it has not been used previously for assessing pesticide risks to soil invertebrates. In this project the SSD and HC_5 were calculated as they are probabilistic approaches that can account for uncertainty in species sensitivities.

Where possible, SSD were fitted to the toxicity data and HC_5 were calculated using the program EtX-2000 version 1.409 (van Vlaardingen et al. 2003). Confidence limits (90%) were calculated to indicate uncertainty of the HC_5 estimate. A species sensitivity distribution can be derived from as few as 2 species' data but to avoid estimates of the HC_5 being influenced by the sample size, at least 10-15 data values (depending on the pesticide) have been recommended (Wheeler et al. 2002). However, it is difficult to meet these requirements with soil invertebrate data. It is also questionable whether as few as two data could capture the interspecific variation in sensitivity of soil invertebrates. Accordingly, an arbitrary minimum of six soil invertebrate species' data was applied when calculating SSD and estimates of HC_5. An arbitrary minimum of six species' data has also been used to derive SSD and HC_5 estimates in aquatic risk assessment (Campbell et al. 1999; Maltby et al. 2005).

For each pesticide that met the minimum requirement of providing six relevant effects concentration estimates, an SSD and estimate of the HC_5 were calculated. In practice, SSD could only be calculated for acute mortality (LC_{50}) as too few data were available for other endpoints.

In cases where several relevant toxicity data were available for the same species, the geometric mean effect concentration was used as the input data. Two types of data analysis were carried out:

- *All-data approach*: All relevant data were used in the calculation of the SSD.
- *Comparable-data approach*: Only data obtained under comparable test conditions were included (for example, in the case of earthworms, only data acquired according to the standard OECD test guidelines 207 and 222 were used). Data were selected if a standard (OECD) soil was used with pH 5.4-7.5, OM content ~10% and moisture 40-60% of maximum water holding capacity).

The all-data approach makes best use of the available data, irrespective of variability, whereas the comparable-data approach should reduce variability but could lead to a smaller sample size.

2.3.2 Higher-tier data

Due to the heterogeneity of higher-tier studies, which vary considerably in their aims, methods, experimental designs and in the way they report results (for example in the amount of detail given), no single effect measure can be clearly compared across all the studies. To overcome this problem, for each pesticide concentration tested the results were assigned to one of five pesticide effect classes, which cover the range from no observed effect to a pronounced long-term effect. These classes are based on initial recommendations of Brock et al. (2000) for assessing higher-tier effects of pesticides in aquatic systems; the recommendations have since been adopted by the European Union for regulatory risk assessment in aquatic systems (EC 2002b). As no clear framework exists for defining effects on soil invertebrates, the aquatic effects classification of Brock et al. (2000) was adapted for use with the soil invertebrate data. The magnitude of effects was assessed according to guidance for non-target arthropods, in which clear pesticide effects are (arbitrarily) defined as those exceeding 30% compared to a control treatment (Barrett et al., 1994). Here, the 30% threshold is used to distinguish small and large effects, whilst statistical significance is used primarily to distinguish small effects from no effects.

The effect classes used for the terrestrial invertebrate data are essentially the same as those proposed by Brock et al. (2000) for structural endpoints, in which the threshold for assessing recovery is 100 days; this is an arbitrary period, used in the absence of clear ecological guidance for harmonising the reporting of results. The effect classes are:

Class 1: *Effect could not be demonstrated*. No clear causal relationship observed between pesticide and control treatments (primarily based on statistical significance).
Class 2: *Slight effect*. Effects of small magnitude (≤ 30%) and short duration (< 100 days).
Class 3: *Pronounced short-term effect*. Effects of large magnitude (> 30%) but short duration (< 100 days).
Class 4: *Pronounced effect in short-term study*. Effects of large magnitude (> 30%) but the study too short (or the sampling interval too long) to demonstrate complete recovery within 100 days.
Class 5: *Pronounced long-term effect*. Effects of large magnitude (> 30%) and long duration (≥ 100 days).

In cases where no distinct classification of effects was possible (for example, when no statistical significance testing was reported or in studies of behaviour) expert judgment was applied; in borderline cases, effects were usually assigned to the higher effect class in order to remain on the protective side when considering risk to soil invertebrates.

The biological data were grouped as Arachnida, Chilopoda, Coleoptera (larval stages), Collembola, Diptera (larval stages), Enchytraeidae, Formicidae, Lumbricidae and Nematoda. If no further discrimination was possible also total Arthropoda and Insecta were included if it was clear that these were recorded by soil sampling. The assessed endpoints were abundance, biomass, mortality, behaviour (e.g., surface migration) and development. Higher-tier risk was explored by plotting the effects class against the pesticide concentration. If possible, a community NOEC was estimated.

2.3.3 Comparison of lower-tier with higher-tier data

For laboratory and field risk estimates to be comparable, laboratory and field effect concentrations must be expressed in the same units. The exposure unit of kg ha^{-1} used in field experiments was therefore converted to mg kg^{-1}. A factor of 1.33 was used, for consistency with the current risk assessment scheme which assumes that the relevant soil exposure depth is 5cm and the soil dry bulk density is 1500 kg m^{-3} (EPPO 2003).

For pesticides that had sufficient data to enable calculation of HC_5, the degree of protection offered by the lower-tier HC_5 estimates was investigated by comparing the HC_5 (± confidence limits) with the field effects data (where available). Comparisons were also made between relevant deterministic lower-tier risk estimates (e.g. the lowest available NOEC for a given invertebrate species or group) and the higher-tier effects.

3. Principal results

A summary of the results is given here; more information on the pesticides and species studied is given by Frampton et al. (2006) and Jänsch et al. (2006).

3.1 Availability of data

The systematic search for information about pesticide effects on soil invertebrates yielded over 1000 publications, of which about 400 papers were selected for data extraction. The majority of the databases examined focused on pesticide effects data for aquatic endpoints, with very little terrestrial invertebrate data given. The data set for WEBFRAM-5 comprises 2920 pesticide effects data records altogether, of which 1950 are from lower-tier studies and 970 from higher-tier studies. After standardising the higher-tier data to 11 higher taxonomic groups (Section 2.3.2), 653 comparable higher-tier data sets are available.

3.1.1 Pesticides

Lower-tier and higher-tier data are available, respectively, for 250 and 71 active substances. About 58% and 42% of the lower-tier data originate from acute and chronic studies respectively. The majority of higher-tier data (79%) originated from field studies, with the remainder from semi-field or terrestrial model ecosystem studies. Data records from both lower-tier and higher-tier studies are available for 45 (16%) of the pesticides. A key feature of the data is that for many pesticides the data are biased either towards lower-tier or higher-tier studies. The largest numbers of pesticide effects data records are for carbendazim and copper compounds; however the copper data is entirely restricted to lower-tier studies (Fig. 1). Carbendazim has the largest overall number of data sets that includes both tiers, with 119 lower-tier and 276 higher-tier records. The majority of pesticides (92%) have fewer than 20 data records.

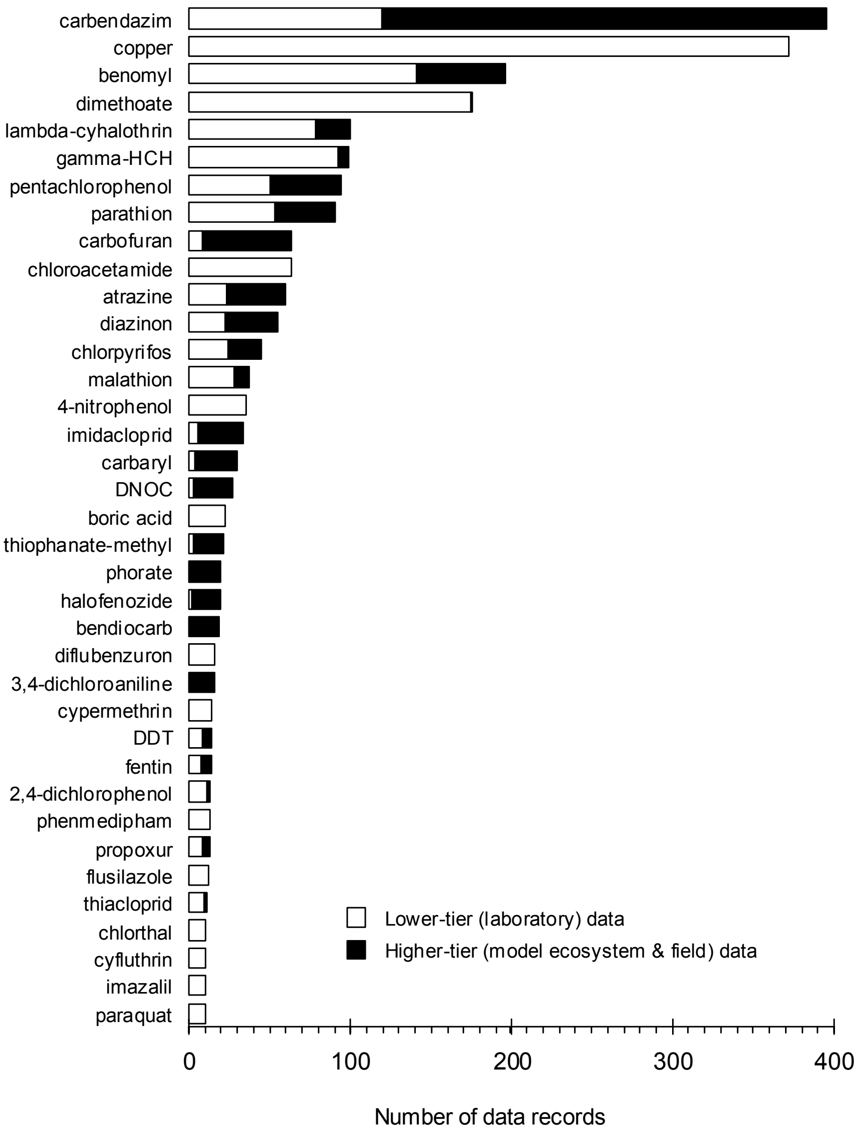

Fig. 1. Availability of soil invertebrate pesticide effects data for pesticides and related substances with 10 or more data records.

3.1.2 Soil invertebrates

Laboratory and higher-tier data are available, respectively, for 67 and 42 invertebrate species or groups. Overall, Lumbricidae, Collembola and Enchytraeidae have been the most frequently tested soil invertebrates in pesticide studies, contributing 90% of the lower-tier data and 65% of the higher-tier data.

Within each of the four main groups of soil invertebrates studied in lower-tier tests (Lumbricidae, Collembola, Enchytraeidae and Acari), more than half of the data records are represented by one genus: *Eisenia* spp. contributes 65% of the Lumbricidae data records, *Folsomia* spp. 59% of the Collembola records, Enchytraeus spp. 78% of the Enchytraeidae records and *Platynothrus peltifer* 52% of the Acari records.

In higher-tier studies, 80% of the data records are represented by Arachnida, Collembola, Lumbricidae and Enchytraeidae. Data were usually reported for broad taxonomic groups; individual species or genera have been identified in a relatively small proportion of the studies. Unidentified earthworms contribute 69% of the Lumbricidae data whilst unidentified Collembola contribute 58% of the Collembola data. The standard collembolan test species *Folsomia candida* (ISO 1999a) accounts for only 3% of the higher-tier Collembola data whilst the standard earthworm test species *Eisenia fetida* + *E. andrei* (ISO 1998) were not identified at all in the higher-tier data for Lumbricidae, despite *E. fetida* + *E. andrei* being the most frequently tested of all soil invertebrate species in laboratory studies. This reflects the ecological requirements of these *Eisenia* species, which occur mainly in man-made accumulations of rich organic material with high moisture content such as compost heaps (Jänsch et al. 2005).

3.1.3 Effects endpoints

Overall, in lower-tier studies the acute LC_{50}, chronic NOEC, acute NOEC, chronic LOEC and chronic EC_{50} are the endpoints most often reported (29%, 18%, 16%, 14% and 9% of the data records respectively). The most frequent measurement endpoints in higher-tier studies (89% of the data sets) are abundance and/or biomass (which were often reported together, with biomass estimated from the abundance and mean weight per species), followed by mortality (10% of the data sets). Very few studies assessed behaviour or development (together less than 1% of the data sets).

3.2 Data reliability

Of the 1950 lower-tier soil invertebrate data records, 175 were classed as ‘reliable without restrictions’, 1104 as ‘reliable with restrictions’, 347 as ‘not reliable’ and 324 as ‘not assignable’. (overall 65% were classed as reliable, 35% unreliable). For the standard earthworm test species (*Eisenia fetida* + *E. andrei*), the corresponding numbers of data records are 164, 621, 83 and 85 (82% classed reliable). For the standard springtail test species *Folsomia candida*, the corresponding numbers are 8, 63, 30 and 5 (67% classed reliable). The modal reliability class for most soil invertebrates is 2 (‘reliable with restrictions’). However, for unidentified earthworms the modal class (97% of the data) is 4 (‘reliability not assignable’). These results suggest that less taxonomic detail is reported in studies that are not carried out according to standard guidelines.

3.3 Sensitivity of the standard test species *Eisenia* spp. and *Folsomia candida* to a range of pesticides

For insecticides, *Eisenia fetida* + *E. andrei* is one of the least sensitive invertebrate groups. *Folsomia candida* is among the most sensitive invertebrates to the insecticides dimethoate, chlorpyrifos and gamma-HCH (Fig. 2).

3.4 Relationship between acute and chronic endpoints for earthworms (*Eisenia fetida* + *E. andrei*)

According to the Guidance Document on Terrestrial Ecotoxicology (EC 2002a) (which is based primarily on experiences with earthworms), it is implied that, for a given species and pesticide, the need for a chronic reproduction (NOEC) test can be identified by the LC_{50} obtained from an acute mortality test. This assumes that an extrapolation can be made from acute mortality to chronic reproduction (the risk assessment does not differentiate separate extrapolations from acute LC_{50} to acute NOEC and from acute NOEC to chronic NOEC).

The cumulative density function of the acute-chronic ratio ($\log_{10}(LC_{50}$ / NOEC)) for earthworm data (Frampton et al., 2006) indicates that in 50% of cases the chronic test will be more than 10 times as sensitive as the acute test and in the majority (85%) of cases be more than twice as sensitive as the acute test. Because the relationship between the TER for acute studies (trigger value 10) and chronic studies (trigger value 5) is two, in 85% of cases the chronic test would be more a sensitive indicator of risk than the acute test (i.e., in 85% of cases the acute mortality test does not provide the most sensitive risk assessment).

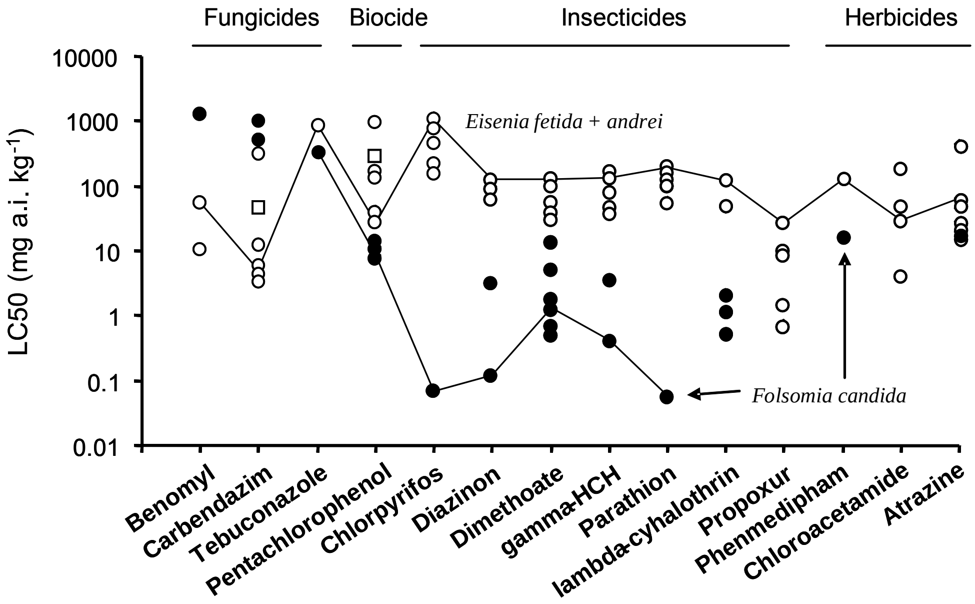

Fig. 2. Sensitivities of different soil invertebrate species (○ = oligochaetes, ☐ = nematodes, ● = arthropods) to pesticides illustrated with acute mortality data. Each data point represents the geometric mean of the LC_{50} for one species. Data points for the standard test species *Eisenia fetida* + *E. andrei* and *Folsomia candida* are connected by lines.

3.5 Lower-tier risk estimates and their uncertainty

Based on the minimum requirement of five input data, using the comparable-data approach, species sensitivity distributions can be derived for only four pesticides: chlorpyrifos, grouped copper compounds, dimethoate and pentachlorophenol (PCP). When all the available data are used, SSD can be derived for eleven pesticides: atrazine, carbendazim, chlorpyrifos, grouped copper compounds, diazinon, dimethoate, gamma-HCH, lambda-cyhalothrin, parathion, PCP and propoxur. However, only ten of these pesticides (excluding copper) have higher-tier data for comparison (Table 1). With the exception of chlorpyrifos and parathion, the sensitivity distributions pass the goodness-of-fit test (Anderson-Darling) for the log-normal distribution. The chlorpyrifos and parathion SSD pass goodness-of-fit tests if based on Lumbricidae data only, but not if the LC_{50} data for *Folsomia candida* is included.

The inclusion of *Folsomia candida* (Collembola) data has a marked influence on the SSD for chlorpyrifos. The 90% confidence interval for the chlorpyrifos HC_5 spans nearly four orders of magnitude, whereas if *Folsomia candida* data is excluded, the resulting Lumbricidae-only HC_5 has an uncertainty range of approximately one order of magnitude (Table 1). Excluding the *F. candida* data causes a marked shift of the SSD curve to the right, in the direction of lower sensitivity, with a resulting increase in the HC_5 from 0.37 mg kg^{-1} to 124.9 mg kg^{-1}.

Clear differences in sensitivity between soil invertebrate groups are found for some of the pesticides, as illustrated by the species sensitivity ranking (where the rank indicates the position in the SSD curve, with 1 = the most sensitive species or group) (SSD curves are given in Frampton et al., 2006). For the fungicide carbendazim, arthropods are less sensitive (rank 8-10) than oligochaete worms (rank 1-4) and nematodes (rank 5). The opposite is true for the insecticides gamma-HCH (arthropods rank 1-3; oligochaetes rank 4-8), chlorpyrifos, diazinon, parathion (arthropods rank 1; oligochaetes rank 2-7) and dimethoate (arthropods rank 1-9; oligochaetes rank 10-14). Collembola are consistently more sensitive than oligochaetes to the mixed-action pesticide pentachlorophenol (Collembola rank 1-3; oligochaetes rank 4-9). The sensitivity differences between arthropods and oligochaete worms suggest that these soil invertebrates might exhibit different sensitivity distributions. However, based on a minimum of five data, too few data would be available to construct separate SSDs for these groups.

Table 1. Summary of lower-tier (laboratory) and higher-tier effects data for those pesticides that have at least five data sets for calculation of HC_5 and also at least one higher-tier data set for comparison. N=number of data sets. The lower and upper estimates of the HC_5 are the lower (10%) and upper (90%) confidence limits respectively. Note that HC_5 estimates for chlorpyrifos and parathion are given both with and without Collembola data included.

Pesticide	Laboratory (SSD) data					Higher-tier data			
			HC_5 (mg kg^{-1})						
	N	Taxonomic composition	Lower	Median	Upper	N	Taxonomic composition	NOEC (mg kg^{-1})	Best LOEC estimate (mg kg^{-1})
Atrazine	7	3 Collembola 5 Lumbricdae	0.76	5.39	13.97	29	6 Arachnida 13 Collembola 2 Coleoptera 2 Enchytraeidae 4 Lumbricidae 2 Nematoda	0.13	0.53
Carbendazim	10	1 Diplopoda 3 Enchytraeidae 3 Lumbricidae[a] 2 Isopoda 1 Nematoda	0.04	0.75	3.90	177	1 Collembola 67 Enchytraeidae 61 Lumbricidae 48 Nematoda	Not estimable	≤ 0.24
Chlorpyrifos	7	1 Collembola 6 Lumbricidae	0.001	0.370	6.040	11	1 Arachnida 6 Collembola 1 Formicidae 3 Lumbricidae	Not estimable	≤ 0.64
	6	6 Lumbricidae	25.8	124.9	252.8				
Diazinon	5	1 Collembola 3 Lumbricidae 1 Isopoda	0	0.06	1.07	16	4 Arachnida 4 Collembola 1 Formicidae 7 Lumbricidae	Not estimable	≤ 5.97
Dimethoate	14	2 Acari 3 Collembola 1 Chilopoda 1 Coleoptera 1 Diplopoda 5 Lumbricidae 1 Isopoda	0.05	0.30	0.95	1	1 Collembola	Not estimable	≤ 0.53
gamma-HCH	8	2 Collembola 1 Enchytraeidae 4 Lumbricidae 1 Isopoda	0.01	0.27	1.66	5	1 Arachnida 1 Collembola 3 Nematoda	Not estimable	≤ 1.30
Lambda-cyhalothrin	5	1 Diplopoda 2 Lumbricidae[a] 2 Isopoda	0	0.09	0.84	4	4 Collembola	Not estimable	≤ 2
Parathion	7	1 Collembola 1 Enchytraeidae 5 Lumbricidae	0	0.27	2.77	13	4 Arachnida 4 Collembola 5 Lumbricidae	Not estimable	≤ 0.35
	6	1 Enchytraeidae 5 Lumbricidae	25.9	57.3	81.7				
PCP	9	3 Collembola 1 Enchytraeidae 4 Lumbricidae[b] 1 Nematoda	0.43	3.74	12.10	26	3 Arachnida 7 Collembola 5 Enchytraeidae 8 Lumbricidae 3 Nematoda	Not estimable	≤ 6.60
Propoxur	5	5 Lumbricidae	0.01	0.36	1.51	2	2 Lumbricidae	Not estimable	≤ 1.34

[a] includes Glossoscolecidae
[b] includes Eudrilidae

The effect of selecting only comparable data as opposed to using all available data to estimate a geometric mean acute LC_{50} varies with the pesticide. For atrazine, the LC_{50} is increased (from 64 to 101 mg kg kg^{-1}) whereas for benomyl it is reduced (from 56 to 21 mg kg^{-1}). However, the difference is within a factor of two for all the pesticides except benomyl (factor 2.7).

Estimates of HC_5 are higher when the analysis is based only on comparable data, although the differences are very small for dimethoate and pentachlorophenol. The largest differences are for copper compounds (comparable-data HC_5 = 352.5 mg kg^{-1}; all-data HC_5 = 183.3 mg kg^{-1}) and chlorpyrifos (comparable-data HC_5 = 94.9 mg kg^{-1}; all-data HC_5 = 0.37 mg kg^{-1}). Using all available data in the sensitivity distribution therefore appears to provide a risk estimate that would be at least as conservative as that based on the comparable data approach. The choice of data selection approach makes little difference to the range of uncertainty of the HC_5 estimate, except in the case where the all-data approach for chlorpyrifos includes Collembola (*F. candida*) data; in this case the all-data approach increases the uncertainty range from approximately one to more than three orders of magnitude. However, the SSD including the Collembola data is not a good fit to the lognormal distribution (does not pass the Anderson-Darling test), reflecting the limitation of including taxa with markedly different sensitivities within the same SSD.

For seven out of the eleven pesticides analysed using the SSD approach (chlorpyrifos, diazinon, dimethoate, gamma-HCH, lambda-cyhalothrin, parathion and propoxur), the difference between the geometric mean LC_{50} for *Eisenia fetida* + *E. andrei* and the median HC_5 is more than a factor of 10. The ratios of $LC_{50\ Eisenia}$ to HC_5 are 2911, 1055, 300, 281, 266, 548 and 28 respectively. The application (uncertainty) factor of 10 used to represent uncertainty in the earthworm acute mortality test (EPPO 2003) would not cover the full range of species sensitivities to these insecticides. However, the ratios for chlorpyrifos and parathion are highly dependent on whether Collembola are included in the data sets; for these pesticides, the ratios of $LC_{50\ Eisenia}$ to HC_5 based only on Lumbricidae data would be 6.0 and 2.8 respectively.

3.6 Higher-tier risk estimates

In most cases the modal effects class is Class 1 (no effect observed), although relatively high numbers of pronounced and persistent effects (Class 5) occurred when Lumbricidae and Enchytraeidae were exposed to fungicides and when Lumbricidae, Collembola and Arachnida were exposed to insecticides (Table 2). The large number of studies reviewed, together with the heterogeneity of study designs and the variety of pesticides within each mode of action places a limitation on interpretation of the summarised data. However, the reported data does indicate a general lack of higher-tier effects of fungicides on Arachnida, Nematoda and Formicidae, and a lack of herbicide effects on Lumbricidae, Nematoda and Formicidae (Table 2).

Table 2. The number of higher-tier pesticide effects data records in each of five effect classes for the six most intensively studied groups of invertebrates. The effect classes are: (1) no observed effect, (2) slight transient effect, (3) pronounced transient effect, (4) pronounced effect in a short-term study, (5) pronounced long-term effect. Mixed action pesticides are those that are not solely a fungicide, herbicide or insecticide.

Mode of action	Effect class	Soil invertebrate group					
		Lumbricidae	Collembola	Enchytraeidae	Arachnida	Nematoda	Formicidae
Fungicides	1	52	6	40	5	50	5
	2	1	2	0	0	0	0
	3	4	0	11	0	0	0
	4	11	0	0	0	0	0
	5	38	1	20	0	0	0
Herbicides	1	12	11	0	6	2	4
	2	0	0	1	0	0	0
	3	0	0	1	3	0	0
	4	0	2	0	0	0	0
	5	0	7	2	4	0	0
Insecticides	1	32	26	7	24	1	11
	2	9	13	0	4	2	0
	3	20	0	0	2	0	0
	4	15	17	1	4	1	4
	5	16	11	1	10	1	0
Mixed action	1	7	7	4	2	4	0
	2	2	0	0	3	0	0
	3	4	4	1	3	0	0
	4	4	2	1	2	0	1
	5	6	5	3	3	1	0

3.6.1 Comparison of field effects concentrations with laboratory-derived HC_5 estimates

In aquatic ecotoxicology, for some pesticides it has been possible to validate HC_5 estimates obtained from mesocosm studies by comparing these with estimates of the higher-tier NOEC or LOEC (Maltby et al. 2005). However, for soil invertebrates the higher-tier NOEC cannot be estimated for nine of the ten pesticides, as effects occurred at the lowest concentrations tested (Table 1).

The median HC_5 exceeds both the higher-tier NOEC and LOEC estimates for atrazine whereas for most pesticides it lies below the lowest concentration at which higher-tier effects have been observed (Table 1). However, when the uncertainty (90% confidence interval) of the HC_5 estimate is taken into account, there are no cases where the HC_5 could be considered clearly protective for soil invertebrates exposed to any of these eleven pesticides. For atrazine, pronounced long-term effects occur in higher-tier studies around 0.5 mg kg^{-1}, which is below the lower confidence limit of the HC_5 (0.76 mg kg^{-1}).

The taxonomic composition of the data differs both between study tiers and pesticides (Table 1). For carbendazim, Collembola are represented only in one higher-tier data set whilst for gamma-HCH Collembola are the only group common to both laboratory and higher-tier analyses. The data for pentachlorophenol has the best taxonomic similarity between the laboratory and higher-tier studies, with four taxonomic groups common to both tiers (Table 1).

3.6.2 Comparison of field effects concentrations with other lower-tier effects estimates

A number of pesticides have relatively good higher-tier effects data but insufficient lower-tier data for comparison based on HC_5 estimates. Of these, the most higher-tier effects data available (> 25 data sets) is for bendiocarb, benomyl, carbaryl, carbofuran, diflubenzuron, halofenozide, imidacloprid and phorate (Jänsch et al., 2006). For these pesticides, estimates of higher-tier NOEC and LOEC are not possible, as slight effects (Classes 2-3) or pronounced effects (Classes 4-5) occurred at the lowest concentrations that have been tested. With the exception of carbofuran, for these pesticides higher-tier effects have been observed at lower concentrations than those that would be predicted from the most sensitive available lower-tier toxicity data, assuming that no application (uncertainty) factor is used.

3.6.3 Comparison of TME and field studies

In the case of carbendazim, enough data is available for a comparison between effects observed in terrestrial model ecosystems (TME) and field studies for Lumbricidae, Enchytraeidae, Nematoda and Collembola. The pattern of effects in both the TME and field studies appears similar, although pronounced long-term (Class 5) effects on oligochaetes have occurred at lower concentrations in the TME studies (Jänsch et al., 2006).

3.7 Case studies illustrating the use of distribution-based and deterministic risk estimates

An objective of WEBFRAM-5 is to illustrate, with one or more case studies, how distribution-based estimates of risk might be used in the existing risk assessment scheme for soil invertebrates, compared with the current deterministic approach. Of the pesticides that have adequate data available at both lower and higher tiers of testing (Table 1), an example of an herbicide (atrazine), a fungicide (carbendazim) and an insecticide (chlorpyrifos) were selected for case studies. The case studies are: (1) A post-emergence application of atrazine to maize at 1.5 kg ha^{-1} once every three years on 1 May; (2) Two applications of carbendazim to winter cereals, each at 0.25 kg ha^{-1}, repeated annually. (3) Two application scenarios for chlorpyrifos: (A): A summer application to winter cereals at 0.48 kg ha^{-1} on 22 June, assuming 50% crop interception; (B): A winter application to winter cereals at 0.72 kg ha^{-1} on 1 February, assuming zero crop interception. For both scenarios, applications are repeated annually. Further details of these case studies are given in the full project report.

Appropriate exposure data for the case studies was obtained using a simple model to obtain relevant predicted environmental concentrations (PEC). The model provides deterministic estimates of the PEC for illustrative purposes only and should not be interpreted as an indication of current regulatory guidance (EPPO 2003). A full explanation of the model is given in Appendix 3 of the full project report. Effects concentration data for the case studies are taken as the most sensitive of the reliable data available for each endpoint. Both effects and exposure data for these case studies can be used as example data to support the internet risk assessment tools developed from this project.

Distribution-based approaches for assessing pesticide risks to soil invertebrates are not currently required within the EU risk assessment scheme (EPPO 2003). Moreover, this project questions the applicability of soil invertebrate SSD for the current risk assessment scheme (Frampton et al., 2006; Jänsch et al., 2006). To ensure that stakeholders can clearly distinguish between the accepted and exploratory risk assessment procedures, we propose that the WEBFRAM internet module for soil invertebrates should in the first instance (1) focus on effects data, and (2) carry a clear statement that the SSD approach is illustrative, not required. Accordingly, the case studies summarized here would only be included in the WEBFRAM internet module for soil invertebrates if the relevance of SSD can be clarified. In addition to addressing the issues of data availability, it would need to be demonstrated how the application of SSD within the current risk assessment compartments for oligochaetes and arthropods would address the overall protection goal (populations of all soil invertebrates; EPPO 2003).

The risk assessment examples for atrazine, carbendazim and chlorpyrifos represent the best comparisons of deterministic and HC_5 approaches for predicting risk to soil invertebrates that can be given at present using empirical data available in the public domain. The key points that emerge from these examples are:

- Availability of effects data to illustrate the case studies is poor, especially for Collembola.
- The HC_5 examined here do not reduce the predicted risk compared to the deterministic approach; however, HC_5 with confidence limits do indicate the certainty of the risk prediction, whereas the certainty of the deterministic approach is unknown.
- None of the HC_5 in these examples could be convincingly validated due to either the HC_5 itself being based on a small and/or unrepresentative data set, or there being insufficient higher-tier effects information.
- It is important to be clear about the protection goal, as HC_5 based on data for one taxonomic group might not protect other soil invertebrate groups. This is illustrated by the chlorpyrifos data in particular and also implicit in the assumptions underpinning the SSD concept.

4. Discussion

4.1 Data availability

Lower-tier pesticide effects data can be generated relatively quickly under standardised conditions, but the availability of higher-tier data is more problematic because field studies are expensive, less easy to control and their success less predictable. Generally, higher-tier studies with soil invertebrates have been carried out with relatively few pesticide concentrations (often only the recommended application rate is used), meaning that acceptable concentrations in the field (e.g., a community NOEC) cannot be determined, or can only be roughly approximated. The relatively low number of field studies carried out combined with the low number of concentrations per study means that even for widely-used broad-spectrum pesticides (e.g., dimethoate), the available effects data is very limited. The comparison of carbendazim effects in terrestrial model ecosystem and field studies shows that TME studies can, at least for certain pesticides and taxonomic groups, predict effects in the field. TME studies could be a potential solution to the difficulty of obtaining higher-tier data for soil invertebrates, provided that their predicative capability is verified for a wider range of pesticides and species.

Another aspect of data availability is the question of whether agrochemicals industry-owned data would significantly improve the existing data set. This is difficult to answer, as it is unclear at present how much relevant data exists (Appendix 1). For present purposes, industry data is excluded from the database because (1) the industry has not reached clear agreement on whether to provide any data (for reasons of commercial confidentiality); (2) it would be difficult to ensure that the data is not biased if provided on an ad hoc basis; (3) coding data to ensure confidentiality is incompatible with the aim of maintaining transparency in the risk assessment (real and hypothetical data could not be distinguished in peer review); (4) the aim of the WEBFRAM projects is to provide an impartial and objective comparison of deterministic and probabilistic risk assessment approaches but a drive to obtain more data might be interpreted as specifically seeking to support the probabilistic approach (which is more data dependent). To enable industry stakeholders to receive an impartial message concerning the relative strengths and weaknesses of the deterministic and probabilistic approaches, we propose not to request industry-owned data unless it is voluntarily donated and can be shown to be unbiased. Industry stakeholders would have an opportunity to run risk assessments such as HC_5 calculations in the internet-based risk assessment module without prejudice.

4.2 Confounding of sensitivity with exposure and other variables

When reviewing a species sensitivity distribution it is important to bear in mind that the distribution reflects not only the variation in sensitivity but also variation in exposure and, perhaps, other sources of variability as a result of differences between tests. Variation in sensitivity is most likely to be confounded with variation in exposure when effects are compared for test organisms with different test conditions (for example, *Eisenia fetida* in a standard earthworm reproduction test compared with *Folsomia candida* in a standard springtail reproduction test). Such variability can be reduced to a certain extent by using only comparable data when deriving species sensitivity distributions but the very principle of SSD – that it contains different species – means that some variability due to the unique test conditions required for each species cannot be excluded.

When soil invertebrate HC_5 estimates based on all available data are compared to those calculated using only comparable data, the difference is usually within a factor of two. In terms of HC_5 estimates, the all-data approach provides at least as conservative a risk estimate as the comparable-data approach. Arguably, the inclusion of all data could be preferable, as the test conditions might be more relevant to field conditions (but such an assumption would need to be justified by reviewing the input data first).

4.3 Data reliability

At present it is difficult to establish the ecotoxicological implications of the variation in data reliability, due to confounding of reliability with other variables. Sensitive species or particular test scenarios, for instance, tend to be over-represented in certain reliability classes; this could lead to an apparent relationship between sensitivity and data reliability that is in fact not directly related to reliability. In the case of copper LC_{50} data, for example, Class 3 organisms were all arthropods whereas Class 4 invertebrates are all Enchytraeidae whilst Class 2 included earthworms (which are highly sensitive to Cu). However, the reliability of lower-tier data assessed according to the classification of Klimisch et al. (1997) mainly reflects formal aspects of study reporting or concurrence with guidelines rather than describing the quality or relevance of the biological information. A strict reliability classification based on biological relevance would probably exclude most of the data, which is dominated by species of Lumbricidae, Enchytraeidae and Collembola that are not typical of soil invertebrate communities in agroecosystems (e.g., Jänsch et al., 2005). The reliability classification sensu Klimisch et al. (1997) does, however, highlight the need for improved standardisation and reporting of studies which is currently inconsistent across taxonomic groups.

4.4 Acute-chronic ratios

Similar comparisons of acute-chronic ratios to those reported here have been made in aquatic studies, for example Länge et al. (1998) reported an acute-chronic ratio range for pesticides from 1.33 to 180 with a median 9.07. However, it should be noted that the acute-chronic ratio can be highly dependent on the soil type (e.g., Lock et al. 2002 observed a range of 6 to 40 when exposing *Folsomia candida* to gamma-HCH (lindane) in three soils).

The acute-chronic ratio could in theory be used to derive an extrapolation factor for the HC_5 estimates, which currently are based only on acute mortality data, to enable HC_5 predictions to take the more sensitive chronic endpoint into account. If it is assumed that the chronic mortality test is at least twice as sensitive as the acute test (which the acute-chronic ratio suggests would be true in 85% of cases), a general estimate of the chronic HC_5 could be obtained by dividing the acute mortality data used in the HC_5 calculation by two. The adjusted HC_5 would not alter the conclusions for atrazine, carbendazim or chlorpyrifos (observed higher-tier effects concentrations would still lie within the 90% confidence interval of the revised HC_5). For dimethoate, gamma-HCH and pentachlorophenol (PCP) however, the revised HC_5 might become protective, as all observed higher-tier effects concentrations would be above the upper 95% confidence limit of the revised HC_5. Whether the revised HC_5 incorporating an extrapolation factor actually is protective for these insecticides cannot be established because effects occurred at the lowest concentrations tested in the higher-tier studies.

The analysis of the acute-chronic ratio reported here is relatively crude, being based on all pesticides, due to a lack of paired acute and chronic data sets for different toxic modes of action. Although the extrapolation factor of two may be appropriate in 85% of cases, clear examples of high acute-chronic ratios exist (e.g., Lock et al. 2002), suggesting that it is inappropriate to calculate a 'general' extrapolation factor across all toxic modes of action. The limitations of the data should also be borne in mind, for example many of the test results were obtained using OECD soil. Paired acute and chronic data obtained from a wider range of test conditions should ideally be used to test the validity of an extrapolation factor for the HC_5. An overriding problem at present, however, is that HC_5 estimates cannot be validated with higher-tier data.

4.5 Implications for regulatory risk assessment

According to the EU Terrestrial Guidance Document on Ecotoxicology (EC 2002a), the aim of the ecological risk assessment for effects of pesticides on the soil environment is the protection of populations of non-target organisms, although only earthworm tests are strictly required as part of the current pesticide risk assessment process for soil invertebrates in the European Union. Our results raise the question of whether information on earthworms alone is sufficient to protect soil invertebrate communities. The work that has investigated the relevance of earthworm tests has focused mainly on extrapolation from one earthworm species to another (e.g., Heimbach 1988) or the use of laboratory earthworm data to predict risks to earthworm populations in the field (e.g., Heimbach 1998; Jones & Hart 1998), rather than using earthworms to predict effects on a wider range of soil fauna. Trigger values for earthworm tests are mainly derived from comparisons involving different tiers of earthworm testing rather than comparisons involving other soil invertebrates.

The application (uncertainty) factor of 10 used to represent uncertainty in the earthworm acute mortality test would not cover the full range of species sensitivities to three insecticides. The explanation may be that for these three insecticides (chlorpyrifos, dimethoate and gamma-HCH) the most sensitive species is an arthropod. The widely different sensitivities to pesticides of arthropods and non-arthropod invertebrates have also been observed in aquatic studies (van den Brink et al. 2002). Of the comparisons that were possible with higher-tier data, earthworms were found to be the most sensitive soil invertebrates to benomyl, carbaryl, imidacloprid and parathion whereas Collembola were more sensitive than earthworms to atrazine, chlorpyrifos and gamma-HCH. It is clear that both earthworms and Collembola are important for the risk assessment of pesticides and that Collembola sensitivity is not restricted to insecticides.

At present, the regulatory risk assessment for soil invertebrates does not support the use of a distribution-based approach to risk assessment because only one species (*Eisenia fetida*) is used routinely in lower-tier tests, whereas most higher-tier data are not amenable to analysis using the SSD approach. To use a species sensitivity distribution based approach, either the species data must come from an external data set (in aquatic ecotoxicology the US EPA AQUIRE database is an example) or it must be generated within the risk assessment procedure. The latter option would require the additional testing of a minimum of five species in order to meet our minimum number of six input data for SSD analysis, however this number is arbitrary and further work would be needed to clarify the most appropriate number of input data for soil invertebrates. Even if there were a consensus of agreement on which additional species should be tested, the development of new standard methods and validated guidelines would take more than a decade. The use of standard single-species test methods also has the potential disadvantage of fixing the taxonomic composition of risk assessment data (which even in the open literature is clearly biased towards the standard test species). This conflicts with the assumptions of SSD that the species are randomly selected and originate from the pool of relevant species in the community to be protected (Forbes & Calow 2002). Higher-tier multi-species approaches such as TMEs could provide more flexibility to vary the taxonomic composition of the data whilst at the same time minimising differences between species in the test conditions.

The species sensitivity distributions calculated in this project are based on acute mortality LC_{50} data because insufficient chronic NOEC data were available. The HC_5 are therefore not based on the most sensitive lower-tier endpoint and consequently their relevance in a tiered risk assessment pathway is unclear. (An extrapolation factor could be applied to the HC_5 to take into account the more sensitive chronic mortality endpoint but, as discussed above, such an approach cannot be validated at present). A more promising direction for the development of species sensitivity distributions in soil invertebrate risk assessment could be to use terrestrial model ecosystems to generate sensitivity distributions. This would have the advantages that individual test guidelines per species would not have to be developed and that test conditions for the different species would be both more realistic and less variable. This approach, with SSD employed at a semi-field tier, is broadly comparable with the use of SSD in aquatic risk assessment. However, terrestrial TME studies would need to (1) incorporate a concentration-response design to allow the derivation of sensitivity distributions, (2) be evaluated for a wider range of pesticide modes of action (currently only carbendazim results are available), and (3) be evaluated for a wider range of soil invertebrates (currently only Enchytraeidae, Lumbricidae and Nematoda data are available, plus a single Collembola data set).

4.6 Uncertainty in higher-tier pesticide effects: Species Effect Distributions

In aquatic ecotoxicology, concentration-response information may be available for all tiers of testing and it may be possible to use the SSD and HC_5 approach to estimate pesticide risks at higher as well as lower tiers (van den Brink et al. 2002). In the SSD approach, the concentration is variable whereas the effect (i.e., HC_5) is effectively fixed. Higher-tier studies with soil invertebrates are, however, generally carried out with a single pesticide concentration (e.g., the recommended field application rate), in which case the concentration is effectively fixed and the effects (on different species) variable. Such single-concentration data do not permit the construction of higher-tier SSD (examples of higher-tier concentration-response data do exist for earthworms but due to commercial confidentiality the data is not available for the current project).

For the SSD approach, the distribution used (e.g., log-normal) has indefinite upper bounds, as the range of possible effect concentrations is unlimited. With single-exposure-concentration data, it is possible to make a distribution of the effects at a given field concentration. However, higher-tier effects are usually expressed as proportions (e.g., percentage effects compared to a control), meaning that the effects data distribution would have defined upper limits. An appropriate distribution to describe such data would be a beta-distribution. In theory, using a beta distribution to describe the distribution of effects would enable uncertainty to be included in the risk assessment and permit the magnitude of effect for a given proportion of species (e.g., 95%) to be estimated with a given degree of certainty (e.g., 95%). Such a distribution is termed a Species Effect Distribution. Although conceptually different to SSD, the Species Effect Distribution is based on similar assumptions, including that the chosen distribution (in this case beta) is the most appropriate and that the input data are based on an appropriate (i.e., random) subset of the of species.

An example scenario for interpreting the risk estimate given by a Species Effect Distribution might concern the acceptability of pesticide effects on earthworm populations. Suppose, based on ecological considerations, we assume that a reduction in the number of earthworms (measured in a field earthworm study) of more than 10% is critical in any field. With the proposed approach we could calculate the maximum effect in any field (with a 95% certainty) for the proposed application rate (i.e., that tested in the earthworm study). The estimated effect (with a 95% certainty) could then be compared to a 10% trigger value.

The Species Effect Distribution concept has not been evaluated in detail. A preliminary demonstration of the possible use of this approach is given in Appendix 2 of the full project report.

4.7 Quality requirements of higher tier data

A problem encountered when examining higher-tier pesticide effects studies is that often there was insufficient information given to enable the reported effects to be interpreted with confidence. For example, the estimation of soil concentrations of pesticides is difficult if only nominal application rates are available without information about the soil characteristics or factors that could affect pesticide deposition. Another problem encountered is that often the taxonomic status of the test invertebrates was not reported in detail. If broad taxonomic groups are reported such as ‘earthworms’, it is not possible to establish the degree of taxonomic overlap of data sets, which restricts the suitability of the data for use in sensitivity distributions. The interpretation of pesticide effects could be improved and perhaps uncertainty in the estimation of effects concentrations reduced if relevant information is more clearly reported in higher-tier studies, for example as recommended by Römbke et al. (2002).

4.8 Internet-based WEBFRAM-5 pesticide risk assessment tools for soil invertebrates

An objective shared by each of the WEBFRAM risk assessment projects is that, for each project, an internet-based risk assessment module will be made available with which stakeholders could compare deterministic and probabilistic risk assessment approaches, either using worked examples based on archived data (in the case of soil invertebrates, for atrazine, carbendazim and chlorpyrifos) or using their own input data. It is anticipated that the risk assessment modules for each project will share certain components including a species sensitivity distribution model; the structure, functioning and design of the modules will be harmonised to provide an integrated risk assessment package tailored to assist risk assessments as carried out under the current European pesticides regulatory scheme.

The internet-based risk assessment modules will have several functions: to assist stakeholders to compare deterministic and probabilistic risk assessment approaches; to provide risk assessment tools that stakeholders may use independently (without risk to commercially sensitive information); and to provide an educational resource so that stakeholders may learn about appropriate risk assessment procedures. The education function of the modules is particularly important, as the use of probabilistic approaches is controversial and probabilistic risk assessments may be complex and difficult to interpret and communicate clearly without adequate training.

The option for stakeholders to independently run their own data analyses has the potential advantages that:

- The generic SSD model is independent of the pesticide and could therefore be applied to new active substances
- Industry-owned data could be analysed confidentially if desired

Opportunities to develop case studies to illustrate the use of deterministic and probabilistic risk assessments for soil invertebrates are clearly limited by the availability of data. Species sensitivity distributions for soil invertebrates could only be derived for eleven pesticide substances, which in all cases were based on acute effects concentrations (LC_{50}) due to a lack of chronic (e.g., NOEC) values; in one case the resulting sensitivity distribution would be clearly biased as it only contains earthworm data and in two cases the laboratory-derived HC_5 estimates lacked higher-tier data for comparison.

Despite the obvious limitations of these data they provide valuable information that should be communicated in the web-based risk assessment module for soil invertebrates: Examples from WEBFRAM-5 can be used to illustrate:

- limitations of the SSD and HC_5 concepts;

- the importance of clear communication (for example, it is essential that the taxonomic basis of a sensitivity distribution is clearly reported, as the example for chlorpyrifos and earthworms illustrates;
- the need for more empirical data to support species sensitivity distribution analyses.

The internet tools arising from this project are currently being finalised in collaboration with DEFRA Central Science Laboratory and the Cadmus Group for uploading to the WEBFRAM website. Stakeholders will have the opportunity of graphically displaying estimates of HC_5 with their confidence intervals. PEC estimates derived from the exposure model case studies (Appendix 3 in the full project report) could be included but we propose that, to avoid possible confusion to stakeholders (between required and exploratory risk assessment procedures), the initial version of the risk assessment module should focus on effects data. The output options available to stakeholders will include both deterministic and distribution-based risk estimates for comparison.

5. Implications of the findings

This is the first project to evaluate in detail the possibility of applying probabilistic approaches to the pesticide risk assessment for soil invertebrates and is based on the largest available number of pesticide effects data sets for soil invertebrates. A key finding is that despite the large number of pesticide data sets collected, the data are biased towards a relatively small proportion of the chemicals and invertebrate taxa. This limits the analyses that can be carried out.

Implications of the project's findings are:

- Lower-tier risk estimates generally cannot be validated against higher-tier data as higher-tier effects thresholds are mostly unavailable.
- HC_5 for soil invertebrates do not appear protective, but can only be based on acute mortality data, which is not the most sensitive endpoint.
- In theory, chronic sensitivity could be taken into account by applying an extrapolation factor to the HC_5 based on the acute-chronic ratio. However, such an extrapolation cannot be validated using existing higher-tier data.
- Care is needed in the selection of data for inclusion in SSD and HC_5 calculations because in some cases the addition or deletion of one species (as in the example with chlorpyrifos) can alter the predicted hazardous concentration by several orders of magnitude.
- Arthropods and oligochaetes together provide valuable information on pesticide risks that is not captured in all cases by monitoring oligochaetes alone. Inclusion of arthropods in the risk assessment could improve risk predictions.
- The data available from independent research studies seems to mirror the regulatory risk assessment data, being dominated by the standard test species and their relatives. This situation seems likely to continue unless research incentives can stimulate the study of more ecologically relevant species.
- Accordingly, a key assumption of the SSD approach, that the species data is representative of the community to be protected, cannot be met for most pesticides.
- The WEBFRAM projects provide an important arena for raising awareness of these issues. Knowledge of the data gaps and limitations is a fundamental prerequisite for the efficient targeting of future research resources.

6. Recommendations and future work

The findings of this work suggest that some improvements could be made in the way pesticide effects studies are carried out, to clarify interpretation of effects and prediction of risk to soil invertebrates. In particular, there is a need to:

(1) investigate pesticide risks for species other than the standard test species;

(2) broaden the range of species sensitivity data to include more species and chemicals; and

(3) address the limitations of current studies, particularly with regard to:

(i) testing appropriate concentration ranges in higher-tier studies (to enable threshold effects concentrations such as community NOEC or LOEC to be identified or, preferably, concentration-response relationships to be investigated);

(ii) improving (and where possible harmonising) the reporting of studies (for example reporting exposure concentrations and identifying soil invertebrates to species where possible);

(iii) evaluating the potential of TMEs for risk assessment with a range of chemicals representing different toxic modes of action (as this tier of study appears the most likely to efficiently generate the type of data required to support distribution-based risk assessments)

7. Knowledge transfer

The following research papers which report aspects of the project work and acknowledge DEFRA funding are in press: Frampton et al. (2006), Jänsch et al. (2006).

The following research paper which reports aspects of the project work and acknowledges DEFRA funding has been accepted for publication: Scott-Fordsmand, J.J. & Damgaard, C. Uncertainty analysis of single concentration exposure data.

Acknowledgements

The authors acknowledge the contribution of the German BVB (Technical Committee: Biological Assessment of Soils) and the German UBA (Federal Environment Agency) for providing soil ecotoxicological data, including the SoilValue database. We also thank Willem Roelofs (DEFRA Central Science Laboratory, CSL) and Jim Siegmund (Cadmus Group) for developing and uploading the internet risk assessment models (www.webfram.com).

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Badejo, M.A. & van Straalen, N.M. (1992). Effects of atrazine on growth and reproduction of *Orchesella cincta* (Collembola). Pedobiologia 26, 221-230.

Barrett, K.L., Grandy, N.,Harrison, E.G., Hassan, S. and Oomen, P. (1994). Guidance Document on Regulatory Testing Procedures for Pesticides with Non-target Arthropods. Society of Environmental Toxicology and Chemistry (SETAC-Europe), Brussels, Belgium. 51pp.

Aldenberg, T. & Slob, W. (1993). Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. Ecotoxicology and Environmental Safety 25, 48-63.

Brock, T.C.M., Lahr, J. & van den Brink, P.J. (2000). Ecological risk of pesticides in freshwater ecosystems. Part 1. Herbicides. Alterra Report 88. Alterra Green World Research, Wageningen, The Netherlands.

Campbell, P.J., Arnold, D.J.S., Brock, T.C.M., Grandy, N.J., Heger, W., Heimbach, F., Maund, S.J. & Streloke, M. (1999). Guidance Document on Higher-tier Aquatic Risk Assessment for Pesticides (HARAP). SETAC-Europe, Brussels, Belgium.

EC (European Commission) (2002a). Draft working document. Guidance document on terrestrial ecotoxicology under Council Directive 91/414/EEC. SANCO/10329/2002 rev 2 final. 39pp.

EC (European Commission) (2002b). Working document. Guidance document on aquatic ecotoxicology in the context of the Directive 91/414/EEC. SANCO/3268/2001 rev 4 final. 62pp.

EPPO (European Plant Protection Organisation) (2003). EPPO Standards. Environmental risk assessment scheme for plant protection products. Chapter 8: Soil organisms and functions. OEPP/EPPO Bulletin 33, 195-209.

Forbes, V.E. & Calow, P. (2002). Species sensitivity distributions revisited: a critical appraisal. Human and Ecological Risk Assessment 8, 473-492.

Frampton, G.K., Jänsch, S., Scott-Fordsmand, J.J., Römbke, J. & van den Brink, P.J. (2006). Effects of pesticides on soil invertebrates in laboratory studies: A review and analysis using species sensitivity distributions. Environmental Toxicology and Chemistry. In press.

Hart, A. (ed.) (2001). EUPRA: Probabilistic Risk assessment in Europe: implementation and research needs. Central Science Laboratory, Sand Hutton,York, UK. 120pp.

Heimbach, F. (1998). Comparison of the sensitivity of an earthworm (Eisenia fetida) reproduction test and a standardized field test on grass land. In: Sheppard, S.C., Bembridge, J.D., Holmstrup, M. & Posthuma, L. (eds) Advances in Earthworm Ecotoxicology. SETAC Press, Boca Raton, USA. pp. 235-245.

Hopkin, S.P. (1993). Ecological implications of '95% protection levels' for metals in soils. Oikos 66, 137-141.

ISO (International Organisation for Standardisation) (1998). Soil quality – Effects of pollutants on earthworms (*Eisenia fetida*). Part 2: Determination of effects on reproduction. ISO 11268-2:1998. ISO, Geneva, Switzerland.

ISO (International Organisation for Standardisation) (1999a). Soil quality – Inhibition of reproduction of Collembola (*Folsomia candida*) by soil pollutants. ISO 11267:1999. ISO, Geneva, Switzerland.

ISO (International Organisation for Standardisation) (1999b). Soil quality – Effects of pollutants on earthworms. Part 3: Guidance on the determination of effects in field situations. ISO 11268-3:1999. ISO, Geneva, Switzerland.

Jänsch, S., Amorim, M.J. & Römbke, J. (2005). Identification of the ecological requirements of important terrestrial ecotoxicological test species. Environmental Reviews 13, 51-83.

Jänsch, S., Frampton, G.K., Römbke, J., van den Brink, P.J. & Scott-Fordsmand, J.J. (2006). Effects of pesticides on soil invertebrates in model ecosystem and field studies: A review and comparison with laboratory toxicity data. Environmental Toxicology and Chemistry. In press.

Jones, A. & Hart, A.D.M. (1998). Comparison of laboratory toxicity tests for pesticides with field effects on earthworm populations: a review. In: Sheppard, S.C., Bembridge, J.D., Holmstrup, M. & Posthuma, L. (eds) Advances in Earthworm Ecotoxicology. SETAC Press, Boca Raton, USA. pp. 247-267.

Klimisch, H.-J., Andreae, M. & Tillmann, U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology 25, 1-5.

Kooijman, S.A.L.M. (1987). A safety factor for LC_{50} values allowing for differences in sensitivity among species. Water Research 21, 269-276.

Länge, R., Hutchinson, T.H., Scholz, N. & Solbé, J. (1998). Analysis of the ECETOC Aquatic Toxicity (EAT) database. II – Comparison of acute to chronic ratios for various aquatic organisms and chemical substances. Chemosphere 36, 115-127.

Lock, K., de Schamphelaere, K.A.C. & Janssen, C.R. (2002). The effect of lindane on terrestrial invertebrates. Archives of Environmental Contamination and Toxicology 42, 217-221.

Maltby, L., N. Blake, T.C.M. Brock and P.J. Van den Brink. (2005). Insecticide species sensitivity distributions: the importance of test species selection and relevance to aquatic ecosystems. Environmental Toxicology and Chemistry 24, 379-388.

Römbke, J., Labes, G., Woiwode, J. (2002): Ansätze für Strategien zur Bewertung des Bodens als Lebensraum für Bodenorganismen. Ergebnisse eines internationalen BMU-Fachgesprächs: Empfehlungen und Forschungsbedarf. Bodenschutz 2: 62-69.

van den Brink, P.J., Brock, T.C.M. & Posthuma, L. (2002). The value of the species sensitivity distribution concept for predicting field effects; (Non-)confirmation of the concept using semi-field experiments. In Posthuma, L., Suter, G.W.I. & Traas, T.P. (eds) Species Sensitivity Distributions in Ecotoxiology. Lewis Publishers: Boca Raton, Florida, USA. pp. 155-193.

van Straalen, N. M. & Denneman, C. A. J. (1989). Ecotoxicological evaluation of soil quality criteria. Ecotoxicology and Environmental Safety 18, 241-251.

van Vlaardingen, P., Traas, T.P. & Aldenberg, T. (2003). EtX-2000. Normal distribution based hazardous concentration and potentially affected fraction. Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Bilthoven, The Netherlands.

Wagner, C. & Løkke, H. (1991). Estimation of ecotoxicological protection levels from NOEC toxicity data. Water Research 25, 1237-1242.

Wheeler, J.R., Grist, E.P.M., Leung, K. M. Y., Morritt, D. & Crane, M. (2002). Species sensitivity distributions: data and model choice. Marine Pollution Bulletin 45, 192-202.