Grant Agreement No: 245163
Project start date: June 2010
Project end date: November 2013

ACROPOLIS

Aggregate and Cumulative Risk Of Pesticides: an On-Line Integrated Strategy

SEVENTH FRAMEWORK PROGRAMME

Deliverable 5.6 A scientific paper on new approaches to uncertainty analyses for use in aggregate and cumulative risk assessment of pesticides

The Food and Environment Research Agency (Fera), Sand Hutton, York, UK
Contact person: Marc C. Kennedy, email: Marc.Kennedy@fera.gsi.gov.uk

Project team Acropolis: Marc C. Kennedy (Fera), Hilko van der Voet (Biometris), Vicki Roelofs (Fera), Willem Roelofs (Fera), C. Richard Glass (Fera), Paul W. Goedhart (Biometris), Waldo de Boer (Biometris), Johannes Kruisselbrink (Biometris), Gerie W.A.M. van der Heijden (Biometris), Andy D.M. Hart (Fera)

Table of contents

New approaches to uncertainty analysis for use in aggregate and cumulative risk assessment of pesticides

Marc C. Kennedy[a*], Hilko van der Voet[b], Vicki Roelofs[a], Willem Roelofs[a], C. Richard Glass[a], Paul W. Goedhart[b], Waldo de Boer[b], Johannes Kruisselbrink[b], Gerie W. A. M. van der Heijden[b], Andy D. M. Hart[a]

[a] *The Food and Environment Research Agency (Fera), Sand Hutton, York, UK*
[b] *Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands*

Abstract

Risk assessments for human exposures to plant protection products (PPPs) have traditionally focussed on single routes of exposure and single compounds. Extensions to estimate aggregate (multi-source) and cumulative (multi-compound) exposure from PPPs presents many new challenges and additional uncertainties that should be addressed as part of risk analysis and decision-making. In this paper we describe a general approach for identifying, classifying and quantifying the relevant uncertainties and variabilities. We describe the implementation of uncertainty methods within the MCRA software developed as part of the EU-funded ACROPOLIS project. In addition, several potential new methods are introduced and illustrated for dietary and non-dietary exposures to the triazole class of compounds. Specific aspects covered are: uncertainties in relative potency factors, derived from dose-response models and used to weight multiple compounds; linking variability and uncertainty generated from an external model for bystander exposure with MCRA dietary exposure assessments; inferring correlations from multiple-compound residue sample data; uncertainties about extreme residues; combining pesticide usage survey information with residue monitoring data, to overcome uncertainty due to non-detect samples. A general discussion of alternative approaches to treat uncertainty, either quantitatively or qualitatively, is included.

Keywords: common assessment group, probabilistic modelling, occupational exposure, consumer exposure, multiple pesticides

Running title: Uncertainty in aggregate & cumulative pesticide exposure

* Corresponding author. Marc C. Kennedy, The Food and Environment Research Agency (Fera), 12G39, Sand Hutton, York, YO41 1LZ. Tel: +44 1904 462175;
E-mail address: marc.kennedy@fera.gsi.gov.uk

1. *Aggregate and cumulative risk assessment of pesticides*

The use of plant protection products (PPPs), containing one or more pesticides, is regulated to ensure any risk to human health is minimised. Exposure is possible during application of PPPs or through contact with treated crops, including dietary consumption. A wide variety of products are licensed for use as PPPs and new pesticide compounds are being developed all the time. It is therefore important to consider in a cumulative risk assessment whether individuals might be exposed to multiple compounds and at what levels, given the combinations of exposure sources. Risk assessment should also consider the overall health impact of the combinations to which individuals could be exposed.

Agricultural produce is routinely sampled and tested for multiple pesticide residues, providing data as input to risk assessment tools. Traditional risk models for pesticide exposure have focussed on single sources and routes of exposure, being restricted for example to dietary exposure or occupational/residential exposure, and considering individual pesticides. This is partly due to the fact that pesticides are regulated individually and also because simpler models were considered desirable in terms of practicality and transparency. Recent advances in risk assessment methodology include extensions to take into account aggregate and cumulative exposures. For the purpose of this paper, *aggregate exposure* refers to exposure from more than one source, where potential sources could include diet, consumer products or occupational activities, for example. *Cumulative exposure* refers to exposure from multiple compounds affecting the same toxicological endpoint (referred to collectively as a cumulative assessment group or CAG). Regulation (EC) No 1107/2009 for approving new PPPs requires assessment of these effects when scientific methods accepted by the relevant authority are available. The modelling of aggregate and cumulative exposure adds many complexities. Given the large range of potential compounds, sources and routes of exposure that can occur in reality, together with their interactions, it is not practical to study these exposures in great detail. Models of the real processes are approximations which inevitably involve relatively simple assumptions. These problems are compounded by limitations in available data. Some activities leading to exposure are not routinely monitored and the collection and detection of multi-compound data is expensive and difficult. To establish trust in any model of aggregate and cumulative exposures, it is therefore essential that the uncertainties attached to the model results are quantified as fully as practically possible and, more generally, that the limitations are made completely transparent (Codex 2011, EFSA 2006, 2009, 2012).

Probabilistic models for aggregate and cumulative exposure have been developed previously. The latest version of the US SHEDS multimedia model (Zartarian *et al*, 2008) includes cumulative dietary (SHEDS-Dietary) and aggregate residential (SHEDS-Residential) modules. The general approach for dealing with variability and uncertainty is the same as that used here, although details of the implementation vary. More detailed comparisons are included in Section 3.

The focus of this paper is the framework for the treatment of uncertainty as developed in the EU project Acropolis (www.acropolis-eu.com). The objectives of Acropolis included the development of a web-based software tool (mcra.rivm.nl) to facilitate Europe-wide risk assessment for aggregate and cumulative modelling. The software includes various options for probabilistic modelling as described in Van der Voet *et al*. (in prep.), and, specifically for the aggregate models in Kennedy *et al*. (in prep.). The first focus in this paper is on the way uncertainty can be addressed in the models

developed by Acropolis using MCRA 8. The second focus is to describe several new methods for dealing with uncertainty that are still in the research stage but may be useful for more specialised analyses.

We begin by describing in Section 2 the general procedure followed to identify and prioritise the key relevant sources of uncertainty. Section 3 begins with a general discussion of the overall uncertainty model framework and introduces some alternative implementations. The main uncertainties implemented in MCRA are then discussed with illustrative examples. In Section 4, additional uncertainty methods are described. We conclude in Section 5 with discussion including suggestions for further development in the area of uncertainty modelling.

2. Identified sources of uncertainty

Uncertainty can arise from natural variability (aleatoric uncertainty) or uncertainty due to imperfect knowledge (epistemic uncertainty). Both types, for short often simply indicated by 'variability' and 'uncertainty', are considered in this paper, and it is often important to report them separately in order to make well-informed decisions and to manage risks more effectively.

The general strategy for prioritising and handling uncertainties in Acropolis involved the following:

- Project partners were invited to list those features and parameters, from within their area of expertise, that are considered to be uncertain and which could have a substantial impact on the assessment
- Considering the limited resources available, the data or other information sources available for analysis, and the likely impact of each source of uncertainty, each source was classified as M, EJ, or UT to indicate the methods employed – M represents quantitative modelling via statistical methods, EJ is quantified expert judgement and UT is the use of uncertainty tables to capture those aspects that are recognised but not quantified. These alternatives are described in Sections 3 and 4, with examples.

The results of the consultation exercise were discussed amongst project partners and a refined prioritised list was produced, leading to the main topics listed below for uncertainty analysis. The aim was to quantify as many important sources of uncertainty as possible with the resources available. Practical considerations and the defined scope of the Acropolis project also had to be respected. For example, the range of suitable modelling options is determined by the quality and quantity of data available. Similarly, expert elicitation methods require access to individuals with relevant expertise. Capturing the remaining unquantified uncertainties, together with an assessment of how those might affect the risk assessment is important for transparency and decision-making (risk management) as well as a useful pointer for future information/data collection and research.

The main sources of variability and uncertainty, considered important for aggregate and cumulative exposure assessment, are listed below under general headings:

Multivariate pesticide residues. A cumulative assessment depends on information about co-occurrence of pesticides, which may be obtained from multi-compound monitoring samples. Monitoring programs vary within Europe, as does the list of compounds registered for use within each Member State. Common problems giving rise to uncertainty include small datasets, large proportions of missing data and non-detects (censored data). In addition, monitoring is often

targeted on the most commonly consumed crops or those previously found to contain high residues of pesticides. Where a compound was not tested for, the reason could be that the compound was known to be unlikely or other reasons including limited resources. But in practice the reasons are often not available. Extrapolating from a small dataset to predict the tails of the residue distribution is important in risk assessment, but is a further source of uncertainty that is problematic even in the single compound case. New methods for uncertainties in residue distributions are covered in sections 4.1 and 4.2.

Multiple food consumption patterns. Information about dietary consumption within a population is obtained from national surveys. Random simulation of records from a survey is used to preserve correlations between foods for individual-days. For acute assessments the total number of survey days is typically large enough that uncertainty is small relative to other sources. For chronic exposure, the observed individual mean (OIM) method can be applied but uncertainty can be substantial. This is because a small number of days are available *per individual* to estimate a long term average total intake, and therefore the observed combinations of food types per individual are limited. New uncertainty methods are not presented here. Details of the existing uncertainty methods can be found in Van der Voet *et al.* (2013), Van Klaveren *et al.* (2012), and Kennedy (2010).

Aggregate (non-dietary) exposure. Non-dietary sources of exposure can arise from various activities and scenarios. Compared to the dietary source there may be many more variables to consider. For example, PPP operators and workers use a wide variety of protective clothing and practices across the EU, and compliance with advice cannot always be guaranteed. These population groups may have a higher exposure from their occupation than from their diet. In any given exposure scenario/route there are often little or no data with which to estimate model parameters, so uncertainty will typically be larger for non-dietary exposure. Examples of relevant data collection and modelling activities include Glass *et al*. (2012), BROWSE (www.browseproject.eu), and the US EPA SHEDS models (Zartarian *et al*. 2000, 2012). The aggregate model developed in Acropolis is described in Kennedy *et al*. (2013), with uncertainty methods and examples presented in Section 3.3.

Toxicology. The treatment of uncertainty in dose-response modelling, and hazard modelling more generally, is an extremely challenging problem and largely outside the scope of the Acropolis project. However, it emerged as an important issue from the prioritisation exercise. Our model combines compounds using a relative potency factor (RPF) approach. This approach to cumulative modelling assumes a common slope of dose-response function, as described for example in Bosgra *et al.* (2009), and models are fitted based on limited datasets. Specific examples of uncertainty relevant to Acropolis arise from estimation of RPF values for each compound, the assumption of parallel slopes, and the use of default extrapolation factors to extrapolate between species. Discussion of these topics and some preliminary research is included in Section 4.3.

3. Models for uncertainty analysis in MCRA

3.1 Framework and uncertainty modelling options

Monte-Carlo (MC) simulation is used to integrate the effects of various sources for both variability and uncertainty. For each source separately the variability can be represented by a parametric distribution or by a series of observed values (empirical distribution). For uncertainty, the empirical option is not relevant, as it is not meaningful to 'observe' realisations of a fixed but unknown

quantity. Each variable involved in the model can be represented in a number of ways, depending on its status:

- Known constant – single value, common across all individuals
- Unknown constant – multiple (N_u) realisations of this single unknown value are generated. In a given iteration of the uncertainty loop, a single realisation of this variable is applied to all individuals
- Variable between individuals within the population, with known distribution – N_v realisations from the variability distribution are generated, usually equal to the number of simulated individuals or individual-days
- Variable between individuals but with uncertainty about the true variability distribution – N_v realisations for each of the N_u simulated variability distributions. Each simulated variability distribution uses parameters generated from an uncertainty distribution

For this final case, second order or 2-dimensional Monte Carlo (2DMC) was adopted for its simplicity and flexibility (Burmaster and Wilson, 1996; Cullen and Frey, 1999). Crucially it allows for uncertainty and variability to be separately quantified and reported, which is important in risk management as it can help in deciding whether to reduce uncertainty (with additional data) or reduce variability (via control measures). Samples are generated according to some assumed probability distribution but the type and source of that distribution can be arbitrarily chosen. The 2DMC algorithm consists of an uncertainty loop ($1 \ldots N_u$) in which alternative realisations of uncertain quantities, e.g. parameters, are generated. Conditional on *each* simulated uncertainty value, a further set of variability realisations $1 \ldots N_v$ are generated from the conditional probability distribution of the variable inputs.

Inputs and outputs, represented as matrices of uncertainty and variability simulations, can easily be chained together even if they are generated using completely different software modules. Hence individual model components can be developed independently. The analyst is free to use whichever method is most convenient or appropriate. Dependencies between parameters can be accounted for by ensuring that features of a given individual are generated from the correct joint distribution, observing any dependency rules and labelled consistently in the various input files for that individual. Similarly, uncertain parameters should be generated from a joint distribution where appropriate and any collection of correlated inputs must be used together in each uncertainty loop. Particular examples of this modular approach relevant to Acropolis are the use of PROAST software for dose-response uncertainty (RIVM, 2013) and the Browse software which quantifies uncertainty in non-dietary exposure sources. PROAST can also be used in combination with dietary exposure assessment to quantify distributions of individual margins of exposure in the integrated probabilistic risk assessment (IPRA) method. Summaries or sensitivity analysis results of interest to a risk assessor can be derived directly from the output. Important examples are mean, standard deviation and high quantiles, relative to the population (e.g. 95th percentile exposure) and uncertainty associated with these summaries.

Aside from the practical benefits, a more philosophical question is how to interpret the probabilities represented by the samples. Frequentist and Bayesian statisticians interpret probability in a fundamentally different way but either approach may be taken in the 2DMC framework. In frequentist or classical statistics it is not meaningful to use a probability distribution to represent uncertainty about a fixed but unknown value. Distributions are used instead to represent variability

due to random sampling of data. Estimators of some feature of the underlying population, such as a fixed unknown mean value, are functions of data, and it is these estimators with their variability under repeated sampling whose distributions are of interest. Confidence intervals for many standard estimators are available in closed form, when assuming particular models. In more complex risk models bootstrapping can also be used as a general purpose method to give approximate intervals (Efron and Tibshirani, 1993). Bootstrapping works by resampling, from the original dataset, samples of the same size, but with replacement. By repeated sampling and subsequent estimation this technique can be used to study the frequentist properties of estimators using random samples of a given size. Essentially the original sample is treated as a surrogate for the total population, and the repeat sampling mimics the sampling process. Naturally, therefore, approximate means and confidence intervals are more accurate for large samples (provided they are random samples) but become highly unreliable with small samples.

From the Bayesian perspective probability quantifies a degree of belief in an event, including those associated with parameter values. Therefore any unknown quantity can be assigned a probability distribution. Bayesian statistics works by assuming prior distributions for a quantity, then using Bayes' Theorem to update beliefs – to derive the conditional probability distribution – given a sample of data. Prior information can be incorporated from carefully elicited expert opinions, although often an approximate flat prior is used instead to represent prior ignorance. By combining conditional probability distributions representing individual model components, very complex models can in principle be built. More freedom is therefore available to combine parametric distributions, opinions or other information sources such as simulations from computer models. For small sample sizes, a Bayesian model can provide more meaningful results than the classical bootstrap alternative. However, in the small sample case care should be taken in selecting a prior, as results will be more driven by this choice. In general, sensitivity to the prior choice should always be investigated.

Examples of both frequentist and Bayesian uncertainty models are illustrated in the following sections. Either can be used with the general framework described, although it is important to be clear about the different interpretations of the outputs and to provide suitable warnings about bootstrapping, or sensitivity to prior assumptions, when samples are small. Each input to the model could in theory be presented in the 2DMC matrix form, but in practice it is restricted to a few key inputs. The US SHEDS models mentioned in Section 1 incorporate the 1D or 2DMC methods to account for variability and/or uncertainty as explained in Zartarian *et al.* (2008). The variability component includes a detailed longitudinal model of activities per individual to capture realistic exposure profiles. Records are drawn from the Consolidated Human Activity Database (CHAD - McCurdy *et al.*, 2000) to represent a realistic range of activity patterns, including variations between weekday/weekend and seasonal behaviours. Uncertainty in input parameters is accounted for through the selection of probability distributions from a standard set. Sampling uncertainty is accounted for using a frequentist bootstrapping approach (Zartarian *et al.*, 2008). The detailed model is expensive in terms of CPU and memory usage. Each run of SHEDS for a population of 1000 individuals can take around an hour, and with uncertainty iterations this can become several days or weeks. Detailed activity records are not available for the EU, so the SHEDS model is not practical for that region.

Often, models include inputs for which only single point estimates are available. The uncertainty associated with these inputs and the impact on the final exposure estimates is very difficult to quantify in general and might be listed as unquantified. More details of our approach are presented in Section 4.4. While not ideal this is important for transparency and much preferable to leaving the (possibly substantial) uncertainty hidden.

Quantifying uncertainty is often technically challenging, especially when the quantity and/or quality of data are limited. Furthermore, the most appropriate approach is dependent on the data available. Therefore, when developing generic guidance for probabilistic exposure assessment for pesticides, EFSA (2012) proposed a tiered approach starting with a relatively simple 'basic' model, in which variability and some uncertainties are quantified probabilistically and other, more difficult uncertainties are addressed by means of rerunning the model with alternative deterministic assumptions, referred to as 'pessimistic' and 'optimistic'. When the results of the pessimistic model raise concern, but the optimistic model run does not, this indicates that refined modelling of one or more of the more difficult uncertainties will be required to support decision-making.

3.2 Cumulative dietary exposure

As part of the Acropolis project, EFSA's (2012) recommended approaches for basic probabilistic modelling of dietary exposure were incorporated in the MCRA software tool where possible. Some aspects of the EFSA guidance involved running the software with alternative datasets for the optimistic and pessimistic scenarios, or extrapolating input values from particular sources (Boon *et al*., 2013). The EFSA guidance also suggests a number of options for more advanced treatment of variability and uncertainty in refined assessments. See Van der Voet *et al.* (2013) for details. Some examples of variability and uncertainty assessments within a cumulative dietary model are presented in Figure 1 for the optimistic and pessimistic selections as defined in EFSA (2012).

Figure 1 here

Figure 1. Results from MCRA for cumulative dietary exposure of women (19-45) with EFSA guidance optimistic settings (left panels) and pessimistic settings (right panels). Estimates based on nominal values and p2.5 - p97.5 uncertainty bands are shown in black and red respectively in the top panels.

For our example scenario we consider acute exposure to the collection of triazoles in the CAG for the cranio-facial developmental effect. As the health effect is relevant to women of child-bearing age, dietary consumption of UK women (aged 19-45) was taken from the NDNS (Hoare *et al*., 2004). We include this example purely as an illustration of hypothetical cumulative exposure and risk assessment conclusions cannot be drawn from these results.

3.3 Aggregate (dietary plus non-dietary) exposure

Extending the example introduced in the previous section, we model an additional non-dietary exposure to Cyproconazole in the same population due to a pass of an arable crop sprayer. To allow for maximum flexibility, non-dietary exposures are calculated outside MCRA and presented using a

standard input format to represent a constant, a 1DMC or a 2DMC simulation as explained in Section 3.1. Outputs from existing exposure models such as ConsExpo (Bremmer *et al*., 2006), BROWSE, BREAM (Kennedy *et al.*, 2012), or empirical samples from databases such as EUROPOEM can all be adapted to produce the required inputs. Multiple exposure values can represent those of individuals within the population. These can be either: randomly assigned to dietary records from the whole population; randomly assigned to dietary records, provided they match particular criteria (e.g. age range, sex); or matched to a specific individual. Several examples are presented in Kennedy *et al*. (2013) demonstrating the use of probabilistic models to quantify variability of non-dietary exposure in this way. An example is also included in which a single conservative, deterministic non-dietary value is used, which is a popular first tier approach to deal with uncertainty. We now present a case that includes both variability and uncertainty in the dietary and non-dietary sources. Non-dietary dermal exposure calculated using the BREAM model. For illustration we suppose that Cyproconazole is sprayed in a nearby field. Variability is included to represent real variations in boom height, wind speed and wind direction during a spray event. Uncertainty is also included, in parameters defining the true relationship between spray concentration and bystander deposits, as those relationships are estimated using limited datasets. A 2DMC algorithm was used with 1000 variability loops and 100 uncertainty loops. The non-dietary results were aggregated with those triazoles in the CAG present in the women's diet. MCRA simulation results are shown in Figure 2 showing a comparison between the optimistic and pessimistic settings for dietary exposure.

Figure 2 here

Figure 2: MCRA uncertainty analysis of aggregate exposure of children (2-6) through dietary and arable crop spraying. Optimistic assumptions (left panel) and pessimistic assumptions (right panel) were used for dietary exposure. Uncertainties and variabilities are included in dietary and non-dietary sources here.

Uncertainty included	**P95 estimate (with 95% intervals)** µg/kg bw/day	**P99 estimate (with 95% intervals)** µg/kg bw/day
Variability only	16.31	20.38
Uncertainty in diet only	…	…
Uncertainty in diet and bystander exposure	…	…

Table 1: Comparison of estimates and uncertainty intervals of aggregate exposure under different uncertainty model configurations. Results for optimistic concentration model shown here.

The similarity between the final 2 rows of Table 1 shows that the uncertainty included in bystander model parameters here does not have a large impact compared to the corresponding variability-only run. The same conclusion was arrived at in the BREAM analysis of Kennedy *et al.* (2012). As seen from Figures 1 and 2, uncertainties are greatest in the distribution tails in each of the models. This is largely as a result of random simulations generated to account for uncertainty in the matching of diet and non-dietary exposures per individual.

4 Advanced uncertainty methods

4.1 Modelling variability and uncertainty of residues

As part of a dietary exposure model, the consumption amount of each single food item is multiplied by the corresponding residue concentration in that item. Depending on the model, the residue selected could be a single deterministic estimate or generated repeatedly from a distribution to represent variability. As explained in Section 3.1, the variability and uncertainty can be modelled using parametric or empirical distributions. In general we will refer to a *residue generator* as any algorithm used to estimate or simulate residues. The true distributions are typically unknown and can also have complex shapes (Roelofs *et al*., 2011; Roelofs and Craig, 2014; Roelofs *et al.*, 2014). In the following sections we illustrate residue generators with different levels of complexity, beginning with an assumption of independence between compounds before considering a multivariate method in which information is shared between pesticides.

4.1.1 Uncertainty in upper tail residues

Often residue generators sample values from existing concentration data sets (see examples in Section 3.2, for example, for the optimistic EFSA guidance scenario). However when the sample size is small this may lead to an underestimation of the upper tail (roughly above 1/(2n), as the highest data point represents the upper 1/n fraction of the distribution) as empirical sampling cannot generate values beyond those observed. We therefore considered modelling the upper tail using techniques from extreme value theory (EVT) (Coles, 2001). These methods use the result that the distribution of values above a fixed threshold (to be chosen by the modeller) tend to a Generalized Pareto Distribution (GPD) as the threshold level increases, regardless of the true distribution of the data (Davison and Smith, 1990). The GPD models can improve on standard parametric models, such as the lognormal (LN), if there is a particular focus on estimating probabilities of extreme residues and there is a reasonably good sample of data above the threshold. To use this approach a threshold first has to be chosen. We set this to be the 95^{th} percentile for our first example here but the impact should be tested for each individual case. The method will perform poorly if there are insufficient values above the threshold, as these are required to fit the separate GPD model. For the values below the threshold we use two approaches, either empirical sampling or empirical sampling followed by fitting a truncated LN distribution. To account for uncertainty about the residue distribution we repeat the sampling and/or fitting process to the values below the threshold and fit a GPD to the data above the 95^{th} percentile multiple times. Results are shown in Figure 3 based on a dataset of 100 simulated points from a truncated LN distribution with GPD tail. Inference was based on 1000 iterations. Uncertainty is captured in the sense that the data points are all contained within the intervals. For simplicity here we have used the assumed fitted model as the true distribution of residues. A more thorough test of a GPD model in the context of pesticide residues is presented in Roelofs *et al*. (2011) using a selection of alternative true distributions.

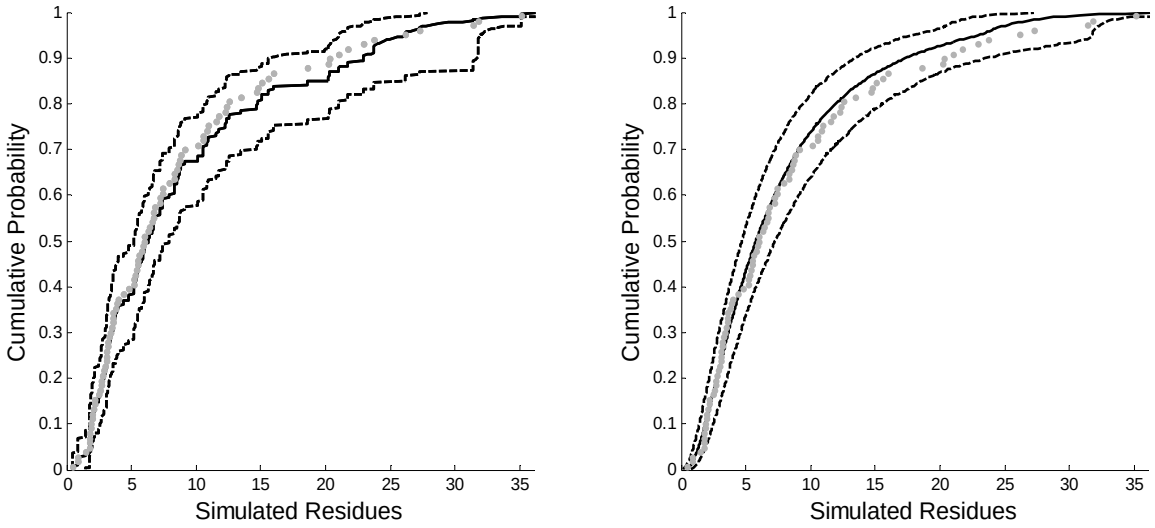

Figure 3: Comparison of the modelled residues with the simulated residues. The left pane shows the results when using empirical sampling for the distribution below the 95th percentile and the right pane shows the results when fitting a truncated LN distribution to the empirical samples for the distribution below the 95th percentile. Both methods then use a GPD to model the residues above the 95th percentile. Datapoints are shown together with pointwise median and 95% confidence intervals

4.1.2 Bayesian mixture of truncated Normal and GPD distributions with uncertain threshold

The model described in Section 4.1.1 was extended by assuming there is an unknown threshold for which the log-residue has a truncated Normal distribution when it is below the threshold, and a GPD distribution when above the threshold. We used a Bayesian framework to account for threshold uncertainty as well as uncertainty about the parameter values of the distributions. The model is similar to that described in Nascimento *et al*. (2012) that assumes a mixture of Gamma distributions for the lower tail. We instead used a truncated Normal distribution as it is unlikely there will be enough data in practice to fit a mixture of distributions for the lower tail. A Metropolis Hastings algorithm was used, with prior and proposal distributions for the GPD parameters, including the threshold, as proposed in Nascimento *et al*. (2012). Results are shown in Figure 4 for large (n = 10,000) and moderate (n = 120) simulated data samples which were sampled from a LN distribution truncated at the 95th percentile and a GPD above the 95th percentile. For these datasets, judging by the uncertainty intervals and median compared to the data, inferences appear to be reasonable.

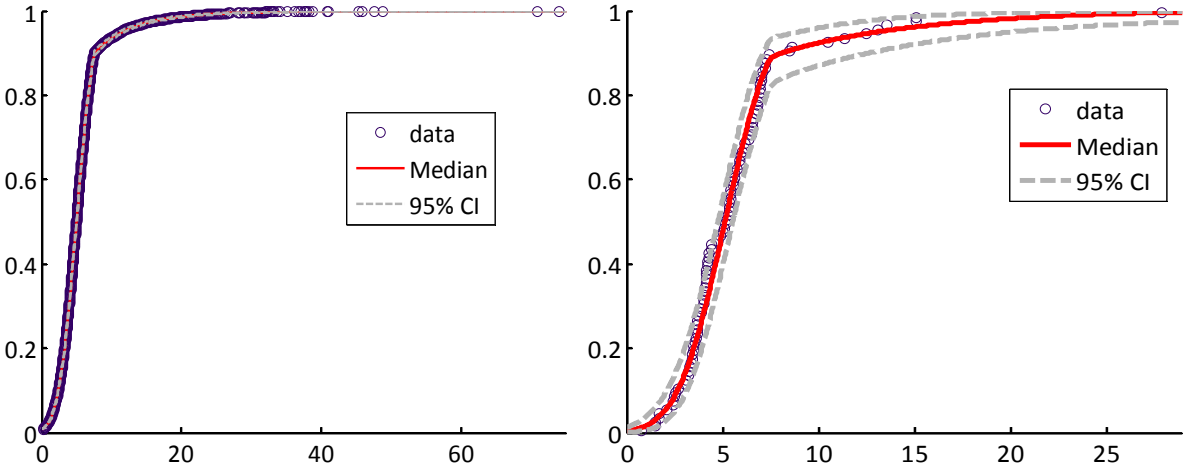

Figure 4: Data compared with the median and 95% credible intervals of the cumulative distribution function of the truncated Normal-GPD mixture distribution. Results are shown for n = 10,000 (left) and n = 120 (right) simulated data samples.

4.1.3 Pooled Bayesian mixture of truncated Normal and GPD distributions with uncertain threshold

With small samples (n = 120) the MCMC chains did not always converge for different simulated datasets. To address this issue, the model was extended following Roelofs and Craig (2014) to use pooled normalised data from multiple datasets where the shape, i.e. the functional form of the probability distributions describing residue variability, is assumed to be the same. Shape information might for example be shared between pesticides and/or foods. This larger pooled dataset, with for example 1000 points in the example below, can then be used to estimate the shape distribution whilst allowing each individual dataset to have its own location and scale. In Roelofs and Craig (2014) a Dirichlet mixture of Normal distributions was used as a semi-parametric method to estimate the shape. We instead used a single truncated Normal distribution for the lower tail but also incorporate a GPD for the upper tail and include threshold uncertainty, as in Section 4.1.2. The following steps are repeated:

Step 1: Normalise each log-residue dataset by its own simulated location and scale parameters, then pool the resulting normalised datasets.

Step 2: Infer the shape distribution of the pooled data using a mixture of a truncated N(0, 1) distribution and a GPD above a simulated threshold.

Step 3: Generate new values of dataset specific location and scale parameters given the assumed LN + GPD mixture distribution for the shape. These values are then used in Step 1 of the next iteration to normalise the data.

The model is described in more detail in Roelofs and Roelofs (2014). An example of the results for a data set consisting of 10 chemicals, for which 100 log-residue values were reported per chemical, can be seen in Figure 5 for one of the chemicals.

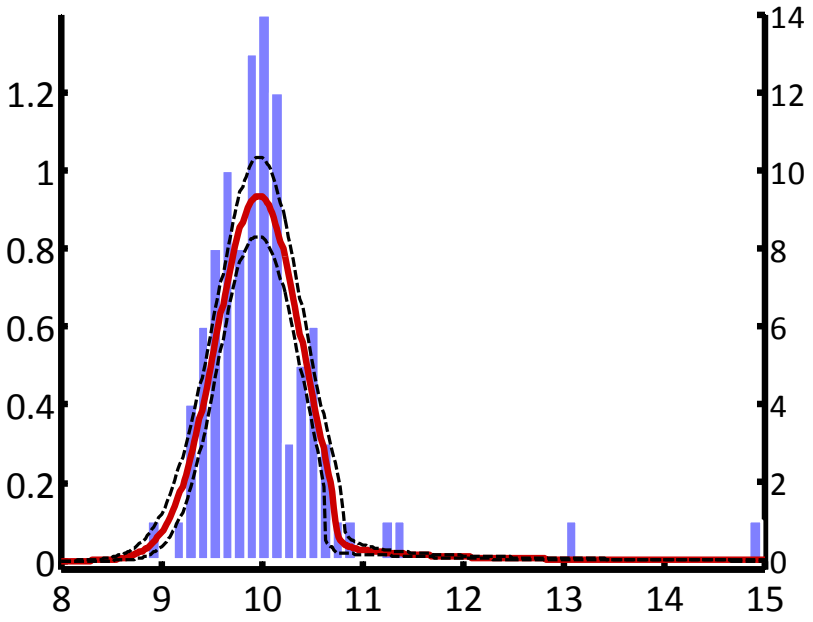

Figure 5 Histogram of data compared with the median and 95% credible intervals of the probability density function of the hierarchical truncated Normal-GPD mixture distribution

4.1.4 Bayesian Bivariate Normal Model

All the models described previously in this section have assumed independence between pesticide residue levels of different compounds. An analysis of the 2010 UK monitoring data suggests that correlations between compounds can occur (Roelofs 2013), so options should be considered for modelling them simultaneously. We present a bivariate Normal model which can be used to infer the pesticide residues of two pesticides on one food item. Higher dimensional extensions are possible but may not always be practical given the limited amount of pesticide residue data available.

The bivariate Normal model works by assuming that each composite monitoring sample comes from one of 4 categories: {x,y}, {<LORx, y}, {x, <LORy}, {<LORx, <LORy} where x and y indicate a measured pesticide residue level above the respective reporting limit (LORx and LORy). <LORx (<LORy) indicates that the sample contained no residue of X (Y) or that the levels were below the limit (LOR might represent a limit of detection or limit of quantification, for example). Using indicator functions we can allocate each data value to a particular category as part of a Bayesian framework. We need to determine the bivariate Normal distribution for X and Y, and a proportion of the composites that are true zeros. Prior means for the weights associated with each distribution/category were estimated from the UK pesticide usage survey (PUS) (Garthwaite *et al*. 2011, 2012). A weight was then selected to indicate the relative influence of the PUS-based prior compared to residue data. The model is described in detail in Roelofs (2013). Results are presented in Figures 6 and 7, based on a validation data set of size 100 with 90% of data below the limit of quantification. Predicted residue levels in Figure 7 show that the model is capable of inferring the dependency in residue levels for those samples that were treated with both chemicals.

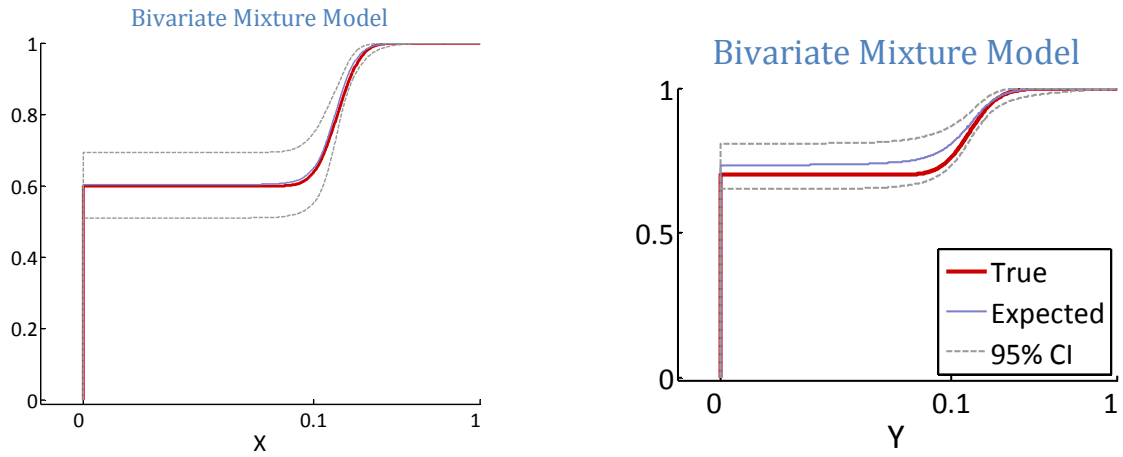

Figure 6: Median and 95% credible intervals of the marginal cumulative distribution function of X and Y obtained using the bivariate mixture model. The red lines indicate the true marginal distributions from which data were generated, which was a mixture of univariate and bivariate lognormal distributions.

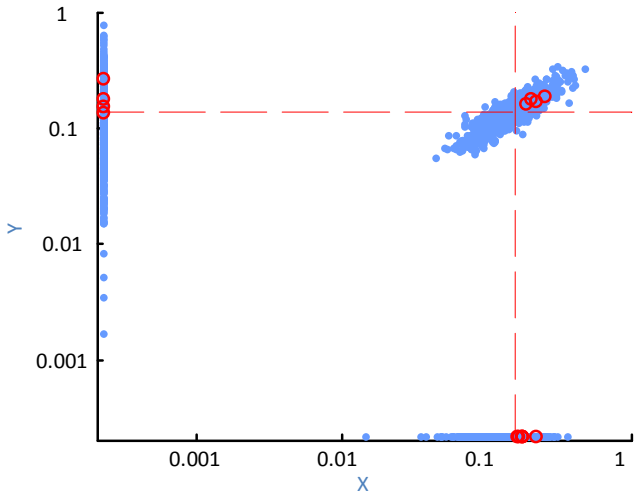

Figure 7: Predictions (blue dots) of residue levels based on a data set with high levels of censoring. The red circles indicate the observed data and the dashed lines indicate the LORs for X and Y.

4.2 Uncertainty in pesticide usage proportions

As mentioned in Section 2, the high proportion of ND values presents a problem in cumulative assessment. Incorporating information about the proportion of crop treated with a given compound can improve the accuracy of probabilistic models and this approach is implemented in MCRA for Acropolis. Estimates of these proportions were obtained from the UK PUS from the most recently available years (Garthwaite *et al*. 2011, 2012). All records of crops treated with one or more triazoles were extracted and summarised in 2 different ways: first as a percentage of grown area where a particular triazole or set of triazoles was applied (possibly independently on distinct dates); and second considering the percentage of individual treatments of that type. The former is used to derive the proportion of true zeros, whereas the latter allows us to assess the proportion of tank mixes, and therefore the proportion of crop with correlated residues. For compound i residues are

assumed to have a mixture distribution with a probability p_{0i} of being untreated, i.e. zero residue, and with probability $(1 - p_{0i})$ of having a $LN(\mu_i, \sigma_i)$ distribution.

A study was carried out to investigate the impact of uncertainty in the parameter estimates given the strength of information available, and to compare alternative strategies for estimating p_0. Some preliminary results are summarised in Table 2 and Figure 8 for carrots. Of the 297 surveyed holdings growing carrots, 4 treatment types were observed (NT = no triazoles, T = tebuconazole only, D = difenoconazole only, TD = tebuconazole and difenoconazole). The uncertainty model accounts for different distributions of field size for different treatment types. Residue data were obtained from UK Pesticides Residue Committee (PRC) monitoring data for 2008 2nd and 4th quarters (PRC, 2009). The UK carrot data comprise 96 samples, each of which is a composite measurement from 10 individual carrots. One residue contained both D and T, 68 are (< LOR, <LOR), 1 contained D, and 26 contained T.

Field treatment type (triazoles)	**Percent of total GB Carrot crop (2007)**		**Number of monitoring samples**
	Area level (whole year) (number of holdings)	**Individual treatment level[1]**	
none	46.08 (165)	46.08	68
Difenoconazole	5.27 (19)	11.27	1
Tebuconazole	33.39 (73)	41.92	26
Difenoconazole, Tebuconazole	15.26 (40)	0.72	1

Table 2. Proportions of field treatment area surveyed in carrot (2007), by triazole treatment type. In parenthesis are the numbers of the surveyed holdings growing carrots and using these triazole treatments

Figure 8 shows the impact on the mean and precision (1/variance) of estimated LN residue distributions from 3 alternative models for the proportion p_0 of crops not treated with Tebuconazole or Difenoconazole:

1. Use residue data only to infer p_0, using the Bayesian model of Paulo *et al.* (2005).
2. As in 1, but with fixed p_0 estimated from Table 2 as the proportion not treated with the given compound. For example, p_0 = (46.08 + 33.39)/100 = 0.7947 for Difenoconazole;
3. Quantify uncertainty in p_0 using our new Bayesian model.

Wider uncertainty intervals are seen when using the Bayesian model including uncertainty in the proportion of true zero residues (Paulo *et al*, 2005). Wider intervals arise generally in cases where $n_0 / n \to 1$ or n is small, where n_0 out of n residue samples are observed as NDs. In such cases, data may be so poor that Bayesian inference is driven by the prior, so it is even more important than usual to select priors carefully. We return to this in the discussion.

[1] Combined with parsnips and celery

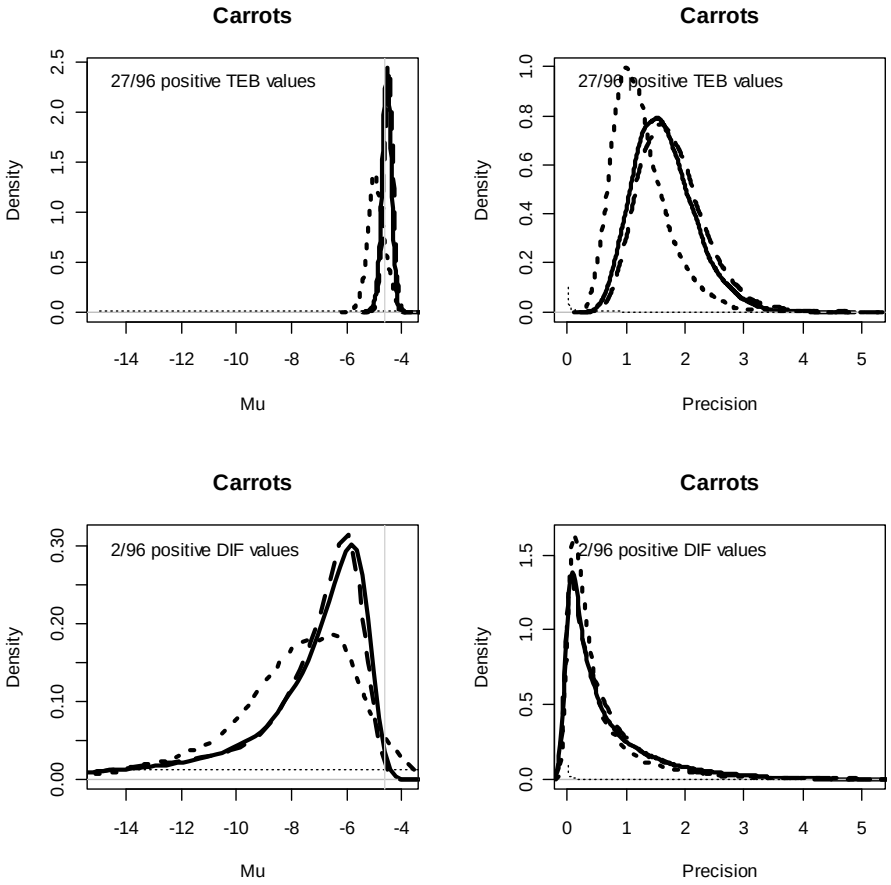

Figure 8: Posterior distribution of lognormal μ and precision = $1/\sigma^2$ parameters from Bayesian analysis of carrots monitoring data (residues of TEB=Tebuconazole and DIF=Difenoconazole), using 3 alternative methods for p_0. With p_0 estimated directly from PUS (solid lines) and fixed; with uncertainty in p_0 analysed using PUS data (dashed lines); p_0 estimated using residue data only (dotted lines). The corresponding numbers of residue measurements and positive values are indicated. Thin dotted lines indicate the prior distributions of the parameters.

4.3 Dose-response models

As consumers may be exposed to multiple pesticides, we need to assess whether the combined exposure could present human health risks. For this purpose, EFSA (2009b) recommend to focus on groups of pesticides that act through a common mechanism of toxicity, often referred to as the cumulative assessment group (CAG). Exposures to those pesticides in the CAG can be expressed in terms of an index chemical using relative potency factors (RPFs) if the slopes of the dose-response functions for each chemical are parallel. Before RPFs should be used, we need to assess whether it is reasonable to assume a common slope for the dose-response functions in the CAG and if so, what the common slope is. As there is generally little data available to estimate the parameters of the dose-response curve for each chemical, these parameters are inherently uncertain. We developed a Bayesian approach to estimate the parameters for all chemicals in a CAG under the assumption of a common but uncertain slope. Figure 9 shows an example of a four-parameter logistic dose-response function fitted using the common-slope assumption and Figure 10 shows the resulting uncertainty in RPF values.

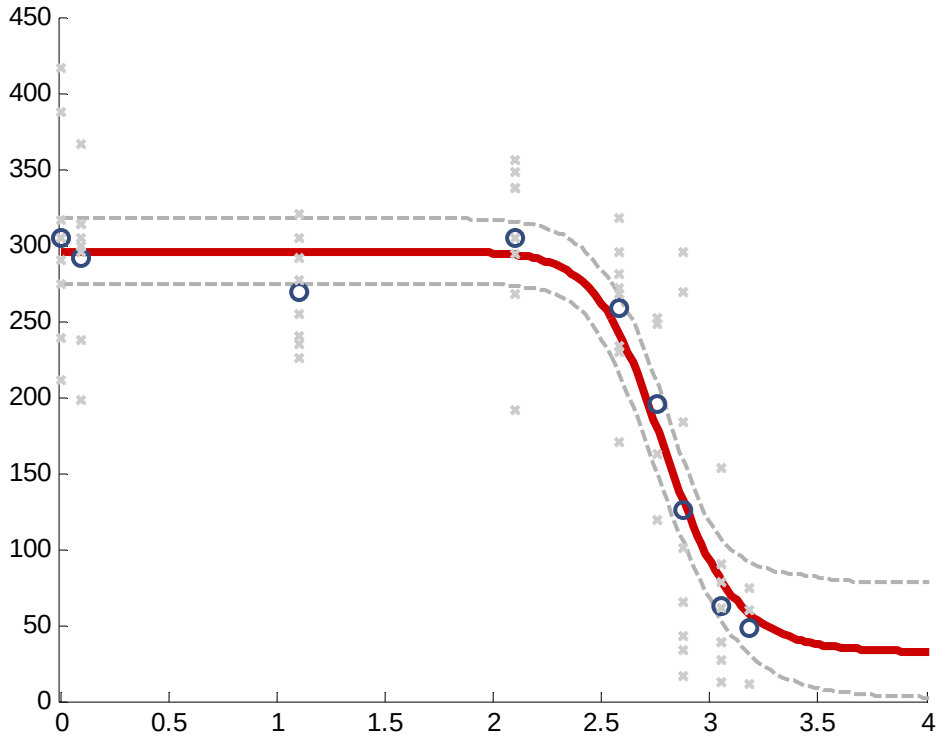

Figure 9: Dose-response function for one of the chemicals in a CAG assuming that the slopes of all chemicals in the CAG are parallel. Data points, pointwise median and 95% CI shown

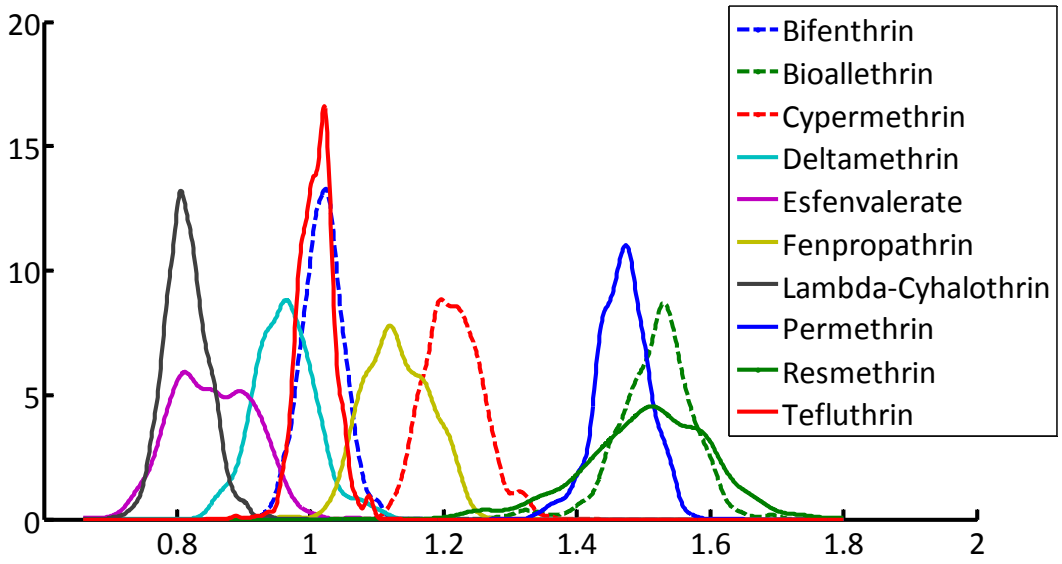

Figure 10 : Posterior distributions of the RPFs for 11 chemicals in a CAG

In keeping with the general modular 2DMC framework outlined in Section 3.1, a user should be able to choose from various options, including the one described above, to deal with uncertainty in the dose response. In MCRA 8.1 users may specify uncertainty in their dose-response functions by supplying either multiple (uncertain) values for the RPFs or 'limit dose'. Limit dose is a generic term

for limiting exposures such as the Critical Effect Dose (CED). When uncertainty analysis is enabled (via the user-interface), for each uncertainty iteration a single RPF/limit dose value will be selected from the vector of possible values - thus propagating uncertainty in dose-response through the exposure calculation. Pre-calculated outputs from the model described above, or alternatives, will be compatible with this implementation.

4.4 Assessment of unquantified uncertainties

A range of approaches have been described for quantifying various uncertainties affecting models of pesticide exposure. However, in any given assessment, only a proportion of the uncertainties will be quantified, while others remain unquantified. Even when a source of uncertainty is quantified, there will be further uncertainty (sometimes referred to as 'secondary uncertainty') about how well it is represented. In the review of the SHEDS model of Zartarian *et al*, (2008), for example, it was noted that more research would be needed to account for model and scenario uncertainty, meaning those aspects not captured in the model itself.

When using an exposure assessment to support decision-making, it is important to consider whether the unquantified uncertainties might be large enough to change the risk management decision. It is therefore necessary to identify and evaluate the unquantified uncertainties in some way. This is implicit in Codex Working Principles for Risk Analysis (Codex, 2011) and EFSA (2009) guidance on transparency, which both emphasise the need to consider uncertainties at each step in the risk assessment and document them in a transparent manner. EFSA's (2006) guidance on uncertainty in dietary exposure assessment suggested a tabular approach for listing the uncertainties and evaluating their individual and combined impact on the estimated exposure. The 2006 guidance suggested an ordinal scale for evaluating the uncertainties (e.g. low, medium or high, represented respectively by one, two or three plus or minus symbols). Subsequently it was recognised that to inform decision-making it is useful to provide at least an approximate quantification of the scale, and this was recommended in EFSA's guidance on probabilistic dietary exposure assessment (page 64 and Table 6 in EFSA 2012).

The format used in Appendix 2 of EFSA (2012) included a general assessment of unquantified uncertainties affecting EFSA's 'basic' probabilistic models and was designed to be used as a starting point or template. This can be adapted to specific applications and models such as those implemented in the MCRA software tool. Accordingly, uncertainty tables have been prepared for the MCRA software tool by starting with the EFSA (2012) tables and adjusting the content to take account of the specific approaches implemented for modelling dietary exposure in MCRA. The resulting generic tables are presented in the Reference Manual of MCRA (2013). Note that the overall assessment is dependent on the case-specific combination of uncertainties and is not included as part of the generic tables. An example for cumulative dietary exposure to triazoles is given in Boon *et al* (in prep.). Unquantified uncertainties affecting non-dietary routes and also the combined impact (overall direction and magnitude) relative to the overall estimate of aggregate exposure should be assessed separately. Uncertainties for non-dietary routes may be assessed using tables of the same format as that used for dietary exposure in MCRA (2013), or using the simpler 2-column format shown in Table 6 of EFSA (2012). Examples of completed evaluations are included in the case studies of aggregate exposure reported by Kennedy *et al* (in prep).

It may be convenient for both assessor and reader to present a separate summary table, as illustrated in Table 3, which captures the overall impact on uncertainties from each of the routes of exposure and the aggregate result. Again, an example may be found in Kennedy et al. (*in prep*).

Table 3. Suggested format for table summarising the assessment of unquantified uncertainties for each route of exposure together with the assessor's subjective evaluation of their combined impact on the estimate of aggregate exposure (bottom row). For example of scale for symbols see Figure 9.

Route of exposure	**Magnitude and direction of unquantified uncertainties affecting the estimated exposure**
Dietary route *Copy here the narrative conclusion from the evaluation of unquantified uncertainties for the dietary route (e.g. bottom row of Table XX in INSERT REFERENCE TO MCRA 8 REFERENCE MANUAL).*	*Symbols to show evaluation of uncertainty for the dietary route, copied from the relevant table (e.g.: +/++)*
Dermal route *Copy here the narrative conclusion from the evaluation of unquantified uncertainties for the dermal route (e.g. bottom row of Table XX in Kennedy et al. INSERT REFERENCE TO AN UNCERTAINTY TABLE FOR DERMAL EXPOSURE IN THE CASE STUDIES).*	*Symbols to show evaluation of uncertainty for the dermal route, copied from the relevant table*
etc.	
Insert additional rows for further routes of exposure, when required.	
Overall evaluation of uncertainty affecting the estimate of aggregated exposure *Add narrative text here, describing the assessor's subjective evaluation of the overall degree of uncertainty affecting the assessment outcome, taking account of the uncertainties for each route as summarised above.*	*Evaluation of overall uncertainty for the estimate of aggregate exposure (e.g., - - - /+)*

As mentioned above, if symbols are used for the evaluation of the uncertainties then it may help to define a quantitative scale for them; an example is shown in Figure 11. Alternatively, the evaluations could be reported as quantitative ranges without using symbols (e.g. 0.5x – 5x).

It is acknowledged that assessing the impact of unquantified uncertainties in this way is subjective and approximate. Some users of the tabular approach have preferred not to give quantitative ranges and, in some cases, give only a narrative evaluation for the combined effect of multiple uncertainties (e.g. EFSA (2013) Opinion on bisphenol A exposure assessment). However, a quantitative judgement will usually be implicit either in the reported conclusion of the exposure assessment or in the risk management decision it informs. For example, if the exposure estimate is reported without reference to uncertainty, this presumably implies the overall uncertainty is judged to be too small to change the estimate by enough to alter the risk management decision. This judgement is implicitly quantitative, and transparency requires that the basis for it is presented explicitly. In doing so, it is of course important to make clear the subjective and approximate nature of the assessment, to avoid it being over-interpreted. If a firmer assessment of the unquantified uncertainties is required, consideration could be given to conducting the evaluation using formal methods for expert elicitation (e.g. O'Hagan *et al*. 2006) or to treating some of the unquantified uncertainties deterministically or probabilistically so they become part of the quantified uncertainty. In the latter case it may be efficient to target quantification on the most important unquantified uncertainties, as indicated by the approximate subjective evaluation. This process may be repeated iteratively until

the characterisation of uncertainty is sufficiently clear to support the risk management decision in hand. For more discussion of this tiered approach to evaluating uncertainty, see EFSA (2006).

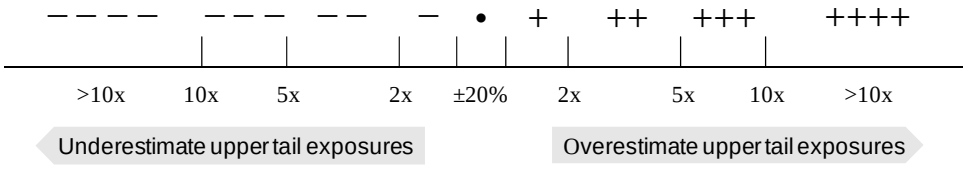

Figure 11. Example of quantitative scale for symbols used to evaluate unquantified uncertainties.

5. Discussion

We have described a general framework for uncertainty identification and modelling, as developed for the Acropolis project and implemented, in part, within MCRA. The flexibility built into the system for aggregating non-dietary information is particularly important, as relevant information available for any given risk assessment can be extremely limited. Uncertainty will therefore be a key consideration in any risk management decisions. Our model allows for simple default point estimates, sampled values from probability distributions, or 2DMC matrices, allowing us to quantify uncertainty, variability or both. The Bayesian and non-Bayesian (Frequentist) statistical alternatives have been discussed. In both cases, an element of subjective judgement is involved in the selection of probabilistic models. In the Bayesian case we have noted that prior distributions for parameters can have a substantial effect on the results when data are sparse. Model diagnostics and sensitivity analysis are important general tools to monitor model fit and investigate impacts due to variations in choices made. For those model components where uncertainty is assessed to be substantial, the assessment should be repeated with alternative plausible alternative associated modelling choices. The aggregation of dietary and non-dietary sources, in which both include variability and/or uncertainty, can be performed simply if the output is generated with variability/uncertainty quantified, in a form that is compatible with MCRA aggregate exposure simulations. The Browse model is being developed in this way.

The optimistic and pessimistic basic probabilistic dietary exposure models as presented in the EFSA (2012) guidance are built into MCRA and examples were presented in Section 3.2. The modular design coupled with a simple interface allows user generated simulation results from external models or databases to be included both for aggregating non-dietary exposures, and for extending residue generator models. This has been illustrated with examples using the BREAM model in Section 3.3. Additional case studies are described in Kennedy *et al* (2013) that use ConsExpo and EUROPOEM models.

Many of the methods presented above require further work before they can be integrated within standard software tools. It is important to ensure these are robust and statistically validated against a range of realistic datasets and sample sizes that would arise in practice. Some methods, such as the multivariate residue generator presented in Section 3.3.4, appear to work well for reasonable sized datasets but are not suitable for small sample sizes. In the case of dietary exposure, MCRA already includes rules that will automatically revert to a simpler probabilistic or deterministic method in cases where the data are insufficient for reliable estimation.

In almost all UK crops surveyed, the proportion of crop treated with triazole combinations as tank mixes was much lower than the percentages applied at different times (for example, 0.72% vs. 15.26% for the combination Difenoconazole/Tebuconazole, see Table 1). Although up to 12 triazoles appeared in the survey as a whole, at most 3 occurred in any given field. Those with 3 triazoles were never observed as tank mixes and were rare. It was concluded from this investigation that correlations in amounts of residues due to simultaneous application are not substantial, due to the rarity of tank mix applications. However, correlations can occur as mentioned in Section 3.3.4 and Roelofs (2013). In cherries, a 7.96% treatment proportion was estimated for tank mixtures involving Fenbuconazole and Myclobutanil, but this is not a major consumed food type. We conclude that uuncertainty in usage proportions of particular combinations can be left out of the model with minimal impact on the results. The practical implementation described in Section 3.2, with fixed proportions for co-usage and independent residue amounts therefore appears to be justified. Further investigation is required to establish whether for other crops the uncertainty may have a more significant impact. Section 4.1.4 illustrates a method that can account for correlations, when enough data are available, and our study of the PUS data suggests it is rare to find more than 2 triazoles in a tank mix.

Situations frequently arise in which there are no data at all, or where additional information can be obtained from a subject expert. In our study for example, information about pesticide combinations applied to overseas crops such as bananas and pineapples, found to be main contributors to triazoles exposure, was not available. Probability distributions, if properly measured to represent the beliefs of an expert, can be used alongside data models as part of a Bayesian analysis as explained in Section 3.1. The uncertainty prioritisation exercise identified the following features as potential candidates for expert judgment:

- For aggregate exposure, usage of a particular consumer product, such as a biocide spray, wood treatment or cosmetics in a given country
- Pesticide usage in countries that export key commodities such as pineapple and bananas. These uncertainties, when quantified, can be included in the methods of Section 4.2.

Relevant methods are described in O'Hagan *et al.*, 2006.

References

Bosgra, S., van der Voet, H., Boon, P. E. and Slob, W., 2009. "An integrated probabilistic framework for cumulative risk assessment of common mechanism chemicals in food: An example with organophosphorus pesticides." Regulatory Toxicology and Pharmacology, 54 124-133.

Bremmer HJ and van Veen MP (2000). General Factsheet. RIVM rapport 612810009 available at http://www.rivm.nl/en/, National Institute for Public Health and the Environment, Bilthoven.

Burmaster, D.E., Wilson, A.M., 1996. An introduction to second order random variables in human health risk assessments. Human and Ecological Risk Assessment 2, 892–919.

Codex, 2011. Working Principles for Risk Analysis. In: Codex Alimentarius Commission, Procedural Manual, 20th edition. FAO/WHO, Rome.
Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Verlag, London.

Coles S. An Introduction to Statistical Modeling of Extreme Values. London: Springer Verlag, 2001.

Cullen, C.A., Frey, H.C., 1999. Probabilistic Techniques in Exposure Assessment: A Handbook for dealing with Variability and Uncertainty in Models and Inputs. Plenum Press, New York.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds (with discussion). J. R. Statist. Soc. B 52, 393–442.

EC (2009). Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:309:0001:0050:EN:PDF

Efron, B. and Tibshirani, R.J. (1993). An introduction to the bootstrap, Chapman and Hall, New York.

EFSA, 2006. Guidance of the Scientific Committee on a request from EFSA related to uncertainties in dietary exposure assessment. EFSA Journal, 438, 1-54.

EFSA, 2009a. Guidance of the Scientific Committee on Transparency in the Scientific Aspects of Risk Assessments carried out by EFSA. Part 2: General Principles. EFSA Journal 1051, 1-22.

EFSA, 2009b. EFSA Panel on Plant Protection Products and their Residues (PPR Panel) Scientific Opinion on risk assessment for a selected group of pesticides from the triazole group to test possible methodologies to assess cumulative effects from exposure throughout food from these pesticides on human health on request of EFSA. 2009; 7 (9); 1167. [104pp.]. Available online: www.efsa.europa.eu

EFSA, 2012. Guidance on the Use of Probabilistic Methodology for Modelling Dietary Exposure to Pesticide Residues. EFSA Journal, 10(10): 2839.

EFSA, 2013. DRAFT Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs – Part: exposure assessment. http://www.efsa.europa.eu/en/consultations/call/130725.pdf

Garthwaite, D.G., Barker, I., Parrish, G., Smith, L., Chippindale, C. & Pietravalle, S. (2011) Pesticide Usage Survey Report 235 - Arable Farm Crops in the United Kingdom, 2010. London: Defra

Garthwaite, D.G., Barker, I., Parrish, G., Smith, L. & Pietravalle, S. (2012) Pesticide Usage Survey Report 243 – Outdoor Vegetable Crops in the United Kingdom, 2011. London: Defra

Glass et al. 2012. (EFSA operator activity surveys.)

Hoare, J., Henderson, L., Bates, C.J., Prentice, A., Birch, M., Swan, G., Farron, M., 2004. The National Diet and Nutrition Survey: adults aged 19 to 64 years. Summary Report. TSO, London. Available online: food.gov.uk.

Kennedy, M. C. (2010). Bayesian modeling of long-term dietary intakes from multiple sources. Food and Chemical Toxicology. 48, 250-263.

Kennedy, M. C., Butler Ellis, M. C., Miller, P. C. H. (2012). BREAM: A probabilistic Bystander and Resident Exposure Assessment Model of spray drift from an agricultural boom sprayer. Computers and Electronics in Agriculture. 88, 63-71.

Kennedy, M. C., Glass, Bokkers, B., Hart, A., Hamey, P., Kruisselbrink, J., de Boer, W., van der Voet, H., Garthwaite, D., Van Klaveren, J. D. (2013). A European model and case studies for aggregate exposure assessment of pesticides, *in prep.*

MCRA (2013). Web-based software for Monte-Carlo risk assessment: User, format and reference manuals. Biometris, Wageningen University and Research centre and RIVM, National Institute for Public Health and the Environment. https://mcra-test.rivm.nl/Support/ReferenceDocuments

McCurdy T., Glen G., Smith L., and Lakkadi Y. (2000). The National Exposure Research Laboratory's Consolidated Human Activity Database. Journal of Exposure Analysis and Environmental Epidemiology, 10:566-578.

Nascimento, F. do, D. Gamerman, and H. Lopes (2012). A semiparametric Bayesian approach to extreme value estimation. Statistics and Computing. 22, 661{675.

O' Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. E., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E. and Rakow, T. (2006). Uncertain Judgements: Eliciting Expert Probabilities. John Wiley and Sons.

Paulo, M., van der Voet, H., Jansen, M.J.W., ter Braak, C.J.F., van Klaveren, J.D., (2005). Risk Assessment of dietary exposure to pesticides using a Bayesian method. Pest Management Science, 61, 759–766.

PRC (2009). Pesticide Residues Committee Reports 2008. Second Quarter (April to June 2008, Published 9th December 2008) and Fourth Quarter results (October to December 2008, Published 25 June 2009) http://www.pesticides.gov.uk/prc.asp?id=2536 Accessed 18-07-2011

RIVM (2013). PROAST - National Institute for Public Health and the Environment. Available online from http://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST [accessed 22-08-13]

Roelofs, W., Kennedy, M.C., Roelofs, V., Craig, P., Anderson, C. W. and Ámbrus, Á, (2011). Final Report for the extension of the CRD/Syngenta project on improved methods for estimating residue distributions (Ref PS2615).. Available online http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=16249 Accessed 25-11-13.

Roelofs, W. (2013). Multivariate modelling of pesticide residues. *In Prep*.

Roelofs, W. and Craig, P. C. (2014). Modelling within-field and between-field variation in pesticide residues. *In Prep.*

Roelofs, W. and Roelofs, V. (2014). Bayesian Hierarchical mixture model for unit variation in pesticide residues. *In Prep.*

Roelofs, W., Craig, P., and Ámbrus, Á., (2014). Describing unit variation in residue data using Dirichlet Process Mixture Models. *In Prep.*

Van der Voet, H., de Boer, W. J., Kruisselbrink, J., Goedhart, P. W., van der Heijden, G. W. A. M, Kennedy, M. C., Boon, P. E., van Klaveren, J. D. The MCRA model for probabilistic single-compound and cumulative risk assessment of pesticides. *In Prep.*

Van Klaveren, J. D, Goedhart, P. W., Wapperom, D. and van der Voet, H. (2012). A European tool for usual intake distribution estimation in relation to data collection by EFSA. Available online: [http://www.efsa.europa.eu/en/supporting/doc/300e.pdf] accessed 25th July 2013.

Xue J., Zartarian V., Özkaynak H., Dang W., Glen G., Smith L., and Stallings C. (2006). A Probabilistic Arsenic Exposure Assessment for Children Who Contact Chromated Copper Arsenate (CCA)-Treated Playsets and Decks: Part 2: Sensitivity and Uncertainty Analyses. Risk Analysis 26: 533-541.

Zartarian V.G., Özkaynak H., Burke J.M., Zufall M.J., Rigas M.L., Furtaw Jr. E.J. A Modeling Framework For Estimating Children's Residential Exposure and Dose to Chlorpyrifos Via Dermal Residue Contact and Non-Dietary Ingestion. *Env Health Perspectives* **2000** 108(6): 505-514.

Zartarian, V.G., Glen G., Smith L., Xue J. Stochastic Human Exposureand Dose Simulation Model for Multimedia, Multipathway Chemicals SHEDS-Multimedia Model Version 3 Technical Manual. U.S. Environmental Protection Agency, EPA/600/R-08/118, 2008. [http://www.epa.gov/heasd/documents/shedsmultimedia3_techmanual.pdf] [accessed 22-08-13]

Zartarian, V. G., J. Xue, G. Glen, L. Smith, N. S. Tulve, and R. Tornero-Velez. Quantifying Children's Aggregate (Dietary and Residential) Exposure and Dose to Permethin: Application and Evaluation of EPA's Probabilistic SHED-Multimedia Model. Journal of Exposure Science and Environmental Epidemiology. Nature Publishing Group, London, Uk, 22(3):267-273, (2012).

Acknowledgements

Acropolis is funded by the European Commission (EU Grant Agreement No. 245163) with additional funding from the UK Food Standards Agency (T10023). Dave Garthwaite provided PUS data for the non-dietary exposure examples and Helen Owen assisted with coding of MCRA aggregate and uncertainty model components.