

REVIEW

Open Access

Population level risk assessment: practical considerations for evaluation of population models from a risk assessor's perspective

Magnus Wang^{1*} and Robert Luttk²

Abstract

Population models are increasingly being considered as a tool for pesticide risk assessment in order to evaluate how potential effects act on the population level and population recovery. While the importance and difficulties of such models have been discussed by various authors during the past decade, mainly with a focus on how to describe or develop such models, several biological and methodological aspects have never been addressed so far, which are relevant for the application of models in risk assessment. These include a critical review of our knowledge of a species, the use of field data by taking methodological constraints into account, how to include uncertainty in model validation or how to measure effects. Although these aspects will be critical for the acceptance of population models by authorities, most of them apply not only to population models, but also to standard risk assessment. In the present article, we give practical recommendations for addressing these questions in population level risk assessments.

Keywords: population models, risk assessment, pesticides, recovery, uncertainty

Introduction

During the last decade, several attempts were made to introduce population modelling in pesticide risk assessment by either the academia or the industry [1-6]. The main advantage of using population models in risk assessments is seen in reaching a higher realism and ecological relevance. Most recently, the European Union [EU]-funded project CREAM [7] was initiated, aiming at the application of population models in chemical risk assessment. This project includes a variety of subprojects ranging from aquatic organisms to polar bears. Despite these attempts to use or establish population models in pesticide risk assessment, the acceptance of such models is still limited in Europe. This is due to two main issues: the complexity of the models and a lack of trust in the models or scenarios. Population models are necessarily complex because they must include all aspects which are relevant for the development of populations, i.e. reproduction, survival and further factors. Even if the simulation of each of these

factors may be easy to understand (e.g. the simulation of survival, which may be as simple as throwing a dice), the combination of various factors makes the models more complex. However, in reality, it's not the model which is complex, but it's the biology and ecology which are complex, and the model has to reproduce this complexity. Still, due to this complexity, it is extremely important to explain how a model works and what it represents. This does not only imply a detailed technical description of how the model works mechanistically (model type or description of processes in the model) and how it was parameterized, but also covers important assumptions used for the model, a description of what the model actually reflects (representativeness) and an evaluation of the realism of the model (validation) and applied scenarios.

While the advantages of using models and how to develop and describe them have been addressed in several previous studies [2,8,9], several aspects which are extremely relevant for the application of population models in pesticide risk assessment have never or marginally been addressed [5]. These include an assessment of our biological knowledge of a species, the critical

* Correspondence: magnus.wang@rifcon.de

¹RIFCON GmbH, Im Neuenheimer Feld 517, Heidelberg, 69120, Germany
Full list of author information is available at the end of the article

review of the quality of the data which are used for parameterisation or validating models or the question on how to assess for which purpose a model can be applied (representativeness of a model and scenario selection). In the following sections, we discuss these aspects which we consider relevant and important for applying population models in pesticide risk assessment. We specifically focus on the view and the needs of risk assessors.

Technical description of models

A major obstacle in the application and acceptance of population models is the communication of the methodology and the complexity of the models. Although parts of population models, e.g. the simulation of developmental stages or the simulation of survival, may be rather easy to understand, the existence and interaction of many such 'modules' make models more difficult to explain. Therefore, a clear description of the applied models is one important prerequisite to make models understandable to risk assessors who have no expertise in modelling. In the past, a few attempts were made to harmonise the way of technically describing population models. Grimm et al. [8] proposed an Overview, Design concepts, Detail [ODD] protocol to describe population models, including submodels and general principles of a model. Later, Schmolke et al. [9] developed the Transparent and Comprehensive Ecological [TRACE] model documentation framework, which extends the idea of ODD to documenting not only a model, but the entire process of formulating, implementing, parameterising, testing and using a model (the so-called 'modelling cycle', [10]). Schmolke et al. [9] suggested to use TRACE as the basis for 'good modelling practise' for models which are to be used for environmental decision making. The proposed description of models following ODD or TRACE has a focus on the technical part of the model and includes aspects such as the purpose of a model, scales or input parameters. Topping et al. [11] proposed that a model description should include a structured description of the object-oriented computer programme code. Accordingly, his proposal for a model description, which partly overlaps with Grimm's ODD protocol [8], includes a description of programming classes and functions according to methods used in information science.

A concise and standardised way to technically describe models would facilitate the application and review of models in risk assessment. However, apart from how a model functions, which parameters have been applied or what purpose a model has, a number of other aspects are relevant for evaluating if a model is suitable and realistic, including also a critical review regarding the biology of the species and available field data.

Complexity

High versus low complexity

From a naïve perspective, one might think that a model which includes all the various processes acting in natural populations, from physiological mechanisms of individual animals or behavioural responses to population level responses, would produce the most realistic population model since all these processes exist in nature. Apart from the fact that such a model would be extremely complex (high model complexity is often regarded a main obstacle of using population models in risk assessment [12]), such a 'meta-model' would usually go hand-in-hand with increased uncertainty [10]. Again, it has to be acknowledged that this phenomenon is not restricted to population models only but applies to any analysis based on several parameters, including standard risk assessments. Of course a minimum number of parameters (or processes) are necessary to produce a realistic model. For example, survival rates are needed in practically all population models since without these parameters, modelling of population dynamics or recovery would not be possible. The increase of model complexity up to a certain degree usually makes a model more realistic and increases performance. However, including more parameters will often also increase the overall uncertainty of a model, especially if parameters cannot be estimated reliably, e.g. due to the lack of studies [13]. Hence, at a certain degree of complexity, the performance (i.e. the accurate representation of population dynamics) of the model will decrease again [14]. For a risk assessor, this implies that with increasing complexity, the chance increases that uncertain data (e.g. incidental information) or mechanisms have to be used in a model and that the model hence produces a prediction with a higher uncertainty. Consequently, an optimal model includes those parameters, which are really needed to simulate the population development, but excludes all parameters or processes which either have no significant influence on the model or which are not reliable. For example, when considering food ingestion, it might be more adequate to use measured food ingestion rates rather than to include an additional submodel for simulating food ingestion. Such a submodel may be based on the metabolic rate of an animal (normally measured in the laboratory), extrapolating this rate to free ranging animals, estimating the diet composition (which will vary over seasons and habitats), food energy content (which won't be available for some food items), modelling plant growth during the seasons (if the food includes plants), etc. Such an approach may inevitably make it necessary to use data from incidental observations or to make assumptions on parameters for which no information is available.

Coping with uncertainty

Parameters taken from the field for use in a population model include some degree of variability and uncertainty resulting from methodological constraints or limited knowledge. While the need of estimating uncertainty and variability was recognised early in the advent of population ecology [15-17], these factors are not commonly addressed explicitly in population models (for exceptions, see the study of Clark [18]). It is important to acknowledge that different types of uncertainty have to be dealt with, such as uncertainty due to sample size and geographical, spatial or other types of variation. The easiest source of uncertainty is the uncertainty caused by sample size. Sample size effects can be addressed by confidence intervals if a sufficient number of single measurements are available. Approaches for addressing this type of uncertainty are discussed in detail in the study of Hart [19]. Often, however, it will be difficult to distinguish between natural variability and uncertainty. For example, if litter size in 10 litters obtained from a field study ranges from four to six pups, we don't know how much this range is influenced by uncertainty due to the small sample size (for larger samples, sample size effects can however be addressed, e.g. by subdividing the sample, bootstrapping or other methods). Often, when only mean values (and standard deviations) or ranges are reported, only a qualitative comparison of the values from different studies is possible, which can, however, also help to evaluate the reliability of the used data (especially if samples sizes are available). A qualitative analysis of uncertainty can be sufficient; see for instance the approach in chapter 6.8 of the Guidance document for risk assessment for birds and mammals [20]. It is better to try to cover all sources of uncertainty instead of reporting one uncertainty in detail but neglecting others (however, in some cases, it may be impossible to address some sources of uncertainty due to the lack of appropriate data). In Table 1, two examples of data sets are shown. In the

first example, a very detailed data set would enable a rather exact estimation of uncertainty due to sample size, while in the second example, much less detailed data is available. However, since data from different studies are listed, uncertainty due to spatial or temporal variation can be addressed too.

Remarkably, when addressing uncertainty in a population model, a high uncertainty of a parameter does not necessarily result also in a high uncertainty of the output of the model, since some parameters have little influence on the model output while others have a strong impact. A sensitivity analysis can clearly show for which parameters additional information would be worthwhile to decrease the uncertainty of the model.

Realism and validation

Models should be validated in order to show that they produce a realistic output. Although true validation in a strict sense is never possible [21], i.e. it can never be shown that a model produces correct results for all possible situations, we here use the term 'validation' as a method which shows that a model produces a realistic prediction for a variety of situations by comparing model results with observational data [22,23].

Even without knowing all technical details of a model, one can get a first impression of the realism of a model by comparing how it performs in comparison to nature. If a model reproduces all important aspects of the behaviour of natural populations (especially those driving population dynamics and recovery), one can be confident that the model is realistic or at least, that a model reflects our current knowledge of a species, which will always be limited.

Validation of population models has received little or almost no attention in the past [24]. A recent review by Schmokle et al. [25] reported that even today only 3% of models are validated using independent data. Examples of model validation can be found in the studies of Wang and Grimm [26] and Preuss et al. [27] and

Table 1 Examples of how to address uncertainty for different datasets

Data type	Available data	Possible estimation of uncertainty and variability	Source of uncertainty which can be addressed
Litter size (common vole, <i>M. arvalis</i>)	1 to 2 ($N = 37$) 3 to 4 ($N = 106$) 5 to 6 ($N = 44$) 7 to 8 ($N = 7$) [68]	Confidence intervals can be calculated after recreation of the original sample underlying the data, and variability can be estimated based on the data distribution (or calculation of a standard deviation)	Only uncertainty due to sample size is included; uncertainty due to study location, climate or other factors is not included
Clutch size (little owl, <i>Athene noctua</i>)	Central Europe, usually 3 to 5 and exceptionally up to 7 [69] France, 3.9 ($N = 80$, [70]) Portugal, 3.3 ± 1.2 (min, 1; max, 5; $N = 15$, [71])	Only qualitative analysis of uncertainty is possible. However, the standard deviation from the study of Tomé et al. [71] and the ranges from Glutz von Blotzheim [69] indicate that an average clutch size between 3 and 4 is realistic. Data from Portugal indicate that smaller clutch sizes might be observed in southern Europe	Variability and uncertainty due to sample size cannot be distinguished in detail, but sample sizes and comparable ranges indicate that data are reliable. Since data from several studies and years are shown, uncertainty due to study location and temporal variability can be estimated at least qualitatively

regarding models that were partially validated in the review of Schmokle et al. [25].

Parameters for validation

While a number of authors have focussed on parameter estimation (e.g. Grimm and Railsback [10], Jakeman et al. [13], Rykiel [22]), only few have discussed which parameters need to be included in validation and how validation shall be done in detail (for example, see the study of Bart [28]). Since the aim of population models in pesticide risk assessment is a realistic representation of the population development of a species under concern, the most logical approach is to validate all aspects of a model which have a significant influence on the simulation of the population development. For the development of a population for an organism with a rather simple biology, it may be sufficient to address *survival* (including predation) and *reproduction*, which are the single most important factors for the development of a population. For many species, also the *spatial behaviour* (such as *dispersal* or *home range behaviour*), *population dynamics* and *habitat preference* will be crucial. For instance, dispersal may be a key element for the recolonisation of habitats. This does not necessarily mean the recolonisation of habitats, which are depleted because of effects by pesticides, but it may also include recolonisation of fields which have been abandoned by animals after harvest.

The validation of each of these factors will usually include several parameters, i.e. when validating reproduction in a *Daphnia* model, one might compare the fraction of animals with eggs, the proportion and the timing of sexually and parthenogenetically reproducing animals, the formation of ephippia and even further parameters [27]. Also, indirect comparisons are possible, for example, a comparison of the age distribution in a model with the age distribution from the field. Evidently, only those parameters are suitable for validation, which are not implicitly defined in a model. If survival is taken from the field and set to this value in the model, then it is useless to try to validate this value. However, many parameters in a model are not implicitly defined but are the result from the interaction of other parameters in the model and the model structure [10,28,29]. For example, in the shrew model in the study of Wang and Grimm [26], one of the factors which was validated was the fraction of lactating or pregnant females (grouped together because in the literature, only data for 'lactating or pregnant females' were available for validation). This parameter was not implicitly set in the model, but it resulted indirectly from the age of sexual maturity of females, gestation length, lactation length, sex ratio at birth and other parameters. Evidently, if one of these parameters was unrealistic, also the fraction of lactating

and pregnant females would show an unrealistic result. Therefore, an efficient way of validating is to focus on parameters which depend on many other parameters (and the model structure). This indirect approach of validation is also known as 'pattern oriented modelling' [30,31], 'indirect' [29] or 'secondary predictions' [28].

When performing a validation analysis of a population model, uncertainty should be addressed as discussed above in more detail to address the question to which extent methodological bias, sample size effects or other factors may cause uncertainty of the parameters used for model parameterisation (and consequently the model predictions) or validation. This can be done using either a quantitative or qualitative approach. While the model output can easily be quantified, the data with which the model results are compared for validation may not be available in sufficient detail to make a quantitative comparison possible. However, a qualitative tabular comparison (model prediction vs. field data) of the output of the model and the values observed in the field may also be appropriate, e.g. addressing uncertainty by showing ranges, means plus standard deviations, confidence intervals or percentiles from different studies and comparing such values with the results from the model.

In Table 2, several parameters are listed which might be used for validation of population models. Implicit parameters are not included. The regulatory important group of terrestrial non-target arthropods is not listed since the possibilities of validation depend very much on the species and their specific features (e.g. developmental stages, modes of dispersal). Population dynamics refer to intra- and inter-annual population fluctuations. Apart from the parameters listed in Table 2, also population genetics can be a tool for validation since the population genetic composition of a population (allele frequencies, formation of genetically distinct local subgroups or subpopulations) is the result of reproduction, survival and dispersal. Strong population fluctuations with marked bottlenecks at some times of the year may, for example, lead to a reduced genetic variability. However, dispersal may help to maintain genetic variability and to produce genetically homogeneous populations. Remarkably, for many species, especially mammals and birds, genetic studies are available, and population genetics has a long history in species conservation, where it is considered a very useful tool [32]. Technically, it is not difficult to include such information in a population model and has been done in some models [33].

Limited knowledge about a species

In most cases, our knowledge about the ecology of a species is limited. Even in the best studied species, there will always be a lack of knowledge, and it is unrealistic

Table 2 Parameters which might be taken into account for population model validation for pesticide risk assessments

Relevant aspects for population development	Aquatic organisms (e.g. crustaceans)	Honeybees	Birds	Mammals
Reproduction	Proportion of adults with eggs Occurrence of ephippia or resting stages Age structure (adults/subadults or moult stages) Onset and speed of recruitment Population growth rates Population dynamics	Number and proportion of eggs, larvae, pupae and adults per season Percentage of drone eggs during the year Hive population dynamics Annual production of honey	Number of clutches per year Number of chicks per lifetime Age structure Proportion of breeding females and males Onset and speed of recruitment Population growth rates Population dynamics	Number of litters per year Number of litters per lifetime Age structure Proportion of breeding females and males Onset and speed of recruitment Population growth rates Population dynamics
Survival	Age structure (adults/subadults) Development of 'new' populations from ephippia Population growth rates Population dynamics	Number and proportion of eggs, larvae, pupae and adults per season Life span of adults Cannibalism of larvae Hive population dynamics Annual production of honey	Age structure Life span Recruitment in the breeding season Population growth rates Population dynamics	Age structure Life span Recruitment in the breeding season Population growth rates Population dynamics
Spatial and other behaviours	Horizontal and vertical distribution Immigration from neighbouring water bodies Response to weather, temperature or environmental changes	Food site selection (e.g. flowering fields) Speed of detection of resources Number of nectar and pollen foragers Time and frequency of foraging trips Influence of weather conditions on foraging activity	Distribution of breeding and foraging territories or home ranges Habitat preference Size of home ranges (per season) Proportion of breeding pairs Dispersal (e.g. after harvesting or high densities)	Distribution of home ranges Habitat preference Habitat specific population densities Size of home ranges (per season) Dispersal (e.g. after harvesting or high densities)

to assume that we will ever know every aspect which may have an influence on the population dynamics or recovery potential of any given species. However, this lack of knowledge is nothing which is specific for population models, but it is equally affecting any other higher tier risk assessments, which are based on field studies or the literature. Even a large-scale field study with many plots may have a limited representativeness and statistical power (for example, see the study of Wang and Rifel [34]). Also, when compiling information from the literature for a particular parameter for use in a risk assessment, it will remain uncertain how representative the values are. Hence, the limited knowledge of the ecology of a species affects not only the realism of model simulations, but also of higher tier risk assessments or the use of field studies in risk assessment. Therefore, for a detailed evaluation of a population model, it may be worth compiling the overall ecology of a species in detail in order to identify possible areas where the lack of knowledge may be problematic. This will also help to understand how realistic a population model is. The lack of knowledge can potentially be a major source of uncertainty; therefore, the potential impact this uncertainty could have on the output of the model needs to be considered.

In some cases, not all data needed for a model may be available, for example, data for a specific geographical region. Hence, a model might, for example, need to be based on data on reproduction from the UK, survival rates from France and habitat preference from Scandinavia. Some data might even be taken from the laboratory, which may sometimes not be easily extrapolated to the field situation [35]. The mixture of data from various origins does not necessarily mean that such a model is unrealistic. However, it should explicitly be discussed if there is evidence that this mixture of data may cause unrealistic model simulations. If field data show that there is no evidence of any geographical trend for a given parameter, then mixing data from different geographic regions may not be problematic. In some cases, it may even be necessary to use data from another species than the modelled one. If a specific behaviour is observed in several related species, then it may be reasonable to assume that this behaviour will also be shown by the modelled species. For example, in several mice and voles [36], 'exploratory excursions' are observed, i.e. short movements outside the home range of an animal with the aim to explore neighbouring habitats. Such excursions are not reported for the common vole (*Microtus arvalis*); however, since exploratory

excursions are common in small mammals, it might be reasonable to assume that they also exist in common voles. This might reduce the vole's susceptibility to herbicides since this behaviour would facilitate movements from areas in which herbicides were applied to areas providing more food.

Taking account of field or lab methodology

When evaluating the realism of a model or performing a validation, a detailed look at the applied methodology for obtaining the data used for parameterisation or validation of a model is required. For example, in birds or mammals, survival rates and population densities are often calculated based on minimum number alive [MNA]. Using this method, animals are captured or trapped in regular intervals, and when an animal is not observed anymore (until the end of the study), it is assumed that this animal has died. However, animals which are not trapped may just have been lucky to miss all traps or may have dispersed. Hence, survival rates based on MNAs are always considerably underestimated. Other capture-recapture methods [CRM] are based on probabilities and include the probability of being captured and of dispersal [37-39]. These methods result in much more accurate estimates of both population size and survival rates. In a study by Höller [40], in which survival rates were calculated with both the MNA method and the CRM, survival rates based on MNA were almost 20% lower than the survival rates based on a CRM. Markedly lower survival rates calculated from MNA compared to rates calculated with a CRM are also reported by Pollock et al. [37]. This demonstrated that field methodology can have a considerable influence on the parameterisation and validation of a population model.

Also, other methodological issues may have a pronounced influence on the results from a laboratory or field study and, hence, also on either the parameterisation or validation of a population model. For example, the size of a trapping grid in a mammal field study may considerably influence population dynamics [41]. The smaller the observed group of animals is, the stronger the stochastic effects will be [42]. For example, in a mammal field survey with a small trapping grid of 100 m², covering much less than a home range size of a wood mouse, only very few animals will be captured. Consequently, the death of a single individual may have already a dramatic effect on the 'observed population'. Hence, rather strong population dynamics will be observed in this study, even though the seemingly dramatic fluctuations are only due to a small trapping grid. In a study with a larger trapping grid, however, much less pronounced fluctuations of population size will be observed since the death or birth of single individuals

has a much smaller effect. In addition, trap type, trap density and study duration may influence trapping success [43,44].

Similarly, the size of a water body or tank in which aquatic organisms are kept during a laboratory or semi-field study determines the amount of population fluctuations which will be found, and it has also been shown [45] that besides carrying capacity, also population growth rate and extinction risk depend on the size of artificial water bodies. Hence, the size of tanks or water bodies needs to be taken into account when using data for validation or parameterisation of a model.

Trapping methods used for arthropods may also include a considerable uncertainty, for example, arthropods being collected using pitfall, yellow sticky or pheromone traps [46]. Other examples are dispersal distances reported for mammals but also for insects. Reported dispersal distances are often limited by the size of the study area. For example, Kölliker-Ott et al. [47] reported dispersal distances in the parasitoid *Trichogramma brassicae* of at most at approximately 8 m. However, no traps were located further away than 8 m.

All these examples stress the importance to carefully consider the methodology used for obtaining field or laboratory data before using them for validation. Without this, a reliable validation is very difficult.

Factors facilitating recovery

The main reason why natural populations can recover is that several mechanisms exist, which help to stabilise populations. Recovery of populations can be caused by various, very different biological mechanisms including intrinsic population recovery, e.g. by increased reproduction or reduced mortality, and recolonisation [48-50].

Each of these factors can be regulated in a rather complex, density dependent way. For example, in mammals, reproduction may be reduced at high density by either reduced litter size, resorption of embryos, complete loss of litters or delayed implantation [51]. When density is low, the absence of all these effects may lead to a rapid increase of population size. However, for a particular species, not all of these processes may be known to occur. One may, for example, have detailed information of survival rates at different densities, indicating that at higher densities, survival is reduced. However, regarding other processes, there may be observations indicating that litters are resorbed at high density, but without giving sufficient detail for including this in a model. Consequently, the model may produce a more conservative output, i.e. underestimate the ability to recover, since one of the existent 'stabilising' factors would not be included.

In order to evaluate if a model is conservative or realistic, it might help to list all known processes which

may act as regulatory mechanisms facilitating recovery and to indicate which of these processes are included in the model and which are not. Then, it is easy to see if real populations can presumably recover more or less easily as the modelled populations. Major processes which may act as regulatory mechanisms in several taxonomic groups are summarised in Table 3. Such processes will facilitate recovery after impacts by agricultural practise, pesticide application or other impacts.

Representativeness

For risk assessors, a technical description of a model and the knowledge that the model produces a realistic output are not sufficient. They also have to know what a model represents. The representativeness of population models for use in risk assessment has so far received little or no focus, although it is crucial to know if a model represents the behaviour of aquatic organisms in, for example, a small artificial water tank in the laboratory or the aquatic populations in a ditch on the side of a field. Similarly, if a model of a bird species was based on data from Germany, the model might or might not be representative for the UK. The description of what a model represents is important since parameters in nature vary geographically or depend on habitat type. For example, in several birds, clutch size varies considerably depending on many different factors, such as habitat type, altitude or food availability [52]; the onset and length of the breeding season may vary according to latitude and climate conditions [52] or the existence and magnitude of population cycles (e.g. in voles, the presence of cycles seems to depend on the geographical location of populations, specifically latitude, [53]). Also, population parameters, such as population growth rates, can differ considerably between different natural populations [18].

For population models of aquatic organisms, it is crucial to know what a model represents too. Williams

et al. [54] and Meester et al. [55] have shown that different types of water bodies (e.g. ponds, streams, etc.) may show a remarkably different species composition and abundance in both plant and invertebrate species. Additionally, Williams et al. [54] have also even found a considerable variation between different water bodies of the same type (e.g. ponds) and identified several factors which influence species diversity, such as shading, seasonality, altitude, water depth and several other factors.

Without the knowledge about such types of variations in nature, one might apply a model inappropriately for a situation which the model may not realistically reproduce. Therefore, it is important to describe which situation a model represents and what a model is actually able to reproduce. In addition, it may be worth to state if there is evidence for any type of variability which is not included in the model. For example, one might construct a model using survival rates from Poland and validating the model based on population dynamics in Germany. If there is no evidence for a geographical variation of survival rates or population dynamics, such an approach may result in a realistic model. However, if there is evidence that population density is much higher in one country than in the other (this is the case, e.g. in the Grey Partridge with an almost ten times larger breeding population in Poland compared to that in Germany, [56]) a model may not be adequate for Germany or Poland. For risk assessors, it might therefore be helpful to systematically describe for all major areas, i.e. survival, reproduction, etc., where data were taken from, what they are representative for and if there is any evidence for geographical or other variation. If variation is present but not included in a model, it might be helpful to state if this will result in a more or less conservative estimation of, for example, recovery. A (shortened) example of how the representativeness of a model could be described is shown in Table 4. In real applications of

Table 3 Processes which may regulate population density in several taxonomic groups

Relevant aspects for population development	Aquatic organisms (e.g. crustaceans)	Honeybees	Birds	Mammals
Reproduction/development	Juvenile production/production of ehippia	Removal of larvae at low food Production of workers/drones Earlier development of foragers Swarming (division of colonies)	Reduced number of clutches Missed breeding season Nest abandonment Hatch success Clutch size	Delayed implantation Resorption of litters Reduced litter size Age at weaning Delayed sexual maturity Loss of litter/embryo survival
Survival	Survival Life span	Survival of larvae/adults Life span	Chick/adult survival Survival of dispersers Life span	Juvenile/adult survival Survival of dispersers Life span
Spatial and other behaviours		Swarming (dispersal by swarming)	Giving up of home ranges Disperal	Giving up of home ranges Disperal

Table 4 Example description of representativeness of different aspects in a population model of the common vole

Relevant aspects for population development	Data source	Representativeness	Evidence for geographical or other sources of variation	Remarks
Reproduction (sex. maturity, gestation and lactation length, litter size)	Various habitats and laboratory studies	Generally representative for the common vole; the breeding season length used in the model is typical for Central Europe but may in reality be longer or shorter	Generally no evidence for geographical or other variability within Central Europe, but breeding season length varies depending on latitude and altitude and between years	Data about sexual maturity of male common voles was not available; therefore, data for <i>Microtus agrestis</i> were used
Survival (survival rates, juvenile survival)	Mixed arable land in Poland and Germany	Mixed arable land in Central Europe	No evidence for regional variation available. Theoretically, influence of climate or predator abundance possible	Data from Poland were chosen for the model since they were based on large sample sizes and since survival rates were calculated using very reliable methods (CRM survival rates). Of all available data, these data were the most reliable ones. Data from Germany were used for habitat-specific survival
Spatial behaviour (home ranges, dispersal)	Various habitats in Germany	Central Europe	No evidence for geographical variation and no mechanistic reason for such variation evident	Home range sizes in the model depend on food availability. In low food habitats, home ranges increase in size

Adapted from the study of M Wang (unpublished work).

population models, further factors would be included, such as population density or dynamics.

Scenarios and protection goals

Scenarios describe the environment which is considered in simulations with a population model, e.g. a specific landscape with a given landscape structure or the dimensions and the location of a surface water body which are considered for modelling aquatic organisms. Also, the timing of the application of a pesticide can be considered as being part of a scenario.

For population-level risk assessments, it is evidently impossible to consider every possible landscape or environment for simulations in a population model. Therefore, the most practical way would be to run simulations for a few selected 'scenarios', which function as surrogates for the entire agricultural area. A pragmatic approach as used in the evaluation of the environmental fate of pesticides could be to produce a set of scenarios which aim to cover 90% of the agricultural area in the EU [57,58].

The selection of such scenarios is critical since, for example, landscape structure may have an important influence on the occurrence of species and also on recovery. For example, Wang and Grimm [5] have demonstrated for the common shrew (*Sorex araneus*) that in a landscape with hedges, recovery is considerably faster than recovery in a landscape without hedges. Also, the timing of application had a remarkable influence on shrew populations: while populations always recovered after repeated application of a hypothetical pesticide in spring, populations became extinct when

the pesticide was applied in summer. Also, Galic et al. [59] analysed the effects of landscape structure on invertebrate populations and found that size and connectivity of habitat patches (source habitats) influenced the recovery of neighbouring depleted areas by immigrating individuals in a model population.

In principle, the selection of scenarios is a question of the protection goal and concerns the questions 'what shall be protected' and 'how protective' the assessment should be. The European Food Safety Authority [EFSA] [60] have proposed to define protection goals based on an 'ecosystem service concept', in which the 'benefit' of a group of animals for society is assessed (e.g. genetic resources, education and inspiration, aesthetic values). Subsequently, the EFSA [60] proposed to define protection goals by the ecological entity (e.g. individual or population), by the attribute (e.g. lethality or abundance) and by three types of scale: magnitude of impact, spatial scale of impact and temporal scale of impact. For vertebrates, the proposed protection goals are, for instance, 'no decline in biodiversity' and 'negligible effects on population structure'. The focus on the protection on the population level may seem to be a clearly defined goal [61]. However, the term 'population' is still not clearly defined in present guidelines. Clarification would be necessary since the definition of what a population is in biology is as controversial as the species definition, and different views exist, depending on the focus on ecology, genetics or evolution [62-64]. The term population is sometimes even used interchangeably with species [64]. The most common definition, which is based on the definition by Mayr [65], is that animals in a

population interbreed at the same time and have a level of genetic similarity. Apart from the issue to define a population, which could be solved by population genetics, another question is where to protect the animals. EFSA [60] proposes to focus on the (edge-of) field or landscape (depending on the home range of a species). One option could be to protect animals in their typical habitats, e.g. to protect wood mice in landscapes with forests, fields and hedges, the habitats where they typically occur [66]; and hares in open landscapes predominated by large fields [67]. That would imply that wood mice would not be considered in risk assessments for landscapes with a low proportion of forests or hedges. On the other hand, if a few wood mice would live in such unfavourable habitats, their small 'stock' (to avoid the word population) would evidently be more susceptible to effects by agricultural practise or pesticides. However, the few animals living in such areas might or might not represent a noticeable fraction of the overall wood mouse population and might possibly disperse to more adequate areas whenever available. This demonstrates the controversy when discussing protection goals. A practical solution might be to protect 'typical' or 'permanent' occurrences only, excluding temporary occurrences of a species.

Endpoints and measurement of effects

Only few studies have focussed on the question how population-level effects or recovery may be measured in a population model [61]. The easiest way may be to assess population size before and after a pesticide application and to evaluate if any differences are observed. If reduced population size is observed after a pesticide application, then one might measure the time needed until the 'normal' population size is reached to estimate the time until recovery. However, it has been rarely investigated how to evaluate if a normal population size is reached after an effect or if population size is the most suitable 'endpoint' for population-level risk assessment. A systematic comparison of possible endpoints (population density and population growth rate) for use in population-level risk assessments has been published by Wang and Grimm [5,46] (but see also the book of Barnthouse et al. [61] for a discussion of further endpoints). The authors applied a population model for the common shrew (*S. aranaeus*) and showed that growth rate was only temporarily affected when applying an additional mortality on the populations (simulating the impact by a hypothetical pesticide), i.e. population growth rate was reduced only at the time of the impact. However, after application, population growth rate did not differ from the growth rate in control simulations, suggesting a very fast recovery. In contrast, when focussing on population density as an endpoint to measure

population-level effects, simulated treatment populations differed significantly from control populations until several months after treatment. These results demonstrated that population density was a much more sensitive endpoint for population level risk assessments. Population size or density is also the ecologically most reasonable parameter since it's the number of animals which provides food for predators, reduces food resources (plant parts, seeds, etc.) or disperses seeds. In contrary, population growth rate does not have a direct influence on any of these factors.

Apart from the selection of endpoints, it has to be defined how to detect effects in a population model or more specifically, how to decide if an effect exists or not. In common field trials, the presence of an effect is currently determined based on a statistical evaluation ('is the survival in the treatment significantly lower than that in the control'). A statistical evaluation of experimental trials is the only reasonable way since sample size is usually limited, and one has to know if a difference between the control and the treatment is caused by a real effect or by sample size effect. In population models, however, the number of simulations can be increased with relatively little effort. Therefore, even the smallest differences could be detected with a high number of simulations by a statistical test. Whether minute differences of population size are biologically relevant is debatable. One practical approach to make model simulations comparable to field studies could be to limit the number of simulations to a sample size which is comparable to the one of field studies (for example, see the study of Wang and Grimm [5]). Alternative approaches would be to use an accepted threshold or trigger for defining effects or recovery. Recovery could then be assumed if the population density consistently reaches, for example, 95% of the density of the control simulations. While the use of a threshold may seem to be an easy and straightforward approach, the amount of effect being acceptable will vary for each species. Therefore, the use of a statistical test using a number of simulations which are comparable to field studies may be the better choice. However, since the performance of a statistical test depends largely on sample size [34], the choice of the number or simulations used would be a choice of which magnitude of effect one wishes to detect.

Conclusions

Population models can provide a useful tool in risk assessment. However, for creating confidence in such models and to make them applicable for risk assessors, several issues should be addressed starting with a detailed technical description of the model, an evaluation of the realism of a model (validation), the description of what a model represents and which scenarios are

considered in a final population-level risk assessment. For covering all these aspects, a thorough knowledge of the wildlife ecology and methodology is essential. However, even after a thorough validation and an evaluation of the representativeness of a model, some degree of uncertainty regarding the prediction will always remain since population models are and will always be based on a limited amount of knowledge. However, the same applies for virtually any risk assessment or field study. Also, higher tier risk assessments are typically based on a limited set of data from the literature, and usually much less literature is reviewed for a risk assessment (depending on what is refined) than for a population model. In a best case, a population model may even reflect all our current knowledge of a species. Also, when performing field studies for answering a specific risk-related question, there is uncertainty. Usually, a study is conducted over a limited time period in a specific geographical area with a limited number of plots. Hence, there is an uncertainty if results from this study are representative for other areas or years (with different population densities, weather conditions or food abundance). In addition, sample size and the number of plots or replicates can cause a considerable amount of uncertainty. Compared to higher tier risk assessments or field studies, population models do also offer the advantage that they are flexible. Instead of focusing on a single field study, they could be used to run several different scenarios, which reflect the conditions in different geographical areas or landscape structures. Also, the detectability of effects can easily be increased by running more simulations (sample size) or increasing the plot or landscape area that is simulated. The final questions which cannot be solved by modelling ecologists but only by risk assessors is the definition of adequate scenarios, acceptable levels of recovery and the measurement of population-level effects.

Author details

¹RIFCON GmbH, Im Neuenheimer Feld 517, Heidelberg, 69120, Germany

²National Institute for Public Health and the Environment, P.O. Box 1, Bilthoven, 3720 BA, The Netherlands

Authors' contributions

Both authors participated in the drafting of the manuscript and approved its final version.

Competing interests

The authors declare that they have no competing interests.

Received: 5 September 2011 Accepted: 17 January 2012

Published: 17 January 2012

References

1. Pastorok RA, Bartell SM, Ferson S, Ginzburg LR: *Ecological Modeling in Risk Assessment - Chemical Effects on Populations, Ecosystems, and Landscapes* Boca Raton: Lewis; 2002.

2. Pastorok RA, Akcakaya HR, Regan H, Ferson SM, Bartell SM: **Role of ecological modeling in risk assessment.** *Hum Ecol Risk Assess* 2003, **9**:939-972.
3. Topping CJ, Dalkvist T, Forbes VE, Grimm V, Sibly RM: **The potential for the use of agent-based models in ecotoxicology.** In *Ecotoxicology Modeling*. Edited by: Devillers J. New York: Springer; 2009:205-237.
4. Forbes VE, Hommen U, Thorbek P, Heimbach F, van den Brink P, Wogram J, Thulke HH, Grimm V: **Ecological models in support of regulatory risk assessments of pesticides: developing a strategy for the future.** *Integr Environ Assess Manag* 2009, **5**:167-172.
5. Wang M, Grimm V: **Population models in pesticide risk assessment: lessons for assessing population-level effects, recovery, and alternative exposure scenarios from modelling a small mammal.** *Environ Toxicol Chem* 2010, **29**:1292-1300.
6. Kramer VJ, Etterson MA, Hecker M, Murphy CA, Roesijadi G, Spade DJ, Spromberg JA, Wang M, Ankley GT: **Adverse outcome pathways and ecological risk assessment: bridging to population-level effects.** *Environ Toxicol Chem* 2011, **30**:64-76.
7. Thorbek P, Forbes V, Heimbach F, Hommen U, Thulke HH, van den Brink PJ, Wogram J, Grimm V: *Ecological Models for Regulatory Risk Assessments of Pesticides: Developing a Strategy for the Future* Boca Raton: CRC; 2010.
8. Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz SK, Huse G, Huth A, Jepsen JU, Jørgensen C, Mooij WM, Müller B, Pe'er G, Piu C, Railsback SF, Robbins AM, Robbins MM, Rossmanith E, Rügen N, Strand E, Souissi S, Stillman RA, Vabø R, Visser U, DeAngelis DL: **A standard protocol for describing individual-based and agent-based models.** *Ecological Modelling* 2006, **198**:115-126.
9. Schmolke A, Thorbek P, DeAngelis DL, Grimm V: **Ecological models supporting environmental decision making: a strategy for the future.** *Trends in Ecology and Evolution* 2010, **25**:479-486.
10. Grimm V, Railsback SF: *Individual-based Modeling and Ecology* Princeton: Princeton University Press; 2005.
11. Topping CJ, Høyea TT, Olesena CR: **Opening the black box - development, testing and documentation of a mechanistically rich agent-based model.** *Ecological Modelling* 2009, **221**:245-255.
12. Forbes VS, Hommen U, Thorbek P, Heimbach F, Van den Brink PJ, Wogram J, Thulke H-H, Grimm V: **Ecological models in support of regulatory risk assessments of pesticides: developing a strategy for the future.** *Integr Environ Assess Manag* 2009, **5**:167-172.
13. Jakeman AJ, Letcher R, Norton J: **Ten iterative steps in development and evaluation of environmental models.** *Environ Model Softw* 2006, **21**:602-614.
14. Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, Thulke H-H, Weiner J, Wiegand T, DeAngelis DL: **Pattern-oriented modeling of agent-based complex systems: lessons from ecology.** *Science* 2005, **310**:987-991.
15. Andrewartha HG, Birch LC: *The Distribution and Abundance of Animals* Chicago, USA: University of Chicago Press; 1954.
16. Strong DR: **Density vagueness: abiding the variance in the demography of real populations.** In *Community Ecology*. Edited by: Diamond J, Case TJ. Cambridge: Harper and Row; 1986:257-268.
17. Turchin P: **Population regulation: old arguments and a new synthesis.** In *Population Dynamics: New Approaches and Synthesis*. Edited by: Cappuccino N, Price PW. San Diego: Academic Press; 1995:19-39.
18. Clark JS: **Uncertainty and variability in demography and population growth: a hierarchical approach.** *Ecology* 2003, **84**:1370-1381.
19. Hart A: **The EUFRAM framework - introducing probabilistic methods into the ecological risk assessment of pesticides.**[<http://www.eufram.com>].
20. EFSA (European Food Safety Authority): **Guidance document on risk assessment for birds and mammals.** *EFSA Journal* 2009, **7**:1438-1577.
21. Oreskes N, Shraderfrechette K, Belitz K: **Verification, validation, and confirmation of numerical models in the earth-sciences.** *Science* 1994, **263**:641-646.
22. Rykiel EJ: **Testing ecological models: the meaning of validation.** *Ecol Modelling* 1996, **90**:229-244.
23. Conroy MJ, Cohen Y, James FC, Matsinos YG, Maurer BA: **Parameter estimation, reliability, and model improvement for spatially explicit models of animal populations.** *Ecol Appl* 1995, **5**:17-19.
24. Grimm V: **Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?** *Ecol Modelling* 1999, **15**:129-148.

25. Schmolke A, Thorbek P, Chapman P, Grimm V: **Ecological models and pesticide risk assessment: current modeling practice.** *Environ Toxicol Chem* 2010, **29**:1006-1012.
26. Wang M, Grimm V: **Home range dynamics and population regulation: an individual-based model of the common shrew.** *Ecological Modelling* 2007, **205**:397-409.
27. Preuss T, Hammerswirth M, Hommen U, Rubach M, Ratte H: **Development and validation of an individual based *Daphnia magna* population model: the influence of crowding on population dynamics.** *Ecological Modelling* 2009, **220**:310-329.
28. Bart J: **Acceptance criteria for using individual-based models to make management decisions.** *Ecol Appl* 1995, **5**:411-420.
29. Guisan A, Zimmermann NE: **Predictive habitat distribution models in ecology.** *Ecol Modelling* 2000, **135**:147-186.
30. Grimm V, Franke K, Jeltsch F, Brandla R, Uchmariskib J, Wissela C: **Pattern-oriented modeling in population ecology.** *The Science of the Total Environment* 1996, **183**:151-166.
31. Wiegand T, Jeltsch F, Hanski I, Grimm V: **Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application.** *Oikos* 2003, **100**:209-222.
32. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C: *Population Genetics for Animal Conservation* New York: Cambridge University Press; 2009.
33. Pertoldi C, Topping C: **The use of agent-based modelling of genetics in conservation genetics studies.** *J Nat Cons* 2004, **12**:111-120.
34. Wang M, Riffel M: **Making the right conclusions based on wrong results and small sample sizes: interpretation of statistical tests in ecotoxicology.** *Ecotoxicol Environ Saf* 2011, **74**:684-692.
35. Seitz A, Ratte H: **Aquatic ecotoxicology: on the problems of extrapolation from laboratory experiments with individuals and populations to community effects in the field.** *Comp Biochem Physiol* 1991, **100C**:301-304.
36. Wolton RJ: **The ranging and nesting behaviour of wood mice, *Apodemus sylvaticus* (Rodentia: Muridae), as revealed by radio tracking.** *J Zool Lond* 1985, **206**:203-224.
37. Pollock KH, Nichols JD, Brownie C, Hines JE: **Statistical inference for capture-recapture experiments.** *Wildlife Monographs* 1990, **107**:1-97.
38. Lebreton J-D, Burnham KP, Clobert J, Anderson DR: **Modelling survival and testing biological hypothesis using marked animals: a unified approach with case studies.** *Ecol Monogr* 1992, **62**:67-118.
39. Schwarz CJ, Seber GA: **Estimating animal abundance: review III.** *Statistical Science* 1999, **14**:427-456.
40. Höller D: **The population dynamics and individual winter survival of rodents in a rich structured farmland in Western Germany.** *Master thesis* University of Bonn, Germany; 2002.
41. Moreau F: **A comparison of trapping methodologies and grid size in for small mammals research.** *Master thesis* University of British Columbia, Canada; 2000.
42. Bondrup-Nielsen S: **Density estimation as a function of live-trapping grid and home range size.** *Can J Zool* 1983, **61**:2361-2365.
43. Burger JR, Chesh AS, Castro RA, Tolhuysen LO, Torre I, Ebensperger LA, Hayes LD: **The influence of trap type on evaluating population structure of the semifossorial and social rodent *Octodon degus*.** *Acta Theriol* 2008, **54**:311-320.
44. Conard JM, Baumgardt JA, Gipson PS, Althoff DP: **The influence of trap density and sampling duration on the detection of small mammal species richness.** *Acta Theriol* 2008, **53**:143-156.
45. Griffen BD, Drake JM: **Effects of habitat quality and size on extinction in experimental populations.** *Proc R Soc B* 2008, **275**:2251-2256.
46. New TR: *Invertebrate Surveys for Conservation* Oxford: Oxford University Press; 1999.
47. Kölliker-Ott UM, Bigler F, Hoffmann AA: **Field dispersal and host location of *Trichogramma brassicae* is influenced by wing size but not wing shape.** *Biol Control* 2004, **31**:1-10.
48. Boer PJ: *Regulation and Stabilization Paradigms in Population Ecology* London: Chapman & Hall; 1996.
49. Gardmark A, Enberg K, Ripa J, Laakso J, Kaitala V: **The ecology of recovery.** *Ann Zool Fenn* 2003, **40**:131-144.
50. Chapman JL, Reiss MJ: *Ecology: Principles and Applications* Cambridge: Cambridge Univ Press; 1999.
51. Flowerdew JR: *Mammals: Their Reproductive Biology and Population Ecology* London: Edward Arnold Pbs. Ltd; 1987.
52. Bairlein F: *Ökologie der Vögel* Stuttgart - Jena - Lübeck - Ulm: Gustaf Fischer Verlag; 1996.
53. Tkadlec E, Stenseth NC: **A new geographical gradient in vole population dynamics.** *Proc R Soc Lond B, Biol Sci* 2001, **268**:1547-1552.
54. Williams P, Whitfield M, Biggs J, Bray S, Fox G, Nicolet P, Sear D: **Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in southern England.** *Biological Conservation* 2004, **115**:329-341.
55. Meester L De, Declerck S, Stoks R, Louette G, Van de Meutter F, De Bie T, Michels E, Brendonck L: **Ponds and pools as model systems in conservation biology, ecology and evolutionary biology.** *Aquatic Conservation: Marine and Freshwater Ecosystems* 2005, **15**:715-725.
56. Burfield I, van Bommel F: *Birds in Europe: Population Estimates, Trends and Conservation Status.* *BirdLife Conservation Series 12* Cambridge: BirdLife International; 2004.
57. FOCUS 2000: **FOCUS groundwater scenarios in the EU review of active substances.** *Report of the FOCUS Groundwater Scenarios Workgroup 202*, EC Document Reference SANCO/321/2000 rev.2.
58. FOCUS 2001: **FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC.** *Report of the FOCUS Working Group on Surface Water Scenarios 245*, EC Document Reference SANCO/4802/2001-rev.2.
59. Galic N, Baveco JM, Hengeveld GM, Thorbek P, Bruns E, van den Brink PJ: **Simulating population recovery of an aquatic isopod: effects of timing of stress and landscape structure.** *Environmental Pollution* .
60. EFSA Panel on Plant Protection Products and their Residues (PPR): **Scientific opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the guidance documents on aquatic and terrestrial ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002).** *EFSA Journal* 2010, **8**:1821-1876.
61. Barnhouse LW, Munns WR, Sorensen MT: *Population-level Risk Assessment* Boca Raton: Taylor & Francis; 2007.
62. Templeton AR: **The meaning of species and speciation: a genetic perspective.** In *Speciation and its Consequences*. Edited by: Otte D, Endler J. Sunderland: Sinauer; 1989:3-27.
63. Godfrey-Smith P: *Darwinian Populations and Natural Selection* Oxford: Oxford University Press; 2009.
64. Wilkins JS: **What is a species? Essences and generation.** *Theory in Biosciences* 2010, **129**:141-148.
65. Mayr E: *Systematics and the Origin of Species from a Viewpoint of a Zoologist* Harvard: Harvard University Press; 1942.
66. Tattersall FH, Macdonald DW, Hart BJ, Manley W: **Balanced dispersal or source-sink - do both models describe wood mice in farmed landscapes?** *Oikos* 2004, **106**:536-550.
67. Corbet GB, Harris S: *The Handbook of British Mammals*. 3 edition. Oxford: Blackwell Scientific Publications; 1991.
68. Reichstein H: **Untersuchungen zum Körperwachstum und zum Reproduktionspotential der Feldmaus.** *Z Wiss Zool* 1964, **170**:112-222.
69. Glutz von Blotzheim UN: *Handbuch der Vögel Mitteleuropas. Bd. 9* Wiesbaden: Aula Verlag; 1980.
70. Labitte A: **Notes biologiques sur la Chouette chevêche, *Carine noctua vidalii* A.E. Brehm, 1857.** *L'Oiseau et R.F.O* 1951, **21**:120-126.
71. Tomé R, Catry P, Bloise C, Korpimäki E: **Breeding density and success, and diet composition of Little Owls *Athene noctua* in steppe-like habitats in Portugal.** *Ornis Fennica* 2008, **85**:22-32.

doi:10.1186/2190-4715-24-3

Cite this article as: Wang and Luttik: Population level risk assessment: practical considerations for evaluation of population models from a risk assessor's perspective. *Environmental Sciences Europe* 2012 **24**:3.