Ecotoxicology (2006) 15: 305–314
DOI 10.1007/s10646-006-0060-x

Making sense of ecotoxicological test results: towards application of process-based models

Tjalling Jager · Evelyn H. W. Heugens · Sebastiaan A. L. M. Kooijman

Accepted: 6 February 2006 / Published online: 20 April 2006

Abstract The environmental risk of chemicals is routinely assessed by comparing predicted exposure levels to predicted no-effect levels for ecosystems. Although process-based models are commonly used in exposure assessment, the assessment of effects usually comprises purely descriptive models and rules-of-thumb. The problems with this approach start with the analysis of laboratory ecotoxicity tests, because only a limited amount of information is extracted. Standard summary statistics (NOEC, EC*x*, LC50) are of limited use in part because they change with exposure duration in a manner that varies with the tested species and the toxicant. As an alternative, process-based models are available. These models allow for toxicity measures that are independent of exposure time, make efficient use of the available data from routine toxicity tests, and are better suited for educated extrapolations (e.g., from individual to population, and from continuous to pulse exposure). These capabilities can be used to improve regulatory decisions and allow for a more efficient assessment of effects, which ultimately will reduce the need for animal testing. Process-based modeling also can help to achieve the goals laid out in REACH, the new strategy of the European Commission in dealing with chemicals. This discussion is illustrated with effects data for *Daphnia magna*, analyzed by the DEBtox model.

Keywords Effects assessment · Toxicity testing · Dose-response modeling · DEBtox · REACH

T. Jager (✉) · S. A. L. M. Kooijman
FALW/Department of Theoretical Biology, Vrije Universiteit Amsterdam, De Boelelaan 1085, NL-1081 HV Amsterdam, The Netherlands
e-mail: tjalling@bio.vu.nl

E. H. W. Heugens
Department of Aquatic Ecology and Ecotoxicology, Faculty of Science, University of Amsterdam, P.O. Box 94084, NL-1090 GB Amsterdam, The Netherlands

Introduction

As presently conducted, environmental risk assessment for the regulation of chemicals compares exposure levels to no-effect levels for ecosystems. In the European risk assessment guidelines (EC, 2003), the predicted environmental concentration (PEC) is divided by the predicted no-effect concentration (PNEC), and the resulting ratio forms the basis for decisions regarding the acceptability of a compound. Input for this assessment is formed by a limited dataset, delivered by the manufacturer of the chemical.

For exposure assessment, the dataset contains physico-chemical properties of the chemical in question, such as hydrophobicity, water solubility, as well as results from (bio-)degradation tests. To calculate the PEC, process-based fate models are commonly used. In such models, the available data are used to estimate environmental process parameters such as volatilization rates, degradation rates and partition coefficients. Using these models, it is possible to simulate different release scenarios without actually measuring concentrations. For effects assessment, the available data in the initial tiers typically include the results from highly standardized toxicity tests, reported as NOECs, LC50s or EC50s (after a specified exposure time). NOECs are derived using a statistical hypothesis test (comparison of effects to the control response), whereas the other

summary statistics are interpolated values based on a purely descriptive regression model. More recently, the EC*x* has been proposed as an alternative for the NOEC (e.g., Van der Hoeven, 1997), where *x* is a small percentage (usually 5% or 10%).

Unlike exposure assessment, the method to derive the PNEC does not use process-based models; the available data are not used to estimate ecotoxicological process parameters. In fact, the approach relies on the use of arbitrary safety factors (generally multiples of 10, see Chapman et al., 1998, for historical overview and discussion) to derive a predicted no-effect concentration (PNEC) as ''a concentration below which an unacceptable effect will most likely not occur'' (EC, 2003). The safety factors should reflect our ignorance about the translation from laboratory tests (short-term, high exposure, one species, and controlled environment) to the field (long-term, low exposure, multiple species, variable environment). Although ecotoxicological experience and common sense forms the basis of these factors, the scientific rigor of the extrapolation from test result is in marked contrast with the process-based models used for the exposure assessment. More recently, species-sensitivity distributions (SSD; see Posthuma et al., 2002) have been proposed as an alternative extrapolation from single species to ecosystems. Although the SSD improves effects assessment in some aspects, serious critique has been raised (Forbes and Forbes, 1993; Smith and Cairns, 1993).

It may be questioned whether we have enough knowledge to extrapolate laboratory test results to field populations. However, we will argue that additional information can be extracted from the same standard ecotoxicological laboratory tests, and that even more progress can be made by extending the standard test protocols. The current methods for analyzing toxicity test results are largely descriptive, and may be based on flawed assumptions; serious objections have been raised against the NOEC (Laskowski, 1995; Crane and Newman, 2000), as well as against the regression-based statistics such as LC50 and EC*x* (Kooijman, 1996). Moreover, these statistics are inefficient in that they do not incorporate all of the data from the toxicity test, and use only the results at the end of the prescribed exposure time. Although these summary statistics form the basis of virtually every effects assessment, the uncertainties they include may well compromise any method for PNEC extrapolation. Quantitative knowledge on the processes underlying the toxic effect is required for an educated extrapolation to the protection goal of ecological risk assessment (generally, field populations or ecosystems). For this reason, alternatives to the statistical methods currently used for the analysis of ecotoxicity tests should be considered. Here, we discuss some shortcomings of current procedures and provide an example of the type of process-based and integrated approach that could be used to replace them.

We restrict our discussion to the derivation of more meaningful information from standard ecotoxicity tests. To illustrate this discussion, we make use of a particular process-based method, DEBtox (Kooijman and Bedaux, 1996a), and use data for a well-known toxicant (cadmium) and a well-known test species (*Daphnia magna*). We are aware that there is a large body of literature on ecological risk assessment that includes process-based approaches to estimate population and food-web effects (see e.g., Bartell et al., 1992), but a thorough review of this literature is outside the scope of the present paper.

Limitations of current procedures

The NOEC has been criticized for being a fundamentally invalid interpretation of hypothesis testing (Laskowski, 1995; Van der Hoeven, 1997; Crane and Newman, 2000). ''No statistically significant effect'' does not mean that there is no effect. In fact, the effect at the NOEC is often 10–34% and in extreme cases can approach 100% (Crane and Newman, 2000). Even though the OECD recommends not using this statistic (OECD, 1998), the NOEC is still used frequently in risk assessment frameworks. The EC*x* (and related statistics, such as EC50 and LC50) is a more robust summary statistic because it, with its confidence interval, results from a curve fitted to all of the data (at a single time point). The log-logistic curve is a popular choice, but there is no biological or toxicological reason to prefer one specific curve over another: the only basis for selection can be goodness-of-fit. For lethality, the curve often is thought to represent the distribution of the sensitivities in the tested population. However, this assumption can be questioned because mortality appears to be largely stochastic at the level of the individual (Kooijman, 1996; Newman and McCloskey, 2000). Regression models are in themselves defensible if the only aim is to estimate the concentration related to a specific quantity of effect at a given time, under two conditions: (1) that the model is used only to interpolate, and (2) the data should fit the curve well. However, a robust estimate of EC*x* requires several concentrations with partial effects, which may be difficult to achieve in practice. The main problem with the regression approaches is their descriptive character: no attempt is made to understand the processes behind the effect. As a result, it is virtually impossible to make consistent use of multiple measurements through time, or of additional sources of information that could help interpret the toxic response (e.g., measurements of internal concentrations or body size). Furthermore, it is impossible to translate the observed effects to other species (e.g., from species used in

laboratory tests to specific species of interest), or to field circumstances such as varying exposure concentrations, differences in temperature, limiting food supply, etc.

Several examples may illustrate limitations of the current standard test protocols and the statistical methods to analyze data obtained using these protocols. The duration of an ecotoxicological test is prescribed by the test protocol. For example, an acute *Daphnia* test takes 2 days, and an acute fish test 4 days. However, LC50s generally decrease through time, because it takes time for the chemical to accumulate in the organism's tissue and exert its toxic effect. The toxicokinetics depend on properties of the organism (especially the surface area/volume ratio) as well as on properties of the chemical (e.g., its hydrophobicity) so it makes little sense to focus on an arbitrarily standardized time period only. This is not a new insight. As early as 1969, Sprague (1969) recognized that a single exposure time for all chemicals could not be specified, and advised continuing an acute test until mortality ceased (yielding the so-called incipient LC50). For some chemicals, the standard test durations may be sufficient to yield an incipient LC50. For others though, it may not, and this imposes severe restrictions on the ability to compare toxicity of different chemicals. Furthermore, the standard test duration for *Daphnia* is relatively long (given its small size and large surface/volume ratio) compared to the duration of the acute fish test. Thus, *Daphnia* are most likely closer to equilibrium than are fish, exposed to the same chemical. As a result, the standardized exposure time actually hampers comparison of the sensitivity of these test species. In chronic tests, the summary statistics also will change through time, although the response may differ between endpoints and the EC*x* may not decrease through time, like the LC50 (as will be demonstrated later in this paper). Clearly, time is an important factor in toxicity tests and simple standardization can restrict the interpretation of the test results and, subsequently, affects the assessment of risk.

Another problem is that current test protocols do not consider the biological background of the endpoint being studied. A NOEC for fish growth, for example, is treated the same as a NOEC for *Daphnia* reproduction: the lowest value of both NOECs is used for PNEC derivation. However, a 10% effect on body size is not equivalent to a 10% effect on reproduction. In many cases, a small effect on growth will be accompanied by a larger effect on reproduction. In extreme cases, a chemical may completely suppress reproduction yet have no effect on body size (for example, chlorpyrifos on springtails, Crommentuijn et al., 1997). Therefore, NOEC or EC*x* values for different endpoints cannot be directly compared.

These two problems, the time dependency of the current test summaries, and failing to consider the biological background, preclude useful extrapolation to field populations. Furthermore, it makes little sense to compare differences in the apparent ''sensitivities'' of different species, or differences in the ''toxicities'' of different chemicals. Apparent differences in sensitivity or toxicity may simply reflect the (arbitrary) choices in testing protocols. The inability to properly compare species using the traditional summary statistics is a serious reason for not using SSDs. If the toxicity of chemicals cannot be compared, priority-setting and comparative risk assessment (one of the explicit aims of the new EU chemical policy REACH, Christensen et al., 2003) are rendered impossible. These problems with current practices were recognized at the 1996 OECD workshop in Braunschweig, Germany (OECD, 1998) on the statistical analysis of aquatic ecotoxicity data. At this workshop, it was concluded that it was time to move away from the NOEC, towards process-based approaches that deal explicitly with exposure time and allow for links to population models.

Process-based analyses

A process-based analysis can address some of the limitations of the currently used statistical analyses of ecotoxicological tests. A viable alternative should provide a process-based description of mortality, growth and reproduction within a single consistent framework. Dealing with sub-lethal effects, it is essential to have a model capable of showing the relations between feeding, maintenance, growth, development and reproduction. This approach allows questions to be rephrased in a more meaningful framework. For example, if a chemical decreases reproduction, the question could be, not how to derive an EC*x*, but rather, why is reproduction decreasing. This question can be approached from the perspective of energy balance; offspring are produced from food, so one possibility is that ingestion has declined as a result of toxic stress. Alternatively, feeding may be unaltered, but there may be an additional energy drain for metabolic repair, leaving less energy for the production of eggs. Other hypotheses as to the observed reproductive effects can be put forward, obviously. The theory of dynamic energy budgets (DEB; see Kooijman, 2000, 2001; Nisbet et al., 2000) provides a framework for addressing this question: the DEB theory describes how individuals acquire and use energy based on a set of simple rules for metabolic organization. Within this theory, organisms are treated as dynamic systems with explicit mass and energy balances.

Based on this theory, DEBtox (Fig. 1) was developed some 10 years ago to analyze toxicity data, as an alternative for descriptive statistics. We use DEBtox here to illustrate what a process-based analysis looks like and how

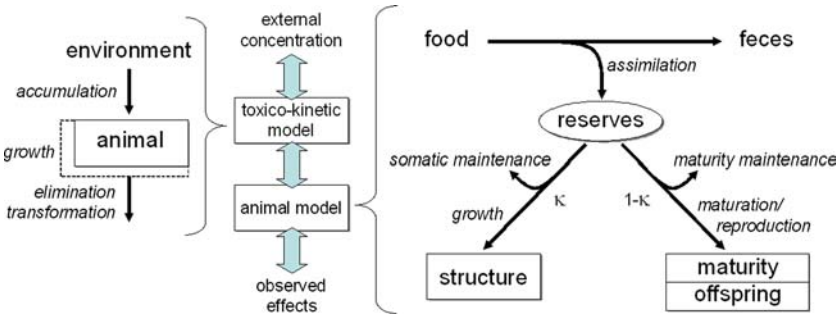

Fig. 1 The hierarchy of processes needed to describe effects observed in a toxicity test, including a schematic overview of the DEB model (right). In the DEB model, κ is a fixed fraction of the reserves that is used for growth and maintenance (which always has priority above growth)

it can be used to analyze a toxic response. Here, we focus on deriving more meaningful information from standard laboratory tests, and to our knowledge, DEBtox is the only approach to work with these data that systematically incorporates exposure time and biology of the organisms, including the natural links between the processes of feeding, maintenance, growth, development and reproduction. We think that these features are essential for any method to analyze toxicity data.

DEBtox has been described in detail elsewhere (Bedaux and Kooijman, 1994; Kooijman and Bedaux, 1996a, b; Kooijman et al., 1996). Here, we reiterate only the most important principles. The basic assumption of DEBtox is that toxicants must be taken up by the organism before they can exert an effect; this conforms with the concept of critical body residues (see Kooijman, 1981; McCarty and Mackay, 1993). Although this starting point is widely accepted, it is seldom applied in the analysis of toxicity test results. The first model in the chain should thus be a toxico-kinetics model (Fig. 1). In its simplest form, a toxico-kinetics model could be as simple as the well-known one-compartment model with first-order kinetics. When the organism is growing during the test, the effects of growth dilution must be included, as well as the effects on the rate constants; rate constants for uptake processes depend on the ratio of the surface area for exchange to the volume of the compartment and therefore decrease with body size (Kooijman and Bedaux, 1996a).

The toxicant, once inside the organism, may increase the probability of death (on the basis of hazard modeling, see Bedaux and Kooijman, 1994) and may affect a parameter of the animal model (Fig. 1). Which parameter is affected depends on the chemical: a chemical might increase maintenance costs, decrease the assimilation of energy from food, increase the energetic costs for growing new body tissue, increase the energetic costs for producing offspring, or pose a direct hazard to the developing embryo. These are the five modes of action that are currently incorporated in DEBtox (see Kooijman and Bedaux, 1996a); other mechanisms may also be explored.

Each of these modes of action is based on resource allocation, which is not directly comparable to the more familiar use of the term (e.g., distinguishing between narcosis, polar narcosis and uncoupling). Furthermore, the modes of action in DEBtox represent quite coarse changes in bioenergetics, and very different molecular mechanisms may lead to the same apparent mode of action. For example, the mode of action ''assimilation'' may result from a decrease in feeding rate, a decrease in assimilation efficiency in the gut, or perhaps even from changes in the cellular machinery responsible for the mobilization of energy. However, standard laboratory tests offer no possibility to identify the toxic mechanism in more detail.

To demonstrate some of the advantages of a DEBtox analysis, we use data sets for cadmium in *Daphnia magna*. This combination of toxicant and test species was chosen because it provides a clear example for a well-known test system, but this does not mean that the approach is limited to this species or this toxicant. Examples for other species and/or toxicants can be found in, among others, Kooijman and Bedaux (1996a, b), Bedaux and Kooijman (1994), Jager et al. (2004), Jager and Kooijman (2005), Alda Álvarez et al. (2005).

In Fig. 2, a simultaneous analysis of internal concentrations and mortality is demonstrated for cadmium in *D. magna* (details in Heugens et al., 2003). Even though DEBtox does not require internal concentrations (which can be treated as a hidden variable; Kooijman and Bedaux, 1996a), internal concentrations can be included when data are available because the toxico-kinetic model is part of the model chain (Fig. 1). Mortality and accumulation are modeled explicitly in time, which implies that the data for all time points can be used. In this case, the model describes the data well (Fig. 2), supporting the link between internal concentrations and effects. The model parameters estimated from this data set are given in Table 1. These parameters can then be used to make educated extrapolations, for example to effects under time-varying concentrations as demonstrated in Fig. 3 (using the results for continuous exposure in Table 1 to predict effects for a 20-h pulse). Interestingly, much higher cadmium concentrations are needed to kill the animals, and mortality does not cease immediately when exposure stops, reflecting the slow toxicokinetics of this chemical.

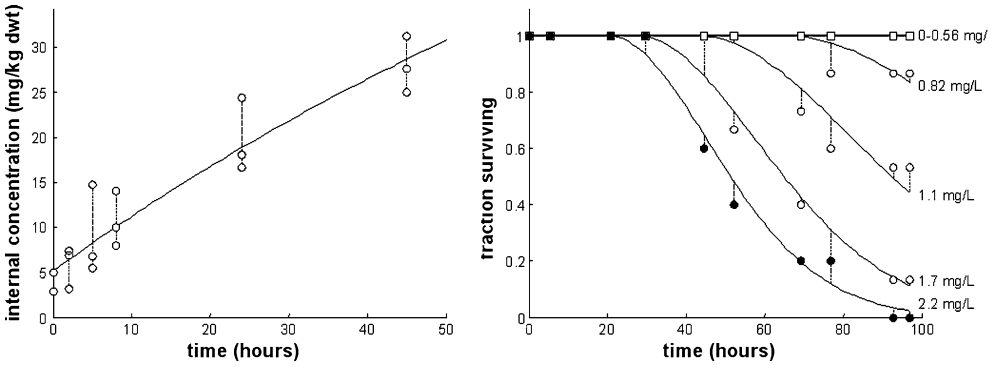

Fig. 2 Simultaneous analysis of body residues (at 0.1 mg/l, left plot) and mortality with DEBtox for cadmium (0–2.2 mg/l, right plot) in *Daphnia magna* at 10°C (no food). In the right plot, model lines indicate different exposure concentrations (squares are control, black dots the highest dose)

Table 1 Parameter estimates for the fit in Fig. 2, with 95% likelihood-based confidence intervals

Parameter	Value (95% conf.)	Unit
Elimination rate constant	0.00876 (0–0.0218)	h^{-1}
Bioconcentration factor	728 (368–∞)	l/kg_{dwt}
Initial body residue	5.20 (3.05–7.35)	mg/kg_{dwt}
NEC (internal basis)	259 (202–321)	mg/kg_{dwt}
Killing rate (internal basis)	1.31×10^{-4} (8.32×10^{-5}–1.89×10^{-4})	kg_{dwt} (mg h)$^{-1}$

Table 1 shows that the confidence interval for the elimination rate includes zero. This condition implies that the data also can be described by assuming zero elimination (i.e., a straight increasing line for the body residues in time), with an infinitely large bioconcentration factor. This causes problems for identifying the NEC, because at an infinitely large bioconcentration factor, even extremely low concentrations will eventually kill all organisms, provided that the organisms live long enough to accumulate substantial amounts of chemical in their bodies. This problem is circumvented in this example by expressing the NEC as an internal concentration (see Heugens et al., 2003), which is made possible by the availability of body residue data.

The iso-effect plots (Fig. 3) reveal that LC*x* values decline gradually with time and eventually converge at the no-effect concentration (NEC). The NEC is a model parameter in DEBtox and represents the concentration of the chemical that does not cause any effects (relative to the control) after prolonged exposure. The NEC is logically equivalent to the LC0 following infinite exposure. Therefore, the NEC is a time-independent summary statistic that can be estimated with a confidence interval, because it is a model parameter.

The second example is of a partial life-cycle test for *Daphnia magna* exposed to cadmium (Heugens, 2003). This test is very similar to the OECD test protocol for this species, but used a shorter test duration. In addition to survival and reproduction, body size was measured at the end of the test. The data for all three endpoints are analyzed simultaneously with DEBtox (Fig. 4), providing an indication for the mode of action and allowing parameter estimates such as the no-effect concentrations for survival and reproduction (Table 2). Note that the elimination rate is not well fixed by the data: indeed no elimination at all is within the 95% confidence interval. A consequence of this result is that other toxicological parameters, namely NECs, killing rate and tolerance, also are poorly identified. In contrast to the acute toxicity data set (Table 1), no internal concentrations were used for this fit, so these parameters cannot be expressed in internal concentration units.

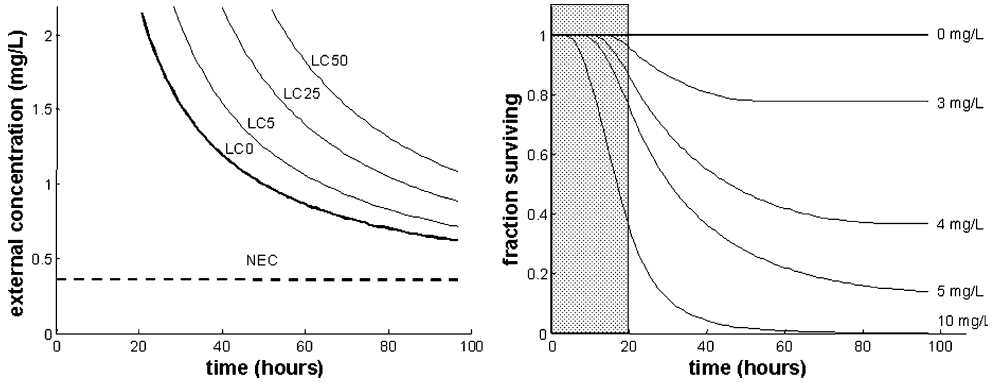

Fig. 3 Results derived from the fits in Fig. 2 for cadmium in *Daphnia magna*. Left plot shows the iso-effect lines in time, right plot shows a model simulation for a 20-h pulse exposure (gray area), based on the parameter estimates of Table 1

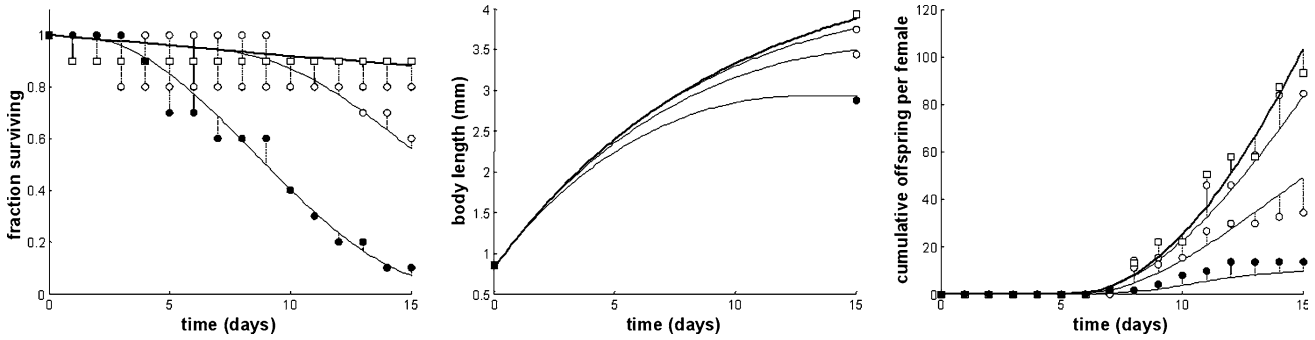

Fig. 4 Simultaneous analysis of survival (left), body size (middle) and reproduction (right) for *Daphnia magna* exposed to cadmium (0–0.23 mg/l) at 20°C. Food was present in the form of algae. Model lines indicate different exposure concentrations (squares are control, black dots the highest dose)

The model parameters can be used to calculate LC*x* and EC*x* versus time (Fig. 5), without additional estimation. This plot shows that each endpoint has its own dynamics for EC*x*, not necessarily decreasing through time. The shapes of the iso-effect lines depend on properties of the chemical (e.g., mode of action and toxico-kinetics), and on specific life-history characteristics of the test species and the effect level selected. This plot also shows that, at least in this case, body size is a less ''sensitive'' endpoint than reproduction (5% effect on body size is accompanied by a larger effect on reproduction). At the 5% effect level, body size is almost as sensitive as survival, but at the 50% effect level, survival is more sensitive than growth. Therefore, no single time point can provide a ''representative'' EC*x*, and no simple extrapolation between endpoints is possible.

Figure 5 also shows the results of a simple analysis of the population consequences (the intrinsic rate of population increase, *r*, using the Euler-Lotka equation). The intrinsic rate of population increase is much more ecologically relevant than any statistic based on a single endpoint (Forbes and Calow, 1999). Using the DEBtox model, population consequences under food limitation can be explored, assuming that the intrinsic sensitivity does not change with food level (Jager et al., 2004). The DEBtox analysis shows that food limitation can aggravate the toxic effects of cadmium.

Advantages of process-based analysis

Process-based modeling has clear advantages for environmental risk assessment. For example, this type of approach facilitates extrapolation between species (e.g., from laboratory species to related field species of interest) because many model parameters vary with body size in a predictable way (Kooijman, 2000). Analogously, test data for different species may be combined to yield a coherent set of information on a chemical. Process-based methods also aid in extrapolating between chemicals, because the model parameters often have simple relationships with chemical properties such as hydrophobicity (e.g., Kooijman et al., 1998; Kooijman et al., 2004). This predictability of effects, as long as the mechanism of action remains the same, is especially valuable in that it allows predictions for untested compounds. Furthermore, process-based modeling facilitates an educated extrapolation from single-species test results to population consequences (Fig. 5), which is impossible using the standard summary statistics.

Table 2 Parameter estimates for the fit in Fig. 4, with 95% likelihood-based confidence intervals (the < sign means that the interval extends to, but does not include, zero). Best fitting mode of action is a decrease in assimilation due to cadmium. N.e. is not estimated

Parameter	Value (95% conf.)	Unit
Elimination rate constant	0.000310 (0–0.00423)	d^{-1}
NEC for survival (external)	0.000274 (<0.0431)	mg/l
Killing rate (external)	512 (3.44–∞)	l (mg d)$^{-1}$
NEC for effect on assimilation (external)	0.000126 (<0.0192)	mg/l
Tolerance concentration (external)	0.00214 (<0.312)	mg/l
Blank hazard rate	0.00851 (0.00218–0.0222)	d^{-1}
Von Bertalanffy growth rate	0.106 (n.e.)	d^{-1}
Initial body size	0.857 (n.e.)	mm
Length at first reproduction	2.80 (2.74–2.87)	mm
Maximum length	4.67 (4.56–4.77)	mm
Maximum reproduction rate	37.8 (33.7–42.2)	juv/d

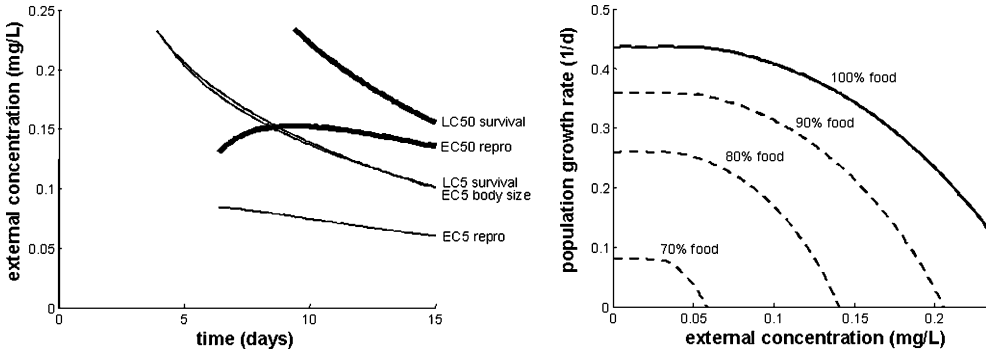

Fig. 5 Results derived from the fits in Fig. 4 for cadmium in *Daphnia magna*. Left plot shows the iso-effect lines through time for various endpoints; the right plot shows predicted intrinsic rate of population increase (r), based on the parameter estimates of Table 2

Because processes are modeled explicitly in time, the results of short-term test can be extrapolated to chronic timescales, and vice versa. This claim rests on two preconditions: (1) that the model parameters must be accurately fixed by the data, and (2) that no other modes of action of the chemical appear after prolonged exposure. It is also possible to make predictions of effects resulting from time-varying concentrations, even when the test is performed under constant exposure (Fig. 3). In the European guidelines (EC, 2003), the local scenarios encounter this problem for non-continuous emissions: surface water concentrations of pollutants are high during the emission episode, but low outside that period. For pesticides, pulsed exposures are even more common and model approaches to predict effects are required (Reinert et al., 2002).

In addition to the possibilities for extrapolation, process-based approaches also can improve the efficiency of the effects assessment. Current test analyses make inefficient use of the available data (OECD, 1998); the standard test protocols prescribe several measurements during the exposure period, but these data are not incorporated into the derivation of the summary statistics. For example, the OECD guideline for the acute *Daphnia* test prescribes that immobility is reported after 24 and 48 h, although only the 48 h results are used in risk assessment. Similarly, the *Daphnia* reproduction protocol recommends daily counting of neonates, but only the cumulative number of juveniles after 21 days is used. With a process-based approach, all observations are used to estimate model parameters, so fewer parameters need to be estimated per data point.

Yet another advantage of process-based approaches is that they can deal with problematic data sets that otherwise might need to be discarded. For example, if the exposure concentration varies through time, this will preclude estimation of an ECx or NOEC. However, the data can still be used in a process-based model (Péry et al., 2001). When concentrations decrease in time, additional assumptions may be needed about the recovery of the organisms (Reinert et al., 2002). Similarly, mechanistic models allow dealing with ionizing compounds (Kooijman, 2000), growing organisms (Kooijman and Bedaux, 1996a), or a decline in body weight in the test. Furthermore, it is possible to combine several studies in a process-based assessment (cf. Fig. 2), multiple chemicals with a similar mode of action (Jager and Kooijman, 2005), or respiration data with data on growth and reproduction (Jager et al., 2005). Even when each individual study is insufficient for proper analysis, existing ''low-quality'' data may be combined to obtain valuable information.

Greater efficiency (using more data points per parameter) results in more accurate predictions, which might allow for lower assessment factors. On the other hand, if greater accuracy is not critical, the number of tests or the number of animals per test may be decreased. Reducing animal testing is important for ethical reasons, and it also can reduce costs for chemical risk assessment. One aim of the new EU chemicals policy is to generate more data on chemical properties while limiting animal testing (Christensen et al., 2003). There is hope that QSARs can be used to fill ecotoxicological data-gaps (Christensen et al., 2003; Bradbury et al., 2004), but QSARs for ecotoxicity suffer from the same drawbacks as NOEC, LC50 and ECx methods. A more promising direction might be to develop more efficient testing strategies and test designs, and to use these in combination with process-based analyses. This approach would allow more information to be extracted from toxicity tests and for more accurate extrapolations, which would reduce animal testing. An outline for such a testing strategy and the consequences of its implementation for risk assessment are discussed in the following section.

Consequences for risk assessment

Would a switch to process-based analysis of toxicity tests require radical changes to the existing test protocols? Not necessarily, because DEBtox can already be used to analyze data from bioassays performed according to existing protocols. However, more radical changes in test protocols may be needed to make better use of the advantages of process-based modeling. Such changes will increase the costs of individual tests, so an effective screening

procedure would be needed in the initial risk assessment tier, to avoid using an extended test protocol for every individual chemical. Presently, acute survival tests are used for this purpose. However, these tests are not useful because acute mortality bears little relation to chronic population effects, which is the endpoint of interest for most risk assessments. Apart from risk assessment situations where acute mortality is important (short-term exposure to high concentrations), this form of testing might be discarded altogether. As an alternative, efficient screening could be achieved by using a low default value for the PNEC (Bradbury et al., 2004). Chemicals for which predicted exposure levels remain well below this threshold would not need to be tested unless there are other considerations, such as structural similarity to known problematic chemicals.

For chemicals that fail the criterion above, life-cycle tests should be considered. Such tests follow more endpoints (at least survival, growth and reproduction) over a longer period of time. Partial life-cycle tests may be sufficient because the early offspring contribute most to the population growth rate (due to the principle of ''interest upon interest''). The costs for a partial life-cycle test are not necessarily much greater than the costs of standard tests. For example, the *Daphnia* 21-day reproduction test can easily be adapted for this purpose. The protocol requires that reproduction and mortality are monitored at least 3 days per week. Additionally, several length measurements through time are sufficient for a useful life-cycle dataset that can be analyzed in a process-based manner (measuring body size is already recommended in the OECD protocol). Although reproduction and survival are the primary parameters that govern population growth, body size helps to identify the mode of action (Kooijman and Bedaux, 1996a), which is critical for extrapolations, such as predicting population response at limiting food levels (Kooijman and Metz, 1984) as demonstrated in Fig. 5.

For other species, extension of the standard chronic test protocols to a partial life-cycle test may be more costly because the test duration must be significantly increased. For example, chronic tests with fish usually consider only growth of early life stages. Such a test is of limited use because growth is not always a sensitive indicator for population effects, as discussed above. However, extending the test duration to a partial life-cycle test will seriously increase costs. To avoid prohibitive costs of testing, as well as animal lives, we have to investigate to what extent invertebrate tests are predictive for vertebrates, such as fish. Furthermore, we need better techniques for determining which chemicals really need testing with vertebrates (cf. Bradbury et al., 2004). For many soil- and sediment-dwelling organisms, a useful life-cycle protocol may also be more costly for a different reason. An appropriate data set contains observations in time, but these organisms and their offspring cannot simply be observed without sacrificing the test system. A more detailed analysis of the costs related to toxicity testing (see Kszos and Stewart, 1991) is needed to optimize the test protocols in terms of observations in time, and number of test animals.

Apart from changes in test design and testing strategy, we also need to consider how process-based analyses can be used to improve risk management decisions. PEC/PNEC ratios, currently used in this context, do not quantify risk; they merely indicate whether a certain (conservative) threshold has been exceeded. The absolute value of a PEC/PNEC ratio cannot be interpreted due to the unknown and varying degree of uncertainty, and because we do not know how this ratio relates to environmental impacts. For effective risk-based decision-making, the environmental impact of a certain exposure level must be estimated. Full or partial life-cycle tests allow calculation of the intrinsic rate of population increase (Fig. 5), a parameter that is time-independent, integrates all endpoints, accounts for the life history of the test organism, and has much greater ecological relevance than the standard test summaries (Forbes and Calow, 1999). Using computer simulations, it is even possible to explore population responses under food limitation (Jager et al., 2004, see Fig. 5) or link a decrease in population growth to the probability of extinction (Snell and Serra, 2000). In this way, effects assessment can quantify the effects that are expected at the predicted exposure level, instead of focusing on a PNEC with an unknown degree of built-in conservatism. Unlike the PEC/PNEC ratio, these ecological impacts can be directly compared between chemicals, which allows for comparative risk assessment as foreseen in REACH (Christensen et al., 2003).

Conclusions and outlook

Standard methods presently used to analyze ecotoxicological tests are very limited with respect to the amount of information that they extract from the data. These methods assess the data from a purely statistical perspective, yet ecotoxicological tests should more correctly be regarded from a biological viewpoint. When the data are considered from the latter perspective, they provide opportunity for estimating the parameters for an organism-focused process-based model. In our opinion, an integrated and process-based analysis would provide more insight into how toxicants affect organisms. A better understanding of toxicity is of scientific interest, but it would also improve the efficiency of toxicity testing, reduce the number of animals used for testing, and result in better informed

regulatory decisions. Process-based models should not be viewed as a complicated way to derive more accurate EC*x* values. Instead, these models will allow for educated extrapolation from laboratory tests to population-directed assessment endpoints.

The 1996 OECD workshop in Braunschweig, Germany (OECD, 1998) made some very clear recommendations: the NOEC should be phased out as a summary statistic and exposure time should be included into the data analysis. Unfortunately, these recommendations seem to have gone largely unnoticed: the procedures for testing, data analysis, and risk assessment have changed little over the last two decades. However, a promising development is the inclusion of the DEBtox method into a recent ISO/OECD document on the statistical analysis of ecotoxicity data, next to the approaches for NOEC and EC*x* (OECD, 2003). The basic ingredients for the new framework are already present, and we have an excellent opportunity to change ecotoxicological effects assessment as a result of the new EU strategy to deal with chemicals.

Acknowledgements This research was supported by the Netherlands Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs (WEB.5509). We thank Kees van Leeuwen (EC-JRC, Ispra) and the two anonymous reviewers for their valuable comments on drafts of this manuscript

References

Alda Álvarez O, Jager T, Kooijman SALM, Kammenga JE (2005) Responses to stress of *Caenorhabditis elegans* populations with different reproductive strategies. Funct Ecol 19:656–64

Bartell SM, Gardner RH, O'Neill RV (1992) Ecological Risk Estimation. Lewis Publishers Chelsea, MI, US

Bedaux JJM, Kooijman SALM (1994) Statistical analysis of bioassays based on hazard modeling. Environ Ecol Stat 1:303–14

Bradbury SP, Feijtel TCJ, Van Leeuwen CJ (2004) Meeting the scientific needs of ecological risk assessment in a regulatory context. Environ Sci Technol 38:463A–70A

Chapman PM, Fairbrother A, Brown D (1998) A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environ Toxicol Chem 17:99–108

Christensen FM, De Bruijn JHM, Hansen BG, Munn SJ, Sokull-Klüttgen B, Pedersen F (2003) Assessment tools under the new European Union chemicals policy. GMI 41:5–19

Crane M, Newman MC (2000) What level of effect is a no observed effect? Environ. Toxicol Chem 19:516–9

Crommentuijn T, Doodeman CJAM, Doornekamp A, Van Gestel CAM (1997) Life-table study with the springtail *Folsomia candida* (Willem) exposed to cadmium, chlorpyrifos and triphenyltin hydroxide. In Van Straalen NM, Løkke H (eds) Ecological Risk Assessment of Contaminants in Soil. Chapman & Hall, London, UK, pp 275–91

EC (2003) *Technical Guidance Documents on Risk Assessment, Part II. EUR 20418 EN/2* (http://ecb.jrc.it/tgdoc). Ispra, Italy: European Commission, Joint Research Centre

Forbes TL, Forbes VE (1993) A critique of the use of distribution-based extrapolation models in ecotoxicology. Funct Ecol 7:249–54

Forbes VE, Calow P (1999) Is the per capita rate of increase a good measure of population-level effects in ecotoxicology? Environ Toxicol Chem 18:1544–56

Heugens, EHW (2003) *Predicting Effects of Multiple Stressors on Aquatic Biota*. Ph.D. Thesis, University of Amsterdam

Heugens EHW, Jager T, Creyghton R, Kraak MHS, Hendriks AJ, Van Straalen NM, Admiraal W (2003) Temperature-dependent effects of cadmium on *Daphnia magna*: accumulation versus sensitivity. Environ Sci Technol 37:2145–51

Jager T, Alda Álvarez O, Kammenga JE, Kooijman SALM (2005) Modelling nematode life cycles using dynamic energy budgets. Funct Ecol 19:136–44

Jager T, Crommentuijn T, Van Gestel CAM, Kooijman SALM (2004) Simultaneous modeling of multiple endpoints in life-cycle toxicity tests. Environ Sci Technol 38:2894–900

Jager T, Kooijman SALM (2005) Modeling receptor kinetics in the analysis of survival data for organophosphorus pesticides. Environ Sci Technol 39:8307–14

Kooijman SALM (1981) Parametric analyses of mortality rates in bioassays. Water Res 15:107–19

Kooijman SALM (1996) An alternative for NOEC exists, but the standard model has to be abandoned first. Oikos 75:310–6

Kooijman SALM (2000) Dynamic Energy and Mass Budgets in Biological Systems. Cambridge University Press Cambridge, UK

Kooijman SALM (2001) Quantitative aspects of metabolic organization: a discussion of concepts. Phil Trans R Soc Lond B 356:331–49

Kooijman SALM, Bedaux JJM (1996a) Analysis of toxicity tests on *Daphnia* survival and reproduction. Water Res 30:1711–23

Kooijman SALM, Bedaux JJM (1996b) Analysis of toxicity tests on fish growth. Water Res 30:1633–44

Kooijman SALM, Bedaux JJM, Gerritsen AAM, Oldersma H, Hanstveit AO (1998) Dynamic measures for ecotoxicity. In Newman MC, Strojan C (eds) Risk Assessment: Logic and Measurement. Ann Arbor Press, Chelsea, MI, US, pp 187–224

Kooijman SALM, Hanstveit AO, Nyholm N (1996) No-effect concentrations in algal growth inhibition tests. Water Res 30:1625–32

Kooijman SALM, Jager T, Kooi BW (2004) The relationship between elimination rates and partition coefficients. Chemosphere 57:745–53

Kooijman SALM, Metz JAJ (1984) On the dynamics of chemically stressed populations: the deduction of population consequences from effects on individuals. Ecotoxicol Environ Saf 8:254–74

Kszos LA, Stewart AJ (1991) Effort-allocation analysis of the seven-day fathead minnow (*Pimephales promelas*) and *Ceriodaphnia dubia* toxicity tests. Environ Toxicol Chem 10:67–72

Laskowski R (1995) Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology. Oikos 73:140–4

McCarty LS, Mackay D (1993) Enhancing ecotoxicological modeling and assessment. Body residues and modes of toxic action. Environ Sci Technol 27:1719–28

Newman MC, McCloskey JT (2000) The individual tolerance concept is not the sole explanation for the probit dose-effect model. Environ Toxicol Chem 19:520–6

Nisbet RM, Muller EB, Lika K, Kooijman SALM (2000) From molecules to ecosystems through dynamic energy budget models. J Anim Ecol 69:913–26

OECD (1998) Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data. Organisation for Economic Cooperation and Development (OECD) Paris, France

OECD (2003) *Draft Guidance Document on the Statistical Analysis of Ecotoxicity Data*. Paris, France: Organisation for Economic Cooperation and Development (OECD), (for ISO as working draft ISO TC 147/SC 5 N 18, ISO/WD 1)

Péry ARR, Bedaux JJM, Zonneveld C, Kooijman SALM (2001) Analysis of bioassays with time-varying concentrations. Water Res 35:3825–32

Posthuma L, Suter GW, Traas TP (2002) Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers, Boca Raton, FL, USA

Reinert KH, Giddings JM, Judd L (2002) Effects analysis of time-varying or repeated exposures in aquatic ecological risk assessment of agrochemicals. Environ Toxicol Chem 21:1977–92

Smith EP, Cairns J (1993) Extrapolation methods for setting ecological standards for water quality: statistical and ecological concerns. Ecotoxicology 2:203–19

Snell TW, Serra M (2000) Using probability of extinction to evaluate the ecological significance of toxicant effects. Environ Toxicol Chem 19:2357–63

Sprague JB (1969) Measurement of pollutant toxicity to fish. I. Bioassay methods for acute toxicity. Water Res 3:793–821

Van der Hoeven N (1997) How to measure no effect. Part III: statistical aspects of NOEC, EC*x* and NEC estimates. Environmetrics 8:255–61