An alternative for NOEC exists, but the standard model has to be abandoned first

S.A.L.M. Kooijman, Dept. Theoretical Biology
Vrije Universiteit, de Boelelaan 1087, 1081 HV Amsterdam
15 aug 1995[1]

NOEC

Ryszard Laskowski discussed the bad properties of the No Observed Effect Concentration (NOEC) in the last issue of Oikos (Laskowski, 1995). I fully agree with him that the NOEC should be banned. Within the OECD and SETAC, discussions have been initiated to replace the NOEC by a 'small'-effect concentration (EC5 or EC10), but I will argue that this alternative is also far from ideal and propose a better alternative: the No Effect Concentration (NEC). It cannot be implemented successfully in the standard log-logistic or log-probit model. This is, however, hardly a handicap, because the biological foundations of those models are anyway extremely shaky.

Standard descriptions

LC50 and gradient

The results of standard tests on the lethality of toxicants usually give the number of surviving animals as a function of the concentration of toxicant, which has been constant (hopefully) during a standardized exposure. The blank survival probability is typically larger than 90% and a sigmoid curve is fitted to the number of survivors as a function of the concentration of toxicant (in the water). This curve is usually the log-logistic or log-probit curve, which both are characterized by a 50% point (the LC50) and a gradient parameter, which represents the maximum slope of the number of survivors as a function of the logarithm of the concentration in the water.

The model is based on the idea that death is certain as soon as the toxicant in the organism exceeds a certain individual-specific threshold value. Individuals vary in physiological condition, and therefore in threshold values. The threshold value of a particular individual is assumed to be a (random) trial from a bell-shaped frequency distribution, which leads to a sigmoid concentration-response curve for the number of survivors. The effect, then, is described deterministically at the level of the individual and stochastically at the level of the cohort of tested organisms. When counts have been made at different observation times, the LC50 generally decreases as a function of exposure time. This phenomenon can be described well on the assumption that the survival probability depends on the concentration in the organism, and that the toxicant follows some simple kinetics (Kooijman,

[1]OIKOS **75** (1996): 310–316

1981). This idea couples the LC50-time behaviour directly to the uptake-elimination behaviour of the compound, which also depends on characteristics of the organism, such as body size.

Problems

Several problems are inherent to this method of description.

- Extreme standardization of culture conditions practically eliminates the physiological differences between individuals, but experimental practice learns that the gradient parameter cannot be increased above a given level. There seems to be an upper limit for the maximum slope of the concentration response curve. In other words, there is a rather substantial variation of threshold values between individuals. It appears that the effect is *stochastic* at the level of the individual, not deterministic.

- The distribution that describes the variation in threshold values (log-logistic, log-probit) represents a rather arbitrary choice from the large set of possible distributions. This is of little relevance of the estimation of the LC50 value itself, as long as the selected curve fits the data. The particular choice of response function is however of great importance when we wish to obtain 'small' effect concentrations, such as the LC1 or LC5, from the estimated parameters (LC50 and the gradient). The smaller the effect level, the larger the confidence interval, and more important, the stronger the result depends on the specific choice for the model. (Parameter estimations based on the logit or probit transformation require exclusion of data with no or full mortality, in which case the LC1 would probably be an extrapolation. This problem can be avoided by using the superior maximum likelihood method with untransformed data, but the above mentioned problems remain.)

- At high concentrations, the standard LC50/EC50 model has a very unrealistic property: it predicts that there are individuals that are not affected by a very high concentration of the very toxic compound (See figure 1). This property is due to the infinitely long upper tail of the distribution of threshold values. This is not only at odds with physical chemistry, but from long experience we simply know that no individual survives prolonged exposure to very toxic compounds.

- Since the gradient parameter reflects the variation of the (logarithms of the) threshold values, it is independent of the exposure time. In practice, however, the gradient tends to increase with the exposure time. We cannot simply parameterize this phenomenon by choosing a time-dependent function for the gradient parameter, because we then run into the problem that for certain (small or large) concentrations, the survival probability will increase in time, which is obviously not possible. The only way to incorporate such phenomena is to go back to a process oriented description for survival.

- Sublethal effects show that the reasoning behind the description of lethal effects is simply wrong. No big variations in the physiological conditions of standardized test organisms have been observed. When a toxicant affects reproduction, for instance, it does so in all individuals to about the same extend. There are no records of individuals that cease reproduction, while others continue at the blank rate.

- As mentioned before, LC50/EC50 values themselves depend on exposure time. The problem is not completely solved by standardizing toxicity tests to a fixed exposure period (The standardized toxicity test with *Daphnia magna* lasts 48 h, independent from the type of compound that is tested). Surfactants react quickly; if no effect shows up after a few hours of exposure, it is unlikely that any effect will show up at that concentration. Things are totally different for toxicants such as cadmium. The LC50 for an animal as small as *Daphnia* still decreases after three weeks of exposure. The LC50-time behaviour depends on properties of the chemical as well as those of the organism (especially body size). This mixture of properties is most unfortunate in application of LC50/EC50 in risk assessment and reduces the comparability of results of standardized tests.

EC50

There is not much to say about models that are used to describe sublethal effects. The standard approach is to relate the quantity of interest, such as the cumulated number of offspring during a standardized test period, to the logarithm of the concentration of toxicant in a logistic way. Apart from the NOEC, this gives three parameters: the blank value, the EC50 and a gradient parameter. This model just serves the purpose to describe a very limited data set, without bothering about the foundations. The aim seems to be to obtain an EC50 value for the quantity of interest and compare it with other EC50 values of other compounds and/or effects. A good example of a frequently applied nonsense parameter is the EC50 for biomass of algae in a 3 days test on growth inhibition. The value depends on the arbitrarily chosen test period and the (blank) population growth rate (and therefore the medium composition, the temperature, light conditions, turbulence and alga species). Although these limitations are known (Nyholm, 1985), this is apparently no reason to abandon such measures. Many people seem to think that standardization solves all problems and leave to the poor administrator the problem of risk assessment which requires an integration of different information.

Conclusion

The standard model is based on assumptions that are not realistic. The fact that LC50 or EC50 values as well as NOECs depend on exposure time in a way that depends on both compound and organism characteristics hampers their application. The use of repeated observations to detect deviations from the blank is problematic without an adequate model for the appearance of effects. In general, such observations are statistically dependent. Moreover, the standard model is inconsistent with the concept of the NOEC, because the

log-logistic as well as the log-probit concentration-response curve approaches the blank response for decreasing concentrations only asymptotically. The standard model can be extended to include a no-effect concentration (NEC) as parameter (Kooijman, 1981). Such a model solves the problem of statistical dependence and the unknown power of the test to spot deviations from the blank response. This is because the null-hypothesis states that the NEC equals zero, while the alternative hypothesis asserts that it is positive. Twenty years of routine application learns that point estimates for NEC are positive in about 50% of the cases and in less than 10% of the cases the NEC differs significantly from zero. This poor performance is due to the gradient parameter, whose unknown value reduces the information content of a concentration that shows an obvious effect to "the NEC is smaller than that concentration". Response curves with a positive NEC and with a NEC of zero are too similar. Today, I consider my previous attempts to improve the standard model a failure.

Small effect concentrations have been proposed to replace the NOEC. Apart from the problem to define 'small', such parameters hardly solve the problem, due to the arbitrariness of the choice of response function. The problems become less pressing for moderate effect concentrations (e.g. LC10 or LC25), but who wants to allow such effects to occur? The larger the effect, the more important it is to have a reliable translation of the effect into consequences to the ecosystem. Such reliable translations do not exist and it is extremely unlikely that they will exist in the near future.

In summary, I have to conclude that the LC50 and EC50 are parameters that have nice statistical properties (Hoekstra, 1993), but nobody should attach much importance to their values. They are hardly relevant for risk assessment and they are based on a model with a shaky basis. Useful descriptions should be process-oriented. The mere fact that extensive data bases for LC50/EC50 values exist should not be a reason to continue the application of the standard model.

Alternative effect models

Organisms evolved in a chemically varying environment, that is they can cope with a varying concentration of any particular compound, as long as the variations are within a certain window. The upper boundary for this window might be zero for particular compounds. Each molecule of such compounds induces effects with a certain probability, but for most compounds, the upper boundary is positive. The lower boundary of the window is zero for most compounds, because they are not necessary for life. Compounds such as copper are necessary, so that the lower boundary for copper is positive. Effects of a shortage of a compound resemble effects of an overdoses in their kinetics. The founder of ecotoxicology, Sprague (1969), studied the effects of toxicants in bioassays, using oxygen shortage as example. Although many interrelationships exist between nutrition and toxic effects, the upper boundary of the tolerance window attracted almost all attention in ecotoxicology, due to its application in risk assessment studies, while ecology focused on the lower boundary.

Each physiological process has its own tolerance window for any compound. The

upper boundaries can be ordered, which means that at the lowest tissue-concentration range that produces effects, only one physiological process is affected, while at high tissue-concentrations many processes are affected. As long as the partitioning of the compound over the various body fractions is rapid with respect to the uptake/elimination kinetics for the whole animal, it is not essential to specify in which tissue or organ the most sensitive physiological process is affected. Specification of tissues only becomes essential if the partitioning is slow, which makes many-compartment models beneath effect studies so much more complex to apply: we have to know a great deal more. (Notice that one-compartment models can handle different concentrations in different organs as long as partitioning is rapid. Observed deviations from one-compartment kinetics with constant coefficients frequently relate to the variations in the coefficients due to changes in feeding conditions, body size, temperature, but not necessarily to the presence of more compartments.)

Basic to the description of small effects of toxicants is the notion that each additional molecule beyond the tolerance window contributes to the effect to the same extent. Interactions between the molecules only occur at higher tissue-concentrations. This means that the effect size is, as a first approximation, linear in the tissue concentration. This point of view relates to the Taylor approximation to describe how effect size relates to tissue-concentrations. This function might be non-linear, but we use only the first term of the Taylor approximation at the upper boundary of the tolerance window. The theorem by Taylor states that we can describe any non-linear function in a given interval arbitrarily well with an appropriate polynomial function if we include enough higher order terms. So when we want to improve the description of effects, if they happen to deviate from a linear relationship with tissue-concentrations, we simply include the squared term, the cubed term, etc. Such improvements will rapidly become counter-productive because we increase the number of parameters that have to be estimated and because higher tissue-concentrations will affect more physiological processes. So we are increasing precision at the wrong points. Practice teaches that very good descriptions can be obtained by just taking effect size linear in the tissue-concentration, even at rather high effect sizes, provided that we focus at the right physiological process.

Some people will argue that it is far from sure that a positive NEC actually exists. I hardly see this as a problem, because we always can (and should) test the hypothesis that the NEC equals zero. If we cannot reject this hypothesis, we have to live with the possibility that each molecule of that compound can have an effect. If this compound still has to be emitted into the environment, this might be a good reason to give priority to research into the ecological effect size of such an emission. The primary purpose of routine toxicity testing is to set priorities to further research, not to predict ecological effect sizes.

I will now discuss a selection of frequently occurring effects of toxicants, all based on the above mentioned principles. The details can be found in Bedaux & Kooijman (1994), Kooijman (1993), Kooijman & Bedaux (1995a, 1995b, 1995c), Kooijman et al (1995).

Effects on survival

The effects of toxicants on survival can be modelled successfully by taking the hazard rate linear in the tissue-concentration, that is proportional to the tissue-concentration

that exceeds the threshold value (i.e. the upper boundary of the tolerance window), see Kooijman, 1993, Bedaux & Kooijman, 1995. The proportionality factor, called the 'killing rate', quantifies the toxicity of the compound in a way that is independent of the exposure time. In (acute) toxicity tests that are started with animals from a blank culture, this cause of mortality results in a survival probability that equals the exponent of (minus) a constant times squared time if the exposure time is short with respect to the inverse elimination rate. When the animals are fully loaded, so that the tissue-concentration does not increase any longer, the survival probability equals the exponent of (minus) a constant times time. Since the hazard rate starts to increase from the blank value at the moment that the NEC value is exceeded, we have to wait longer for effects at lower concentrations.

The mortality is thus taken to be a function of the environment-concentration and the exposure time. The parameters of this function are the NEC, the killing rate, the elimination rate and the blank mortality rate. Even if the number of survivors are counted just once, at the end of the experiment, we have to estimate four parameters. If the exposure time is short with respect to the inverse of the elimination rate, we can only estimate the product of the elimination and the killing rate and are unable to translate the NEC for that test to an ultimate NEC. If the exposure time is large with respect to the inverse of the elimination rate, we end up with an exponential survival model that has the (ultimate) NEC, the killing and blank mortality rates as parameters. So we can sandwich the four parameter model between two three parameter models. If the cultures are in good condition and the test has been done carefully, we can avoid blank mortality, which further reduces the number of parameters. (It is bad practice to accept any blank mortality in acute toxicity tests.) This low number of parameters is essential for routine applications because the data do not allow the estimation of a larger number of parameters. Although the model is better underpinned mechanistically than the standard model, the number of parameters is smaller, because the gradient parameter and the LC50 are both replaced by the killing rate. An essential difference with the standard model is that the hazard-based model treats mortality for a particular individual as a stochastic event, rather than a deterministic one. The individuals are treated as identical (stochastic) copies, rather than different (deterministic) copies.

It is of course possible to account for differences between individuals in the hazard based model, which then appear as differences in parameter values for each individual. This makes sense for animals that are collected from the field, where differences in health, age, size, feeding conditions, sex, all contribute to differences in sensitivity. The way to proceed is to describe this variation in the set of four parameters by some (multivariate) scatter distribution and obtain what is called a 'mixture' in applied probability theory. The number of parameters in this scatter distribution is obviously larger than four and the resulting survival model can easily become complex. One must be prepared to estimate these extra parameters by increasing the number of observation times and tested concentrations. When the individual sizes are measured and the mortality observations specify the individuals, it is possible to correct for these differences in size in a rather straightforward way, which does not increase the number of parameters. The rate of uptake and elimination is proportional to a surface area, which implies that the waiting time to effects is proportional to a length measure

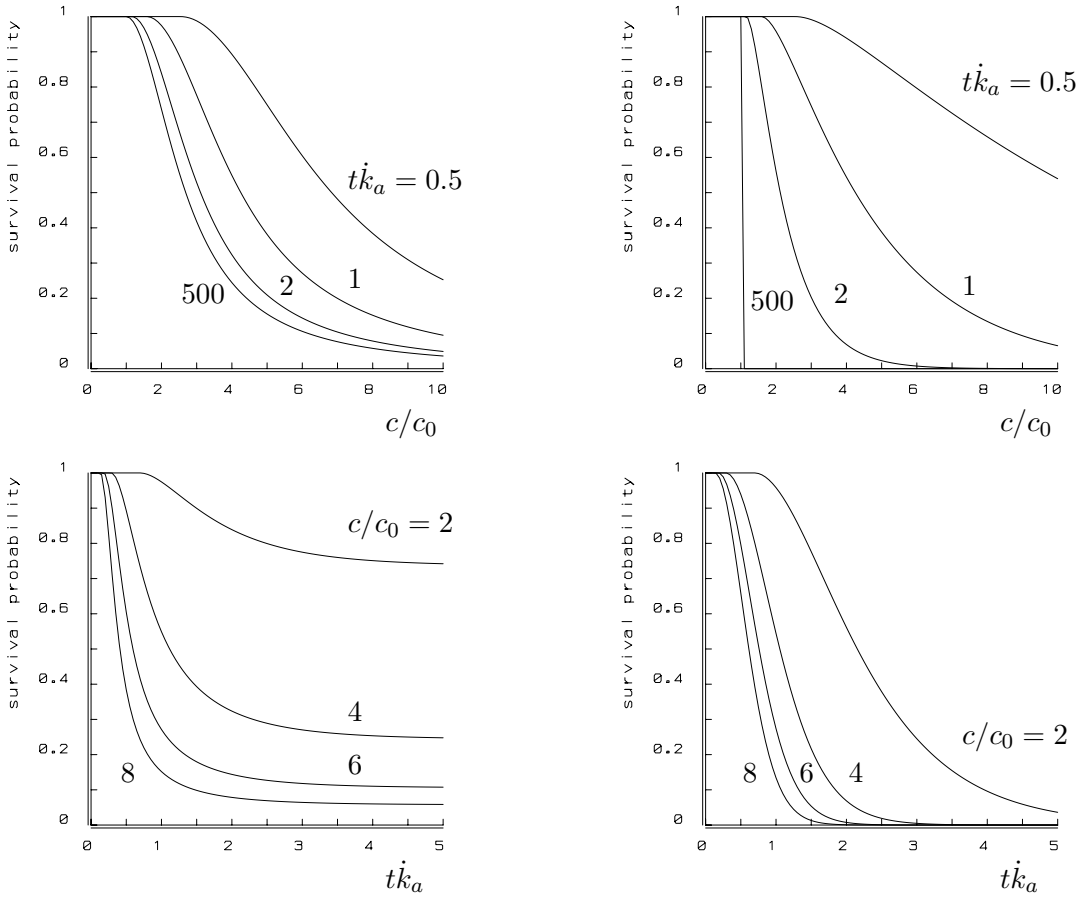

Figure 1: The survival probability as functions of the concentration of toxicant (top) and exposure time (bottom) for the standard log-logistic model (left) and the hazard model (right). The log-logistic model is extended to include a no-effect level c_0, while uptake/elimination follows a simple one-compartment model with elimination rate $\dot{k}_a$ in both models. The parameters $c_{L50.\infty}$ and β of the log-logistic model have been chosen such that both models have the same c_{L50} and c_{L25} values for exposure time $\dot{k}_a^{-1}$ for the choice $\dot{k}_\dagger = \dot{k}_a/c_0$. This is the case when $c_{L50.\infty} = 2.69c_0$ and $\beta = 0.51$. In the log-logistic model, some individuals will survive forever, even if the concentration exceeds the no-effect level. In the hazard model, all individuals will eventually die if the concentration exceeds the no-effect level. From Kooijman (1993).

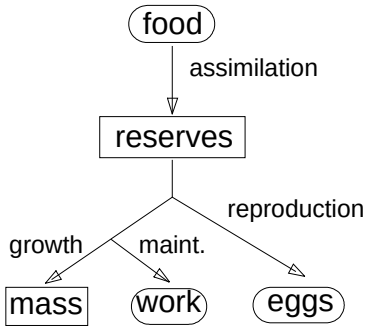

Figure 2: The energy fluxes as specified quantitatively by the DEB model for an ectotherm with body size and reserve density as state variables. Toxic compounds that affect reproduction can do so directly, or indirectly via assimilation, growth and maintenance. The rounded boxes indicate sources or sinks.

The elegance of hazard-based modelling is that other causes of mortality can easily be built in, so that the significance of toxicant induced mortality can be evaluated. This fits into the theory of competing risks (David and Moeschberger, 1978), which is based upon the fact that an animal can die only once, despite many possible causes of death.

Effects on reproduction

For a description of sublethal effects, it is essential to have a model that tells how the processes of feeding, digestion, storage dynamics, maintenance, growth and reproduction are interrelated. This interrelationship entails that processes can also be affected indirectly. The Dynamic Energy Budget model (Kooijman, 1993) has been designed specifically for this purpose. The DEB model describes how resources (i.e. products derived from food) are allocated to reproduction in heterotrophs. Toxic effects of chemicals change the allocation via the parameter values. Since the processes of assimilation (i.e. the combination of feeding and digestion), growth, maintenance and reproductions are intimately interlinked, changes in any of these processes will result in changes in reproduction. Figure 2 gives a simplified diagram for the relationships between assimilation, growth, maintenance and reproduction on which the DEB model is based. See e.g. Hallam et al. 1990 and McCauley et al. 1990 for alternatives for *Daphnia*. Two classes for the mode of action of compounds can be distinguished: direct and indirect effects on reproduction.

When reproduction is affected directly, assimilation, growth and maintenance are not affected (Kooijman 1993, Kooijman & Bedaux, 1995b). There exist two closely related routes within the DEB framework to affect reproduction directly. One is via survival of each ovum, one via the energy costs of each egg.

The survival probability of each ovum is affected very similarly to what is discussed in the previous section on effects on survival, except that the sensitive period is taken to be (relatively) short and fixed. (The case in which the sensitive period is long has already been discussed in the previous section.) The combination of an effect on the hazard rate of the ovum and a fixed sensitive period results in a survival probability that depends on the (local) environment of the ovum. This leads to another important difference with the previous section: the local environment of the ovum is the tissue of the mother, rather than the environment-concentration. So the relevant concentration changes in time. The toxicity parameters that appear in the survival probability of an ovum are the NEC, as before, and the tolerance concentration, which is inverse to the product of the killing rate and the length of the sensitive period. The elimination rate defines how the effect builds

up during exposure.

The reproduction rate, in terms of number of eggs per time, equals the ratio of the energy allocated to reproduction and the energy costs of an egg. If the compound affects the latter, it can be modelled by taking the energy costs a linear function of the tissue-concentration. The model is mathematically different from the hazard model, but behaves quantitatively rather similar. It has the same three toxicity parameters: the NEC, a tolerance concentration and the elimination rate. One possible interpretation of the tolerance concentration is the EC50 for the energy costs per egg minus the NEC.

Allocation to reproduction is initiated as soon as the cumulated investment into the increase of the state of maturity exceeds some threshold value. Since direct effects on reproduction only affect the translation from energy allocated to reproduction into number of offspring, these modes of action do not affect the time of the onset of reproduction. Indirect effects on reproduction via assimilation, maintenance and growth do delay the onset of reproduction. The occurrence of such delays is the best criterion to distinguish direct from indirect effects.

Indirect effects on reproduction all follow the same basic rules: the relevant parameter (which may be surface-specific assimilation rate or the volume-specific maintenance costs or the volume-specific cost of growth), is taken to be a linear function of the tissue-concentration. Since the assimilation rate represents a source of income, rather than costs, it is assumed to *decrease* linearly with the tissue-concentration, rather than increase. The effect on the reproduction rate as a function of environment-concentration and exposure time, all effects work out rather similar and all have the same three toxicity parameters. If growth has been measured during exposure, or if the size of the animals at the end of the exposure period has been measured, it is possible to identify the mode of action. The differences in effects on reproduction are too small to do this on the basis of effects on reproduction.

Toxicants sometimes stimulate reproduction at low concentrations, rather than reduce it; a phenomenon known as 'hormesis'. The actual cause is largely unknown and therefore difficult to model. For some compounds that showed hormesis at high feeding levels, I have been able to avoid hormesis in *Daphnia* reproduction tests by reducing the feeding levels. This points to an explanation in terms of suppression of a secondary stress by the toxicant at low concentrations. It is far from obvious to what extent this explanation is general. A wise strategy to deal with hormesis is the choose test conditions such that hormesis is avoided, if possible.

Conclusion

The wish to have estimates for concentrations of toxicants that do not have effects is legitimate. The NOEC, however, can be easily misleading, because of its bad statistical properties. The NEC serves the same purpose as the NOEC, but it has good statistical properties. The essential difference amounts to an exchange of the null and the alternative hypotheses. The inability to reject the null hypothesis that the response in the treatment is different from that in the blank, leads to a positive NOEC, thus an unsafe situation.

The inability to reject the null hypothesis that the NEC is zero leads to the safe conclusion that each molecule of the compound might have an effect.

The NEC cannot be built successfully into standard response models. They have to be replaced by mechanistically underpinned models. I have shown that this does not need to end in very complex and therefore inapplicable models. The mechanistic models are in fact simpler, both conceptually and in terms of the number of parameters. Experience with the application of NEC estimates so far learns that the NEC estimates are rather insensitive to the choice of model, within the context of the DEB theory. So the proper identification of the mode of action seems not to be essential for NEC values. The evaluation of ecological consequences does require such an identification, of course. This might serve as a sound motivation to avoid ecological effects of pollutants.

References

Bedaux, J.J.M. & Kooijman, S.A.L.M. 1994. Statistical analysis of bioassays, based on hazard modeling. *Envir. & Ecol. Stat.* (to appear)

David, H.A. & Moeschberger, M.L. 1978. *The theory of competing risks*, Griffin's Statistical Monographs & Courses **39**, C. Griffin & Co Ltd.

Hallam, T.G., Lassiter, R.R., Li, J. & Suarez, L.A. 1990. Modelling individuals employing an integrated energy response: application to *Daphnia*. *Ecology* **71**: 938-954.

Hoekstra, J.A. 1993. *Statistics in ecotoxicology.* PhD-thesis, Vrije Universiteit, Amsterdam, the Netherlands.

Kooijman, S.A.L.M. 1981. Parametric analyses of mortality rates in bioassays. *Water Res.* **15**: 107-119.

Kooijman, S.A.L.M. 1993. *Dynamic Energy Budgets in Biological Systems. Theory and applications in ecotoxicology.* Cambridge University Press, 350 pp.

Kooijman, S.A.L.M. & Bedaux, J.J.M. 1995a. Statistical properties of no effects levels. (subm.)

Kooijman, S.A.L.M. & Bedaux, J.J.M. 1995b. Analysis of toxicity tests on *Daphnia* survival and reproduction. (subm.)

Kooijman, S.A.L.M. & Bedaux, J.J.M. 1995c. Analysis of toxicity tests on fish growth. (subm.)

Kooijman, S.A.L.M., Hanstveit, A.O. & Nyholm, N. 1995. No-effect concentrations in alga growth inhibition tests. (subm.)

Laskowski, R. 1995. Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology. *OIKOS* **73**: 140-144.

McCauley, E., Murdoch, W.W., Nisbet, R.M. & Gurney, W.S.C. 1990. The physiological ecology of *Daphnia*: Development of a model of growth and reproduction. *Ecology* **71**: 703-715.

Nyholm, N. (1985) Response variable in algal growth inhibition tests - Biomass of growth rate? *Water Res.* **19**: 273-279.

Sprague, J.B. 1969. Measurement of pollutant toxicity to fish I. Bioassay methods for acute toxicity. *Water Res.* **3**: 793-821.