Ecotoxicology and Environmental Safety 114 (2015) 212–221

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv

Hierarchical modelling of species sensitivity distribution: Development and application to the case of diatoms exposed to several herbicides

Guillaume Kon Kam King [a,*], Floriane Larras [b], Sandrine Charles [a,c], Marie Laure Delignette-Muller [a,d]

[a] CNRS, UMR5558, Laboratoire de Biométrie et Biologie Évolutive, F-69622 Villeurbanne, France
[b] Institut National de la Recherche Agronomique, UMR 0042, Carrtel, Thonon, France
[c] Institut Universitaire de France, 103 bd Saint-Michel, 75005 Paris, France
[d] VetAgro Sup, Campus Vétérinaire de Lyon, 69280 Marcy l'Étoile, France

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form
15 January 2015
Accepted 21 January 2015
Available online 3 February 2015

Keywords:
Hazardous concentration
Global response
Bayesian statistics
Uncertainty
Diatoms
Ecological risk assessment

ABSTRACT

The species sensitivity distribution (SSD) is a key tool to assess the ecotoxicological threat of contaminants to biodiversity. For a contaminant, it predicts which concentration is safe for a community of species. Widely used, this approach suffers from several drawbacks: (i) summarizing the sensitivity of each species by a single value entails a loss of valuable information about the other parameters characterizing the concentration–effect curves; (ii) it does not propagate the uncertainty on estimated sensitivities into the SSD; (iii) the hazardous concentration estimated with SSD only indicates the threat to biodiversity, without any insight about a global response of the community related to the measured endpoint. To remedy these drawbacks, we built a global hierarchical model including the concentration–effect model together with the distribution law of the SSD. We revisited the current SSD approach to account for more sources of variability and uncertainty into the prediction than the traditional analysis and to assess a global response for the community. Working within a Bayesian framework, we were able to compute an SSD taking into account the uncertainty from the original raw data. We also developed a quantitative indicator of a global response of the community to the contaminant. We applied this methodology to study the toxicity and the risk of six herbicides to benthic diatoms from Lake Geneva, based on the biomass endpoint. Our approach highlighted a wide variability within the set of diatom species for all the parameters of the concentration–effect model and a potential correlation between them. Remarkably, variability of the shape parameter of the model and correlation had not been considered before. Comparison between the SSD and the global response of the community revealed that protecting 95% of the species might preserve only 80–86% of the global response. Finally, propagating the uncertainty on the estimated sensitivity showed that building an SSD on a low level of effect, such as EC_{10}, might be unreasonable as it induces a large uncertainty on the result.

1. Introduction

1.1. General introduction to SSD

The species sensitivity distribution (SSD) is a cornerstone of ecological risk assessment. Among other uses, it serves to predict concentrations in contaminant which are safe for a community. SSD is essentially an extrapolation of the sensitivity of a community of species from monospecific laboratory tests. The most standard approach (Aldenberg and Slob, 1993; Aldenberg and Jaworska, 2000; Posthuma et al., 2001) models the interspecific sensitivity variability in an assemblage of tested species in three steps. In the first step, the sensitivity of each species is summarized by a single critical effect concentration (CEC). This CEC can be a no observed effect concentration (NOEC) or a lowest observed effect concentration (LOEC). It can also be a no effect concentration (NEC) or an effective concentration at $x\%$ (EC_x), which are obtained by fitting a model to the concentration–effect curve. In the second step, the CECs in the community are assumed to follow a distribution law. Common choices for the distribution law include log-normal, log-logistic, and BurrIII. The chosen distribution is then fitted to the CECs of the sample of tested species. In the third step, the hazardous concentration to $p\%$ of the community (HC_p) is computed as a percentile of the previous fitted distribution.

The HC_p represents the concentration which affects $p\%$ of the

* Corresponding author.
E-mail address: guillaume.kon-kam-king@univ-lyon1.fr (G. Kon Kam King).

http://dx.doi.org/10.1016/j.ecoenv.2015.01.022
0147-6513/

community. The term "*affect*" is directly linked to the type of CEC in terms of level of effect (for example the x of the EC_x) and of biological effect (lethal, non-lethal, acute, and chronic). If NOEC or NEC were true no effect concentrations, one would expect the HC_p to leave $(100-p)$% of the community species completely unharmed. Using EC_{50} however, which is a level of effect commonly selected, one expects $(100-p)$% of the community to remain unaffected, which means that they suffer a reduction of less than 50% to their measured endpoint. But it is not possible to determine the reduction suffered by the unaffected species, which could lie anywhere between 0% and 50%.

SSD essentially carries information about the structural response of a community to a contaminant, i.e. the fraction of species affected at a certain level. The HC_p for small p, such as the HC_5, is ultimately used as a risk indicator. It is compared to the actual concentration of contaminant in an environmental setting to determine if the community living there is at risk, or to define an acceptably safe concentration for that community.

Several sources of uncertainty enter at the various steps of the SSD approach and have an influence on the predicted HC_p value. First, there is an uncertainty on the estimate of the CEC from the experimental data: when the CEC is estimated from a concentration–effect curve or more generally from any model, it comes with a confidence interval. Second, uncertainty arises from the fitting of a distribution to the CECs: parameters of the distribution also have their own confidence intervals. This adds to the total uncertainty on the HC_5. The uncertainty of this second step has already been studied and methods have been found for specific distribution laws (Aldenberg and Slob, 1993; Aldenberg and Jaworska, 2000; Wagner and Lokke, 1991). For other types of distributions, it is possible to use bootstrap (Efron and Tibshirani, 1994) to obtain confidence intervals, as described by Shao (2000) for the BurrIII distribution or in previous work by Kon Kam King et al. (2014). This uncertainty was also investigated with non-parametric approaches in the estimation of the SSD (Jagoe and Newman, 1997; Verdonck et al., 2001; van der Hoeven, 2001; Grist and Leung, 2009). However, there are currently very few attempts to include together all the sources of uncertainty into the final prediction of the SSD (Aldenberg and Rorije, 2013).

1.2. Several flaws of current SSD methodology

The classical SSD approach described in the previous section and its many variants present a number of flaws (Forbes and Calow, 2002; Power and McCarty, 1997) ranging from ecotoxicological concerns (use of laboratory data to predict field effects, inferring community sensitivity from monospecific sensitivities and chronic vs. acute effects) to statistical issues (fitting a distribution on a small dataset, distributional assumptions and treatment of the uncertainty). This paper focuses on several of these: first, the classical SSD approach does not propagate the uncertainty on the CEC to the prediction. This is a source of concern, because following this approach, the uncertainty on the HC_p depends on the number of species, but not on the quality of the data used. Second, the CEC retains only a fraction of the information originally present in the data. Since the aim of SSD is to model the variability in sensitivity in the community, it is important to consider all the information available in the data to obtain the best estimation of that variability. Indeed, there is relevant biological information in all the parameters of the concentration–effect curve and their potential correlations. Third, providing an HC_p, the classical SSD approach outputs information about a structural response of the community only. It essentially yields the proportion of affected species for a given concentration in contaminant. It does not give information about the global response of the community (Forbes and Calow, 2002; Kefford et al., 2012; De Laender et al., 2008), i.e. a response of the same nature as the measured endpoint. For instance, when using EC_{50} for biomass reduction as input, the SSD does not say anything about the change in the biomass of the community. In other words, the SSD aims to protect the structure of the community, but does not consider the effect on the community endpoint linked to the tested species which could be growth, reproduction, biomass, respiration, photosynthesis or any ecosystem process.

To address such issues, we revisited the current SSD approach to account for more sources of variability and uncertainty into the prediction than the traditional analysis and to assess the risk for the community from a global point of view. For this purpose, we built a hierarchical model inspired by Moore et al. (2010) including the concentration–effect model together with the distribution law of the SSD. From this hierarchical model, we were able to develop: (1) an indicator for the global response of the community, which we compared to the structural response predicted by the classical SSD; and (2) an SSD calculated on any level of effect (x of the EC_x) including correlation among the parameters of the concentration–effect model and the uncertainty from the original data.

2. Materials and methods

2.1. Diatoms sensitivity dataset

Our work was developed on a previously published dataset (Larras et al., 2012) containing 11 diatom species exposed to six herbicides: atrazine, terbutryn, diuron, isoproturon, metolachlor and dimethachlor. Between five and ten species were tested per herbicide. Benthic diatoms are unicellular microalgae which form a group of high diversity and which are often used to monitor water quality. They are well known to evolve in the biofilm matrix, at the interface of water column and substrata. The chosen diatom species were representative of Lake Geneva benthic diatoms communities and covered a great diversity in terms of taxonomy, morphology, herbicide sensitivity and ecological traits. More details about chosen diatoms are presented in Larras et al. (2012). Then, a panel of herbicides was selected considering their occurrence in Lake Geneva, their hazard to microalgae and their mode of action. Atrazine, terbutryn (triazine family), diuron and isoproturon (phenylurea family) prevent the photosynthesis at the level of the photosystem II, but with different mechanisms. Metolachlor and dimethachlor (chloroacematide family) inhibit especially the biosynthesis of very long chains of fatty acids. The sensitivity of the species was determined by assessing the growth over four days as endpoint, based on chlorophyll a fluorescence (the part of light which is absorbed by chlorophyll molecules then re-emitted at a defined wavelength), a proxy of the biomass. Bioassays were conducted in triplicates on diatom strains in their exponential growth phase, when the daily growth ratio is approximately constant. Seven to ten herbicide concentrations were tested. Chlorophyll a fluorescence was measured using Fluoroskan (Fluoroskan Ascent, Thermo-scientific, Finland) at the beginning and at the end of the experiment. More details about the tested species are presented in the first section of the Supplementary Information and in Larras et al. (2012).

2.2. Concentration–effect model

Contrasting with Larras et al. (2012), the response of each set (herbicide, species, and replicate) was defined as the ratio:

$$R = \frac{\beta_f}{\beta_0} \quad (1)$$