Rebuttal related to "Traits and Stress: Keys to identify community effects of low levels of toxicants in test systems" by Liess and Beketov (2011)

Matthias Liess · Mikhail A. Beketov

Identifying long-term community effects of low toxicant concentrations is one of the major challenges in ecotoxicology at present. Mesocosm investigations, which are one valuable tool to identify such long-term community effects, face the problem that results are obscured by confounding factors and high variability between replicates (Sanderson et al. 2009). This problem is even greater when assessing long-term effects of toxicants because inter-replicate variation increases with time. Under such conditions, multivariate analyses on species data such as usually conducted with PRCs dramatically lose power to identify statistical links. With the $SPEAR_{mesocosm}$ approach based on the investigations of Liess and von der Ohe (2005), we are aiming at resolving these shortcomings. We suggested a trait based aggregation of species that reduces the variability of community measures to reveal causality between exposure and effect for mesocosm experiments. We also show that this approach offers advantages in detecting long-term community effects compared to the traditional species-based PRC approach of (Van den Brink and Ter Braak 1999). Van den Brink and Ter Braak stated that the $SPEAR_{mesocosm}$ approach (Liess and Beketov 2011) "may not offer the level of improvement suggested", is not "an entirely appropriate approach for the evaluation mesocosm studies", and "the generality of the proposed method is at best rather questionable". We had to come to the conclusion that most of their critics are based on a lack of understanding of the $SPEAR_{mesocosm}$ approach. We are happy to assist in gaining further insights to our approach with comments targeted directly to the critics of Van den Brink and Ter Braak.

Understanding $SPEAR_{mesocosm}$

We agree with Van den Brink and Ter Braak that "PRC is a multivariate statistical method and $SPEAR_{mesocosm}$ is a univariate method." However a more fundamental difference is that $SPEAR_{mesocosm}$ is not solely a statistical method but (i) is using a priori knowledge to identify the most vulnerable taxa and (ii), is aggregating these taxa to reduce between replicate variability. There is often much variability between mesocosms under the same treatments, the different control mesocosms can have quite different communities (Sanderson et al. 2009), as can replicates of the various treatments. Under multivariate methods this adds noise and effects of the treatment will only be detected if they are greater than this noise. But methods like $SPEAR_{mesocosm}$ which aggregate all taxa into sensitivity categories (at risk or not at risk in the case of SPEAR) reduce this noise and then only compare whether the proportions of the categories differ between treatments. For the same reason SPEAR in general was also highly successful in identifying effects of pesticides on the ecosystem level (Liess and von der Ohe 2005; Liess et al. 2008; Schäfer et al. 2011).

Van den Brink and Ter Braak miss the point when they write "PRC diagrams show the contrasting responses of different taxa, very much like the contrasting response of the sensitive univoltine species and the other taxon groups as displayed in Fig. 2 of Liess and Beketov (2011)." The groups displayed in Fig. 2 are obtained using the respective a priori knowledge, not by statistical analyses. Van den Brink and Ter Braak write "The agreement of Fig. 2 with

M. Liess (✉) · M. A. Beketov
Department System Ecotoxicology, UFZ-Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany
e-mail: matthias.liess@ufz.de

PRC would have been even greater if the percentage change would have been plotted on a logarithmic scale." Presenting the results on a logarithmic scale or a linear scale will not change statistical power. The fact is that PRC does not detect statistically significant long-term effects even at the highest concentration of 100 µg/l Thiacloprid (neither 1st nor 2nd PRCs, see below) whereas $SPEAR_{mesocosm}$ detected changes at the lowest concentration tested of 0.1 µg/l, which is a factor 1,000 lower.

To identify the performance of the different approaches, it is crucial to compare the respective outcomes. However, Van den Brink and Ter Braak question a comparison of $SPEAR_{mesocosm}$ versus PRC. Their reasoning is that "the PRC diagram shows the dominant response present in the whole invertebrate community, $SPEAR_{mesocosm}$ only takes (presumed) sensitive species into account". We compare the 2 approaches because they are applied for the same purpose: the detection of adverse toxicant effects for risk assessment. For this, it is crucial to identify short-term effects, and perhaps even more important, long-term effects of toxicants. Regarding the difference of endpoints, it is more logical to describe the community in relation to its proportion of sensitive species instead of its dominant response, with respect to levels of protection and acceptability of effects in ecological risk assessment. The vulnerable species are those most determinant in risk assessment; as they will be more threatened by toxicants and thus will lead the risk assessment. Similarly to this approach, PRC also attempts to identify species that are sensitive towards the toxicant applied. The difference between the approaches is that (i) $SPEAR_{mesocosm}$ groups the vulnerable species and can therefore also include those that are only present in low numbers and show a great variability between replicates, and (ii) which results in $SPEAR_{mesocosm}$ detecting long-term effects that the PRC approach based on all species using PRCs is not able to identify.

Effect identification of single taxa is suggested by Van den Brink and Ter Braak as a possibility to get a better insight and improve understanding on relevant effect concentrations. We agree that this approach is reasonable and in many cases more sensitive compared to the PRC approach. That may be the reason that it is commonly required in the evaluation of microcosm and mesocosm studies performed for registration purposes. However, the optimism regarding effect assessment of single taxa of Van den Brink and Ter Braak is misleading. They state that "Such analyses would no doubt also highlight the sensitive responses at the population level as are presented by the SPEAR method." But we showed previously that Thiacloprid effects at low concentrations (0.1 µg/l) could only be observed for one species at some time points and long-lasting effects—as shown by $SPEAR_{mesocosm}$—were not

identified (Beketov et al. 2008). The cause for the difficulty to identify low-level effects for single taxa lies in the low numbers of replicates available for all mesocosm investigations and the high variance generally present in such complex systems (Sanderson et al. 2009). As stated repeatedly, only the grouping of taxa—as done in $SPEAR_{mesocosm}$—allows this problem to be tackled.

Improving PRC

Van den Brink and Ter Braak suggested several improvements for the standard PRC as described in (Leps and Smilauer 2003). This includes an a priori classification into "PRC analysis only using the sensitive univoltine taxa". We implemented this idea already in a mesocosm investigation quantifying recovery times for species with contrasting life cycles (Beketov et al. 2008). The results revealed that taxa characterised by a long life cycle need a prolonged time for recovery. However, this "improved PRC" approach again did not reveal statistically significant effects as low as SPEARmesocosm.

Additionally Van den Brink and Ter Braak suggested applying a higher order PRC even they "acknowledge that testing the second PRC on its significance and presenting it when it is significant is not common practice". We are fully aware of the potential of the second, third and further PRCs, but we cannot share the optimism of van den Brink and Ter Braak regarding this technique. When assessing effects using second and further PRCs, it is not possible to assess effects of separate concentrations, e.g. by supplementary RDAs, and therefore to derive NOECs and LOECs. Furthermore, for such axes it is problematic to unequivocally attribute the observed effects to the tested toxicant and not to other, frequently unknown factors. This makes this approach an exploratory tool that may be used to delineate possible gradients and generate hypotheses (similarly to unconstrained ordination techniques) (e.g. Leps and Smilauer 2003). In our case, the second PRC was statistically significant, and regarding the species scores, clustered the SPEAR-species together. However, the patterns expressed by the second PRC itself did not show long-term effects. This again confirms that it is difficult to interpret results obtained by this technique, especially when compared with simple and transparent grouping of most vulnerable species into the $SPEAR_{mesocosm}$ index.

Technical questions on $SPEAR_{mesocosm}$

The general correctness of species classification into sensitive and insensitive taxa was questioned by Van den Brink and Ter Braak. Pesticide effects "shown by Figs. 2A

and C of Liess and Beketov (2011) which indicate direct effects on taxa that are classified as insensitive" was used as an argument to support their critics. Here we suggest to apply the general knowledge of Paracelsus who realised already in the 16th century the basic principle of the dose response relationship, stating that the strength of effect is dose depending "All Ding' sind Gift und nichts ohn' Gift; allein die Dosis macht, das ein Ding kein Gift ist". Coming back to our example, we can expect to see effects on insensitive species at 100 µg/l of Thiacloprid when simultaneously sensitive species show effects at 0.1 µg/l, i.e. a factor 1,000 lower!

In particular the species classification of *Gammarus* and *Chironomidae* was questioned by Van den Brink and Ter Braak. Here our paper states that a taxon is regarded as a "species at risk" only if the generation time is equal or more than 1 year. Hence, both Chironomidae and *Gammarus* sp. cannot be classified as species at risk (i.e. SPEAR-species) because they are multivoltine.

Regarding the statement "$SPEAR_{mesocosm}$ indicator does not allow for unforeseen sensitivities or life cycle characteristics of taxa", we would like to draw the attention to the following statement in our publication: "We suggest using the $SPEAR_{mesocosm}$ approach as well as the PRC approach in order to obtain a comprehensive assessment of the toxicant induced community effect". Hence, we fully agree that PRC should be also used in concert with $SPEAR_{mesocosm}$ to not miss any chance to identify unforeseen effects. However, to our experience until now, $SPEAR_{mesocosm}$ was identifying effects unforeseen by the PRC approach.

It was criticised that "the reader cannot ascertain whether Fig. 2D is based on high or low abundance values (even single individuals), which is of crucial importance in any robust evaluation of the effects on sensitive univoltine species as compared to the whole community." But to inform the reader about the robustness of an evaluation, the use of statistical test in combination with certain levels of significance (i.e. $P < 0.05$) is generally widely accepted. This also enables a fast and reproducible estimation. Hence, to inform the reader, we provided information of the tests used and levels of significance applied. For the SPEAR-species, with the average of 6.2 individuals collected in the time period before first contamination/mesocosm (i.e. equals 23 individuals/square meter. This ranks the SPEAR-species exactly in the average abundance for all species; roughly 50 times less abundant than the dominant species (*Simulium*) and 50 times more abundant than the largest predator species (*Aeshna*).

Regarding the statement that of Van den Brink and Ter Braak that $SPEAR_{mesocosm}$ index "focuses on sensitivity and voltinism, it also ignores indirect effects, which are a key consideration for performing microcosm and mesocosm tests", we again have to draw attention to the content of our publication which clearly shows that the $SPEAR_{mesocosm}$ index covers the overall results of indirect effects.: "The $SPEAR_{mesocosm}$ index was computed as the relative abundance of sensitive univoltine species… as detailed in the following formula…" (page 1332). Hence, as the $SPEAR_{mesocosm}$ index is calculated as relative abundance of species at risk to the abundance of species not at risk, it indeed accounts for indirect effects. The decline of sensitive species increases the development of insensitive species as also shown for ecosystem level effects of pesticides (Liess and von der Ohe 2005). Also interactions of toxicant stress with abiotic stress (Duquesne and Liess 2003), predation stress (Beketov and Liess 2006) and also even subtle stress leading to behavioural responses (Reynaldi et al. 2011) will be included into the altered ratio of SPEAR, as sensitive species are affected more than insensitive species (Foit et al. 2011). Therefore we conclude that of course the $SPEAR_{mesocosm}$ approach includes indirect effects into its response.

Toxicants with different mode of actions can have very different toxicity profiles. To account for this fact we stated in our paper that "…classification of taxon sensitivity was adapted… to produce a ranking of taxon sensitivity to this specific toxicant according to the available knowledge." We believe that Van den Brink and Ter Braak understood this approach as they state that "This suggests that one could require a different $SPEAR_{mesocosm}$ for each new compound to be tested…" In this context it is incomprehensible to us that Van den Brink and Ter Braak state "The use of a single indicator of sensitivity neglects the fact that pesticides with different mode of actions can have very different toxicity profiles." This is exactly what we stated in our paper and the reason for us to adopt the sensitivity ranking for Thiacloprid.

Concluding remarks

Moving forward from the rather technical discussion of this paper, we would like to draw the attention of the readers to the implications of our findings that may have sparked this discussion. Since decades, there is a heated debate about extrapolating effects of chemicals observed in lower- and higher-tier tests to make predictions so that the aquatic communities of natural ecosystems are not endangered. Within this context, one important question is the assessment factor that should be used to make sure that no unacceptable effects on the ecosystem will occur, following the use of pesticides when extrapolating the endpoint from a SSD based on acute laboratory LC50 data. Amongst other papers with the participation of van den Brink, Maltby et al. (2005) compared single-species acute toxicity

data with effects observed in (micro)mesocosm. Reviewing information for single applications of 7 insecticides, they concluded that the median HC5 derived from a SSD based on acute laboratory LC50 information was generally protective for communities in (micro)mesocosms.

We would like to challenge this "rule" based on our observations. In the study of Liess and Beketov (2011) we identified long-term alterations of community structure with $SPEAR_{mesocosm}$ at 0.1 µg/l. This concentration is 7 times below the concentration identified as relevant endpoint from a SSD based on acute laboratory LC50 information for Thiacloprid (i.e. HC5 LC50) (Beketov and Liess 2008). As an additional support we re-evaluated with $SPEAR_{mesocosm}$ the study of Van den Brink et al. at (1996) investigating the effects of chlorpyrifos in mesocosms. We are, unfortunately, not in the position to present the results here—supporting our claims—since no authorisation for their presentation was given, despite a request to the authors.

We conclude, that trait based methods such as $SPEAR_{mesocosm}$ enable a realistic assessment of long-term community effects. This allows in concert with other methods (e.g. adoption of SPEAR to available toxicity information for particular compound) to derive safe concentrations for effects in complex mesocosm communities and eventually the field.

References

Beketov MA, Liess M (2006) The influence of predation on the chronic response of *Artemia* sp. populations to a toxicant. J Appl Ecol 43(6):1069–1074

Beketov MA, Liess M (2008) Acute and delayed effects of the neonicotinoid insecticide Thiacloprid on seven freshwater arthropods. Environ Toxicol Chem 27(2):461–470

Beketov M, Schäfer RB, Marwitz A, Paschke A, Liess M (2008) Long-term stream invertebrate community alterations induced by the insecticide Thiacloprid: effect concentrations and recovery dynamics. Sci Total Environ 405:96–108

Duquesne S, Liess M (2003) Increased sensitivity of the macroinvertebrate *Paramorea walkeri* to heavy-metal contamination in the presence of solar UV radiation in Antarctic shoreline waters. Mar Ecol Prog Ser 255:183–191

Foit K, Kaske O, Liess M (2011) Competition increases toxicant sensitivity and delays the recovery of two interacting populations. Aquat Toxicol. doi:10.1016/j.aquatox.2011.09.012

Leps J, Smilauer P (2003) Multivariate analysis of ecological data using CANOCO. University Press, Cambridge

Liess M, Beketov M (2011) Traits and stress-keys to identify community effects at low toxicant level. Ecotoxicology 20:1328–1340

Liess M, von der Ohe PC (2005) Analyzing effects of pesticides on invertebrate communities in streams. Environ Toxicol Chem 24(4):954–965

Liess M, Schäfer RB, Schriever CA (2008) The footprint of pesticide stress in communities-species traits reveal community effects of toxicants. Sci Total Environ 406(3):484–490

Maltby L, Blake N, Brock TCM, Van Den Brink PJ (2005) Insecticide species sensitivity distributions: importance of test species selection and relevance to aquatic ecosystems. Environ Toxicol Chem 24(2):379–388

Reynaldi S, Meiser M, Liess M (2011) Effects of the pyrethroid fenvalerate on the alarm response and on the vulnerability of the mosquito larva Culex pipiens molestus to the predator Notonecta glauca. Aquat Toxicol 104(1–2):56–60

Sanderson H, Laird B, Brain R, Wilson CJ, Solomon KR (2009) Detectability of fifteen aquatic micro/mesocosms. Ecotoxicology 18(7):838–845

Schäfer RB, Kefford B, Metzeling L, Liess M, Burgert S, Marchant R, Pettigrove V, Goonan P, Nugegoda D (2011) A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in South-East Australia. Sci Total Environ 409(11):2055–2063

Van den Brink PJ, Ter Braak CJF (1999) Principal response curves: analysis of time-dependent multivariate responses of biological community to stress. Environ Toxicol Chem 18(2):138–148

Van den Brink PJ, Van Wijngaarden RPA, Lucassen WGH, Brock TCM, Leeuwangh P (1996) Effects of the insecticide Dursban 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: II invertebrate community responses and recovery. Environ Toxicol Chem 15(7):1143–1153