

Updating statistical practice in ecotoxicology: reflections and recommendations

S. Jannicke Moe^{1,2,*}, David R. Fox³, and Raoul Wolf⁴

¹Norwegian Institute for Water Research (NIVA), Økernveien 94, 0579 Oslo, Norway

²Senior Editor, *Integrated Environmental Assessment and Management*

³Environmetrics Australia, Beaumaris, Victoria, 3193, Australia

⁴Norwegian Geotechnical Institute (NGI), Sandakerveien 40, 0484 Oslo, Norway

*Corresponding author: S. Jannicke Moe. Email: Jannicke.Moe@niva.no

Approaches to statistical analysis of data from ecotoxicity testing emerged in the 1980s and have grown over the decades, but with only intermittent involvement of statisticians. Consequently, statistical practices in this field have evolved in various directions. Fragmented, inconsistent, and outdated use of statistical methods for ecotoxicology by authorities and jurisdictions have further contributed to a landscape which can be confusing to navigate. A prominent example is the issue of whether no-observed effect concentrations (NOECs) should be banned from or used in regulatory ecotoxicology, which has been debated for more than 30 years (e.g., Laskowski, 1995; van Dam et al., 2012).

Other scientific disciplines, e.g., ecology, psychology, and medicine, have refined and optimized their statistical toolboxes in closer collaboration with statisticians in recent decades. While ecotoxicology has benefitted from the progress of these related disciplines, regulatory risk assessments are still largely based on statistical principles and approaches that can no longer be considered state-of-the-art.

Within the Society for Environmental Toxicology and Chemistry (SETAC), recent initiatives related to statistics with regulatory relevance have experienced a high level of interest and discussion. Activities initiated in 2024 include the proposal to establish a SETAC interest group for statistics and the launch of a special series on statistical methods in *Integrated Environmental Assessment and Management*. Moreover, many ecotoxicologists have expressed interest in the revision of OECD document no. 54, “Current approaches in the statistical analysis of ecotoxicity data: a guidance to application” (OECD, 2006). The document is no longer considered reflective of contemporary statistical methods or computational platforms available to ecotoxicology, and a revision is planned for 2026. The German Environment Agency (UBA) will coordinate scientific contributions to the revision process (Daniels et al., 2024), and organized the “1st UBA expert workshop on the OECD No. 54 revision” in September 2024 focusing on three chapters from OECD document no. 54: chapter 4, “Hypothesis testing,” chapter 5, “Dose-response models,” and chapter 6, “Biological effect models.” While a summary from the workshop by the organizers can be expected at a later point, we will here share some reflections on chapters 4 and 5.

In its current version, OECD document no. 54 proposes a distinction between “hypothesis testing” (ANOVA-type models) and

“dose-response modeling” (regression models). This dichotomy neglects that common statistical concepts are underlying both approaches. Both ANOVA-type models and regressions are variants of linear models in statistics. Both involve estimation of parameters, and both use likelihood or information criteria to compare hypotheses. The main difference in this context is that “hypothesis testing” treats concentrations as categories, while “dose-response models” use the concentration as a continuous predictor variable to estimate parameters of a regression equation.

Ecotoxicologists now have access to a wide variety of statistical tools for combined hypothesis testing and dose-response modeling via generalized linear models (GLMs) as well as more advanced and flexible nonlinear models. Basic GLMs can make use of link functions (generalization) instead of data transformation to mimic normal-based models, for example, for logistic regression. Hierarchical versions (mixed-effect GLMs) can better capture nested structures and variability in data (Green, 2015). Furthermore, generalized additive models (GAMs) allow description of relationships between predictor variables and the response by smooth curves (Pedersen et al., 2019); these can therefore be powerful tools for exploring nonlinear patterns in dose-response data. Open-source statistical software, such as R (R Core Team, 2024), has made these methodologies and frameworks readily available to scientists and regulators worldwide.

Dose-response curves can be fitted with a range of models containing 2–5 parameters (Ritz et al., 2015), for obtaining parameter values or derived values such as the EC_x (concentration associated with x% effect). Alternative metrics such as the benchmark dose (BMD; EFSA Scientific Committee et al., 2022) and the no-significant effect concentration (NSEC; Fisher & Fox, 2023) have been proposed more recently. These alternatives have different properties, e.g., regarding sensitivity to low sample size, computational complexity, and suitability as input to higher-level assessment such as species sensitivity distributions. Further exploration and comparison of these methods can help us gain more insight into these alternatives, and, for example, establish criteria for recommending one method over another.

Our reflections can be concluded with a “wish list” for the OECD document No. 54 revision process:

- Clarify the connections between hypothesis testing (chapter 4) and dose-response models (chapter 5). Let continuous regression-based models be the default choice whenever possible and ANOVA-type models be the special case when strictly necessary.
- Include contemporary and well-established statistical tools, e.g., GLMs, in the recommended toolbox for ecotoxicologists, ideally supported with references to well-established statistical software, packages, and online platforms.
- Present Bayesian methods as an alternative to their frequentist counterparts; both frameworks come with strengths and weaknesses, and both should adhere to well-established best practices.

In the longer term and more generally, we would welcome these developments:

- Stronger collaboration with regulatory authorities and jurisdictions for better clarification of issues such objectives of ecotoxicity testing, statistical design considerations, recommendations on modes of analysis, and terminology.
- Higher awareness regarding the importance of statistical design and methodology for reducing animal testing in ecotoxicology.
- More investment in training in statistical science and data literacy for ecotoxicologists, beyond sporadic courses and specialized fields of research within ecotoxicology. This can include topics such as experimental design, visual inspection, best practices for all types of regression models, and knowledge of Bayesian frameworks.
- Better integration of process-based and statistical models for obtaining toxicity estimates.

In moving forward, progress should integrate learning from the past decades (Chapman et al., 1996). Improved statistical practices for regulatory ecotoxicology can support a better assessment and decision-making by authorities regarding chemical hazard and risk. With significant interest from industry, academia, and regulators aligning in recent years, the time seems ripe for an overhaul.

Author contributions

S. Jannicke Moe (Conceptualization, Investigation, Methodology, Writing—original draft, Writing—review & editing), David R. Fox

(Conceptualization, Methodology, Writing—original draft, Writing—review & editing), Raoul Wolf (Conceptualization, Methodology, Writing—original draft, Writing—review & editing)

References

- Chapman, P. F., Crane, M., Wiles, J., Noppert, F., & McIndoe, E. (1996). Improving the quality of statistics in regulatory ecotoxicity tests. *Ecotoxicology (London, England)*, 5, 169–186.
- Daniels, B., Gräff, T., Kotschik, P., & Walter-Rohde, S. (2024). Activities to revise the OECD document no. 54 on statistical analysis of ecotoxicity data. Poster presentation. SETAC Europe 34th Annual Meeting, Spain.
- EFSA Scientific Committee; S. J., More, V., Bampidis, D., Benford, C., Bragard, T. I., Halldorsson, A. F., Hernández-Jerez, S. H., Bennekou, K., Koutsoumanis, C., Lambré, K., Machera, W., Mennes, E., Mullins, S. S., Nielsen, D., Schrenk, D., Turck, M., Younes, M., Aerts, L., Edler, S., Sand, ... J., Schlatter (2022). Guidance on the use of the benchmark dose approach in risk assessment. *EFSA Journal*, 20, e07584.
- Fisher, R., & Fox, D. R. (2023). Introducing the no-significant-effect concentration. *Environmental Toxicology and Chemistry*, 42, 2019–2028.
- Green, J. W. (2015). The challenge: Statistical challenges in ecotoxicology. *Environmental Toxicology and Chemistry*, 34, 2437–2437.
- Laskowski, R. (1995). Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology. *Oikos*, 73, 140.
- OECD. (2006) *Current approaches in the statistical analysis of ecotoxicity data: A guidance to application* (annexes to this publication exist as a separate document). OECD Series on Testing and Assessment, no. 54. OECD Publishing. <https://doi.org/10.1787/9789264085275-en>
- Pedersen, E., Miller, D., Simpson, G., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, 7, e6876.
- Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-response analysis using R. *PLoS One*, 10, e0146021.
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- van Dam, R. A., Harford, A. J., & Warne, M. S. (2012). Time to get off the fence: The need for definitive international guidance on statistical analysis of ecotoxicity data. *Integrated Environmental Assessment and Management*, 8, 242–245.