

# STATISTICAL ECO(-TOXICO)LOGY

IMPROVING THE UTILISATION OF DATA FOR  
ENVIRONMENTAL RISK ASSESSMENT

by

EDUARD SZÖCS

from ZĂRNEȘTI / ROMANIA

Submitted Dissertation thesis for the partial fulfillment of the requirements for a  
Doctor of Natural Sciences  
Faculty 7: Natural and Environmental Sciences  
University Koblenz-Landau

20. November 2016



## Galileo

Few may hear Galileo's song (calling)  
A tribulation  
Adversities  
Fuel for a living, feeds us all

Spirit is fire  
Uncompromising  
Hidden hand, protect us from  
The dead and dying

Echo his madness  
His heresy feeds us all

Spirit is fire  
Uncompromising  
Hidden hand, protect us from  
The dead and dying

Few may

Spirit is fire  
Feed on the senseless ending

Spirit is fire  
Feed on the senseless ending

*- Maynard James Keenan*



# GENERAL DISCUSSION AND OUTLOOK

## TOPICS IN STATISTICAL ECOTOXICOLOGY

The simulation study performed in chapter ?? clearly showed that common experimental designs in ecotoxicology exhibit unacceptably low statistical power (Szöcs and Schäfer, 2016; Van Der Hoeven, 1998). This underpins the criticism accumulated over the last 30 years towards the usage of NOEC as an endpoint for ERA (Fox and Landis, 2016). Nevertheless, the NOEC is still one of the standard endpoints of mesocosm experiments in higher tier risk assessment (EFSA, 2013). Therefore, further advances in the statistical evaluation of mesocosm experiments is needed.

Recently, *a posteriori* calculations of statistical power have been proposed to counteract these limitations and aid the interpretation treatment-related effects in model ecosystems (Brock et al., 2015). The "minimum detectable difference" (MDD) estimates the difference between two means that must exist in order to produce a statistically significant result ( $p < 0.05$  (Gelman and Stern, 2006)) and could be used to interpret NOEC. However, *a posteriori* calculations have been shown to have logical flaws when used for interpretation of non-significant results (Hoenig and Heisey, 2001; Nakagawa and Foster, 2004). However, conducting and reporting of *a priori* power calculations, as performed in chapter ??, might provide researchers important information to optimise their study designs, ensuring that their experimental designs have appropriate power and can lead to interpretable results (Johnson et al., 2015).

Moreover, similar simulations could not only be used to study factorial but also regression designs. Indeed, simulations could be used to determine the optimal experimental design for dose-response models and  $EC_x$  determination, balancing precision and usage of resources. Regression designs are generally more powerful and provide more information than factorial designs (Cottingham et al., 2005). In mesocosm experiments, such designs, assigning the replicates to more tested concentrations, might also provide additional insights. Currently, statistical tools to analyse a community dose-response relationship are not well explored and no equivalent  $EC_{x,mesocosm}$  can be derived. On possibility could be to fit separate dose-response models to each species, leading

to a  $EC_x$  for each species in a mesocosm study. Subsequently, these  $EC_x$  values could be combined and summarised using Species Sensitivity Distributions (Posthuma et al., 2002), providing a single measure for the community response, e.g. a hazardous concentration ( $HC_{x,mesocosm}$ ) for  $x$  % of species affected in mesocosms (Maltby et al., 2005). Another possibility would be to use a logistic type of ordination (van den Brink et al., 2003). Reduced-Rank vector generalised linear models (RR-VGLM) could be used to fit such type of models (Yee, 2015; Yee and Hastie, 2003), but they have not been applied in ecotoxicology yet.

In a similar vein, community ecology is currently experiencing a shift towards a new class of multivariate methods, incorporating statistical models for abundances across many taxa simultaneously (ter Braak and Šmilauer, 2015; Warton et al., 2015a; Warton et al., 2015b; Warton et al., 2012). However, these methods have not been applied frequently and their applicability to ecotoxicological data is currently unclear (Szöcs et al., 2015). All these models have in common, that the choice of statistical model is primarily based on the grounds of data properties. In chapter ?? we showed, that using statistical models that fit the type of data analysed, can provide higher statistical power. Simultaneously, Ives (2015) published a study reaching contradictory conclusions (*"For testing the significance of regression coefficients, go ahead and log-transform count data"*). It must be noted, that the simulation designs differed significantly between both studies: We used a low-replicated factorials design, whereas Ives (2015) simulated a well-replicated regression designs with two predictors. We both found that the negative-binomial GLMs were surprisingly prone to Type I errors, although the assumptions of this model closely matched the data. Nevertheless, as we show in chapter ??, the parametric bootstrap might provide a solution to this problem, but is computationally intensive and not widely used. The parametric bootstrap is akin to Bayesian methods (Gelman et al., 2014), which might provide an alternative inference method for inference. The main point, leading Ives (2015) to his conclusions, was that GLM showed undesirable Type I errors in case of correlated predictors, a case not commonly encountered in ecotoxicology and not studied in chapter ??. Recently, the current state-of-the-art was discussed by Warton et al. (2016): i) choose the statistical model based on the grounds of data properties; ii) fix Type I errors using parametric bootstrap or resampling; iii) take mean-variance relationship into account, which is in line with the findings of chapter ??. However, there are still open questions regarding the use of GLMs for count data (see e.g. raised by Prof. John Maindonald, <http://uni-ko-ld.de/fb>). To diagnose issues such as overdispersion

and excess of zeros in count data models new tools like the recently developed "*Rootograms*" (Kleiber and Zeileis, 2016) provide useful additions.

In chapter ?? we applied new statistical modelling techniques that explicitly consider the limit of quantification. The currently most often used methods to deal with such censored data is to omit or substitute non-detects. Censoring is very common when dealing with chemical and ecological datasets, but is rarely taken into account (Fox et al., 2015). Recent examples from ecotoxicology and environmental chemistry show that omission (Hansen et al., 2015), randomization (Goulson, 2015) or substitution by a fixed value (Helsel, 2010; Helsel, 2006) can lead to biased results. Hansen et al. (2015) used a Tobit regression (Tobin, 1958) that takes the amount of censored data into account, assuming a (log-) normal distribution of concentrations. In chapter ?? we used a slightly different approach, using a zero adjusted gamma model (ZAGA). We modelled measured concentrations as two separate processes, generating i) zero values and ii) non-zero values assuming a gamma distribution of concentrations. In ecological statistics this type models is also known as *hurdle* models (Martin et al., 2005). The log-normal Tobit model has no probability mass at zero, whereas ZAGA model has a probability at zero. Generally, the difference between Tobit and two-part models are small (Min and Agresti, 2002) and the same holds for differences between the log-normal and Gamma distribution. Indeed, a Tobit-like model could be also fitted assuming a Gamma distribution (Sigrist and Stahel, 2010).

Grab sampling likely underestimates chemical concentrations because of short term peak concentrations (Stehle et al., 2013; Xing et al., 2013). Although this leads to an increased variation in chemical measurements, we still can learn from the process generating values above LOQ, even if the absolute value is subject to error. This is also highlighted by the results of chapter ??, with estimated coefficients for the absolute concentration showing much larger uncertainty than coefficients for the probability of exceeding LOQ (Figure ??). Currently, models explicitly taking the censored nature of chemical monitoring data are not well explored and seldom applied. Further research on those is needed and might provide useful information for analysing monitoring data, assessing the chemical status and trends thereof.

## LEVERAGING MONITORING DATA FOR ECOLOGICAL RISK ASSESSMENT

In chapter ?? we compiled and analysed monitoring data leading to the currently most extensive dataset on pesticide exposure available for Germany. We demonstrated, that within the current monitoring scheme small streams below 10 km<sup>2</sup> are underrepresented (Figure ??, top). Given their importance, we must admit that we currently do not have much knowledge about these and their threats (Biggs et al., 2016; Lorenz et al., 2016). To fill these gaps, monitoring networks need to be adapted to give a better representation of small streams.

We analysed only data from small streams, however, data on lentic systems is even more scarce. Although, more the 95% of German standing waters are lentic small waterbodies a recent meta-analysis revealed that only 5% of studies investigating pesticides in freshwaters were performed on lentic small water bodies Lorenz et al. (2016). The data compiled in chapter ?? comprised also lentic small water bodies. However, the query to the federal states revealed that there are currently no such monitoring data available (Brinke et al., 2016). This highlights the urgent need to adapt monitoring schemes to also include small standing waters.

Our results revealed that chemical monitoring schemes within Germany differed largely in terms of spatio-temporal coverage and compound spectra between federal states. Similarly, Malaj et al. (2014) showed big differences between European countries. Overall, a homogenisation and standardisation of chemical monitoring programs would enhance the comparability and the possibility for a large-scale assessment.

We found that the signal from agricultural pesticides can be detected down to a small percentage of agriculture within the catchment (Figure ??). Thus, we can conclude that if there is agriculture with a catchment, it is very likely that pesticides will be applied, enter the streams and are detected. This has implications for selection of reference sites for environmental monitoring, that need to have no agricultural influence. We studied only the influence of agricultural non-point sources, however, point-sources like wastewater treatment plants can also contribute to pesticide pollution of streams (Bunzel et al., 2014).

We were able to detect a small, but distinct increase of risks after precipitation events. This is in line with findings that pesticides mainly enter surface waters via edge-of-field runoff (Schulz, 2001). Moreover, our results suggest that abso-



lute measured concentrations are subject to a high error due to the sampling process and adds evidence that current monitoring schemes, largely unconnected from precipitation events, underestimate pesticide risks (Stehle et al., 2013; Xing et al., 2013). Automatic event-driven samplers in small streams could provide knowledge on pesticide risk dynamics that are currently unknown.

Monitoring data can provide an opportunity to inform ERA after authorization and could possibly trigger a refinement of the assessment (Knauer, 2016). However, current monitoring mainly addresses streams bigger than those considered in ERA. Our results indicated that small streams are frequently exposed to high risks from pesticides. Moreover, to provide a suitable feedback for ERA small agricultural streams must be integrated into environmental monitoring schemes. As the measurements within the current monitoring schemes provide an underestimation, all exceedances of risk thresholds represent an unacceptable risk and indicate that current ERA might miss potential risks and further enhancements of the current authorisation procedure are needed.

Risk thresholds in chapter ?? were especially exceeded for the organophosphate Chlorpyrifos and neonicotinoid substances. This adds to the existing evidence that this particular class of insecticides poses currently a high threat to freshwaters and stricter regulations are warranted (Goulson, 2013; Morrissey et al., 2015). The high number of exceedances shows that ERA for these substances was not accurate enough and lead to risks for the environment. Recent studies investigating large-scale pesticide risks did not consider neonicotinoid insecticides (Malaj et al., 2014; Stehle and Schulz, 2015) and therefore likely underestimated the risks to freshwaters. However, this also shows that the analysed spectrum is an important driver of detected compounds (Malaj et al., 2014; Schreiner et al., 2016) and must be taken into account when evaluating monitoring data for risk assessment. The WFD currently considers on a few, relatively well-known substances (European Union, 2013) and a status assessment based on EQS likely missed the actual chemical pollution (Moschet et al., 2014). Compared to the data presented in chapter ?? the WFD considers currently only 4% of the pesticides (19 out of 478 pesticides). Recently, neonicotinoid substances have been incorporated in the watch list of substances for Union-wide monitoring (European Union, 2015), so that more information on the environmental fate of compound group will be available. Monitoring data provide also valuable information for the prioritisation of emerging pollutants and future monitoring with the WFD (Brack et al., 2017) and the compiled data we compiled could be a valuable input for such a prioritisation.

Monitoring under the WFD is also performed for biological components of freshwaters and a combination with pesticide exposure data might provide valuable insights into large-scale field effects of chemical substances (Schipper et al., 2014). Currently, chemical and biological monitoring are not synchronised. On a continental scale, Malaj et al. (2014) was able to compile ecological status data only for 5% of sites with chemical measurements. For the dataset presented in chapter ?? we found a spatial match with biological monitoring for 60% of sites (Brinke et al., 2016). However, as biological data in Germany is sampled at lower frequencies (often less than once per year) a spatio-temporal match would result in much less accordance. Synchronising these samplings in a future monitoring would possibly enable to assess large-scale post-authorization field effects of chemical substances.

## CHALLENGES UTILISING 'BIG DATA' IN ECOLOGICAL RISK ASSESSMENT

Effect assessment and environmental monitoring produce huge amounts of data. However, the profoundness of ecological risk assessment often determined by the available data (Van den Brink et al., 2016). Useful data for ERA is currently spread over several largely unconnected databases. E.g. ecotoxicity data is spread over database maintained by the U.S. EPA (ECOTOX, U.S. EPA (2016)), the University of Hertfordshire (PPDB, Lewis et al. (2016)), the German Environment Agency (ETOX, Umweltbundesamt (2016)) and others. Chemical information is similarly spread over several databases, like PubChem (Kim et al., 2016) or Chempider (Pence and Williams, 2010). Additional complications arise because these databases use different identifiers for chemical substances. The U.S. EPA (U.S. EPA, 2016) uses solely the CAS-Number for identification, whereas other databases uses SMILES (Weininger, 1990) or InChI and InChIKeys (Heller et al., 2015). Integrating these databases is currently a challenge in ERA, which is complicated by ambiguous identifiers, e.g. should different salts be considered separately for aquatic risk assessment? Projects like the NORMAN EMPO-DAT database (Brack et al., 2012) or the STOFF-IDENT (Huckele and Track 2013, <http://uni-ko-ld.de/fc>) are first attempts for such an integration. Integration monitoring data and risk assessment data is a mandatory requirement for landscape level ecotoxicology and risk assessment (Focks, 2014) and needed for an improved model development and validation (Brock et al., 2006; Knäbel et al.,

2012). Chapter ?? is an example for such an integration, but represents only a preliminary assessment and spatial-temporal risk dynamics should be further investigated.

The webchem package, presented in chapter ??, can foster such an integration. However, data must also be accessible in order to be retrievable by webchem. Unfortunately, major parts of data produce for environmental risk assessment are not available (Dafforn et al., 2015; Schäfer et al., 2013). Recently, it has been demonstrated that data from the European Registration, Evaluation, Authorisation, and Restriction of Chemicals (REACH) database can be used to improve the characterisation of ecotoxicity in life cycle assessment (LCA) (Müller et al., 2016). Although this database hosts humongous amounts of data used risk assessment, it is currently not available in a convenient way. Indeed, a systematic data collection contravenes the legal usage of the REACH database (<http://uni-ko-ld.de/fd>). This may be also the reason, why the quality of chemical property data submitted this database is currently unknown (Müller et al., 2016; Stieger et al., 2014). Webchem currently can retrieve data from 11 data sources. However many other data sources are available and implementation of more than 15 other data sources in the pipeline will be implemented in the future (<http://uni-ko-ld.de/fi>) and collaborators are welcome.

The software tools described in chapters ?? and ?? assist researchers handling and cleaning their data. Aggregating taxonomic data to a higher taxonomic level is a common task when analysing data from mesocosm experiments or from field sampling. Taxize facilitates the retrieval of taxonomic classification, which is the basis also for more sophisticated aggregation methods (Cuffney et al., 2007). Today, taxize has been used in more than thirty scientific publications. Recent applications of the webchem package, have been demonstrated by Münch and Galizia (2016) and Ranke (2016): Münch and Galizia (2016) compiled a database for odorant responses of *Drosophila melanogaster* and webchem "likely saved [him] hundreds of working hours". Ranke (2016) is using webchem to compile and store chemical information for further analyses. The analyses performed in chapter ?? needed to integrate monitoring, chemical and risk assessment data which would not have been difficult without the webchem package. These examples show that researchers have been missing such tools in the past. If they can reduce their time spend on data retrieval and handling, they could focus more on the quintessence of their research.

## CONCLUSIONS

In the near future, big amounts of data will be available for environmental risk assessment. Integration and analysis of these data are forthcoming challenges in ecotoxicology (Dafforn et al., 2015; Van den Brink et al., 2016). This thesis provides insights into statistical analyses and experimental designs for a future effect assessment. Statistical ecotoxicology is just at the beginning and many problems are pending to be solved.

Environmental monitoring data can provide an important feedback to improve environmental risk assessment. The integration of both in this thesis showed that in Germany highly toxic insecticides pose a major threat to freshwaters. We demonstrated current problems integrating these and further improvements of those are needed to safeguard freshwaters from chemicals. It needs to be re-evaluated if the current use of neonicotinoid insecticides is within the safe operating space to provide long-term health of ecosystems upon which humanity depends (Rockström et al., 2009).

Big data and modern statistical tools are a means of improving the accuracy and reducing the uncertainty of environmental risk assessment (Van den Brink et al., 2016). The software described in this thesis contribute its part to these improvements. Nevertheless, science and politics need to develop a culture of openness to promote the safety of our environment.

## REFERENCES

- Biggs, J., S. von Fumetti, and M. Kelly-Quinn (2016). "The Importance of Small Waterbodies for Biodiversity and Ecosystem Services: Implications for Policy Makers". *Hydrobiologia*.
- Brack, W., V. Dulio, M. Ågerstrand, I. Allan, R. Altenburger, M. Brinkmann, D. Bunke, R. M. Burgess, I. Cousins, B. I. Escher, F. J. Hernández, L. M. Hewitt, K. Hilscherová, J. Hollender, H. Hollert, R. Kase, B. Klauer, C. Lindim, D. L. Herráez, C. Miège, J. Munthe, S. O'Toole, L. Posthuma, H. Rüdél, R. B. Schäfer, M. Sengl, F. Smedes, D. van de Meent, P. J. van den Brink, J. van Gils, A. P. van Wezel, A. D. Vethaak, E. Vermeirssen, P. C. von der Ohe, and B. Vrana (2017). "Towards the Review of the European Union Water Framework Management of Chemical Contamination in European Surface Water Resources". *Science of The Total Environment* 576. 00000, 720–737.
- Brack, W., V. Dulio, and J. Slobodnik (2012). "The NORMAN Network and Its Activities on Emerging Environmental Substances with a Focus on Effect-Directed Analysis of Complex Environmental Contamination". *Environmental Sciences Europe* 24 (1), 1–5.
- Brinke, M., E. Szöcs, K. Foit, M. Liess, R. B. Schäfer, B. Bänsch-Baltruschat, and M. Keller (2016). *Umsetzung des Nationalen Aktionsplans zur nachhaltigen Anwendung von Pestiziden – Bestandsaufnahme zur Erhebung von Daten zur Belastung von Kleingewässern in der Agrarlandschaft*. Endbericht FKZ 3714 67 4040/1, unpublished. Umweltbundesamt.
- Brock, T. C. M., G. Arts, L. Maltby, and P. van den Brink (2006). "Aquatic Risks of Pesticides, Ecological Protection Goals and Common Aims in EU Legislation". *Integrated Environmental Assessment and Management* 2 (4), e20–e46.
- Brock, T. C. M., M. Hammers-Wirtz, U. Hommen, T. G. Preuss, H.-T. Ratte, I. Roessink, T. Strauss, and P. J. Van den Brink (2015). "The minimum detectable difference (MDD) and the interpretation of treatment-related effects of pesticides in experimental ecosystems". *Environmental Science and Pollution Research* 22 (2), 1160–1174.

- Bunzel, K., M. Liess, and M. Kattwinkel (2014). "Landscape Parameters Driving Aquatic Pesticide Exposure and Effects". *Environmental Pollution* 186, 90–97.
- Cottingham, K. L., J. T. Lennon, and B. L. Brown (2005). "Knowing When to Draw the Line: Designing More Informative Ecological Experiments". *Frontiers in Ecology and the Environment* 3 (3), 145–152.
- Cuffney, T. F., M. D. Bilger, and A. M. Haigler (2007). "Ambiguous Taxa: Effects on the Characterization and Interpretation of Invertebrate Assemblages". *Journal of the North American Benthological Society* 26 (2), 286–307.
- Dafforn, K. A., E. L. Johnston, A. Ferguson, C. Humphrey, W. Monk, S. J. Nichols, S. L. Simpson, M. G. Tulbure, and D. J. Baird (2015). "Big Data Opportunities and Challenges for Assessing Multiple Stressors across Scales in Aquatic Ecosystems." *Marine and Freshwater Research*.
- EFSA (2013). "Guidance on Tiered Risk Assessment for Plant Protection Products for Aquatic Organisms in Edge-of-Field Surface Waters". *EFSA Journal* 11 (7). 00014, 3290.
- European Union (2013). "Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 Amending Directives 2000/60/EC and 2008/105/EC as Regards Priority Substances in the Field of Water Policy". *Official Journal of the European Union* L226, 1–17.
- European Union (2015). "Commission Implementing Decision (EU) 2015/495 of 20 March 2015 Establishing a Watch List of Substances for Union-Wide Monitoring in the Field of Water Policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council (Notified under Document C(2015) 1756)". *Official Journal of the European Union* L28, 40–42.
- Focks, A. (2014). "The Challenge: Landscape Ecotoxicology and Spatially Explicit Risk Assessment". *Environmental Toxicology and Chemistry* 33 (6), 1193–1193.
- Fox, D. R. and W. G. Landis (2016). "Comment on ET&C Perspectives, November 2015-A Holistic View". *Environmental Toxicology and Chemistry* 35 (6), 1337–1339.

- Fox, G. A., S. Negrete-Yankelevich, and V. J. Sosa, eds. (2015). *Ecological Statistics: Contemporary Theory and Application*. 00011. Oxford: Oxford University Press. 389 pp.
- Gelman, A. and H. Stern (2006). "The difference between "significant" and "not significant" is not itself statistically significant". *The American Statistician* 60 (4), 328–331.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2014). *Bayesian Data Analysis*. CRC Press.
- Goulson, D. (2013). "An Overview of the Environmental Risks Posed by Neonicotinoid Insecticides". *Journal of Applied Ecology* 50 (4), 977–987.
- Goulson, D. (2015). "Neonicotinoids Impact Bumblebee Colony Fitness in the Field; a Reanalysis of the UK's Food & Environment Research Agency 2012 Experiment". *PeerJ* 3, e854.
- Hansen, C. T., C. Ritz, D. Gerhard, J. E. Jensen, and J. C. Streibig (2015). "Re-Evaluation of Groundwater Monitoring Data for Glyphosate and Bentazone by Taking Detection Limits into Account". *Science of the Total Environment* 536, 68–71.
- Heller, S. R., A. McNaught, I. Pletnev, S. Stein, and D. Tchekhovskoi (2015). "InChI, the IUPAC International Chemical Identifier". *Journal of Cheminformatics* 7 (1).
- Helsel, D. (2010). "Much Ado about next to Nothing: Incorporating Nondetects in Science". *Ann Occup Hyg* 54 (3), 257–62.
- Helsel, D. R. (2006). "Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it". *Chemosphere* 65 (11), 2434–2439.
- Hoening, J. M. and D. M. Heisey (2001). "The abuse of power". *The American Statistician* 55 (1), 19–24.

- Huckele, S. and T. Track (2013). "Risk Management of Emerging Compounds and Pathogens in the Water Cycle (RiSKWa)". *Environmental Sciences Europe* 25 (1), 1.
- Ives, A. R. (2015). "For testing the significance of regression coefficients, go ahead and log-transform count data". *Methods in Ecology and Evolution* 6 (7), 828–835.
- Johnson, P. C. D., S. J. E. Barry, H. M. Ferguson, and P. Müller (2015). "Power Analysis for Generalized Linear Mixed Models in Ecology and Evolution". *Methods in Ecology and Evolution* 6 (2), 133–142.
- Kim, S., P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S. H. Bryant (2016). "PubChem Substance and Compound Databases". *Nucleic Acids Research* 44 (D1), D1202–D1213.
- Kleiber, C. and A. Zeileis (2016). "Visualizing Count Data Regressions Using Rootograms". *The American Statistician*, 1–25.
- Knäbel, A., S. Stehle, R. B. Schäfer, and R. Schulz (2012). "Regulatory FOCUS Surface Water Models Fail to Predict Insecticide Concentrations in the Field". *Environmental Science & Technology* 46 (15), 8397–8404.
- Knauer, K. (2016). "Pesticides in Surface Waters: A Comparison with Regulatory Acceptable Concentrations (RACs) Determined in the Authorization Process and Consideration for Regulation". *Environmental Sciences Europe* 28 (13).
- Lewis, K. A., J. Tzilivakis, D. J. Warner, and A. Green (2016). "An International Database for Pesticide Risk Assessments and Management". *Human and Ecological Risk Assessment: An International Journal* 22 (4), 1050–1064.
- Lorenz, S., J. J. Rasmussen, A. Süß, T. Kalettka, B. Golla, P. Horney, M. Stähler, B. Hommel, and R. B. Schäfer (2016). "Specifics and Challenges of Assessing Exposure and Effects of Pesticides in Small Water Bodies". *Hydrobiologia*, 1–12.
- Malaj, E., P. C. von der Ohe, M. Grote, R. Kühne, C. P. Mondy, P. Usseglio-Polatera, W. Brack, and R. B. Schäfer (2014). "Organic Chemicals Jeopardize



- the Health of Freshwater Ecosystems on the Continental Scale". *Proceedings of the National Academy of Sciences* 111 (26), 9549–9554.
- Maltby, L., N. Blake, T. C. M. Brock, and P. J. Van Den Brink (2005). "Insecticide Species Sensitivity Distributions: Importance of Test Species Selection and Relevance to Aquatic Ecosystems". *Environmental Toxicology and Chemistry* 24 (2), 379–388.
- Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham (2005). "Zero Tolerance Ecology: Improving Ecological Inference by Modelling the Source of Zero Observations". *Ecology Letters* 8 (11), 1235–1246.
- Min, Y. and A. Agresti (2002). "Modeling Nonnegative Data with Clumping at Zero: A Survey". *Journal of Iranian Statistical Society* 1 (1), 7–33.
- Morrissey, C. A., P. Mineau, J. H. Devries, F. Sanchez-Bayo, M. Liess, M. C. Cavallaro, and K. Liber (2015). "Neonicotinoid Contamination of Global Surface Waters and Associated Risk to Aquatic Invertebrates: A Review". *Environment International* 74, 291–303.
- Moschet, C., I. Wittmer, J. Simovic, M. Junghans, A. Piazzoli, H. Singer, C. Stamm, C. Leu, and J. Hollender (2014). "How a Complete Pesticide Screening Changes the Assessment of Surface Water Quality". *Environmental Science & Technology* 48 (10), 5423–5432.
- Müller, N., D. de Zwart, M. Hauschild, G. Kijko, and P. Fantke (2016). "Exploring REACH as a Potential Data Source for Characterizing Ecotoxicity in Life Cycle Assessment: Exploring REACH for Characterizing Ecotoxicity in LCA". *Environmental Toxicology and Chemistry*.
- Münch, D. and C. G. Galizia (2016). "DoOR 2.0 - Comprehensive Mapping of *Drosophila Melanogaster* Odorant Responses". *Scientific Reports* 6, 21841.
- Nakagawa, S. and T. M. Foster (2004). "The Case against Retrospective Statistical Power Analyses with an Introduction to Power Analysis". *Acta Ethologica* 7 (2), 103–108.

- Pence, H. E. and A. Williams (2010). "ChemSpider: An Online Chemical Information Resource". *Journal of Chemical Education* 87(11), 1123–1124.
- Posthuma, L., G. W. Suter, and T. P. Traas (2002). *Species Sensitivity Distributions in Ecotoxicology*. Environmental and ecological risk assessment. Boca Raton and Fla: Lewis.
- Ranke, J. (2016). *Jranke/Chents*. URL: <https://github.com/jranke/chents>.
- Rockström, J., W. Steffen, K. Noone, A. Persson, 3. Chapin F. S., E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sorlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. A. Foley (2009). "A Safe Operating Space for Humanity". *Nature* 461(7263), 472–5.
- Schäfer, R. B., M. Bundschuh, A. Focks, and P. C. von der Ohe (2013). "Letter to the Editor". *Environmental Toxicology And Chemistry* 32(4), 734–735.
- Schipper, A. M., L. Posthuma, D. de Zwart, and M. A. J. Huijbregts (2014). "Deriving Field-Based Species Sensitivity Distributions (F-SSDs) from Stacked Species Distribution Models (S-SDMs)". *Environmental Science & Technology* 48(24), 14464–14471.
- Schreiner, V. C., E. Szöcs, A. K. Bhowmik, M. G. Vijver, and R. B. Schäfer (2016). "Pesticide mixtures in streams of several European countries and the USA". *Science of The Total Environment* 573, 680–689.
- Schulz, R. (2001). "Comparison of spray drift-and runoff-related input of azinphos-methyl and endosulfan from fruit orchards into the Lourens River, South Africa". *Chemosphere* 45(4), 543–551.
- Sigrist, F. and W. A. Stahel (2010). "Using the Censored Gamma Distribution for Modeling Fractional Response Variables with an Application to Loss given Default". *arXiv preprint arXiv:1011.1796*. URL: <http://arxiv.org/abs/1011.1796>.

- Stehle, S., A. Knäbel, and R. Schulz (2013). "Probabilistic Risk Assessment of Insecticide Concentrations in Agricultural Surface Waters: A Critical Appraisal". *Environmental Monitoring and Assessment* 185 (8), 6295–6310.
- Stehle, S. and R. Schulz (2015). "Pesticide Authorization in the EU—environment Unprotected?" *Environmental Science and Pollution Research* 22 (24), 19632–19647.
- Stieger, G., M. Scheringer, C. A. Ng, and K. Hungerbühler (2014). "Assessing the Persistence, Bioaccumulation Potential and Toxicity of Brominated Flame Retardants: Data Availability and Quality for 36 Alternative Brominated Flame Retardants". *Chemosphere* 116, 118–123.
- Szöcs, E. and R. B. Schäfer (2016). "Statistical Hypothesis testing—To Transform or Not to Transform?" *Integrated Environmental Assessment and Management* 12 (2), 398–400.
- Szöcs, E., P. J. van den Brink, L. Lagadic, T. Caquet, M. Roucaute, A. Auber, Y. Bayona, M. Liess, P. Ebke, A. Ippolito, C. J. F. ter Braak, T. C. M. Brock, and R. B. Schäfer (2015). "Analysing Chemical-Induced Changes in Macroinvertebrate Communities in Aquatic Mesocosm Experiments: A Comparison of Methods". *Ecotoxicology* 24 (4), 760–769.
- Ter Braak, C. J. and P. Šmilauer (2015). "Topics in constrained and unconstrained ordination". *Plant Ecology* 216 (5), 683–696.
- Tobin, J. (1958). "Estimation of Relationships for Limited Dependent Variables". *Econometrica* 26 (1), 24.
- Umweltbundesamt (2016). *ETOX: Information System Ecotoxicology and Environmental Quality Targets*. URL: <http://webetox.uba.de/webETOX/index.do>.
- U.S. EPA (2016). *ECOTOX Database*. URL: <http://cfpub.epa.gov/ecotox/>.
- Van den Brink, P. J., C. B. Choung, W. Landis, M. Mayer-Pinto, V. Pettigrove, P. Scanes, R. Smith, and J. Stauber (2016). "New Approaches to the Ecological Risk Assessment of Multiple Stressors". *Marine and Freshwater Research* 67 (4), 429.

- Van den Brink, P. J., N. W. van den Brink, and C. J. F. ter Braak (2003). "Multivariate Analysis of Ecotoxicological Data Using Ordination: Demonstrations of Utility on the Basis of Various Examples". *Australasian Journal of Ecotoxicology* 9, 141–156.
- Van Der Hoeven, N. (1998). "Power Analysis for the NOEC: What Is the Probability of Detecting Small Toxic Effects on Three Different Species Using the Appropriate Standardized Test Protocols?" *Ecotoxicology* 7 (6), 355–361.
- Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. C. Hui (2015a). "So Many Variables: Joint Modeling in Community Ecology". *Trends in Ecology & Evolution* 30 (12), 766–779.
- Warton, D. I., S. D. Foster, G. De'ath, J. Stoklosa, and P. K. Dunstan (2015b). "Model-Based Thinking for Community Ecology". *Plant Ecology* 216 (5), 669–682.
- Warton, D. I., M. Lyons, J. Stoklosa, and A. R. Ives (2016). "Three Points to Consider When Choosing a LM or GLM Test for Count Data". *Methods in Ecology and Evolution*.
- Warton, D. I., S. T. Wright, and Y. Wang (2012). "Distance-based multivariate analyses confound location and dispersion effects". *Methods in Ecology and Evolution* 3 (1), 89–101.
- Weininger, D. (1990). "SMILES. 3. DEPICT. Graphical Depiction of Chemical Structures". *Journal of Chemical Information and Computer Sciences* 30 (3), 237–243.
- Xing, Z., L. Chow, H. Rees, F. Meng, S. Li, B. Ernst, G. Benoy, T. Zha, and L. M. Hewitt (2013). "Influences of Sampling Methodologies on Pesticide-Residue Detection in Stream Water". *Archives of Environmental Contamination and Toxicology* 64 (2), 208–218.
- Yee, T. W. (2015). *Vector Generalized Linear and Additive Models*. Springer Series in Statistics. New York, NY: Springer New York.
- Yee, T. W. and T. J. Hastie (2003). "Reduced-Rank Vector Generalized Linear Models". *Statistical modelling* 3 (1), 15–41.



## COLOPHON

This document was typeset in  $\text{\LaTeX}$  using a slightly customized version of the classicthesis typography by Prof. Miede (<http://www.miede.de/>). Most of the graphics in this thesis are generated using pgf/tikz. The bibliography is typeset using biblatex. This thesis and the source code is also available from [https://github.com/EDiLD/phd\\_thesis](https://github.com/EDiLD/phd_thesis).