

Statistical Eco(-toxico)logy

Improving the Utilisation of Data for
Environmental Risk Assessment

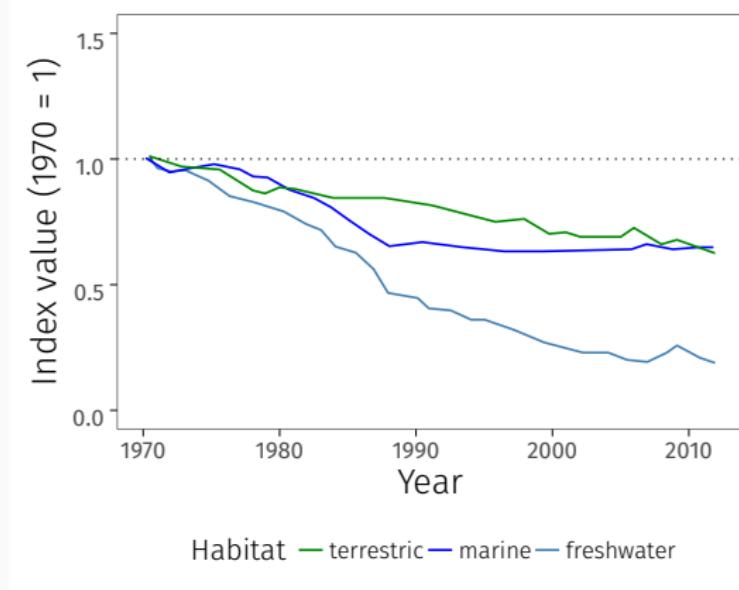
Eduard Szöcs

25th January 2017

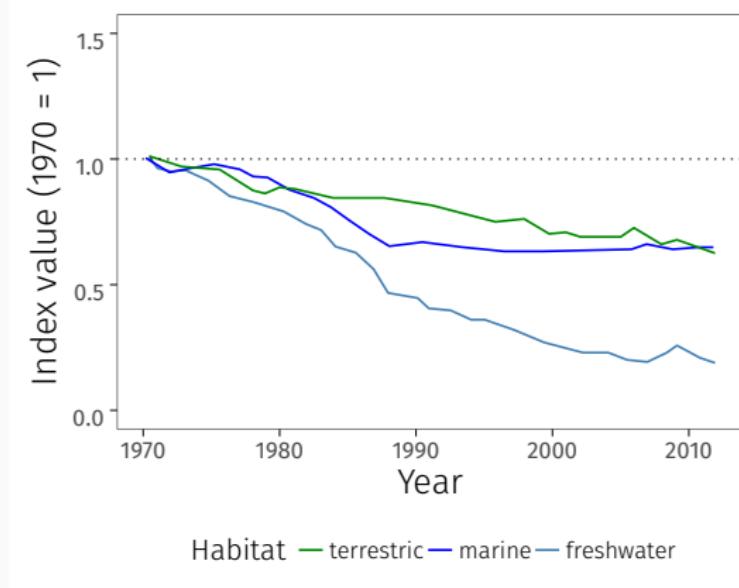
Table of contents

1. Environmental Risk Assessment (ERA) and Monitoring
2. Improving Statistics in ERA
3. Exploring Monitoring Data for ERA
4. Solutions for Linking Data in ERA

Freshwater biodiversity is strongly declining



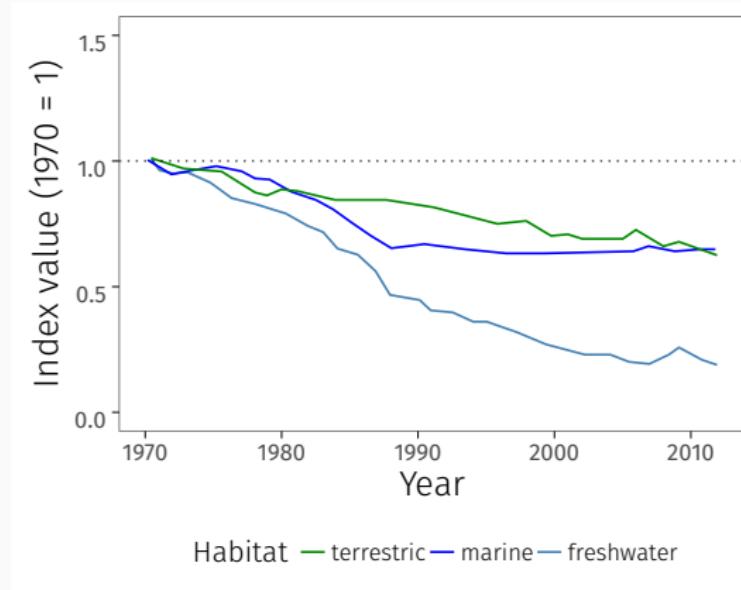
Freshwater biodiversity is strongly declining



Reasons

- Habitat loss
- Overexploitation
- Pollution
- Invasive species

Freshwater biodiversity is strongly declining



Reasons

- Habitat loss
- Overexploitation
- **Pollution**
- Invasive species

Environmental Risk Assessment and Environmental Monitoring

Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

Water Framework Directive
2000/60/EC

Environmental
Risk
Assessment

Environmental
Monitoring

Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

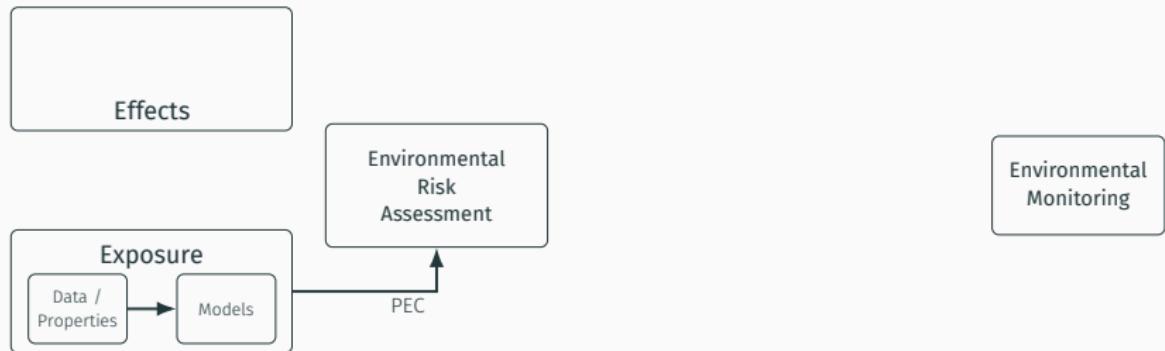
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

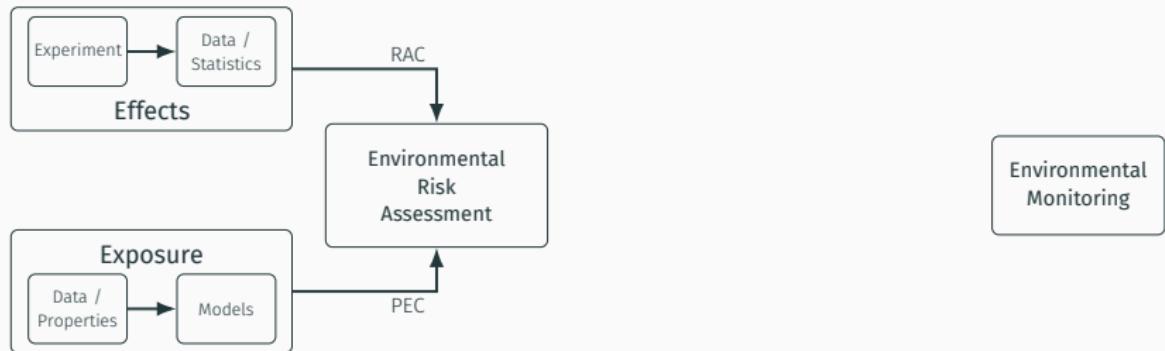
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

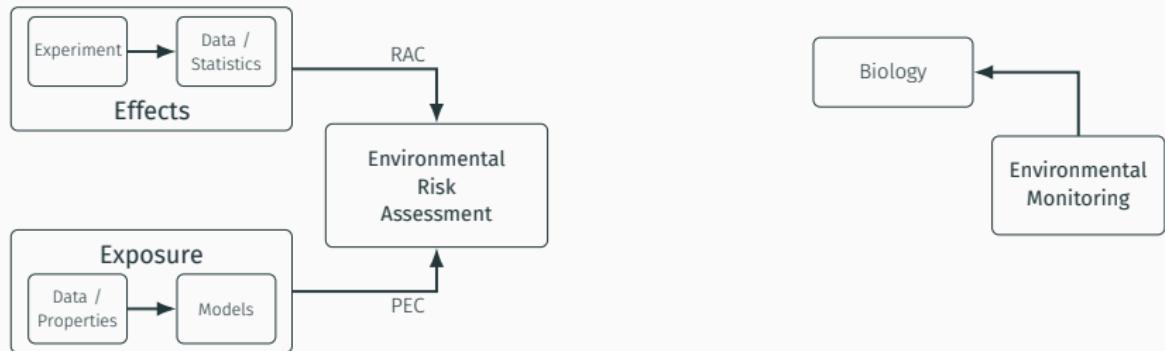
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

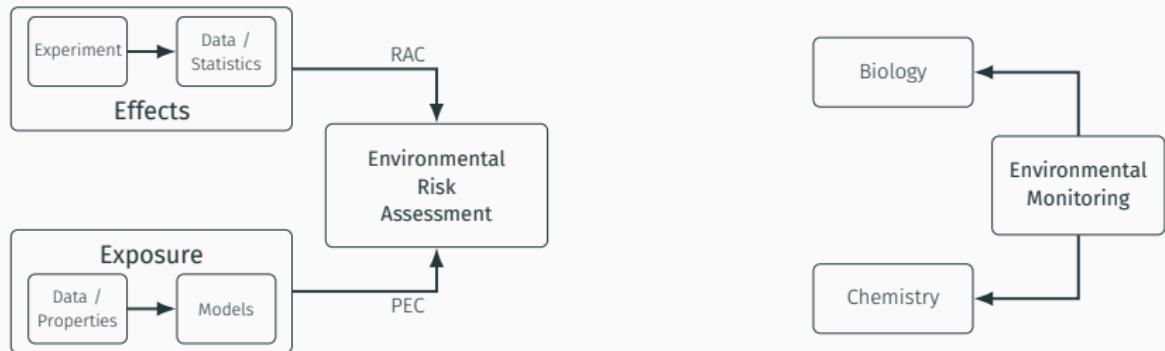
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

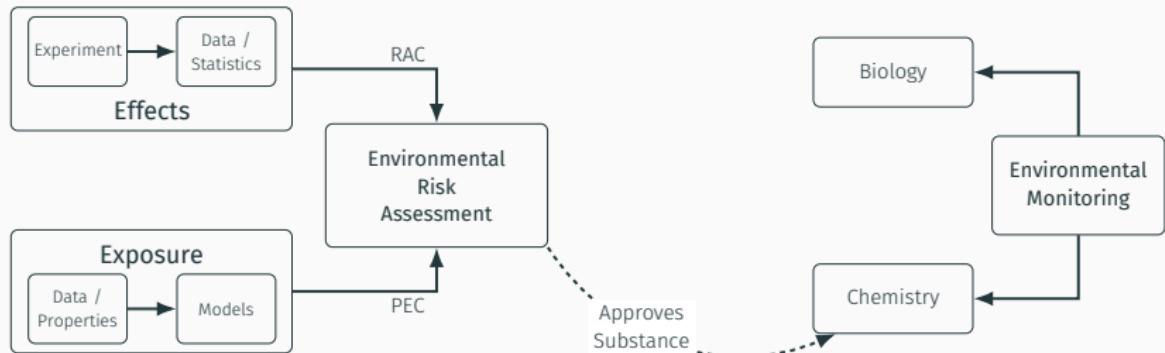
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

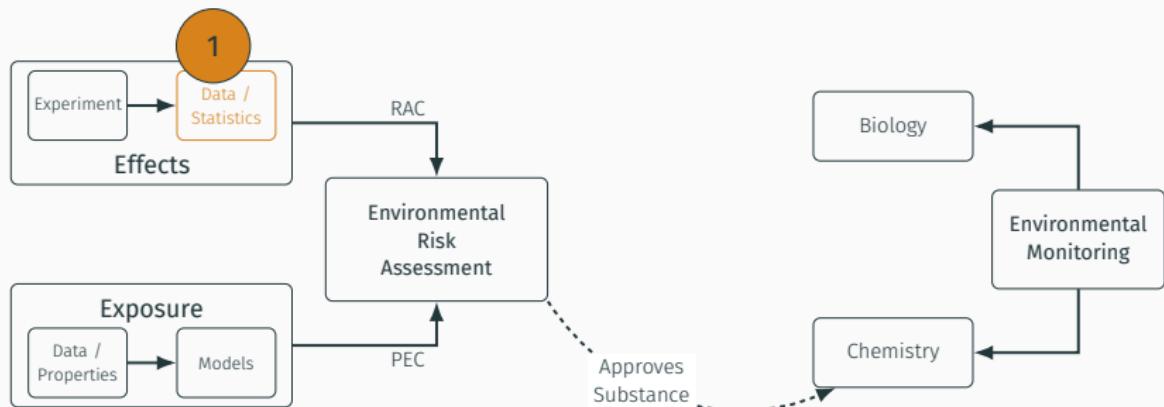
Water Framework Directive
2000/60/EC



Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

Water Framework Directive
2000/60/EC



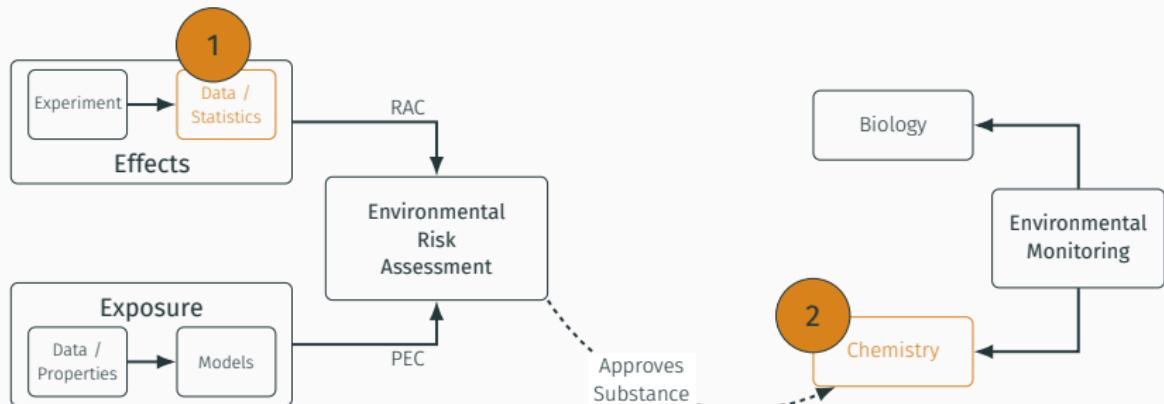
1

Szöcs & Schäfer (2015). "Ecotoxicology is not normal". *ESPR* 22(18), 13990–13999.

Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

Water Framework Directive
2000/60/EC



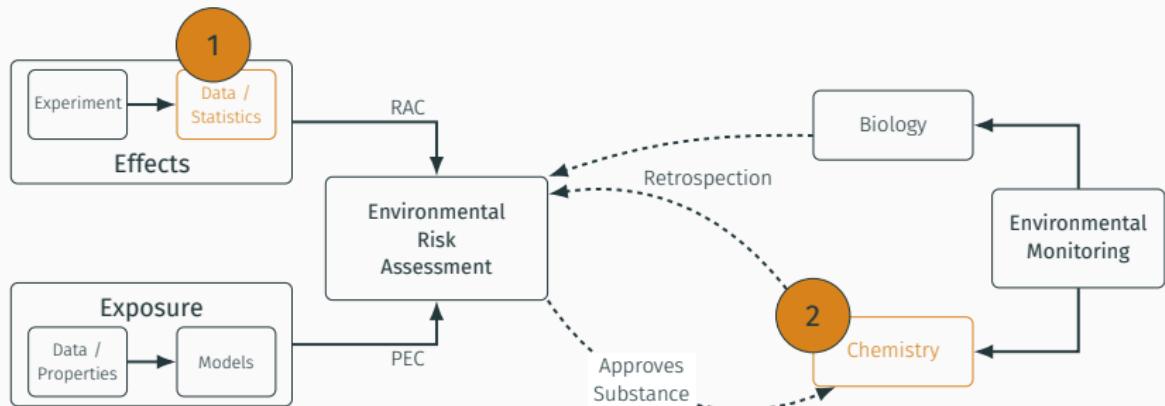
2

Szöcs, Brinke, Karaoglan & Schäfer (submitted). "Large scale risks from pesticides in small streams". ES&T.

Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

Water Framework Directive
2000/60/EC



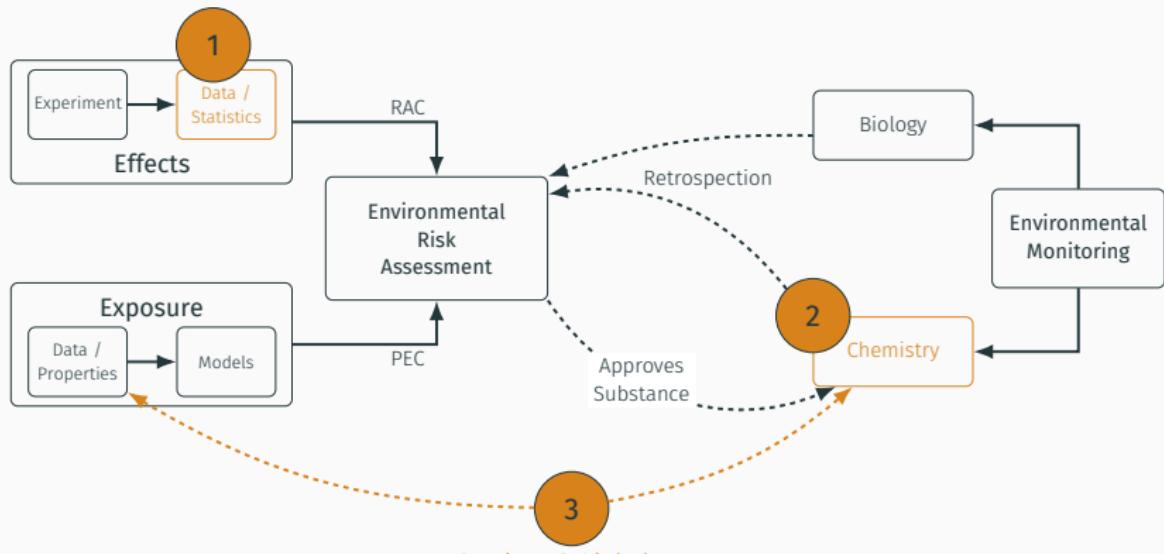
2

Szöcs, Brinke, Karaoglan & Schäfer (submitted). "Large scale risks from pesticides in small streams". ES&T.

Environmental Risk Assessment and Monitoring

Plant Protection Products
1107/2009

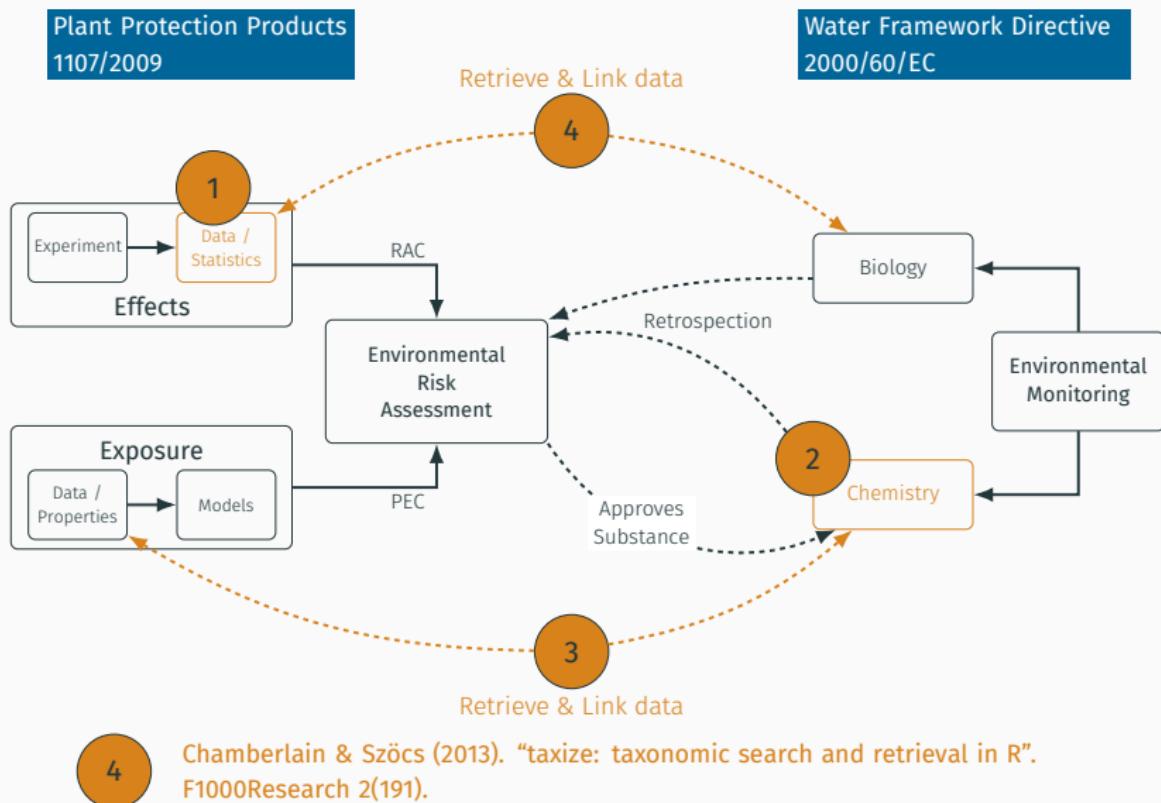
Water Framework Directive
2000/60/EC



3

Szöcs & Schäfer (accepted). "webchem: An R Package to Retrieve Chemical Information from the Web". JSS.

Environmental Risk Assessment and Monitoring



Improving Statistics in ERA

Experiments in Effect Assessment



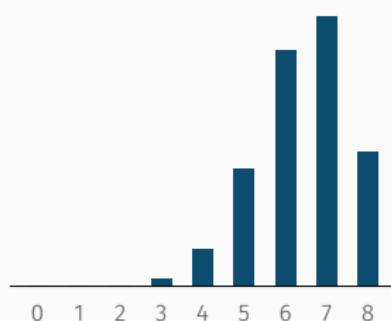
- Daphnia Test
- Lower Tier
- "*x out of n survived*"

Experiments in Effect Assessment

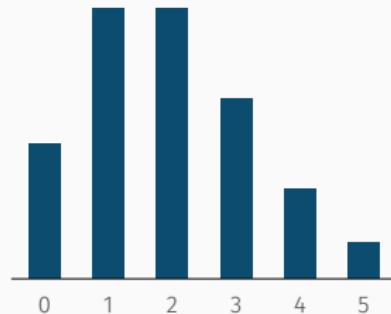


- Daphnia Test
- Lower Tier
- “*x out of n survived*”
- Mesocosm
- Higher Tier
- “*number of animals*”

Ecotoxicology is not normal

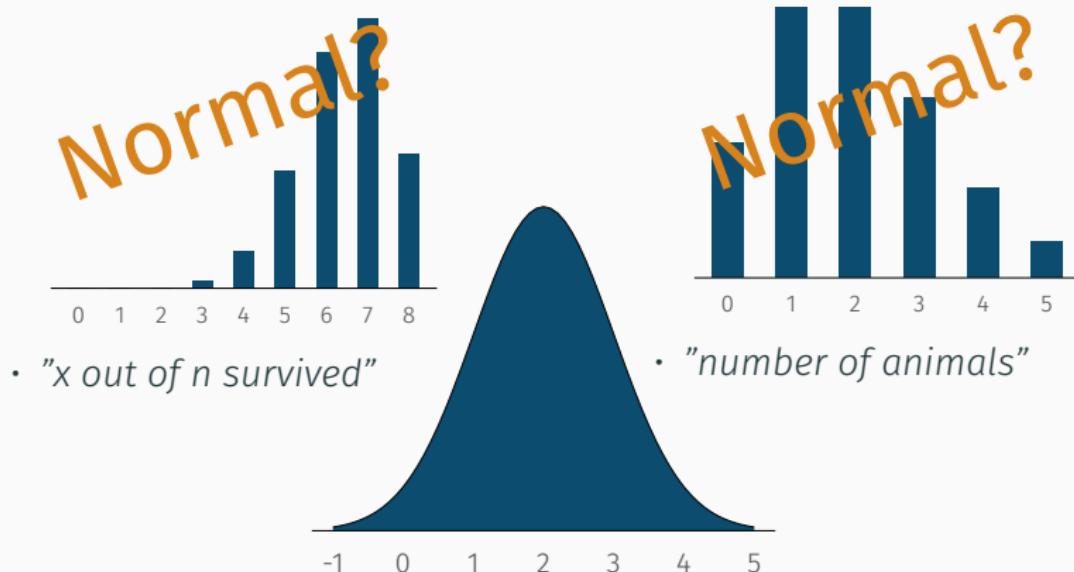


- " x out of n survived"

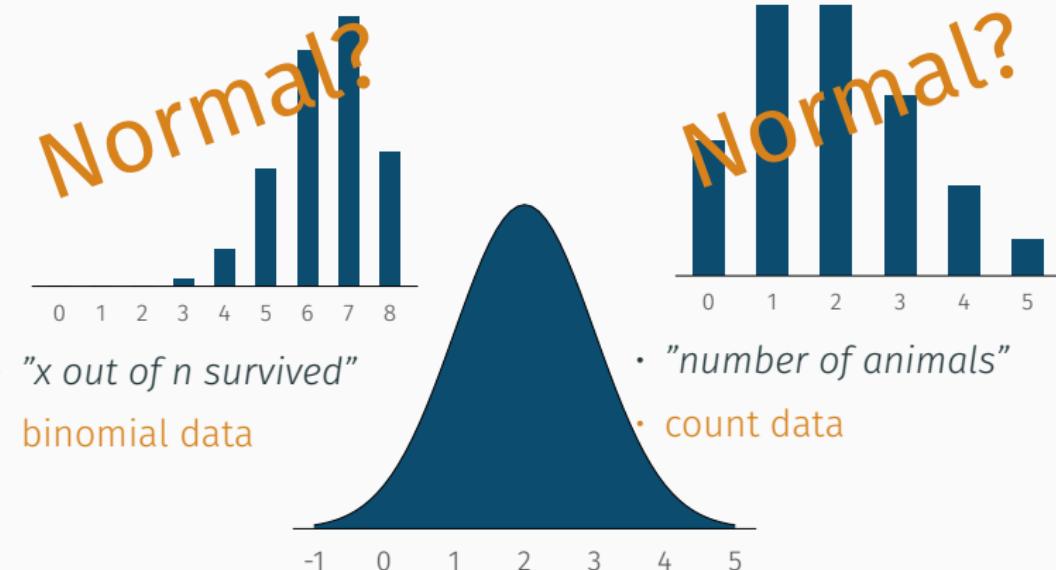


- "number of animals"

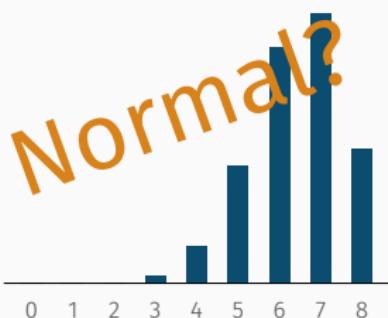
Ecotoxicology is not normal



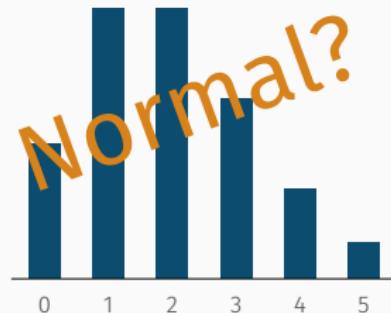
Ecotoxicology is not normal



Ecotoxicology is not normal



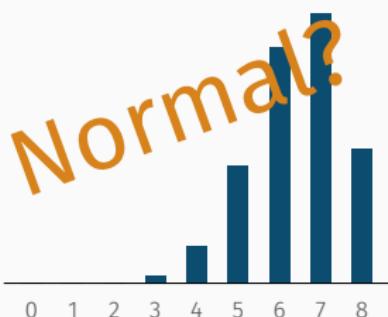
- " x out of n survived"
- binomial data



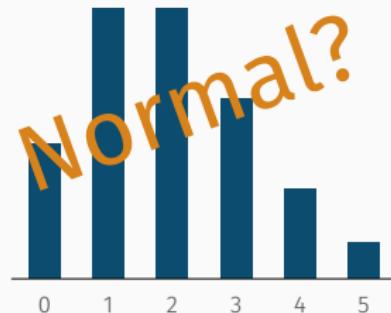
- "number of animals"
- count data

- ignore?

Ecotoxicology is not normal



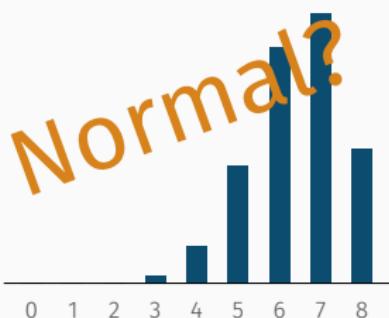
- " x out of n survived"
- binomial data



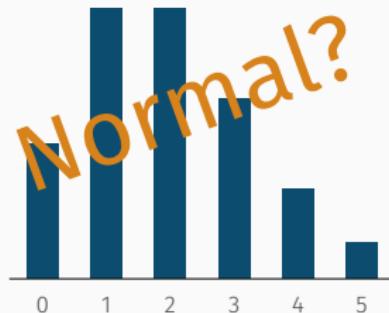
- "number of animals"
- count data

- ignore?
- transform?

Ecotoxicology is not normal



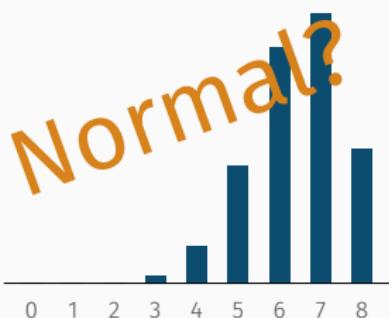
- " x out of n survived"
- binomial data



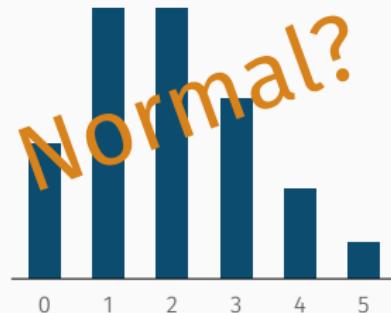
- "number of animals"
- count data

- ignore?
- transform?
- non-parametric?

Ecotoxicology is not normal



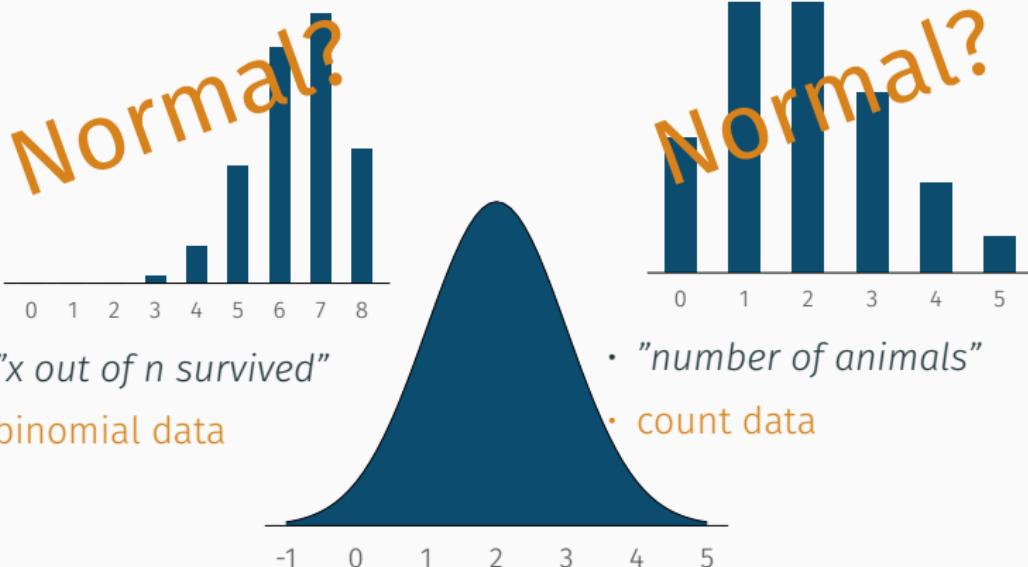
- " x out of n survived"
- binomial data



- "number of animals"
- count data

- ignore?
- transform?
- non-parametric?
- Generalized Linear Model (GLM)

Ecotoxicology is not normal



- "x out of n survived"
- binomial data
- "number of animals"
- count data
- ignore?
- transform?
- non-parametric?
- Generalized Linear Model (GLM)

A recent history of GLM (uncomprehensive) in ecology

J. R. Statist. Soc. A,
(1972), **135**, Part 3, p. 370

370

Generalized Linear Models

By J. A. NELDER and R. W. M. WEDDERBURN

Rothamsted Experimental Station, Harpenden, Herts



1972

Methods in Ecology and Evolution



Methods in Ecology & Evolution

doi: 10.1111/j.2041-210X.2010.00021.x

Do not log-transform count data

Robert B. O'Hara^{1*} and D. Johan Kotze²

¹Biodiversity and Climate Research Centre, Senckenberganlage 25, D-60325 Frankfurt am Main, Germany and

²Department of Environmental Sciences, PO Box 65, University of Helsinki, Helsinki FI-00014, Finland



A recent history of GLM (uncomprehensive) in ecology

Ecology, 92(1), 2011, pp. 3–10
© 2011 by the Ecological Society of America

The arcsine is asinine: the analysis of proportions in ecology

DAVID I. WARTON^{1,2,3} AND FRANCIS K. C. HUI¹

¹School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW 2052 Australia
²Evolution and Ecology Research Centre, The University of New South Wales, Sydney, NSW 2052 Australia



Methods in Ecology and Evolution



Methods in Ecology and Evolution

doi: 10.1111/j.2041-210X.2011.00127.x

Distance-based multivariate analyses confound location and dispersion effects

David I. Warton^{1*}, Stephen T. Wright¹ and Yi Wang^{1,2}

¹School of Mathematics and Statistics and Evolution & Ecology Research Centre; and ²School of Computer Science and Engineering, The University of New South Wales, NSW 2052, Australia



A recent history of GLM (uncomprehensive) in ecology

Ecotoxicology

DOI 10.1007/s10646-015-1421-0

Analysing chemical-induced changes in macroinvertebrate communities in aquatic mesocosm experiments: a comparison of methods

Eduard Szöcs · Paul J. Van den Brink · Laurent Lagadic · Thierry Caquet ·
Marc Roucaute · Arnaud Auber · Yannick Bayona · Matthias Liess ·
Peter Ebke · Alessio Ippolito · Cajo J. F. ter Braak · Theo C. M. Brock ·
Ralf B. Schäfer



A recent history of GLM (uncomprehensive) in ecology

Environ Sci Pollut Res
DOI 10.1007/s11356-015-4579-3

RESEARCH ARTICLE

Ecotoxicology is not normal

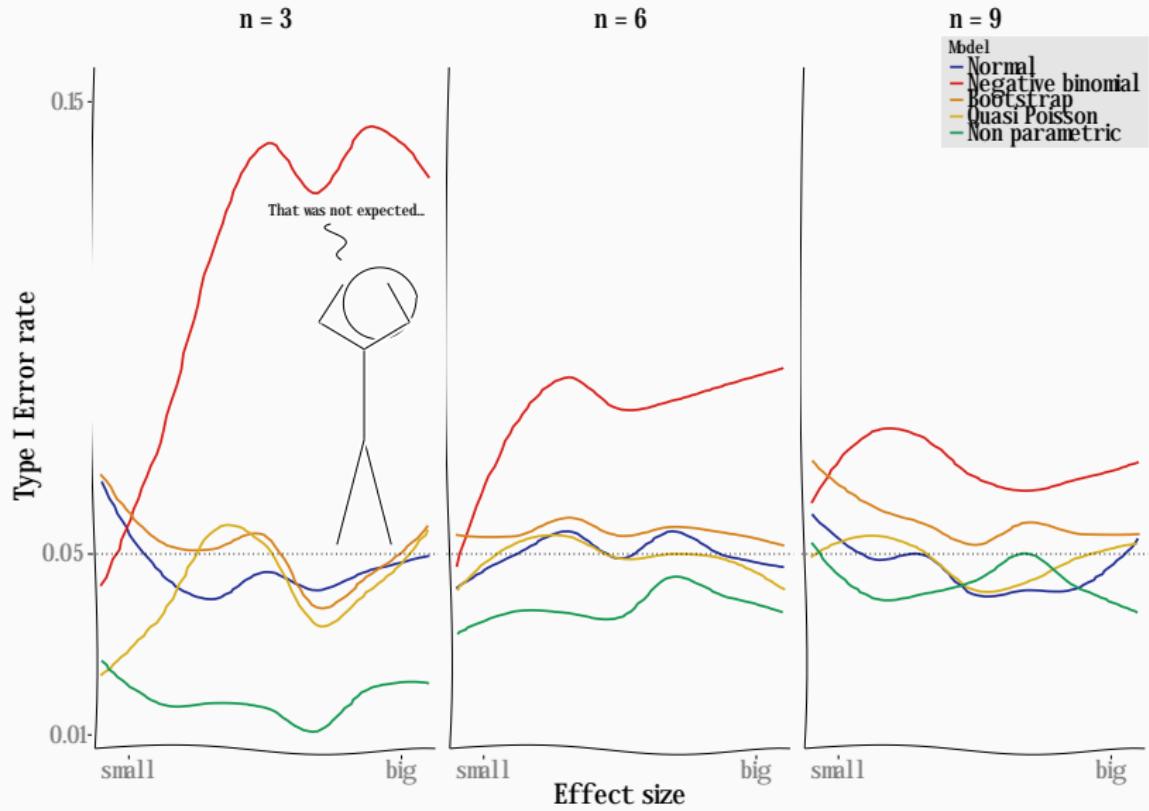
A comparison of statistical approaches for analysis of count
and proportion data in ecotoxicology

Eduard Szöcs¹ · Ralf B. Schäfer¹

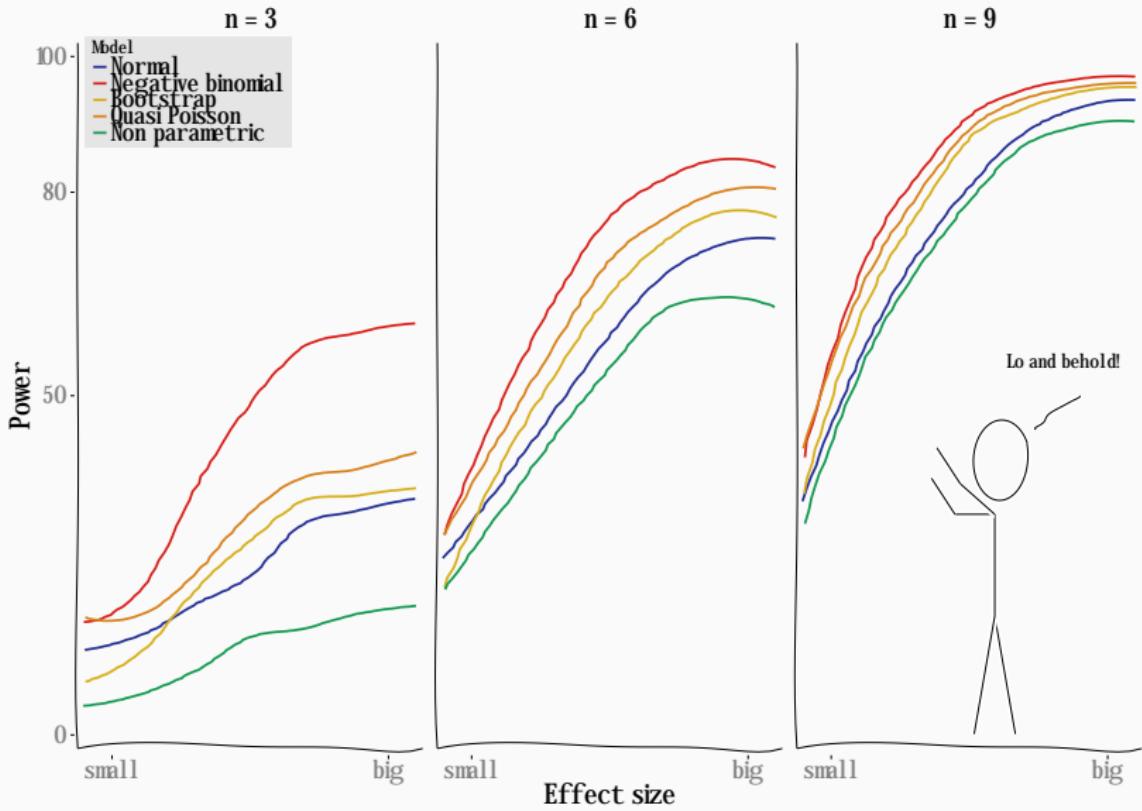


A simulation study

GLMs do not always work



Power is low, GLM can do better



What we learned from this study

1. Negative-binomial GLM show increased Type I errors
2. Can be fixed via bootstrap
3. Power in ecotoxicological experiments generally low
4. NOECs are not reliable
5. GLMs can increase this power

Where are we today?

Three days earlier...

Methods in Ecology and Evolution



Methods in Ecology and Evolution 2015, **6**, 828–835

doi: 10.1111/2041-210X.12386

For testing the significance of regression coefficients, go ahead and log-transform count data

Anthony R. Ives*

Department of Zoology, University of Wisconsin-Madison, Madison, WI 53706, USA



Methods in Ecology and Evolution



Methods in Ecology and Evolution 2016, 7, 882–890

doi: 10.1111/2041-210X.12552

FORUM

Three points to consider when choosing a LM or GLM test for count data

David I. Warton^{1*}, Mitchell Lyons², Jakub Stoklosa¹ and Anthony R. Ives³

¹School of Mathematics and Statistics and Evolution & Ecology Research Centre, University of New South Wales, NSW 2052, Australia; ²School of Biological, Earth and Environmental Sciences, University of New South Wales, NSW 2052, Australia; and

³Department of Zoology, University of Wisconsin-Madison, Madison, WI 53706, USA



Where are we today?

Three points to consider ...

1. Choose your model based on data properties
2. Fix Type I errors by resampling
3. Models that better fit the data have better power properties

Exploring Monitoring Data for ERA

Environmental Monitoring

Overview on data compiled

Thresholds

Statistics with chemical measurements

Dynamics

Check: What did go wrong with Neonics during ERA? Exposure wrong? Effect wrong (e.g. did not consider sensitive species)? => Quite sure that Effect assessment missed it... => show SSD, highlighting standard test species (from EPA) * old guidance document did not enforce insect. * new one (from 2013) enforces additional insect data

RAC is landscape dependend (=> exposure?)

Example for power (experimental setup)

What we learned

Solutions for Linking Data in ERA

Biologists & Chemists face the same problems

Names

Osmia rufa, *Osmia bicornis*, *Osmia ruffa*, *Osmia unilandauis*, *Osmia spec.*

Chlorpyrifos, Chlorpyriphos,
Chlorphyrifos, Chlorpyrifos-ethyl,
Chlorpypifot

Biologists & Chemists face the same problems

Names

Osmia rufa, *Osmia bicornis*, *Osmia ruffa*, *Osmia unilandauis*, *Osmia spec.*

Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot

Hierarchies

Hymenoptera/ Apoidea/
Megachilidae/ *Osmia/ rufa*

organophosphate, ester, insecticide

Biologists & Chemists face the same problems

Names

<i>Osmia rufa</i> , <i>Osmia bicornis</i> , <i>Osmia ruffa</i> , <i>Osmia unilandauis</i> , <i>Osmia spec.</i>	Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot
----------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ <i>Osmia/ rufa</i>	organophosphate, ester, insecticide
-----------------------------------------------------------	-------------------------------------

Traits / Properties

Wing length, Mass, Season	Mass, <i>Kow</i> , <i>LC₅₀</i>
---------------------------	-------------------------------------------

Biologists & Chemists face the same problems

Names

<i>Osmia rufa</i> , <i>Osmia bicornis</i> , <i>Osmia ruffa</i> , <i>Osmia unilandauis</i> , <i>Osmia spec.</i>	Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot
----------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ <i>Osmia/ rufa</i>	organophosphate, ester, insecticide
-----------------------------------------------------------	-------------------------------------

Traits / Properties

Wing length, Mass, Season	Mass, <i>Kow</i> , <i>LC₅₀</i>
---------------------------	-------------------------------------------

Identifiers

NCBI, ITIS, EOL, ...	2921-88-2, Clc1c(OP(=S)[...], InChI=1S/C9H11C[...], SBPBAQFW[...], CSID,...
----------------------	-----------------------------------------------------------------------------

Biologists & Chemists face the same problems

Names

<i>Osmia rufa</i> , <i>Osmia bicornis</i> , <i>Osmia ruffa</i> , <i>Osmia unilandauis</i> , <i>Osmia spec.</i>	Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot
----------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ <i>Osmia/ rufa</i>	organophosphate, ester, insecticide
-----------------------------------------------------------	-------------------------------------

Traits / Properties

Wing length, Mass, Season	Mass, <i>K_{ow}</i> , <i>LC₅₀</i>
---------------------------	------------------------------------------------------

Identifiers

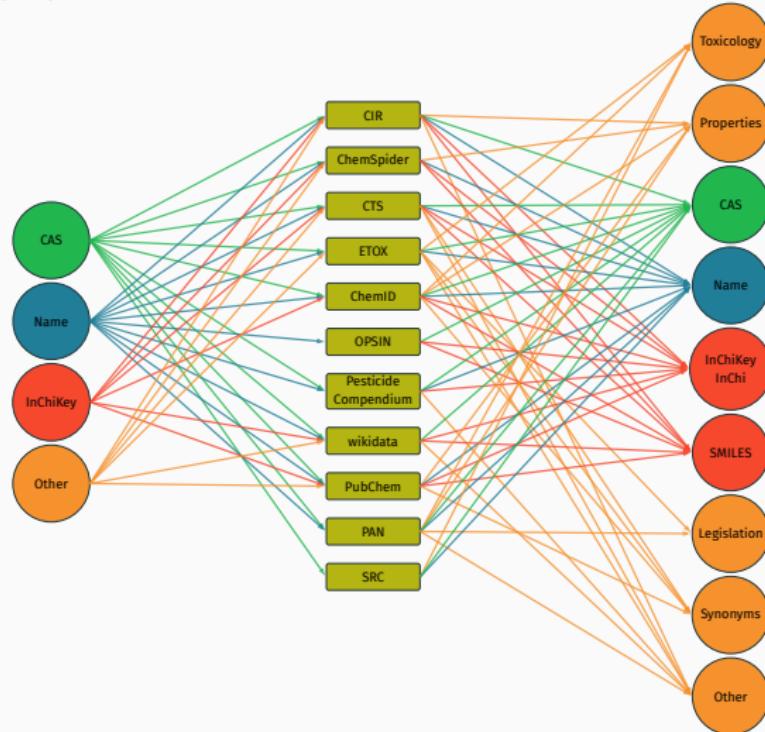
NCBI, ITIS, EOL, ...	2921-88-2, Clc1c(OP(=S)[...], InChI=1S/C9H11C[...], SBPBAQFW[...], CSID,...
----------------------	-----------------------------------------------------------------------------

Amount of data

2993 taxa	489 pesticides (+ 590 other organics)
-----------	------------------------------------------

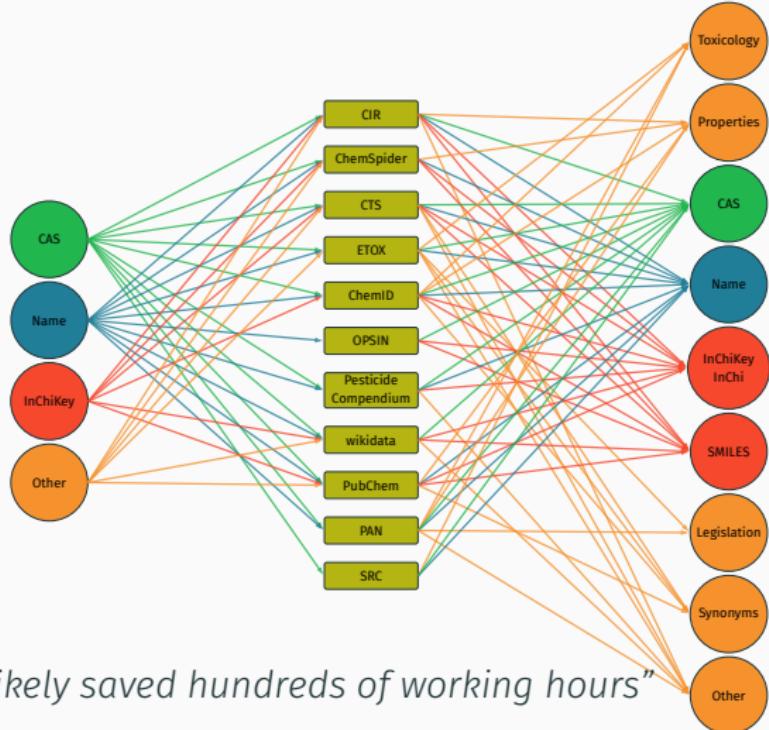
Instead of wasting time...

... use webchem!



Instead of wasting time...

... use webchem!



"webchem ...likely saved hundreds of working hours"

Instead of wasting time...

... use taxize!

A grid of logos for various biodiversity databases and projects:

- ITIS
- iPlant Collaborative
- Plantminer
- Catalogue of Life
- Tropicos
- GBIF
- NCBI
- eOL Encyclopedia of Life
- gnl
- RED LIST
- ubio
- Canadensys
- ThePlantList

Instead of wasting time...

... use taxize!

The image displays a collection of logos for various biological databases and projects, arranged in a grid. The logos include:

- ITIS**: Global Invasive Species Database
- iPlant Collaborative**
- Catalogue of Life**
- Plantminer**
- Tropicos**
- RED LIST**
- uBio**
- ThePlantList**
- Canadensys**
- eol** Encyclopedia of Life
- gni**
- NCBI**
- GBIF**

"Days of searching done during my morning coffee. Amazing. **taxize**."

Recap

What we learned

✓ Improving Statistics in ERA

- Change your model, not your data
- Ultimately ban NOEC
- Take LOQ into account

What we learned

✓ Improving Statistics in ERA

- Change your model, not your data
- Ultimately ban NOEC
- Take LOQ into account

✓ Exploring Monitoring Data for ERA

- Risk drivers and dynamics
- Agricultural Small streams at risk & neglected
- Neonicotinoids
- Feedback for ERA

What we learned

✓ Improving Statistics in ERA

- Change your model, not your data
- Ultimately ban NOEC
- Take LOQ into account

✓ Exploring Monitoring Data for ERA

- Risk drivers and dynamics
- Agricultural Small streams at risk & neglected
- Neonicotinoids
- Feedback for ERA

✓ Solutions for Linking Data in ERA

- Handling big eco(toxico-)logical data not easy
- Now easier

Statistical Ecotoxicology

Improving the Utilisation of Data for
Environmental Risk Assessment

Eduard Szöcs

💻 <http://edild.github.io/>

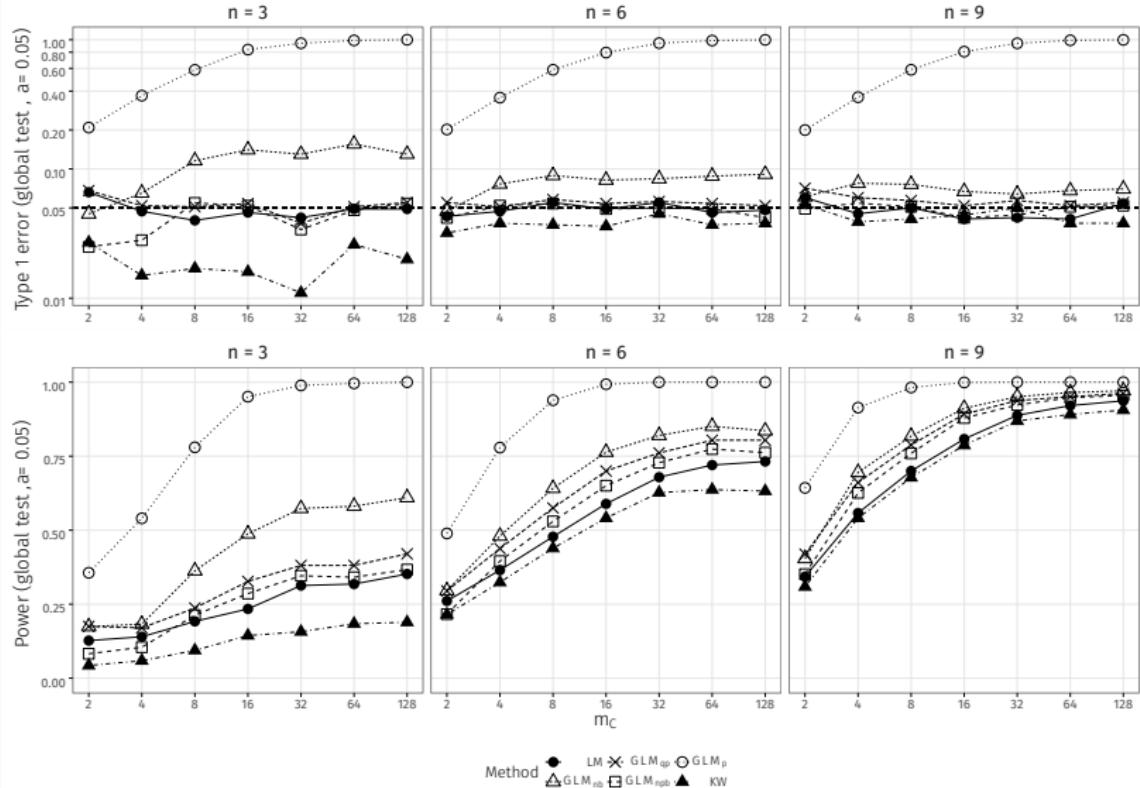
🐦 [@EduardSzoebs](https://twitter.com/EduardSzoebs)

📄 https://github.com/edild/phd_defense

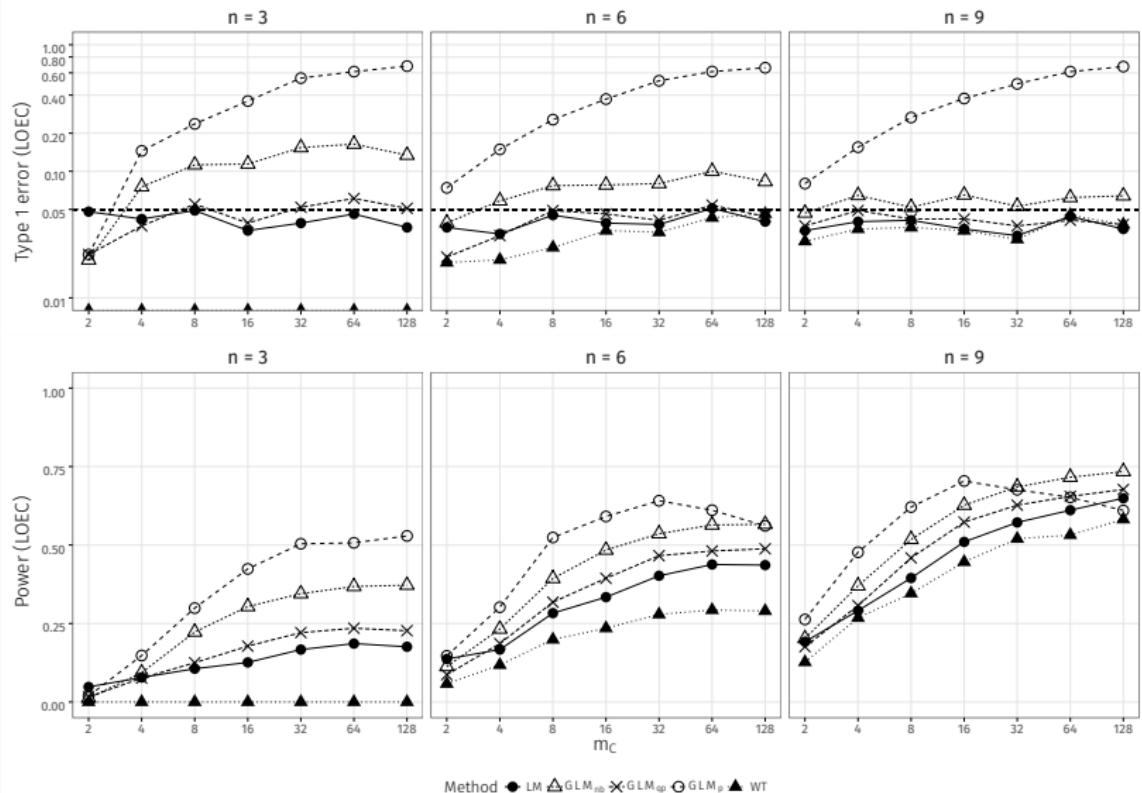
📄 https://github.com/edild/phd_thesis



Power en detail



For LOEC it is even worse



Comparison with Ives...

Szöcs (2015)

- factorial design
- one predictor
- low replicated
- LM, GLM, bootstrap
- High T1 error of NB
- Quasi-Poisson worked well
- Bootstrap fixes the problems

Ives (2015)

- continuous design
- two predictors
- well replicated
- LM, GLM
- High T1 error of NB
- Quasi-Poisson has problems with multiple predictors
-

ZAGA what...?

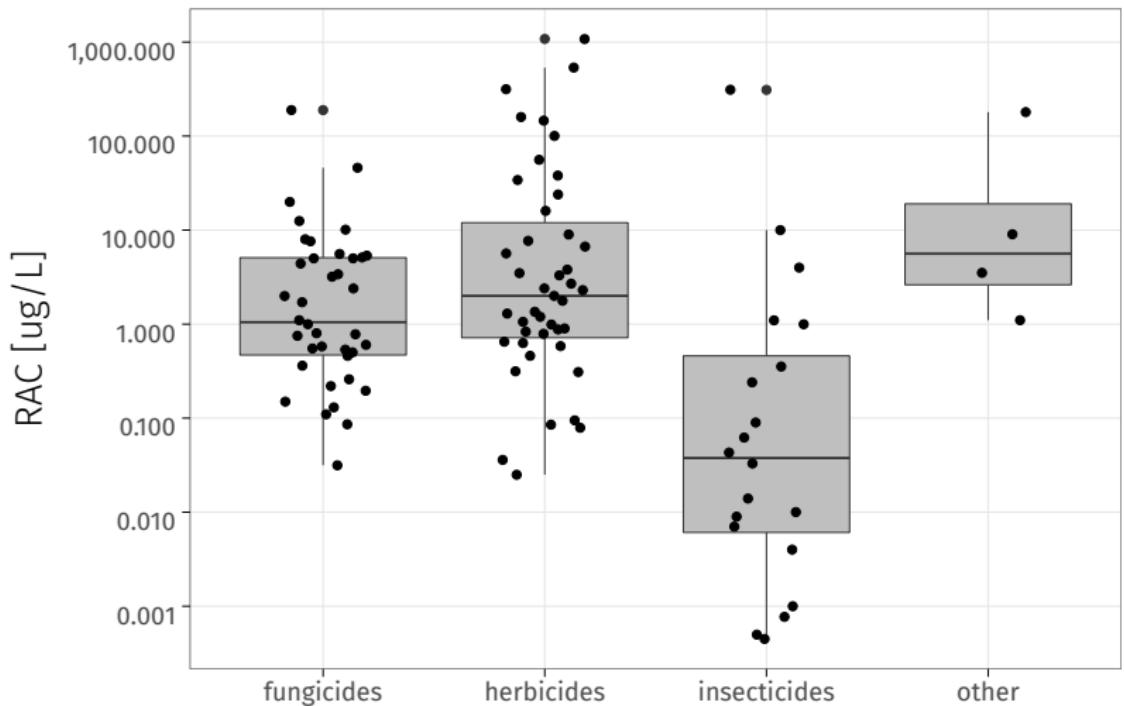
shiny app: <http://uni-ko-ld.de/g4>

Comparison with recent studies

<u>Szöcs (2016)</u>	<u>Stehle (xxx)</u>	<u>Knauer (xxx)</u>
Germany	Europe / Global	Switzerland
Monitoring	Scientific Publications	Monitoring
Grab sampling	Grab & Event driven sampling	Grab sampling
% LOQ:	% LOQ:	% LOQ:

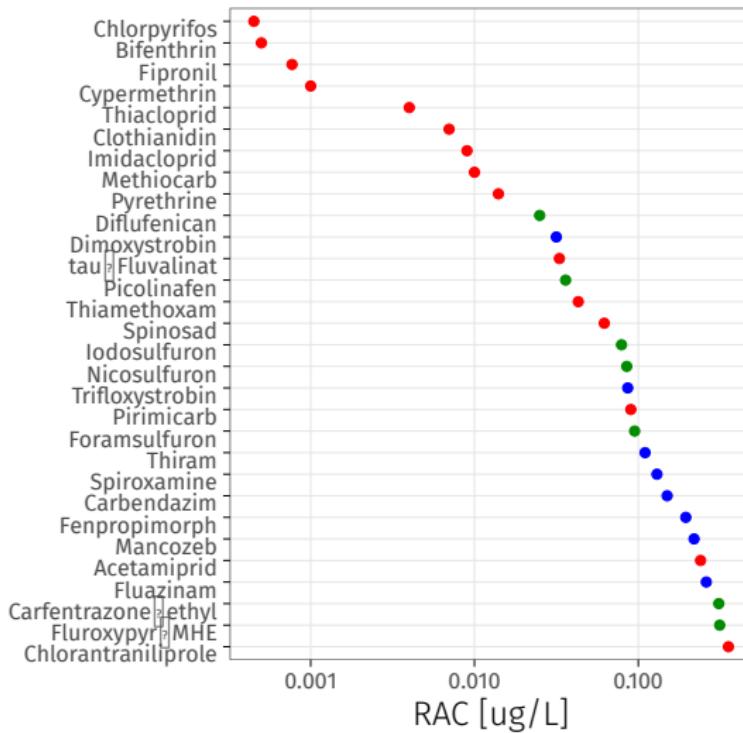
RACs by Type

105 RACs provided by UBA splitted by group



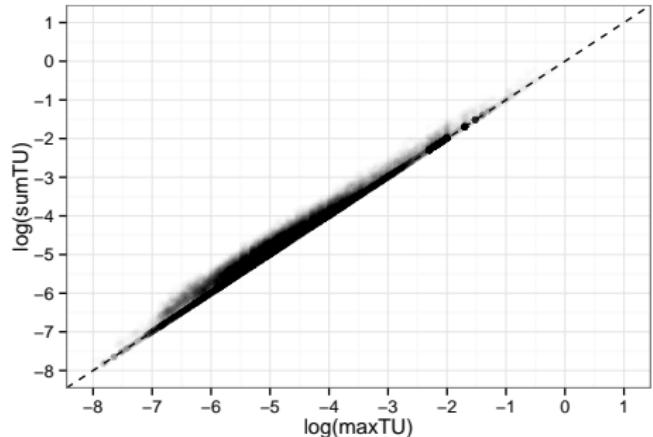
RACs by Compound

30 lowest RACs



Type • fungicides ● herbicides • insecticides

Mixtures are common, but one compound dominates the risk

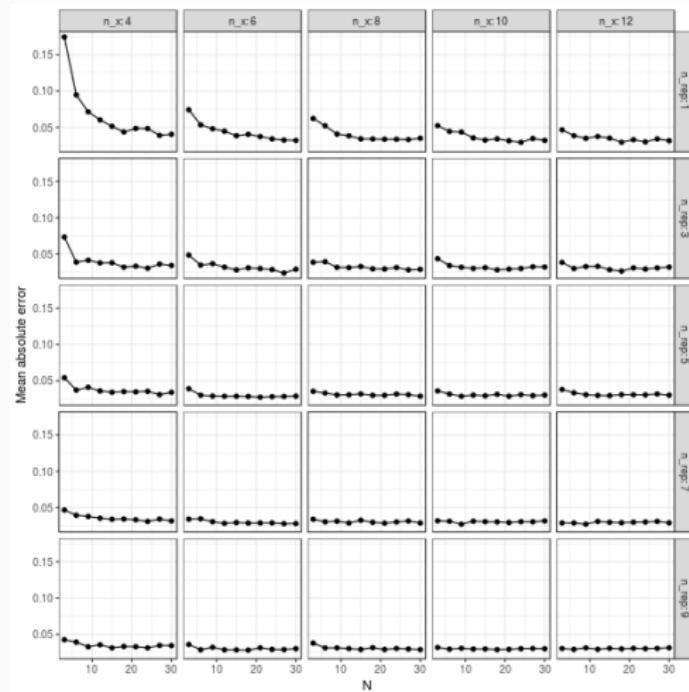


- up to 50 compounds in one sample
- high correlation
- ~ 0.5 TU increase
- mainly one compound responsible for risk

Simulations are worth their work, use them *a priori*!

Experimental design for dose-response experiments - a simulation

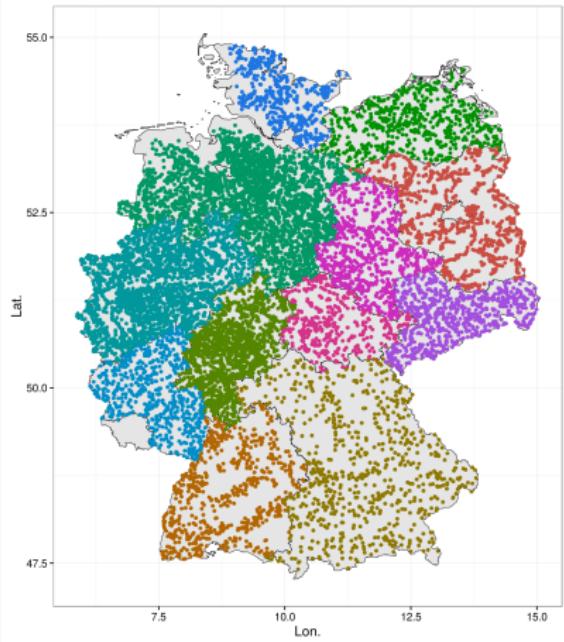
webchem http://edild.github.io/lc50_bias_sim/



Reasons for observed RAC exceedances

see notes on lecture on 16.12.2016

Biotic field effects



- biological data with good spatial coverage
- 60% of spatial congruence
- Large scale effects largely unknown.
- Some work left...
- Future....

Software

webchem <https://github.com/ropensci/webchem>

taxize <https://github.com/ropensci/webchem>

Stable versions also on CRAN...