Statistical Ecotoxicology - Improving the utilization of data for ecological risk assessment

Eduard Szöcs

Institute for Environmental Sciences, University of Koblenz-Landau

Landau, 22.09.2016

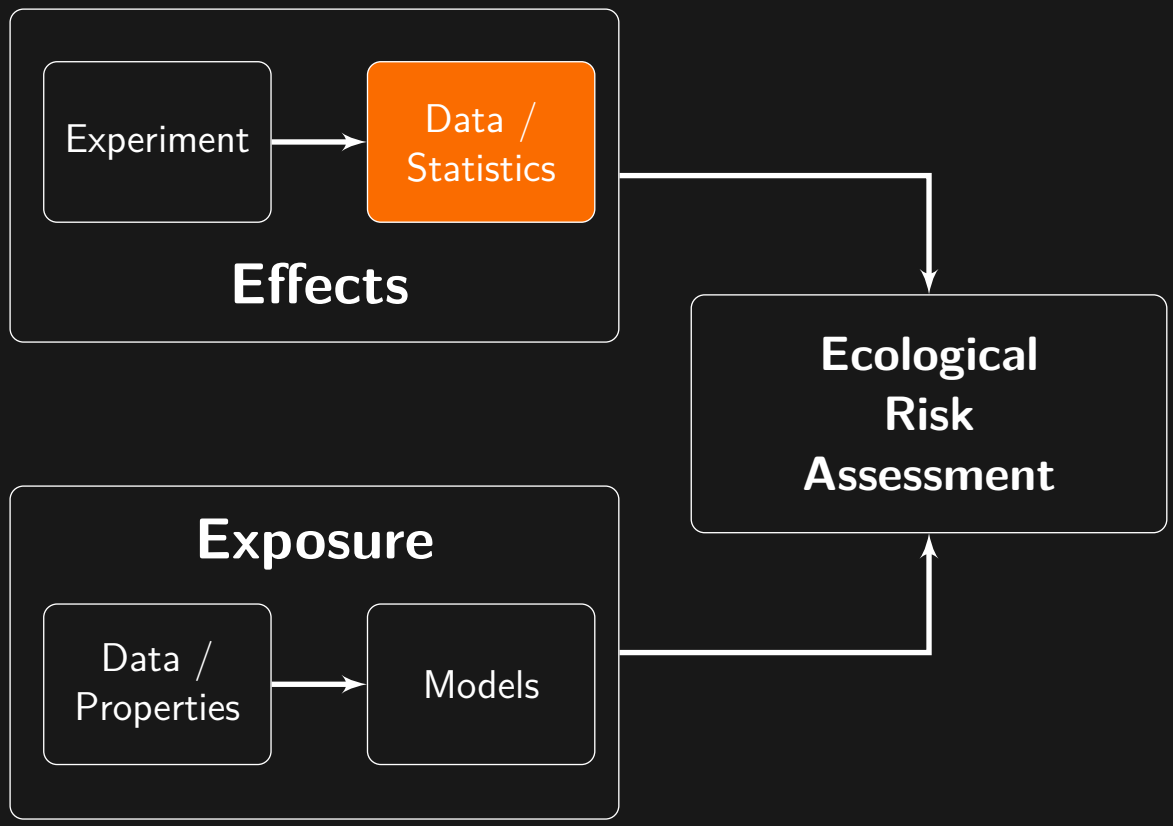
My field of research is somewhere between...

... Eco(-toxico)logy, Data Analysis & Programming

Statistical Ecotoxicology

Current use in ecotoxicology

- Ecological risk assessment (ERA) relies on statistics

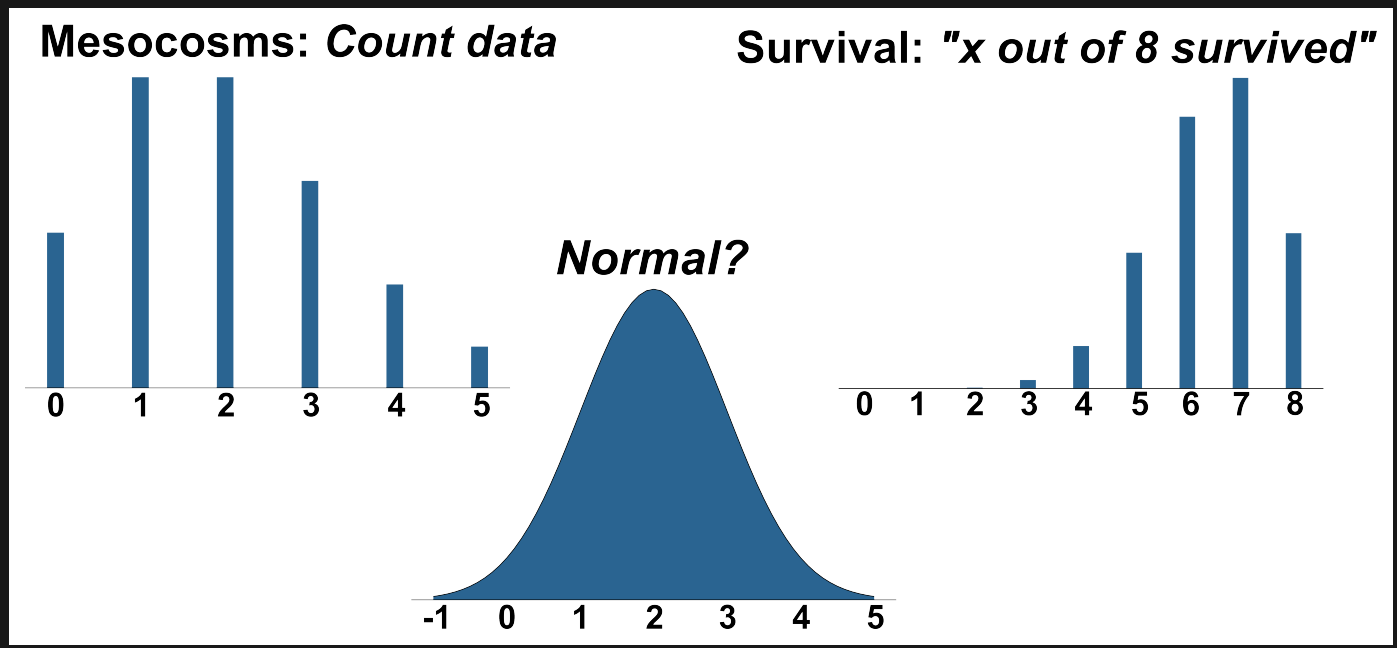


Current use in ecotoxicology

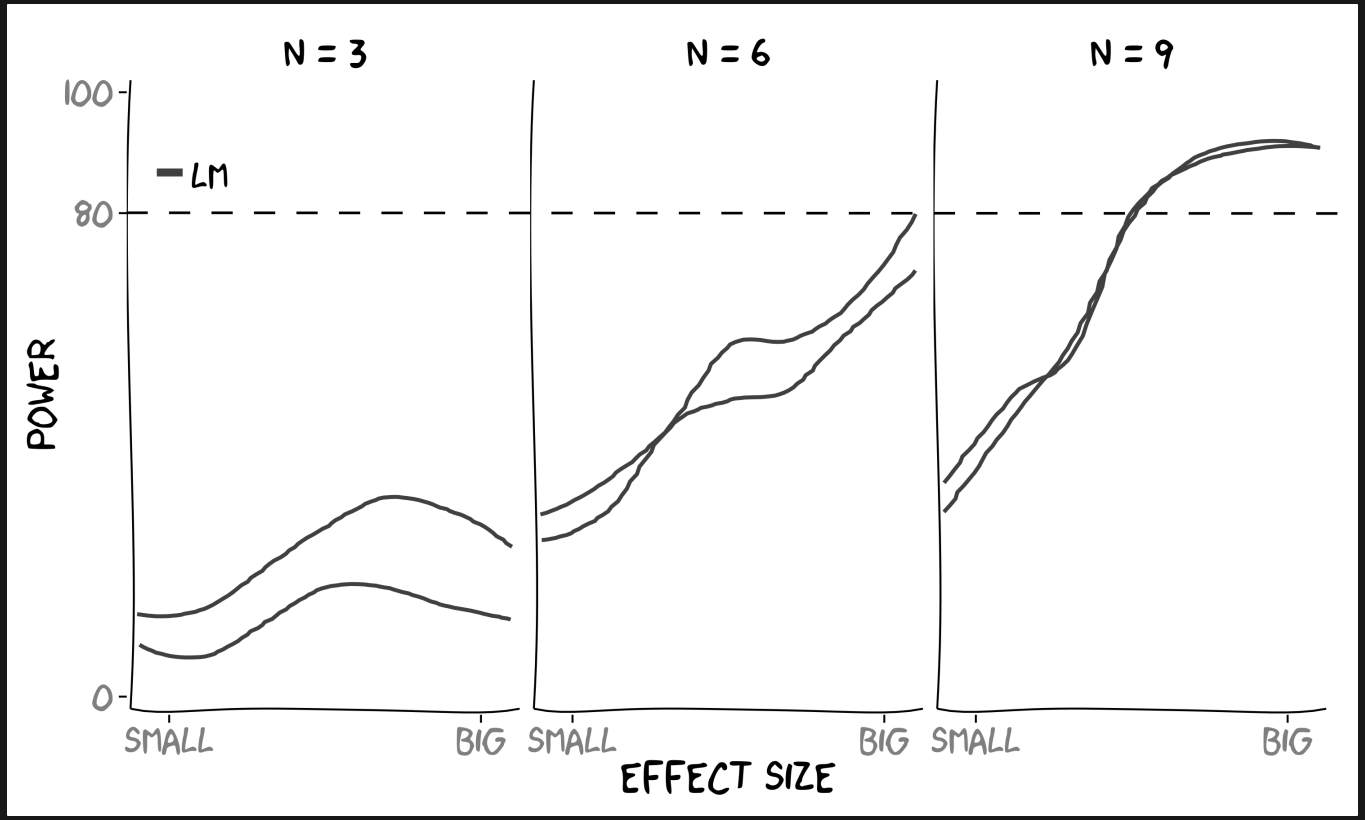
- Ecological risk assessment (ERA) relies on statistics
- Experiments with low replication

Current use in ecotoxicology

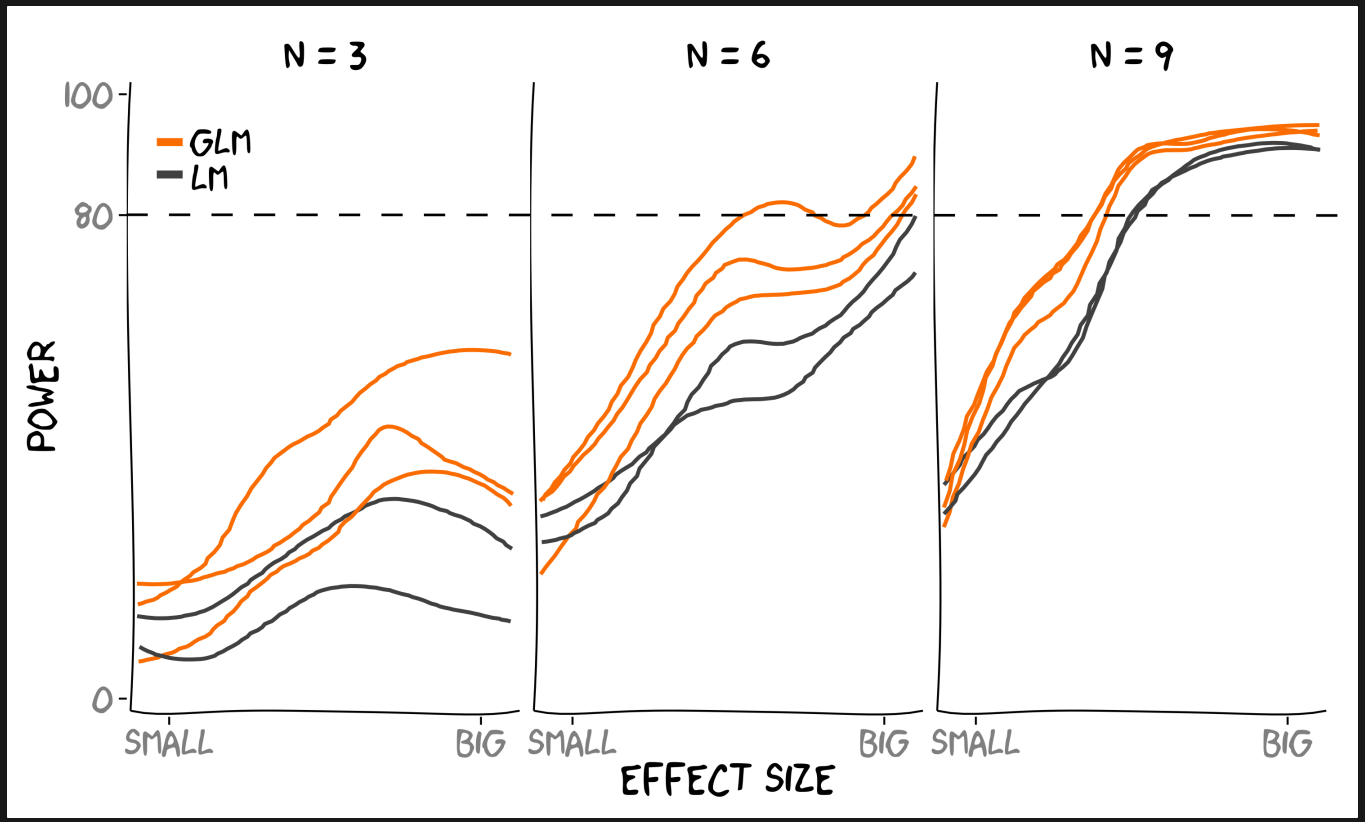
- Ecological risk assessment (ERA) relies on statistics
- Experiments with low replication
- Usually analysed using Linear Models of transformed data
- Null Hypothesis Significance Testing (=> NOEC)



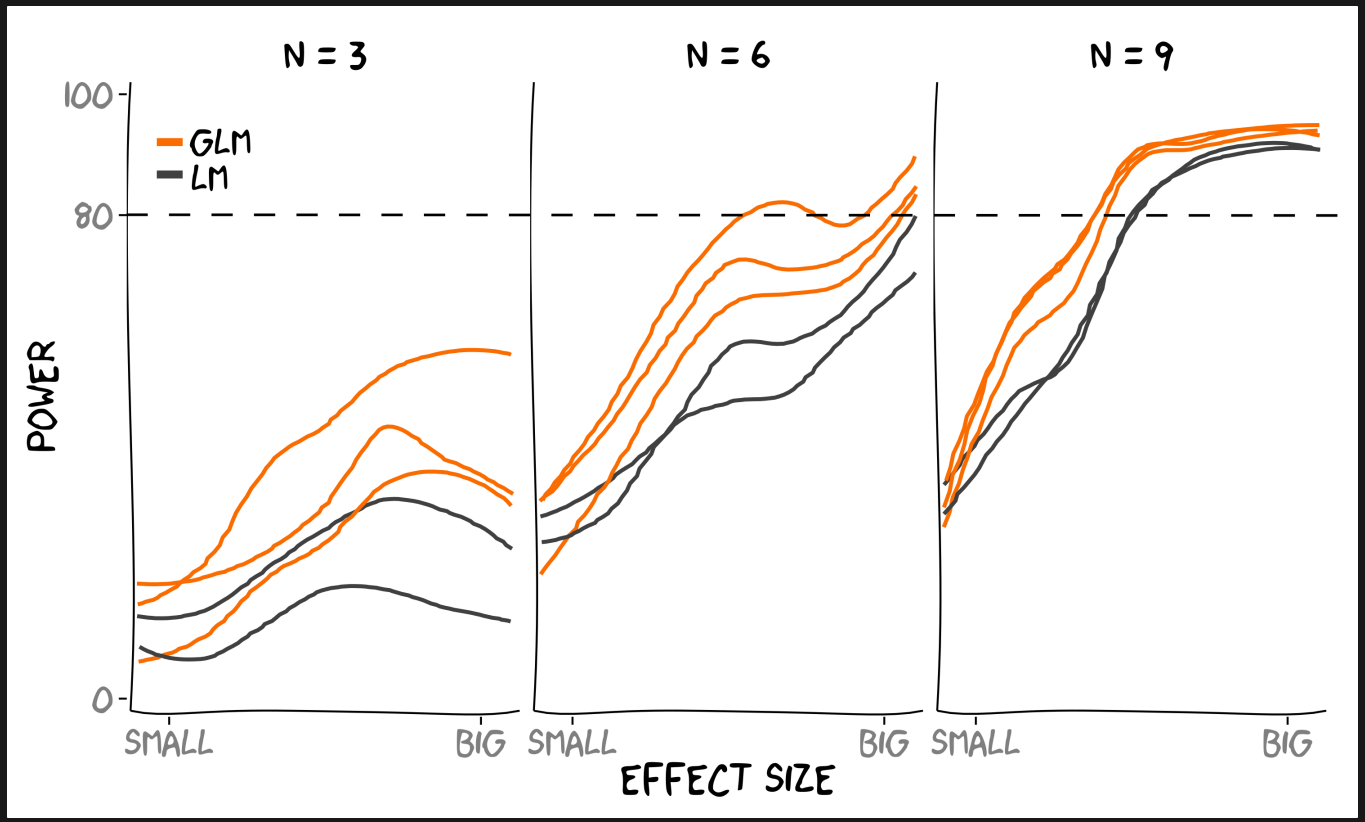
Statistical Power in current experimental designs in ecotoxicology is unacceptably low



Generalized Linear Models can do better



Generalized Linear Models can do better



Better abandon NOEC and use a regression design [1]...

A priori power analysis for better design.

[1] debated since 30 years.

Monitoring Data

Monitoring data...

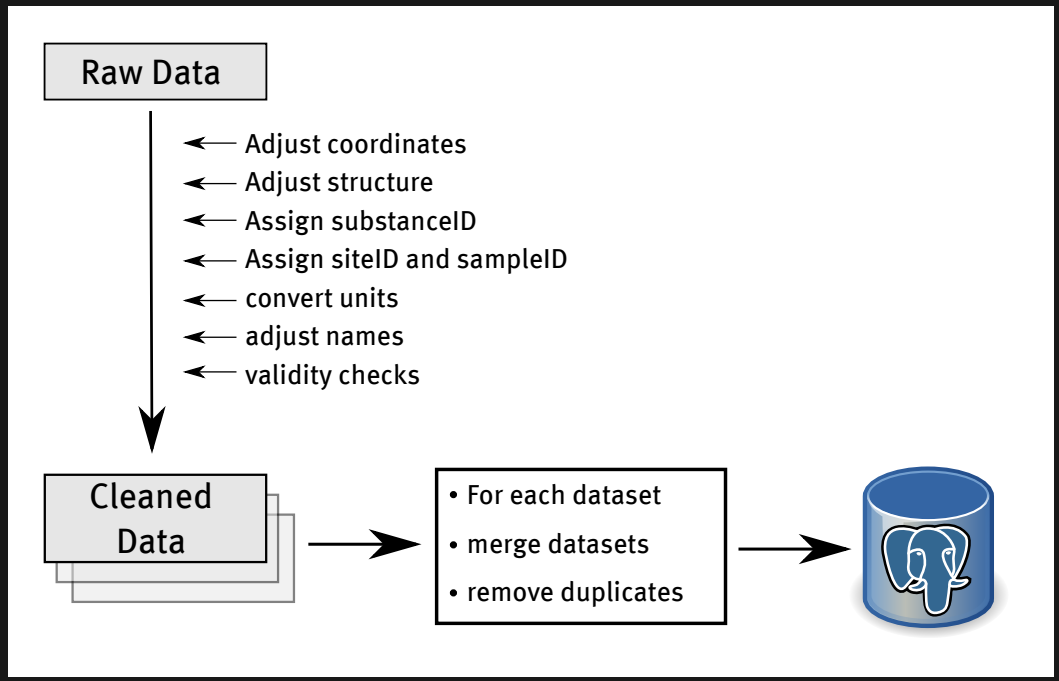
- ... provides the biggest amount of data available on pesticides in the environment

Monitoring data...

- ... provides the biggest amount of data available on pesticides in the environment
- ... provides an opportunity to study large-scale dynamics of pesticides

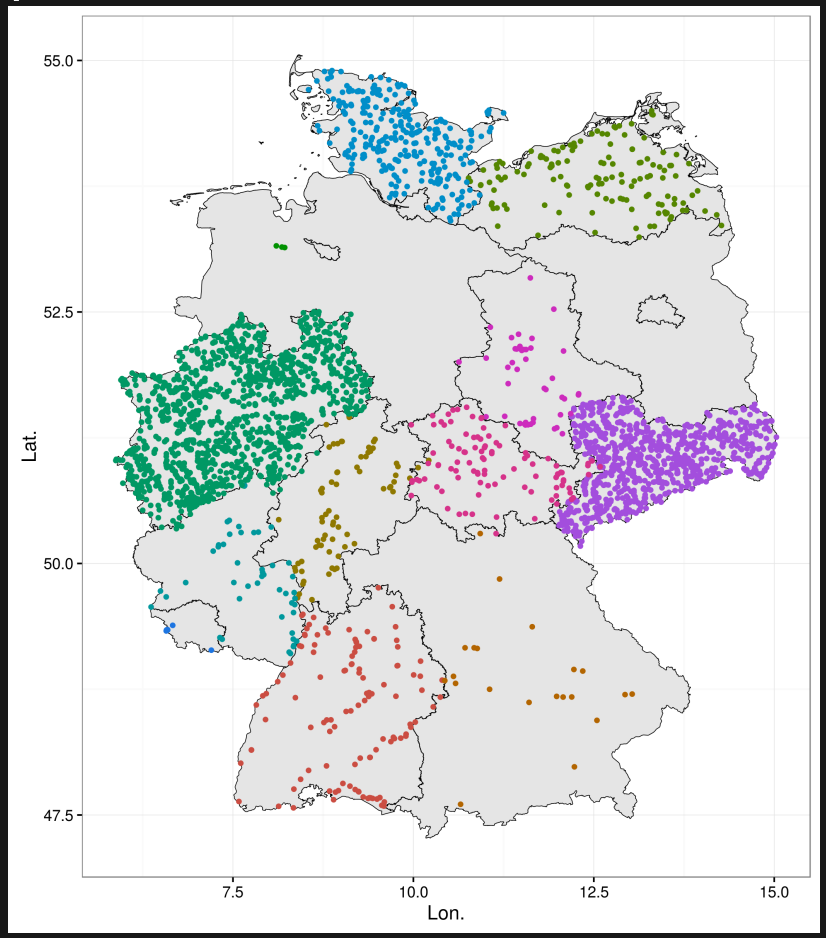
Monitoring data...

- ... provides the biggest amount of data available on pesticides in the environment
- ... provides an opportunity to study large-scale dynamics of pesticides
- ... is really messy



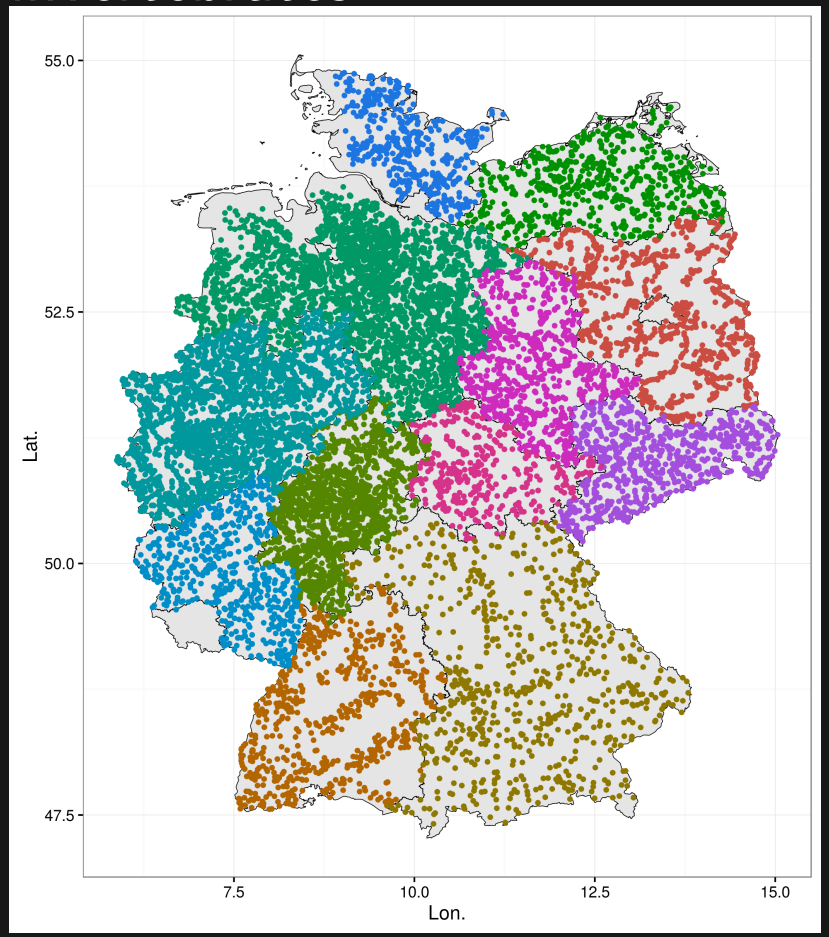
The biggest currently available dataset on

pesticides



3,000 sites, 45,000 samples, 500 pesticides

invertebrates



14,000 sites, 27,000 samples, 3000 taxa

Additional data on

Sites

- catchment size
- agriculture within catchment

Additional data on

Sites

- catchment size
- agriculture within catchment

Samples

- daily precipitation

Additional data on

Sites

- catchment size
- agriculture within catchment

Samples

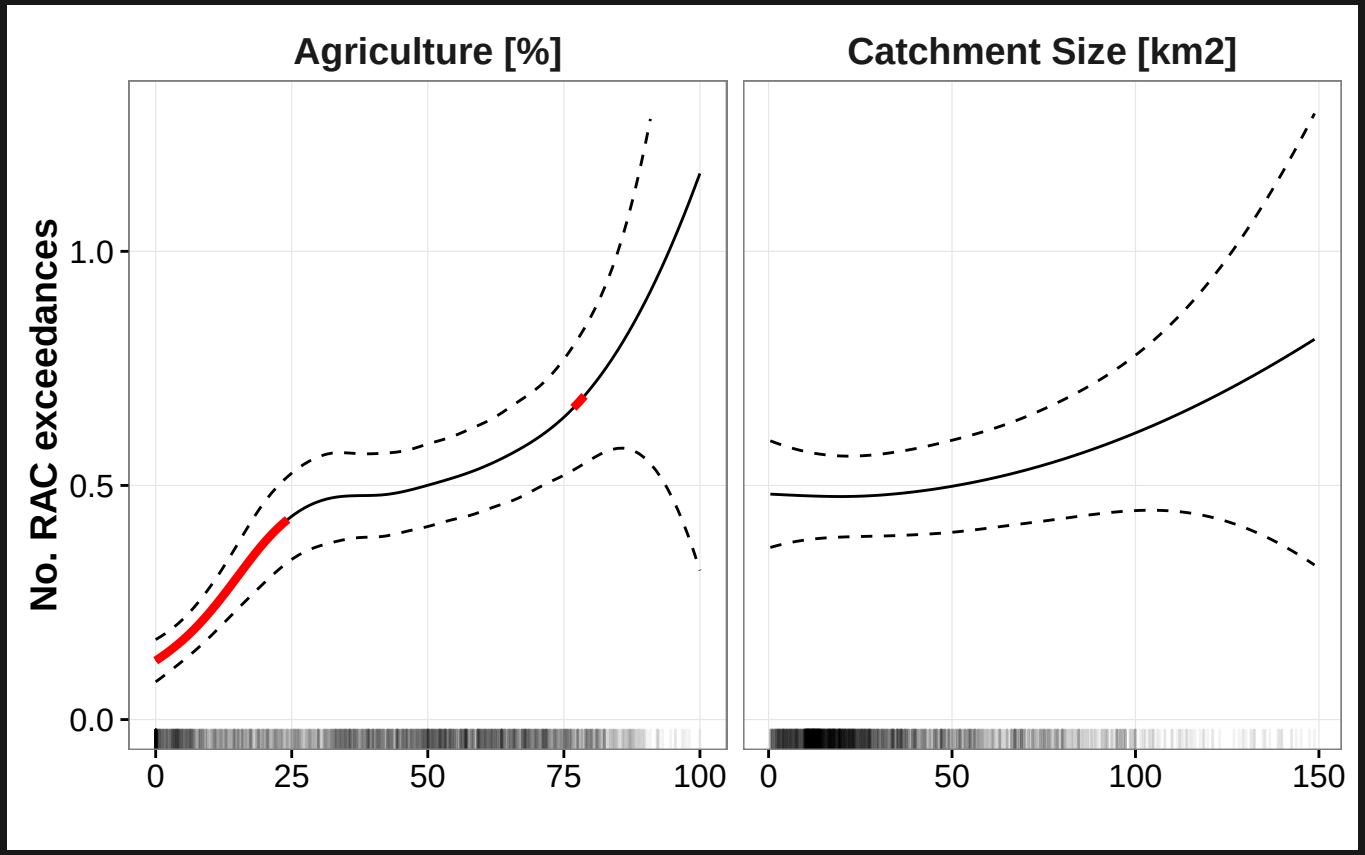
- daily precipitation

Compounds

- RAC, LC50, EQS
- chemical group
- identifiers
- properties

Results - Thresholds

Results - Thresholds



Results - Precipitation & Seasonality

Results - Precipitation & Seasonality

- Used a mixture model

$$RQ_i \sim ZAGA(\mu_i, \sigma, \pi_i) = \begin{cases} (1 - \pi_i) & \text{if } y < LOQ \\ \pi_i \times f_{Gamma}(\mu_i, \sigma) & \text{if } y \geq LOQ \end{cases}$$

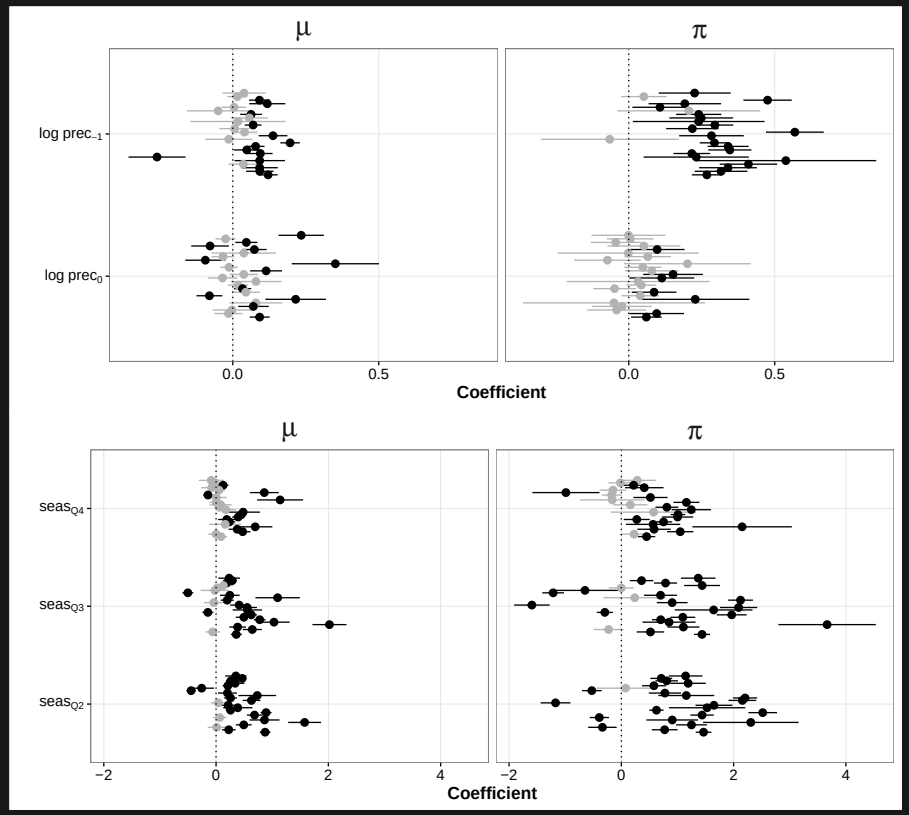
- Precipitation and Quarter as predictors
- Site within state as random intercept

Results - Precipitation & Seasonality

- Used a mixture model

$$RQ_i \sim ZAGA(\mu_i, \sigma, \pi_i) = \begin{cases} (1 - \pi_i) & \text{if } y < LOQ \\ \pi_i \times f_{Gamma}(\mu_i, \sigma) & \text{if } y \geq LOQ \end{cases}$$

- Precipitation and Quarter as predictors
- Site within state as random intercept
- Precipitation before sampling increases RQ

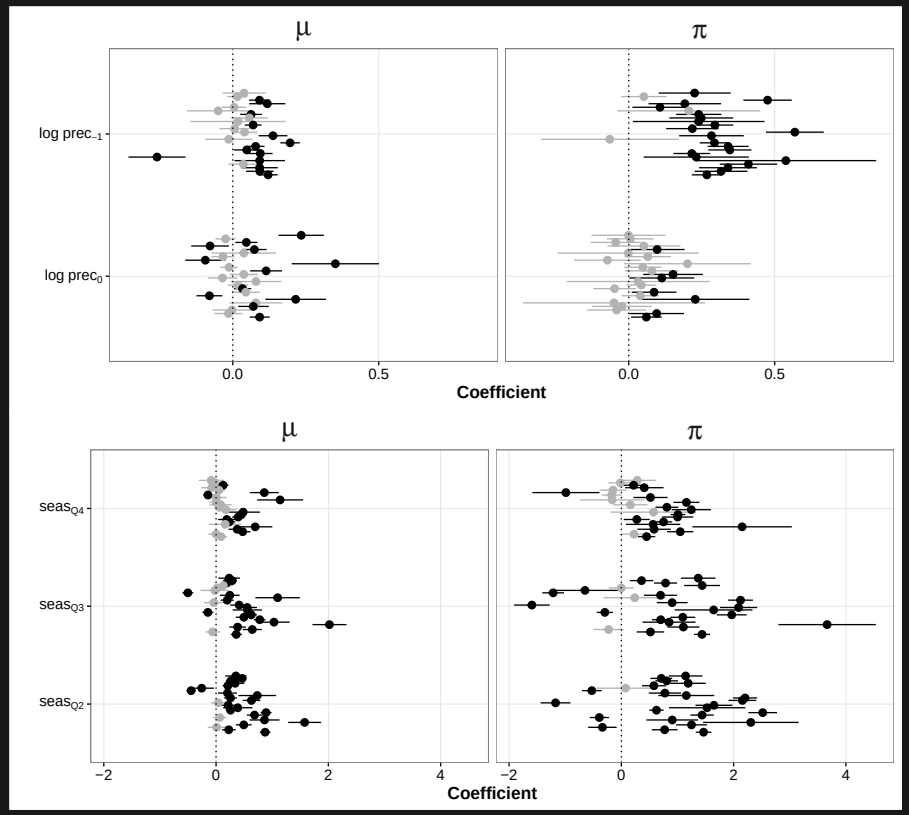


Results - Precipitation & Seasonality

- Used a mixture model

$$RQ_i \sim ZAGA(\mu_i, \sigma, \pi_i) = \begin{cases} (1 - \pi_i) & \text{if } y < LOQ \\ \pi_i \times f_{Gamma}(\mu_i, \sigma) & \text{if } y \geq LOQ \end{cases}$$

- Precipitation and Quarter as predictors
- Site within state as random intercept
- Precipitation before sampling increases RQ
- Summer higher RQ, but compound specific



Results - Small Water Bodies (SWB)

- most streams are *small*

Results - Small Water Bodies (SWB)

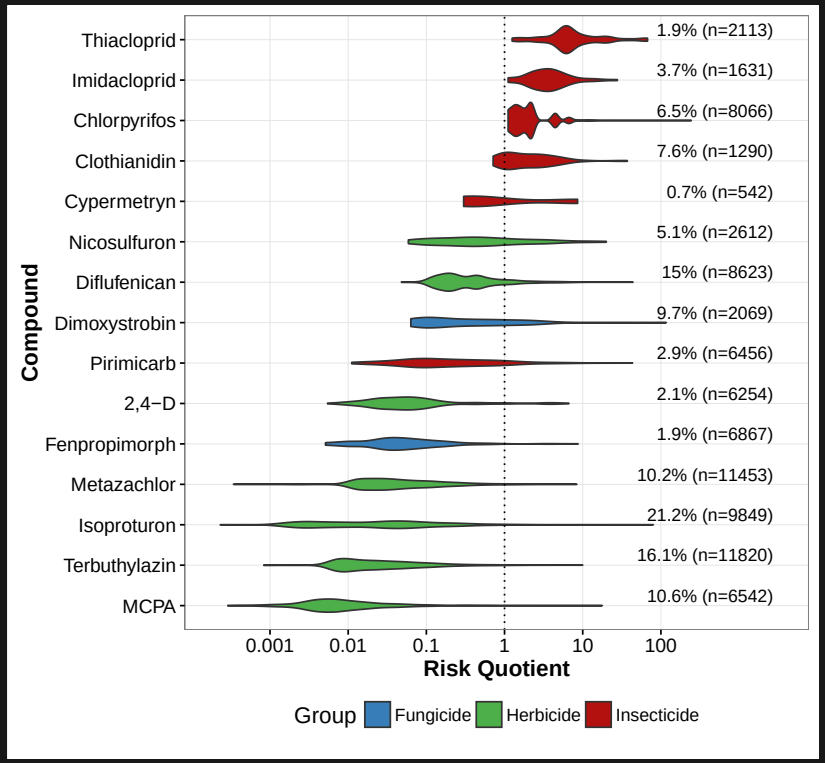
- most streams are *small*
- High risk of pollution
  - adjacency to fields
  - low dilution

Results - Small Water Bodies (SWB)

- most streams are *small*
- High risk of pollution
  - adjacency to fields
  - low dilution
- refuge of biodiversity

Results - Small Water Bodies (SWB)

- most streams are *small*
- High risk of pollution
  - adjacency to fields
  - low dilution
- refuge of biodiversity
- Neonicotinoids
- up to 244x RAC
- ecological effects likely



Software

Biologists and Chemists face similar problems...

Biologists and Chemists face similar problems...

Names

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

Chlorpyrifos, Chlorpyriphos,
Chlorphyrifos, Chlorpyrifos-ethyl,
Chlorpypifot

Biologists and Chemists face similar problems...

Names

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

Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ Osmia/ rufa

organophospate, ester, insecticide

Biologists and Chemists face similar problems...

Names

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

Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ Osmia/ rufa

organophospate, ester, insecticide

Attributes

Wing length, Mass, Season

Mass, K_{OW}, LC_{50}

Biologists and Chemists face similar problems...

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

Hymenoptera/ Apoidea/
Megachilidae/ Osmia/ rufa

Wing length, Mass, Season

NCBI, ITIS, EOL, ...

Names

Chlorpyrifos, Chlorpyriphos,
Chlorphyrifos, Chlorpyrifos-ethyl,
Chlorpypifot

Hierarchies

organophospate, ester,
insecticide

Attributes

Mass, K_{OW}, LC_{50}

Identifiers

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

Biologists and Chemists face similar problems...

Names

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

Chlorpyrifos, Chlorpyriphos, Chlorphyrifos, Chlorpyrifos-ethyl, Chlorpypifot

Hierarchies

Hymenoptera/ Apoidea/ Megachilidae/ Osmia/ rufa

organophospate, ester, insecticide

Attributes

Wing length, Mass, Season

Mass, K_{OW}, LC_{50}

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)

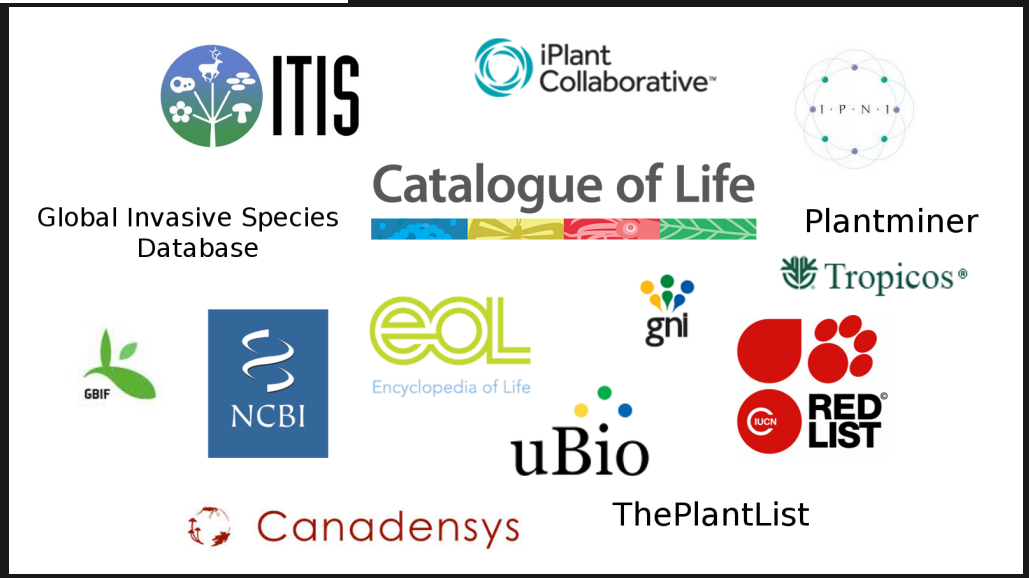
Instead of wasting time...

Instead of wasting time...

taxize - taxonomic search and retrieval in R

Instead of wasting time...

taxize - taxonomic search and retrieval in R



Instead of wasting time...

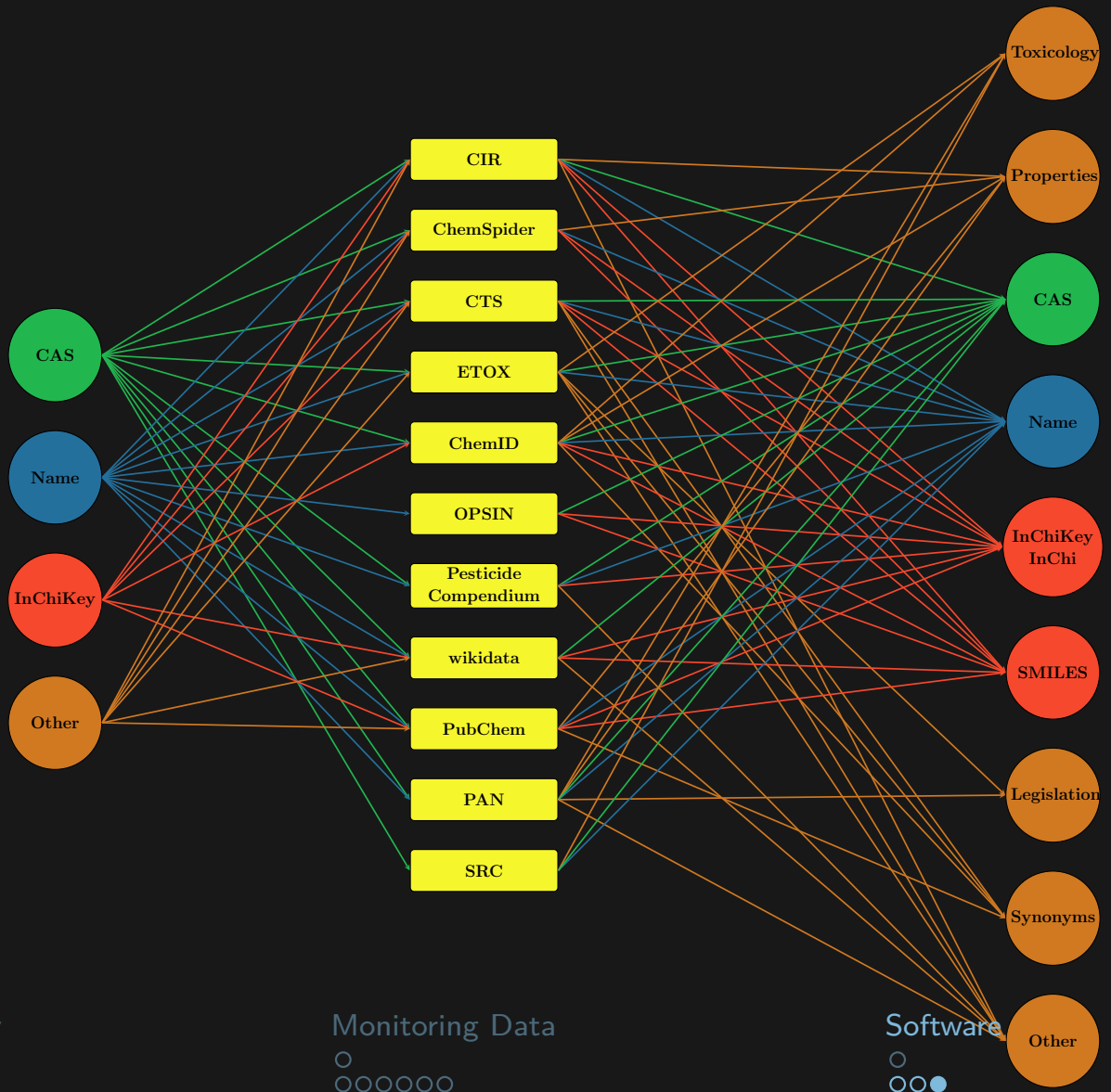
Instead of wasting time...

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

Münch (2016)

Instead of wasting time...

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



Münch (2016)

Conclusions from my PhD

- Change your model, not your data

Conclusions from my PhD

- Change your model, not your data
- Ultimately ban NOEC

Conclusions from my PhD

- Change your model, not your data
- Ultimately ban NOEC
- Monitoring data can be used to
  - study pesticide dynamics
  - inform ERA

Conclusions from my PhD

- Change your model, not your data
- Ultimately ban NOEC
- Monitoring data can be used to
  - study pesticide dynamics
  - inform ERA
- Agricultural SWB at risk from pesticides

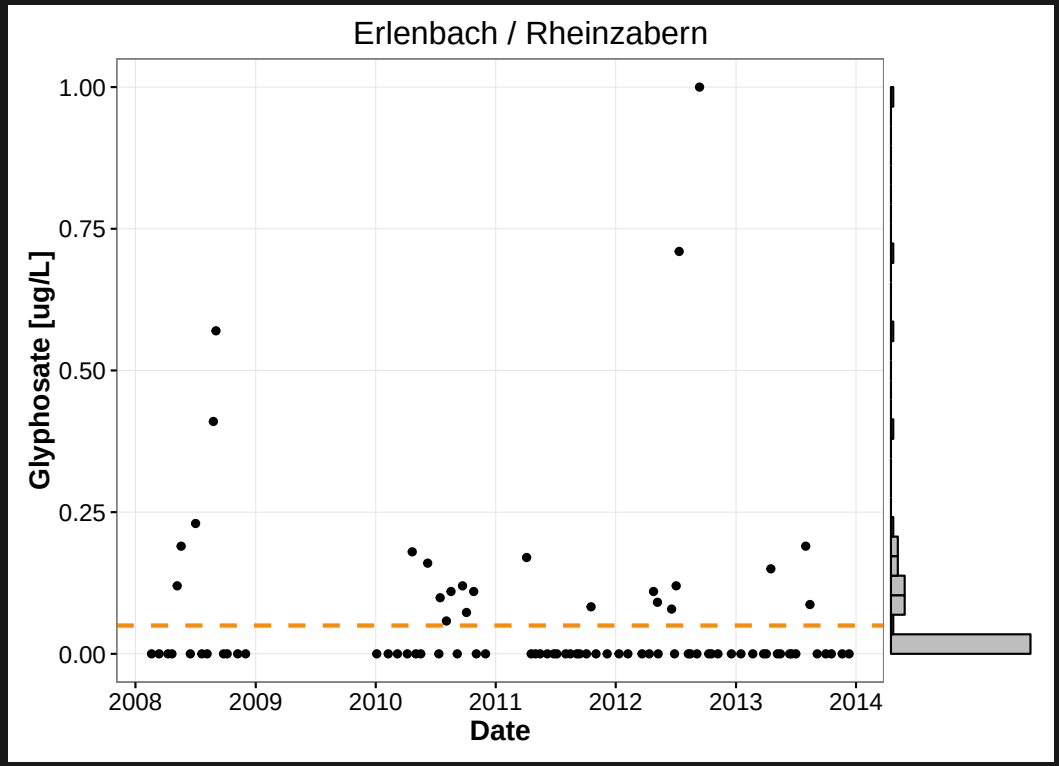
Conclusions from my PhD

- Change your model, not your data
- Ultimately ban NOEC
- Monitoring data can be used to
  - study pesticide dynamics
  - inform ERA
- Agricultural SWB at risk from pesticides
- Handling big eco(toxico-)logical data not easy
  - now easier

Outlook

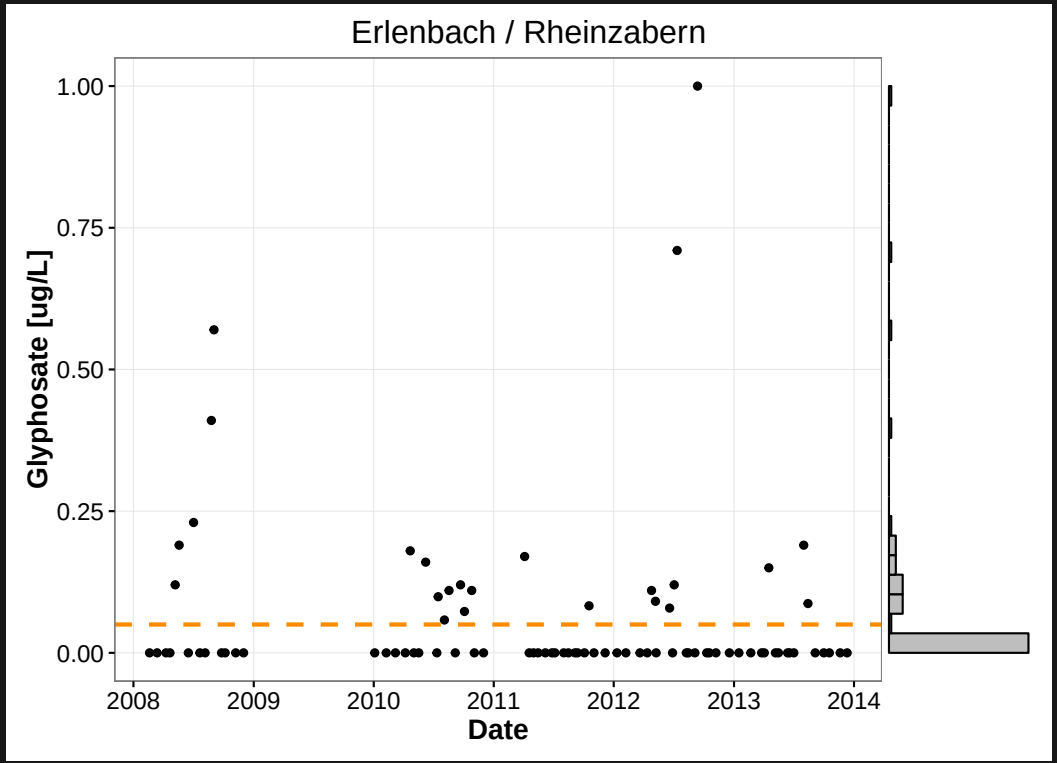
Analysing chemical concentrations is not easy, because of

- continuous distribution in $\mathbb{R}_0^+$
- censoring (x $<$LOQ)

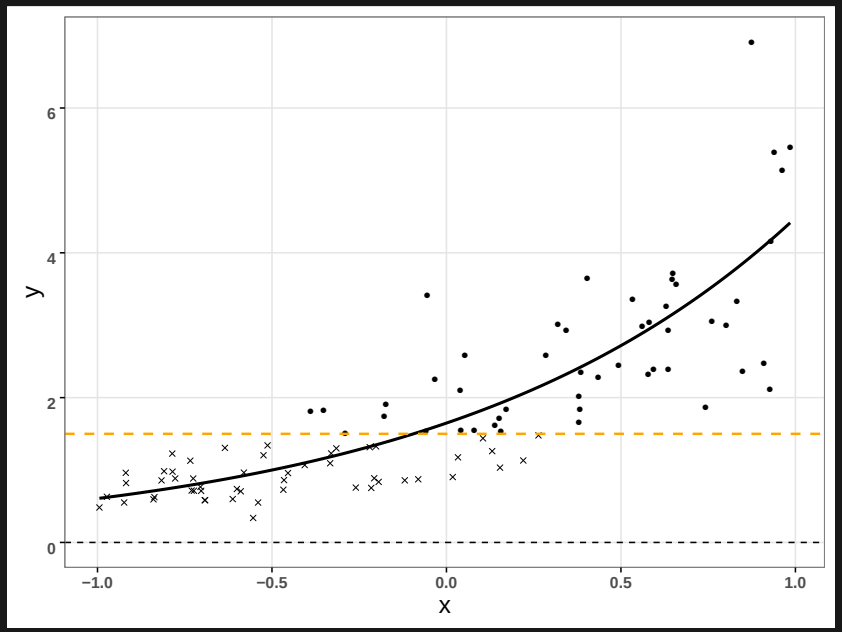


Analysing chemical concentrations is not easy, because of

- continuous distribution in $\mathbb{R}_0^+$
- censoring (x $<$LOQ)
- non-linearity (season, trends)
- dependency (spatial, temporal)
- missing data

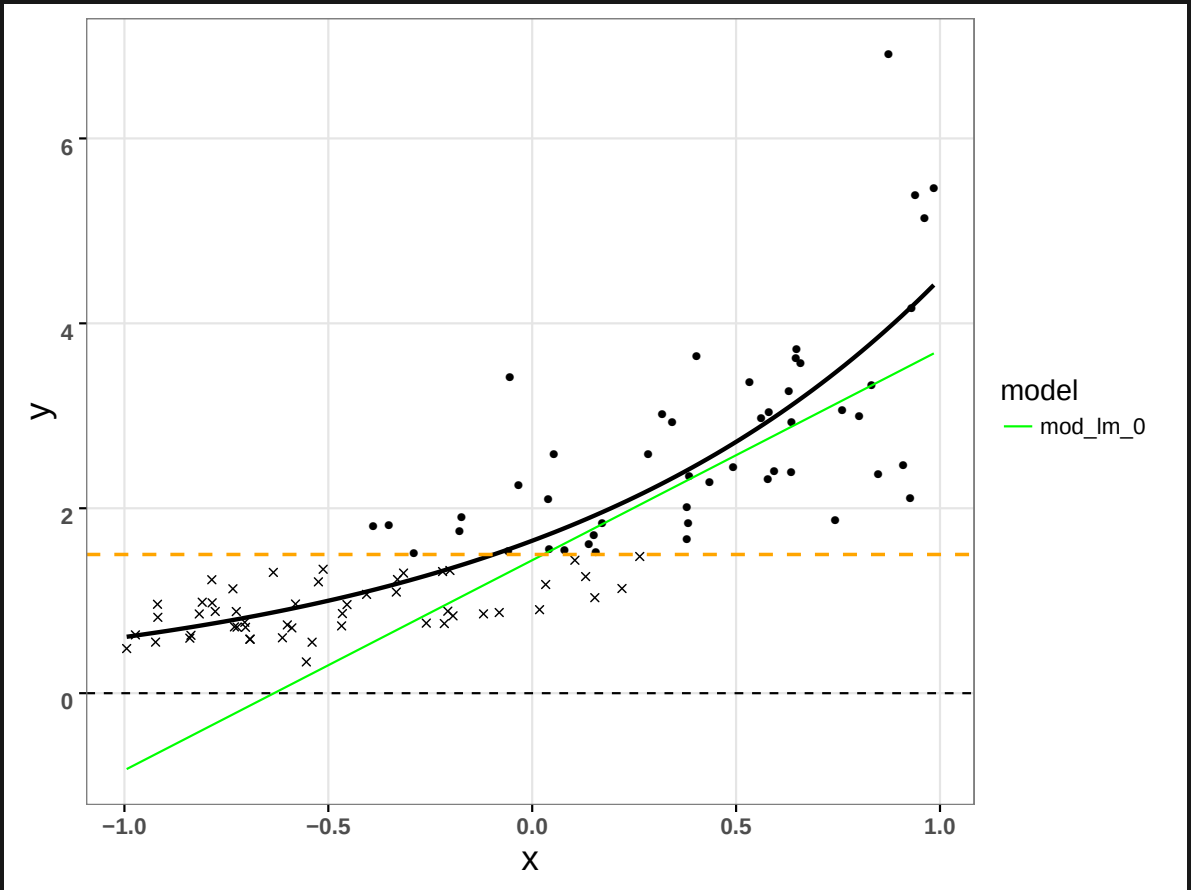


Dealing with censored, non-normal data

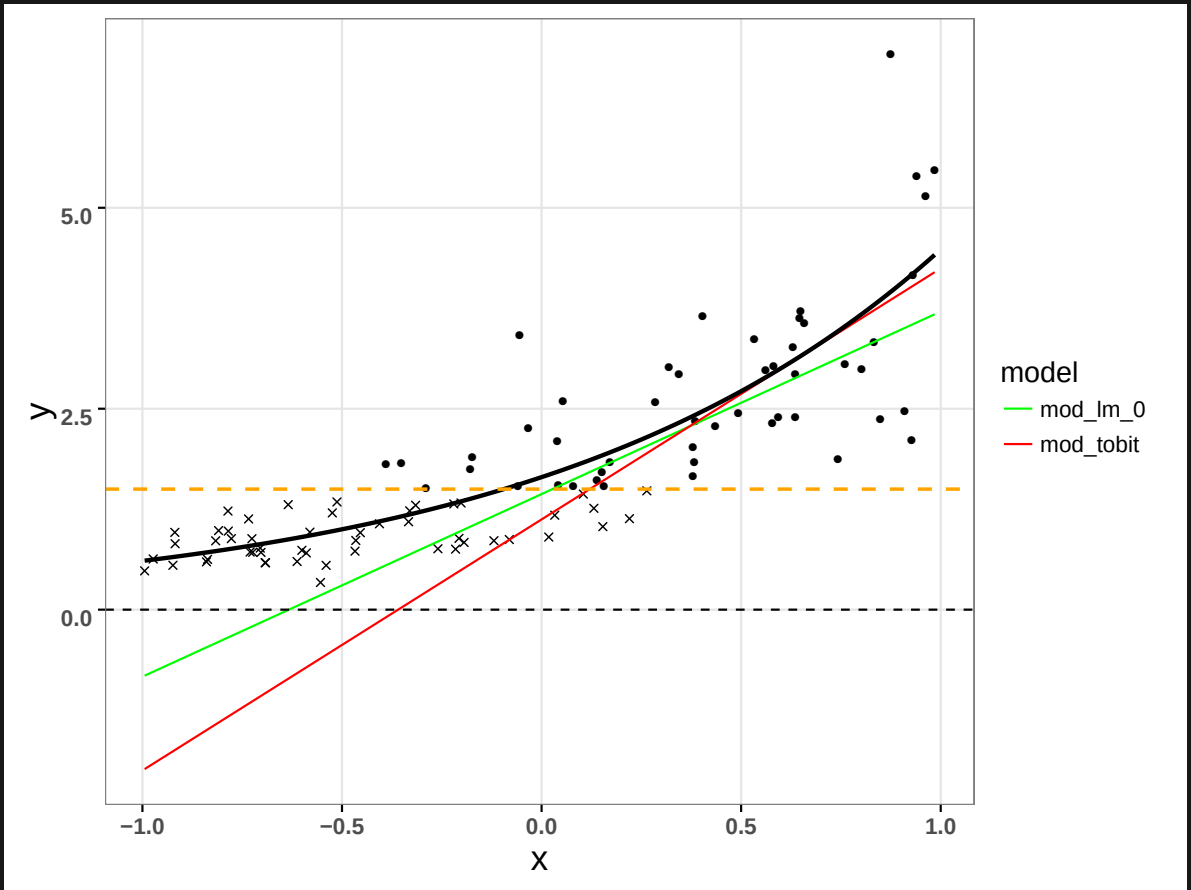


- Simulated data
- $y \sim Gamma(\mu, \kappa = 10)$
- $\mu = e^{0.5+x}$
- censoring at $c = 1.5$

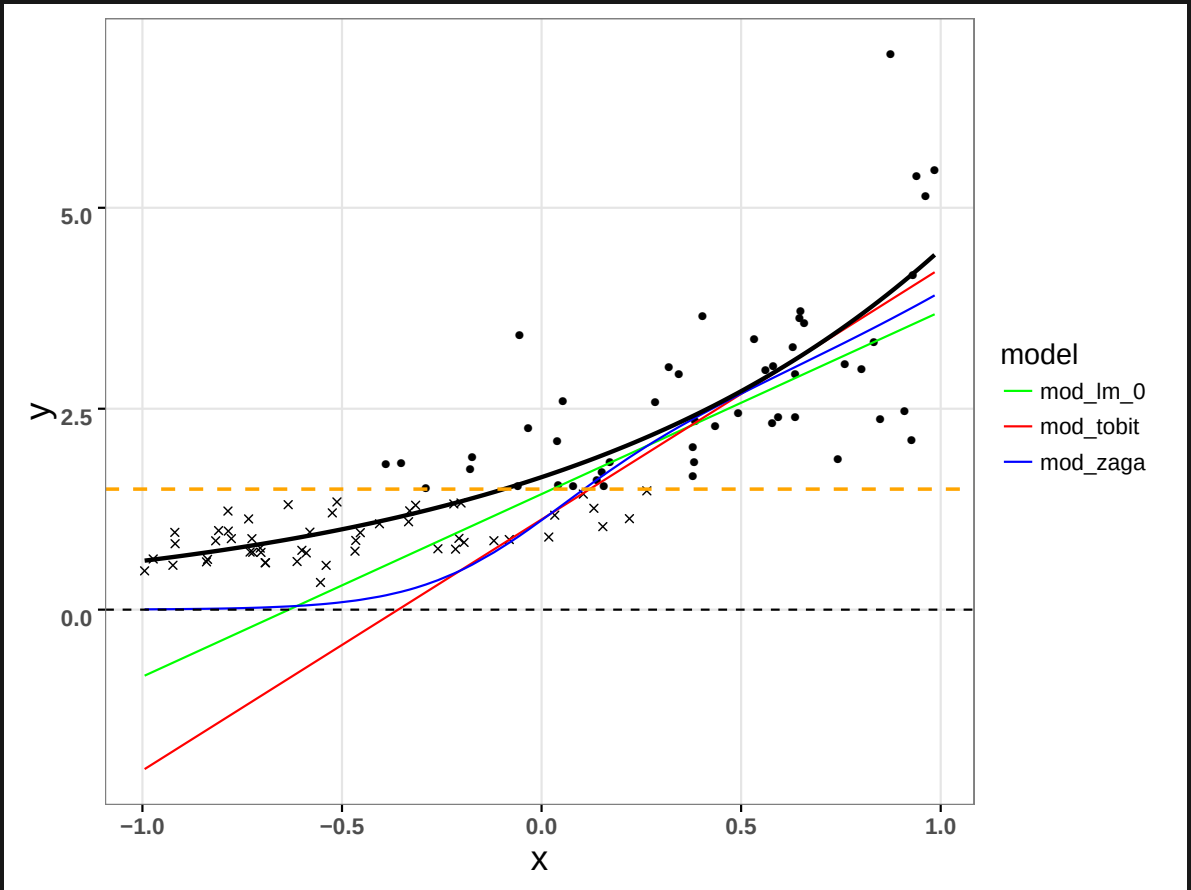
Dealing with censored, non-normal data



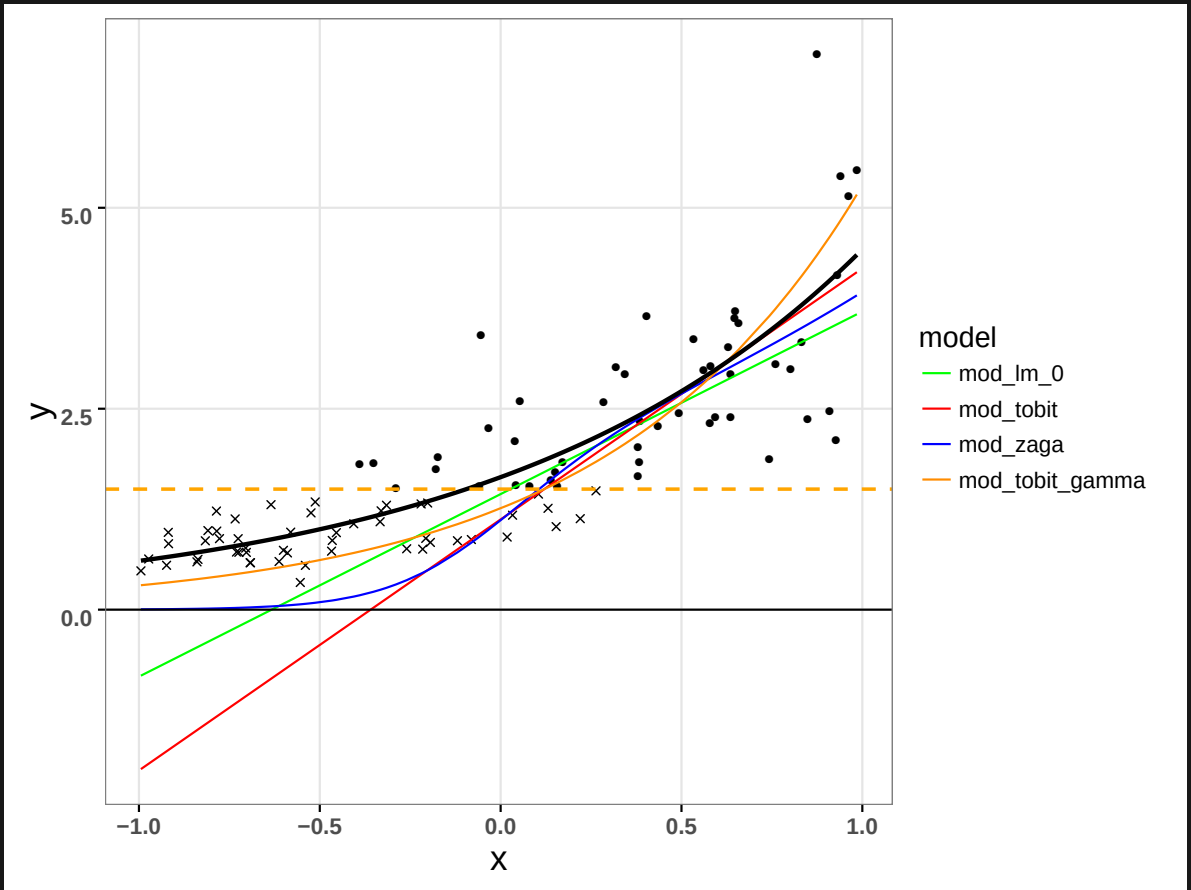
Dealing with censored, non-normal data



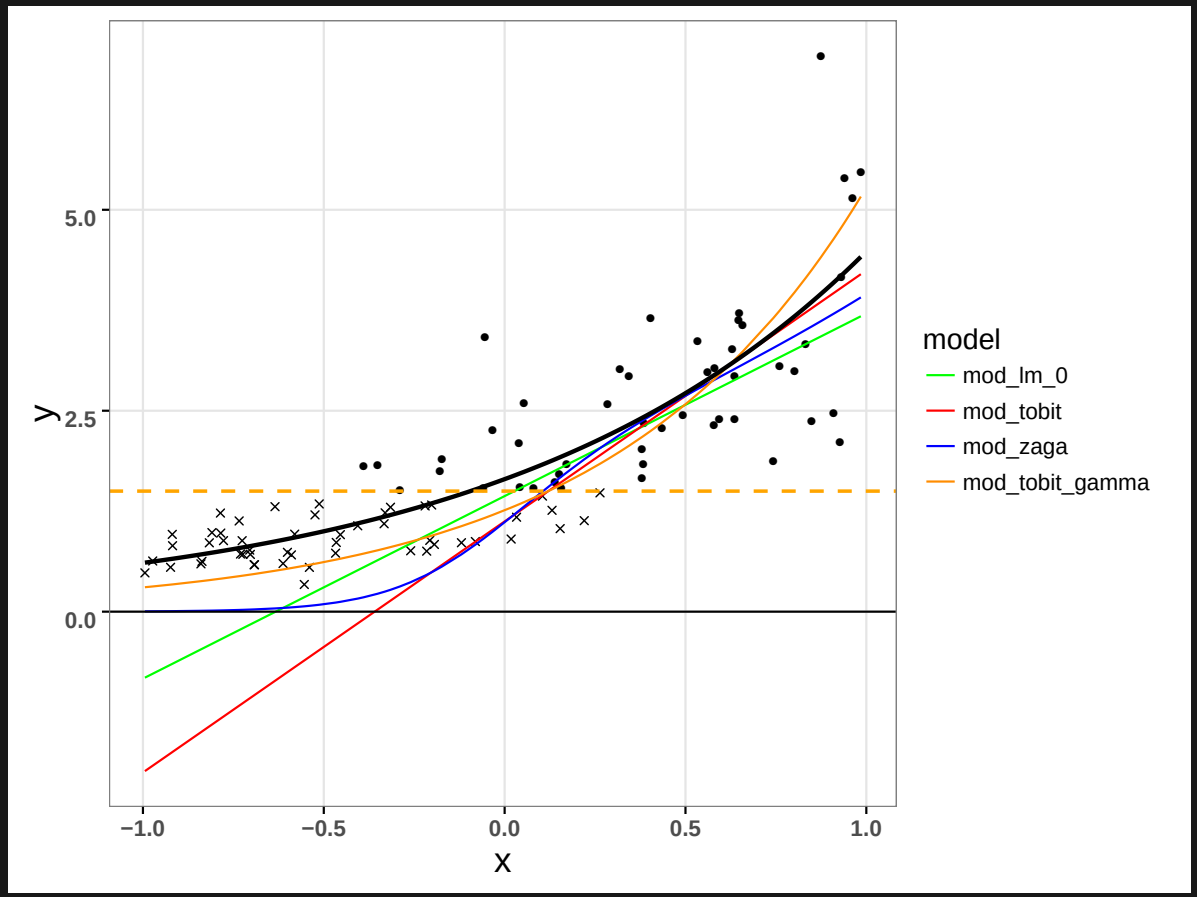
Dealing with censored, non-normal data



Dealing with censored, non-normal data



Dealing with censored, non-normal data



Guidance how to model environmental concentrations is missing

Temporal dynamics of pesticide occurrence

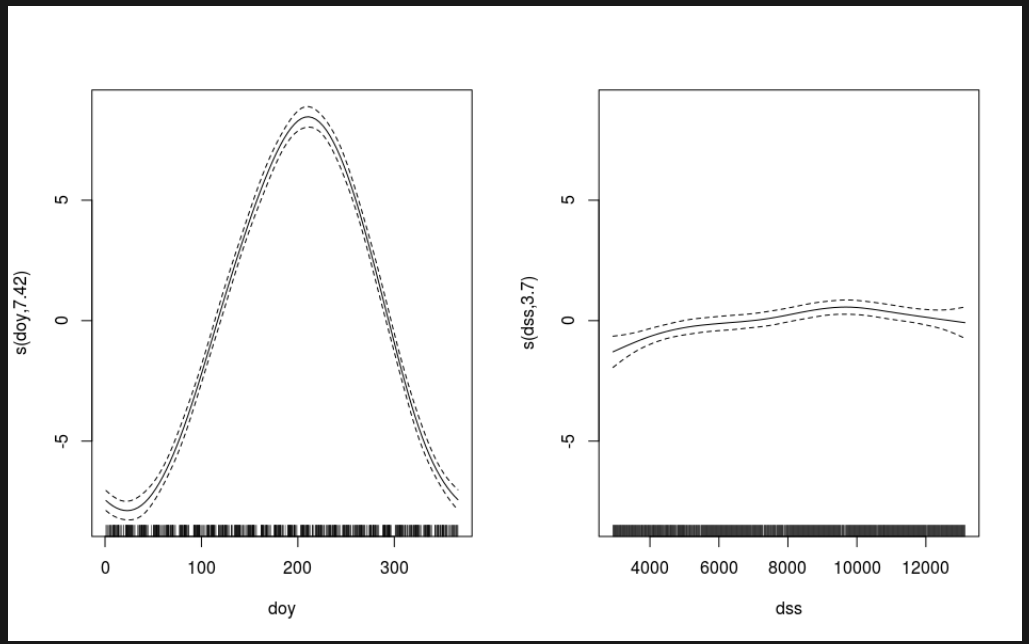
- Pesticides show compound specific dynamics
- Mixture dynamics? - Multivariate response.

Temporal dynamics of pesticide occurrence

- Pesticides show compound specific dynamics
- Mixture dynamics? - Multivariate response.
- Seasonality, Trends (Fade out...)?

Temporal dynamics of pesticide occurrence

- Pesticides show compound specific dynamics
- Mixture dynamics? - Multivariate response.
- Seasonality, Trends (Fade out...)?
  $y = \beta_0 + f_{seasonal}(x_1) + f_{trend}(x_2) + \epsilon; \epsilon \sim ???$



Statistical Ecotoxicology - Improving the utilization of data for ecological risk assessment

Eduard Szöcs

Institute for Environmental Sciences, University of Koblenz-Landau

http://edild.github.io/

@EduardSzoecs

szoecs@uni-landau.de

https://github.com/edild/talk_work2