# Slide 1

**BIOMATH**
Department of Applied Mathematics,
Biometrics and Process Control

UNIVERSITEIT GENT

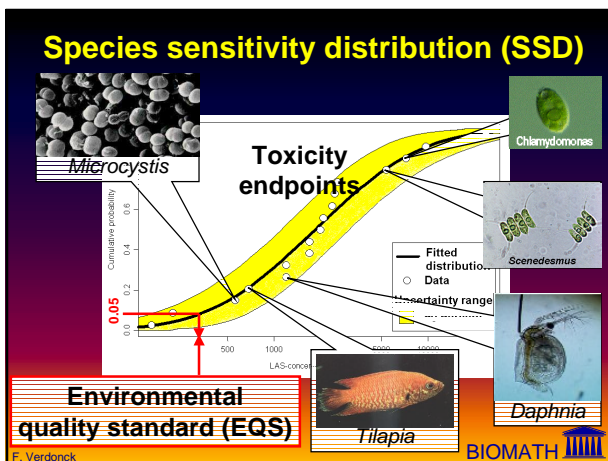## Added value of a (hierarchical) bootstrap model in environmental standard setting

Frederik A.M. Verdonck,
O. Thas, J. Jaworska & P. Vanrolleghem

RUG-Biomath, Coupure 653, 9000 Gent, Belgium (e-mail frederik.verdonck@rug.ac.be)

# Slide 2

## Outline

- Introduction
  - Environmental standard setting
  - Hierarchical structure of a toxicity data set
- Solutions:
  - Weighted hierarchical (bootstrap) model
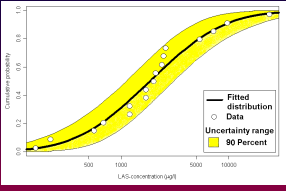  - Weighted non-hierarchical (bootstrap) model
- Discussion
- Conclusions

F. Verdonck

BIOMATH

# Slide 3

## Species sensitivity distribution (SSD)



*Microcystis*

**Toxicity endpoints**

*Chlamydomonas*

*Scenedesmus*

Fitted distribution
○ Data
Uncertainty range

*Tilapia*

*Daphnia*

**Environmental quality standard (EQS)**

0.05

F. Verdonck

BIOMATH

# Slide 4

## Variability <> uncertainty

Variability:
- true heterogeneity
- not reducible through further measurements

Uncertainty:
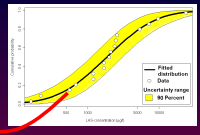- ignorance, partial knowledge
- partly reducible through further measurements



Fitted distribution
○ Data
Uncertainty range
■ 90 Percent

F. Verdonck
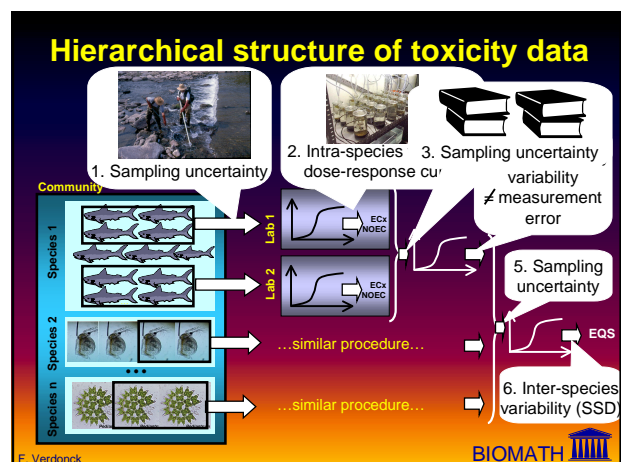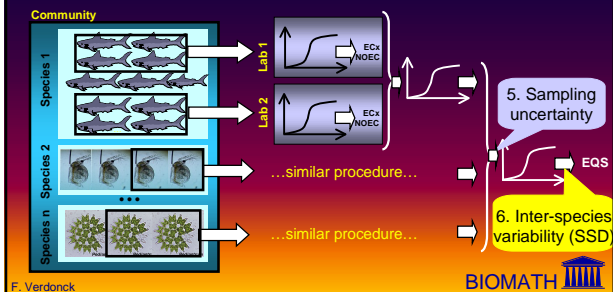
BIOMATH

# Slide 5

## Introduction

- Toxicity data are hierarchical in structure:
  - NOEC: No Effect Concentration
  - ECx: x % Effect Concentration
  - Other summary statistics ...

  

- A lot of information is discarded/lost

- Goal:
  - How to account for this extra info in a SSD?
  - Should a hierarchical or nonhierarchical model be used?

F. Verdonck

BIOMATH

# Slide 6

## Hierarchical structure of toxicity data



1. Sampling uncertainty
2. Intra-species dose-response curve
3. Sampling uncertainty
   variability ≠ measurement error
4.
5. Sampling uncertainty
6. Inter-species variability (SSD)

Community
Species 1
Lab 1
Lab 2
ECx NOEC
Species 2
Species n

…similar procedure…
…similar procedure…

EQS

F. Verdonck

BIOMATH

1

**'Usual' approach based on summary data**

Community
Species 1 | Lab 1 → ECx/NOEC | Lab 2 → ECx/NOEC
Species 2 …similar procedure…
Species n …similar procedure…

5. Sampling uncertainty
EQS
6. Inter-species variability (SSD)

F. Verdonck — BIOMATH



**Goal: base on all raw data**

1. Sampling uncertainty
2. Intra-species dose-response curve
3. Sampling uncertainty
4. Inter-laboratory variability

Community
Species 1 | Lab 1 → ECx/NOEC | Lab 2 → ECx/NOEC
Species 2 …similar procedure…
Species n …similar procedure…

5. Sampling uncertainty
EQS
6. Inter-species variability (SSD)

F. Verdonck — BIOMATH



**Weighting toxicity data**

> To avoid over/under-representation of a species, laboratory, individual!

Community
Species 1 | Lab 1
Lab 2 | $L_{ij}$ indivi-duals (4-8)
$m_i$ labs (1-5)
n species (5-20)

$$w_{ijk} = \frac{1}{L_{ij} \cdot m_i \cdot n}$$

F. Verdonck — BIOMATH
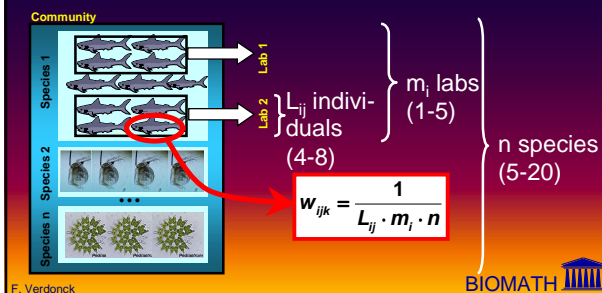


**Weighting toxicity data (cont'd)**

unweighted | weighted
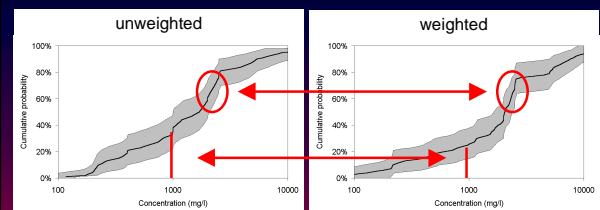
- Shape of the curve changes (shift)
- Uncertainty band changes in width

F. Verdonck — BIOMATH



**First solution: hierarchical models**

Several types:
- nonparametric <> parametric
- Bayesian <> frequentist
- analytical <> numerical

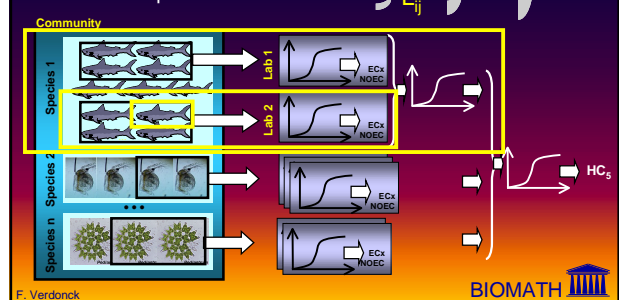All these have their advantages and disadvantages.
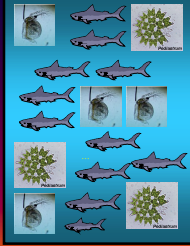
Main focus here is hierarchical or NOT?

F. Verdonck — BIOMATH



**Weighted hierarchical bootstrap**

1. Sample the species
2. Sample the laboratories
3. Sample the individuals

$L_{ij}$ $m_i$ n

Community
Species 1 | Lab 1 → ECx/NOEC | Lab 2 → ECx/NOEC
Species 2 → ECx/NOEC
Species n → ECx/NOEC

$HC_5$

F. Verdonck — BIOMATH

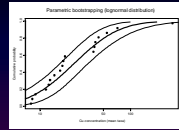## Second solution: non-hierarchical model

**Community**



- Ignore hierarchical structure of variabilities and treat all **raw** data (no summary statistics!) on same level
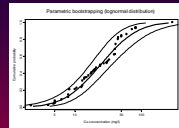- No hierarchical method but easier to understand and apply

F. Verdonck

BIOMATH

---

## Example: Cu



'Usual approach'
- Inter-species variability
- Sampling uncertainty (only 5.)



Hierarchical Model
- Inter-species + inter-laboratory (+ intra-species variability)
- 'Aggregated' sampling uncertainty

Non-hierarchical model
- Inter-species + inter-laboratory (+ intra-species variability)
- Uncertainty difficult to interpret

F. Verdonck
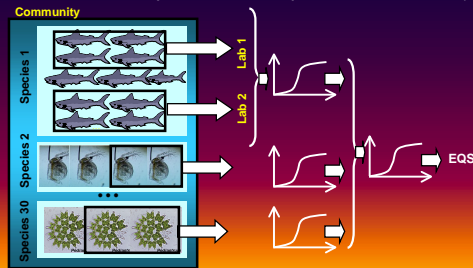
BIOMATH

---

## Validation

- Test
  - bias = estimated 5th percentile - true 5th percentile
  - coverage of confidence/uncertainty interval

- Simulation study with realistic, hypothetical data set:
  - 30 species from which 10 were sampled
  - Each species has specified inter-lab distribution
  - No assumption on inter-species distribution (top level)

F. Verdonck

BIOMATH

---

## Validation

- Simulation study with realistic, hypothetical data set:
  - 30 species from which 10 were sampled
  - Each species has specified inter-lab distribution
  - No assumption on inter-species distribution (top level)



F. Verdonck

BIOMATH

---

## Preliminary results

| | Raw/all data | |
| --- | --- | --- |
| | Hierarchical Model | Non-hierarchical Model |
| Mean bias | -0.3 | -0.9 |
| Variation bias | 1.4 | 1.4 |
| Coverage of 90% CI | ~ 93% | ~ 80% |

F. Verdonck

BIOMATH

---

## Preliminary results

| | Summary data | Raw/all data | |
| --- | --- | --- | --- |
| | 'Usual' approach | Hierarchical Model | Non-hierarchical Model |
| Mean bias | 19.3 | -0.3 | -0.9 |
| Variation bias | 6.1 | 1.4 | 1.4 |
| Coverage of 90% CI | ~ 0% | ~ 93% | ~ 80% |

Validation problem

F. Verdonck

BIOMATH

3

## Issue validation

- Can bias be separated in $bias_{model}$ and $bias_{summary<>all\ data}$?
- Model assumption on top level?



Community
Species 1
Species 2
Species 30
Lab 1
Lab 2
EQS
F. Verdonck
BIOMATH

## Issue validation (true 5th percentile?)

> implications for the hypothetical assumptions
  – bottom-up: is inter-lab truth and resulting inter-species following from that?
  – top-down: is inter-species truth and inter-lab additional?



Community
Species 1
Species 2
Species 30
Lab 1
Lab 2
EQS
F. Verdonck
BIOMATH

## Discussion

|  | Summary data | Raw/all data | |
|---|---|---|---|
|  | 'Usual' approach | Hierarchical Model | Non-hierarchical Model |
| 5th-percentile | + | +++ | ++ |
| Practical Use | +++ | + | +++ |
| Interpretation of Confidence band | + | +++ | - |

F. Verdonck
BIOMATH

## Conclusions

- To incorporate intra-species and inter-laboratory variability into a Species Sensitivity Distribution, a weighted non-hierarchical model is sufficiently accurate and easy-to-use.

- To incorporate several levels of sampling uncertainties, a weighted hierarchical model is more appropriate.

- Difficult to validate.

F. Verdonck
BIOMATH

---

UNIVERSITEIT GENT

BIOMATH
Department of Applied Mathematics, Biometrics and Process Control

### Added value of a (hierarchical) bootstrap model in environmental standard setting

Frederik A.M. Verdonck,
O. Thas, J. Jaworska & P. Vanrolleghem

RUG-Biomath, Coupure 653, 9000 Gent, Belgium (e-mail frederik.verdonck@rug.ac.be)