Combining exposure assessment and effect assessment by TKTD models

Andreas Focks, Maria Arena, Theo Brock, Nina Cedergreen, Sandrine Charles, Sabine Duquesne, Alessio Ippolito, Michael Klein, Melissa Reed, and Ivana Teodorovic

Brussels, 23. October 2018

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

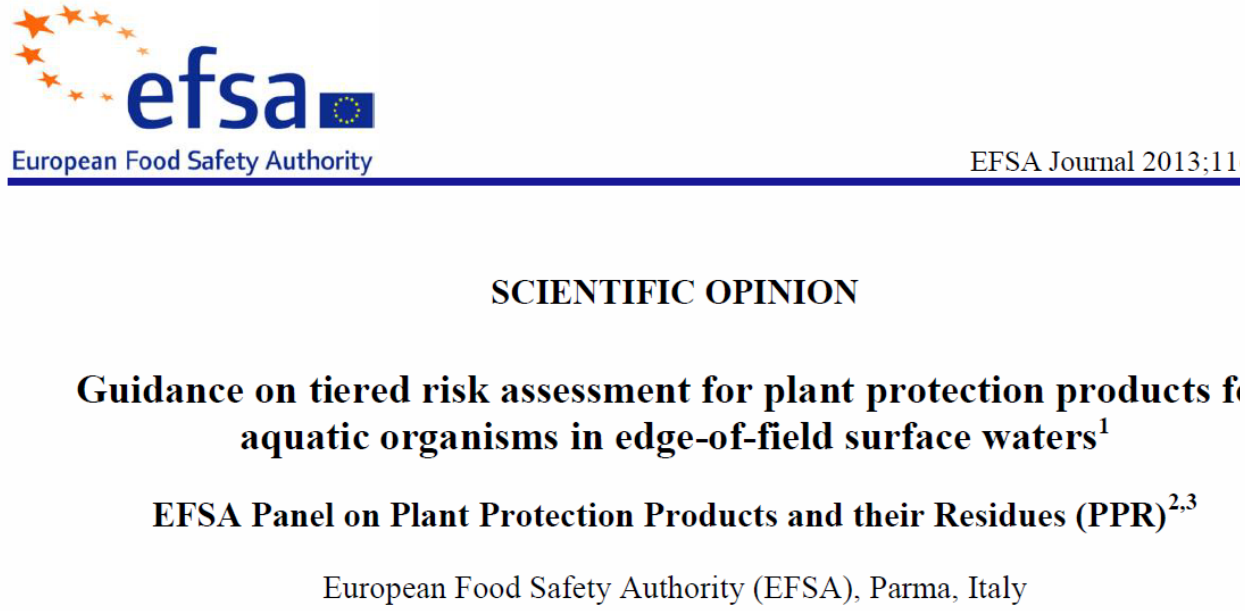
SETAC Europe 13th Special Science Symposium
Extrapolation of Effects Across Biological Levels:
Challenges to Implement Scientific Approaches in Regulation
23–24 October 2018 I Brussels, Belgium

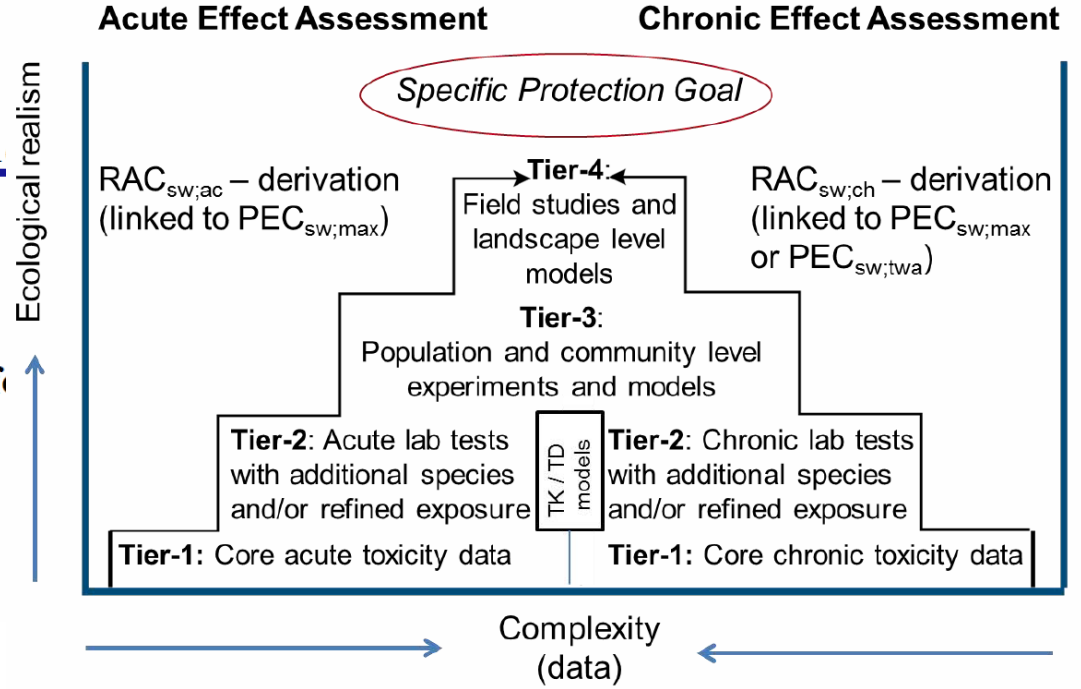
Overview

- Regulatory environment
- Short overview general TKTD
- Evaluation of GUTS usage in the scope of regulatory risk assessment
- Linking GUTS to existing guidance
- Outlook

Regulatory background

- 2008: the Panel on Plant Protection Products and their Residues (PPR) was tasked by EFSA with the revision of the Guidance Document on Aquatic Ecotoxicology under Council Directive 91/414/EEC
- In 2013, "Aquatic guidance document" was published, focus on experimental approaches for pelagic water organisms, already indicating that mechanistic effect models could be used within the tiered approach





WAGENINGEN UNIVERSITY & RESEARCH

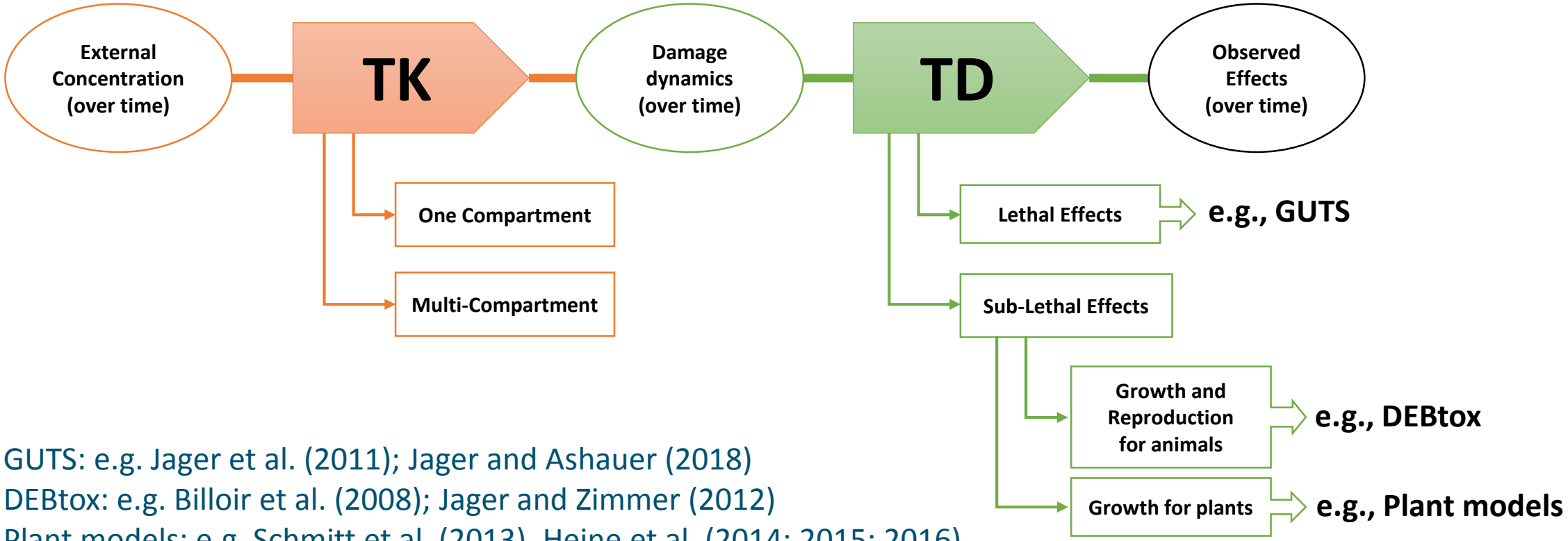
100years 1918 — 2018

Scientific opinion (SO) on modelling - which models?

- Initially, SO should cover general mechanistic effect models (MEM) as tools for the prospective effect assessment procedures for aquatic organisms – on all higher tiers – for individuals, populations, communities...
- Due to huge variety of MEM, their different developmental stages, and open issues with the use of such models for higher tier ERA (e.g. competition, alternative stable states, etc) constrained focus on TKTD models as Tier-2 tools

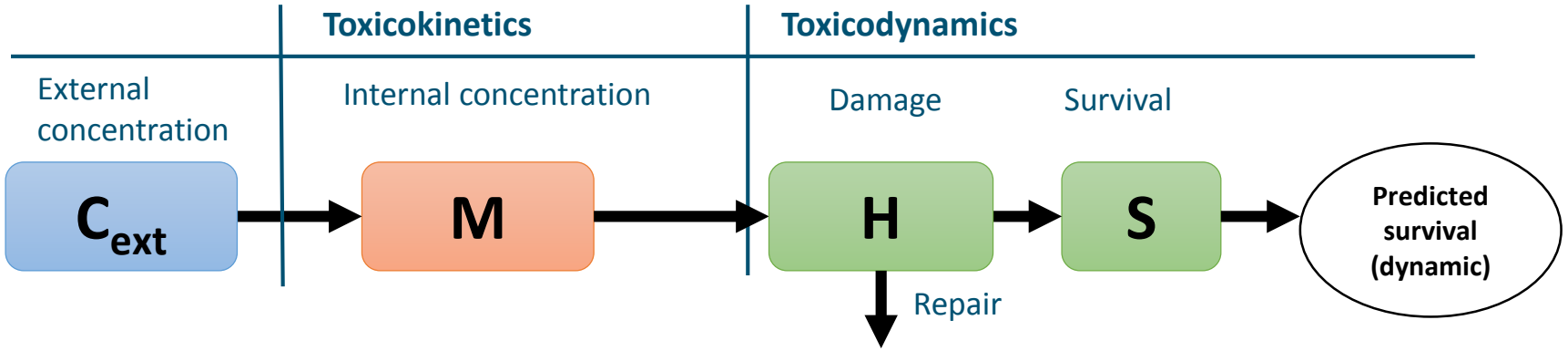
- Stage of development of TKTD models was considered being close to allowing appropriate use in the prospective environmental risk assessment for pesticides, particularly to predict potential risks of time-variable exposures on aquatic organisms
- An EFSA working group developed a scientific opinion about TKTD models for aquatic organisms between December 2016 and May 2018

TKTD models - principles



GUTS: e.g. Jager et al. (2011); Jager and Ashauer (2018)
DEBtox: e.g. Billoir et al. (2008); Jager and Zimmer (2012)
Plant models: e.g. Schmitt et al. (2013), Heine et al. (2014; 2015; 2016).

GUTS: General Unified Threshold models of Survival



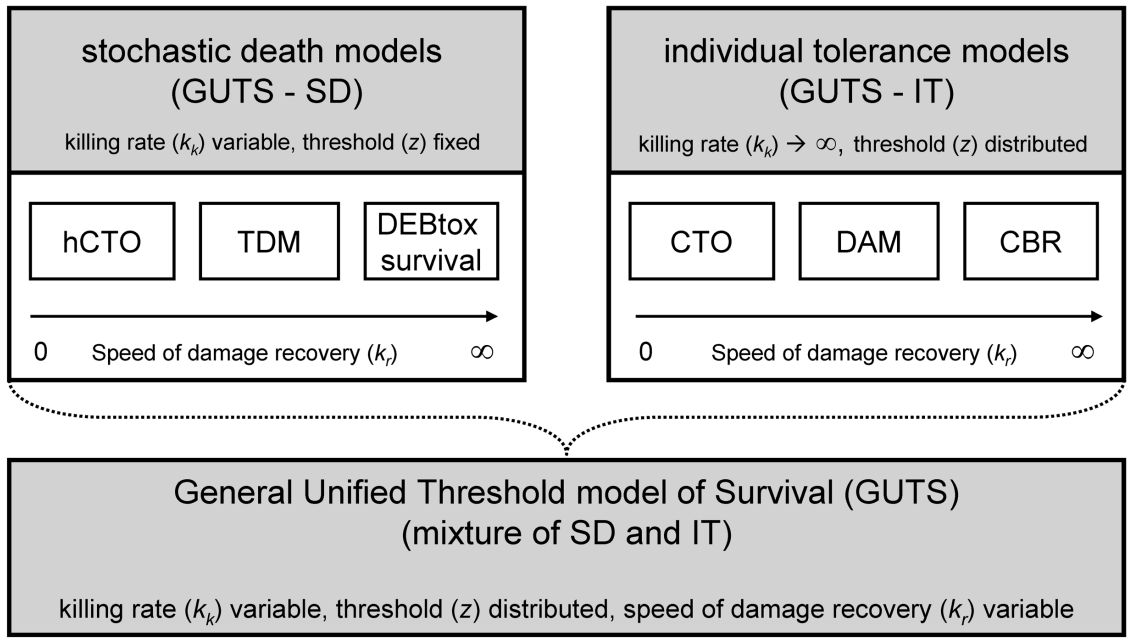
- Uptake processes:
  - Diffusion across membranes
  - Filtration via breathing organs
- Distribution
- Biotransformation
- Excretion

- Damage
- Repair/recovery
- Mortality: Probability not to survive increases with increasing damage

Jager, Albert, Preuss, Ashauer (ES&T 2011): Development of General Unified Threshold models of Survival as comprehensive theoretical foundation for TKTD models of survival

Toxicodynamics-extreme cases:
Stochastic death (SD)
Individual tolerance (IT)

Two models for survival – why?



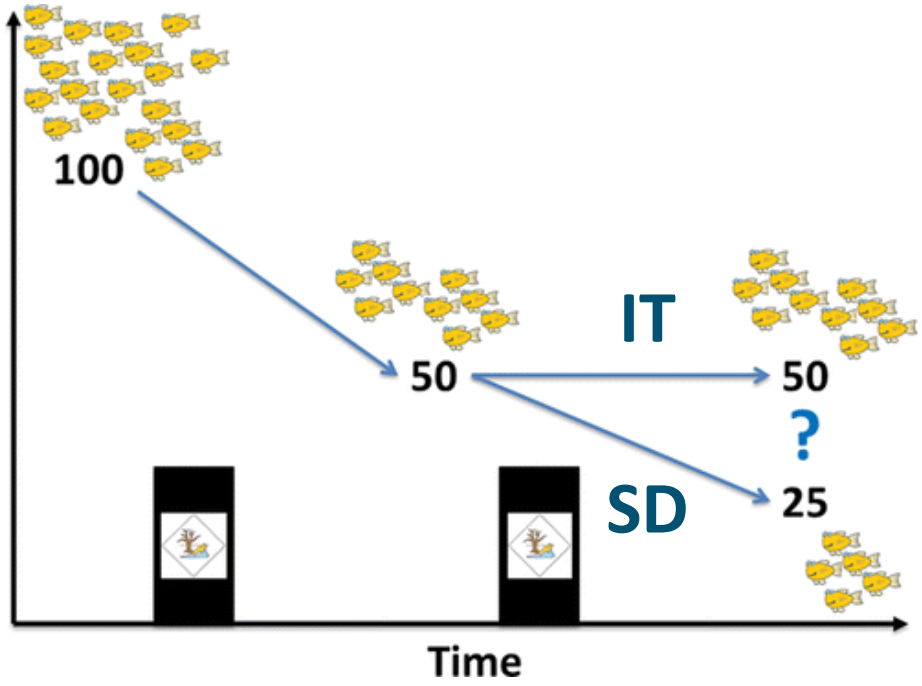
From Jager et al. (2011):

Toxicodynamics-extreme cases:

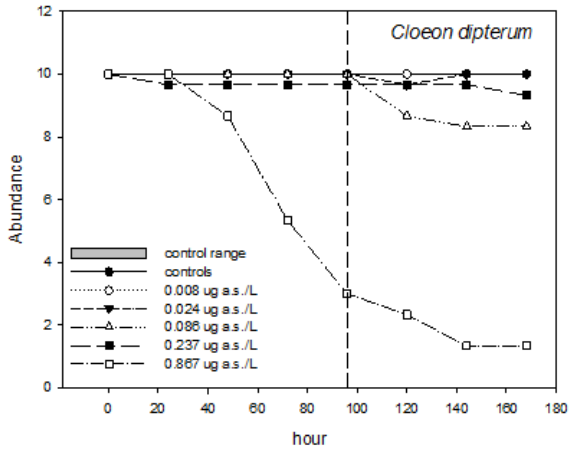
Stochastic death (SD): Threshold fixed, killing rate variable

Individual tolerance (IT): Threshold distributed, immediate killing

Death Dilemma and Organism Recovery in Ecotoxicology (Ashauer et al., 2015)

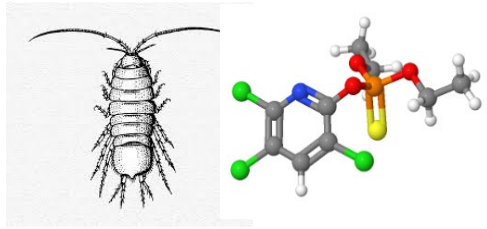


How GUTS modelling works: Calibration, prediction, validation



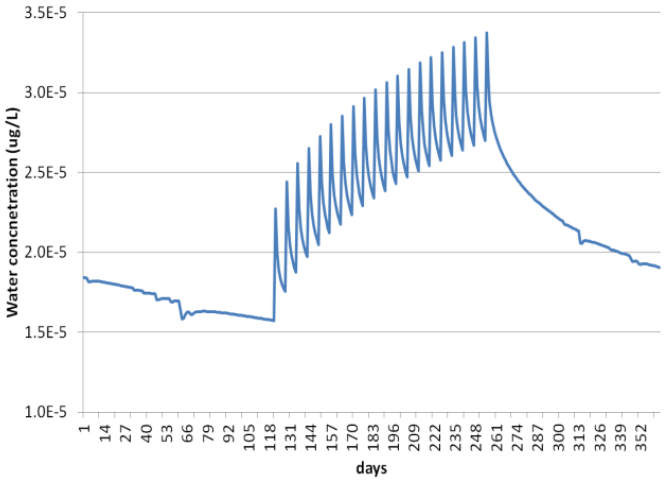
Survival under static exposure: standard ecotox tests

Calibration



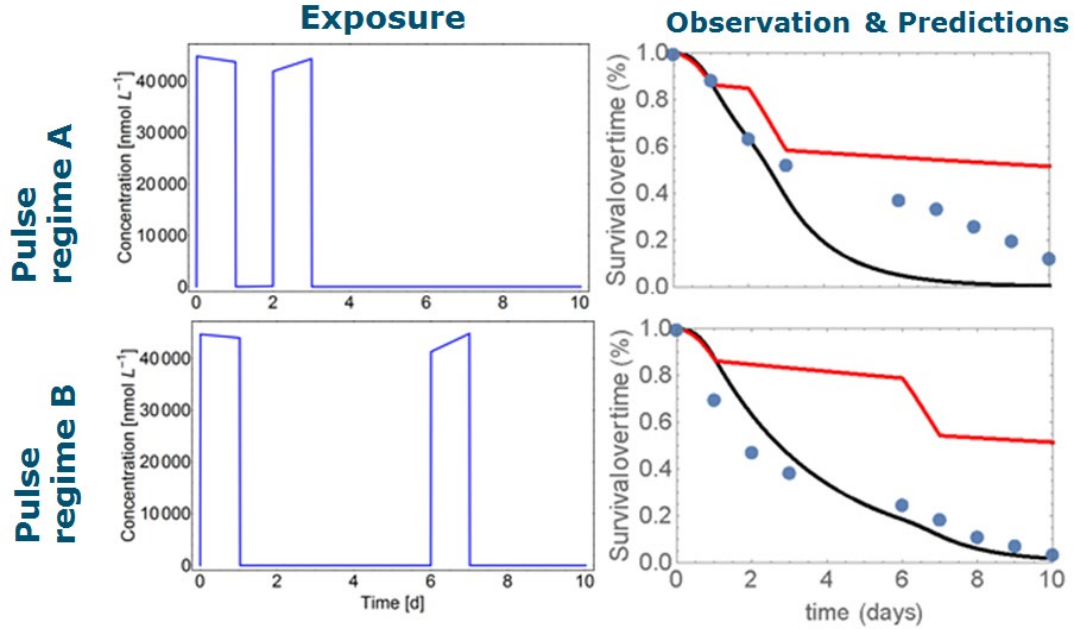
Species-substance specific parameterisation

Prediction



Survival under time-variable exposure

Validation: specific experiments, Testing model predictions of survival in experiments



Evaluation of GUTS usage in regulatory risk assessment

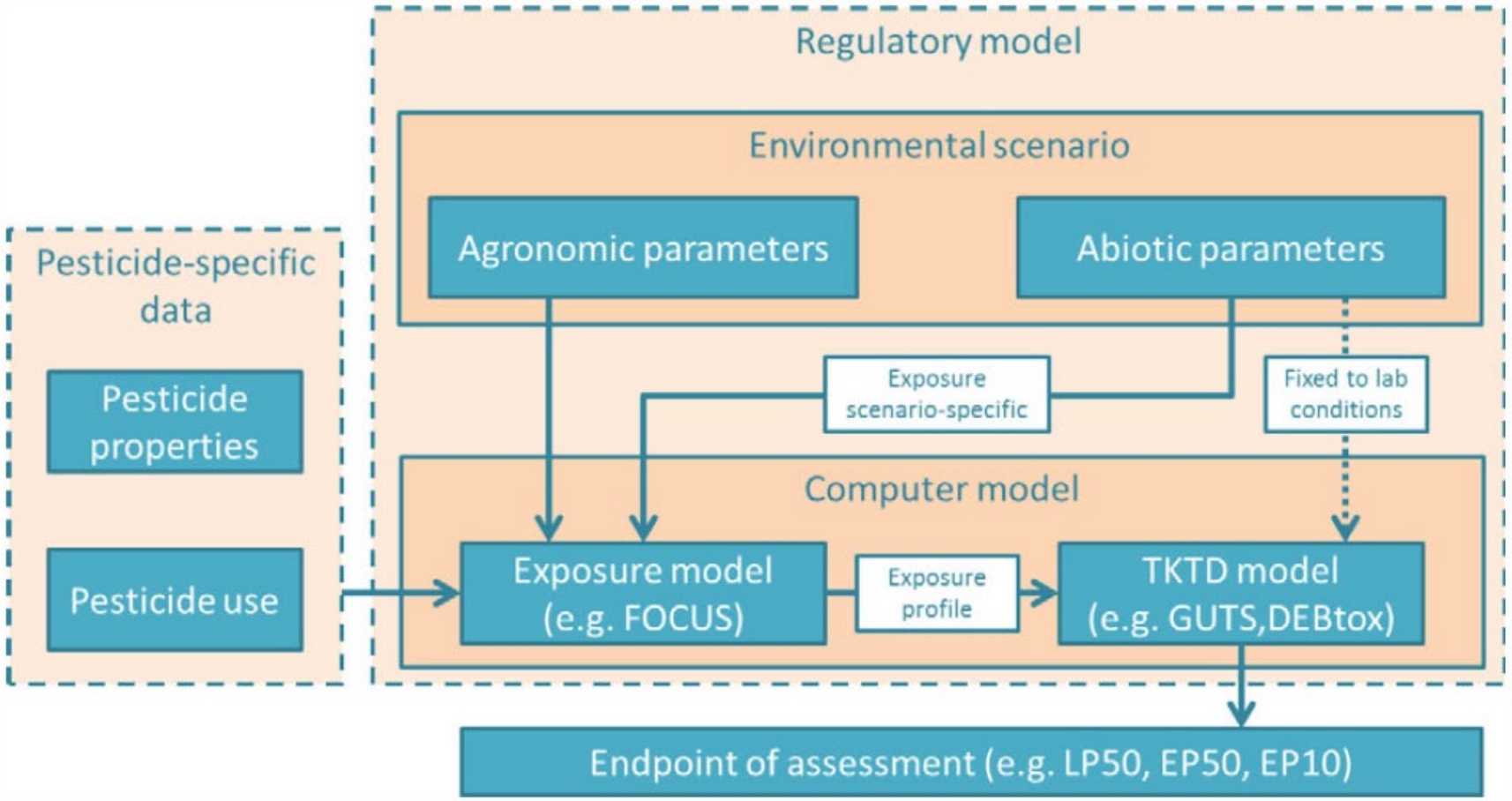
- Follows structure as elaborated in EFSA scientific opinion on 'Good Modelling Practice':
- TKTD SO formulates methods and examples for the evaluation of
  - Problem definition
  - Quality of the supporting experimental data
  - Conceptual model
  - Formal model
  - Computer model
  - Regulatory model (environmental scenarios, parameter estimation)
  - Model analysis (Sensitivity and uncertainty analysis, validation)
  - Model use

Some aspects are in the TKTD SO evaluated for GUTS models **in general**

Other aspects have to be tested and documented per GUTS **implementation**

Some aspects remain to be tested and documented for each GUTS **application**

Regulatory model for GUTS

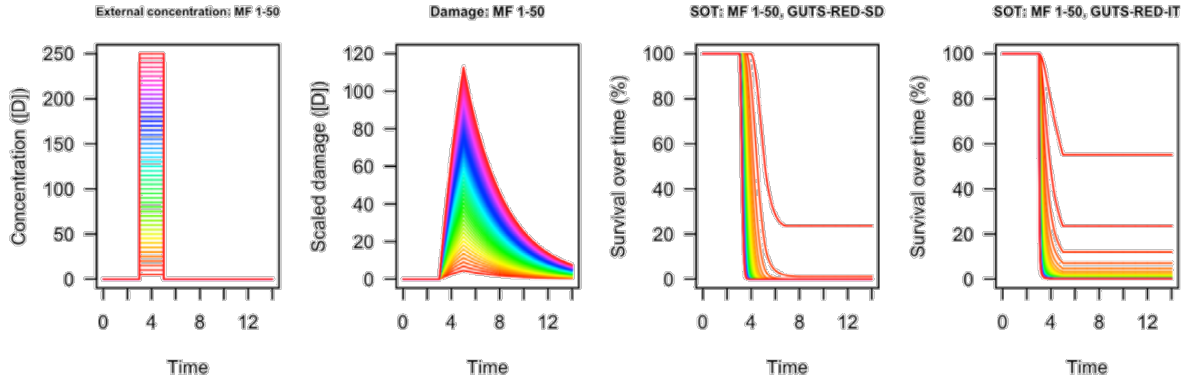


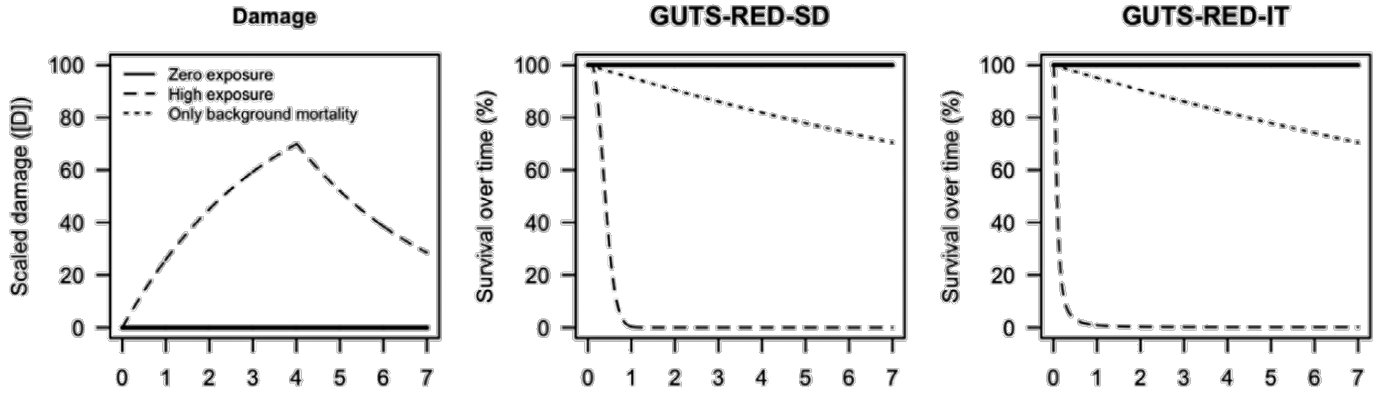
- Environmental scenario feeds into exposure model.
- Exposure profile is used as input by the TKTD models.
- TKTD output gives endpoint of assessment

Checking the implementation of GUTS models: three lines of evidence

1. Test against the ring-test data set (Jager and Ashauer, 2018).
2. Test of a set of scenarios (default, pulsed and 'extreme' cases);
3. Test model output with an independent implementation of GUTS

In addition, availability of the computer code allows further implementation check by experts



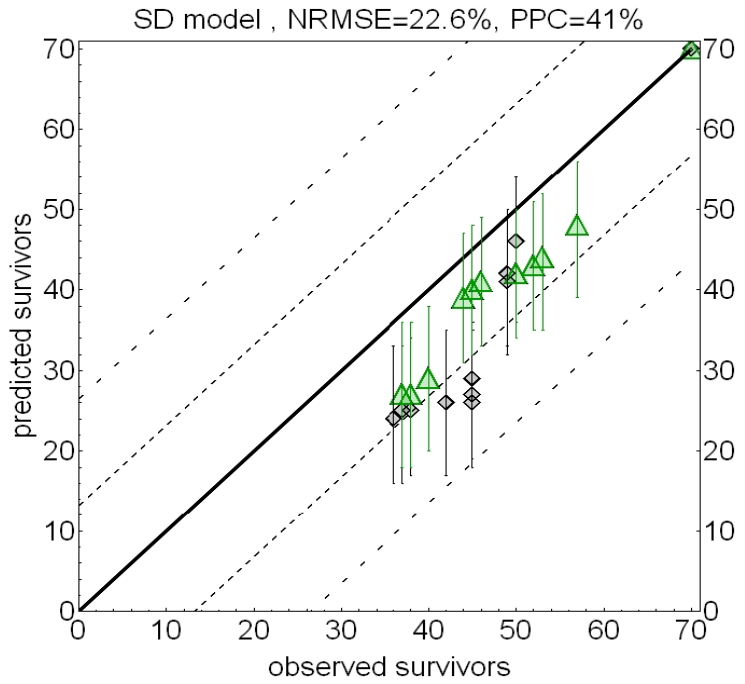


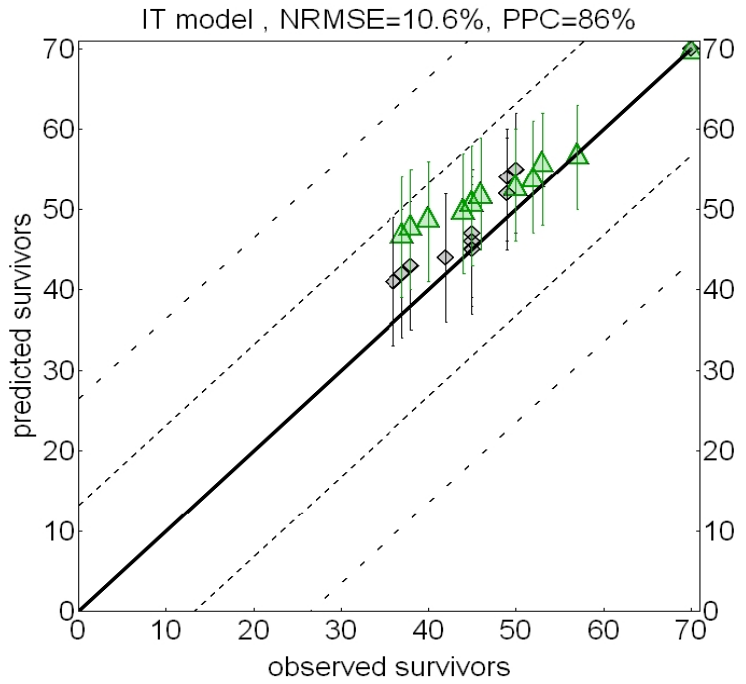
Model validation

- Validation data: have not been used for model calibration, provide relevant output (for GUTS simulated mortality probability over time and LPx/EPx values).
- Special consideration of vertebrates (reduction of vertebrate testing).
- Three different quantitative criteria suggested, to be considered in combination, applicable for both frequentist and Bayesian approaches
    1. Posterior predictive check (PPC): uncertainty
    2. Normalised Root Mean Square Error (NRMSE): match over time
    3. Survival Probability Prediction Error (SPPE): match of final survival
- All criteria deliver absolute indicator values (percentages) which can be interpreted and compared with thresholds

Quantification of prediction quality: Posterior predictive check (PPC)

- PPC based on Bayesian statistics
- Compares predicted mean number of survivors with observed numbers under specific consideration of uncertainty in the model predictions.





PPC between 50 and 90% indicates appropriate uncertainty ranges

Quantification of prediction quality : NRMSE und SPPE

- Normalised RMSE :

$$NRMSE = \frac{RMSE}{\bar{Y}} = \frac{1}{\bar{Y}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{obs,i} - y_{pred,i}\right)^2}$$

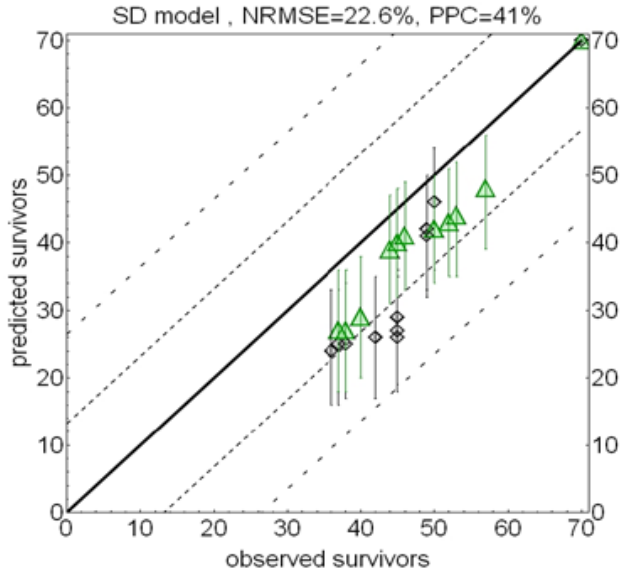
NRSME below 30-50% indicate good prediction of survival over time

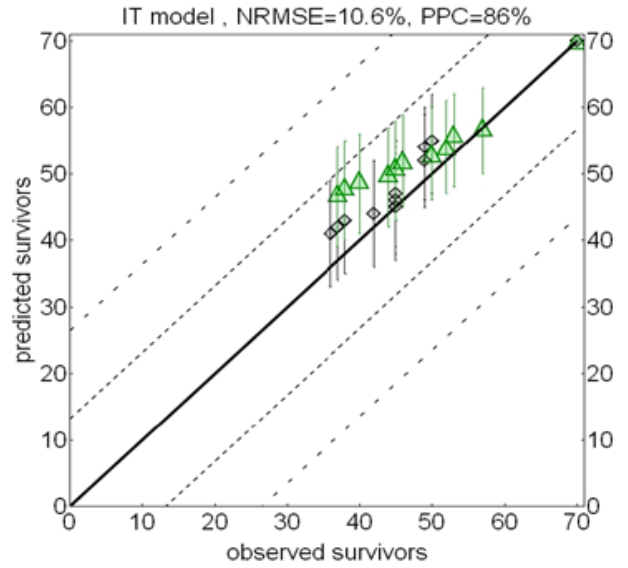
with $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_{obs,i}$ as mean of *n* observed survivors $y_{obs,i}$.

- Survival probability prediction error (SPPE): Evaluation of survival probabilities between beginning and end of validation experiments

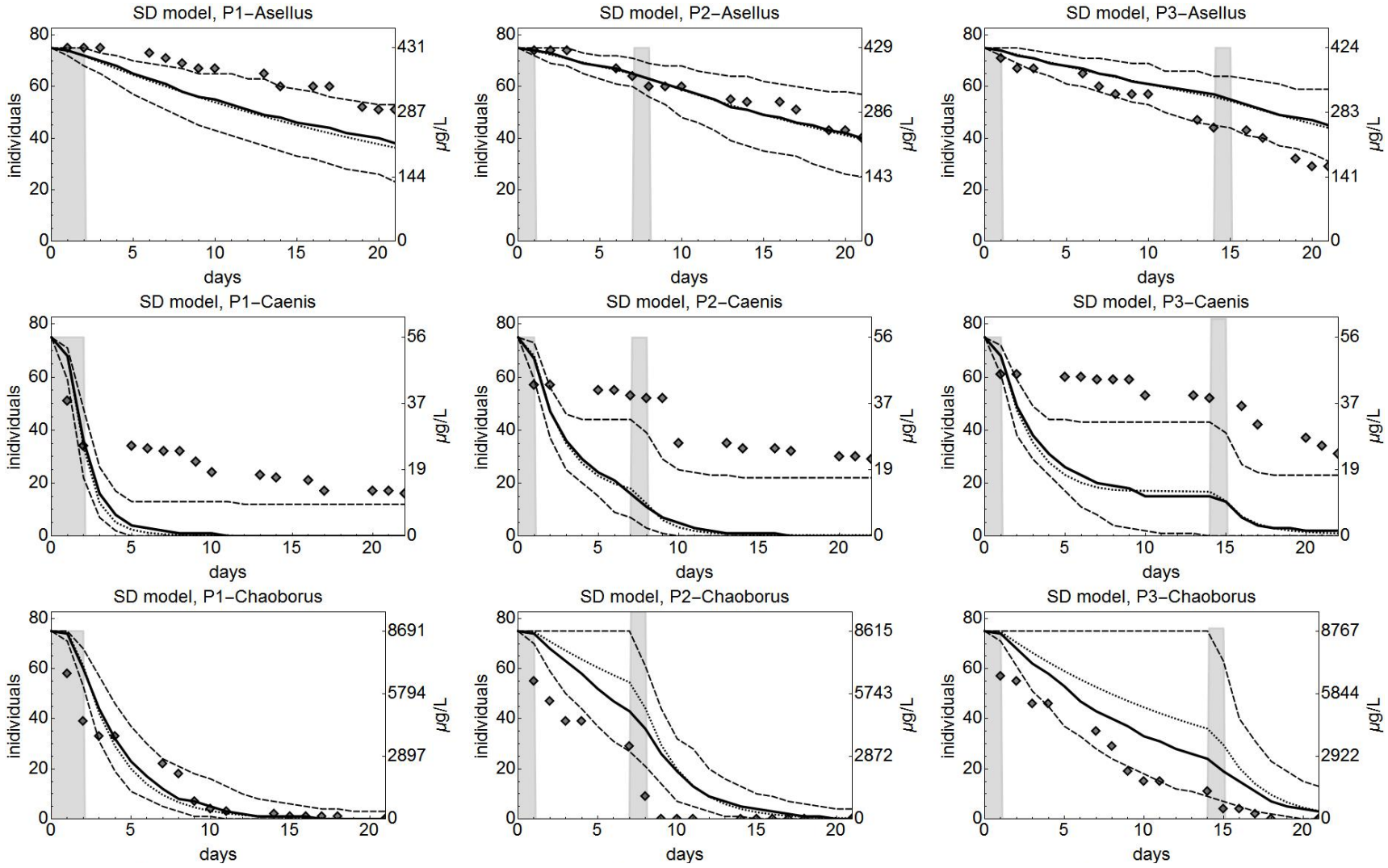
$$SPPE = \left(\frac{y_{obs,tend}}{y_{init}} - \frac{y_{modelled,tend}}{y_{init}}\right) * 100 = \frac{y_{obs,tend} - y_{modelled,tend}}{y_{init}} * 100$$

SPPE below 0% indicate underestimation of mortality.
SPPE between 0 und 30-50% can be considered to show good prediciton of final mortalities.





Example validation results: Imidacloprid



Asellus aquaticus

PPC	93%
NRSME	15%
SPPE	17 / 0 / -21%

Caenis horaria

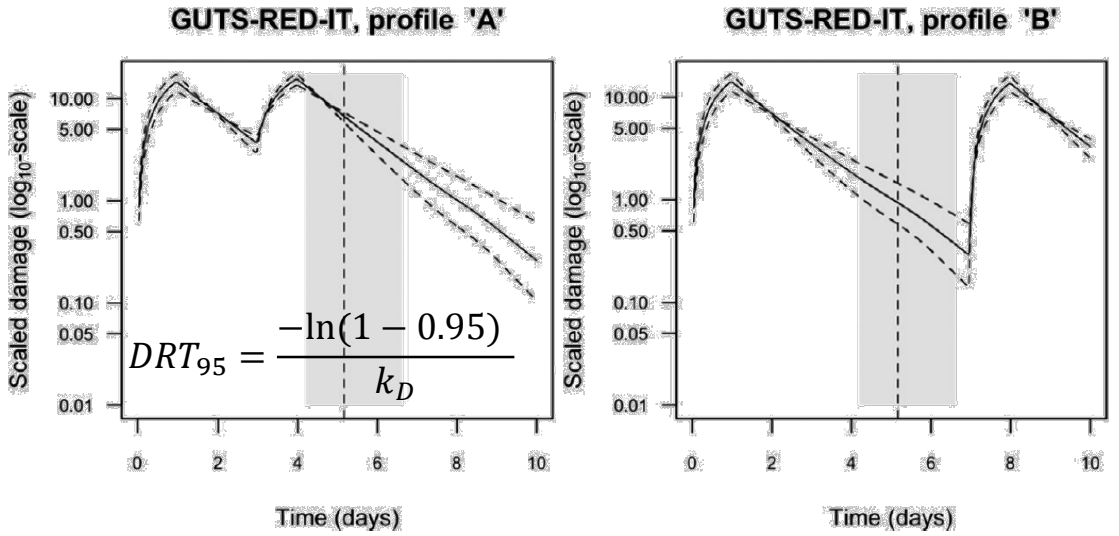
PPC	50%
NRSME	69%
SPPE	21 / 39 / 39%

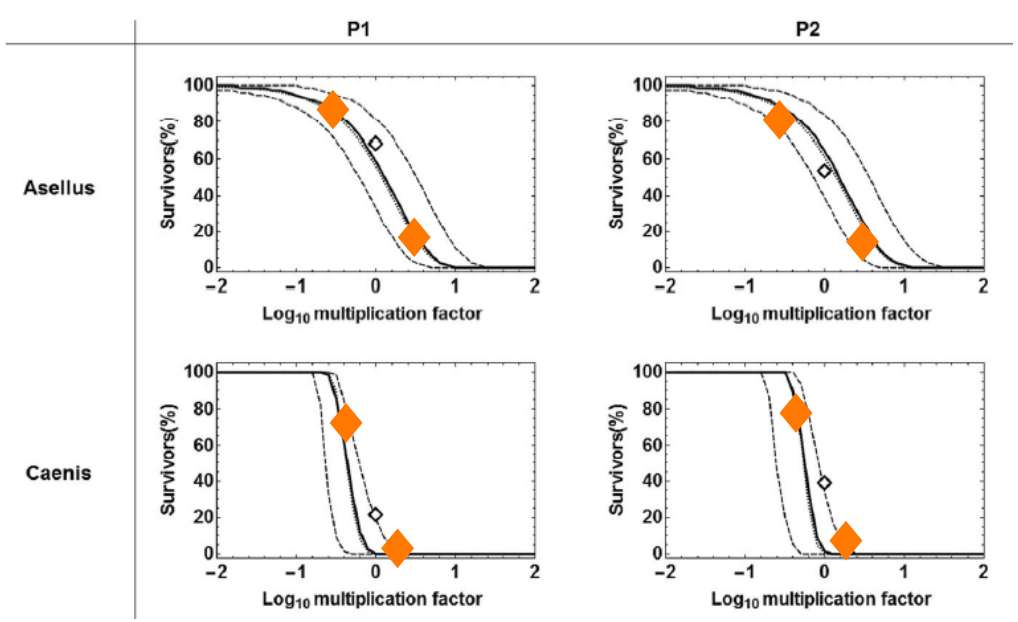
Chaoborus obscuripes

PPC	78%
NRSME	60%
SPPE	0 / 0 / -4 %

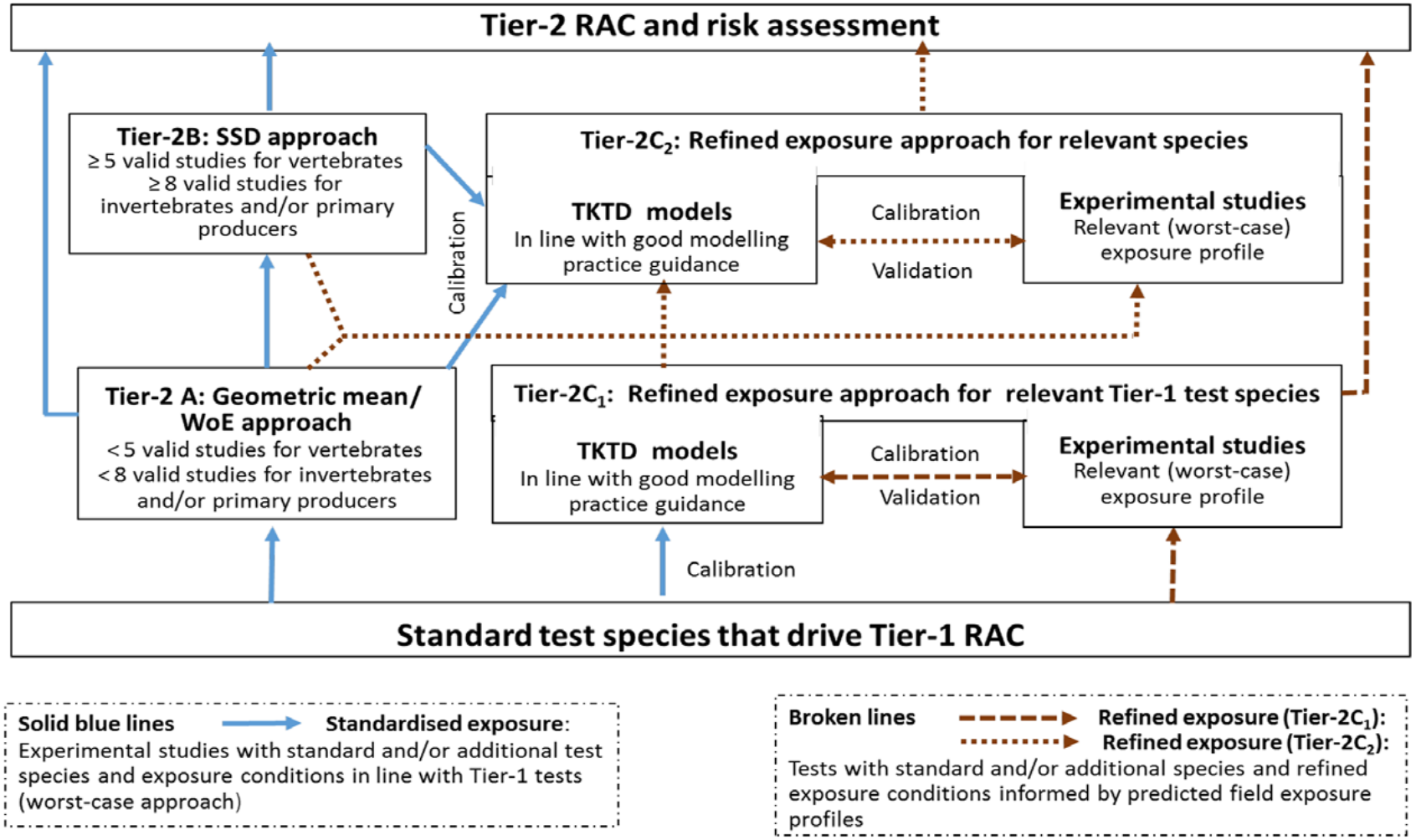
Requirements for validation data

- Effect data from experiments under time-variable exposure
- At least two exposure profiles with at least two pulses each, separated by no-exposure intervals of different duration length; defined based on DRT95
- Exposure specific dose–response curves are at least tested at three concentration levels (low, medium, high)





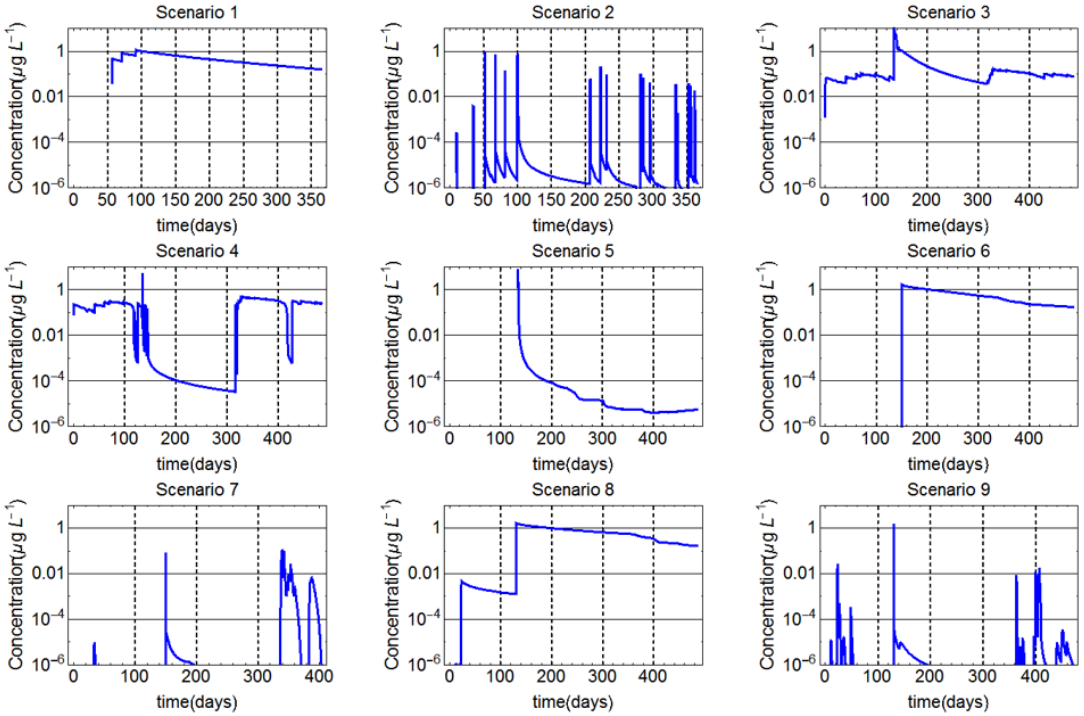
Linking GUTS with current risk assessment guidance



- Predictions of validated GUTS as alternative to experiments with refined exposure in Tier 2

Relevant endpoints

- Exposure profile specific effects on survival over time: final mortality
- In almost all cases, no effects of time-variable exposure scenarios as of exposure assessment in GUTS evaluations



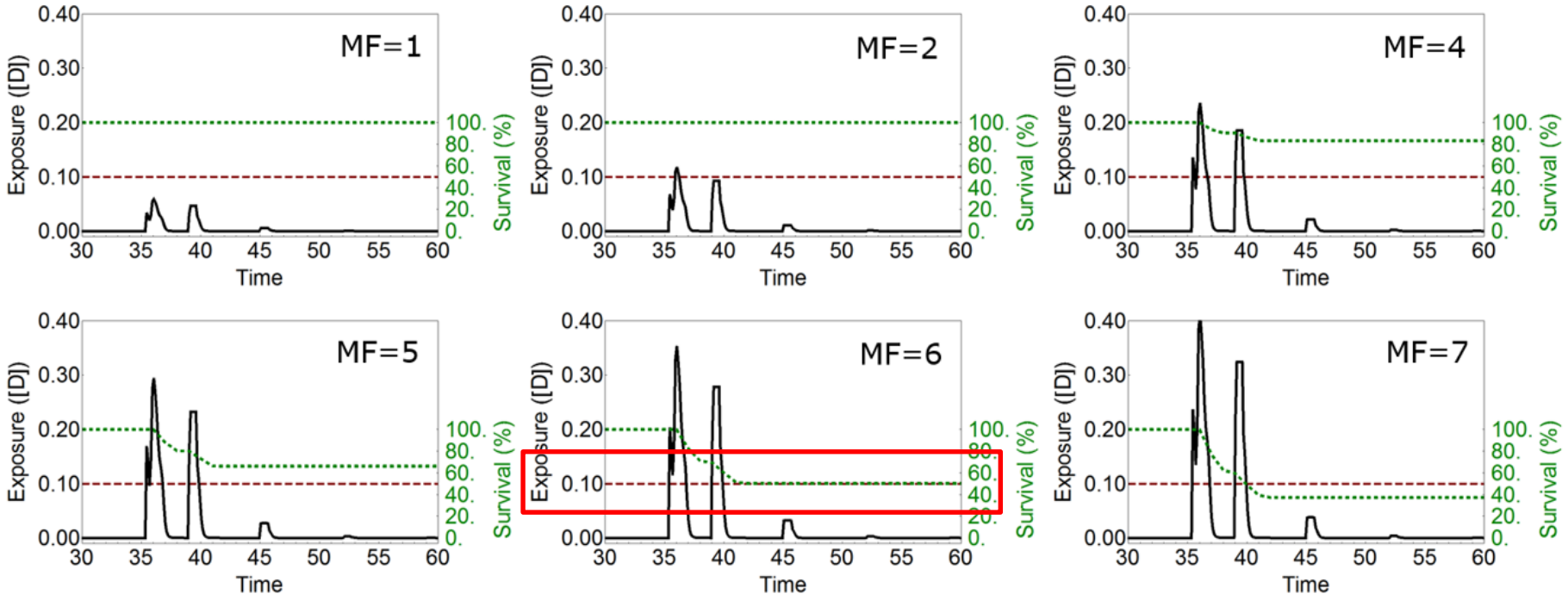
Exemplary concentrations over time for 9 exposure scenarios

No effects of propiconazole on survival of *Gammarus pulex* predicted by calibrated and validated GUTS

– End of the assessment?

Profile-specific multiplication factors

- How 'far' is the exposure profile from causing a defined effect?



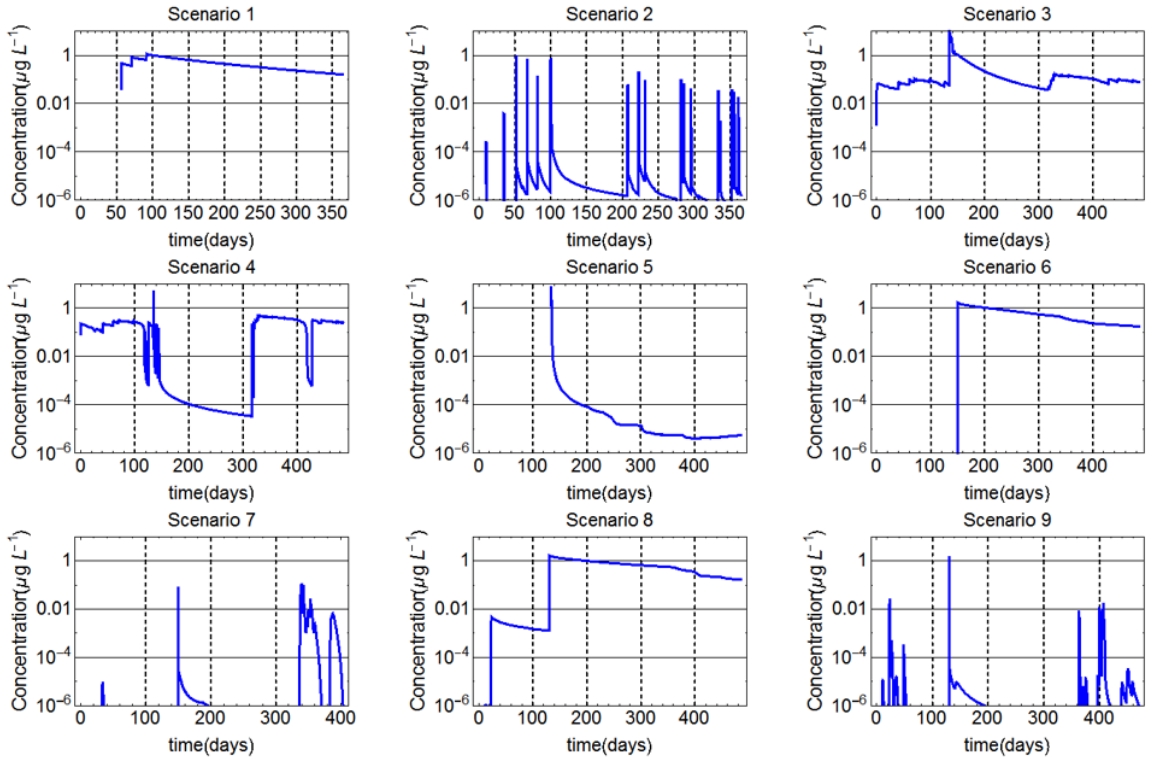
Profile-specific factor leading to 50% mortality at the end:
$LP_{50} = 6$

Multiplication of whole exposure times series with factors results in that factor leading to a certain effect level, e.g. 50%, at the end of the tested profile: LP_{50}.

Idea: "Margins of Safety" (Ashauer et al., 2013)

Application of the LPx concept

Lethal profile (LP_{50}) for mortality, Effect profile (EP_{50}) for immobility. Analogy to LC_{50} or EC_{50} of lab test under static exposure, is intended, but LC_X/EC_X are concentrations, while LP_X/EP_X are factors
→ can be compared with toxicity exposure ratios.



Scenario	1	2	3	4	5	6	7	8	9
Application	apple	apple	cereals	cereals	cereals	cereals	cereals	cereals	cereals
FOCUS SW scen.	R1 pond	R2 stream	D1 ditch	D1 stream	D3 ditch	D4 pond	D4 stream	D5 pond	D5 stream
PEC_{max}	1.130	2.007	10.564	8.063	9.083	1.668	2.268	1.670	2.401
'1st tier' TER	**17**	**10**	**2**	**2**	**2**	**12**	**8**	**11**	**8**
MF_{50} SD model	**17**	**44**	**3**	**20**	**12**	**12**	**205**	**12**	**195**
MF_{50} IT model	**17**	**49**	**3**	**24**	**16**	**12**	**250**	**12**	**237**

LP_{50} values differentiate risk between constant and highly variable exposure profiles!

Evaluation of GUTS applications

- Identification of application specific and unspecific elements in the GMP documentation (unspecific = can be documented as default)
- Clear criteria for the evaluation of specific documentation elements, checklist and examples for risk assessors

Annex A – Checklist for GUTS models

ASPECT OF THE MODEL TO BE EVALUATED BY THE RISK ASSESSOR – GUTS model application for lethal effects		Yes	No
1. Evaluation of the problem definition The problem definition needs to explain how the modelling fits into the risk assessment and how it can be used to address the specific protection goals. For GUTS, questions to be answered are likely to be those that are set out in Chapter 3. Nevertheless, the problem definition should make clear the following points:			
(a)	Is the regulatory context for the model application documented?		
(b)	Is the question that has to be answered by the model clearly formulated?		
(c)	Is the model output suitable to answer the formulated questions?		
(d)	Was the choice of the test species clearly described and justified, also considering all the available valid information (including literature)?		
(e)	Is the species to be modelled specified? – Is it clear whether the model is being used with a Tier-1 test species i.e. Tier-2C_1 or with one or more relevant species (which might include the Tier-1 species) i.e. Tier-2C_2?		
2. Evaluation of the quality of the supporting experimental data In this part of the evaluation, it is checked whether the experimental data with which the model is compared (both calibration and validation data sets) have been subjected to quality control. The focus is on the data quality, i.e. the laboratory conditions, set-up, chemical analytics and similar. Additional specific criteria for the suitability of the data sets for model calibration and validation are evaluated later in more detail (Sections 7 and 9 of this checklist).			
(a)	Has the quality of the data used been considered and documented? (see list of OECD test guidelines in Chapter 7, Table 6)		
(b)	Have all available data been used (either for calibration or for validation)? If not, is there a justification why some information has not been used?		
(c)	Is it checked whether the actual exposure profile in the study matches the intended profile in the test (+/– 20%); if not, are then measured concentrations used for the modelling, instead of nominal ones?		
3. Evaluation of the conceptual model Providing GUTS models are being used to address mortality/immobility effects in fish or invertebrates, the conceptual model will be suitable to address the specific protection goals; so, no further evaluation is required (see Chapters 2.1, 2.2 and 4.1).			
4. Evaluation of the formal model The formal model contains the equations and algorithms to be used in the model. For GUTS models, the equations are standardised, so that no further check is necessary (see Chapter 4.1.1). It has to be documented, however, which GUTS model version is used (e.g. full or reduced model).			
5. Evaluation of the computer model The formal model is converted into a model that can run on a computer (the computer model). For GUTS models, the computer model can be tested by showing the model performance for the GUTS ring-test data and performing some further checks (see Section 7.5).			
(a)	Is the used implementation of GUTS tested against the ring-test data set (see Section 4.2)?		
(b)	Were GUTS parameters estimated for the ring-test data and compared to the reference values, including confidence or credible intervals (Appendices B.6 and B.7)?		
(c)	Is a set of default scenarios (e.g. standard scenarios, extreme cases, see Section 4.1.2) simulated and checked?		
(d)	Are all data and parameters provided to allow an independent implementation of GUTS to be run?		

Recent examples of implementation of science into regulation: what makes the TKTD SO a good example?

- Interdisciplinary and open-minded working group
- Mature, nearly standardised GUTS theory and relatively simple formal model
- A certain number of application examples for GUTS and pesticides
- Definition of modelled endpoints which tie directly to the existing regulatory system (e.g. same assessment factors are suggested)
- Outlook: Calibration and fine-tuning of the suggested procedures for GUTS, e.g. practicability of the validation data suggestions, threshold values for model validation quality indicators

Outlook - DEBtox

- DEBtox models are promising and seen as very relevant for assessment of 'chronic'/sublethal effects under time-variable exposure
- DEBtox applications are still developed on a more case-by-case basis, there is no 'standard DEBtox' terminology nor formal model
- The number of application examples of DEBtox modelling for assessing pesticide effects in the literature is increasing, but still very low
- Due to the model complexity, model calibration is a rather demanding task which requires in-depth statistical knowledge
- Evaluation of any application is complicated by DEBtox models having a DEB component and a TKTD component – it is unclear who could be able and eligible to evaluate the DEB model part

Outlook – models for primary producers

- **Algae model**: model is simple, but flow-through experimental setup used for validation is not standardised, nor robustness check of the setup been ring-tested.
- The **Lemna model** can, when properly tested and documented, be used to evaluate effects of predicted exposure profiles in Tier-2C, if Tier-1 Lemna is the only standard test species triggering a potential risk
- Published **Myriophyllum** modelling approach may be a good basis to further develop TKTD models for rooted submerged macrophytes, needs further standardisation, documentation, calibration and validation.
- Growth models, particularly for Myriophyllum, would benefit from detailed experimental analysis of uptake, transport and elimination processes of organic contaminants.
- A modification of the standard Lemna and Myriophyllum experimental tests by including more frequent monitoring of growth and a recovery phase would provide valuable data for initial fits of plant models.

Acknowledgements

Dutch Minority for Economic Affairs for funding

Paul van den Brink, Mazhar Zafar, Mascha Rubach, Thomas Preuss, Roman Ashauer, Tjalling Jager for support by data, ideas, discussion....

All members of the EFSA working group WG on TKTD modelling

Thank you for your attention!

SETAC Europe 13th Special Science Symposium

Extrapolation of Effects Across Biological Levels:
Challenges to Implement Scientific Approaches in Regulation

23–24 October 2018 I Brussels, Belgium

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018