Dose-response modelling

Cleo Tebby
Alexandre Péry
Unit METO, Modelling in Ecotoxicology and Toxicology

INERIS

Dose-response modelling

Model one dose-response relationship

- Focus on the Hill model

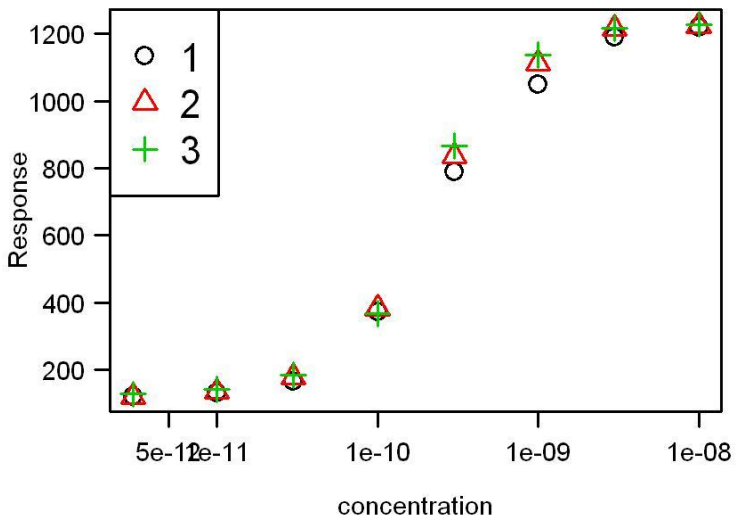
Model several dose-responses for comparisons

- Quantifying relative potency
- Non-constant potency
- Partial agonists

INERIS

Single dose-response curve modelling

Response for several concentrations
Replicates.



INERIS

The Hill model

Commonly used model based on the Hill equation in biochemistry : binding of a ligand to a macromolecule (2 parameters)

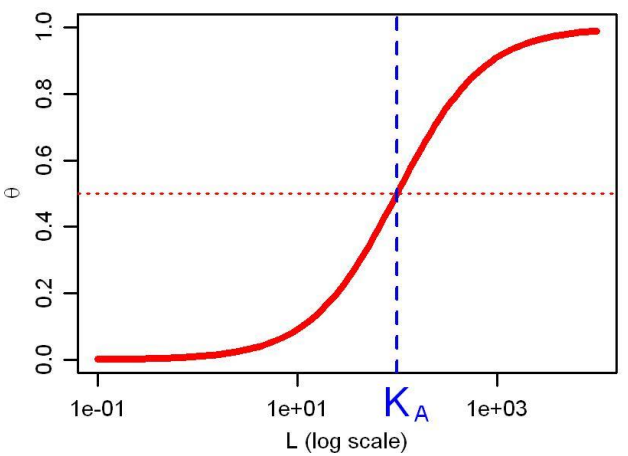
$$\theta = \frac{L^n}{K_A{}^n + L^n}$$

θ: fraction of occupied sites
L^n: free (unbound) ligand concentration
$K_A{}^n$: Ligand concentration producing half occupation,
= microscopic dissociation constant.

n represents the degree of cooperativeness of the ligand binding to the enzyme or receptor (n>1: positive cooperativity)

Bounded between 0 and 1.

Sigmoid on log scale:

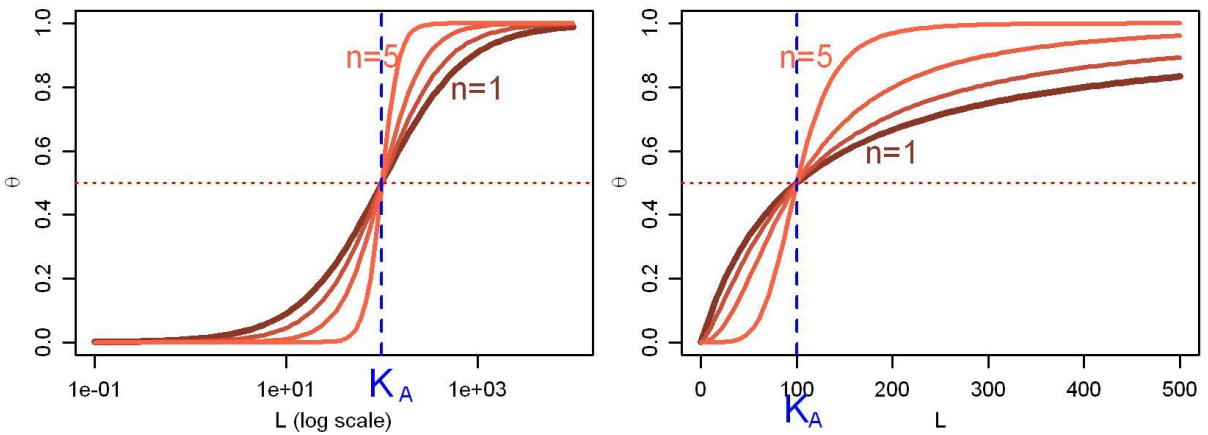


INERIS

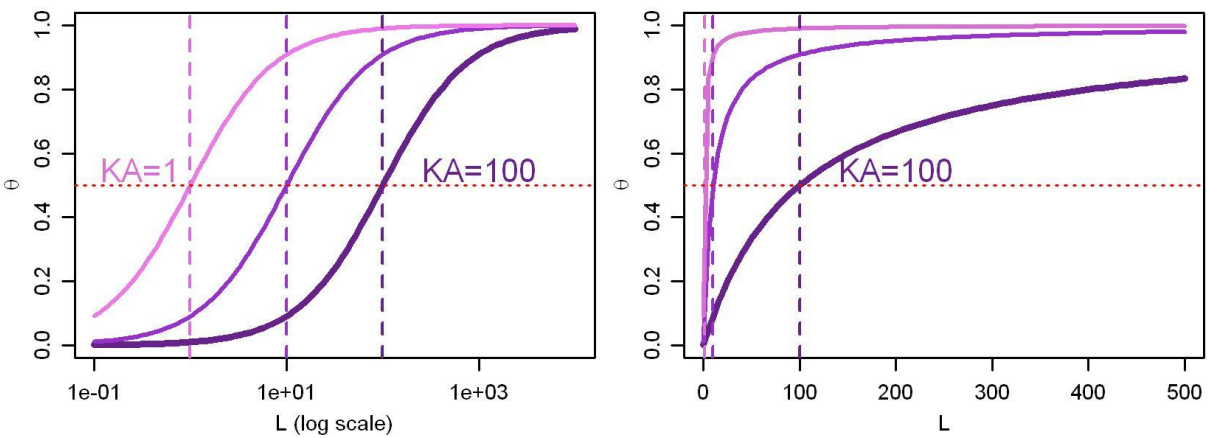
The Hill model

$$\theta = \frac{L^n}{K_A{}^n + L^n}$$

Variations on n (cooperativeness):



Variations on K_A



The Hill model

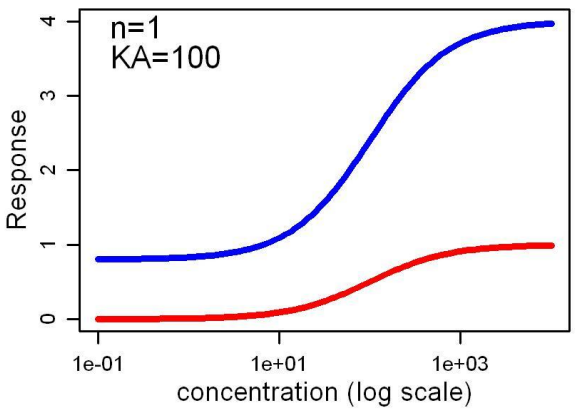
When modelling dose responses, 1 or 2 extra parameters are commonly added:
Maximum and optionnally Minimum.

Ka replaced by EC50

L replaced by nominal concentration c

n called the slope

$$\Phi(c) = Min + \frac{(Max - Min) \times c^n}{EC50^n + c^n}$$



Note: In drug-ligand modelling, often n=1 (fixed)

INERIS

The Hill model – data requirements

Need data that enables estimation of all parameters:

- background level should be visible
- maximum level (plateau): beware of cytotoxicity, receptor activation due to oxydative stress!
- slope : at least two intermediate expression levels

Geometrical progression (serial dilutions)

Need for calibration

Example: D-optimal designs.

Khinkis LA, Levasseur L, Faessel H, Greco WR. Optimal design for estimating parameters of the 4-parameter hill model. Nonlinearity in biology, toxicology, medicine 2003; 1: 363-77.

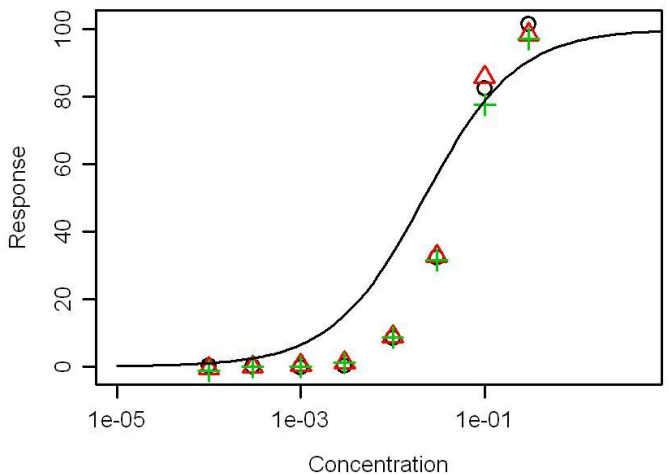
INERIS

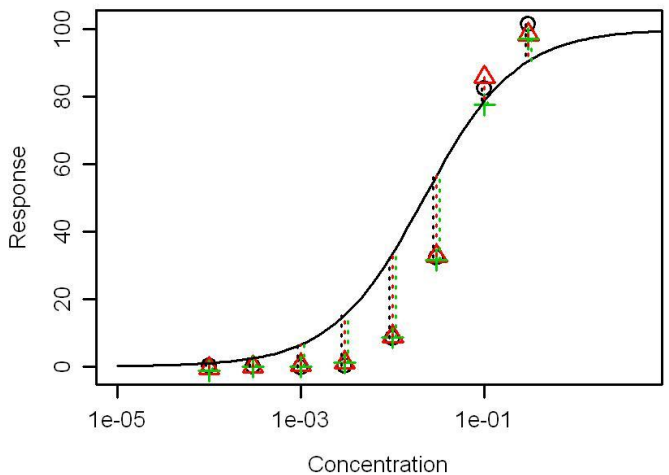
How do I fit a curve?

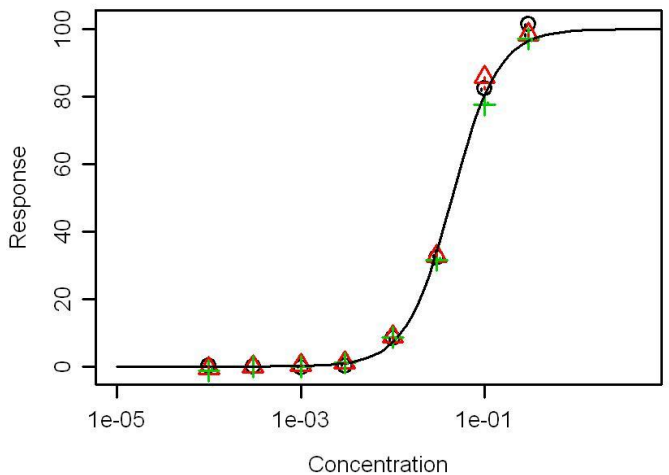
Select relevant model.

Determine combination of parameters that optimises a quality criteria: Minimise the residual sum of squares (Least Squares Regression)

non-linear: no matricial solution, need to use an optimisation algorithm.







INERIS

How do I know whether the model is acceptable?

First look at the data and the curve to check.

Use a statistical test to compare the fit with a different model, typically an analysis of variance. Main idea: a model with enough parameters will fit the data well, but might be unneccessarily complex.

Trade-off between model complexity and quality.

Lack-of-fit test: null hypothesis (H0) “the proposed statistical model fits well”

The F-ratio (F-statistic) **for nested models** takes into account the number of parameters and the quality criterion.

$$F = \frac{(RSS_2 - RSS_1) \times (n - p_1)}{(p_1 - p_2) \times RSS_1}$$

Compare with the F distribution

INERIS

How do I know whether the model is acceptable?

Example: 8 data points

Analysis of variance with 1 parameter for each data point = no particular curve shape.

At each data point, prediction= mean for that data point.

Hill model with 4 parameters = we expect a Hill shape.

fewer parameters -> the predictions of the data points will not be as good, but might be worth it.

Non-nested: need approximate F.

Compare with the F distribution

H0: Hill fits well.

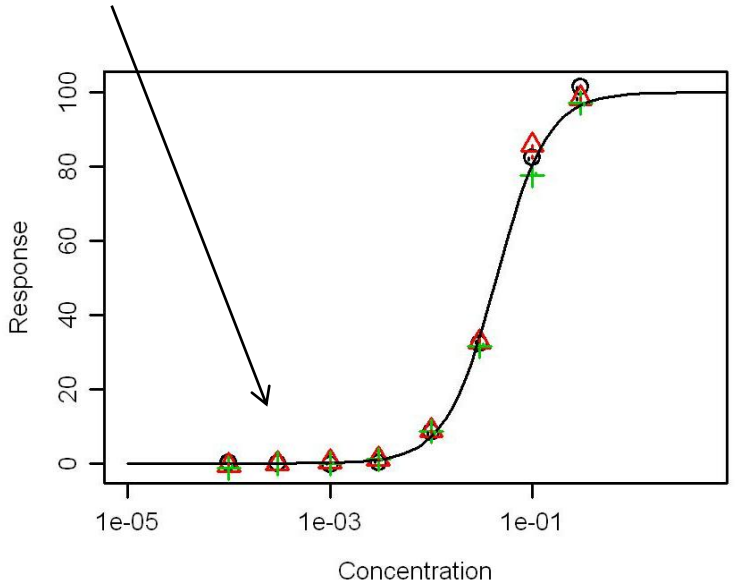
low p-value -> there is a significant difference in quality, although difference is adjusted for the number of parameters (Hill does not fit well)

high p-value->not worth the additionnal anova parameters (Hill fits well)

INERIS

How do I know whether the model is acceptable?

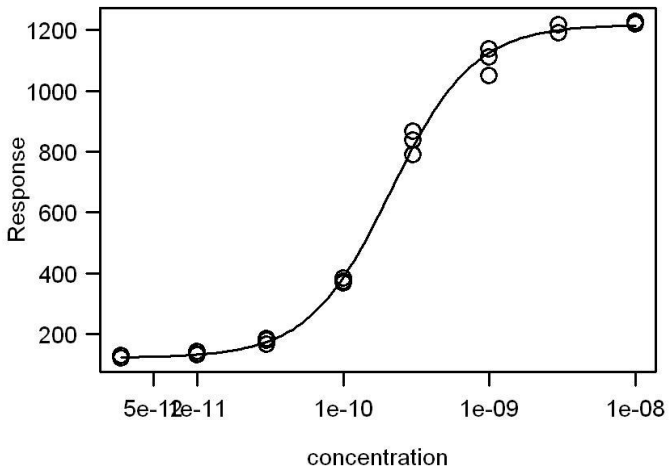
This might be cheating a bit, it is easy to fit a hill model through the background…

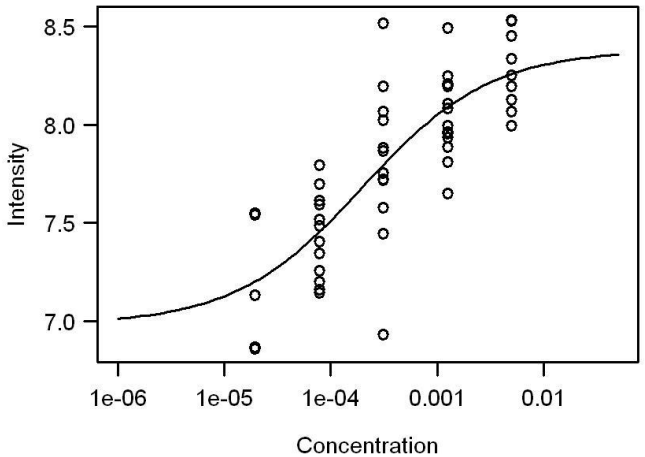


INERIS

Variability

Affects confidence intervals on parameters, acceptability of Hill model





Lack-of-fit test to compare with analysis of variance:

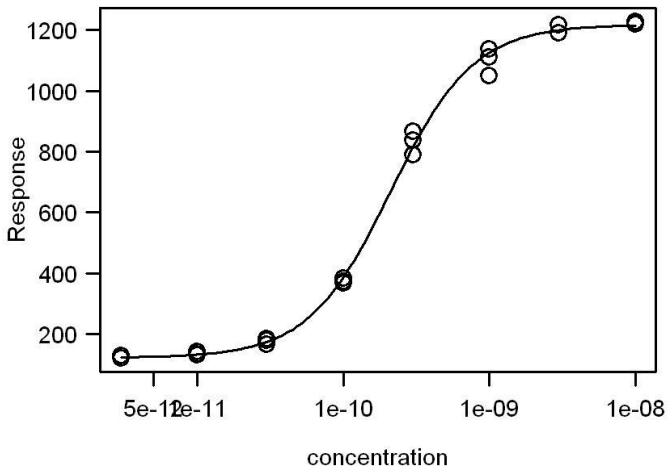
	ModelDf	RSS	Df	F value	p value
ANOVA	16	7971			
DRC model	20	11958	4	2.0004	0.143

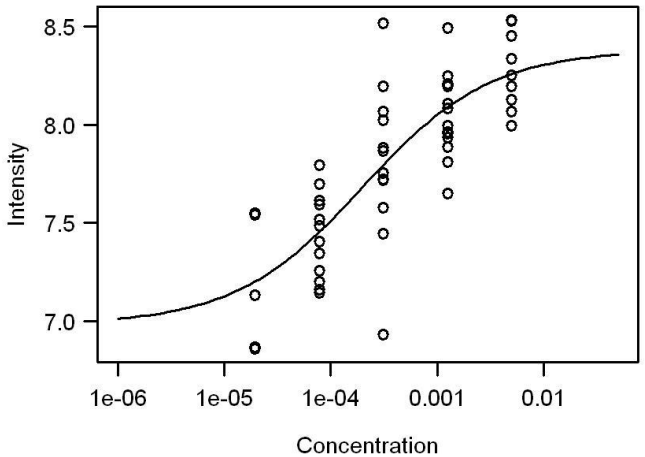
	ModelDf	RSS	Df	F value	p value
ANOVA	65	5.44			
DRC model	67	5.48	2	0.24	0.79

INERIS

Variability

Affects confidence intervals on parameters, acceptability of Hill model





Confidence intervals on parameters based on asymptotic normality, t distribution:

	Estimate	2.50%	97.50%
n	-1.54	-1.71	-1.37
min	123	102	143
max	1218	1195	1241
EC50	2.11E-10	1.96E-10	2.27E-10

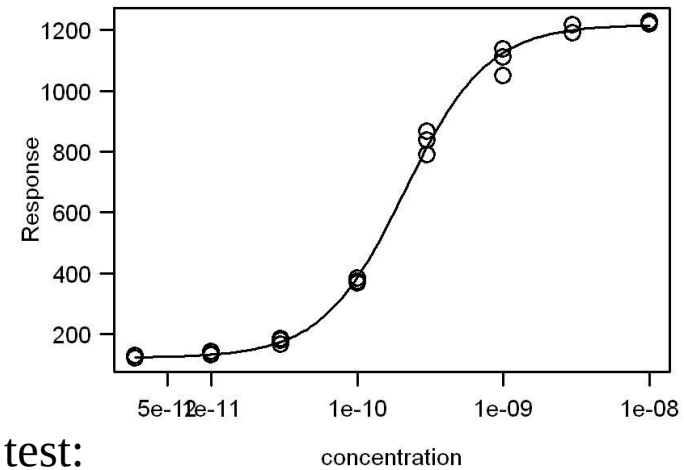
	Estimate	2.50%	97.50%
n	-0.72	-1.2	-0.24
min	7	6.8	7.1
max	8.4	8	8.8
EC50	2.0E-04	-2.9E-05	4.3E-04

CI on curves can also be calculated by bootstrap for pointwise estimation, functional CI could also be calculated – tricky business.

INERIS

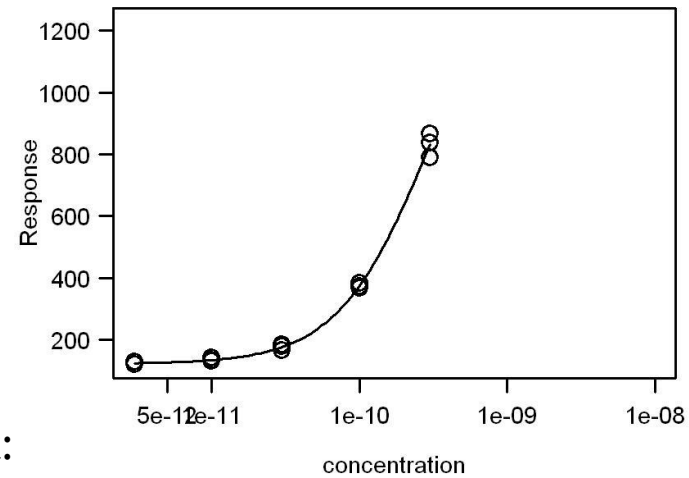
Lack of data

Affects confidence intervals on parameters, acceptability of Hill model



F test:
p=0.14

	Estimate	2.50%	97.50%
n	-1.54	-1.71	-1.37
min	123	102	143
max	1218	1195	1241
EC50	2.11E-10	1.96E-10	2.27E-10



F test:
p=0.84

	Estimate	2.50%	97.50%
n	-1.40	-2.04	-0.761
min	123	100	145
max	1530	292	2767
EC50	2.97E-10	-7.11E-10	6.65E-10

INERIS

Comparison of dose-response curves:

Normalisation

Relative potency

Toxic equivalents

INERIS

Raw data – normalisation - transformation

How do I make the data comparable?

- Position on plate effect?
- Plate effect?
- Block effect? batch of cells etc.
- Experimental effect? Who?
- Measurement effect? Machine wear? Delay?

Background: Subtract or divide (fold induction)?

Guidelines: subtract. But check fold induction for validity…

What does the background represent? Is it **not** related to the chemical.

- We want to get rid of it as we do with confounding factors in multi-way anova.

Estrogenicity assays: also an E2 control representative of the plateau.

- Need to fix the E2 plateau for all datasets.

INERIS

Raw data – normalisation - transformation

In receptor-ligand relationships, the Hill slope is characteristic of each receptor-ligand couple.

With the Hill model, aX+b transformations do not change the slope (cooperativity) or EC.

The background and E2 plateau can therefore be taken into account by subtractions or divisions without changing the characteristics of the dose-response relationships.

INERIS

Raw data – normalisation - transformation

If fold induction, use a log transformation (log-fold induction) to stabilize variance? Gennings C et al. Analysis of resulting data from estrogen receptor reporter gene assays. Journal of Agricultural Biological and Environmental Statistics 2003; 8: 84-104.

Log transformation changes slope and EC.

Note: Variance stabilization: large issue in large datasets like micro-array

Huber W et al. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 2002; 18 Suppl 1: S96-104.

Heteroscedasticity often overlooked in dose-response analysis?

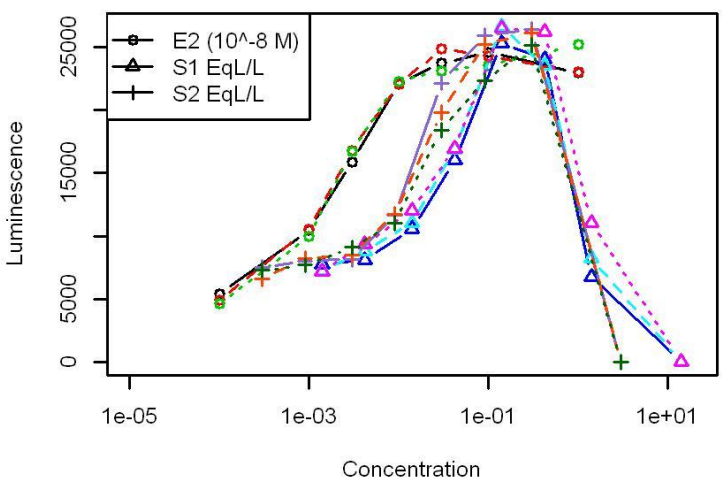
Scholze M et al. A general best-fit method for concentration-response curves and the estimation of low-effect concentrations. Environmental Toxicology and Chemistry 2001; 20: 448-457.

INERIS

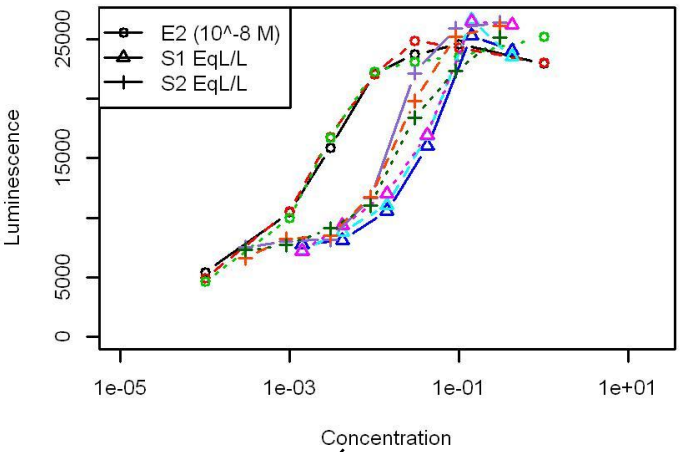
Raw data – normalisation - transformation

Example of normalisation:
Can also remove suspicious data points

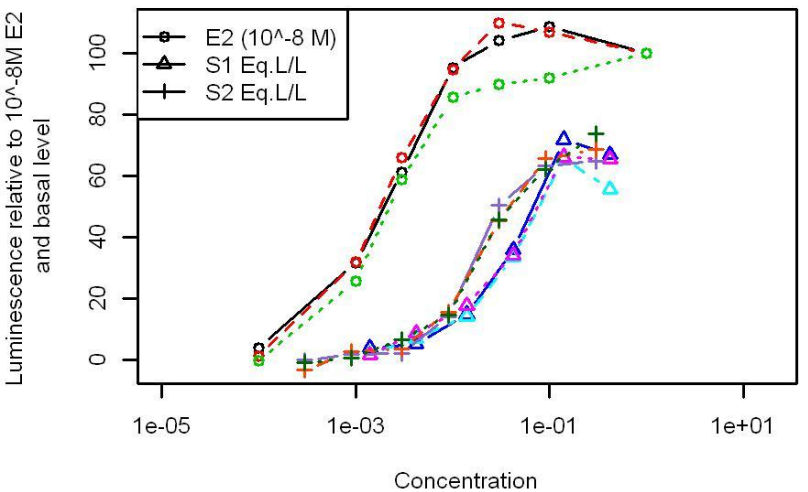
Raw data



Raw data



- background,
/E2 control



INERIS

Relative Potency

2 chemicals A and B

if equal slope, min, and max, parallel dose-response curves on log-scale, only difference is EC50.

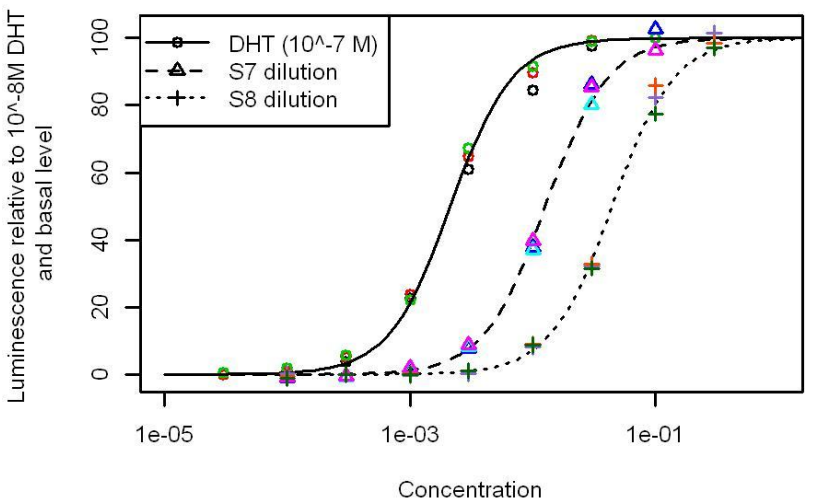
EC50 ratio

- quantifies distance between curves
- used to calculate equivalencies

$$C_A = \frac{EC_{50\,A} C_B}{EC_{50\,B}}$$

Example in hazard assessment:

Toxic equivalency factor (TEF) expresses the toxicity of dioxins, furans and PCBs relative to 2,3,7,8-TCDD



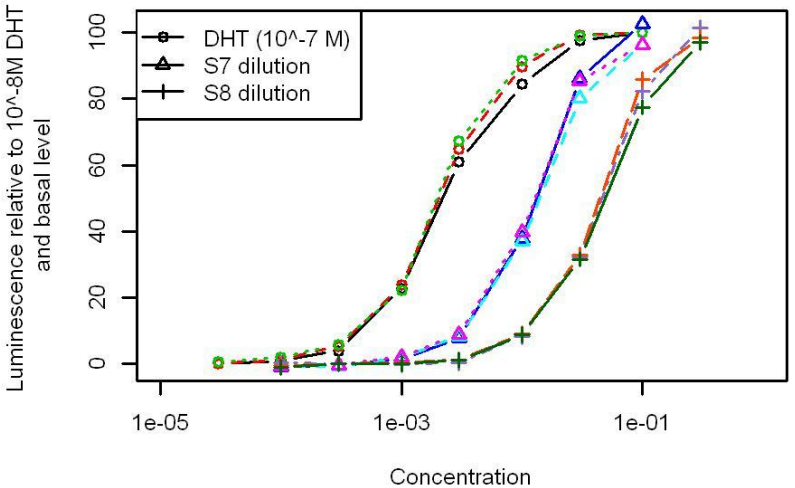
Basis for concentration addition in mixtures

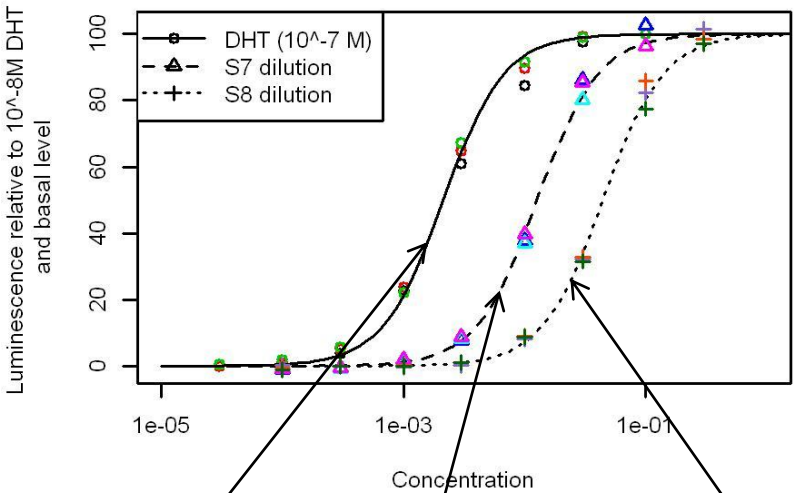
(2 components at half their EC50-> EC50 of mixture)

INERIS

Examples of modelled curves with same slopes, max and min

Common parameters determined by optimisation. Min and Max can be fixed. F-test.





$EC50_1$=0.0021 (reference), $EC50_2$=0.013 , $EC50_3$=0.044

$TEF_2=EC50_1/EC50_2=0.2=1/5$

$TEF_3=EC50_1/EC50_3=0.05=1/20$

Reference

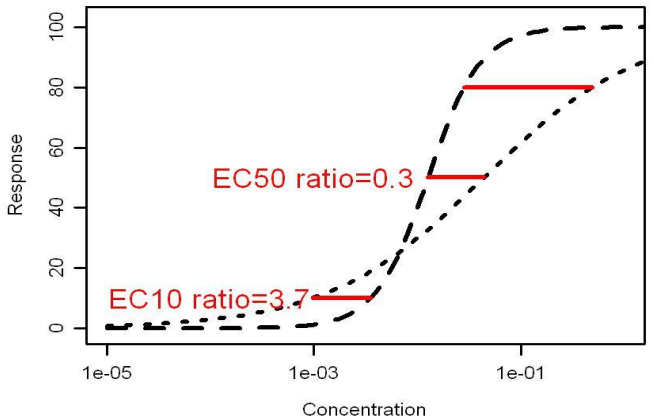
20 times less toxic than reference

5 times less toxic than reference

INERIS

Non constant relative potency: different slopes

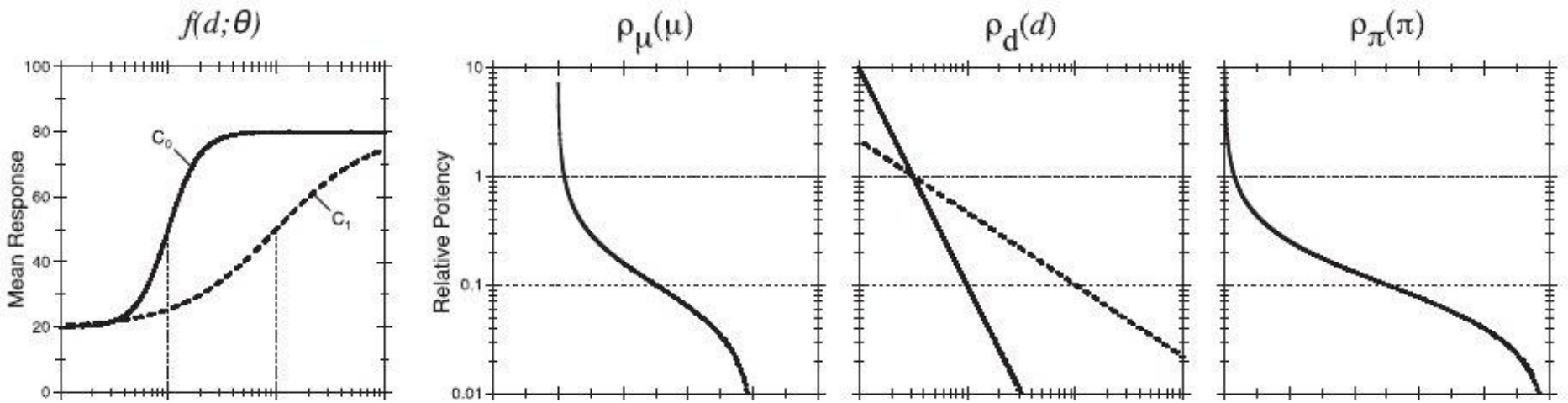
Sometimes common slope does not fit.
EC50 ratio =/= EC10 ratio



Validity of TEFs?

Dinse GE, Umbach DM. Characterizing non-constant relative potency. Regulatory Toxicology and Pharmacology 2011; 60: 342-353.

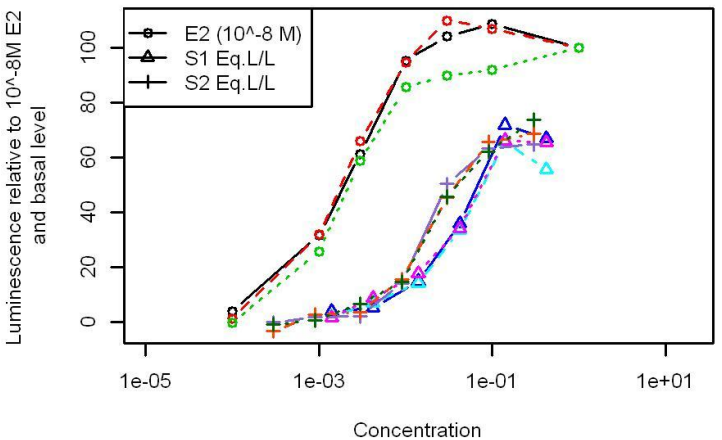
Relative potency functions: Mean response, dose equivalence, proportion of response

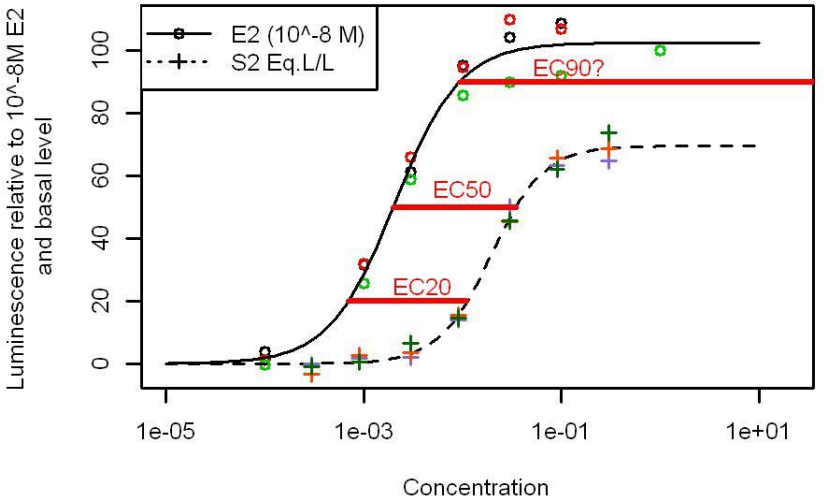


INERIS

Modelled curves with different max

Relative potency as a function of mean response:
Model each curve separately.





limited domain of applicability

There is no way to use a Hill model with different max and have constant relative potency (no use fitting same slope).

EC20/E2 relative potency: 6.3e-10M

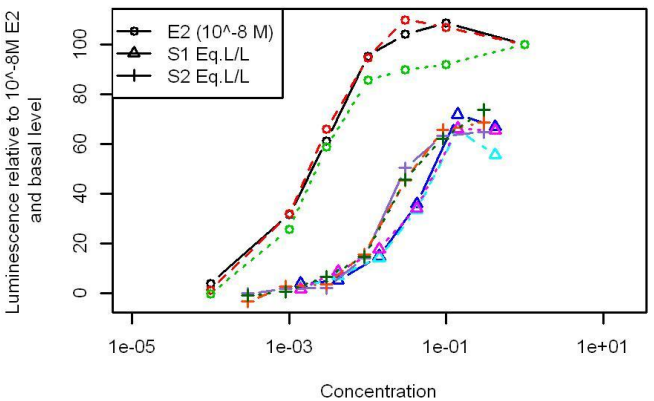
Could may be ignore data points near the plateau and set any identical max? Better not

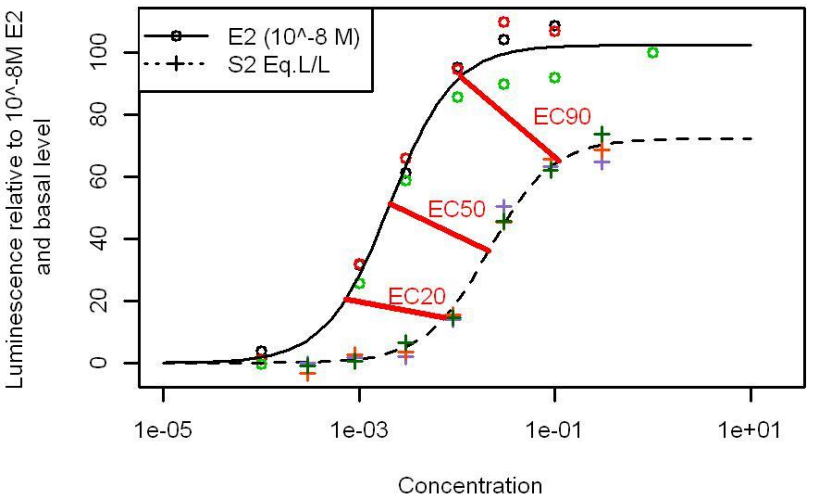
INERIS

Modelled curves with different max

Relative potency as a function of response quantile:

If the curves can be modelled with a same slope, constant relative potency.





Slope looks different but slope parameter is actually the same

Relative potency: 9.8e-10, for any response quantile

loss of information about partial agonism

INERIS

Modelled curves with different max

Relative potency as slopes at low dose.

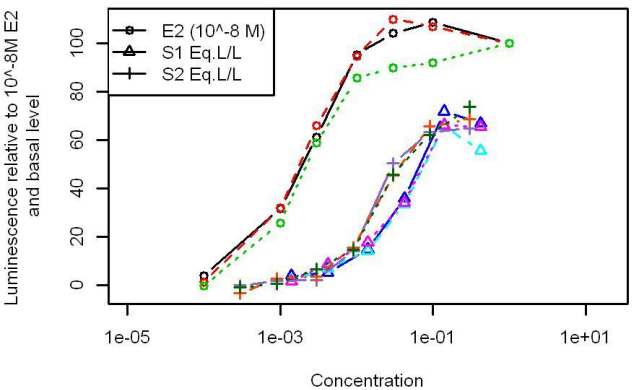
When c<<EC50, Hill can be approximated by

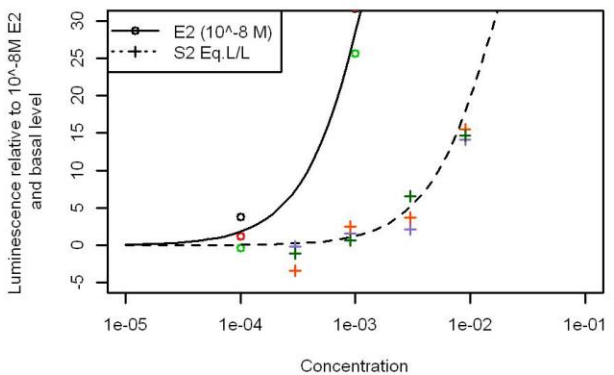
$$\Phi(c) = \frac{Max}{EC50^n} \times c^n$$

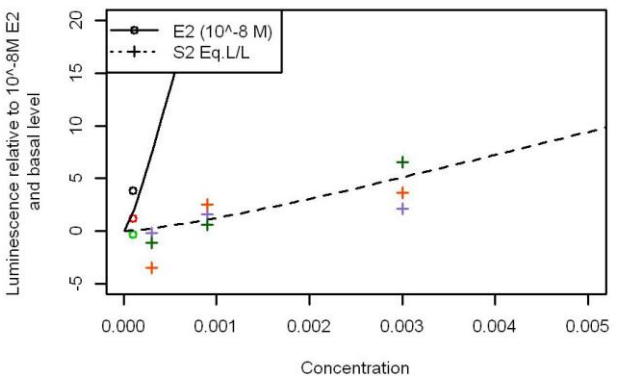
If identical slope parameter, response can be compared with the ratio of

$$\sqrt[n]{\frac{Max}{EC50^n}}$$

Uses the whole of the dose response rather than truncating data.







Relative potency: 7.6e-10 at low doses

See also Audebert M, Zeman F, Beaudoin R, Pery A, Cravedi JP. Comparative potency approach based on H2AX assay for estimating the **genotoxicity** of polycyclic aromatic hydrocarbons. Toxicology and applied pharmacology **2012**; 260: 58-64.

INERIS

Modelled curves with different max

No universal solution!
Use approximate solutions, confidence intervals?

Interpretation of partial agonism?
Interpretation of sub-or supramaximal responses in other types of experiments? Genotoxicity, in vivo…

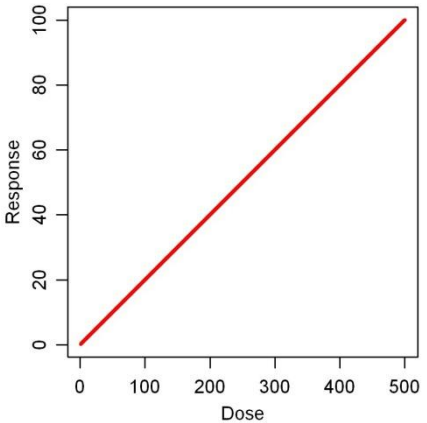
Also a problem in predicting responses to mixtures of partial and full agonists.

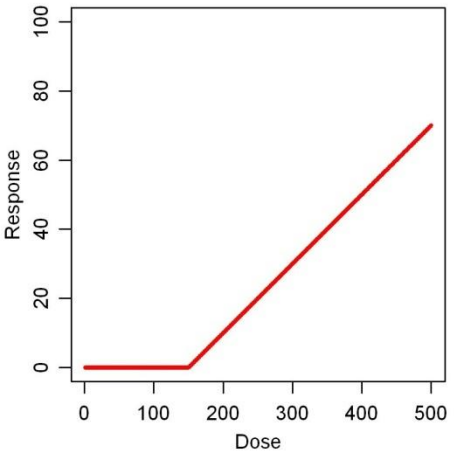
INERIS

Other dose-response models (1)

Common models:

Linear model



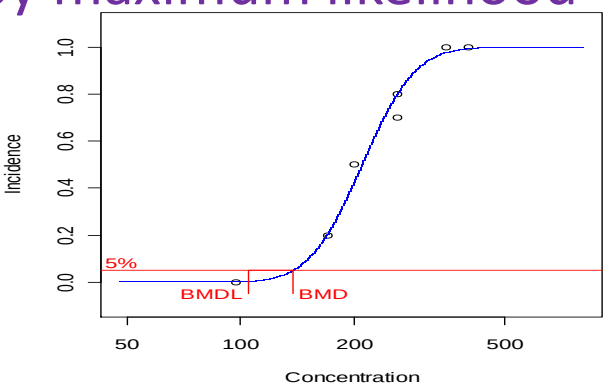


Hockey-stick model (threshold)

…

Probit model (binomial data e.g. survival rate in groups exposed to different concentrations or doses), fit by maximum likelihood

$$\varphi(\alpha + \beta \ln(dose)) \quad \sim N(0;1)$$



INERIS

Other dose-response models (2)

Logistic model (also a probabilistic model)

$$P(x) = \frac{1}{1 + \exp\left[-(a + bx)\right]}, b \geq 0$$

= log-logistic model if using ln(dose).

The 4-parameter (and smaller) log-logistic model is also called the…

Hill model.

The 5-parameter log-logistic model is asymmetric.

Other sigmoïdal variants and boxcox variants, see:

Scholze M, Boedeker W, Faust M, Backhaus T, Altenburger R, Grimme LH. A general best-fit method for concentration-response curves and the estimation of low-effect concentrations. Environmental Toxicology and Chemistry 2001; 20: 448-457.

INERIS

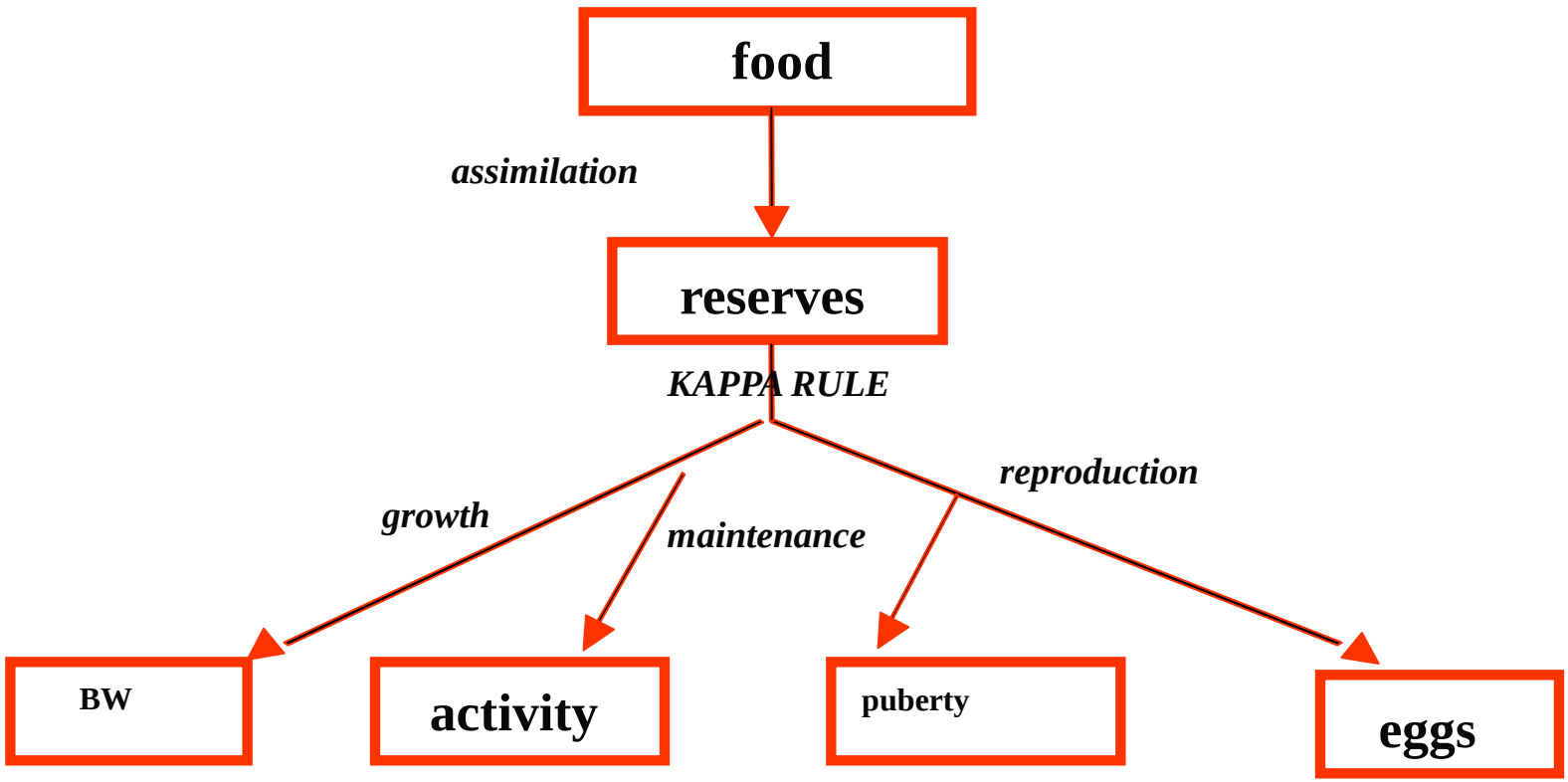
Energy-based modelling to analyze ecotoxicity data

- Mathematical models exist to analyze effects on endpoints such as reproduction and growth (for instance, the 21d daphnids test.
- These models are TK/TD (toxicokinetics/toxicodynamics) models because they relate effects to internal concentration. This internal concentration is deduced from exposure concentration through a kinetics model.
- The most known ones are the DEBtox models. They permit to estimate toxicity parameters that do not depend on the duration of exposure, to extrapolate to time-varying concentrations and to get first insides relative to mode of action.
- Using these models requires specific training.

INERIS

The DEB theory

DEBtox models are based on the DEB theory (Kooijman, 2000).



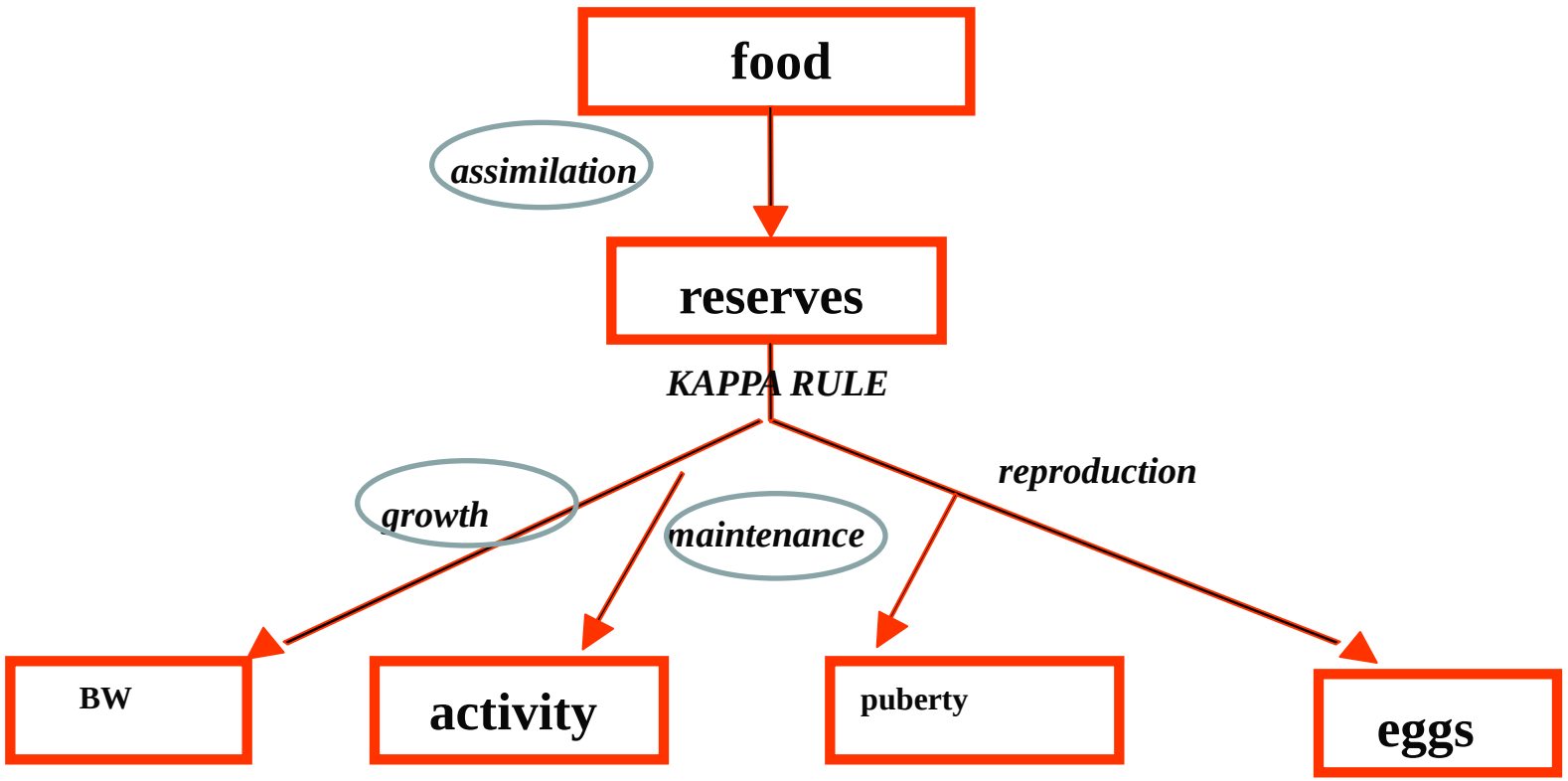
Somatic uses

Reproductive uses

INERIS

Effects on growth

DEBtox models are based on the DEB theory (Kooijman, 2000).

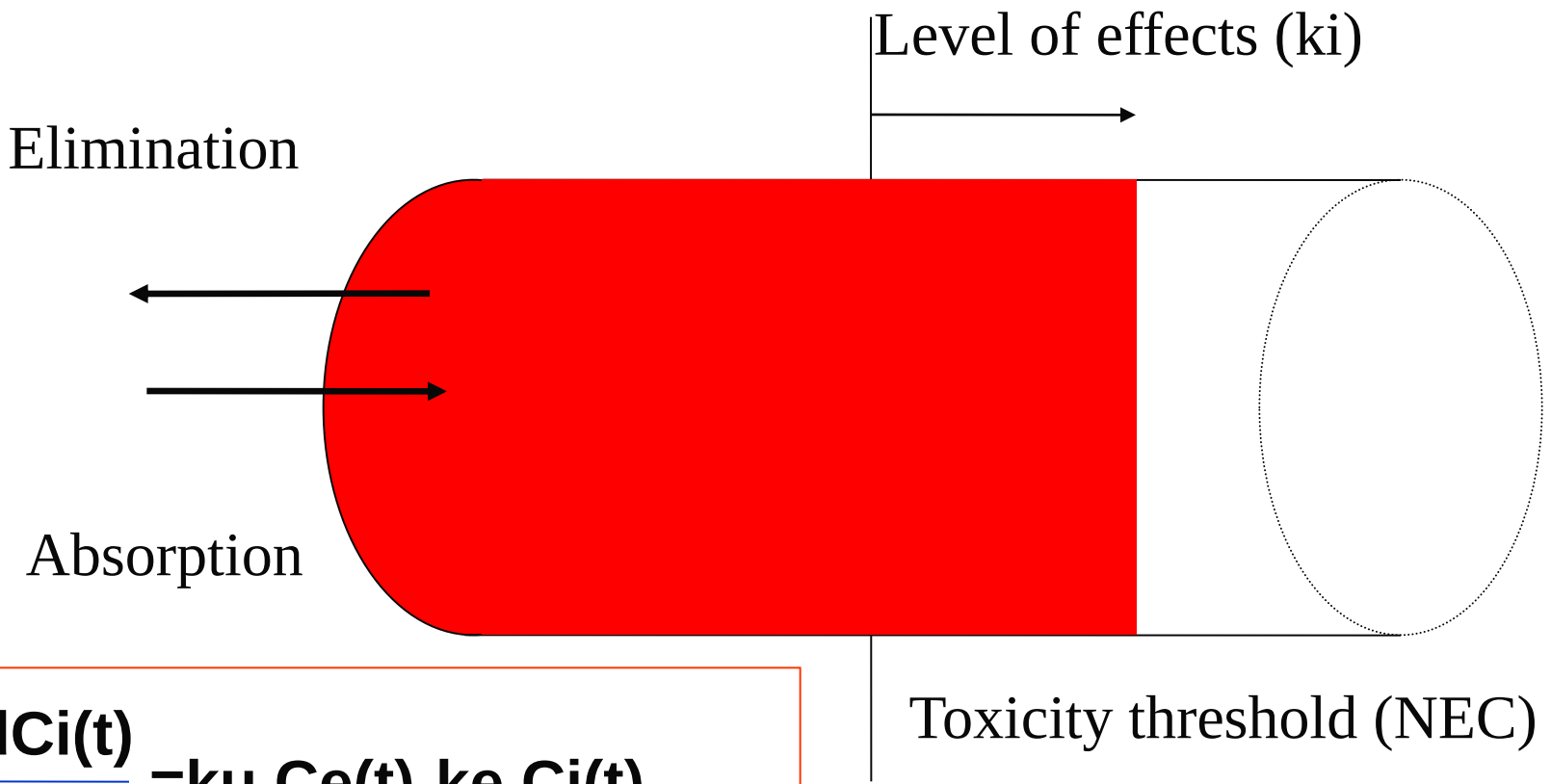


Somatic uses

Reproductive uses

INERIS

A simple kinetics models to relate exposure and internal concentrations



$$\frac{dCi(t)}{dt} = ku\ Ce(t) - ke\ Ci(t)$$

INERIS

DEBtox models to analyze effects on growth

$$L(t) = L_{max} - (L_{max} - L(0))\exp(-\gamma t)$$

Model	L_{max}	γ
Assimilation	1-ki(ci-NEC)	
Growth		2/(2+ki(ci-NEC))
Maintenance	1/(1+ki(ci-NEC))	1-ki(ci-NEC)

INERIS

Illustration with daphnids exposed to uranium

Massarin S, Beaudouin R, Zeman F, Floriani M, Gilbin R, Alonzo F, Péry ARR. 2011. Biology-Based Modeling To Analyze Uranium Toxicity Data on Daphnia magna in a Multigeneration Study. *Environmental Science and Technology* 45, 4151–4158.

DEBtox analysis showed effects on assimilation and a rapid kinetics (effects occur before significant accumulation)

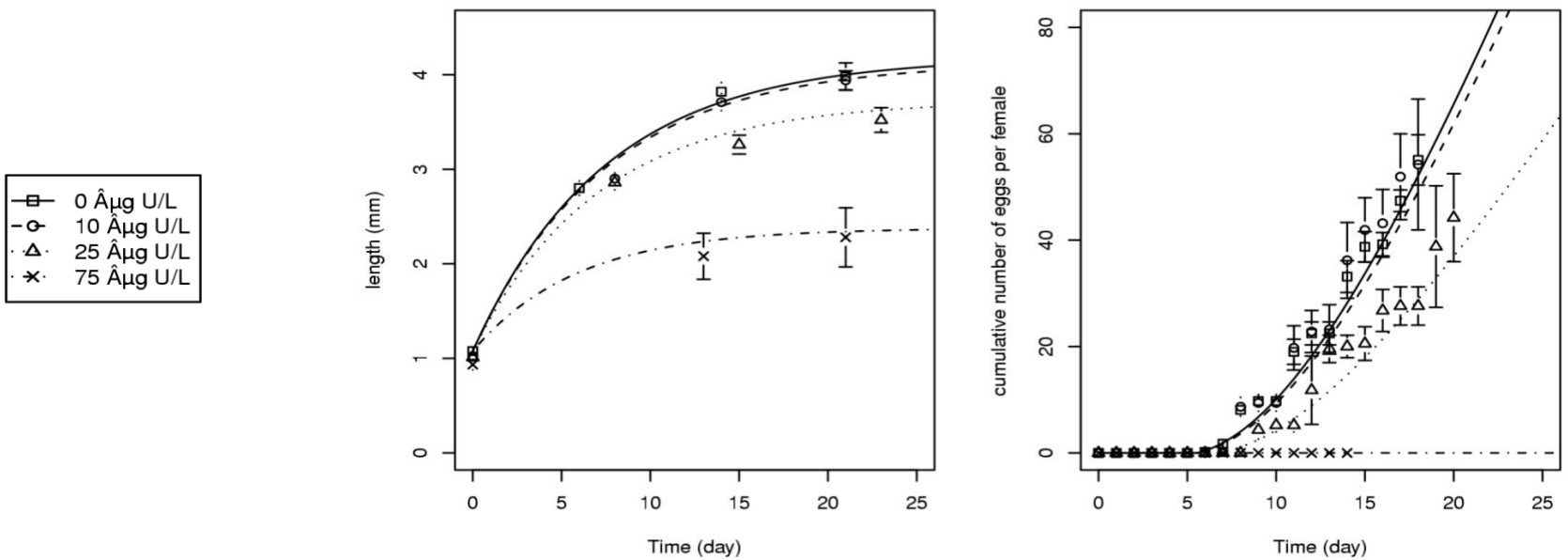


Illustration with daphnids exposed to uranium

Impairment of the digestive process. In order to confirm the effect of uranium on ingestion processes, we investigated the cellular effects of uranium in the digestive tract of *D. magna*, specifically in the midgut, which is involved in nutrient absorption and enzyme secretion in cladocerans.