Modeling microbial dynamics in food processes

An experiment design approach to predictive microbiology

Eva Van Derlinden, Jan F. Van Impe

BioTeC – Chemical and Biochemical Process Technology and Control

Department of Chemical Engineering, KU Leuven, Belgium

Outline

1. **Predictive microbiology**
2. **(Predictive) model building cycle**
3. **Kinetic models**
   - *Optimal Experiment Design (OED)*
     - <u>Case study 1</u>: effect of T on μ_{max}
   - *Design Of Experiments (DOE)*
     - <u>Case study 2</u>: effect of several factors on μ_{max}

1. **Probabilistic models**
   - *Data collection*
   - *Model building*

KATHOLIEKE UNIVERSITEIT LEUVEN

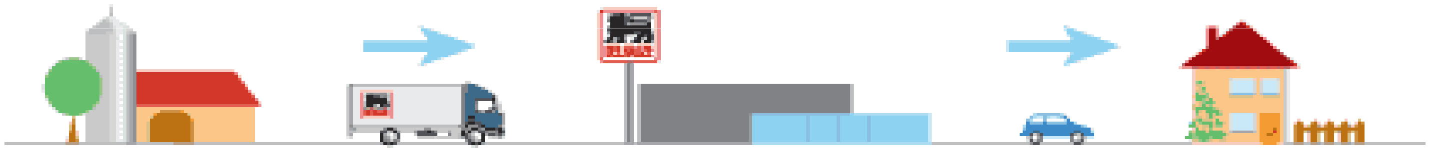
BioTeC

AGRO-INDUSTRIE et MÉTHODES STATISTIQUES

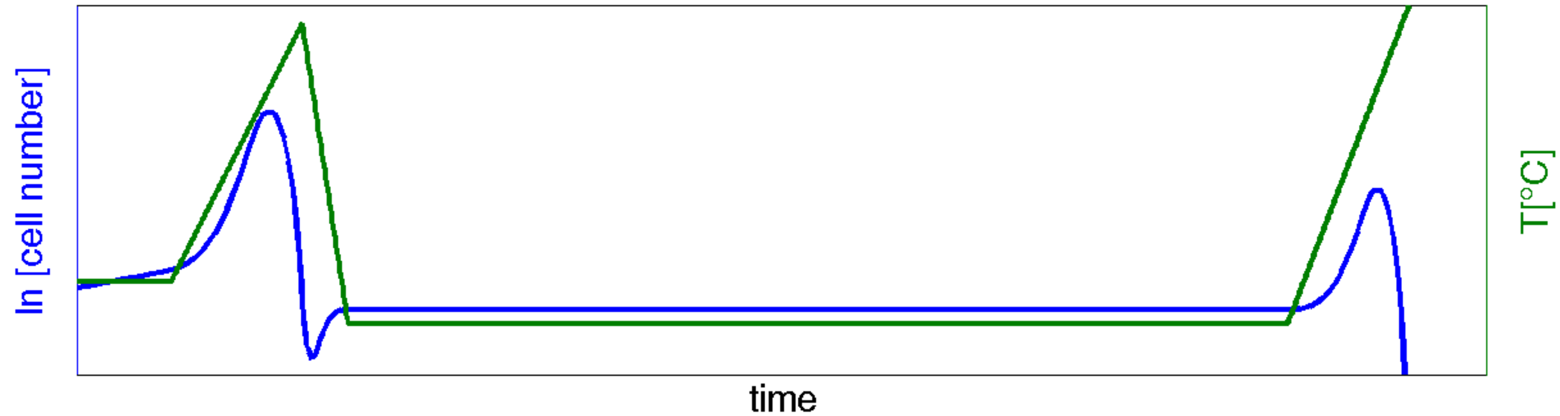
Predictive microbiology

Microbial food safety and quality	➡	Reduction and/or prevention of microbial growth

Microbial dynamics are determined by the environmental factors. ⇒ temperature



processing transportation & storage consumption



Predictive microbiology

Predictive microbiology

⇒ mathematical modelling of **microbial dynamics** (growth, survival, inactivation, growth/no growth) under **changing environmental conditions** (e.g., temperature, pH, a_w)

⇒ predict microbial dynamics in **foods**

⇒ control food safety & food quality

Specific use

- **food process industry** (e.g. formulation of new products)
- **regulatory bodies** (e.g. storage guidelines, food handling)
- **risk assessment** in the food trade

Predictive microbiology

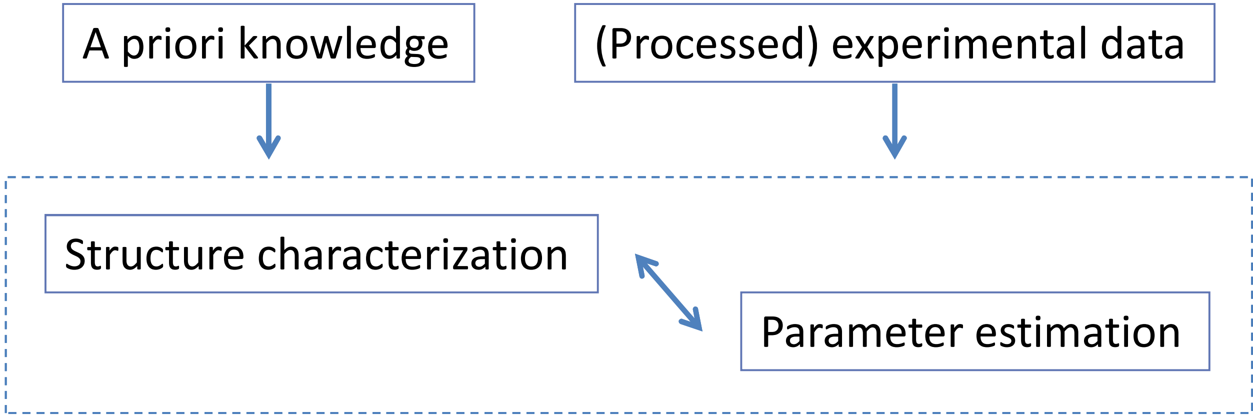
The model building cycle

A priori knowledge ↓

(Processed) experimental data ↓

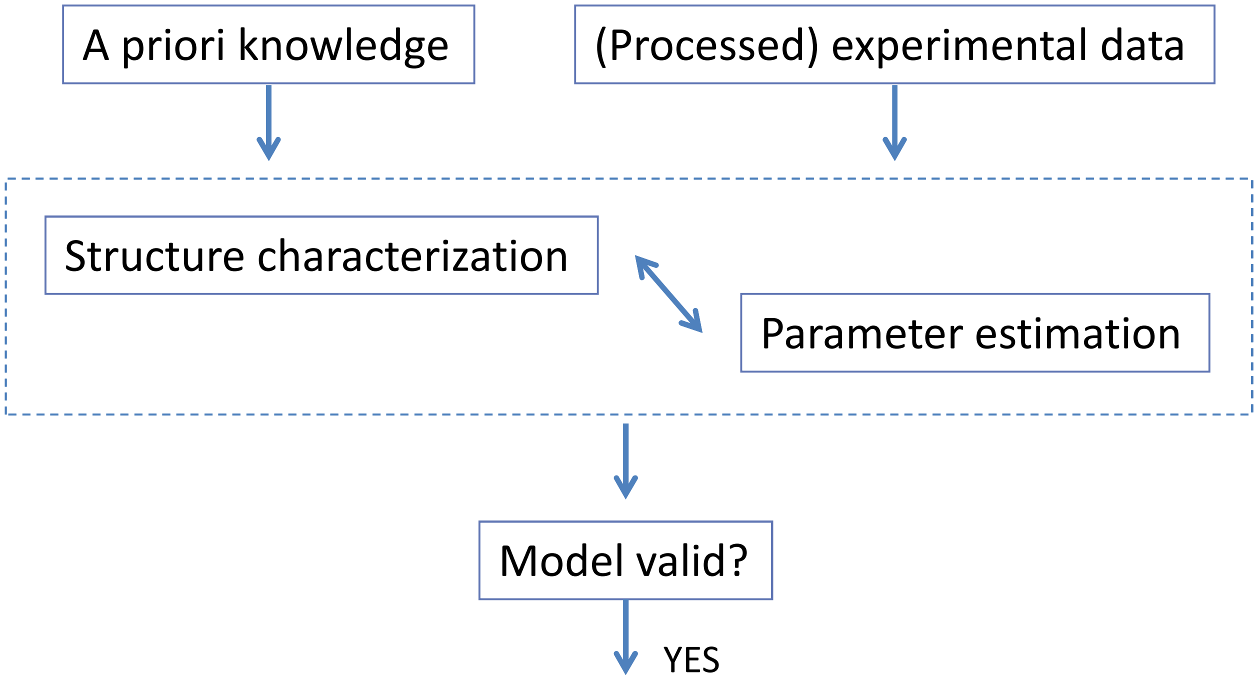
Predictive microbiology

The model building cycle



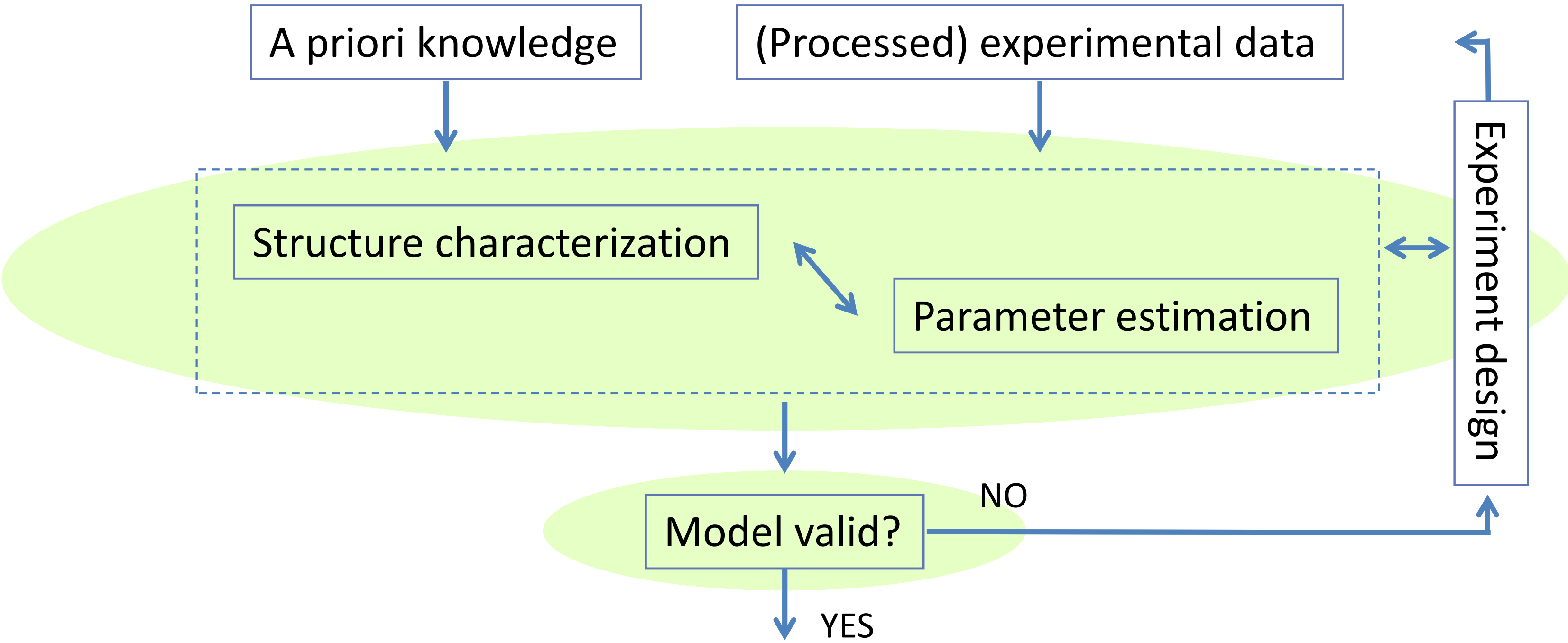
Predictive microbiology

The model building cycle



Predictive microbiology

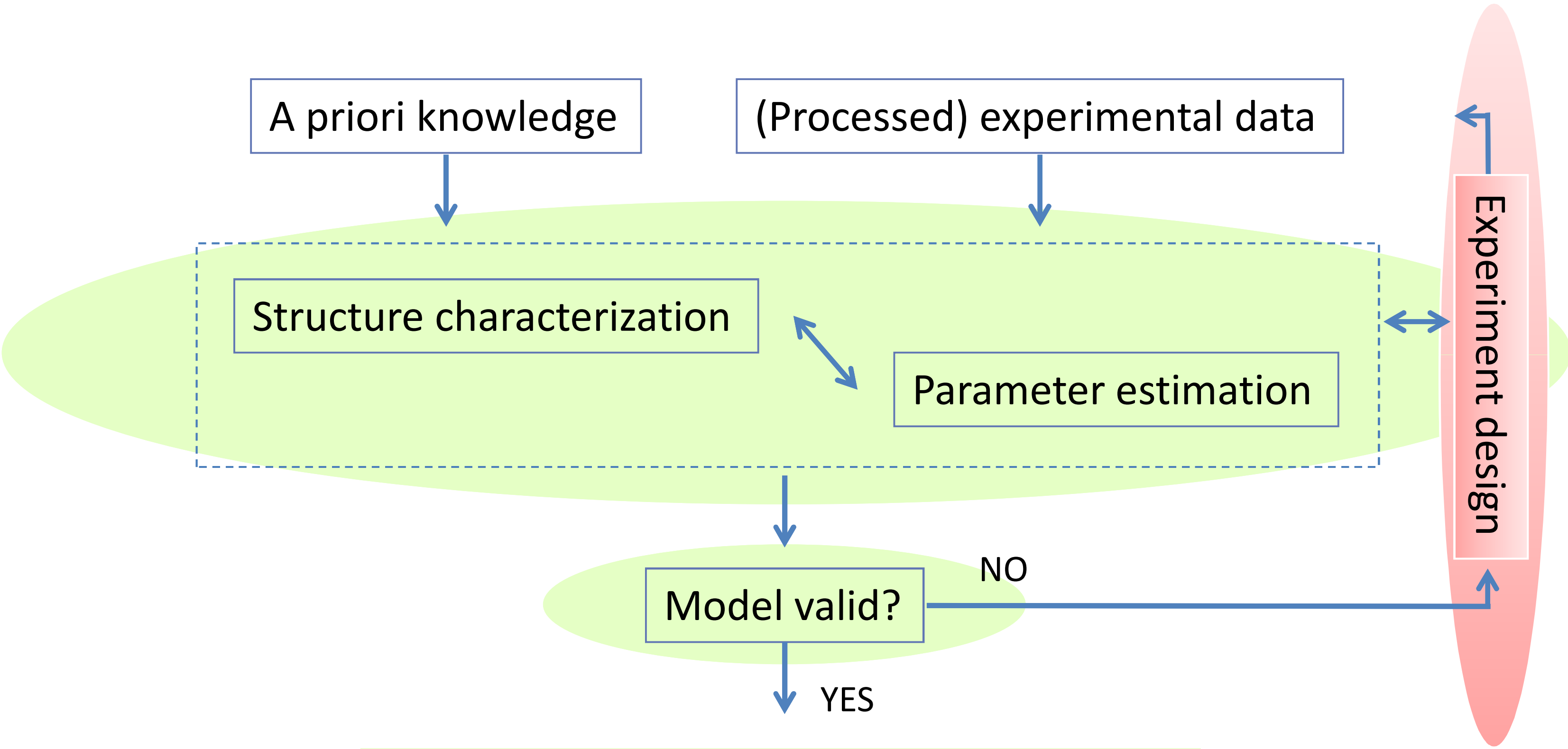
The model building cycle



=> Model building is a data-driven process

Predictive microbiology

The model building cycle



=> Model building is a data-driven process

=> Well-founded design improves model quality and validity

Predictive microbiology

Importance of data acquisition

This presentation will outline the importance of *adequate* data collection when building/improving predictive models for food processes.

1. **Kinetic models**
   - ➢ Techniques to improve parameter estimation accuracy
     - Optimal experiment design
     - Design of Experiments

2. **Probabilistic models**
   - ➢ Guidelines for model building

Kinetic models

Kinetic models

Definition these models quantify the effect of one or multiple intrinsic or extrinsic factors on the microbial evolution rate

- Inactivation during processing
- Growth during storage

Kinetic models

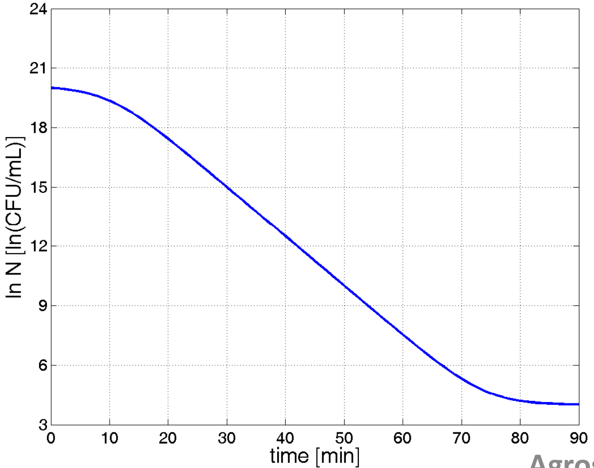
Growth & inactivation

General expression for microbial evolution

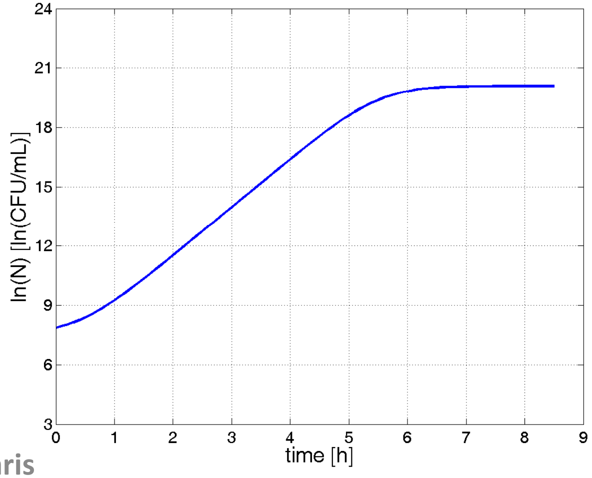
$$\frac{dN(t)}{dt} = \mu \cdot N(t)$$

Primary model

$\mu < 0$



$\mu > 0$



Kinetic models

Growth & inactivation

General expression for microbial evolution

$$\frac{dN(t)}{dt} = \mu(<env>, <phys>, S, P, \ldots) \cdot N(t)$$

$\mu < 0$ $\mu > 0$

Secondary model

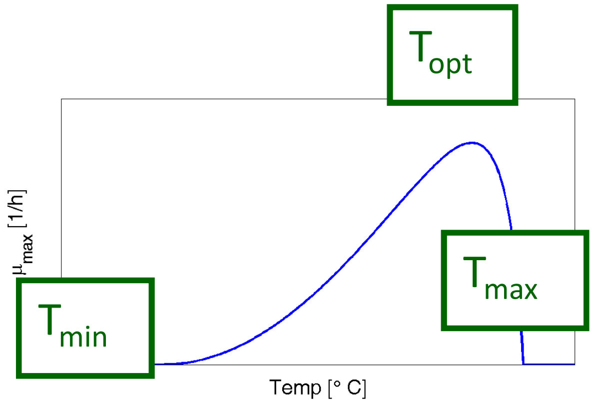
models to describe the effect of the influencing factors as a function of time

Kinetic growth rate models

Case study 1. Optimal Experiment Design (OED)

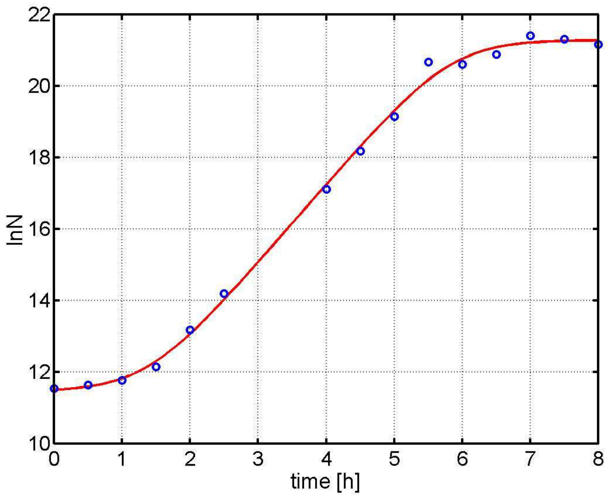
- Parameter estimation
- Model structure selection

<u>Kinetic model</u>: effect of temperature on μmax



- Different areas of interest
  - Fermentation processes
  - Soil microbiology: compost processes
  - Microbial food safety

Microbial dynamics & temperature



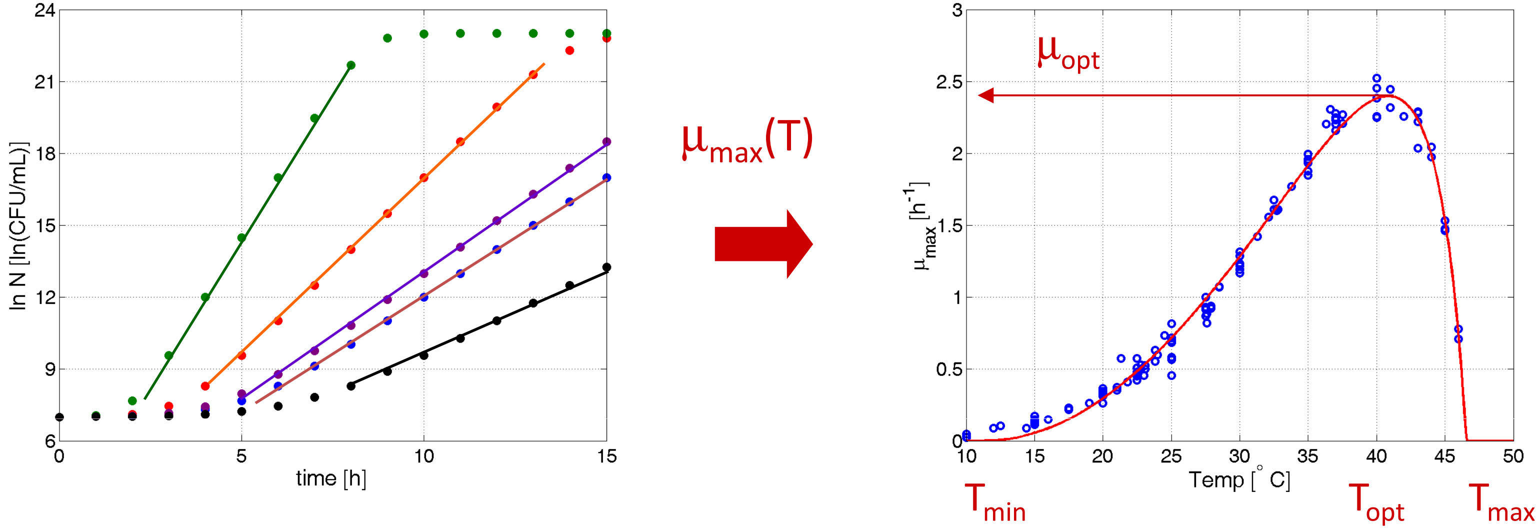
Primary growth model

Baranyi and Roberts (1994)

$$\frac{dN(t)}{dt} = \boxed{\frac{Q(t)}{Q(t)+1}} \cdot \boxed{\mu_{\max}} \cdot \boxed{\left(1-\frac{N(t)}{N_{\max}}\right)} \cdot N(t)$$

$$\frac{dQ(t)}{dt} = \mu_{\max} \cdot Q(t)$$

Microbial dynamics & temperature



Secondary model

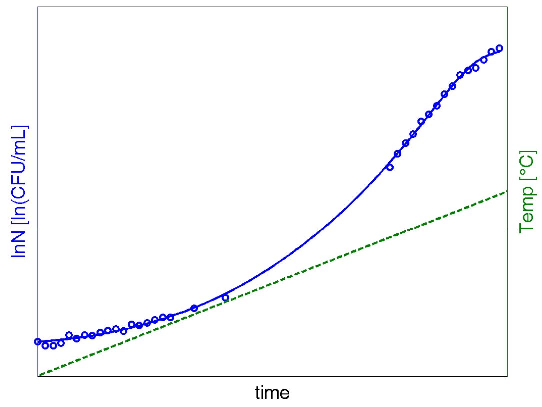
Cardinal temperature model with inflection - CTMI (Rosso et al. 1993)

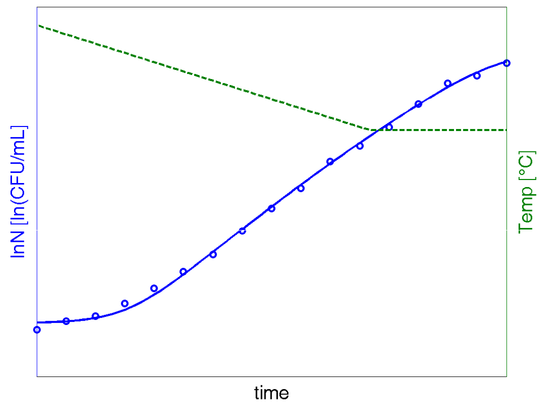
$$\mu_{max} = \mu_{opt} \cdot \gamma(\mathrm{T})$$

$$\gamma(\mathrm{T}) = \frac{(T - T_{min})^2 (T - T_{max})}{(T_{opt} - T_{min})((T_{opt} - T_{min})(T - T_{opt}) - (T_{opt} - T_{max})(T_{opt} + T_{min} - 2T))}$$

Microbial dynamics & temperature

Primary growth model & Secondary CTMI model

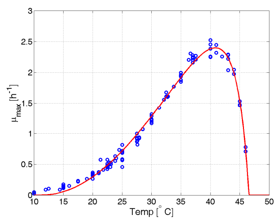




Predictive quality depends on …

- quality of the *model structure*
- quality of the *model parameter estimates*

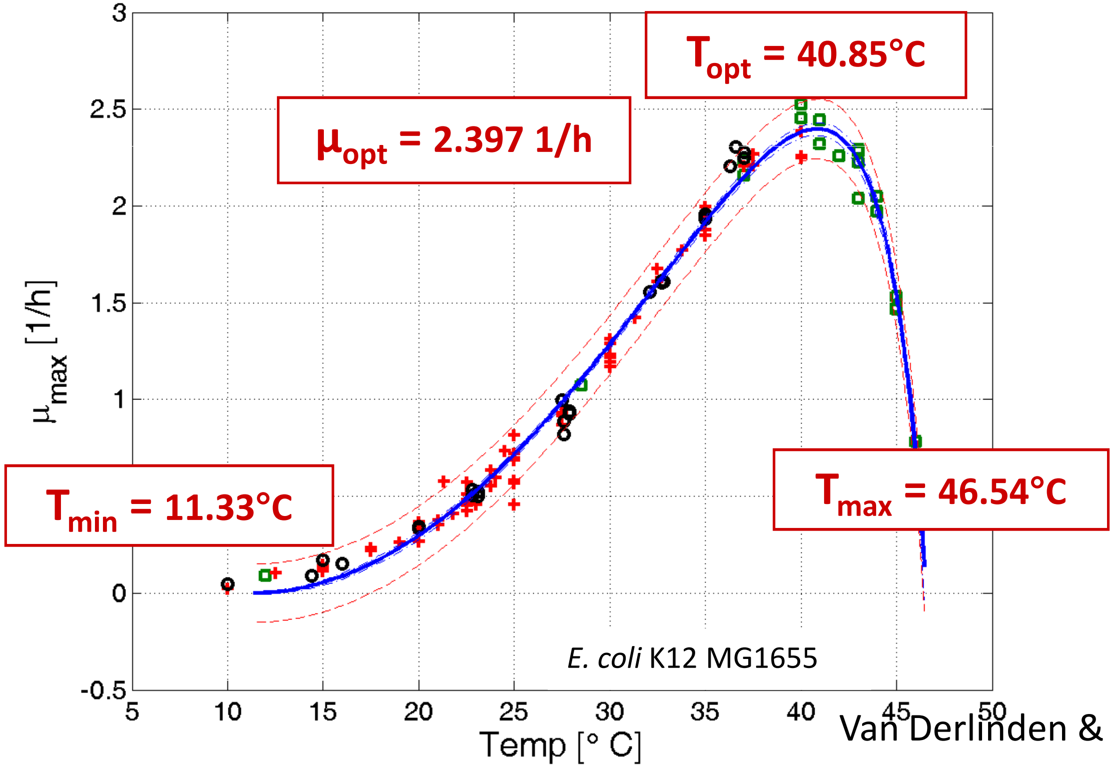
  $\rightarrow T_{min}, T_{opt}, T_{max}$ and μ_{opt}



Microbial dynamics & temperature

Efficient and accurate estimation of
the four CTMI parameters T_{min}, T_{opt}, T_{max} & μ_{opt}

1. Static data collection (= constant T)



Van Derlinden & Van Impe (2012, IJFM)

Microbial dynamics & temperature

2. Dynamic data collection (= dynamic T(t))

$\rightarrow$ evaluate model validity under
time-varying environmental conditions
$\rightarrow$ efficient ~ time & labor

selection of dynamic experiments via **OED/PE**

$\Rightarrow$ optimal dynamic temperature profile which maximizes the information content in the experimental data (= n(t))

$\Rightarrow$ selection of informative temperature zones: best **p**-estimation

> reliable and accurate **p** estimates

> limited set of dynamic experiments

OED/PE framework

Optimization stategy

Efficient and accurate estimation of
the four CTMI parameters T_{min}, T_{opt}, T_{max} & μ_{opt}
via
Optimal Dynamic Experiment Design for
Parameter Estimation (OED/PE)

(1) Single OED/PE

(2) Sequential OED/PE

(3) Global OED/PE

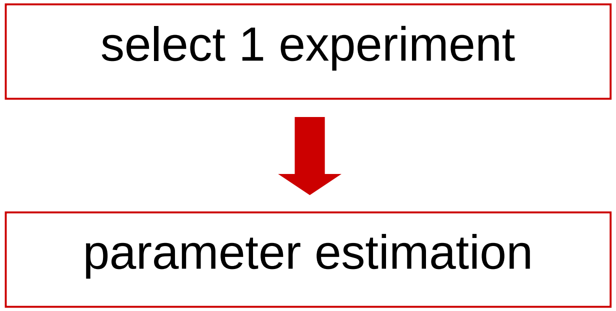
OED/PE framework

Optimization stategy

(1) Single OED/PE

(2) Sequential OED/PE

(3) Global OED/PE



- ? 4 $\mathbf{p}_{CTMI}$ from 1 experiment

OED/PE framework

Optimization stategy

(1) Single OED/PE

(2) Sequential OED/PE $\Rightarrow$ two-by-two estimation

(3) Global OED/PE

select 6 experiments

Intermediate up-date of
- nominal values
- parameter estimates

Exp.	considered parameters	nominal parameters
1	(p_1,p_2)	$(p_1^0, p_2^0, p_3^0, p_4^0)$
2	(p_1,p_3)	$(p_1^1, p_2^1, p_3^0, p_4^0)$
3	(p_1,p_4)	$(p_1^2, p_2^1, p_3^1, p_4^0)$
4	(p_2,p_3)	$(p_1^3, p_2^1, p_3^1, p_4^1)$
5	(p_2,p_4)	$(p_1^3, p_2^2, p_3^2, p_4^1)$
6	(p_3,p_4)	$(p_1^3, p_2^3, p_3^2, p_4^2)$

OED/PE framework

Optimization stategy

(1) Single OED/PE

(2) Sequential OED/PE

(3) Global OED/PE $\Rightarrow$ two-by-two estimation

based on common nominal values

↓

select 6 experiments

↓

global parameter estimation

parameters considered	parameters 'known'
(T_{max}, μ_{opt})	(T_{min}, T_{opt})
(T_{max}, T_{min})	(T_{opt}, μ_{opt})
(T_{max}, T_{opt})	(T_{min}, μ_{opt})
(T_{min}, μ_{opt})	(T_{max}, T_{opt})
(T_{min}, T_{opt})	(T_{max}, μ_{opt})
(T_{opt}, μ_{opt})	(T_{max}, T_{min})

OED/PE framework

Fisher information matrix

Fisher information matrix: information quantification

$$\mathbf{F} = \int_0^{t_f} \left(\frac{\partial n(t)}{\partial \mathbf{p}} \right)_{\big|\mathbf{p}=\mathbf{p}^o} \cdot \mathbf{Q} \cdot \left(\frac{\partial n(t)}{\partial \mathbf{p}} \right)^T_{\big|\mathbf{p}=\mathbf{p}^o} \cdot dt$$

$\longrightarrow$ Dynamic optimization problem

$$T_{input}(t) = \arg\min / \max[scal(\mathbf{F}(T_{input}(t), \mathbf{p}))]$$

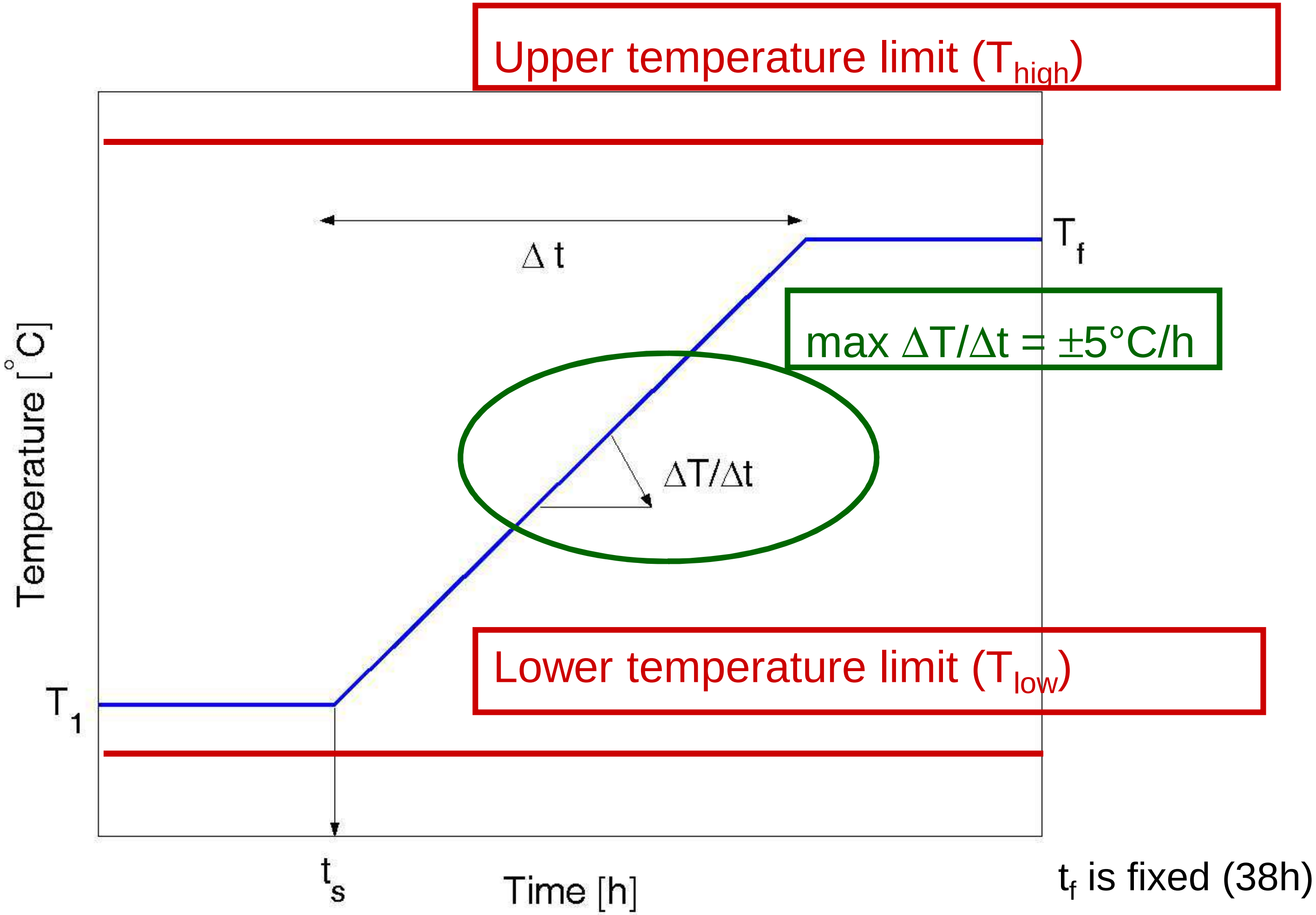
- *Output*: n(t) = ln(N(t))
- *Parameter vector:* $\mathbf{p}_{CTMI} = [T_{min}\ T_{opt}\ T_{max}\ \mu_{opt}]^T$
- *Criterion:* D-criterion (max[|F|])
- *Nominal values*: **pº**

Van Derlinden et al. (2010, JTB)

Telen, Logist, Van Derlinden & Van Impe – Agrostat (2012) On the trade-off between experimental effort and information content in optimal experiment design

OED/PE framework

Control vector parameterization



OED/PE

Comparison of three optimization strategies

1. **Simulation study**

2. **Experimental study**
   - Bioreactor experiments
   - *Escherichia coli* K12

PART I.

Comparison of three OED/PE strategies via a simulation study

OED/PE Part I. Simulation study

Methodology

Three steps

1. Design of D-optimal experiments

2. Simulation of growth curves
   $\rightarrow$ based on the true parameters
   $\rightarrow$ error ~ measurement error and biological variability

3. Estimation of parameters from simulated experimental data
   $\rightarrow$ minimization of the sum of squared errors

- three different sets of true parameters (**p***)
- **SiOED/PE & GlOED/PE** $\rightarrow$ identical nominal parameters
- **SeOED/PE** $\rightarrow$ different order of **p**-combinations

OED/PE Part I. Simulation study

Results

KATHOLIEKE UNIVERSITEIT LEUVEN

BioTeC

AGRO-INDUSTRIE et MÉTHODES STATISTIQUES

(Initial) nominal values

T_{min}° = 11.33 T_{max}° = 46.54

T_{opt}° = 40.85 μ_{opt}° = 2.396

Van Derlinden et al. (2010, JTB)

Results

	$\mathbf{p}^*$	SiOED/PE		SeOED/PE		GlOED/PE	
T_{min}	8.000	8.073	(4.371×10^{-1})	6.163	(1.852×10^{-1})	7.822	(1.072×10^{-1})
T_{opt}	40.00	40.10	(5.288×10^{-1})	40.84	(1.241×10^{-1})	40.12	(7.278×10^{-2})
T_{max}	46.00	46.17	(4.112×10^{-1})	45.51	(3.278×10^{-2})	45.92	(3.549×10^{-4})
μ_{opt}	2.200	2.184	(7.795×10^{-2})	2.368	(2.305×10^{-2})	2.204	(1.110×10^{-2})
T_{min}	6.500	6.464	(4.385×10^{-1})	4.256	(5.445×10^{-1})	6.491	(1.146×10^{-1})
T_{opt}	39.80	40.00	(5.601×10^{-1})	42.04	(6.403×10^{-2})	40.09	(8.706×10^{-2})
T_{max}	45.70	45.78	(2.184×10^{-1})	45.19	(1.012×10^{-4})	45.55	(2.437×10^{-5})
μ_{opt}	2.000	2.007	(5.656×10^{-2})	2.173	(1.698×10^{-2})	2.051	(1.246×10^{-2})
T_{min}	10.00	10.08	(3.683×10^{-1})	9.828	(2.779×10^{-1})	9.889	(8.118×10^{-2})
T_{opt}	40.50	40.47	(5.744×10^{-1})	40.61	(1.123×10^{-1})	40.58	(1.576×10^{-1})
T_{max}	45.30	45.40	(1.635×10^{-1})	45.27	(1.063×10^{-2})	45.29	(3.415×10^{-2})
μ_{opt}	2.350	2.318	(9.301×10^{-2})	2.392	(7.863×10^{-2})	2.348	(2.823×10^{-2})

PART II.

Comparison of SiOED/PE and GlOED/PE strategies based on experimental data

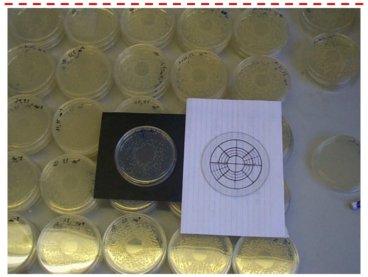
OED/PE Part II. Experiments

Experimental set-up

Optimal experiment design for parameter estimation

- *Escherichia coli*

Experimental implementation



<u>cell density</u>

plate counting

microorganism is grown in laboratory medium

<u>measurements & control</u>

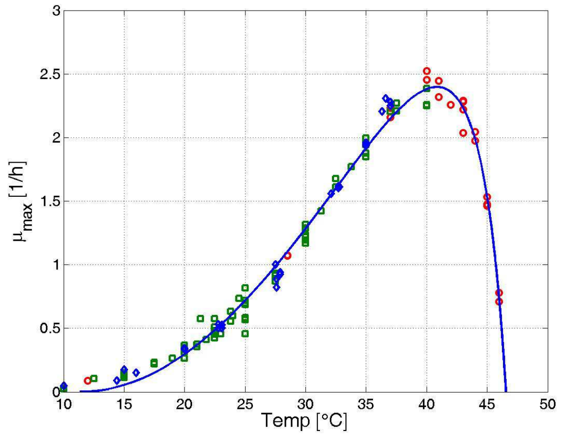
T, pH, aeration & dissolved oxygen, rpm

OED/PE Part II. *Escherichia coli*

Background

- ✓ T effect on *E. coli* dynamics extensively studied @ KULeuven/BioTeC
  - Swinnen et al. (2005, 2006)
  - Bernaerts et al. (2000, 2002)
  - Valdramidis et al. (2006, 2007)

- ✓ advantage
  - design of constraints
  - selection of realistic nominal values



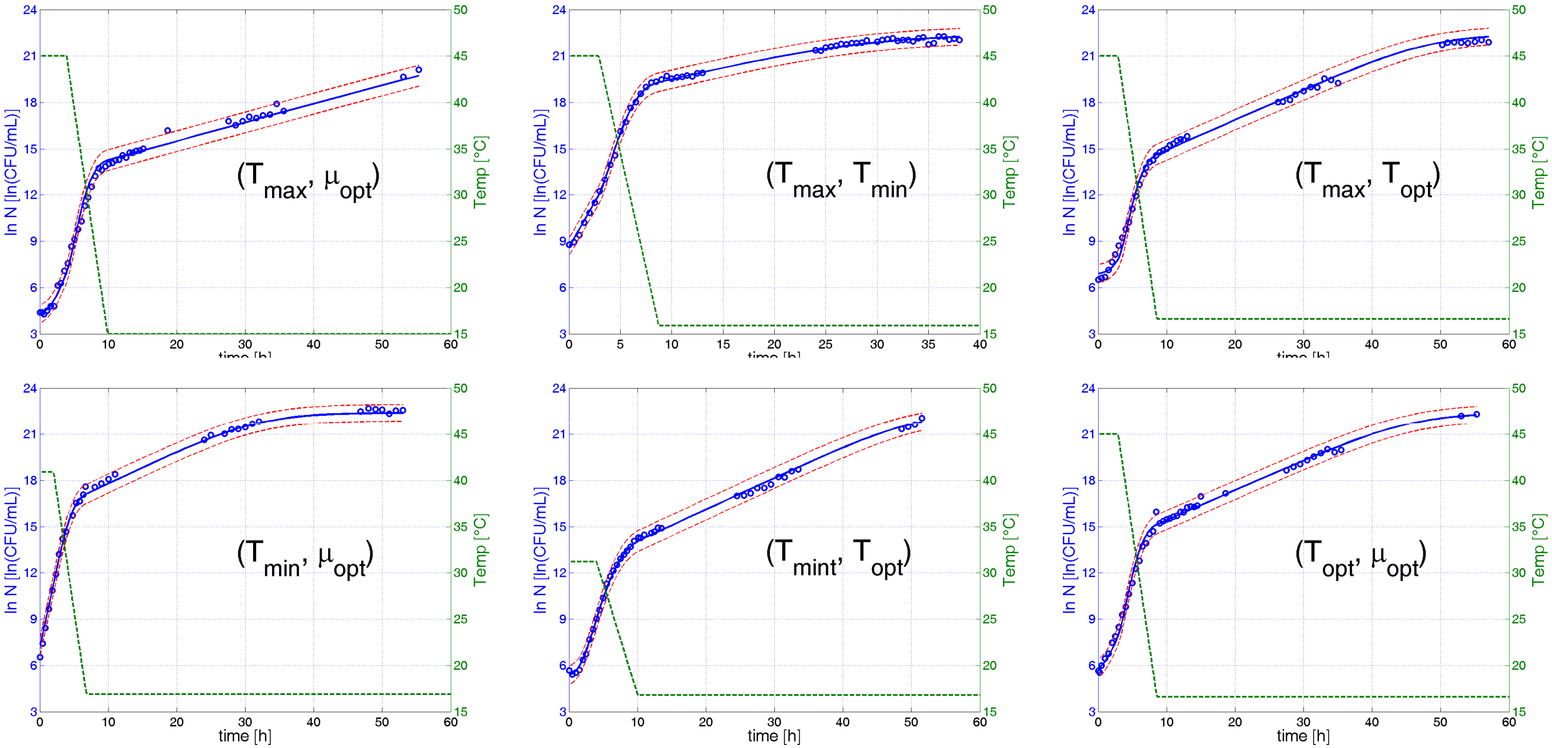
Nominal values

- T_{min} = 11.33°C
- T_{opt} = 40.85°C
- T_{max} = 46.54°C
- μ_{opt} = 2.397°C

- ✓ $(\Delta T/\Delta t)_{max}$ = 5°C/h, t_f = 38h

OED/PE Part II. *Escherichia coli*

Results – Global strategy (GlOED/PE)

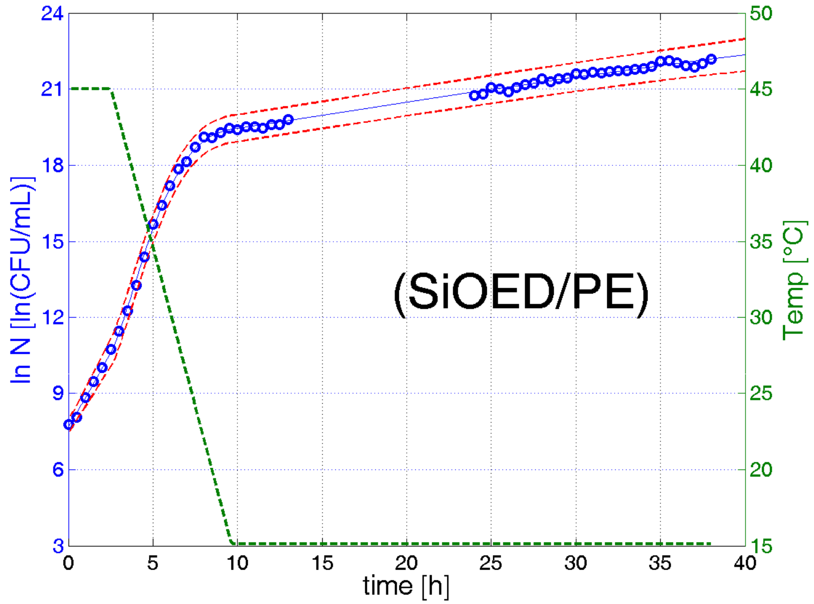


- Reliable estimates for Tmin, Topt, Tmax and μopt
- Similar to values obtained from static experiments
- Low uncertainty

Van Derlinden et al. (2010, JTB)

OED/PE Part II. *Escherichia coli*

Results – Single strategy (SiOED/PE)



Van Derlinden et al. (2010, JTB)

- Reliable and accurate CTMI parameter estimates

OED/PE

General conclusions

Three optimization strategies

- SiOED/PE
- SeOED/PE
- GlOED/PE

1) simulation study

→ SeOED/PE repeatedly in unrealistic $\mathbf{p}_{CTMI}$ estimates

2) reliable and accurate CTMI parameter estimates derived from the six GlOED/PE optimal experiments

→ each CTMI parameter is considered three times

3) accurate $\mathbf{p}_{CTMI}$ estimation not guaranteed via SiOED/PE

→ Information is limited

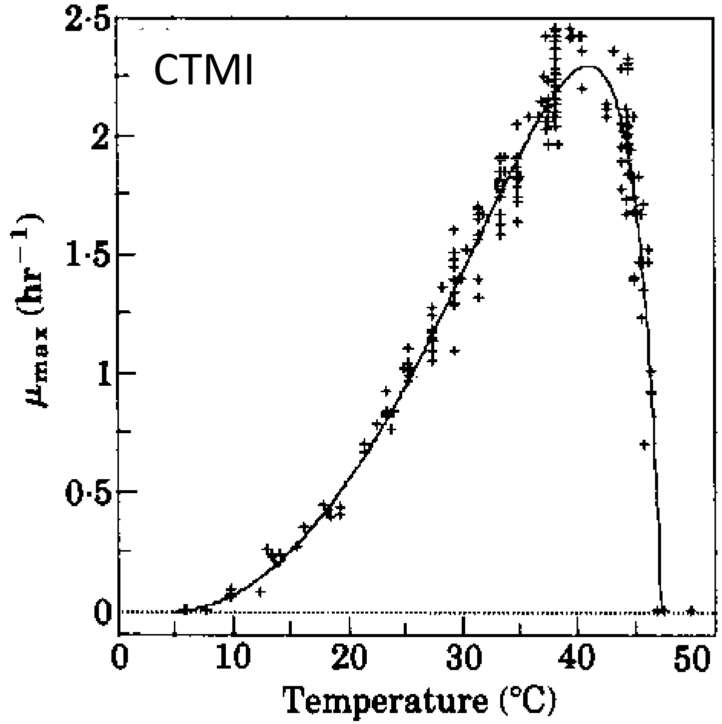
→ Parameter only considered once

→ Highly sensitive to experimental errors

Microbial dynamics & temperature

Multiple model structures

- **Multiple models exist to describe effect of T on μmax**
  - Cardinal Temperature Model with Inflection

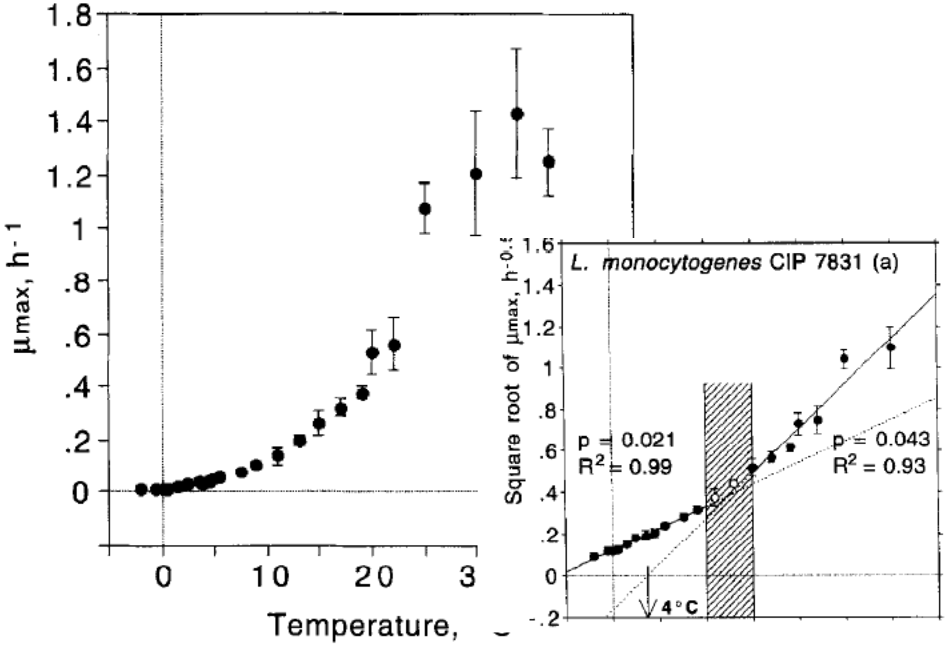


Rosso et al. (1993, JTB)

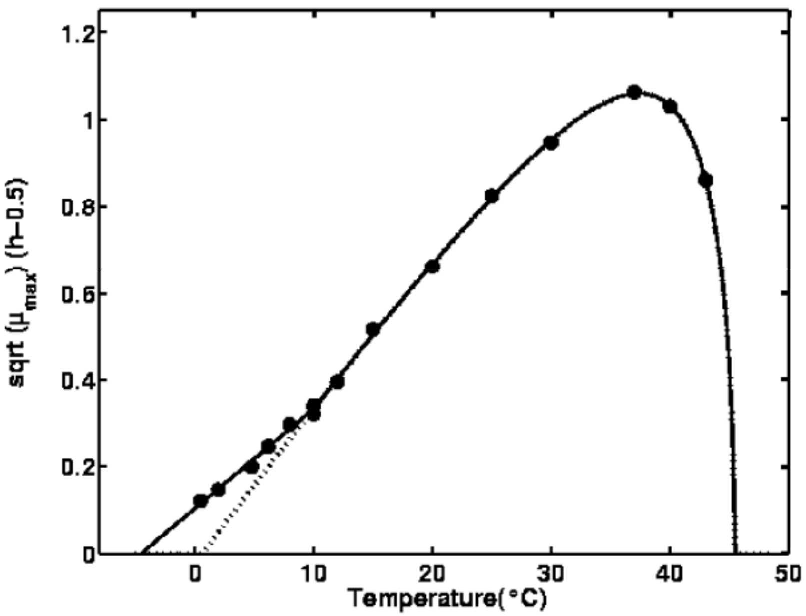
OED/MD

Microbial dynamics & temperature

- Deviating behavior observed for *Listeria*



Listeria monocytogenes (Bajard et al. 1996, IJFM)



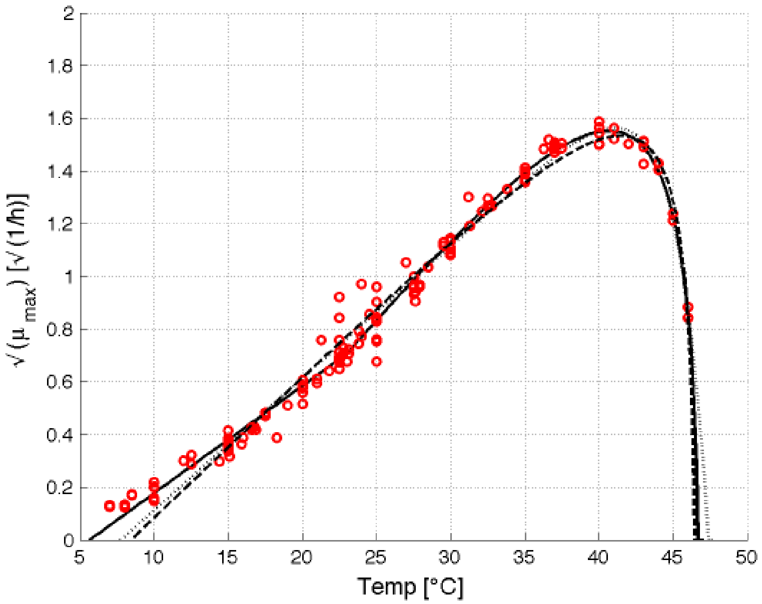
Listeria innocua (Le Marc et al. 2002, IJFM)

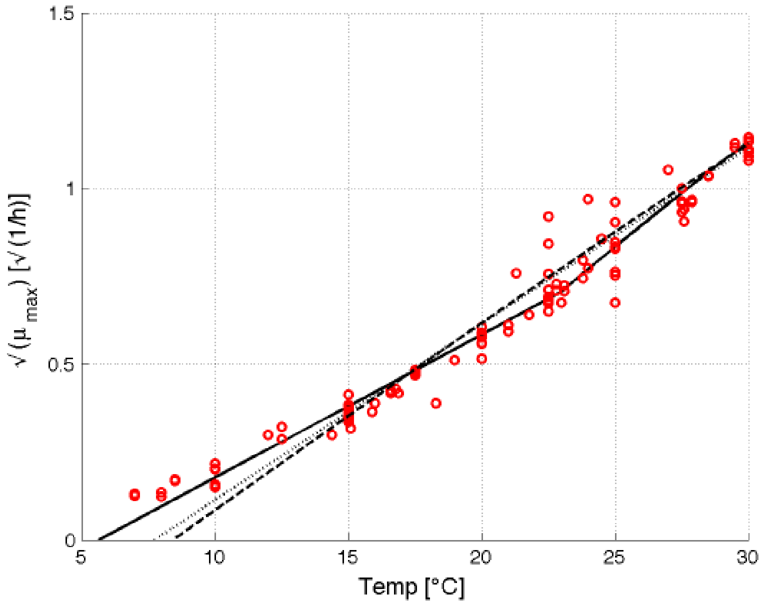
=> Adapted CTMI model structure by Le Marc et al. (2002, IJFM)

OED/MD

Microbial dynamics & temperature

- Deviating behavior possible also valid for other strains
- *Escherichia coli* K12





Van Derlinden & Van Impe (2012, IJFM)

OED/MD

Microbial dynamics & temperature

Work in progress

question: what is the correct model?

optimal experiment design for model structure discrimination

- **Aim** =
  - to design a (set of) dynamic temperature experiments that shows the maximum difference in model output
  - identify which models corresponds most with reality

Kinetic growth rate models

Case study 2. Design Of Experiments (DOE)

- Parameter estimation

Kinetic model: effect of temperature, pH, lactic acid and a_w on μ_{max}

Design Of Experiments

Design of experiments includes different - statistically inspired - approaches that result in the structured collection of information with respect to one or more variables, their variability and interactions.

- Identify variables that mainly control the studied system
- How these variables interact.

➢ **the basic idea behind DOE is to hold certain factors constant and to alter the levels of one or more other variables.**

Design Of Experiments

Implementations in predictive microbiology

1. To **scan an extended region**.
   – Probabilistic model building; quantifiy the chance of survival and/or growth of pathogens or spoilage microorgansims given certain processing conditions
   (Belletti et al., 2007; Gao et al., 2006; Uljas et al., 2001).

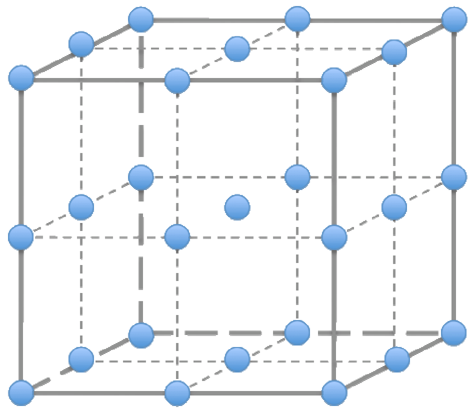
1. Extended data collection to build **response surface models**.
   – via polynomial regression,
   – effect of intrinsic/extrinsic factors on selected variables describing microbial dynamics
   (inactivation parameters - Maks et al., 2010).

1. Accurate and reliable parameter **estimates of existing models.**
   (Jagannath et al., 2005; Miller et al., 2009).

Design Of Experiments

Design schemes

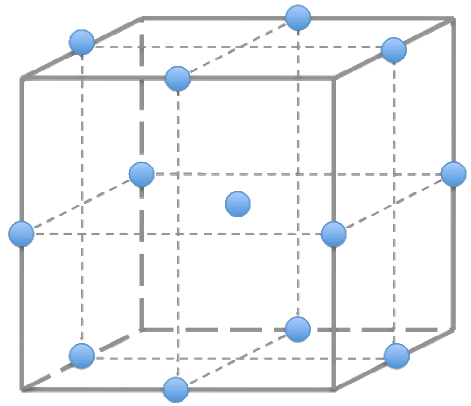
- <u>Experimental design schemes</u>
  - *Full Factorial*
  - Fractional factorial



Design Of Experiments

Design schemes

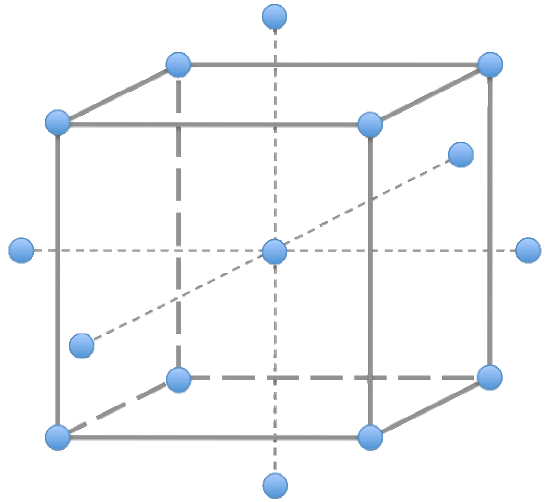
- <u>Experimental design schemes</u>
  - Full Factorial
  - Fractional factorial
  - *Box Behnken*



Design Of Experiments

Design schemes

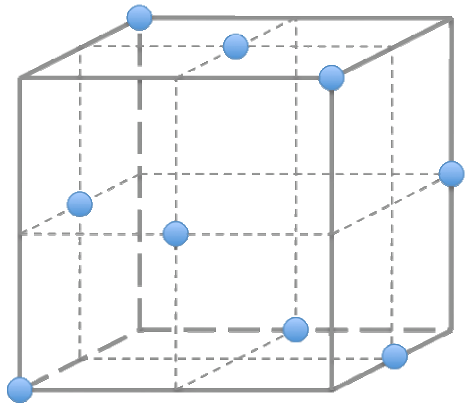
- <u>Experimental design schemes</u>
  - Full Factorial
  - Fractional factorial
  - Box Behnken
  - *Central Composite*



Design Of Experiments

Design schemes

- <u>Experimental design schemes</u>
  - Full Factorial
  - Fractional factorial
  - Box Behnken
  - Central Composite
  - *Latin Square*



Case study

Research objective

Assess the performance of experimental design schemes with respect to parameter estimation of a secondary model

Effect on microbial growth rate (*Escherichia coli)*

- Water activity
- Temperature
- pH
- Lactic acid

$$\sqrt{\mu_{\max}} = c \cdot (T - T_{min}) \cdot (1 - \exp(d(T - T_{max}))) \cdot \sqrt{a_w - a_{wmin}} \cdot \sqrt{1 - 10^{pH_{min} - pH}} \cdot \sqrt{1 - 10^{pH - pH_{max}}} \cdot \sqrt{1 - \frac{LAC}{U_{min} \cdot \left(1 + 10^{pH - pK_a}\right)}} \cdot \sqrt{1 - \frac{LAC}{D_{min} \cdot \left(1 + 10^{pK_a - pH}\right)}}$$

Ross et al. (2002)

Case study

Simulation study

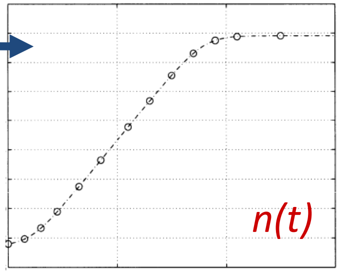
- **Stepwise procedure**

Secondary model

Original parameter estimates

μ_{max} (T, pH, a_w)

Baranyi & Roberts (1994)



Case study

Simulation study

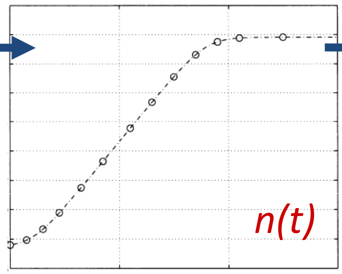
- **Stepwise procedure**

<u>Secondary model</u>

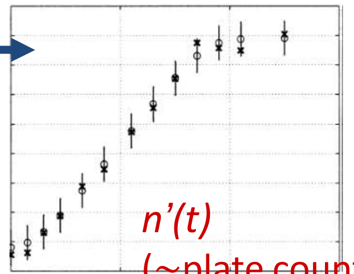
Original parameter estimates

μ_{max} (T,pH,a_w)

Baranyi & Roberts (1994)



perturbation of data



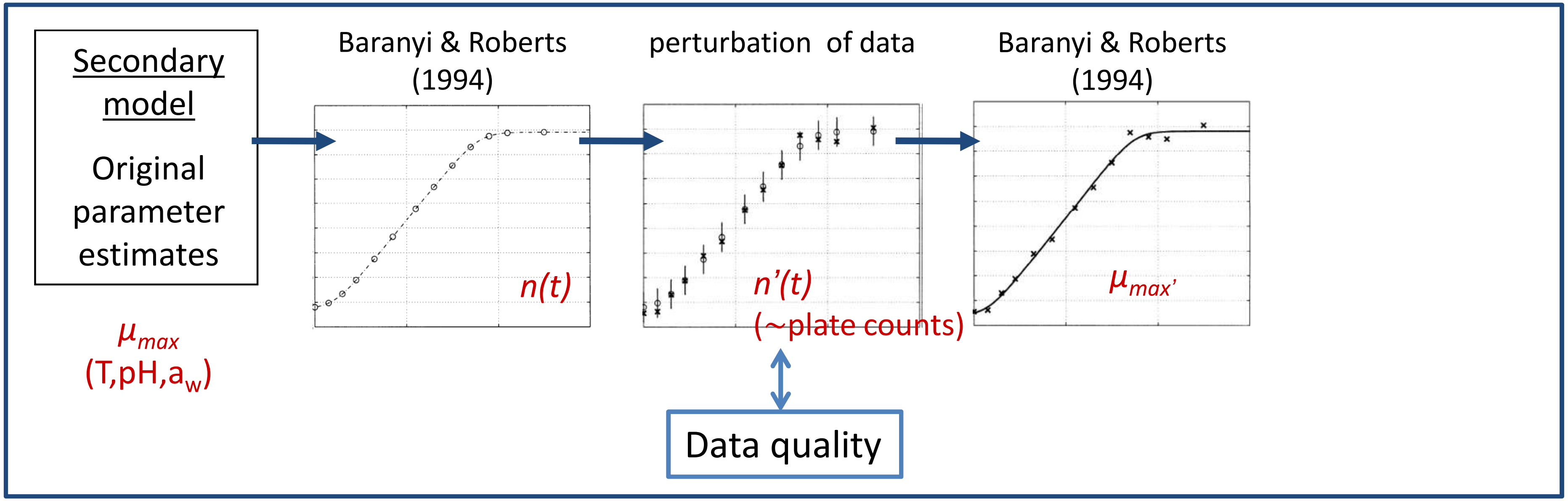
(~plate counts)

Data quality

Case study

Simulation study

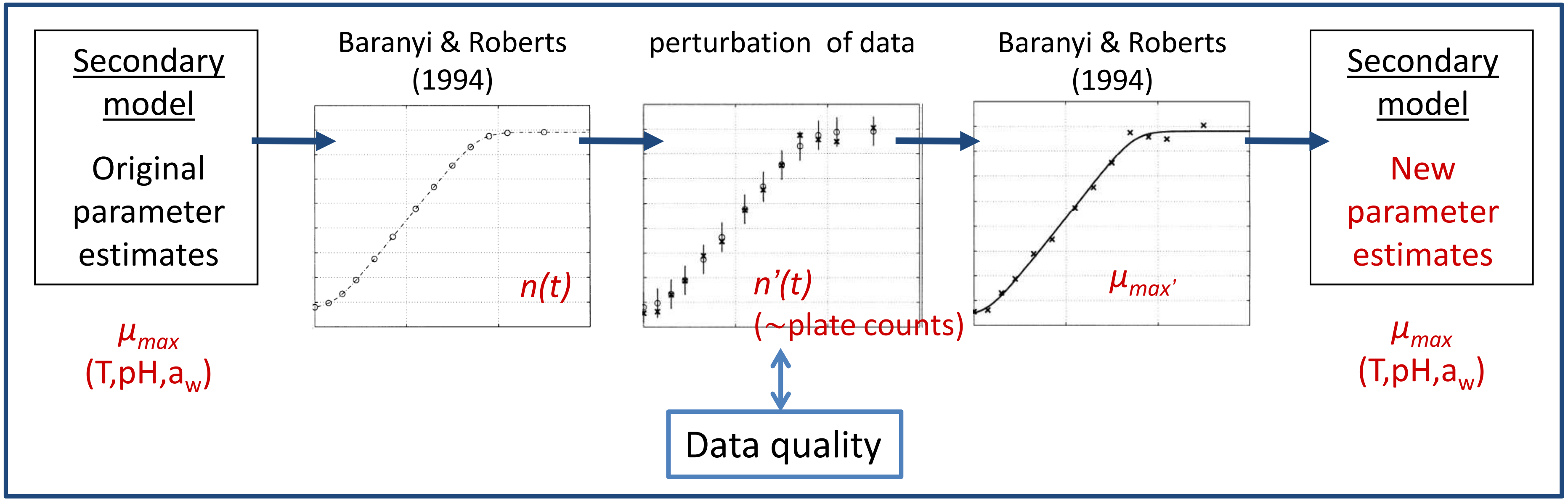
- **Stepwise procedure**



Case study

Simulation study

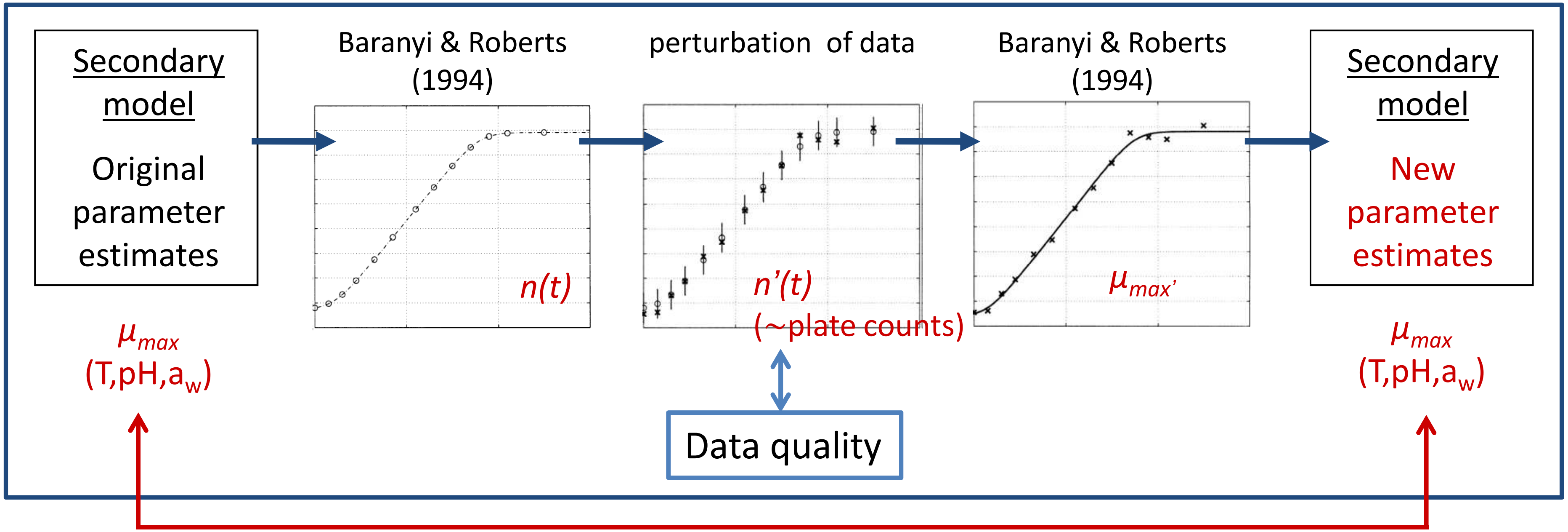
- **Stepwise procedure**



Case study

Simulation study

- **Stepwise procedure**



Performance criteria

- # experiments
- goodness-of-fit statistics
- reliability of parameter estimates

DOE Simulation study

Overview designs

Design scheme	# levels per factor	# experiments	RMSE
FF5 Full Factorial	5	625	0.009/0.030
RFF Reduced Full Factorial	5*	90	0.009/0.029
LS9 Latin-Square	9	81	0.010/0.029
CC Central Composite	5	36	0.010/0.034
LS5 Latin-Square	5	25	0.008/0.031

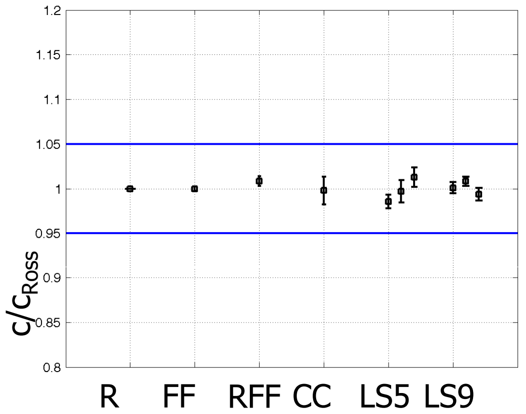
RFF = 5 T, 3 pH, 3 LAC, 2 a_w

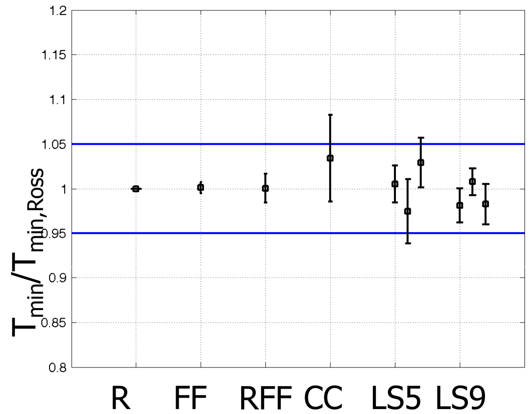
DOE Simulation study

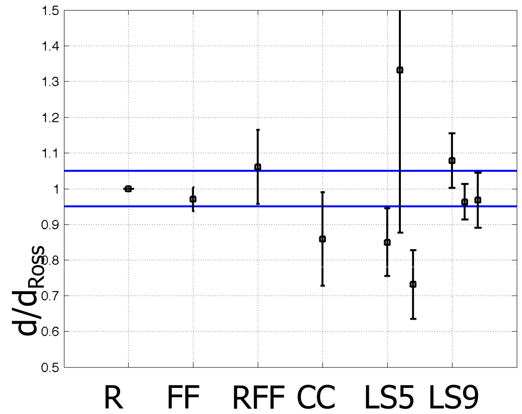
Model parameter estimation results

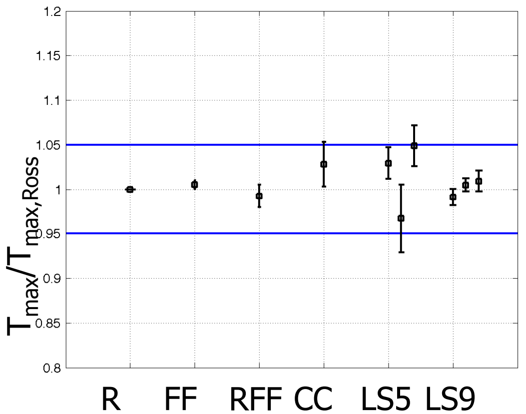
- **Low level of noise on data**

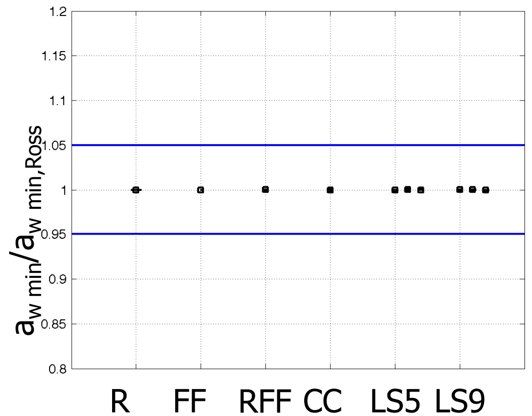
$$\sqrt{\mu_{\max}} = c\cdot(T-T_{min})\cdot(1-\exp(d(T-T_{max})))\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot(1+10^{pH-pK_a})}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot(1+10^{pK_a-pH})}}$$

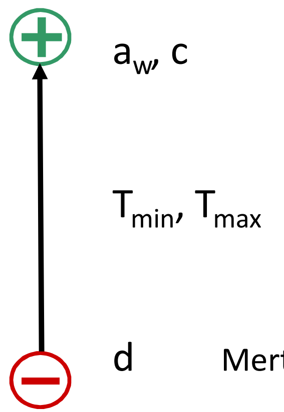












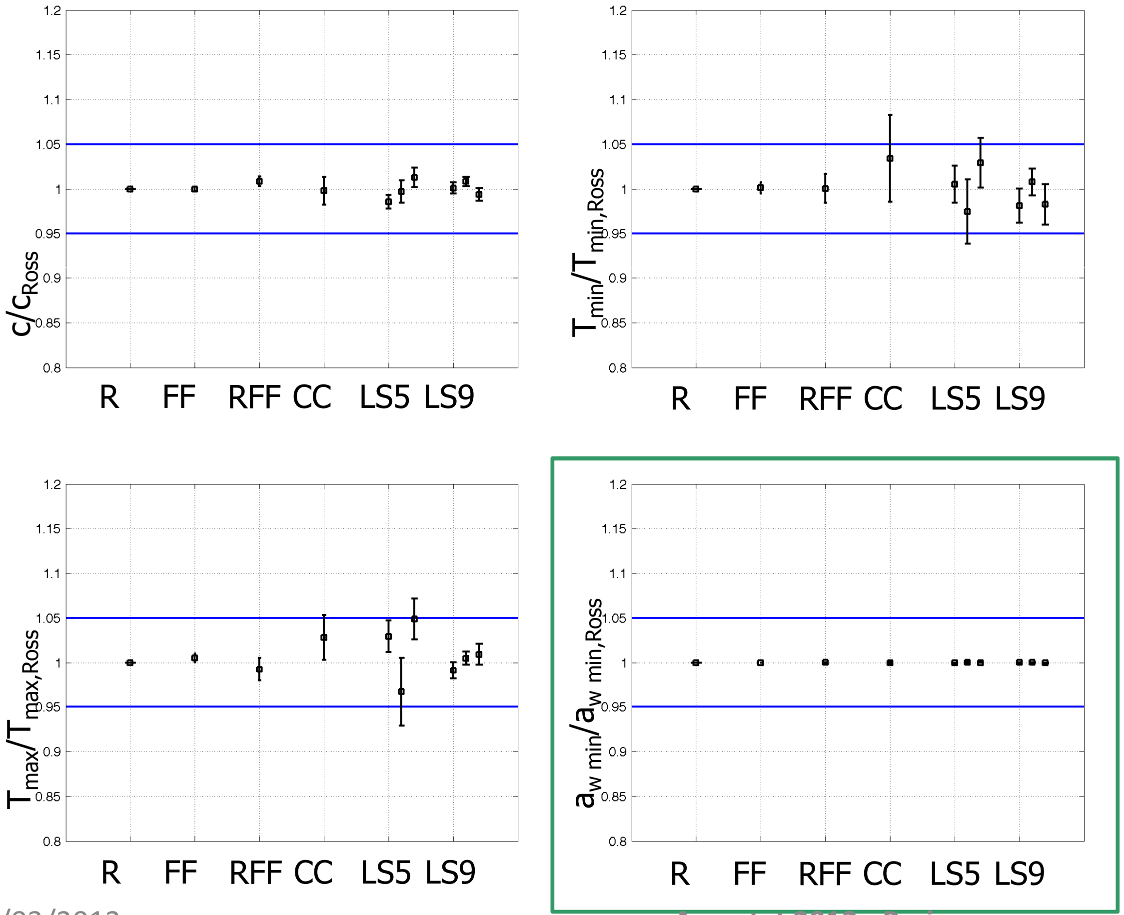
Mertens et al. (2012, FC)

DOE Simulation study

Model parameter estimation results

- **Low level of noise on data**

$$\sqrt{\mu_{\max}} = c\cdot(T-T_{min})\cdot\left(1-\exp\left(d\left(T-T_{max}\right)\right)\right)\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot\left(1+10^{pH-pK_a}\right)}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot\left(1+10^{pK_a-pH}\right)}}$$



a_w, c

T_{min}, T_{max}

d

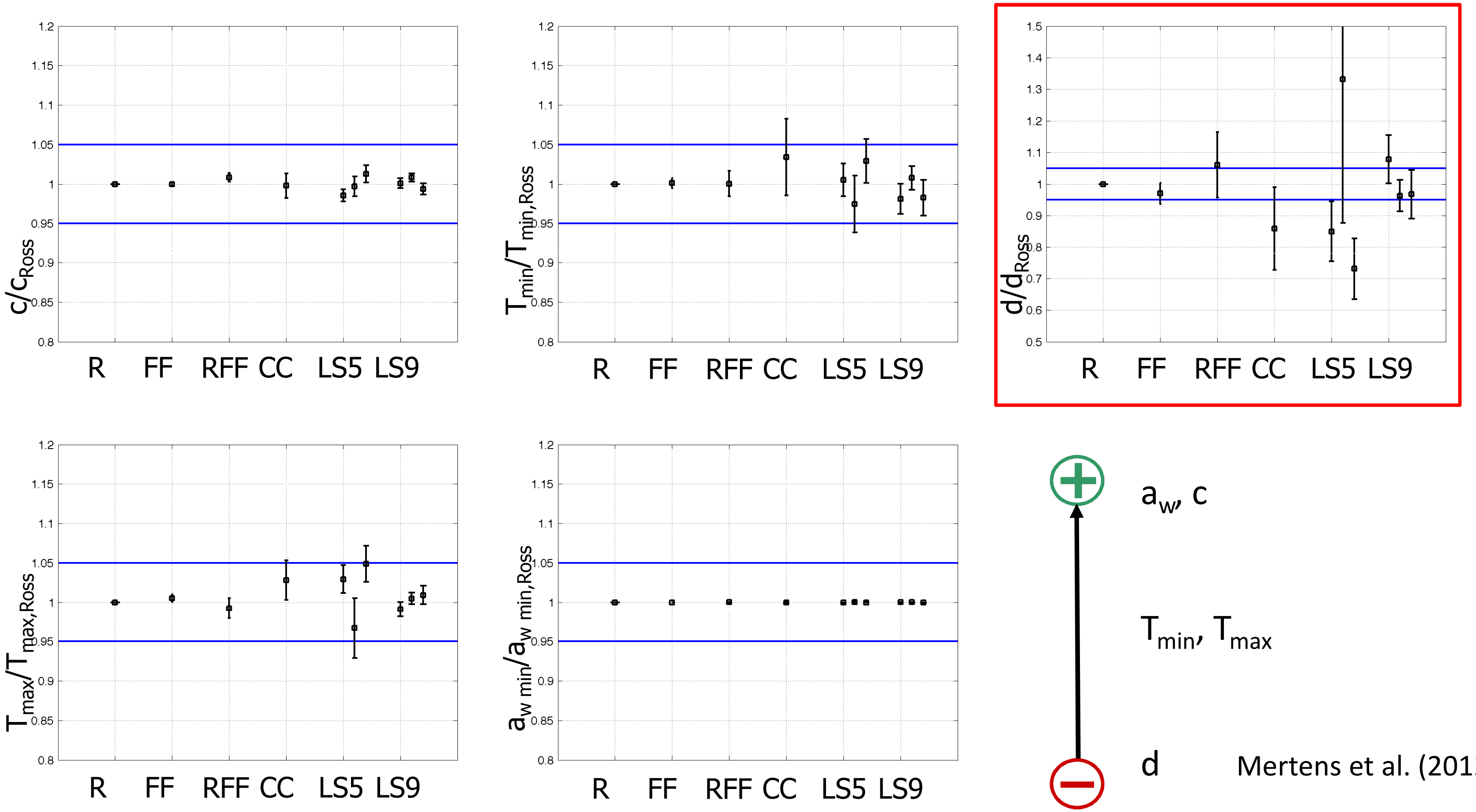
Mertens et al. (2012, FC)

DOE Simulation study

Model parameter estimation results

- **Low level of noise on data**

$$\sqrt{\mu_{\max}} = c \cdot (T - T_{min}) \cdot (1 - \exp(d(T - T_{max}))) \cdot \sqrt{a_w - a_{wmin}} \cdot \sqrt{1 - 10^{pH_{min} - pH}} \cdot \sqrt{1 - 10^{pH - pH_{max}}} \cdot \sqrt{1 - \frac{LAC}{U_{min} \cdot (1 + 10^{pH - pK_a})}} \cdot \sqrt{1 - \frac{LAC}{D_{min} \cdot (1 + 10^{pK_a - pH})}}$$



Mertens et al. (2012, FC)

DOE Simulation study

Model parameter estimation results

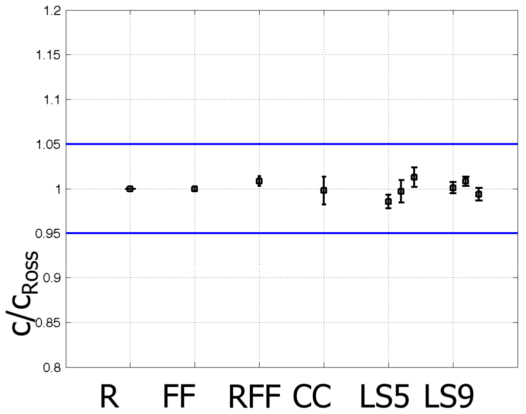
KATHOLIEKE UNIVERSITEIT LEUVEN

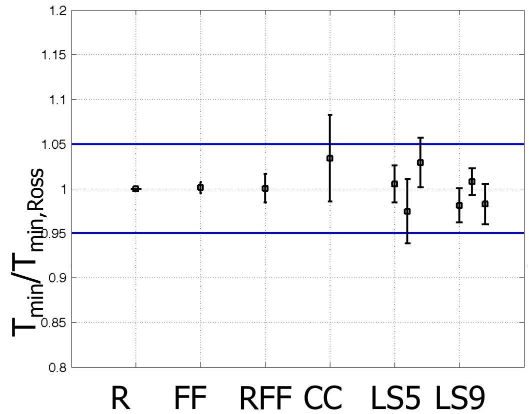
BIOPROCESS TECHNOLOGY & CONTROL BioTeC

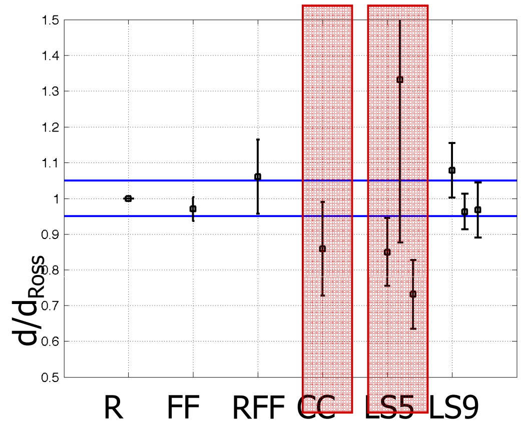
AGRO-INDUSTRIE et MÉTHODES STATISTIQUES

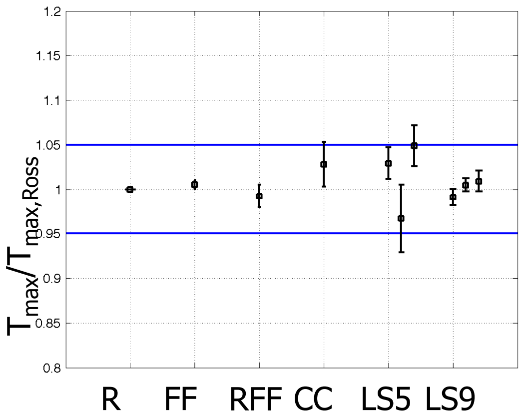
- **Low level of noise on data**

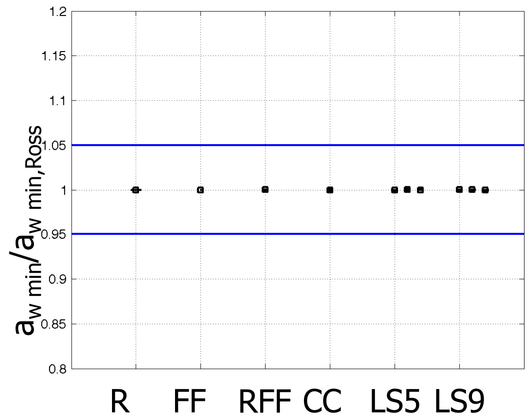
$$\sqrt{\mu_{max}} = c\cdot(T-T_{min})\cdot(1-\exp(d(T-T_{max})))\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot(1+10^{pH-pK_a})}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot(1+10^{pK_a-pH})}}$$

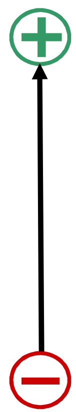












a_w, c

T_{min}, T_{max}

d

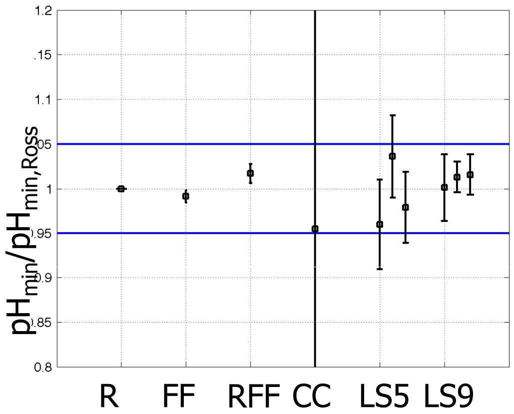
Mertens et al. (2012, FC)

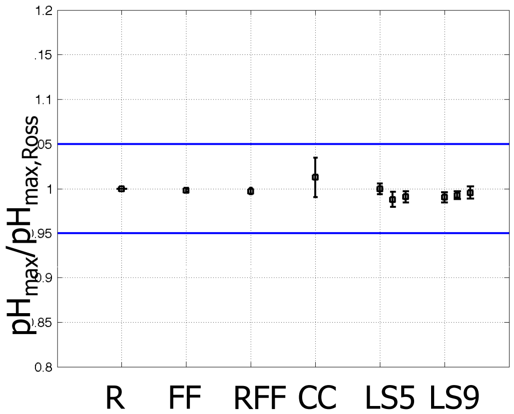
DOE Simulation study

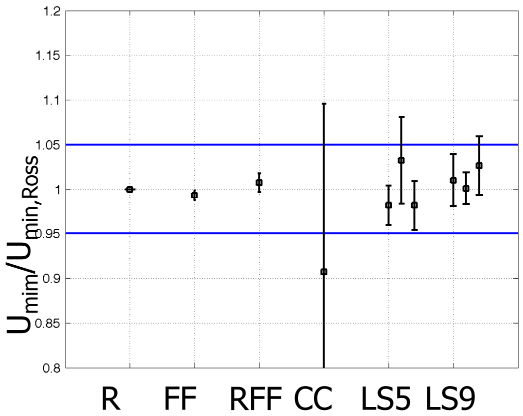
Model parameter estimation results

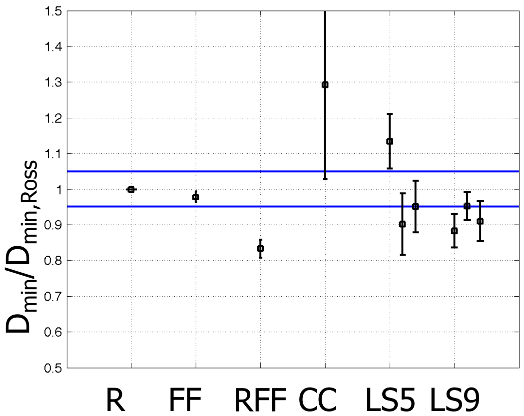
- **Low level of noise on data**

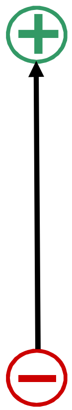
$$\sqrt{\mu_{\max}} = c\cdot(T-T_{min})\cdot(1-\exp(d(T-T_{max})))\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot(1+10^{pH-pK_a})}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot(1+10^{pK_a-pH})}}$$











pH_{max}

pH_{min}, U_{min}

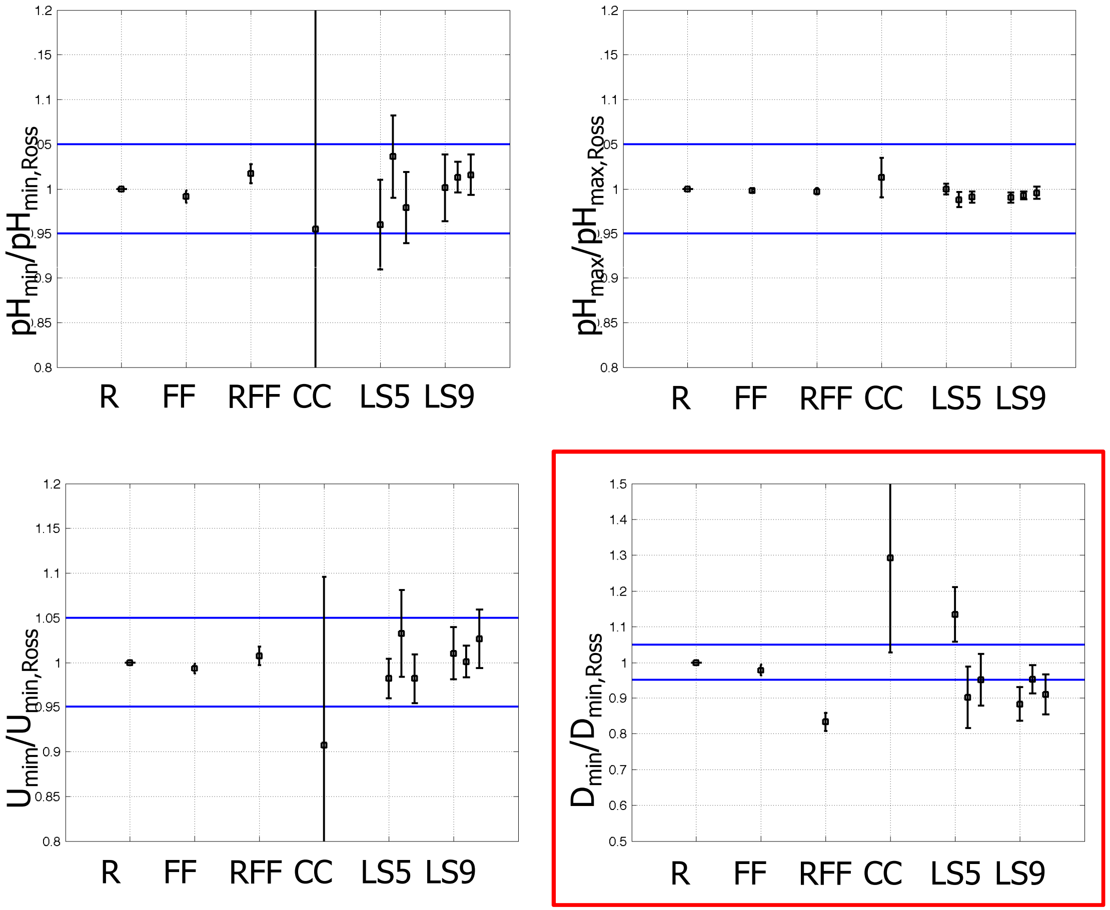
D_{min} Mertens et al. (2012, FC)

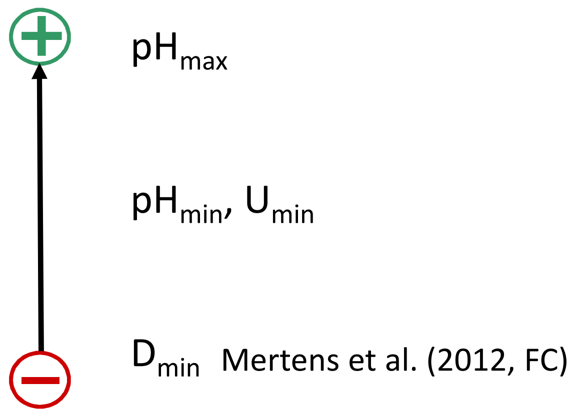
DOE Simulation study

Model parameter estimation results

- **Low level of noise on data**

$$\sqrt{\mu_{max}} = c\cdot(T-T_{min})\cdot(1-\exp(d(T-T_{max})))\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot(1+10^{pH-pK_a})}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot(1+10^{pK_a-pH})}}$$





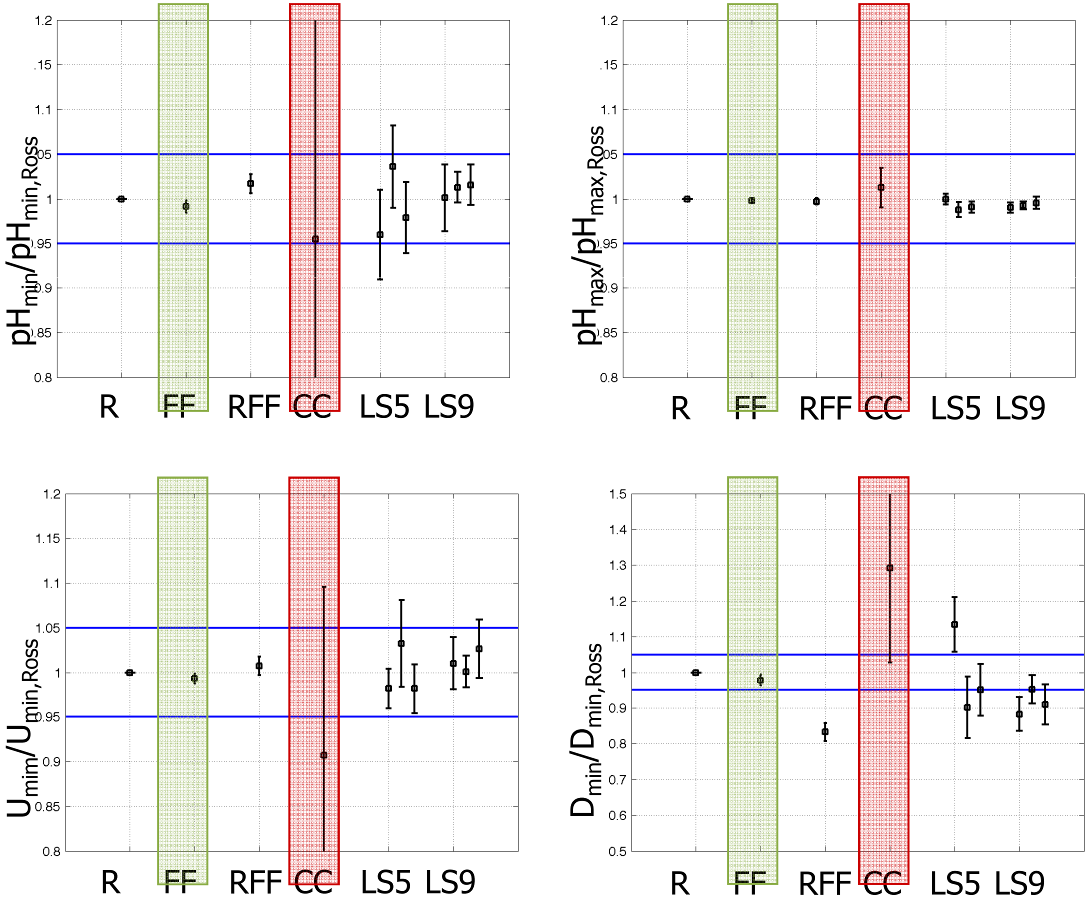
DOE Simulation study

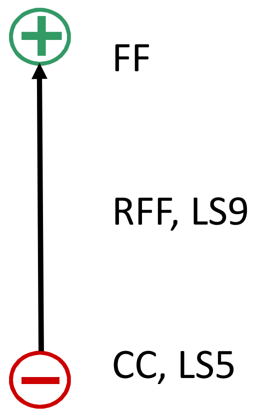
Model parameter estimation results

- **Low level of noise on data**

$$\sqrt{\mu_{\max}} = c\cdot(T-T_{min})\cdot(1-\exp(d(T-T_{max})))\cdot\sqrt{a_w-a_{wmin}}\cdot\sqrt{1-10^{pH_{min}-pH}}\cdot\sqrt{1-10^{pH-pH_{max}}}\cdot\sqrt{1-\frac{LAC}{U_{min}\cdot(1+10^{pH-pK_a})}}\cdot\sqrt{1-\frac{LAC}{D_{min}\cdot(1+10^{pK_a-pH})}}$$

Mertens et al. (2012, FC)





DOE Simulation study

Conclusions

Accuracy & uncertainty of parameter estimates

- (+) a_w
- pH
- (−) Temperature, Lactic acid concentration

Selecting a design = **trade-off** (data quantity, # levels & factors)

- (+) Full Factorial
  - Reduced Full Factorial — based on *a priori* knowledge
  - Latin-Square
- (−) Central Composite/Box Behnken

Probabilistic models

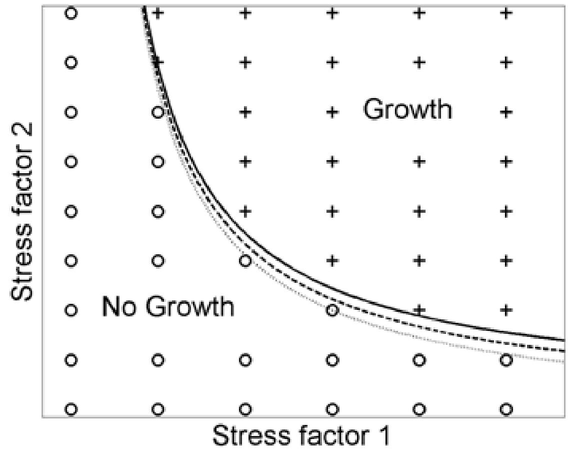
Probabilistic models

Definition: describe the chance that a certain event (growth, survival, toxin production) occurs in a time span given the intrinsic and extrinsic factors.

- a logic continuation of the kinetic growth model
- @ harsher environmental conditions:
  - growth rate is decreasing until zero and
  - lag phase is increasing until eternity
- the boundary between growth and no growth has been reached
- a kinetic model => a probabilistic model
- model output = chance that a certain event will take place within a certain given time

Probabilistic models

- Responses, under the defined experimental conditions and in a certain period of time, will be coded
  - as either 0 (response not observed)
  - or 1 (response observed).
- Repeated observations @ identical conditions
  - a probability: 0 and 100%



Probabilistic models

Applications in food processing

- **In relation to the (food) composition**
  - G/NG models for **pathogens**
    - Low infection dose
    - Growth ability > growth rate
  - G/NG models for **food spoilage** microorganisms
  - Toxin/no toxin production, e.g., *Clostridium perfringens*
- **In relation to the applied process**
  - Recovery/no recovery

Probabilistic models

Model structures

Gysemans et al. (2007, IJFM) defines four major model types

1. **Deterministic models**
2. **Based on the Minimum Convex Polyhedron concept**
3. **Logistic regression models**
4. **Artificial neural networks**

Probabilistic models

Logistic regression models

Ordinary logistic regression model

$$\mathrm{logit}(p) = \left(\frac{p}{1-p}\right)$$

with p the probability of growth, and x & y the variables considered

$$= a_1 \cdot x + a_2 \cdot y + a_3 \cdot x^2 + a_4 \cdot y^2 + a_5 \cdot x \cdot y$$

+ Easy to fit
+ Can take many variables + interactions

– Data over-fitting is plausible

Gysemans et al. (2007, IJFM)

Probabilistic models

Overall applicability

- Due to the mainly data-driven approach of G/NG models, the properties of the experimental data are of utmost importance, i.e., the overall validity of the model is determined by
  - The number of data/repetitions
  - The experimental region considered
  - The experimental time covered
  - The initial cell count

=> Well-considered experimental design is required as this will define the final applicability of the probabilistic model!

Probabilistic models

Design Of Experiments

- Probabilistic models => polynomial
- Use of DOE designs for response surface models

- Polynomial (RSM) => specific DOE
  - Box-Behnken
    - Battey et al. (2002, AEM): probability of yeast spoilage of cold-filled ready-to-drink beverages

Probabilistic models

Design of Experiments

- Probabilistic models => polynomial
- Use of DOE designs for response surface models

- Polynomial (RSM) => specific DOE
  - Box-Behnken
  - Central Composite
    - Hwang (2009, FM): growth probability of *L. monocytogenes* in salmon (salt, smoke (phenol) and T)

Probabilistic models

Design of Experiments

- Probabilistic models => polynomial
- Use of DOE designs for response surface models

- Polynomial (RSM) => specific DOE
  - Box-Behnken
  - Central Composite
    - Hwang (2009, FM): growth probability of *L. monocytogenes* in salmon (salt, smoke (phenol) and T)

- Almost all = full factorial design

Probabilistic models

Design of Experiments

↔ Focus is on the **event** (G/NG) boundary

- Specific location
- Large number of samples
- Large number of replicates

→ a two-step approach

1. First approximation of the G/NG region
2. Perform large # of experiments & replicates @ G/NG

Probabilistic models

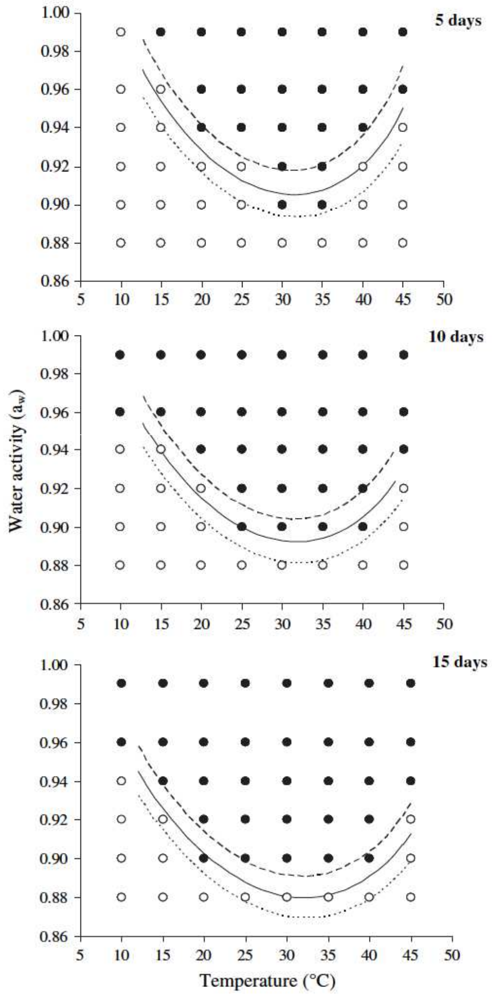
Important model building aspects

- Experimental time
- Inoculum level
- Data collection & Number of repetitions
- Model complexity: polynomial models
- Identifying the studied action: detection limit (OD)
- System characteristics: Liquid-solid

Probabilistic models

Experimental time

- Experimental time taken highly defines the final validity of the model
- Growth boundaries tend to shifted to harsher conditions when the time span increases



G/NG of *Byssochlamys fulva* DSM 1808 on malt extract agar at P values of 0.9 (dashed line), 0.5 (solid line), and 0.1 (dotted line). Solid symbols indicate fungal growth; open symbols indicate no growth.

Panagou et al. (2010, IJFM)

Probabilistic models

Inoculum level

KATHOLIEKE UNIVERSITEIT LEUVEN

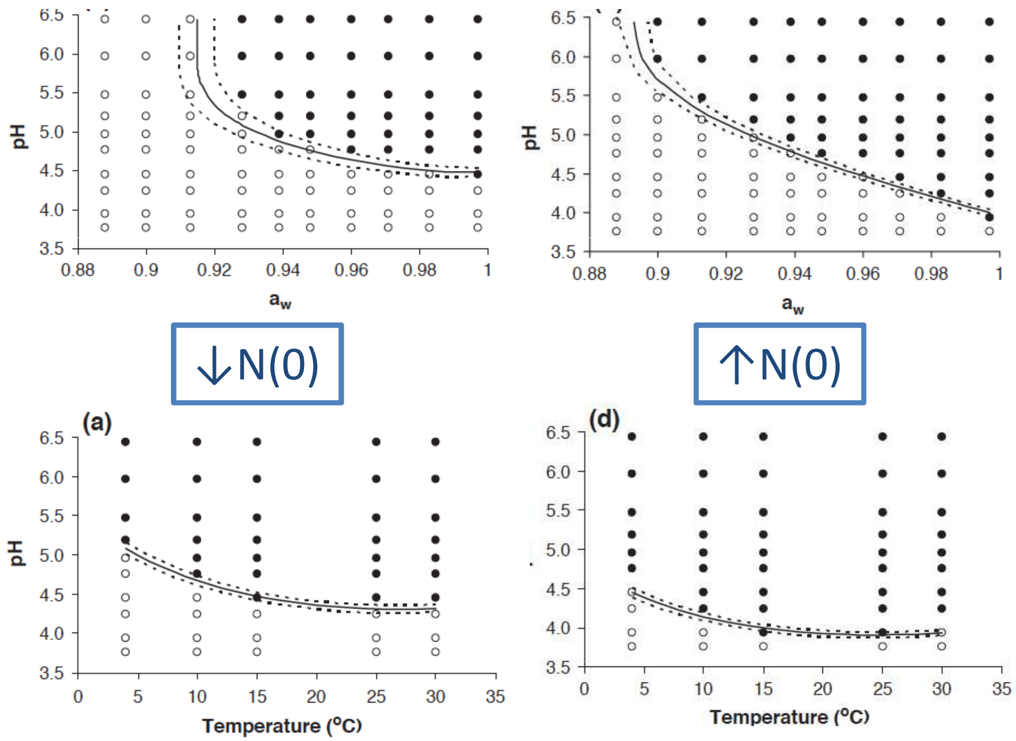
BioTeC

AGRO-INDUSTRIE et MÉTHODES STATISTIQUES

G/NG of *L. monocytogenes* => f(T, pH and a_w)

- ≠ N(0) => different G/NG
- ↓ N(0) => require less harsh conditions for NG
- @↓ N(0): more effect of cell heterogeneity

- For pathogens: ↓ N(0)
- For spoilage: ↑ N(0)



Koutsoumanis & Sofos (2005, IJFM)

Probabilistic models

Data collection & repetitions

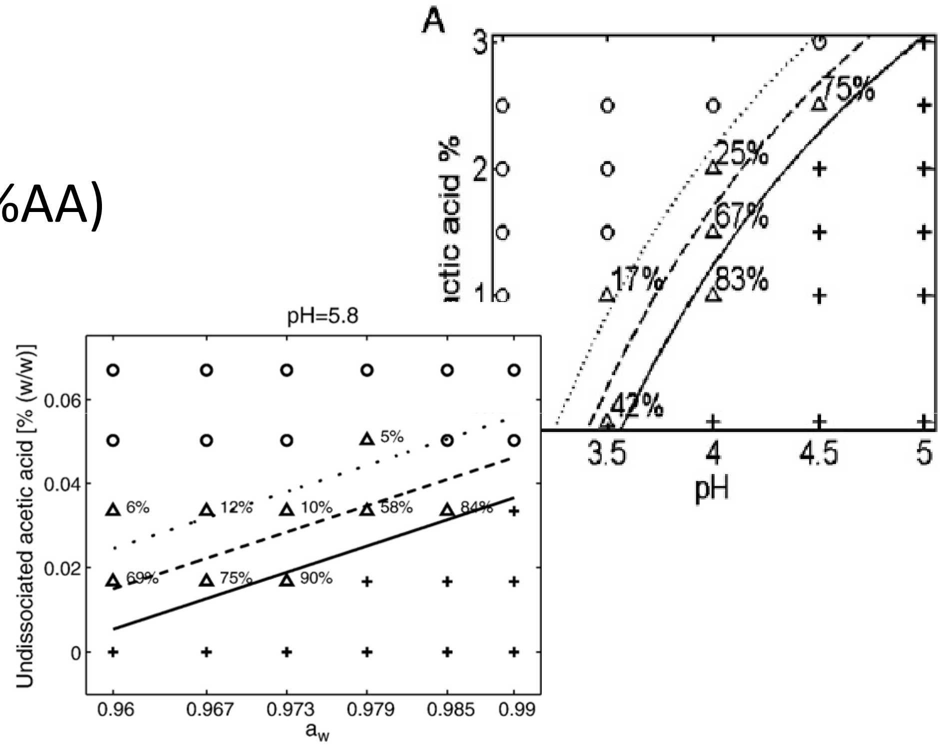
- Number of experimental data & number of repetitions highly determines the model characteristics.
- ↑ # repetitions => a more accurate G/NG boundary

Probabilistic models

Data collection & repetitions

- Vermeulen et al. (2009, IJFM)
  - 12 replicates => 9%
  - *Lactobacillus fructivorans* (0%AA)

- Gysemans et al. (2007, IJFM)
  - 20 replicates => 5%
  - *L. monocytogenes* (pH=5.8)



=> the experimental scheme should be a balance between

- Practical feasibility
- Required accuracy

Probabilistic models

Model complexity

- Logistic regression models = polynomials
  - ↑**# parameters** => can fit any relation
  - Highly sensitive to parameter correlation & **over-fitting**

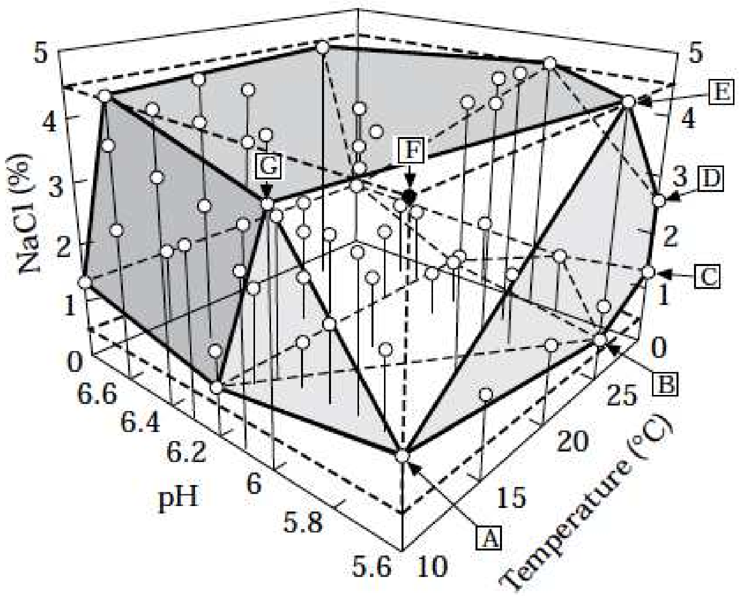
➢ Limit the order of the polynomial
  - Imposing constraints based on microbial knowledge (Geeraerd et al. (2004, IJFM))
  - selection of **significant model parameters**
    - Forward selection
    - Backward elimination
    - Stepwise selection

Probabilistic models

Extrapolation region

- Validity model - range of experimental data
- Straightforward for FFD
- More complex when applying another design
  - Region not easy to identify
  - Can be small
  - ER = minimum convex polyhedron

Baranyi et al. (1996, FM)

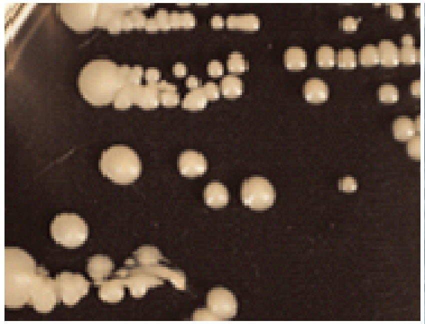


Probabilistic models

Food (model) system limitations

Liquid → **Solid**

IMMOBILIZATION



- *planktonic* growth
- *optimal* transport:
  - nutrient ⟶ cell
  - cell ⟶ metabolites

- growth in *colonies*
- diffusion *limited* transport:
  - local nutrient depletion
  - local metabolite accumulation

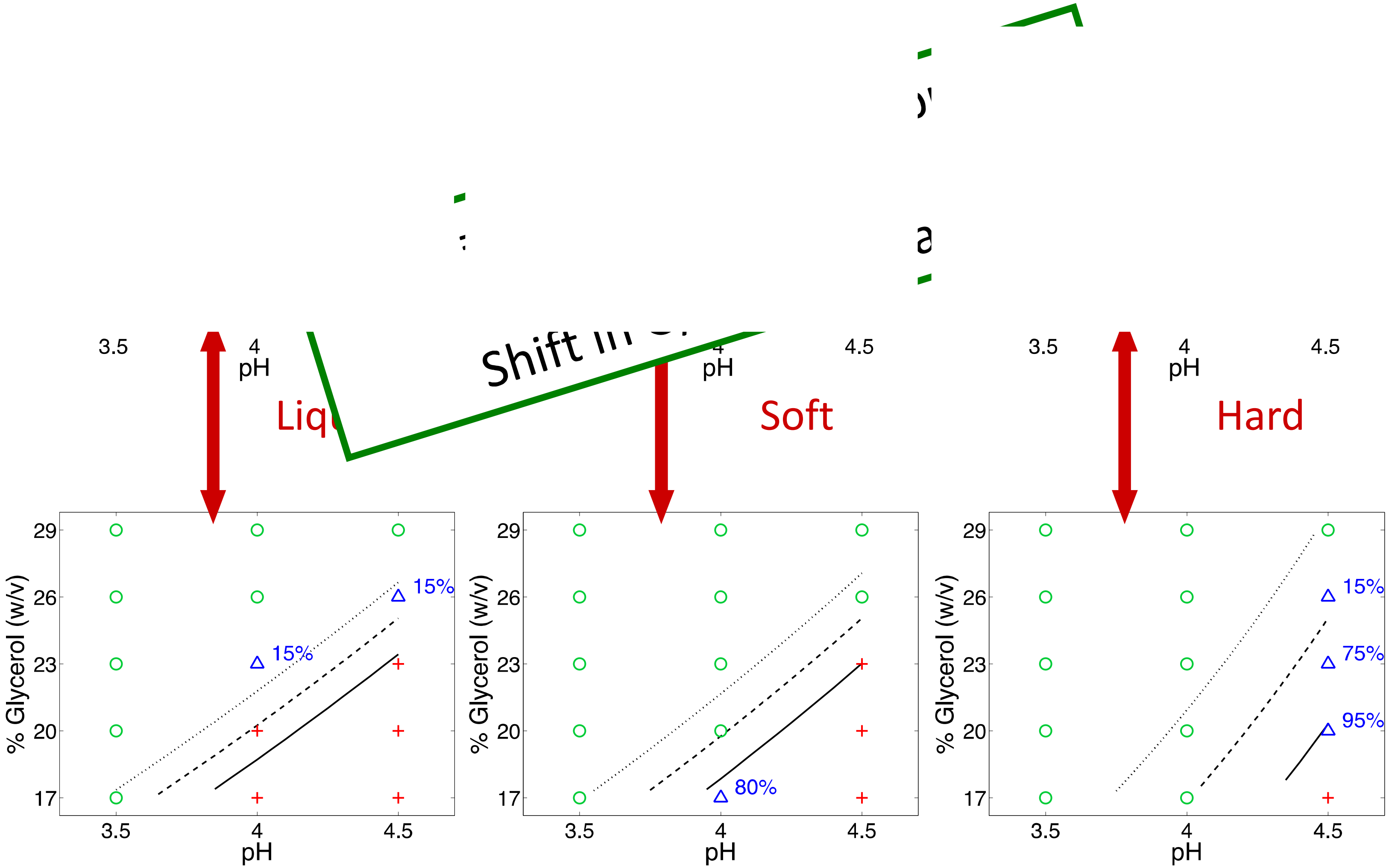
Probabilistic models

Food (model) system limitations

2,0% Acetic acid, 45 days

22°C

30°C



Mertens et al. (2010, FM)

Probabilistic models

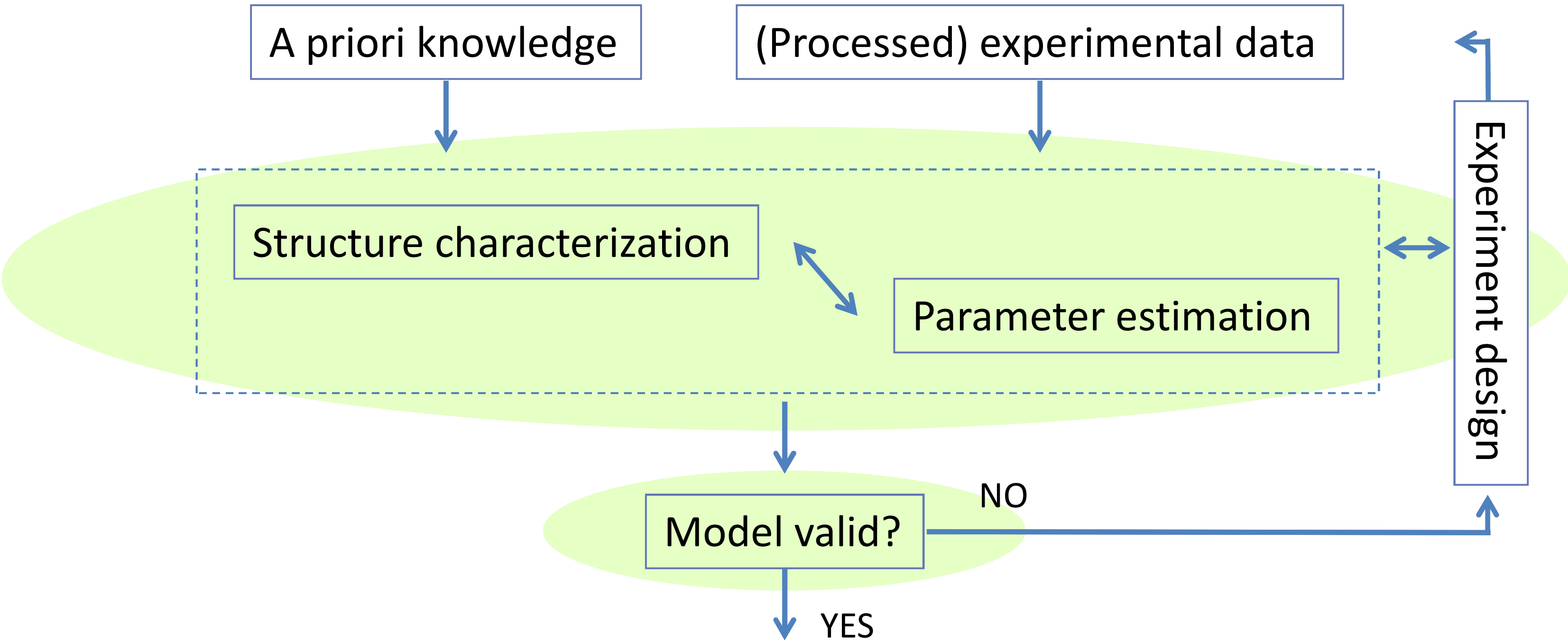
Conclusions

- PB models have a limited transferability
  - Specific strain/cocktail
  - Initial cell level => lower inoculum will require more time to reach the detection limit + different distribution of specific phenomena, e.g., the lag
  - Time span considered => longer times might finally support growth/survival/recovery/toxin production
  - Food product

- Recent publications are limited due to
  - Static character
  - High experimental load
  - Product/case specificity: liquid vs. solid character

Conclusion

Importance of data collection/experimental plan



=> Model building is a data-driven process

Modeling microbial dynamics in food processes

An experiment design approach to predictive microbiology

Eva Van Derlinden, Jan F. Van Impe

BioTeC – Chemical and Biochemical Process Technology and Control

Department of Chemical Engineering, KU Leuven, Belgium