UNIVERSITEIT GENT

BIOMATH

Department of Applied Mathematics, Biometrics and Process Control

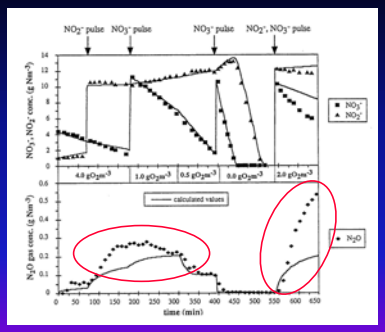
A fully automated, iterative procedure to design and (remotely) perform experiments for accurate estimation of nonlinear model parameters

Peter A. Vanrolleghem and Dirk J.W. De Pauw

Cesame Seminar (February 1, 2005)

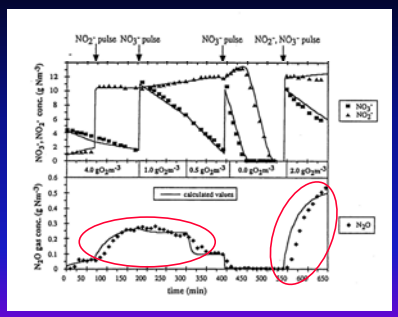
Biomath, Coupure Links 653, 9000 Ghent, Belgium (e-mail dirk.depauw@biomath.ugent.be)

Understanding N_2O emissions



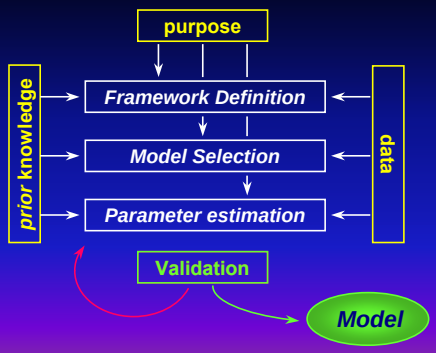
Model is insufficiently "flexible"

Understanding N_2O emissions



Model is much better !

Model building: Subtasks



Outline

- Optimal experimental design (OED)
- Optimization issues and Genetic Algorithms
- An automated procedure
- Practical application to a hybrid respirometer
- A posteriori analysis of results
- Conclusions

Optimal experimental design

- What ?

  A mathematical technique to design an experiment in such a way that it produces data with a maximal information content

- Why ?
  - More accurate parameter estimation
  - Limited resources (time and money)
  - Large # of experimental degrees of freedom

Optimal experimental design

- How ?

Fisher Information Matrix (FIM)
(measure for the information content of the experiment)

$$FIM = \sum_{i=1}^{N} \left(\frac{\partial y}{\partial \theta}\right)_i^T Q_i^{-1} \left(\frac{\partial y}{\partial \theta}\right)_i$$

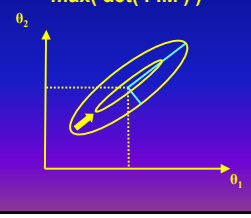
(local) sensitivity functions

measurement error covariance matrix

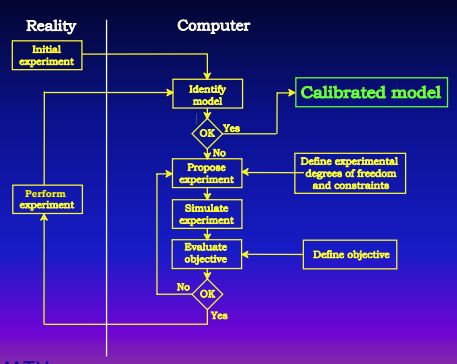
Define objective

Minimizing / maximizing properties of the FIM
➡ **relates to parameter estimation confidence information**

D criterion (volume):
max(det(FIM))



Classical iterative design procedure



(adapted from Dochain and Vanrolleghem, 2001)

UNIVERSITEIT GENT

BIOMATH
Department of Applied Mathematics, Biometrics and Process Control

The use of genetic algorithms in parameter estimation and optimal experimental design problems

Dirk J.W. De Pauw, Michael Rademaker and Peter A. Vanrolleghem

Biomath, Coupure Links 653, 9000 Ghent, Belgium (e-mail dirk.depauw@biomath.ugent.be)

Outline

- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

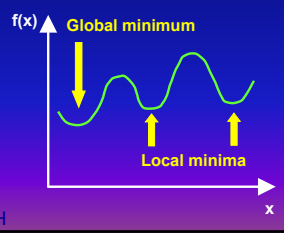
Outline

- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

Classical genetic algorithms

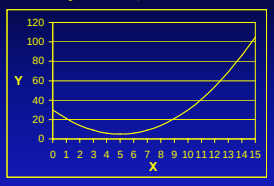
- What ?

Global optimization algorithms inspired on the genetic processes of natural evolution



Problem formulation

- Example: $y = x^2 - 10x + 30$



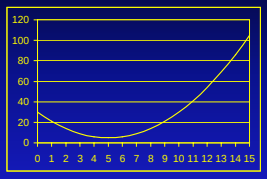
Find integer x which minimizes y

- Encoding:

Problem 0-15 ➡ GA (binary) 0000-1111

Initial population

	x	y
1101	13	69
0001	1	21
0111	7	9
1100	12	54
1111	15	105



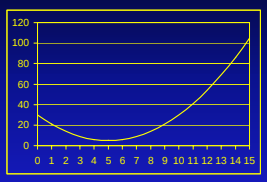
1

Selection, crossover and mutation

	x	y				x	y
1101	13	69					
0001	1	21	➡	0001	➡ 0011	3	9
0111	7	9	➡	0111	➡ 0101	5	5
1100	12	54					
1111	15	105	➡	1111	➡ 0111	7	9

Next generation

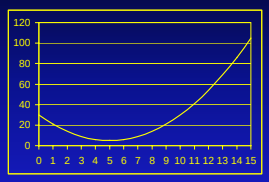
	x	y
0011	3	9
0101	5	5
0111	7	9
1010	10	30
0011	4	6



1
2

Final generation

	x	y
0101	5	5
0101	5	5
0101	5	5
0101	5	5
0101	5	5



1
2
:
N

Binary encoding problems

- Difficulties associated with binary encoding:
  - continuous search spaces
  - large dimension
  - great numerical precision
- Example:
  - [-500, 500]
  - 100 variables
  - 6 digits after the decimal point

➡ chromosome of size 3000

➡ search space $\pm 10^{1000}$

Outline

- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

Real-value genetic algorithms

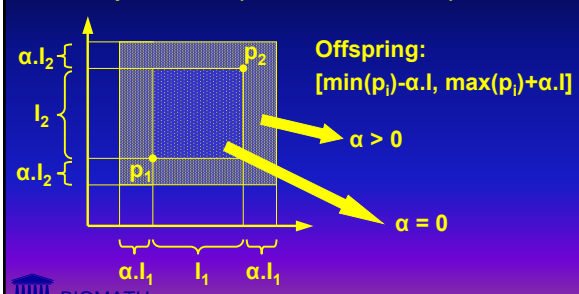
- Directly encode variables as real-values

➡ Each gene represents one variable

- Advantages:
  - large domains
  - high precision / local tuning
  - close to the problem formulation
- Disadvantage:
  - redefine crossover and mutation operators

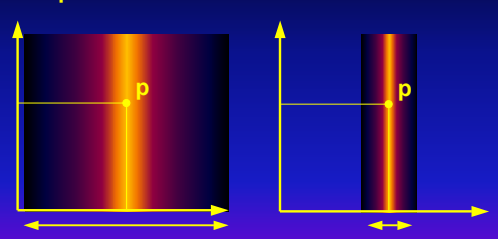
Real-value crossover

- Many operators exist (± 10-20)
- Example: BLX-α (blended crossover)



Real-value mutation

- Also many operators exist (± 10)
- Example: Non-uniform mutation



also adaptive: based on generation number

Practical considerations

- GA tuning:
  - genetic algorithm type
  - genetic algorithm encoding
  - population size
  - choice of crossover / mutation operator
  - crossover / mutation probabilities
  - crossover / mutation tuning parameters
- GA control:
  - change GA settings during execution

➡ More art than science !!!

Outline

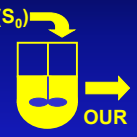
- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

Optimization example

- OUR example:

Batch reactor: S + O --> X

Substrate (S_0)



$$\frac{dS}{dt} = -\frac{1}{Y}\frac{\mu S}{K_S + S}X$$

$$OUR = \frac{1-Y}{Y}\frac{\mu S}{K_S + S}X$$

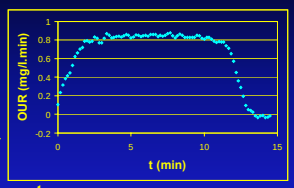
➡ Estimate Monod kinetics parameters μ_{max} and K_S

Optimization example

- OUR example:

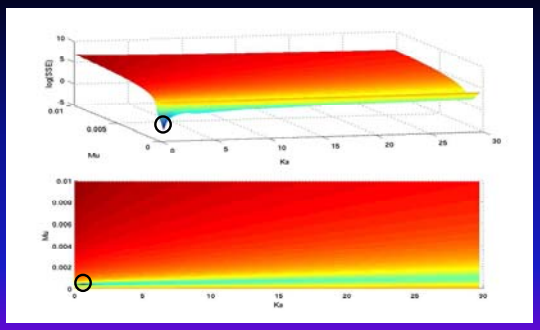
Batch reactor:

Substrate (S_0)



➡ Estimate Monod kinetics parameters μ_{max} and K_S

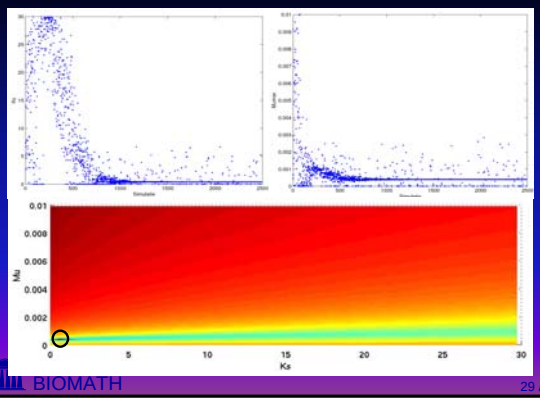
OUR objective surface



Optimum: μ_{max}: 0.00042 min^{-1}

K_S: 0.44 $mg.l^{-1}$

GA performance



Outline

- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

Illustration

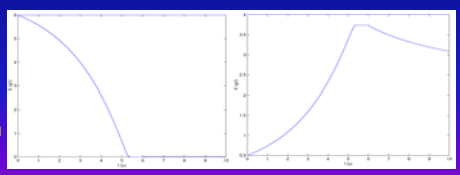
- Fed-batch reactor

Substrate (S_0)
$t = 0$

$t > t_f$: Pump speed Q_f
Feed concentration S_f

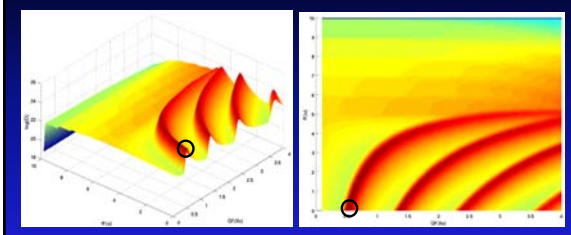
Measure S and X
(interval = 1 h)

$S_0 = 6\ g.l^{-1}$
$t_f = 6\ h$
$Q_f = 4\ l.h^{-1}$
$S_f = 10\ g.l^{-1}$



D-criterion: 1.29E+22

D-criterion objective surface



Optimum: $t_f = 0\ h$
$Q_f = 0.53\ l.h^{-1}$

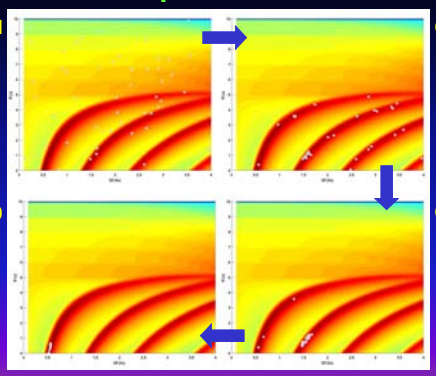
GA performance

Gen.1

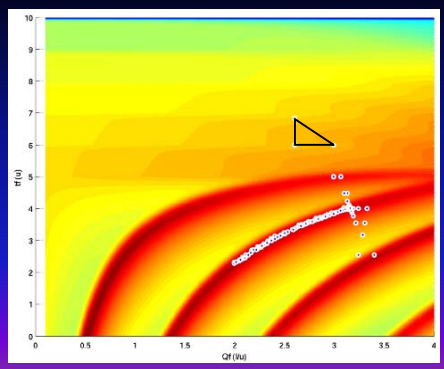
Gen. 9

Gen. 39

Gen. 17



Simplex performance



Outline

- Classical genetic algorithms
- Real-value genetic algorithms
- Use of GA's in calibration
- Use of GA's in experimental design
- Conclusions

Conclusions

- Binary vs. real encoding:
  Real encoding outperforms binary encoding for complex problems
- GA's in general:
  + global search !!!
  + can also be used with discrete variables
  - difficult to tune: case specific
  - large number of objective evaluations
  - slow convergence ➡ hybrid optimization algorithms

UNIVERSITEIT GENT

BIOMATH

Department of Applied Mathematics, Biometrics and Process Control

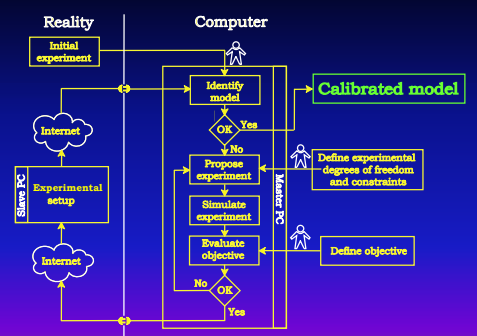
A fully automated, iterative procedure to design and (remotely) perform experiments for accurate estimation of nonlinear model parameters

Peter A. Vanrolleghem and Dirk J.W. De Pauw

Cesame Seminar (February 1, 2005)

Biomath, Coupure Links 653, 9000 Ghent, Belgium (e-mail dirk.depauw@biomath.ugent.be)

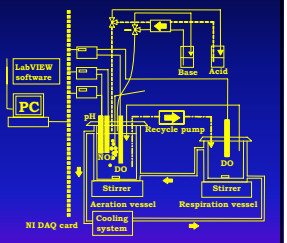
An automated iterative procedure



BIOMATH (adapted from Dochain and Vanrolleghem, 2001) 38 / 25

Experimental setup

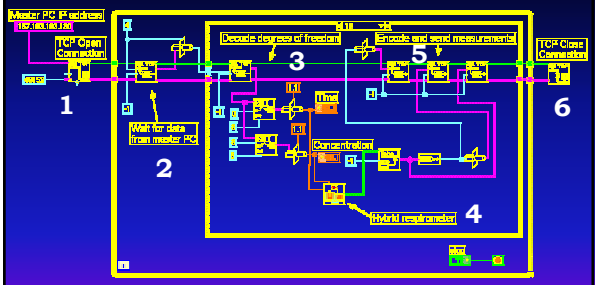
Hybrid respirometer (Sin *et al*. 2003)



DO measurements → Oxygen uptake rate (OUR)

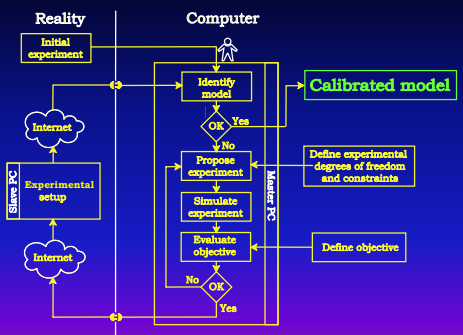
BIOMATH 39 / 25

Communicating with LabVIEW



BIOMATH 40 / 25

An automated iterative procedure

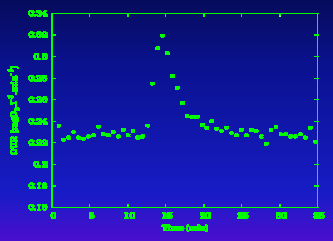


BIOMATH (adapted from Dochain and Vanrolleghem, 2001) 41 / 25

Identify model from initial experiment

- **Initial experiment:**

pulse of 0.5 mg NH_4^+ added to the aeration vessel at 12 minutes after the start of the experiment



BIOMATH 42 / 25

Identify model from initial experiment

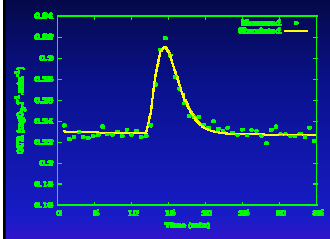
- Model: one-step nitrification

exogenous OUR

$$OUR = \frac{4.57 - Y_{NH4}}{Y_{NH4}} \times \left(1 - e^{-\frac{t - t_{pulse}}{\tau}}\right) \times \mu_{max,NH4} \times \frac{S_{NH4}}{K_{NH4} + S_{NH4}} \times X_{NH4}$$

Identify model from initial experiment

- Model calibration:

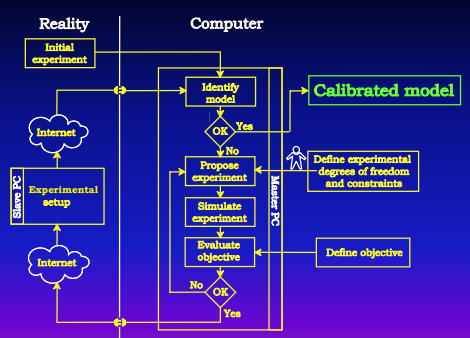


Parameter estimation errors:

Name	Estimation errors (% of parameter value)
$\mu_{max,NH4}$	927
K_{NH4}	5118
Y_{NH4}	4682
τ	106

➡ Optimal experimental design is needed

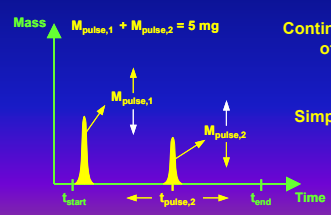
An automated iterative procedure



(adapted from Dochain and Vanrolleghem, 2001)

Define experimental degrees of freedom and constraints

- Two pulse additions of NH_4^+ added by a pump to the aeration vessel
- First pulse always added at start of experiment
- Total added mass = 5 mg NH_4^+ (Grady *et al.*, 1996)
- Total experimental time = 110 minutes

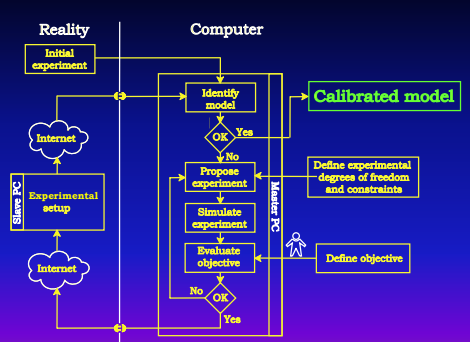


Continuous degrees of freedom

➡

Simplex optimizer

An automated iterative procedure



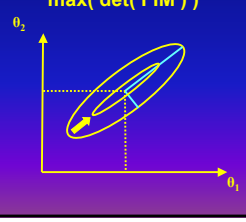
(adapted from Dochain and Vanrolleghem, 2001)

Define objective

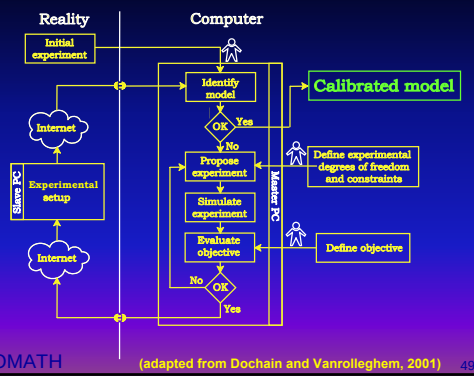
Minimizing / maximizing properties of the FIM

➡ relates to parameter estimation confidence information

D criterion (volume):
max(det(FIM))



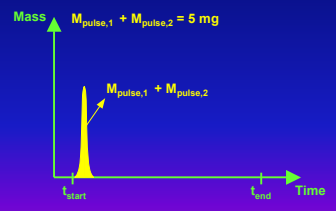
An automated iterative procedure



(adapted from Dochain and Vanrolleghem, 2001)

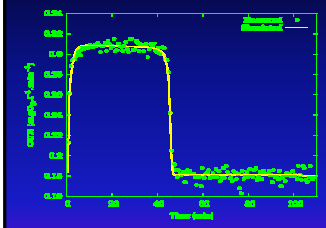
First designed experiment

- **Optimal experiment:**
  - All 5 mg NH_4^+ in first pulse
  - D-criterion value = 2.07E+15



First designed experiment

- **Calibration results:**

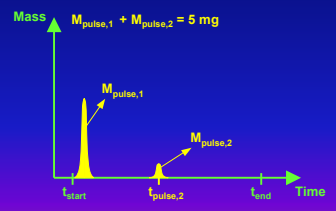


Parameter estimation errors:

Name	Estimation errors (% of parameter value)
$\mu_{max,NH4}$	3.44 [927]
K_{NH4}	21.1 [5118]
Y_{NH4}	3.40 [4682]
τ	10.9 [106]

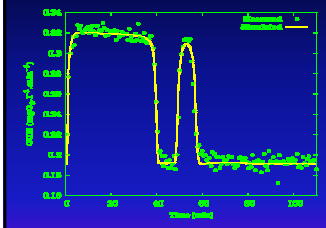
Second designed experiment

- **Optimal experiment:**
  - 4.17 mg NH_4^+ at start and 0.83 mg at 48.32 minutes
  - D-criterion value = 1.26E+23 [2.07E+15]



Second designed experiment

- **Calibration results:**

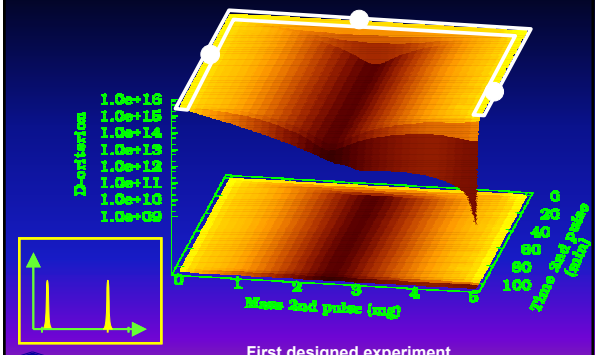


Parameter estimation errors:

Name	Estimation errors (% of parameter value)
$\mu_{max,NH4}$	6.47 [3.44]
K_{NH4}	23.3 [21.1]
Y_{NH4}	6.50 [3.40]
τ	11.6 [10.9]

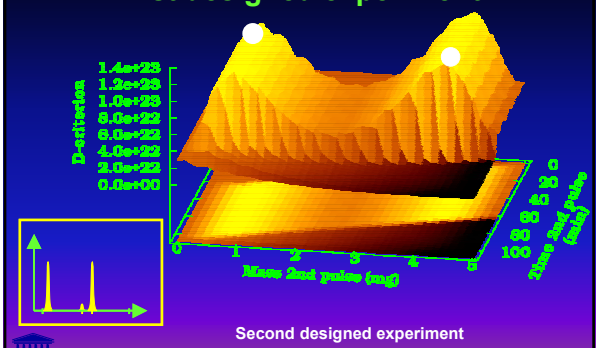
➡ End of experimental design procedure

Objective surface for initial experiment



First designed experiment

Objective surface for first designed experiment



Second designed experiment

Conclusions

- An automated iterative procedure for optimal experimental design was developed and successfully applied.
- The procedure can easily be applied to any experimental setup controlled by a LabVIEW VI.
- The iterative nature of the procedure has shown to be very robust to the nominal parameter values.

Any questions ?