UNIVERSITEIT GENT

BIOMATH
Department of Applied Mathematics, Biometrics and Process Control

Trends in R&D - Module 2: Modelling Biosystem Dynamics

Peter A. Vanrolleghem
December 2000

Trends in Research & Development
(Academic Year 2000-2001)

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

Models in Control...

Definitions

- System
  Part of reality that is separated from its environment on the basis of a purpose defined by the researcher
- Model
  An approximate description of a part of reality considering only those aspects of interest
- Simulation
  = Virtual Experimentation: Manipulation of a model to gain insight in the "behaviour" of the real system

Types of Models

- Mental models (ideas, concepts, ...)
- Verbal models ("description in words")
- Scale models ("house in cardbord")

Scale model (1/72)
of the Lancaster bomber
of Florent Van Rolleghem

Types of Models

- Mental models (ideas, concepts, ...)
- Verbal models ("description in words")
- Scale models ("house in cardbord")
- Computer models ("house in AutoCad")

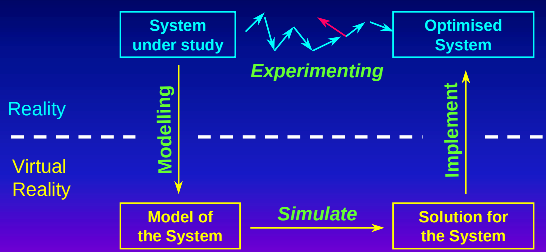
Types of Models

- Mental models (ideas, concepts, ...)
- Verbal models ("description in words")
- Scale models ("house in cardbord")
- Computer models ("house in AutoCad")

- Mathematical models ("equations")

Why Modelling ?

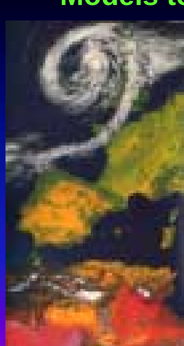
Solving Problems for complex systems



Models to the General Public

$$E = mc^2$$

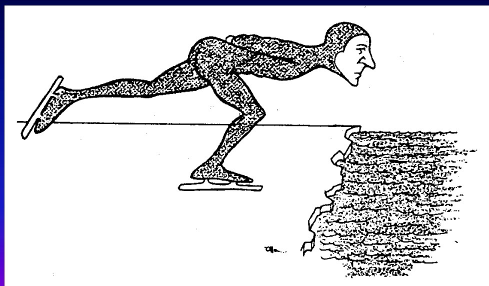
Models to the General Public



Weather Forecasting

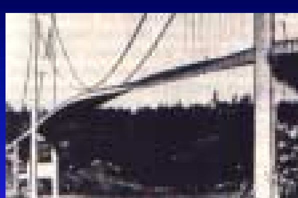
(Reading, UK)

“Do not extrapolate too far with your model”



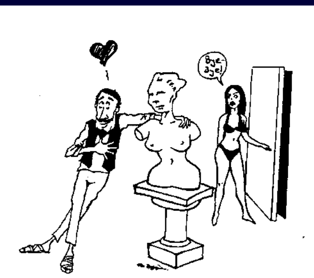
Model (mis)use in Civil Engineering

- Tiny (modelling) errors, significant consequences



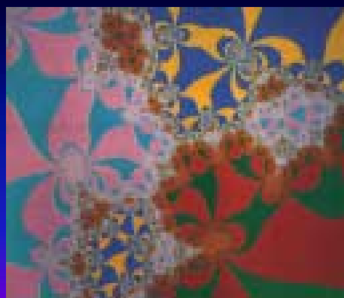
Bridge resonating with wind (+1940)

“Do not fall in love with your model”

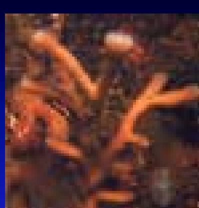


Pygmalion

Fractal models



Fractal models



Growth of “Geweispons” in the Oosterschelde river

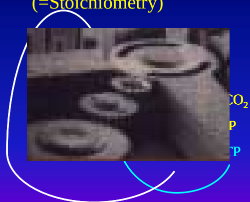
Molecular Models

Final result of folding

Structure from cristalography

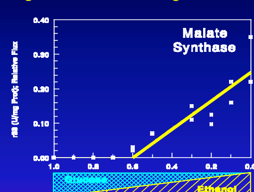
Metabolic Models

Metabolic “gearbox”:
The Boehringer poster:
imposes ratios between reaction rates (=Stoichiometry)



Prediction by Metabolic Models

Saccharomyces cerevisiae (baker’s yeast)
growing on a mixture of glucose and ethanol



Enzymes appear in the cell when metabolism requires them
=> Steering of production possible *(“Metabolic Engineering”)*

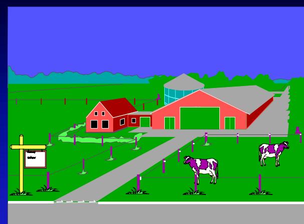
Queuing Models in Food industry

Problem: Minimisation of required storage *(“the queue”)*

e.g.:
- Bottling plant
- Cheese ripening
- Glass washing
- Milk sterilisation
- Sausage fermentation

Models in Agriculture

Model of a dairy farm

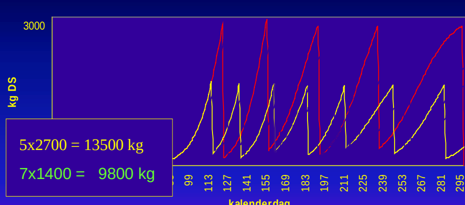


Submodels:

- Milk production
- Fodder intake
- Growth of young cattle
- Growth of grass under mowing or grazing

Question: optimal eco(l/n)o(g/m)ical management?

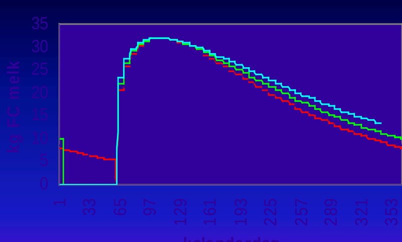
Growth of grass when mowing

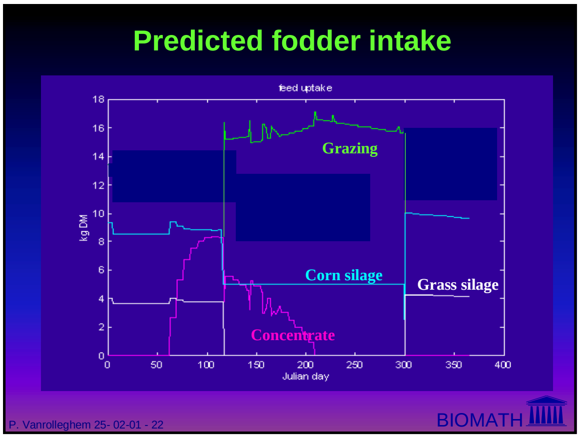


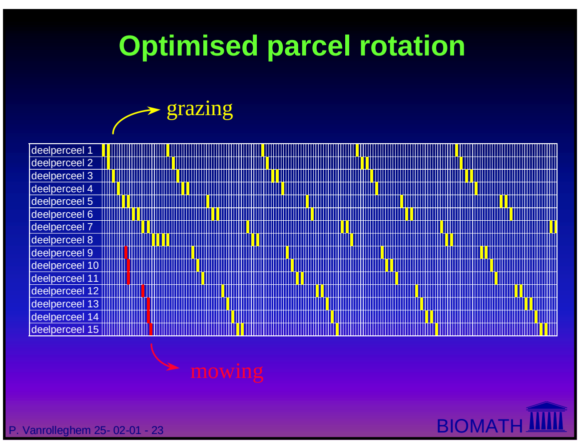
5x2700 = 13500 kg

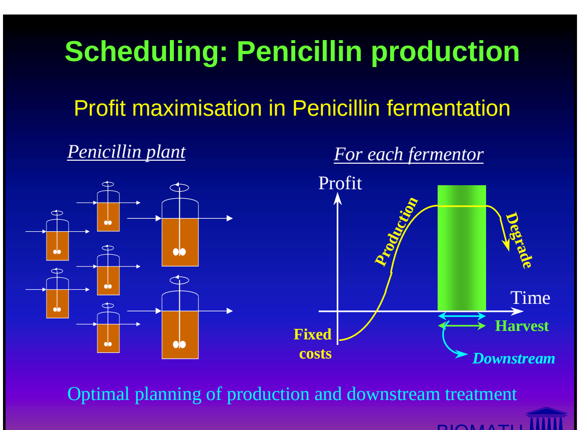
7x1400 = 9800 kg

Model of milk production




Predicted fodder intake
feed uptake
kg DM
Julian day
Grazing
Corn silage
Grass silage
Concentrate
P. Vanrolleghem 25- 02-01 - 22
BIOMATH

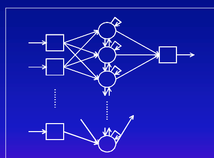

Optimised parcel rotation
grazing
deelperceel 1
deelperceel 2
deelperceel 3
deelperceel 4
deelperceel 5
deelperceel 6
deelperceel 7
deelperceel 8
deelperceel 9
deelperceel 10
deelperceel 11
deelperceel 12
deelperceel 13
deelperceel 14
deelperceel 15
mowing
P. Vanrolleghem 25- 02-01 - 23
BIOMATH


Scheduling: Penicillin production
Profit maximisation in Penicillin fermentation
Penicillin plant
For each fermentor
Profit
Production
Degrade
Time
Harvest
Fixed costs
Downstream
Optimal planning of production and downstream treatment
P. Vanrolleghem 25- 02-01 - 24
BIOMATH

Scheduling: Penicillin production

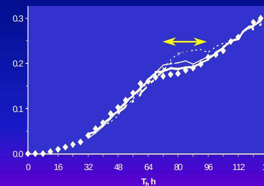
Software sensor for Penicillin concentration

Model:
Artificial Neural Network



"Trained" with data of previous fermentations

8 and 24h ahead predictions
-> 5% accurate



Detection of process upsets

Mathematical Modelling: Overview

- Modelling terminology
  model types, model attributes (strong/weak)
- Model building methodology
  model selection, calibration and validation
- Simulation tools
  solving different types of models, software tools
- Model analysis: Sensitivity and uncertainty
  types of uncertainty, uncertainty propagation

Modelling terminology

- System
  Part of reality that is separated from its environment on the basis of a purpose defined by the researcher
- Model
  An approximate description of a part of reality considering only those aspects of interest
- Simulation
  = *Virtual Experimentation: Manipulation of a model to gain insight in the "behaviour" of the real system*

Model structure

- Model = machine transforming inputs to outputs by a set of equations

  output = variable I'm interested in
  input = any variable that affects the output

- transfer function models between inputs/outputs
  = input/output models
- state space models
  state variables = minimum set of "help" variables

Model constituents

- Variables = constituent whose value changes with the independent variables (time, space)
- Constants = constituent that has always the same value
- Parameters = constituent that may change its value for different applications

Data types

- Deterministic data "the value is …"
- Stochastic data "probability of taking that value is…"
- Fuzzy data "a little this, a little that value"

Model building: Starting points

- Purpose of the model
  - Increasing understanding of a system *(Think tank)*
  - Summary of knowledge/data *(Communication)*
  - Prediction of future behaviour *(Control, Design)*
- *Prior* knowledge
  - Experience
  - Existing models
  - Literature (facts, phenomena, theories, ...)
- Data
  - Existing data
  - New data collected in view of model building

Model attributes

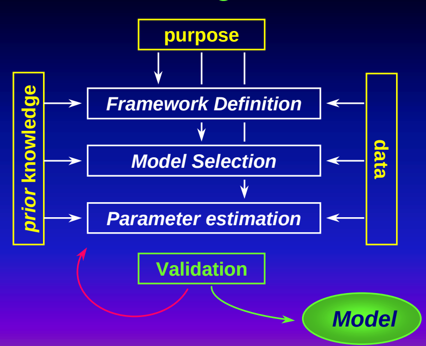
Strong attributes

linear	non-linear
dynamic	static
time-invariant	time-variant
distributed parameter	lumped parameter
discrete time	continuous time
discrete space	continuous space
stochastic	deterministic

Weak attributes

simple	complex
mechanistic	phenomenological

Model building: Subtasks



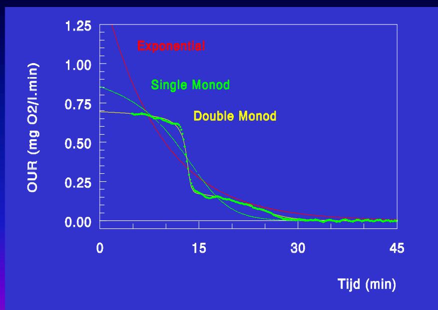
Model selection

- Compare fit quality: (SSR=Sum of squared errors)

$$F_w = \frac{\left(\dfrac{SSR(M_j) - SSR(M_i)}{\dim(\theta_j) - \dim(\theta_i)}\right)}{\dfrac{SSR(M_j)}{Ndata - \dim(\theta_j)}}$$

and compare with F-value to determine significance

Model Selection



BIOMATH

Simulation tools

- Ordinary differential equations (ODE's)

$$\frac{dy_i}{dt} = f_i(t, y_1, \ldots, y_N, x_1, \ldots, x_m) \qquad i=1,N$$

  - Finite difference approximation (Euler):

$$\frac{dy_i}{dt} \cong \frac{y_i(t+h) - y_i(t)}{h} = f_i(t, y_1, \ldots, y_N, x_1, \ldots, x_m)$$
$$\Rightarrow y_i(t+h) = h \times f_i(t, y_1, \ldots, y_N, x_1, \ldots, x_m) + y_i(t)$$

  - initial values must be given
  - sometimes only a boundary value is available ==> "shooting method"

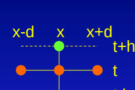
BIOMATH

Simulation Tools

- Partial Differential Equations (PDE's)

$$\frac{\partial u}{\partial t} = -v\frac{\partial u}{\partial x} + D\frac{\partial^2 u}{\partial x^2}$$

  - Finite difference approximation :
    - 3 boundary conditions necessary



$$\frac{u(t+h,x) - u(t,x)}{h} = -v\frac{u(t,x) - u(t,x+d)}{d} + D\frac{u(t,x-d) - 2u(t,x) + u(t,x+d)}{d^2}$$

BIOMATH

Software Tools

- General purpose (open) simulation software
  - solution methods are available and models can be entered at will
  - Excel, Matlab/Simulink, Mathematica, MathCad
  - for WWTP: GPS-X, Simba, WEST
  - for Rivers: Aquasim, Duflow
- Closed model structures
  - specific models are implemented together with their solution methods
  - for WWTP: BioWin, Stoat, EFOR
  - for Sewers: Hydroworks, Mouse
  - for Rivers: ISIS, Mike11
- Software libraries
  - subroutines, procedures that can be combined at will
  - Fortran, Pascal, C++

Sensitivity Analysis

Evaluate the magnitude of change of a variable as a result of a change in another variable or parameter

- Mathematically:

$$Sensitivity = \frac{dY}{d\theta}$$

- when model is specified: $Y=f(x,Y,\theta)$

=> $$Sensitivity = \frac{dY}{d\theta} = \frac{df(x,Y,\theta)}{d\theta}$$

- numerically: $$Sensitivity = \frac{Y(\theta + d\theta) - Y(\theta)}{d\theta}$$

Uncertainty

- Errors may appear in the calculations because
  - data quality is not good enough
  - model structure is wrong
  - information is not sufficient to estimate model parameters
  - calculation procedures are not accurate
- Uncertainty in the outputs is a combination of:
  - uncertainty on input variables
  - uncertainty in the model structure
  - uncertainty in the model parameters
  - mathematical uncertainty

Uncertainty propagation

- Linear model:

$$N_{tot} = N_{NH4} + N_{org}$$

linear uncertainty propagation rule:

$$\sigma_{tot}^2 = \sigma_{NH4}^2 + \sigma_{org}^2 + 2 \cdot \rho \cdot \sigma_{NH4} \sigma_{org}$$

Uncertainty propagation

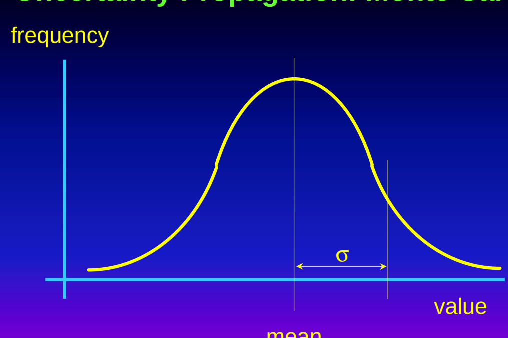
- Nonlinear model --> linearise (Taylor series expansion)

$$r_{nit}(Y_A, \mu_A, K_{NH}, K_O) \approx r_{nit}(Y_{A,0}, \mu_{A,0}, K_{NH,0}, K_{O,0}) + \frac{dr_{nit}}{dY_{A,0}}(Y_A - Y_{A,0}) + \frac{dr_{nit}}{d\mu_{A,0}}(\mu_A - \mu_{A,0}) + \frac{dr_{nit}}{dK_{NH,0}}(K_{NH} - K_{NH,0}) + \frac{dr_{nit}}{dK_{O,0}}(K_O - K_{O,0})$$

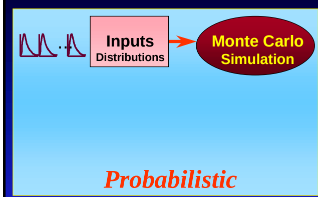
- now apply the linear uncertainty propagation rule:

$$\sigma_{nit}^2 \approx \frac{dr_{nit}}{dY_{A,0}} \sigma_{Y_A}^2 + \frac{dr_{nit}}{d\mu_{A,0}} \sigma_{\mu_A}^2 + 2 \frac{dr_{nit}}{d\mu_{A,0}} \cdot \frac{dr_{nit}}{dY_{A,0}} \cdot \rho_{Y_A \mu_A} \cdot \sigma_{Y_A} \cdot \sigma_{\mu_A} + \frac{dr_{nit}}{dK_{NH,0}} \sigma_{K_{NH}}^2 + \frac{dr_{nit}}{dK_{O,0}} \sigma_{K_O}^2 + \ldots$$

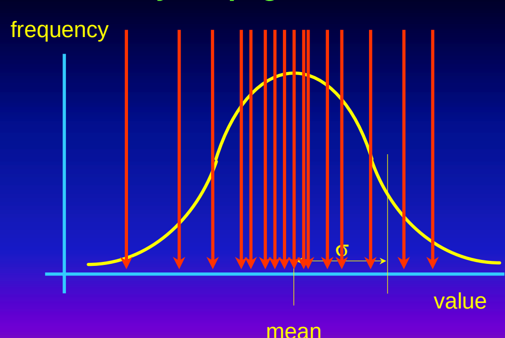
Uncertainty Propagation: Monte Carlo



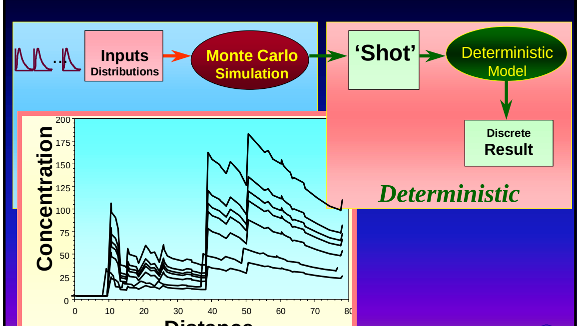
Uncertainty Propagation: Monte Carlo

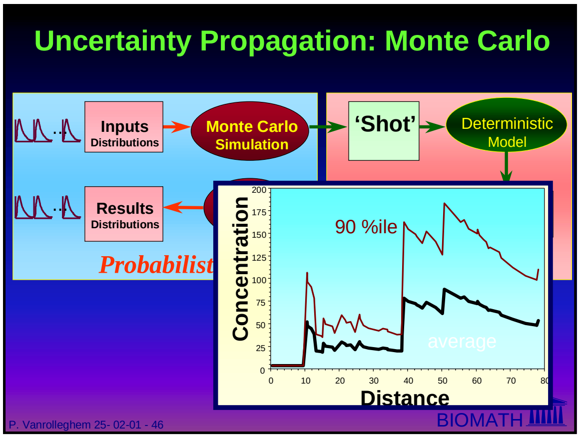


Uncertainty Propagation: Monte Carlo



Uncertainty Propagation: Monte Carlo




Uncertainty Propagation: Monte Carlo
Inputs
Distributions
Monte Carlo
Simulation
'Shot'
Deterministic
Model
Results
Distributions
Probabilist
90 %ile
average
Concentration
Distance
0
25
50
75
100
125
150
175
200
0
10
20
30
40
50
60
70
80
P. Vanrolleghem 25- 02-01 - 46
BIOMATH