

## EVOLUTIONARY STRATEGY IN ITERATIVE EXPERIMENT DESIGN

J. MADÁR, B. BALASKÓ, F. SZEIFERT AND J. ABONYI\*

Department of Process Engineering, University of Veszprém,  
Veszprém, Egyetem u. 10, H-8200, HUNGARY,  
www.fmt.vein.hu/softcomp, abonyij@fmt.vein.hu

Process models play important role in computer aided process engineering, since most of advanced process monitoring, control, and optimization algorithms rely on a model of the process. In most of the cases, some parameters of the model should be estimated based on some experiments. One of the factors affecting the model prediction quality is the accuracy of these estimated parameters. Establishing *optimal experiment design* can maximise the confidence on the parameters, hereby increasing the confidence on the model prediction. The aim of this paper is to work out a modern experiment design tool to minimize the number of experiments while maximizing of their information content. This paper illustrates the applicability of ES for the design of feeding profile for a fed-batch biochemical reactor. The results illustrate that if the model structure is not accurate, the evolutionary strategy can result in more satisfactory parameter values than the classical sequential quadratic programming and nonlinear least squares algorithms.

**Keywords:** Experiment design, model identification, fed-batch bioreactor

### Introduction

Process models play important role in computer aided process engineering since most of advanced process monitoring, control, and optimization algorithms rely on a model of the process. Unfortunately often some of the parameters of these models are not known *a priori*, so they must be estimated from experimental data.

The accuracy of these parameters largely depends on the information content of the experimental data presented to the parameter identification algorithm [1]. Establishing *optimal experiment design* (OED) can maximise the confidence on the parameters. For the identification of the parameters of dynamic models this approach has been utilized in [2-5]. In these studies experiment design is concerned with the following questions: How does one adjust time-varying controls, initial conditions, and/or other design parameters of the experiments to generate the maximum amount of information for the purpose of estimating the parameters with greatest precision.

For nonlinear models, OED is based on an iterative algorithm because the optimal parameters of the experiments depend on the model parameters that are going to be estimated based on the result of the designed experiment. Consequently, OED estimates the model parameters and designs the experiment iteratively. Both parameter estimation and experiment design are based on nonlinear optimization of certain cost-functions. In practice, the applied nonlinear optimization algorithms have great influence on the whole procedure, because for nonlinear dynamical

models the design of the experiment is a hard optimization problem. As an effective optimization algorithm, this paper proposes the application of *evolutionary strategy* (ES) for this purpose. ES is a stochastic optimization algorithm that uses the model of natural selection [7].

In this paper, OED are applied for fed-batch biochemical reactor. One of the factors affecting the modelling of biochemical systems is that accurate description of biochemical reaction is generally not available a priori. Hence usually a simplified kinetic model, e.g. Monod model, is used to describe the microbial dynamics. Some results were presented for experiment design of biochemical systems in [4-6], but these works assumes that the model structure is perfectly known. This paper discusses the application of OED based on models that have structural uncertainty. Our results illustrate that although the model structure used for the design of the experiments is not accurate, OED with ES can result in satisfactory parameter values.

The paper organized as follows: The first section reviews the theory of optimal experiment design. The second section proposes the application of evolutionary strategy for OED. The third section presents the application example. Finally, conclusions are given in the fourth section.

### Optimal input design for parameter estimation

The case study considered in this paper belongs to the following general class of process models:

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{f}(\mathbf{x}(t), u(t), \mathbf{p}) \\ \mathbf{y}(t) &= \mathbf{g}(\mathbf{x}(t)) \end{aligned} \quad (1)$$

where  $u$  is the manipulated input,  $\mathbf{y}$  is the output (vector),  $\mathbf{x}$  is the state (vector) of the system and  $\mathbf{p}$  denotes the unknown model parameters. The  $\mathbf{p}$  parameters are unknown and should be estimated based on data taken from experiments. For the estimation of these parameters classical parameter identification approach is used that is based on the minimization of the square error between the output of the system and the output of the model:

$$\min_{\mathbf{p}} J_{mse}(u(t), \mathbf{p}) \quad (2)$$

where

$$J_{mse}(u(t), \mathbf{p}) = \frac{1}{t_f} \int_{t=0}^{t_f} (\mathbf{e}^T(t) \cdot \mathbf{Q}(t) \cdot \mathbf{e}(t)) dt \quad (3)$$

$$\mathbf{e}(t) = \tilde{\mathbf{y}}(u(t)) - \mathbf{y}(u(t), \mathbf{p})$$

in which  $\tilde{\mathbf{y}}$  is the output of the system for a certain  $u(t)$  input profile, and  $\mathbf{y}$  is the output of the model for the same  $u(t)$  input profile with  $\mathbf{p}$  parameters,  $\mathbf{Q}$  is a user supplied square weighting matrix that represents the (variance of the) measurement error.

The accuracy of the parameter estimation depends on the applied  $u(t)$  input profile. The goal of the experiment design is to determine an optimal input profile in the sense of the parameter estimation leads to optimal parameters with maximal confidence.

The basic element of the experiment design methodology is the Fisher information matrix  $\mathbf{F}$ , which combines information on (i) the output measurement error and (ii) the sensitivity of the model output with respect to the model parameters:

$$\mathbf{F}(\mathbf{p}^0) = \frac{1}{t_f} \int_{t=0}^{t_f} \left( \frac{\partial \mathbf{y}}{\partial \mathbf{p}}(t, \mathbf{p}) \Big|_{\mathbf{p}=\mathbf{p}^0} \right)^T \cdot \mathbf{Q}(t) \cdot \left( \frac{\partial \mathbf{y}}{\partial \mathbf{p}}(t, \mathbf{p}) \Big|_{\mathbf{p}=\mathbf{p}^0} \right) dt \quad (4)$$

in which  $\mathbf{p}^0$  is the nominal parameter vector. The Fisher information matrix  $\mathbf{F}$  provides an approximate quantification of the attainable parameter estimation quality in the neighbourhood of the nominal parameter vector  $\mathbf{p}^0$  for a particular experiment as the inverse of matrix  $\mathbf{F}$  approximates the parameter estimation covariance matrix.

The optimal design criterion aims the minimization of a scalar function of the  $\mathbf{F}$  matrix. Among the existing criteria, the D-optimal criterion and the modified E-optimal criterion suggested by Bernaerts *et al.* [1] are studied in this work.

i. The D-optimal criterion minimizes the determinant of the covariance matrix, and thus minimizes the volume of the joint confidence region:

$$\begin{aligned} \max_{u(t)} J_D(\mathbf{F}) \\ J_D = \det(\mathbf{F}) \end{aligned} \quad (5)$$

ii. E-optimal criterion minimizes the condition number of  $\mathbf{F}$ , i.e. the ratio of the largest to the smallest eigenvalue of  $\mathbf{F}$ :

$$\begin{aligned} \min_{u(t)} J_E(\mathbf{F}) \\ J_E(\mathbf{F}) = \frac{\lambda_{\max}(\mathbf{F})}{\lambda_{\min}(\mathbf{F})} \end{aligned} \quad (6)$$

These values correspond to the uncertainty of the parameter estimation problem. Fig. 1 and 2 illustrate the effect of the input profile on the model output in case of the case study that will be presented later in the Application Example session. The contour plots show the square error of model output with respect to its parameters around the  $\mathbf{p}^0$  nominal parameters:

$$J_i(\mathbf{p}) = \frac{1}{t_f} \int_{t=0}^{t_f} (y(t, \mathbf{p}^0) - y(t, \mathbf{p}))^2 dt \quad (7)$$

One can see that when an E-optimal input profile is used, then the parameter uncertainty region is smaller. This means that, if the  $\mathbf{p}^0$  nominal parameters are close to the optimal  $\mathbf{p}^*$  parameters, the parameter estimation based on this profile (Fig. 1) most likely results in accurate parameters than the estimation based on a manually selected profile (Fig. 2).

These figures suggest that when one has only a draft estimate on the parameters of a complex dynamical model, he/she should use it to design an  $u(t)$  input profile for a parameter estimation procedure (eq. 2-3) rather than to use data taken from a non-optimized input profile for the identification.

In the following session a new optimization algorithm will be presented for the effective design of the experiments.

For nonlinear models, OED results in an iterative procedure due to the fact that the parameters of the designed experiment depends on the model parameters itself, see Fig. 3.

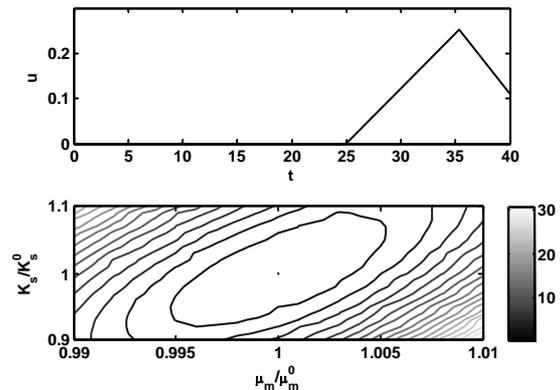


Fig. 1 Contour plots of the identification cost ( $J_i$ ) with respect to parameters for an E-optimized feeding profile

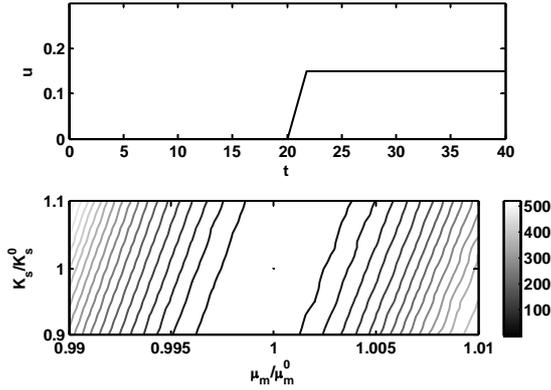


Fig. 2 Contour plots of the identification cost ( $J_i$ ) with respect to parameters for a manually selected feeding profile

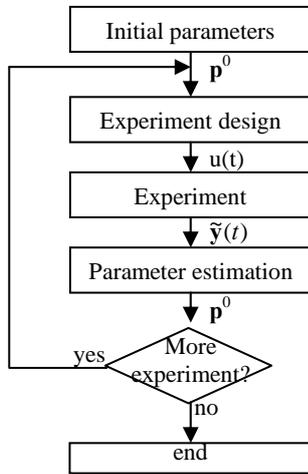


Fig. 3 Scheme of parameter estimation with OED

Both the parameter estimation and the experiment design steps of this iterative scheme represent a complex nonlinear optimization problem, hence the effectiveness of the applied optimization algorithms have great influence on the performance of the whole procedure. The classical solution is to use *nonlinear least squares* (NLS) algorithm for parameter estimation eq. 2-3, and *sequential quadratic programming* (SQP) for the experiment design eq. 5 or eq. 6.

### Evolutionary Strategy

This paper proposes the application of *evolutionary strategy* (ES) instead of the utilization of NLS and SQP. ES is a stochastic optimization algorithm that uses the model of natural selection. The advantage of ES is that it has proved particularly successful in problems that are highly nonlinear, that are stochastic, and that are poorly understood [7].

Evolution strategy is the member of evolutionary algorithms. The design variables in ES are represented by  $n$ -dimensional vector  $\mathbf{x}_j = [x_{j,1}, x_{j,2}, \dots, x_{j,n}]^T$ , where  $\mathbf{x}_j$  represents the  $j$ th potential solution, i.e. the  $j$ th member of the population. The mutation operator adds

$z_{j,i}$  normal distributed random numbers to the design variables:  $x_{j,i} = x_{j,i} + z_{j,i}$ , where  $z_{j,i} = N(0, \sigma_{j,i})$  is a random number with  $\sigma_{j,i}$  standard deviation. To allow better adaptation to the objective function's topology, the design variables are accompanied by these standard deviation variables which are so-called strategy parameters. Hence the  $\sigma_j$  strategy variables control the step size of standard deviations in the mutation for  $j$ th individual. So an ES-individual  $\mathbf{a}_j = (\mathbf{x}_j, \boldsymbol{\sigma}_j)$  consist of two components: the design variables and the strategy variables.

Before the design variables are changed by mutation operator, the standard deviations  $\sigma_j$  are mutated using a multiplicative normally distributed process:

$$\sigma_{j,i}^{(t)} = \sigma_{j,i}^{(t-1)} \exp(\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)) \quad (8)$$

The  $\exp(\tau' \cdot N(0,1))$  is a global factor which allows an overall change of the mutability, and the  $\exp(\tau \cdot N_i(0,1))$  allows individuals to change of their mean step sizes  $\sigma_{j,i}$ . So  $\tau'$  and  $\tau$  parameters can be interpreted as global learning rates. Schwefel suggests setting them as [8]:

$$\tau' = \frac{1}{\sqrt{2n}}, \quad \tau = \frac{1}{\sqrt{2}\sqrt{n}} \quad (9)$$

Throughout this work discrete recombination of the object variables and intermediate recombination of the strategy parameters were used:

$$\begin{aligned} x_{j,i} &= x_{F,i} \text{ or } x_{M,i} \\ \sigma_{j,i} &= (\sigma_{F,i} + \sigma_{M,i})/2 \end{aligned} \quad (10)$$

where  $F$  and  $M$  denotes the parents,  $j$  is the index of the new offspring.

### Application Example

The case study of this paper is a fed-batch bioreactor with non-monotonic kinetics [6]. The following equation describes the mass balance of the reactor:

$$\frac{d}{dt} \begin{bmatrix} S \\ X \\ V \end{bmatrix} = \begin{bmatrix} -\sigma X \\ \mu X \\ 0 \end{bmatrix} + \begin{bmatrix} C_{S,in} \\ 0 \\ 1 \end{bmatrix} u \quad (11)$$

where  $S$  is the mass of the substrate [g],  $X$  is the mass of the micro-organism [g DW],  $V$  is the volume [L],  $u$  is the inlet flowrate [L/h],  $C_{S,in} = 500$  g/L is the substrate concentration in the inlet feed,  $\sigma = \mu/Y_{XS} + m$  is the specific substrate consumption rate, where,  $Y_{XS} = 0.47$  g DW/g,  $m = 0.29$  g/g DW h, while  $\mu$  [1/h] is the kinetic rate. The initial conditions:  $S(t=0) = 500$  g,  $X(t=0) = 10.5$  g DW,  $V(t=0) = 7$  L. The maximum volume is  $V_{max} = 10$  L, and the maximum inlet flowrate is  $u_{max} = 0.3$  L/h.

Two kinetic models of the  $\mu$  kinetic rate were considered:

*Monotonic kinetic (Monode model):*

$$\mu^M(C_S) = \mu_{max} \frac{C_S}{K_S + C_S} \quad (12)$$

*Non-monotonic kinetic (Haldane model):*

$$\mu^H(C_S) = \mu_m \frac{C_S}{K_P + C_S + C_S^2/K_I} \quad (13)$$

where  $C_S = S/V$ .

The majority of the OED applications in the literature assume that the structure of the model used for the design of the experiment is perfectly known. However, the model structure is often inaccurate in practice. For example, a simplified kinetic model is often used to describe the microbial dynamics due to the lack of accurate knowledge of the microbial dynamics.

The purpose of this study is to illustrate the complications that may arise from this *structural uncertainty*. Therefore, in this simulated example, we use different kinetic models during the simulation of the system (considered as a real, unknown process) and in the model itself. It is assumed that the ‘true’ process can be described by a non-monotonic kinetic equation (eq. 13), while the model contains a monotonic kinetic model (eq. 12).

The system was simulated with  $\mu_m = 0.1$  1/h,  $K_p = 1$  g/L and  $K_I = 500$  g/L. The goal was to find the unknown parameters of the model:  $\mu_{max}$  and  $K_S$  (the other parameters were assumed to be accurately known). The  $\mu_{max}$  and  $K_S$  parameters were estimated by optimization, see eq. 2 and 3. Because the measurement of the micro-organism concentration  $C_X$  is quite difficult in practice, it was assumed that only the substrate concentration  $C_S$  is measured. Consequently, in this application example, the system output is  $\tilde{y} = \tilde{S}$  [g], which was generated by the simulation of equations (11) and (12), and the model output is  $y = S$  [g], which it was calculated with the use of equations (11) and (13).

We applied the iterative OED methodology to design the feeding profile  $u(t)$  with E-optimal criterion equations (4) and (6). The D-optimal criterion was not used, because our experience showed that the E-optimality is better suited for this problem. Because the number of experiments and the length of experiments are limited in practice, the number of iterations was limited to 3 and the length of one experiment was 40 h. The sample time was 4 h.

In this example, three methods were examined:

*Method 1:* Manually selected input profiles. To compare and analyze the effectiveness of OED, firstly the parameters were estimated with two feeding profiles. The first profile was:  $u(t) = 0.077$  L/h,  $0 < t < 40$  h, while the second profile was:  $u(t) = 0$  L/h,  $t < 20$  h and  $u(t) = 0.15$  L/h,  $20 < t < 40$  h.

*Method 2:* Iterative OED with NLS in the parameter estimation step and SQP in the experiment design step.

*Method 3:* Iterative OED with ES in both of the parameter estimation and experiment design steps. Because the result depends on the initial parameter estimation, two initial estimations were applied:  $\mu_{max}^{init} = 0.05$  1/h,  $K_S^{init} = 0.5$  g/L and  $\mu_{max}^{init} = 0.1$  1/h,  $K_S^{init} = 1$  g/L.

Certainly, there is no ‘perfect’ solution for this problem. To analyze the results, the obtained monotonic

$\mu^M(C_S)$  functions were compared to the ‘true’ non-monotonic  $\mu^H(C_S)$  function:

$$E_{mse} = \frac{1}{50} \int_{C_S=0}^{50} (\mu^M(C_S) - \mu^H(C_S))^2 dC_S \quad (14)$$

As Table 1 and Fig. 4 show, *Method 3* proved to be the best, it found relatively good solutions. One can see that *Method 2* was sensitive to the initial parameter estimation. If the initial parameter were  $\mu_{max}^{init} = 0.05$  1/h,  $K_S^{init} = 0.5$  g/L this method got stuck into a local minima and resulted in rather wrong parameter values. If  $\mu_{max}^{init} = 0.1$  1/h,  $K_S^{init} = 1$  g/L initial estimations were used, *Method 2* resulted in better solution. In contrast, *Method 3* always resulted in a relatively good solution independently the initial parameter values. Because ES is a stochastic optimization algorithm six independent runs were performed and we got very similar parameters in each case.

Fig. 5 and 6 demonstrate that the obtained models output (by *Method 2* and *Method 3*) and the system output for two input profiles. One can see that the obtained models have relatively good prediction capability.

The results support the conclusion that if the model structure is not accurate, the OED with evolutionary strategy can result in more satisfactory parameter values than the classical OED with sequential quadratic programming and nonlinear least squares algorithms.

Table 1 Estimated parameters and the cost values of the obtained kinetic functions

	$(\mu_{max}, K_S)$	$E_{mse} \cdot 10^6$
Method 1*	(0.0893, 0)	98.0
	(0.0893, 0)	98.0
Method 2*	(0.0881, 0)	96.0
	(0.0894, 1.02)	35.0
Method 3**	(0.0895, 0.416)	15.6

\* These methods were initialized with two parameter vectors, see above.

\*\* The means of the  $E_{MSE}$  values of six independent runs. Two initial parameter vectors were used for 3-3 runs.

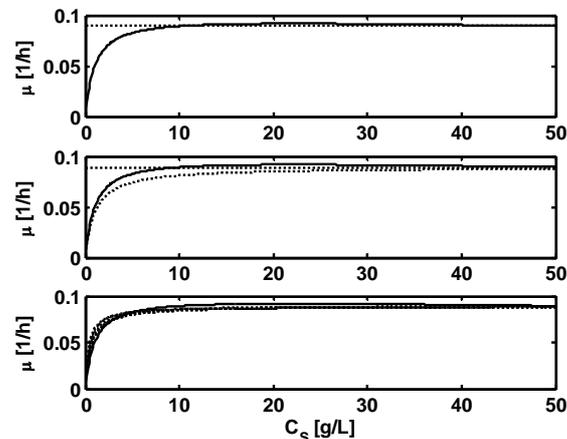


Fig. 4 The obtained kinetic functions vs. the ‘true’ kinetic top: Method 1, middle: Method 2, bottom: Method 3 solid line: ‘true’ kinetic, dashed line: obtained kinetic

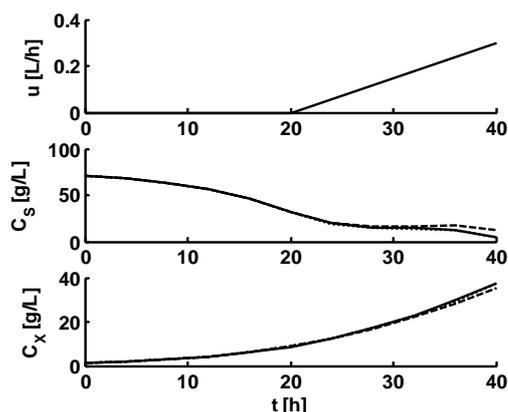


Fig. 5 The obtained model output vs. the 'true' solid: system, dashed: Method 2 dotted: Method 3

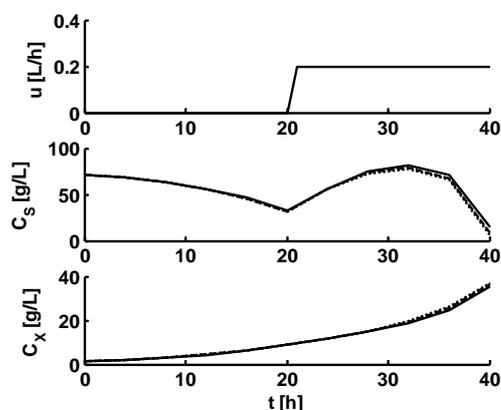


Fig. 6 The obtained model output vs. the 'true' solid: system, dashed: Method 2 dotted: Method 3

### Conclusions

The estimation of model parameters largely depends on the information content of the experimental data presented to the parameter identification algorithm. Establishing *optimal experiment design* (OED) can maximise the confidence on the model parameter, hereby increasing the confidence on the model prediction.

Both the parameter estimation and experiment design tasks represent a complex nonlinear optimization problem, hence the effectiveness of the applied optimization algorithms has great influence on the performance of the whole procedure. This paper proposes the application of *evolutionary strategy* (ES) for this purpose. ES is a stochastic optimization algorithm that uses the model of natural selection.

In this paper, this ES based OED technique has been applied for fed-batch biochemical reactor. One of the

factors affecting the modelling of biochemical systems is that accurate description of biochemical reaction is generally not available *a priori*.

Hence, this paper addressed the application of OED for models that have inaccurate structure. We illustrated that although the model structure used for the design of the experiments is not perfectly known, OED can result in satisfactory parameter values. Our results support the conclusion that when the model structure is not accurate, the OED with evolutionary strategy can result in more satisfactory parameter values than the classical OED with sequential quadratic programming and nonlinear least squares algorithms.

### Acknowledgements

The authors would like to acknowledge the support of the Cooperative Research Centre (VIKKK) (project 2003-I), and founding of the Hungarian Research Found (OTKA T037600).

### REFERENCES

1. BERNAERTS, K., SERVAES, R.D., KOOYMAN, S. and VERSYCK, K.J. and VAN IMPE, J.F.: Optimal temperature design for estimation of the Square Root model parameters: parameter accuracy and model validity restrictions, *Int. Jour. of Food Microbiology*, 2002, 73, 145-157
2. ASPREY, S.P. and MACCHIETTO, S.: Designing robust optimal dynamic experiments, *Journal of Process Control*, 2002, 12, 545-556
3. ESPIE, D.M. and MACCHIETTO, S.: The optimal designs of dynamic experiments, *AIChE J.*, 1989, 35, 223-229
4. VERSYCK, K.J., BERNAERTS, K., GEEARERD, A.H. and VAN IMPE J.F.: Introducing optimal experiment design in predictive microbiology: a motivating example. *Int. Journ. of Food Microbiology*, 1999, 51(1), 39-51
5. ASPREY, S.P. and MACCHIETTO, S.: Statistical tools for optimal dynamic model building, *Comp. Chem. Eng.*, 2000, 24, 1261-1267
6. SMETS, I.Y.M, VERSYCK, K.J.E and VAN IMPE, J.F.: Optimal control theory: A generic tool for identification and control of (bio-)chemical reactors, *Annual Reviews in Control*, 2002, 26, 57-73
7. MADÁR, J. and ABONYI, J.: *Evolutionary Algorithms*, Chapter 2.10. in *Instrument Engineers' Handbook*, 4th Edition, Volume 2 - Process Control, editor: B. Liptak, CRC Press, 2005
8. SCHWEFEL, H.: *Numerical optimization of computer models*. Wiley, Chichester, 1995