Summary

Succesful calibration of bioprocess models can only be achieved when information rich data is available. Therefore, it is desired to design experiments in such a way that the data which will be collected meet this objective. In order to design an optimal experiment, different choices have to be made. It has to be decided whether, where and how the system under study will be manipulated and where, how and when measurements will be performed on this system.

The non-linear and dynamic nature of bioprocess models makes the application of optimal experimental design techniques far from straightforward. During this PhD thesis, several issues related to non-linear dynamic optimal experimental design were identified and solutions proposed.

Optimal experimental design involves a series of complicated steps including parameter estimation, sensitivity analysis, non-linear optimization. However, no simulation package exists which combines all these methods. Hence, extensions were programmed to an existing modelling and simulation package, WEST. This modified simulation software, called EAST, is able to solve optimal experimental design problems in a general way and is applicable to every model available within WEST. In order to maintain generality, much emphasis was put on numerical rather than on analytical techniques.

Local sensitivity functions are an important component of the Fisher Information Matrix (FIM) which is used as the basis for optimal experimental design for model calibration. A much used method to calculate these local sensitivity functions is the finite difference technique. However, the practical application of this technique to a non-linear dynamic model poses a problem: a correct perturbation factor needs to be chosen in order to prevent numerical errors or errors related to the non-linearity of the model to influence the sensitivity analysis results. Therefore, a semi-automatic method to detect wrongly calculated sensitivity functions was developed which is based on the quantification of the difference between two sensitivity functions calculated with opposite perturbation factors. In order

to eliminate the error prone and laborious choice of the perturbation factors, a technique based on complex-number calculations (the complex-step derivative approximation method) was also investigated. Using this technique, the user is no longer required to specify perturbation factors and completely reliable results are obtained. However, a significant simulation speed decrease could be observed when using this technique.

In order to quantify the quality of a parameter estimation exercise, the parameter estimation error covariance matrix should be available. This can be calculated based on the Hessian matrix or the FIM. Calculating the Hessian matrix requires the numerical and computationally intensive evaluation of the second derivatives of the objective function with respect to the model parameters. The Richardson's extrapolation technique proved very useful in calculating these derivatives in a correct and automatic way.

Performing optimal experimental design for complex models containing many parameters requires an a priori choice to be made about the number of parameters that will be taken into account for the design. Therefore, techniques are needed to select identifiable parameter sets based on existing data. Making use of the relationship between the FIM and the Hessian, a new method, based on FIM related properties, was proposed to select identifiable parameter subsets which requires only limited user interaction.

The optimization problem related to optimal experimental design for model calibration can be very complex, especially when many experimental degrees of freedom and constraints are considered. The objective surfaces of the FIM criteria typically show a large number of local minima. In order to increase the probability of finding the global optimum, real-coded genetic algorithms (GAs) were successfully applied to the design of measurement campaigns for sequencing batch reactors.

It was also shown that GAs can be used to solve optimization problems which involve combinations of continuous and discrete optimization variables (experimental degrees of freedom).

In order to decrease the optimization computational demand, the experimental design problem was split up into parts: an inner and an outer loop for the optimization of measurement experimental degrees of freedom and manipulation experimental degrees of freedom respectively.

Another well-known problem of experimental design based on FIM properties is that the FIM optimal design criteria are often conflicting and that an experiment which is optimal for a certain design criterion is far from optimal for another design criterion. Finding experimental designs which are optimal with respect to several criteria can be accomplished by performing a multi-objective optimization of the experimental degrees of freedom. Multi-objective GAs proved ideal candidates for solving optimal experimental design problems dealing with several FIM criteria and also considering experimental costs.

The classical iterative optimal experimental design procedure (iterative experimentation, calibration and experimental design) involves different, rather complicated mathematical and practical steps in order to obtain the cali-

brated model. These steps require much user interaction and expert knowledge. As a solution, an automatic optimal experimental design procedure was developed in which user interaction and expert knowledge is only required at the beginning. All subsequent steps of the iterative search for the best parameter estimates can be performed automatically. These steps include (1) finding the optimal experiment, (2) performing the experiments in practice and (3) recalibrating the model. The proposed procedure was successfully applied to calibrate a one-step nitrification model using a respirometric experiment.

An important aspect of optimal experimental design for parameter estimation for non-linear models is the dependency of the design on the model parameters. This is caused by the fact that the FIM is calculated from sensitivity functions (partial derivatives of the model variables to the parameters) which, for non-linear models, are still function of the model parameters. This causes designs based on the FIM to be only locally optimal (for the particular values of the parameters for which the experimental design is performed) which may lead to unsatisfactory results when the used model parameters are too different from the "real" parameter values. Several known techniques for robust experimental design were studied and implemented in EAST together with a newly proposed technique based on the sensitivity of the FIM criteria to parameter changes. It was shown that the designs based on known robust experimental design criteria produced the best results. It was also found that robust experimental design is very computationally demanding. At this moment, the computational requirements limit the application of these techniques to small, simple models with a limited set of experimental degrees of freedom.

As a concluding remark, it can be said that optimal experimental design is an essential tool if high quality knowledge and models for complex dynamic non-linear systems like bioprocesses are to be obtained.