UNIVERSITEIT GENT

Ghent University

Faculty of Bioscience Engineering

Faculty of Bioscience Engineering

OPTIMAL EXPERIMENTAL DESIGN TO DISCRIMINATE AMONG RIVAL DYNAMIC MATHEMATICAL MODELS

ir. Brecht Donckels

Thesis submitted in fulfillment of the requirements for the degree of

Doctor (Ph.D) in Applied Biological Sciences

Academic year 2008-2009

Summary

A mathematical model can be defined as a mathematical representation of the mechanisms that govern the behavior of a process being studied, and the aim of a modeling exercise is to obtain a mathematical model that adequately describes and even predicts the process behavior. Once a proper model has been identified, it becomes a powerful tool for both scientists and engineers. However, it is important to realize that the lack of insight in the modeled process may result in the proposal of several so-called rival models, each of which represents a certain hypothesis of how the process works. Obviously, one is especially interested in the model that describes the process behavior in the best way. The problem of identifying the best model from a set of rival models, often referred to as the *problem of model discrimination*, was dealt with in this dissertation.

A general procedure to deal with the problem of model discrimination consists of four steps that are performed in an iterative manner until a stopping criterion is met. In the first step, the parameters of the rival models are estimated from all the data that is available. A second step involves an evaluation of the model's adequacy to describe the available data and is thus performed in order to find out which models are able to describe the data in a reasonable manner and which ones do not. Models that pass this test are used in a third step, where an optimal discriminatory experiment is designed. This experiment is then performed in a fourth and last step, after which the loop is closed by re-estimating the parameters of all rival models using all data available at that time.

To identify the best model from the set of rival models, it is thus necessary to collect new information about the process, and thus additional experiments have to be performed. Since performing experiments can be time and money consuming, carefully designing them can significantly reduce the required experimental effort. To achieve model discrimination in a minimal number of experiments, a number of design criteria have been developed. A detailed study of these design criteria revealed that the experimental design basically

comes down to finding that experiment that maximizes the difference between the model predictions, preferably taking into account the uncertainty associated with this difference. Indeed, when designing the experiment, one assumes that one the rival models is the *true* model and that the outcome of the designed experiment can be predicted by this model. Under this assumption, one expects that it will be possible to identify the most appropriate model when the other models predict this experiment in a totally different way.

The uncertainty associated with the difference in the model predictions originates from two sources: the uncertainty on the measurements and the uncertainty on the model predictions. Indeed, the uncertainty on the measurements, which can be seen as a measure of the reproducibility of the experiment(al) data, has to be taken into account when designing a discriminatory experiment, as well as the uncertainty on the model predictions since the evaluation of the discriminatory potential of an experiment is based on how it is predicted by the rival models.

These aspects are incorporated in the design criterion proposed by Buzzi-Ferraris et al. (1984), which was taken as a starting point because it is appealing from a conceptual point of view. However, this design criterion only uses the information content of the already performed experiments to evaluate the discriminatory potential of the designed experiment and, therefore, a modified design criterion was proposed in this dissertation, where the expected information content of the newly designed experiment is considered, even before the experiment is performed. Hence, this approach was called the anticipatory approach to optimal experimental design for model discrimination.

After applying the approaches to design optimal discriminatory experiments to a case study, one could conclude that the anticipatory approach performed better than the original approach of Buzzi-Ferraris et al. (1984). The reason why the anticipatory approach performs better than the original approach of Buzzi-Ferraris et al. (1984) is related to the uncertainty on the parameter estimates, and can be explained by the similarity between the anticipatory approach and optimal experimental design for parameter estimation (both design criteria benefit from a larger information content of the designed experiment with regard to the model parameters).

It is clear that high model prediction uncertainties hamper the efficacy and efficiency of the model discrimination procedure. These model prediction uncertainties are determined by the quality of the available data, since low quality data will result in poorly estimated parameters, which in turn result in uncertain model predictions. The discrimination among several rival models may thus become more efficient and effective if this uncertainty could be reduced prior to the start of the model discrimination procedure. Reducing the uncertainty on the model predictions can be achieved by designing and performing experiments dedicated to reducing the uncertainty on the parameter estimates. However, performing an additional experiment for each rival model may undermine the overall goal of optimal experimental design, since this would require at least as many experiments as the number of rival models. Therefore, the possibility to design a so-called compromise experiment was investigated. Such a compromise experiment is defined as an experiment that is not optimal for one or more of the individual rival models, but is sufficiently informative to improve the overall precision of the parameters of all rival models.

To design compromise experiments, two methods were developed: the kernel-based method and the ideal point method. The kernel-based method was developed to design a compromise experiment for experimental design problems where only the sampling times are to be optimized. Because the use of this method is limited to such experimental design problems, the idea of designing a compromise experiment was further explored by treating it as a multi-objective problem. As a result, the so-called ideal point method was proposed, which can be used for experimental design problems with experimental degrees of freedom of any type (manipulations, initial conditions and sampling times).

Finally, three approaches to integrate optimal experimental design for parameter estimation and model discrimination were investigated. In a first procedure, both aspects are dealt with sequentially, that is, the model discrimination procedure is performed first, and then the parameters of the selected model are further refined through the design of optimally informative experiments. The second procedure, is similar, except that a compromise experiment is designed and performed prior to the start of the model discrimination procedure to improve the quality of the parameter estimates. In the third procedure, both issues are dealt with simultaneously. For this purpose, the joint design criterion proposed by Hill et al. (1968) is modified such that the anticipatory approach can be used to quantify the discriminatory potential of the proposed experiments. The results obtained after applying the three procedures to a case study showed that, although model discrimination was not achieved in less experiments compared to the sequential procedure without a compromise experiment, the quality of the parameter estimates improved faster when a compromise experiment was performed first. The performance of the simultaneous procedure appeared to be worse than the performance of the sequential procedures, because the wrong model was identified as the most appropriate more frequently than with the sequential procedures.

In addition, the results confirmed the similarity between the anticipatory approach to design optimal discriminatory experiments and optimal experimental design for parameter estimation.