Chemometrics and Intelligent Laboratory Systems 110 (2012) 20–31

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Performance assessment of the anticipatory approach to optimal experimental design for model discrimination

Brecht M.R. Donckels [a,b,*], Dirk J.W. De Pauw [b], Peter A. Vanrolleghem [c], Bernard De Baets [b]

[a] BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure links 653, B-9000 Ghent, Belgium
[b] KERMIT, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure links 653, B-9000 Ghent, Belgium
[c] modelEAU, Département de génie civil et génie des eaux, Université Laval, 1065, Avenue de la Médicine, Québec, QC, G1V 0A6, Canada

ARTICLE INFO

Article history:
Received 29 July 2010
Received in revised form 3 June 2011
Accepted 26 June 2011
Available online 12 July 2011

Keywords:
Dynamic modeling
Mathematical models
Model discrimination
Optimal experimental design
Rival models
Anticipatory design
Systems biology
Synthetic biology

ABSTRACT

The problem of model discrimination arises when several models are proposed to describe one and the same process, a situation encountered in many research fields. To identify the best model from the set of rival models, it may be necessary to collect new information about the process, and thus additional experiments have to be performed. Several approaches have been described in literature to design optimal discriminatory experiments. The anticipatory approach is one of them and is very appealing from a conceptual point of view because the expected information content of the newly designed experiment is considered, even before the experiment is performed (anticipatory design). In this paper, the performance of this approach is evaluated by comparing it with the performance of other, established approaches to optimal experimental design for model discrimination. To conduct this comparison four performance measures were defined: (1) whether the most appropriate model could be identified, (2) the number of additional experiments that have to be designed and performed to achieve model discrimination, (3) the quality of the parameter estimates of the model that is eventually identified as the most appropriate one, and (4) the rate at which the inadequate models are identified. The results clearly indicate that the anticipatory approach has its benefits and may be the preferred approach in many applications in (bio)chemical engineering and *in-silico* biology.

1. Introduction

Mathematical models are frequently used for the design, optimization and control of sometimes complex (bio)chemical processes. They also have the potential to be(come) very valuable tools to organize data and to consider interactions in complex systems in a rational way. In fact, they are increasingly used for this purpose in many research areas, especially in the emerging fields of systems biology [3,13,19] and synthetic biology [1,21]. In his influential overview paper [23], Kitano stresses that although the advances made in molecular biology to accurately gather quantitative experimental data have been enormous and will certainly continue, insights into the functioning of biological systems will not result from educated guesses alone, because of their intrinsic complexity. Instead, a combination of experimental and computational approaches is expected to resolve this challenging problem and, consequently, experimental design techniques will become increasingly important, as recognized by many researchers in the field [5,18,26,33,36,40].

The methods to design experiments that allow discriminating among rival models in an effective and efficient way, often referred to as optimal experimental design for model discrimination (OED/MD) or optimal experimental design for (model) structure characterization [41], will be the main focus of this paper. Indeed, when insight in a process is insufficient, several hypotheses can be postulated on how the process actually works. Each of these hypotheses can subsequently be translated into a unique model structure, and a set of rival models for the process arises. Obviously, one is especially interested in the model that describes the process under study in the most appropriate way. To identify this model from the set of rival models, it may be necessary to collect new information about the process, and thus additional experiments have to be performed.

The problem of model discrimination has been addressed in a number of ways (see [14] for a review), but common to all design criteria is the fact that the problem is tackled as and translated into an optimization problem. In the pioneering work of Hunter and Reiner (1965) [22], the difference between the model predictions is simply maximized (as explained in more detail further on). Although this approach does not take into account the uncertainties inherently involved in both the modelling phase and the experimentation phase, the basic rationale is still present in (as good as) all design criteria for OED/MD developed so far.

Buzzi-Ferraris and co-workers presented a design criterion that takes into account both the uncertainty on the measurements and the (resulting) uncertainty on the parameter estimates and model

* Corresponding author.
E-mail address: brecht.donckels@biomath.ugent.be (B.M.R. Donckels).

0169-7439/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2011.06.008

predictions [8]. The latter was further refined in [15] and [38], where the so-called anticipatory approach to OED/MD was introduced. In this approach, the expected information content of the newly designed experiment is considered, even before the experiment is performed (whence the term anticipatory design). In this way, a better estimate of the uncertainties can be achieved and an experiment with an increased discriminatory potential can be obtained. Because of the similarity of this approach with the conventional, state-of-the-art design criteria for optimal experimental design for parameter estimation (OED/PE), improved parameter estimates can be obtained in addition to model discrimination [15,38].

The objective of this paper is to determine whether different approaches to OED/MD differ in their ability to bring forth a series of (informative) discriminatory experiments. In the evaluation of their performance, four aspects are considered after applying them to a case study with nine rival models that was used in previous work on the subject [15–17]. The first aspect has to do with the outcome of the model discrimination procedure, that is, how the procedure ends (most appropriate model identified or all rival models rejected). The second aspect is related to the number of additional experiments that have to be (designed and) performed before the most appropriate model can be identified. It is clear that this is very important, as one obviously wants to minimize the number of additional experiments. The third aspect is related to the quality of the parameter estimates of the model that is eventually identified as the best one (if any). The fourth and last aspect of the performance evaluation is related to the rate at which the inadequate models are identified.

This paper is organised as follows. In Section 2, the basic rationale of optimal experimental design for model discrimination is explained and formalized in a mathematical manner. The section also presents the approaches to design optimal discriminatory experiments that were considered in this paper, as well as the four performance measures that were used in their evaluation. To conclude, this section explains how a case study was designed to investigate the performance of the OED/MD methods. The results obtained after applying these methods to the case study are presented and discussed in Section 3 and the conclusions are presented in Section 4.

2. Methods

2.1. Mathematical model representation

In what follows, general deterministic models in the form of a set of (possibly mixed) differential and algebraic equations are considered, using the following notations:

$$\dot{x}(t) = f(x(t), u(t), \theta, t); \quad x(t_0) = x_0 \tag{1}$$

$$\hat{y}(t) = g(x(t)) \tag{2}$$

where $x(t)$ is an n_s-dimensional vector of time-dependent state variables, $u(t)$ is an n_u-dimensional vector of time-varying inputs to the process, θ is an n_p-dimensional vector of model parameters taken from a continuous, realizable set Θ, and $\hat{y}(t)$ is an n_m-dimensional vector of measured response variables that are function of the state variables, $x(t)$. An experiment will be denoted as ξ, and is determined by the experimental degrees of freedom such as measurement times, initial conditions and time-varying or constant process inputs.

2.2. Parameter estimation

The values of the model parameters, which by definition do not change during the course of the simulation, have to be determined from experimental data. This process is called parameter estimation, and typically consists of minimizing the weighted sum of squared errors (WSSE) functional through an optimal choice of the parameters θ. The WSSE is calculated as follows

$$\mathrm{WSSE}\left(\hat{\theta}\right) = \sum_{k=1}^{n_e} \sum_{l=1}^{n_{sp_k}} \Delta\hat{y}\left(\xi_k, \hat{\theta}, t_l\right)' \cdot Q \cdot \Delta\hat{y}\left(\xi_k, \hat{\theta}, t_l\right), \tag{3}$$

where

$$\Delta\hat{y}\left(\xi_k, \hat{\theta}, t_l\right) = y(\xi_k, t_l) - \hat{y}\left(\xi_k, \hat{\theta}, t_l\right) \tag{4}$$

represents the difference between the vector of the n_m measured response variables and the model predictions at time t_l ($l = 1, \ldots, n_{sp_k}$) of experiment ξ_k ($k = 1, \ldots, n_e$). Further, n_e represents the number of experiments from which data is used for estimating the parameters, n_{sp_k} represents the number of sampling points in experiment ξ_k, and Q is an n_m-dimensional matrix of user-supplied weighing coefficients. Typically, Q is chosen as the inverse of the measurement error covariance matrix $\sum$ [28,35,42]. In this way, the measurement uncertainty is incorporated in the WSSE.

2.3. Model adequacy testing

To test a model's adequacy, a lack-of-fit test, as outlined for instance in [9–11], can be used. This test is based on the property of the WSSE-functional being a sample from a χ^2 distribution with $n - n_p$ degrees of freedom. However, this property only holds under two assumptions [11]: (i) the measurements are disturbed with random zero mean normally distributed noise with known (or a priori estimated) variance, and are not subject to systematic errors; and (ii) no model errors are present.

In this work, the data to which the models are fitted (see below) is generated by adding noise to the simulation results of which the characteristics are known, so the first assumption is always valid. Consequently, when the WSSE is significantly larger than the expected value of the appropriate $\chi^2_{n-n_p}$ distribution, one can conclude that the model is not able to describe the experimental data in a reasonable manner and the model can thus be rejected.

2.4. Optimal experimental design for model discrimination

In general, optimal experimental design is an optimization problem, where the optimum of a well-defined objective function is sought by varying the experimental degrees of freedom. This can be formalized as follows

$$\xi^{\star} = \arg\max_{\xi \in \Xi} T(\xi). \tag{5}$$

The experimental degrees of freedom, ξ, are restricted by a number of constraints that define a set of possible experiments, denoted as Ξ. These constraints are determined by the experimental setup and are specified before the start of the experimental design exercise. Note that in this context, the objective functions are also called design criteria, and these terms will be used as synonyms in the following.

2.4.1. Design criterion of Hunter and Reiner (1965)

Suppose, for simplicity, that one has to design an experiment to discriminate between two rival models ($m = 2$). It is clear that the data expected from the designed experiment should be predicted differently by the two models to allow for model discrimination. Hunter and Reiner translated this heuristic into an objective function [22] given by

$$T_{ij}(\xi) = \sum_{l=1}^{n_{sp}} \Delta\hat{y}_{ij}\left(\xi, \hat{\theta}_i, \hat{\theta}_j, t_l\right)' \cdot \Delta\hat{y}_{ij}\left(\xi, \hat{\theta}_i, \hat{\theta}_j, t_l\right), \tag{6}$$

where

$$\Delta\hat{y}_{ij}\left(\xi,\hat{\theta}_i,\hat{\theta}_j,t_l\right)=\hat{y}_i\left(\xi,\hat{\theta}_i,t_l\right)-\hat{y}_j\left(\xi,\hat{\theta}_j,t_l\right) \tag{7}$$

represents the difference between the n_m-dimensional vectors of the predicted outcomes of experiment ξ by model i and model j at time t_l, and n_{sp} represents the number of sampling points. Note that this notation will be simplified to $\Delta\hat{y}_{ij}(\xi,t_l)$ in the following.

It is important to point out that this objective function does not take into account the uncertainty on the measurements, nor on the model predictions. However, it is important to do so. Indeed, the difference in the model predictions may be high for a particular subset of the experimental degrees of freedom, but when those experimental conditions result in a situation that is characterized by a high measurement error, discrimination may not be possible after all. Furthermore, when the uncertainty on the model predictions is high, the predicted difference in the model predictions may be less pronounced than expected and the value of the designed experiment with regard to model discrimination may ultimately be limited.

2.4.2. Modified design criterion of Hunter and Reiner (1965)

Incorporating the uncertainty on the measurements can easily be done using the measurement error covariance matrix in a similar way as in the objective function used for parameter estimation (Eq. (3)). This results in the following objective function:

$$T_{ij}(\xi)=\sum_{l=1}^{n_{sp}}\Delta\hat{y}_{ij}(\xi,t_l)'\cdot\sum(\xi,t_l)^{-1}\cdot\Delta\hat{y}_{ij}(\xi,t_l). \tag{8}$$

Here, $\sum(\xi,t_l)$ represents the measurement error covariance matrix at time t_l of experiment ξ.

2.4.3. Design criterion of Buzzi-Ferraris (1984)

The design criterion proposed by Buzzi-Ferraris and co-workers [8] builds further on the modified version of Hunter and Reiner's design criterion (Eq. (8)) and also incorporates the uncertainty on the model predictions. This is done by weighing the difference in the predicted outcomes of an experiment, denoted as $\Delta\hat{y}_{ij}(\xi,t_l)$, with the uncertainty associated with them. This uncertainty originates from two sources: the uncertainty on the measurements on the one hand, and the uncertainty on the predictions of both models on the other hand.

Also for the model predictions, a covariance matrix is used to quantify the associated uncertainty. The model prediction error covariance matrix associated with time t_l of experiment ξ_k, denoted as $\Omega(\xi_k,t_l)$, is calculated by propagating the uncertainty on the parameter estimates according to [35]. This is formalized as

$$\Omega(\xi_k,t_l)=\left(\left.\frac{\partial\hat{y}(\xi_k,\theta,t_l)}{\partial\theta}\right|_{\hat{\theta}}\right)\cdot\Phi\cdot\left(\left.\frac{\partial\hat{y}(\xi_k,\theta,t_l)}{\partial\theta}\right|_{\hat{\theta}}\right)', \tag{9}$$

where $\frac{\partial\hat{y}}{\partial\theta}$ represents the sensitivities of the response variables ($\hat{y}$) to changes in the parameters (θ) and Φ represents the parameter estimation error covariance matrix, that can be approximated by the inverse of the so-called Fisher information matrix [2,27,29,43]:

$$\Phi^{-1}=\sum_{k=1}^{n_e}FIM(\xi_k). \tag{10}$$

This parameter estimation error covariance matrix is then used in the calculation of the uncertainty on the model prediction, as shown in Eq. (9). For more detailed information on the calculation of the Fisher information matrix and the uncertainties, the reader is referred to [15] where the anticipatory approach is described in more detail and where the same notations were used.

Now, when $\sum(\xi,t_l)$ and $\Omega(\xi,t_l)$ are assumed to be independent, the uncertainty on the predicted outcome of an experiment can be estimated as $\sum(\xi,t_l)+\Omega(\xi,t_l)$. However, this assumption is not entirely valid. Indeed, the measurement uncertainties are used when estimating the model parameters (Eq. (3)), which are on their turn used in the calculation of the uncertainties on the model predictions (Eq. (9)). Still, it is a reasonable one in this context because it is a theoretical dependence and the objective function is practically useful and helps to identify the most appropriate model. Under a similar assumption of independence, the uncertainty on the difference between the predicted outcomes of an experiment ξ by model i and j, denoted as $\Psi_{ij}(\xi,t_l)$, is given by

$$\begin{aligned}\Psi_{ij}(\xi,t_l)&=\sum(\xi,t_l)+\Omega_i(\xi,t_l)+\sum(\xi,t_l)+\Omega_j(\xi,t_l)\\&=2\cdot\sum(\xi,t_l)+\Omega_i(\xi,t_l)+\Omega_j(\xi,t_l).\end{aligned} \tag{11}$$

The objective function thus becomes

$$T_{ij}(\xi)=\sum_{l=1}^{n_{sp}}\Delta\hat{y}_{ij}(\xi,t_l)'\cdot\Psi_{ij}(\xi,t_l)^{-1}\cdot\Delta\hat{y}_{ij}(\xi,t_l), \tag{12}$$

where $\Psi_{ij}(\xi,t_l)$ represents the uncertainty on the difference between the predicted outcomes of an experiment by models i and j at time t_l.

2.4.4. Anticipatory approach to OED/MD

From a conceptual point of view, the design criterion proposed by Buzzi-Ferraris and co-workers (Eq. (12)) is superior to the other ones because of the importance it gives to the uncertainty when performing model discrimination. Still, this design criterion can be further improved by recalculating the parameter estimation error covariance matrix for each proposed experiment by also including the expected information content of this new experiment. In this way, the expected information content of the newly designed experiment is accounted for, even before the experiment is performed (hence the term anticipatory design). This so-called anticipatory approach to OED/MD was developed simultaneously and independently in [38] and [15].

The objective function or design criterion used in this approach is basically the same as Eq. (12), but the difference lies in the calculation of the parameter estimation error covariance matrix (denoted as Φ), which is recalculated for each proposed experiment by also including the expected *FIM* associated with this new experiment. This can be formalized as follows

$$\Phi^{-1}=\sum_{k=1}^{n_e}FIM(\xi_k)+FIM\left(\xi_{n_e+1}\right), \tag{13}$$

where the expected information content of the newly designed experiment is represented by $FIM(\xi_{n_e+1})$.

2.4.5. OED/MD for more than two rival models

The design criteria described above were developed for model discrimination problems with two rival models. However, when the number of rival mathematical models is larger than two (as in the case study presented below), several strategies are thinkable to steer the model discrimination procedure, irrespective of which design criterion is chosen. Here, the so-called pairwise strategy is chosen, where an optimal discriminatory experiment is designed for each model pair, and the experiment with the largest T_{ij} value (as defined in Eq. (5)) is eventually performed. Other design strategies are possible, but for this the interested reader is referred to [9,14,37].

2.5. Selection of the most promising approaches to OED/MD

In this paper, the performance of the four objective functions described above and summarized in Table 1 will be compared. The approach proposed by [22] is selected for its simplicity and because it represents the basic idea behind OED/MD. In a way, this approach can be seen as a reference approach with which the effects of the conceptual improvements incorporated in the other approaches can be compared. Also, the modified version of this approach is selected because this approach is the only one that takes into account the measurement error without considering the uncertainty on the model predictions as well. The notations T_a and T_b will be used to indicate these approaches, which, for clarity, use the objective functions given by Eqs. (6) and (8).

The design criterion of Buzzi-Ferraris (1984), described by Eq. (12), is also selected because it incorporates both the uncertainty on the measurements and on the model predictions and because it has already been used successfully by other researchers (for instance in [6,7,24,37]). The anticipatory approach is the fourth selected approach, and these will respectively be referred to as T_c and T_d, in the discussion below.

2.6. Performance measures to evaluate OED/MD design criteria

As stated in the introduction, the objective of the case study described and discussed below is to determine whether the four selected approaches differ in their ability to bring forth a series of (informative) discriminatory experiments. In the evaluation of their performance, four aspects are considered:

1. The outcome of the model discrimination procedure, that is, how the procedure ends: either the most appropriate model is identified or all rival models are rejected (discussed in Section 3.1).
2. The number of additional experiments that has to be (designed and) performed before the most appropriate model can be identified. It is clear that this is very important, as one obviously wants to minimize the number of additional experiments (discussed in Section 2).
3. The quality of the parameter estimates of the model that is eventually identified as the best one (if any). Indeed, when the most appropriate model has been identified, its parameter estimates may have to be further improved in order to get reliable model predictions. Good quality parameter estimates after model discrimination is of course a beneficial and desired property of the OED/MD approach (discussed in Section 3).
4. The rate at which the inadequate models are identified (discussed in Section 4).

2.7. Description of the case study

To evaluate the performance of the different approaches, the same case study was used as in other papers on the same topic [15–17]. In this case study, nine models are defined to describe the kinetic behavior of the enzyme glucokinase (*glk*, EC: 2.7.1.2), which catalyzes the conversion of glucose (GLU) and ATP to glucose-6-phosphate (G6P) and ADP. This reaction is the first reaction of the glycolysis pathway and it was recently suggested that glucokinase may be inhibited by phosphoenolpyruvate (PEP) [34]. Each model structure represents a specific hypothesis on how the reaction process works (as shown in Table 2), but for brevity the details on the model structures are not given here and the reader is referred to the cited papers if interested.

Table 1
Overview of the selected approaches to design optimal discriminatory experiments (denoted as T_a, T_b, T_c and T_d).

	T_a	T_b	T_c	T_d
Experimental design driven by the difference in the model predictions	×	×	×	×
Uncertainty on the measurements taken into account	–	×	×	×
Uncertainty on the model predictions taken into account	–	–	×	×
Information content of the designed experiment taken into account	–	–	–	×

2.7.1. Design of the optimal discriminatory experiments

For the design of the optimal discriminatory experiments, the sampling times and the initial concentrations of glucose, ATP and PEP were chosen as experimental degrees of freedom, and were thus optimized. Boundary conditions were defined for the initial concentrations and ten optimal sampling times were determined with the constraint that the time between two consecutive samples was at least 15 s. As in the other papers where this case study was used [15–17], the χ^2 lack-of-fit test described above was used to evaluate the adequacy of the rival models.

2.7.2. Design of the case study

Because the performance of the different experimental design methods might be influenced by the (information content of the) preliminary experiment used to initiate the model discrimination procedure, five different scenarios were worked out (as shown in Fig. 1). Each scenario starts from a randomly generated preliminary experiment, which is denoted as ξ_1^i ($i=1, \ldots, 5$). The experimental data were generated by simulating the experiment with the (arbitrarily chosen) real model (m_5) and the measurement error was simulated by adding random noise as suggested by [39]:

$$\sigma_y = \hat{y}\cdot\varsigma_y\cdot\left(1+\frac{1}{\left(\frac{\hat{y}}{lb_y}\right)^2+\frac{\hat{y}}{lb_y}}\right). \tag{14}$$

Here, ς_y and lb_y respectively represent a constant minimal relative error and a lower accuracy bound on the measurement of y. In this way, the standard deviations of the measurements are proportional to the value of $\hat{y}$, but increase when the latter approaches the detection limit or the lower accuracy bound of the measured state variable.

Although the first discriminatory experiment that is designed in the model discrimination procedure is the same for each repetition of the procedure (within a given scenario i), the data obtained from this experiment are different due to these randomly generated measurement errors. Consequently, the discriminatory experiments designed in the following iterations of the procedure are different as well, because the differences in the experimental data sets lead to differences in the parameter estimates and their uncertainties. To account for this, each model discrimination exercise is repeated thirty times (as indicated in Fig. 1). In total, the model discrimination procedure is thus performed $5\times4\times30=600$ times. Indeed, there are 5 scenarios (each with a different preliminary experiment), 4 different approaches for OED/MD

Table 2
Overview of the nine different rival models used to investigate the performance of the selected approaches to design optimal discriminatory experiments.

Rival model	Random	Ordered		Inhibition by PEP	
		Glucose	ATP	Glucose	ATP
m_1	×	–	–	–	–
m_2	×	–	–	–	×
m_3	×	–	–	×	–
m_4	–	×	–	–	–
m_5	–	×	–	–	×
m_6	–	×	–	×	–
m_7	–	–	×	–	–
m_8	–	–	×	–	×
m_9	–	–	×	×	–

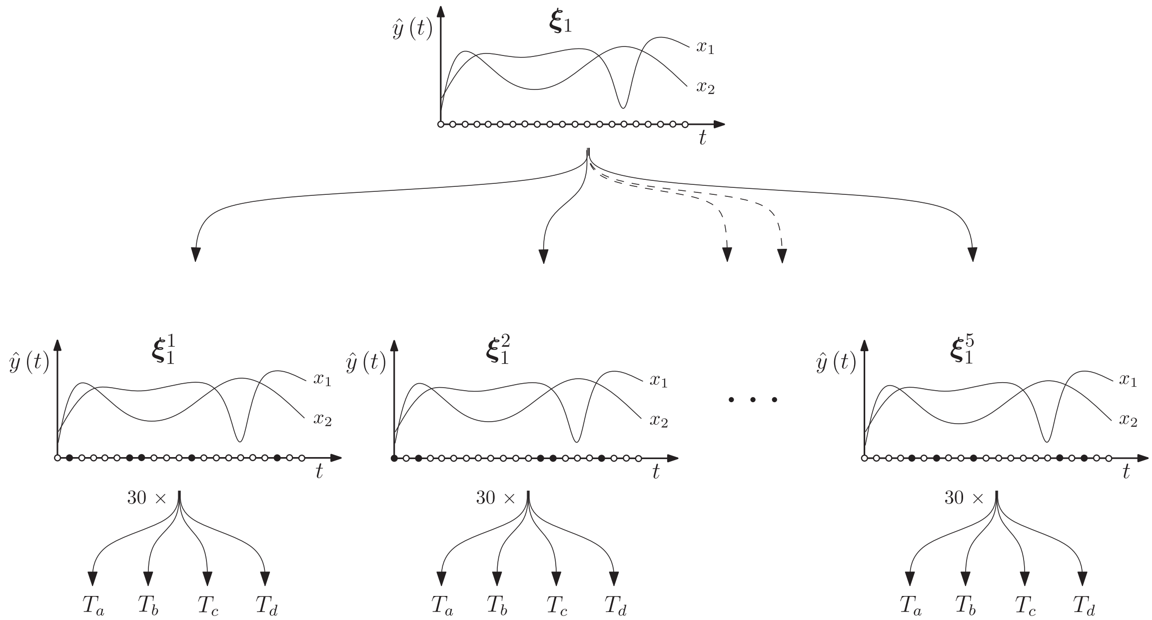

Fig. 1. Illustration of the design of the case study where the aim was to evaluate and compare the performance of four selected experimental design methods to discriminate among rival models (denoted as T_a, T_b, T_c and T_d).

(T_a, T_b, T_c and T_d) and the model discrimination procedure was repeated 30 times to account for the different instances of the measurement noise.

Note that each application of the model discrimination procedure consists of an iterative sequence of experiments, which differs from one repetition to another. In every iteration, an experiment is designed and performed, and the number of iterations is equal to the number of additional experiments that have to be performed before the most appropriate model can be identified. The data obtained from this simulation study will be presented and discussed in the following sections.

3. Results and discussion

In this section, the results obtained after applying the four approaches to the case study are discussed. To organize the discussion, each of the four performance measures described in Section 2.6 will be dealt with successively.

3.1. Outcome of the model discrimination procedure

Ideally, the model discrimination procedure ends when one of the rival models is identified as the most appropriate one. In this case study, the experimental data were generated using model m_5 and it can thus be expected that model m_5 is identified as the most appropriate model in the majority of the runs, regardless of the approach used to design the discriminatory experiments. However, the possibility that another model is identified as the most appropriate one cannot be excluded. A second possible outcome of the model discrimination procedure is that all models appear to be inadequate. Indeed, the adequacy of the rival models is evaluated based on the WSSE value, and even for the true model (m_5) this WSSE value can in some situations be larger than the reference value ($\chi^2_{n-n_p}$) because of the (simulated) error on the measurements.

Note that a third possibility, where the discriminatory potential of the designed experiment is conceived as too low to enable further discrimination among the remaining rival models, is not considered here as the correctness of the currently used criterion to evaluate the discriminatory potential of an experiment [10] is questionable (as discussed in [14]).

The results obtained in this case study are presented in Table 3. For the T_d approach, for instance, one can see that model m_5 was identified as the most appropriate one in 134 of the 150 applications (or runs) of the model discrimination procedure (5 scenarios with a different preliminary experiment and 30 repetitions of each scenario). The results indicate that also for the other approaches the *true* model (m_5) is found in most of the runs. From the results obtained for the T_d approach, one can also see that another model was identified as the best model in 7% of the runs of the model discrimination procedure. For brevity, it is not indicated which of the other rival models is eventually identified as the best model, but, in the majority of the runs where this occurred, model m_2 was selected. This is not surprising because models m_2 and m_5 only differ by the fact that the former

Table 3
Overview of the observed outcomes of the 150 runs of the model discrimination procedure for the different approaches (T_a, T_b, T_c and T_d) and the five scenarios (each with a different preliminary experiment, ξ_1^i with $i = 1, \ldots, 5$).

		T_a	T_b	T_c	T_d
Model m_5	ξ_1^1	30	30	30	27
	ξ_1^2	30	27	30	29
	ξ_1^3	28	27	30	25
	ξ_1^4	30	20	30	27
	ξ_1^5	28	24	30	26
2–6	$\xi_1^1-\xi_1^5$	146	128	150	134
		97%	85%	100%	89%
Other model	ξ_1^1	0	0	0	0
	ξ_1^2	0	3	0	1
	ξ_1^3	0	3	0	4
	σ_1^4	0	5	0	3
	ξ_1^5	1	2	0	2
2–6	$\xi_1^1-\xi_1^5$	1	13	0	10
		1%	9%	0%	7%
All models rejected	ξ_1^1	0	0	0	3
	ξ_1^2	0	0	0	0
	ξ_1^3	2	0	0	1
	ξ_1^4	0	5	0	0
	ξ_1^5	1	4	0	2
2–6	$\xi_1^1-\xi_1^5$	3	9	0	6
		2%	6%	0%	4%

assumes a random binding mechanism, whereas the latter assumes an ordered reaction mechanism (see Table 2). In other words, the models are very similar. In the other runs of the procedure, all rival models were rejected. It is, however, noteworthy that the true model was always identified as the most appropriate one when the T_c approach was used. Although a profound explanation for this observation cannot be given, it might be the result of the conservative character of the T_c approach, which will be discussed in the following.

3.2. Number of additional experiments to achieve model discrimination

The number of additional experiments that have to be performed before the most appropriate model can be identified is an important aspect of the performance of a given OED/MD approach. Before starting the discussion of the obtained results, it is interesting to note that even with a set of fifty randomly generated experiments it was not possible to identify the most appropriate model (results not shown), while model discrimination could be achieved in far less experiments with any of the selected OED/MD approaches. The fifty random experiments were generated by randomly choosing the sampling times (between 0 and 20 min) and the initial concentrations of glucose, ATP and PEP (between 0 and 2 mM). This result clearly illustrates the necessity or at least the importance of performing experiments that are designed in a rational way, that is, designed with the aim to achieve model discrimination.

As explained earlier (Section 2), the model discrimination procedure was initiated with one of the five preliminary experiments and each of these was repeated thirty times to account for the influence of the measurement error. The number of experiments that were required in the different model discrimination runs are presented as boxplots in Fig. 2. This figure contains five subfigures with a white background (entitled ξ_1^i, with $i = 1, \ldots, 5$) and one subfigure with a gray background (entitled $\xi_1^1 - \xi_1^5$). The former presents the results obtained for the simulations where the model discrimination procedure was initiated with the preliminary experiment indicated in the title of the corresponding subfigure, whereas the one with the gray background gives an overall picture of the number of required experiments and presents the values of all model discrimination runs ($4 \times 150 = 600$ in total). Note that in these figures, the preliminary experiment corresponds to experiment number one. In other words, Fig. 2 shows the number of required experiments, and not the number of required *additional* experiments. Also note that the median of the number of required experiments, which will be used frequently in the discussion below, cannot always be determined unambiguously from these boxplots (more precisely, when the horizontal line that indicates the median coincides with one of the edges of the box). Therefore, the median of the number of required experiments is also given in the upper right corners.

From the results shown in Fig. 2, one can see that the highest number of required experiments occurs when the T_a approach is used, regardless of the information content of the preliminary experiment. As discussed earlier, the T_a approach is the most naive one. Therefore, the discriminatory potential of the proposed experiments is often misjudged, and it is not surprising that the T_a approach is the worst performing one when it comes to the required number of experiments. This also clearly illustrates that it is important to take into account the uncertainties when designing optimal discriminatory experiments. This was advocated before from a theoretical or conceptual point of view in [8,10,15,38] (among others), but was never shown so unambiguously. In addition, one can see that the number of required experiments varies significantly among the repetitions. The latter can be observed for each of the starting situations, except for ξ_1^1. This large variability can be explained by the fact that the T_a approach allows taking samples even when the measurement error is high. Indeed, although the first discriminatory experiment is the same for each of the runs that were initiated with a certain preliminary experiment, the generated data sets of the other experiments can differ significantly because of this measurement error. Since the model parameters are estimated from these data sets, large differences in these parameter estimates can be expected. These obviously affect the model discrimination runs that follow and lead to the observed variability among the different repetitions. This variability also reflects the importance to include the information on the measurement errors when designing the discriminatory experiment.

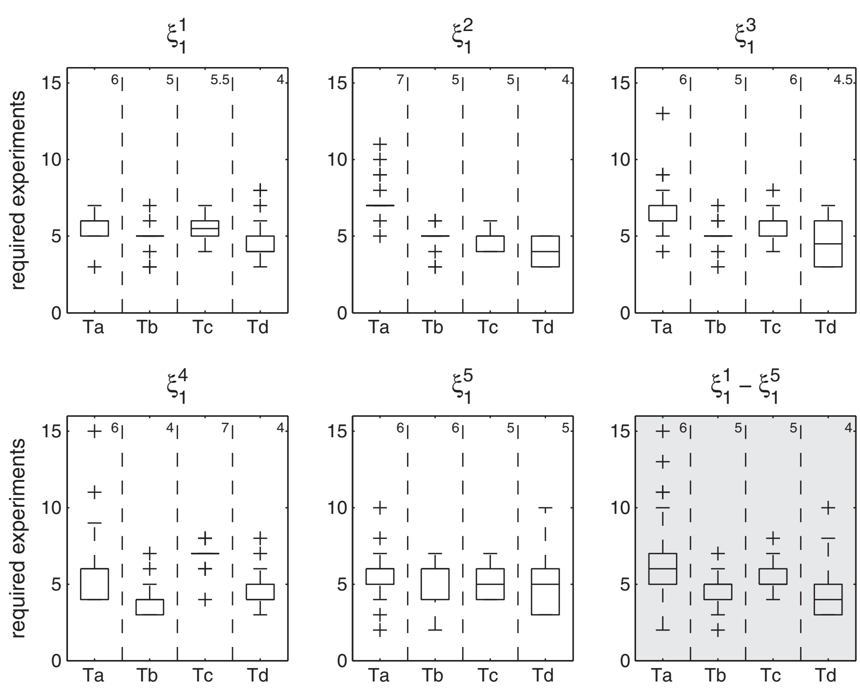

Fig. 2. Boxplots showing the number of experiments that are required to achieve model discrimination starting from each of the five preliminary experiments (ξ_1^1 till ξ_1^5) (white background). The subfigure with the gray background gives an overall picture of the number of required experiments and was made using the results obtained from all starting situations. The numbers in the upper right corners represent the median of the number of required experiments.

For both the T_b and the T_c approaches, the median of the number of required experiments equals five and one could conclude that both approaches perform equally well. That these approaches perform better than the T_a approach can be seen as an illustration of the importance of considering the measurement errors in the design of the experiments, and it confirms some of the conclusions drawn from the results for the T_a approach discussed above. However, a close(r) investigation of the boxplots in Fig. 2 and especially of the histograms in Fig. 3 indicates that the T_b approach is preferred over the T_c approach. Although the median of the number of required experiments is the same, the distributions are clearly different and in favor of the T_b approach. Knowing that the T_c approach has frequently been applied in literature (for instance in [6,7,24,37]) and that it was originally introduced as a conceptual improvement of the T_b approach, this result is somewhat surprising and may be related to the fact that the information content of the designed experiment is not fully considered during the design. This will be further discussed below, together with the results of the T_d approach.

The anticipatory approach (T_d) does take the information content of the to-be-performed experiment into account and performs better than the other approaches, regardless of the information content of the preliminary experiment. Indeed, for each of the starting situations, model discrimination was achieved with the least amount of experimental effort. However, one can see that the variability among the different repetitions is slightly larger than the variability observed for the T_c approach. The latter can be explained as follows. Both approaches use the currently available parameter estimates to predict the outcome of the proposed experiment and the uncertainty associated with it, but the T_c approach is more conservative than the T_d approach because it only uses the information of the already performed experiments to evaluate the proposed experiment for its discriminatory potential. In other words, the T_d approach is more sensitive to the accuracy of the available parameter estimates, but when the available parameter estimates are close to their actual values, the discriminatory potential of the designed experiment is assessed in a better way compared to the other approaches (as discussed in [15]). However, when the parameter estimates used in the experimental design differ significantly from the ones obtained after performing the designed experiment, the discriminatory potential of the experiment may not be as good as one expected. If this occurs, the designed experiment may not significantly contribute to the discriminatory potential of the experiments performed so far, and, in the end, an increased number of experiments will be required before the most appropriate model can be identified. This might explain why a large tail can be observed for the histogram of the T_d approach. Nevertheless, the results in Figs. 2 and 3 show that the T_d approach generally results in faster model discrimination compared to the T_c approach, which may indicate that, at least in this case study, the parameter estimates did not change considerably during the model discrimination procedure.

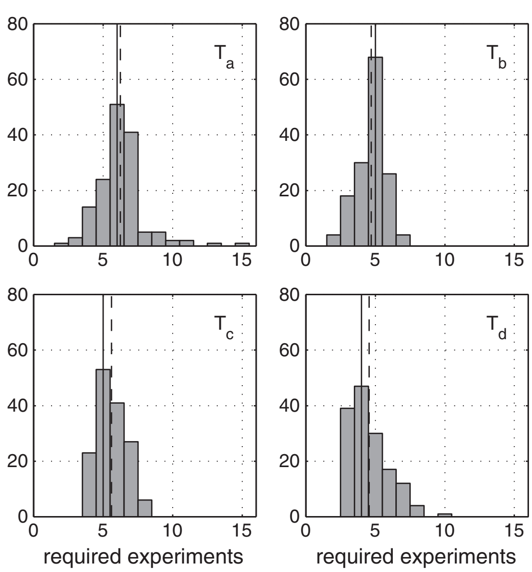

Fig. 3. Histograms showing the number of experiments required to achieve model discrimination for each of the selected approaches (T_a, T_b, T_c and T_d), thereby using the results obtained from all starting situations (gray figure from Fig. 2). The solid vertical lines indicate the median of the number of required experiments, while the dashed lines indicate the mean number of required experiments.

Note that Figs. 2 and 3 represent the number of experiments that are performed until the model discrimination procedure stops, as explained in Section 3.1, regardless of the outcome. In this respect, one could argue that only those runs should be considered in the evaluation where model m_5 was identified as the most appropriate model, but the results nor the discussion are significantly influenced when doing so (results shown in [14]). In other words, the runs in which, for instance, all rival models are rejected are not systematically those for which a high or low number of experiments are required.

To conclude this discussion, note that, recently, the T_a approach was reintroduced [30] to address model discrimination problems in the study of complex biological networks (systems biology). The authors' rationale is that the optimal discriminatory experiment is found after maximizing the differences between the outputs (or predictions) of the rival models (which is basically the same as in the design criterion of Hunter and Reiner (1965)) to ensure that even a noisy measurement has a good chance of discriminating between the models. This is somewhat conflicting with the rationale behind the T_b, T_c and T_d approaches and advocated in previous work on this subject (for instance in [8,10,15]), where the message was to incorporate these uncertainties in the design criterion. Because the authors of the referred paper [30] mention the high costs of performing experiments as one of the reasons why one should carefully design experiments, the results presented above may be relevant in future applications of OED/MD in the field of systems biology, because they clearly showed that model discrimination could be achieved in less experiments when the different sources of uncertainty are considered.

3.3. Evaluation of the quality of the parameter estimates

A third aspect that has to be considered in the evaluation of the approaches is the quality of the parameter estimates obtained at the end of the experiment sequence. The high dependency of the approaches on the quality of the parameter estimates is mentioned as one of the major drawbacks of OED/MD in [31]. Nevertheless, the T_c and T_d provide a way to deal with this by incorporating the uncertainty on the parameter estimates in the design criterion. Here, we are especially interested in the quality of the parameter estimates of the model that is eventually identified as the most appropriate one. Indeed, model discrimination is only one step of a more general model building procedure, and once an appropriate model is identified through model discrimination, the quality of its parameter estimates often has to be improved before the model can actually be applied for its intended use. This is important because inaccurate parameter estimates result in inaccurate (or uncertain) model predictions, which are obviously not desired. To increase the quality of the parameter estimates, dedicated experiments can be designed (using the experimental design techniques (OED/PE) explained for instance in [4,12,20,25,32,42]) and performed. In this respect, it would be interesting to see whether there is a difference among the approaches concerning the evolution of the quality of the parameter estimates throughout the model discrimination procedure, because this may have an influence on the overall required number of experiments (model discrimination and parameter estimation). In other words, if the parameter estimates are already of a high quality when the

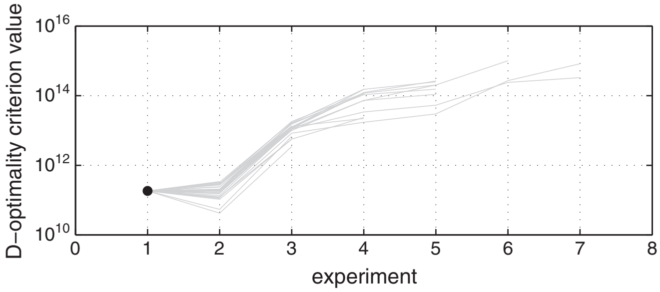

Fig. 4. Evolution of the D-optimality criterion values of model m_5 (in gray) for each of the (thirty) runs of the model discrimination procedure that was initiated with experiment ξ_1^1.

model discrimination procedure ends, less additional experiments will be required to refine the parameter estimates afterwards.

To evaluate the quality of the parameter estimates, the value of the D-optimality criterion is used. This criterion is frequently used in optimal experimental design for parameter estimation (OED/PE) to quantify the information content of an experiment with regard to the parameters of a particular model. Its value is inversely proportional to the volume of the confidence region of the parameter estimates, and experiments that are characterized by a large criterion value are thus expected to bring forth more accurate parameter estimates than experiments with a small criterion value. For more information on the D-optimality criterion, the reader is referred to references cited above where this and other design criteria for OED/PE purposes are described in more detail.

In Fig. 4, the evolution of the D-optimality criterion values of model m_5 are shown in gray for each of the (thirty) runs that were initiated with experiment ξ_1^1 and where the T_d approach was used to design the experiments. At first sight, it may seem strange that the criterion values are not (always) monotonically increasing with the number of performed experiments. Indeed, the D-optimality criterion value represents the information content of the set of experiments and one would expect that this information content can only increase when new information is collected from an additional experiment. However, the decreases in the criterion values can be explained by the fact that the model parameters are re-estimated after performing the new experiment. Since the D-optimality criterion value is dependent on these parameter estimates, a non-monotonic profile can be obtained.

From Fig. 4 and from the discussion held in the previous section, it is clear that the number of required experiments differs among the runs. When the results are represented as in Fig. 4, their interpretation would be hampered and it would be difficult to compare the results obtained for the different starting situations and with the different approaches to OED/MD. Therefore, the results will be presented differently in the following. The median of the D-optimality criterion values will be used to visualize how the quality of the parameter estimates changes during the model discrimination procedure. The number of criterion values from which the median was determined will be indicated and will also be reflected by the size of the bullet symbol (•). The results obtained when the model discrimination procedure was initiated with experiments ξ_1^1 and ξ_1^4 are presented in Figs. 5 and 6, respectively. The results obtained starting from the other preliminary experiments are similar (see [14]), but are not shown here for brevity. In these figures, the median of the D-optimality criterion values for a particular OED/MD approach is shown in black, while the values of the other approaches are shown in gray to facilitate their mutual comparison.

From these figures, one can conclude that the T_a, T_b and T_d approaches perform better than the T_c approach, in the sense that the rate at which the D-optimality criterion values (or the quality of the parameter estimates) increases is higher. Indeed, one can see that the T_c approach is the worst performing one, and its performance is significantly worse (Figs. 5 and 6) when the model discrimination procedure is initialized with preliminary experiments ξ_1^1 and ξ_1^4 (this was not the case for the runs starting from experiment ξ_1^2 (results in [14])). The similarity in the performance of the T_a, T_b approaches on the one hand and T_d approach on the other hand, indicates that the

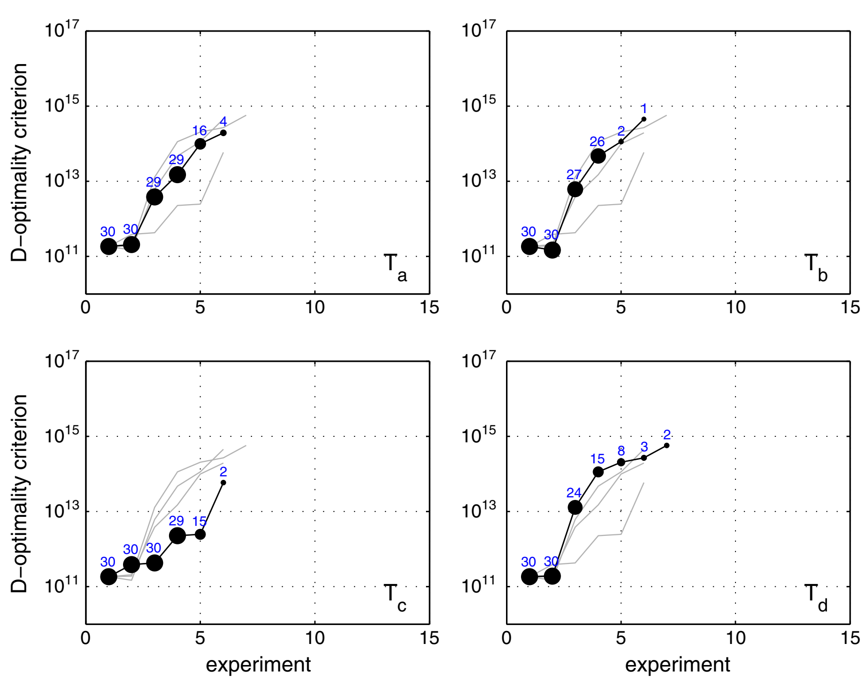

Fig. 5. Evolution of the median D-optimality criterion values of model m_5 for the (thirty) runs of the model discrimination procedure that was initiated with experiment ξ_1^1, for each of the selected approaches for OED/MD. The evolution of the median criterion values of the other approaches is shown in gray to ease the comparison. The number of criterion values from which the median was determined will be indicated by the size of a bullet symbol and the corresponding integer.

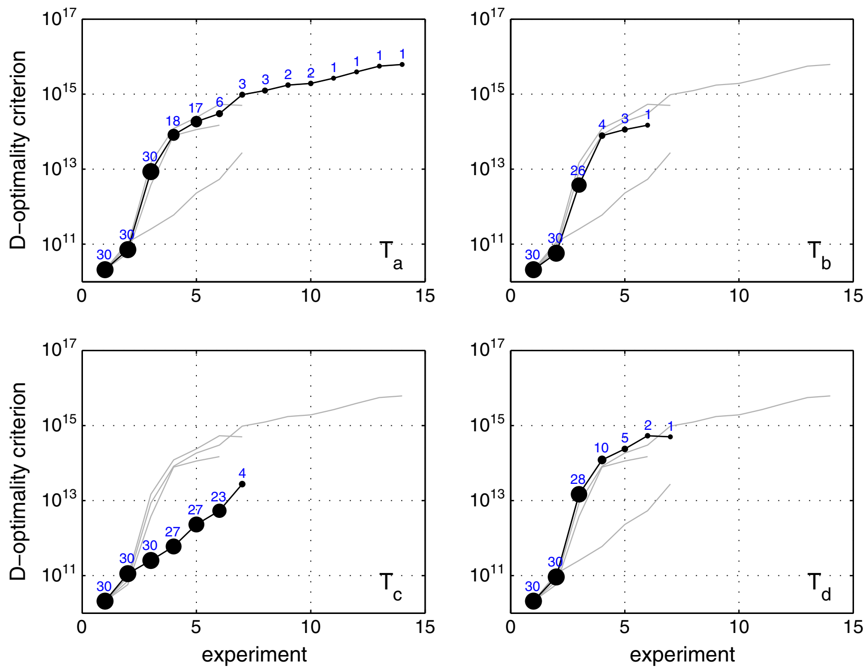

Fig. 6. Evolution of the median D-optimality criterion values of model m_5 for the (thirty) runs of the model discrimination procedure that was initiated with experiment ξ_1^4, for each of the selected approaches for OED/MD. The evolution of the median criterion values of the other approaches is shown in gray to ease the comparison. The number of criterion values from which the median was determined will be indicated by the size of a bullet symbol and the corresponding integer.

information one expects to collect by performing the experiment designed using the T_d approach will reduce the uncertainty on the model predictions such that the experimental design becomes primarily driven by the difference in the model predictions (and the uncertainty on the measurements), as already observed before in [15]. Still, one can state that the T_d approach performs slightly better than the other ones, or, in other words, the T_d approach generally results in experiments with a larger information content with regard to the parameter estimates compared to the other approaches.

That the T_c approach can result in a poor performance with regard to the quality of the parameter estimates, is in agreement with the concepts from which it is derived. As the T_c approach seeks a balance between the difference in the model predictions and the uncertainty associated with it, it will obviously avoid to take samples where the uncertainty on the model predictions is (too) large. However, since uncertainty is to a large extent determined by the sensitivities of these predictions to the values of the model parameters (see Eq. (9)), this also has an impact on the information content of the designed experiment. Indeed, the highest information content with regard to the model parameters is found where these sensitivities are large (as explained in the already cited references on OED/PE). In other words, the T_c approach will not exploit the information present in regions where the model prediction uncertainty is large, unless the difference in the model predictions is significantly larger at these points.

When the T_d approach is used, the information that will be collected on the model parameters when performing the designed experiment is already considered in the evaluation of its discriminatory potential. This (small) conceptual difference with the T_c approach has important consequences. Indeed, experiments that are informative with regard to the model parameters will indirectly contribute to a reduction of the uncertainty on the model predictions. Therefore, the balance between the difference in model predictions and the uncertainty associated with it will shift towards the former compared to the T_c approach. In other words, the regions where the information content with regard to the model parameters is highest, will more likely be exploited by the T_d approach, whereas they will be avoided by the T_c approach.

That the use of the T_d approach may result in improved parameter estimates, was also noted in [38], where it is stated that "*the use of [the anticipatory approach] allows for simultaneous improvement of model discrimination and parameter estimation, as pursued by many researches in the field.*" However, although this statement is in agreement with what was observed in this work, it should not be seen as an absolute truth. Indeed, although the uncertainty on the parameter estimates plays its role in the design of discriminatory experiments, the latter is primarily driven by the difference in the model predictions. It may thus well be that the informative regions with regard to the parameters do not coincide with the regions that are interesting with regard to model discrimination. In such case, the optimal discriminatory experiment may not be informative at all with respect to the parameters and no significant improvement of the parameter estimates would be observed. In addition, it should be kept in mind that the experiments are designed based on the predictions of both rival models. The regions (or experiments) that are informative for model m_i may not be informative at all for model m_j.

3.4. Rate at which inadequate models are identified

In the discussion above, it was assumed that both time and money were available to perform experiments until the most appropriate model was identified. However, in practice, these resources may be limited and the model discrimination procedure may have to be stopped after a particular number of experiments. In this respect, it is important to look at the rate at which inadequate models are identified. Indeed, when the model discrimination procedure is stopped before the most appropriate model is identified, the modeller will have to select it from the models that could not yet be invalidated. It is clear that model selection (where the best model is selected from a set of rival models without performing additional experiments) is less challenging when the number of models to choose from is limited.

For this aspect of the performance evaluation, the number of models that could not yet be invalidated after each step of the sequential procedure will be investigated. The median value of the number of remaining models obtained for the different runs of the model

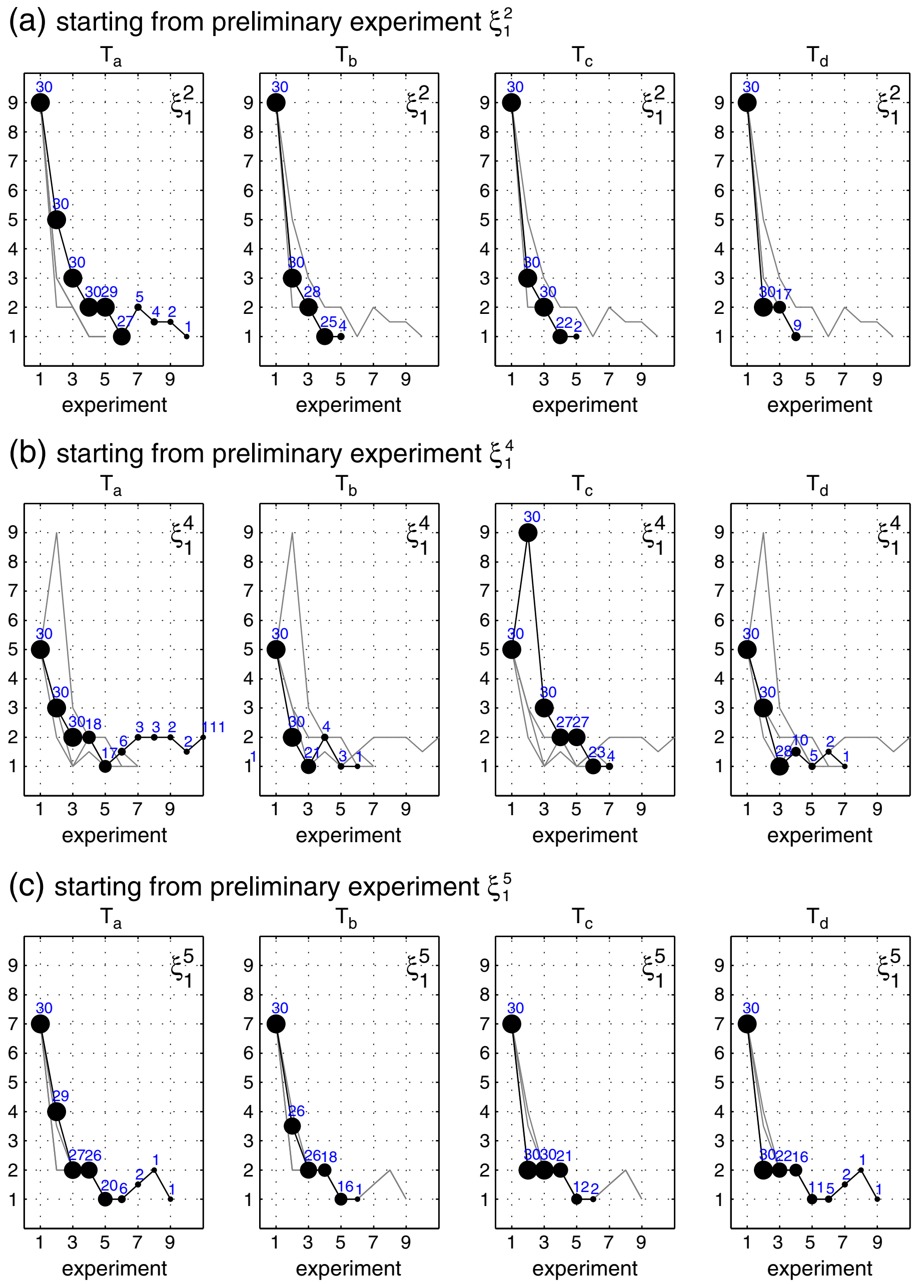

Fig. 7. The median values of the number of adequate models as a function of the number of experiments that have been performed, starting from preliminary experiment ξ_1^2 (a), ξ_1^4 (b) and ξ_1^5 (c). The number of values (runs) from which the median was determined will be indicated by the size of a bullet symbol and the corresponding integer.

discrimination procedure are shown in Fig. 7a, b and c for the scenarios with preliminary experiments ξ_1^2, ξ_1^4 and ξ_1^5, respectively. Note that the results obtained when the model discrimination procedure is initialized with the other preliminary experiments are similar (see [14]), but are not shown here for brevity.

From the results shown in Fig. 7, one can clearly see that, in general, most inadequate models are identified after performing the first optimal discriminatory experiment. Therefore, the median of the number of adequate models remaining after performing the first designed experiment is tabulated in Table 4 and will be used to facilitate the discussion. Although the results presented in Fig. 7 and Table 4 are rather inconclusive, one can still observe that the T_a approach is the worst performing approach in most of the cases (although its performance is not bad). The performance of the T_b and the T_c approaches is comparable, while the T_d approach performs slightly better than the other ones.

Note that an interesting remark can be made when looking at the results presented in Fig. 7b. Here, one can see that the number of adequate models increases from five to nine after performing the

Table 4
The median value of the number of remaining models obtained after performing the first discriminatory experiment designed using the different approaches (T_a, T_b, T_c and T_d). To increase the interpretability of these results, the lowest median values for a given preliminary experiment are indicated in bold.

Preliminary experiment	T_a	T_b	T_c	T_d
ξ_1^1	**2**	**2**	**2**	**2**
ξ_1^2	5	3	3	**2**
ξ_1^3	3	**2**	**2**	**2**
ξ_1^4	3	**2**	9	3
ξ_1^5	3	3.5	**2**	**2**

first discriminatory experiment designed according to the T_c approach. Although this observation only has a minor influence on the performance evaluation of the different approaches, it is an interesting one because it clearly indicates the need to reconsider all models when new experimental data becomes available and illustrates that one should be aware of the possibility that a good model can accidentally be appointed as inadequate. It is therefore recommended to reconsider all models after performing new(ly designed) experiments.

4. Conclusions

The performance of the recently presented anticipatory approach for OED/MD was compared with three older approaches to design optimal discriminatory experiments and it was evaluated by looking at four aspects of discrimination performance: (1) the outcome of the model discrimination procedure, (2) the number of experiments that were required before the model discrimination procedure ended, (3) the evolution of the uncertainty on (or the quality of) the parameter estimates during the model discrimination procedure, and (4) the rate at which the number of adequate models decreases.

The results have shown that it definitely makes sense to design discriminatory experiments (regardless of the approach used to design them), as model discrimination could not be achieved from random experiments. The results also clearly showed that it is important to consider the uncertainties on the measurements and the model predictions when designing the discriminatory experiments. One could also conclude that the approach proposed by Buzzi-Ferraris appeared to be a rather conservative one. Although the true model was always identified as the most appropriate one, more experiments were required compared to the other approaches for parameter estimation. In addition, the information content (with regard to the parameter estimates) of the experiments designed using this approach was often lower than that obtained from the other approaches. With the anticipatory approach, on the other hand, model discrimination was achieved in the lowest number of experiments, and it generally resulted in experiments with a larger information content for parameter estimation. In addition, the rate at which the inadequate models were identified was largest for the anticipatory approach. Based on the results obtained in this case study, one can conclude that the anticipatory approach to design optimal discriminatory experiments is preferred.

Acknowledgments

The authors want to thank the Institute for the Promotion of Innovation by Science and Technology in Flanders for financial support in the framework of SBO-project 040125 (MEMORE). Peter Vanrolleghem holds the Canada Research Chair on Water Quality Modelling.

References

[1] E. Andrianantoandro, S. Basu, D.K. Karig, R. Weiss, Synthetic biology: new engineering rules for an emerging discipline, Molecular Systems Biology (2006) [2:2006.0028].
[2] A.C. Atkinson, A.N. Donev, Optimum experimental design, Oxford University Press, New York, 1992 328 pages.
[3] J.E. Bailey, Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities, Biotechnology Progress 14 (1) (1998) 8–20.
[4] M. Baltes, R. Schneider, C. Sturm, M. Reuss, Optimal experimental design for parameter estimation in unstructured growth models, Biotechnology Progress 10 (5) (1994) 480–488.
[5] J.R. Banga, Optimization in computational systems biology, BMC Systems Biology 2 (2008) 47.
[6] A.L. Burke, T.A. Duever, A. Penlidis, Model discrimination via designed experiments: discrimination between the terminal and penultimate models based on rate data, Chemical Engineering Science 50 (10) (1995) 1619–1634.
[7] A.L. Burke, T.A. Duever, A. Penlidis, An experimental verification of statistical discrimination between the terminal and penultimate polymerization models, Journal of Polymer Science: Part A: Polymer Chemistry 34 (1996) 2665–2678.
[8] G. Buzzi-Ferraris, P. Forzatti, G. Emig, H. Hofmann, Sequential experimental design procedure for model discrimination in the case of multiple responses, Chemical Engineering Science 39 (1) (1984) 81–85.
[9] G. Buzzi-Ferraris, P. Forzatti, P. Canu, An improved version of a sequential design criterion for discriminating among rival multiresponse models, Chemical Engineering Science 54 (2) (1990) 477–481.
[10] B.H. Chen, S.P. Asprey, On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations, Industrial and Engineering Chemistry Research 42 (2003) 1379–1390.
[11] A. de Brauwere, F. De Ridder, R. Pintelon, M. Elskens, J. Schoukens, W. Baeyens, Model selection through a statistical analysis of the minimum of a weighted least squares cost function, Chemometrics and Intelligent Laboratory Systems 76 (2004) 163–173.
[12] D.J.W. De Pauw, P.A. Vanrolleghem, Designing and performing experiments for model calibration using an automated iterative procedure, Water Science and Technology 53 (1) (2006) 117–127.
[13] B. Di Ventura, C. Lemerle, K. Michalodimitrakis, L. Serrano, From *in vivo* to *in silico* biology and back, Nature 443 (2006) 527–533.
[14] B.M.R. Donckels (2009). Optimal experimental design to discriminate among rival dyanamic mathematical models. PhD Thesis. Faculty of Bioscience Engineering. Ghent University. pp. 287.
[15] B.M.R. Donckels, D.J.W. De Pauw, B. De Baets, J. Maertens, P.A. Vanrolleghem, An anticipatory approach to optimal experimental design for model discrimination, Chemometrics and Intelligent Laboratory Systems 95 (1) (2009) 53–63.
[16] B.M.R. Donckels, D.J.W. De Pauw, P.A. Vanrolleghem, B. De Baets, A kernel-based method to determine optimal sampling times for the simultaneous estimation of the parameters of rival mathematical models, Journal of Computational Chemistry 30 (13) (2009) 2064–2077.
[17] B.M.R. Donckels, D.J.W. De Pauw, P.A. Vanrolleghem, B. De Baets, An ideal point method for the design of compromise experiments to simultaneously estimate the parameters of rival mathematical models, Chemical Engineering Science 65 (5) (2010) 1705–1719.
[18] F.J. Doyle III, J. Stelling, Systems interface biology, Journal of the Royal Society, Interface 3 (10) (2006) 603–616.
[19] J. Fisher, T.A. Henzinger, Executable cell biology, Nature Biotechnology 25 (2007) 1239–1249.
[20] G. Franceschini, S. Macchietto, Model-based design of experiments for parameter precision: state of the art, Chemical Engineering Science 63 (19) (2008) 4846–4872.
[21] M. Heinemann, S. Panke, Synthetic biology – putting engineering into biology, Bioinformatics 22 (22) (2006) 2790–2799.
[22] W.G. Hunter, A.M. Reiner, Designs for discriminating between two rival models, Technometrics 7 (1965) 307–323.
[23] H. Kitano, Computational systems biology, Nature 420 (2002) 206–210.
[24] A. Kremling, S. Fischer, K. Gadkar, F.J. Doyle, T. Sauter, E. Bullinger, F. Allgöwer, E.D. Gilles, A benchmark for methods in reverse engineering and model discrimination: problem formulation and solutions, Genome Research 14 (2004) 1773–1785.
[25] C. Kreutz, J. Timmer, Systems biology: experimental design, FEBS Journal 276 (2009) 923–942.
[26] L. Kuepfer, M. Peter, U. Sauer, J. Stelling, Ensemble modeling for analysis of cell signaling dynamics, Nature Biotechnology 25 (2007) 1001–1006.
[27] L. Ljung, System Identification, Theory for the User, Prentice Hall, 1999 608 pages.
[28] S. Marsili-Libelli, S. Guerrizio, N. Checchi, Confidence regions of estimated parameters for ecological systems, Ecological Modelling 165 (2003) 127–146.
[29] R. Mehra, Optimal input signals for parameter estimation in dynamic systems – survey and new results, IEEE Transactions on Automatic Control 19 (6) (1974) 753–768.
[30] B. Mélykúti1, E. August, A. Papachristodoulou, H. El-Samad, Discriminating between rival biochemical network models: three approaches to optimal experiment design, BMC Systems Biology 4 (1) (2010).
[31] C. Michalik, M. Stuckert, W. Marquardt, Optimal experimental design for discriminating numerous model candidates: the AWDC criterion, Industrial and Engineering Chemistry Research 49 (2) (2010) 913–919.
[32] A. Munack, Some improvements in the identification of bioprocesses, in: M.N. Karim, G. Stephanoloulos (Eds.), Modelling and Control of Biotechnical Processes, Pergamon Press, Oxford, 1992, pp. 89–94.
[33] S. Nelander, W. Wang, B. Nilsson, Q.-B. She, C. Pratilas, N. Rosen, P. Gennemark, C. Sander, Models from experiments: combinatorial drug perturbations of cancer cells, Molecular Systems Biology 4 (2008) 216.
[34] T. Ogawa, H. Mori, M. Tomita, M. Yoshino, Inhibitory effect of phosphoenolpyruvate on glycolytic enzymes in *Escherichia coli*, Research in Microbiology 158 (2007) 159–163.
[35] M. Omlin, P. Reichert, A comparison of techniques for the estimation of model prediction uncertainty, Ecological Modelling 115 (1) (1999) 45–49.
[36] B. Palsson, The challenges of *in silico* biology, Nature Biotechnology 18 (2000) 1147–1150.
[37] M. Schwaab, F.M. Silva, C.A. Queipo, A.G. Barreto Jr., M. Nele, J.C. Pinto, A new approach for sequential experimental design for model discrimination, Chemical Engineering Science 61 (2006) 5791–5806.
[38] M. Schwaab, J.L. Monteiro, J.C. Pinto, Sequential experimental design for model discrimination. Taking into account the posterior covariance matrix of differences between model predictions, Chemical Engineering Science 63 (2008) 2408–2419.
[39] M.A.B. Ternbach, C. Bollman, C. Wandrey, R. Takors, Application of model discriminating experimental design for modeling and development of a fermentative fed-batch L-valine production process, Biotechnology and Bioengineering 91 (3) (2005) 356–368.

[40] N.A.W. Van Riel, Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments, Briefings in Bioinformatics 7 (4) (2006) 364–374.
[41] P. Vanrolleghem, M. Van Daele, Optimal experimental design for structure characterization of biodegradation models: on-line implementation in a respirographic biosensor, Water Science and Technology 30 (4) (1994) 243–253.
[42] P.A. Vanrolleghem, D. Dochain, Bioprocess model identification, in: J.F.M. Van Impe, P.A. Vanrolleghem, D.M. Iserentant (Eds.), Advanced Instrumentation, Data Interpretation, and Control of Biotechnological Processes, Kluwer AcademicPublishers, Dordrecht, 1998, pp. 251–318.
[43] E. Walter, L. Pronzato, Identification of parametric models from experimental data, Springer–Verlag, Berlin, Heidelberg, New York, 1997 413 pages.