Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

A framework for model-based design of experiments in the presence of continuous measurement systems

Federico Galvanin, Massimiliano Barolo, Fabrizio Bezzo

CAPE-Lab – Computer-Aided Process Engineering Laboratory
Dipartimento di Principi e Impianti di Ingegneria Chimica, Università di Padova
via Marzolo, 9 35131 Padova, Italy (e-mail: federico.galvanin@unipd.it; fabrizio.bezzo@unipd.it)

Abstract: Model-based design of experiments (MBDoE) techniques are a useful tool to maximise the information content of experimental trials when the purpose is identifying the set of parameters of a deterministic model in a statistically sound way. When samples are collected in a discrete way, the formulation of the optimal design problem is based on the maximisation of the expected information, usually calculated from discrete forms of the Fisher information matrix. However, if a continuous measurement system is available, information can be acquired gradually in a continuous way, and a new MBDoE approach is required to take into account the specificity of the measurement system. In this paper a novel design criterion is formulated by optimising a continuous measurement of the Fisher information matrix, with the purpose of reaching a statistically satisfactory estimation of model parameters in the easiest and quickest way. The benefits of the proposed strategy are discussed through a simulated case study, where the effectiveness of the design is assessed by comparison to a standard MBDoE approach.

Keywords: model based design of experiments, parameter estimation, model identification, continuous measurement systems

1. INTRODUCTION

Dynamic deterministic models are used to describe physical systems through the statement of laws and correlations in the form of a system of differential and algebraic equations (DAEs). Once a dynamic model structure is found adequate to represent a physical system, a series of identification experiments need to be carried out to estimate the set of parameters of the model in the most precise and accurate way. The identification procedure can be costly and very time consuming because the system may exhibit identifiability issues, or there may be a mismatch between the model and the actual system to be represented, or data may be difficult and expensive to obtain. Model-based design of experiments (MBDoE) techniques (Mehra, 1974; Titterington, 1980; Pukelsheim, 1993) represent a valuable and consolidated tool for the rapid assessment and development of dynamic deterministic models, allowing for the maximisation of the information content of the experimental trials in order to assist the parameter identification task. MBDoE techniques are usually carried out in a sequential way, and three fundamental steps are required to determine the model parameters: *i*) the design of the experiment, usually carried out by maximising a measurement function of the expected information (i.e. the information as predicted by the model); *ii*) the execution of the experiment according to the planned experimental conditions; *iii*) the estimation of the model parameters, which provides the statistical assessment of the estimate and where the actual information (i.e. the information as provided by the experiment in the form of collected data) is exploited. This procedure, leading to a progressive reduction of the uncertainty region of the model parameters, can be iterated until a satisfactory parameter estimation is achieved. The effectiveness of model-based design procedures has been demonstrated in several applications (Franceschini and Macchietto, 2008).

In practical experimental environments information is usually acquired from the experiment through discrete collection (sampling) of data. Consequently, when planning an experiment with an MBDoE technique, the expected information being maximised is expressed through a discrete form of the Fisher information matrix (Zullo, 1991), which allows for the optimal allocation of sampling points. However, the rate at which information is acquired depends on the measurement equipment and on the capability to represent and exploit the system dynamics in terms of experimental data. The effect of sampling rate and measurement precision on information evaluation and design effectiveness has been extensively studied in the literature (Emery *et al.*, 2002). Moreover, recently the management of the information dynamics has been discussed to assess how it can influence the duration and the effectiveness of an identification experiment planned by MBDoE techniques (Galvanin *et al.*, 2009a).

The rate of information acquisition may be limited by the experimental budget (e.g., number and duration of experiments, number and type of measurements) and/or by the specific choices and possibilities set by the experimenter. As a result, the experiment design activity usually takes into account the limitations on laboratory facilities and is managed accordingly. When monitoring a process, a number of system outputs (e.g. temperatures and pressures) are measured (in practice) in a continuous way, while other responses, typically concentration measurements, can only be acquired by discrete sampling at a significantly reduced sampling frequency. However, recent advances in sensors technology allow for the development of continuous monitoring systems particularly suitable for concentration measurements. These measurement systems become particularly suitable for monitoring systems where complex dynamics and poor observability can make both the model identification and the system control procedures very complicated tasks. For instance, continuous measurement systems have been developed for monitoring concentrations in biological processes adopting near-infrared spectroscopy (Tosi *et al.*, 2008) or on line respirometry techniques (Dias *et al.*, 2009).

When performing an MBDoE activity, the mathematical formulation of the expected information being maximised should take into account the specificity of the given sampling system. If the samples are collected very frequently, the measure of the actual information gained from the experiment can be approximated by a continuous profile over the experimental horizon. Following this premise, a novel design criterion involving a dynamic MBDoE (DMBDoE) approach is presented and discussed in this paper. The proposed design technique is suitable for systems in which continuous (or highly frequent) measurements are available. The optimal design problem is formulated by optimising a continuous dynamic measurement function of the Fisher information matrix with the purpose of reaching a statistically satisfactory estimation of model parameters in the easiest and quickest way. The applicability to nonlinear dynamic systems and the effectiveness of the proposed DMBDoE approach are illustrated via a simulated case study.

2. THE METHODOLOGY

Let us consider a process described by the set of DAEs of the form:

$$\begin{cases} \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \mathbf{w}, \theta, t) = 0 \\ \hat{\mathbf{y}}(t) = \mathbf{g}(\mathbf{x}(t)) \end{cases} \tag{1}$$

with the set of initial conditions $\mathbf{x}(0) = \mathbf{x}_0$, where $\mathbf{x}(t)$ is the N_x-dimensional vector of time-dependent state variables, $\mathbf{u}(t)$ and $\mathbf{w}$ are the time-dependent and time-invariant control variables (of dimensions N_u and N_w), respectively, $\boldsymbol{\theta}$ is the N_θ-dimensional set of unknown model parameters to be estimated, and t is time. The symbol ^ is used to indicate the estimate of a variable (or of a set of variables): thus, $\mathbf{y}(t)$ is the vector of measured values of the outputs, while $\hat{\mathbf{y}}$ is the vector of the corresponding values estimated by the model. Model-based experiment design procedures aim at decreasing the model parameter uncertainty region predicted by model (1) as the solution of the optimisation problem

$$\boldsymbol{\varphi}^{\text{opt}} = \arg\min_{\boldsymbol{\varphi}} \left\{ \psi\left[\mathbf{V}_\theta\left(\boldsymbol{\theta}, \boldsymbol{\varphi}\right) \right] \right\} = \arg\min_{\boldsymbol{\varphi}} \left\{ \psi\left[\mathbf{H}_\theta^{-1}\left(\boldsymbol{\theta}, \boldsymbol{\varphi}\right) \right] \right\} \tag{2}$$

by acting on the experiment design vector $\boldsymbol{\varphi}$:

$$\boldsymbol{\varphi} = \left[\mathbf{y}_0, \mathbf{u}(t), \mathbf{w}, \mathbf{t}^{sp}, \tau \right]^T \tag{3}$$

which includes the N_y-dimensional set of initial conditions $\mathbf{y}_0$ on the measured variables, the duration of the experiment τ, the continuously manipulated inputs $\mathbf{u}(t)$, usually approximated by a discrete (piecewise constant or piecewise linear) function, and the set of time invariant manipulated inputs $\mathbf{w}$. The set of time instants at which the output variables are sampled is also a design variable, and is expressed through the n_{sp}-dimensional vector $\mathbf{t}^{sp}$ of sampling times. The experiment is designed so as to minimise a measurement function ψ of $\mathbf{V}_\theta$ (the variance-covariance matrix of model parameters) or, equivalently, to maximise a measurement function ψ of $\mathbf{H}_\theta$ (the dynamic information matrix). The particular form of the measurement function represents the chosen design criterion in order to maximise the expected information content of the experiment as predicted by the model. The most common design criteria are the so-called alphabetical ones, i.e. A-, D-, E-optimal criteria (Pukelsheim, 1993), or they are based on singular values decomposition (Galvanin *et al.*, 2007; Zhang *et al.*, 2009). The dynamic information matrix is usually expressed by the discrete dynamic form of the Fisher information matrix as proposed by Zullo (1991):

$$\mathbf{H}_\theta\left(\boldsymbol{\theta}, \boldsymbol{\varphi}\right) = \mathbf{H}_\theta^0 + \sum_{k=1}^{n_{sp}} \sum_{i=1}^{N_y} \sum_{j=1}^{N_y} s_{ij} \left[\frac{\partial \hat{y}_i\left(t_k\right)}{\partial \theta_l}^{\mathrm{T}} \frac{\partial \hat{y}_j\left(t_k\right)}{\partial \theta_m} \right]_{l,m=1\ldots N_\theta} = \mathbf{H}_\theta^0 + \sum_{k=1}^{n_{sp}} \mathbf{M}_k \tag{4}$$

In (4) s_{ij} is the ij-th element of the $N_y \times N_y$ inverse matrix of measurements error, $\mathbf{M}_k$ represents the amount of information that can be recovered from the k-th sample and $\mathbf{H}_\theta^0$ is the prior dynamic information matrix, taking into account the preliminary statistical information about the parametric system before each trial is carried out. Considering an A-optimal design criterion (i.e. focusing on the trace tr of the dynamic information matrix), an upper limit curve on the expected information (Fig. 1) can be characterized as the number of samples becomes very large:

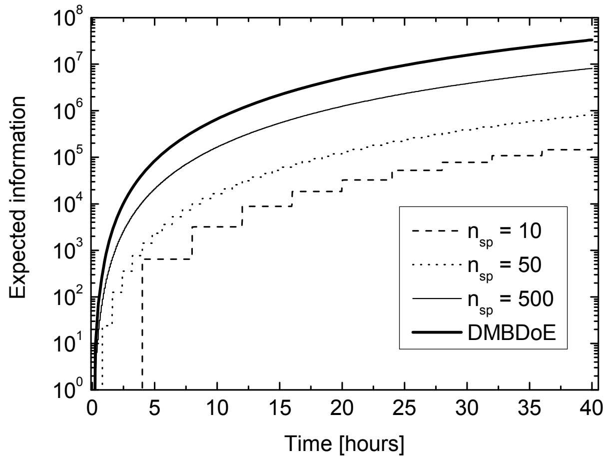

Fig. 1. Effect of the number of samples on the evaluation of the expected information.

$$\lim_{n_{sp}\to\infty} tr\left(\sum_{k=1}^{n_{sp}}\left[\mathbf{M}_k\Delta t\right]+\mathbf{H}_0^0\right) = \lim_{n_{sp}\to\infty}\sum_{k=1}^{n_{sp}} tr\left[\mathbf{M}_k\Delta t\right]+K$$
$$= \int_0^{\tau} tr\left[\mathbf{M}(t)\right]dt + K \tag{5}$$

where Δt is the sampling interval, and here is assumed to be fixed (and thus not optimised by design). In (5) K is a constant term quantifying the prior information, while the trace of $\mathbf{M}(t)$ allows for the dynamic evaluation of the expected information. This new metric of the expected information is suitable for systems where the measurements can be deemed continuous (i.e., where information can be collected at a frequency that is much higher than the characteristic frequency of the process). A novel design criterion for the dynamic model based design of experiments can thus be introduced:

$$\boldsymbol{\varphi}^{\mathrm{opt}} = \arg\max_{\boldsymbol{\varphi}}\left[\int_0^{\tau} tr\left[\mathbf{M}(t)\right]\mathrm{d}t\right] \quad . \tag{6}$$

Basically, the DMBDoE criterion aims at maximising the area underneath the curve of the dynamic expected information, while a standard A-optimal MBDoE criterion aims at maximising the sum of the information content of each single sampling point (Fig. 2).

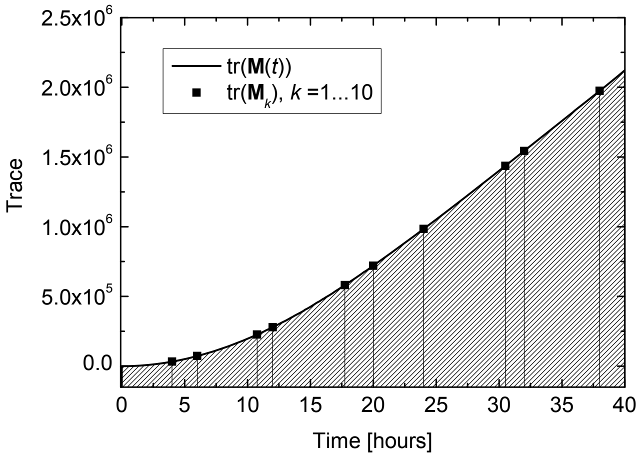

Figure 2. Dynamic evaluation of expected information.

The benefit of adopting (6) is that the design objective function becomes inherently dynamic as is the experiment itself, thus allowing for a continuous exploitation of the available information (from the very beginning of the experiment). This design criterion can be usefully exploited also in a sequential MBDoE framework or by adopting a redesign strategy (Galvanin *et al.*, 2009b) within a proper time window. After maximising the expected information with (2) or (6) by acting on the components of the design vector $\boldsymbol{\varphi}$, the experiment is performed and the actual information is exploited by carrying out a parameter estimation on the collected data. A new information profile will be generated (the "actual information" profile), usually different from the expected profile maximised by design. If the model is a reliable representation of the process, the mismatch between the expected and the actual information can be exclusively attributed to the parametric mismatch between the model and the real system.

3. CASE STUDY

3.1 Bioreactor model

The methodology discussed in the previous section is applied to a simulated biomass fermentation process that appeared in several papers on the subject (Espie and Macchietto, 1989; Asprey and Macchietto, 2002). Assuming Monod-type kinetics for biomass growth and substrate consumption, the system is described by the following set of DAEs:

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= \left(y-u_1-\theta_4\right)x_1 \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= -\frac{yx_1}{\theta_3}+u_1\left(u_2-x_2\right) \\
y &= \frac{\theta_1 x_2}{\theta_2+x_2}
\end{aligned} \tag{7}$$

where x_1 is the biomass concentration (g/L), x_2 is the substrate concentration (g/L), u_1 is the dilution factor (h^{-1}), and u_2 is the substrate concentration in the feed (g/L). The experimental conditions that characterise an experiment are the dilution factor u_1 (range 0.05-0.20 h^{-1}) and the substrate concentration in the feed u_2 (range 5-35 g/L). These conditions are approximated by piecewise constant profiles over 8 switching intervals (the duration of each interval is allowed to be between 1 and 20 h). The initial biomass and substrate concentration $x_1(0)$ and $x_2(0)$ are set to 1.4 g/L and 0 g/L, respectively. It is assumed that both x_1 and x_2 can be measured during the experiment. The final objective is to design a single experiment (lasting τ = 40 h) to yield the best possible information for the estimation of the four parameters θ_i.

3.2 Experiment design configurations

Two experiment design configurations are considered and compared in this study:

1. MBDoE: a standard E-optimal designed experiment with (2) as the objective function; the design also optimises the allocation in time of n_{sp} = 10 samples (the elapsed time between any two sampling points is allowed to be between 1 and 20 h);
2. DMBDoE: a dynamic experiment design is performed by adopting (6) as the objective function; it is supposed that the measurements are available frequently (every 10 min).

Synthetic "experimental" data are obtained by simulation with $\boldsymbol{\theta}$ = [0.310, 0.180, 0.550, 0.050]$^{\mathrm{T}}$ as the true parameters and adding Gaussian noise with a constant relative variance of 0.03 (case A, "moderate noise level") and 0.20 (case B, "high noise level") to the output measurement. The initial guess for the model parameters' values is set to $\boldsymbol{\theta}^0$ = [1.000, 1.000, 1.000, 1.000]$^{\mathrm{T}}$. Since $\boldsymbol{\theta}$ is obviously unknown in practice, results of the parameter estimation are given in terms of the *a-posteriori* statistics obtained after performing a maximum likelihood parameter estimation. The quality of the final estimates is assessed by observing for each parameter the interval of estimation confidence and the t-value statistic obtained after the optimally designed experiments have been executed and the model parameters re-estimated with the new data. For a reliable parameter estimation the t-value must be greater than a computed reference value derived from a Student t-distribution (t-test) with $n_{sp}-N_\theta$ degrees of freedom. Although it has been shown in literature that an E-criterion can be extremely inefficient when dealing with the identification of linear regression models (Dette, 1997), this criterion was used in the standard MBDoE because it was proven as the most effective design approach for this specific nonlinear case study (Asprey and Macchietto, 2000). Even if not shown here for the sake of brevity, it has nonetheless been verified that an A-optimal design criterion would provide very similar optimal excitation patterns.

3.2 Case A: moderate noise level

When measurements are available with moderate noise, both MBDoE approaches allow reaching a statistically satisfactory parameter estimation (Table 1), but DMBDoE ensures a dramatically better confidence on the final estimate, thanks to the higher rate of information acquisition.

Table 1. Case A: comparison of parameter estimations for different design configurations (the reference t-value is 1.74 for MBDoE and 1.65 for DMBDoE estimation).

	MBDoE-A	DMBDoE-A
Estimate	[0.3064 0.2015 0.4955 0.0448]^T	[0.3154 0.1762 0.5792 0.0520]^T
Conf. Interval (95%)	[±0.0118 ±0.0540 ±0.1250 ±0.0134]	[±0.0010 ±0.0014 ±0.0048 ±0.0005]
t-values	[26.04 3.73 3.96 3.350]	[297.20 125.10 119.80 103.50]

Interestingly (Fig 3. and Fig. 4), in a DMBDoE approach the design is such that the system is excited at the very beginning of the experiment in order to increase the information content of the samples being acquired as soon as the experiment starts. On the contrary, the excitation provided by MBDoE is mainly concentrated in the second half of the trial. This clear difference on the excitation policy has a significant effect on the distribution of information along the experimental horizon. In fact, the dynamics of the actual information is completely different for the two design configurations (Fig. 5). A minimum required information limit based on the A-optimal design criterion can be defined by considering a mean standard deviation of 10% on the final estimate of model parameters.

It can be noticed that the second half of the experiment as planned by DMBDoE does not deliver an appreciable contribution to the overall information, which is concentrated at the very beginning of the trial. A maximum on the actual information is reached around t = 11 h, and subsequently the increment on information is negligible. As a result, the experiment planned by DMBDoE could be stopped well before $t = \tau$ as a statistically satisfactory parameter estimation would be reached much earlier.

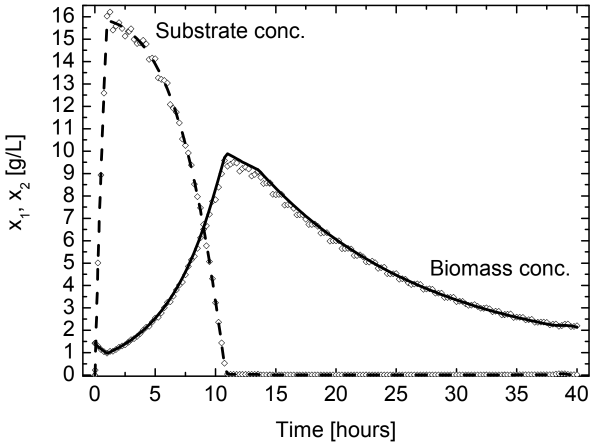

Figure 3. Case A: biomass and substrate concentration profiles as predicted by the model after the parameter identification of the DMBDoE planned experiment; biomass and substrate concentration measurements are indicated by diamonds.

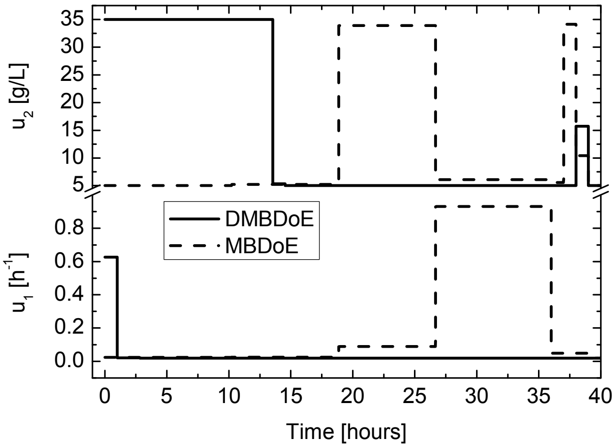

Figure 4. Case A: profiles of the manipulated inputs as optimised by the two different design strategies.

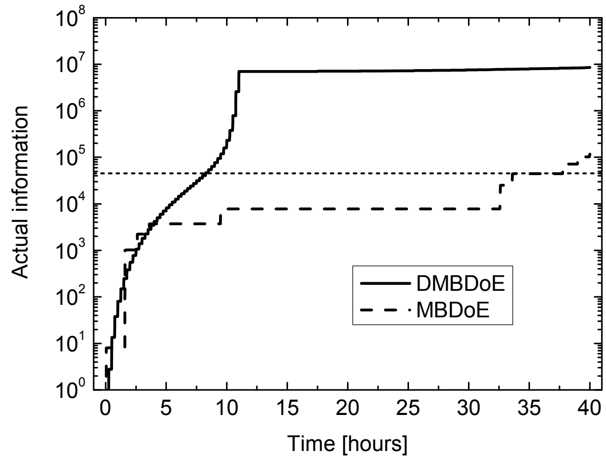

Figure 5. Case A: profiles of actual information for a standard MBDoE and for DMBDoE as given by the summation term of (5); the dotted line represents the A-optimal information limit for a 10% deviation on the final estimate.

On the other hand, the experiment planned by a standard MBDoE technique requires approximately the full length of the experiment for a statistically sound parameter estimation. These results clearly show how a DMBDoE approach can be usefully adopted to shorten the experiment duration if a measurement system capable of providing precise and frequent measurements is available.

3.3 Case B: high noise level

When a high noise level is present in the measurements, a standard MBDoE approach is not sufficient to provide a statistically sound parameter estimation (Table 2) with a single experiment.

Table 2. Case B: comparison of parameter estimations for different design configurations. Superscript * indicates *t*-values failing the test (reference *t*-value is 1.74 for MBDoE and 1.65 for DMBDoE estimation).

	MBDoE-B	DMBDoE-B
Estimate	$[0.3047\ 0.1970\ 0.5099\ 0.0436]^T$	$[0.3040\ 0.1757\ 0.5327\ 0.0491]^T$
Conf. Interval (95%)	[±0.2000 ±0.3369 ±0.8502 ±0.1903]	[±0.0292 ±0.0137 ±0.0745 ±0.0087]
t-values	$[1.52^*\ 0.58^*\ 0.59^*\ 0.23^*]$	[10.40 12.85 7.15 5.62]

On the contrary, the DMBDoE strategy appears to be less sensitive to the level of the measurement noise and provides a statistically satisfactory estimation for all parameters after a single experiment. The level of excitation provided by MBDoE is significantly higher than the one provided by DMBDoE (Fig. 6 and Fig. 7), but it is still concentrated (as in case A) in the second part of the experiment (after 10 h). Analysing the actual information profiles (Fig. 8) it can be noticed how the information acquired through discrete samples is not sufficient to guarantee a statistically sound parameter estimation. Conversely, when a dynamic design is carried out, the information exploited at the very beginning of the experiment is sufficient to reach a statistically sound parameter estimation in the first half of the experiment.
Additionally, it can be observed that even if measurements are noisy, the new approach (DMBDoE-B, Table 2) can provide a sounder parameter estimation of θ_2, θ_3 and θ_4 than the one provided by a standard MBDoE with clean measurements (MBDoE-A, Table 1). Thus, particular attention should be made by the experimenter on choosing the proper measurement system: a continuous measurement system, even if providing noisy data, might be more suitable for model development and validation than a more precise but discrete approach.

3.4 Additional discussion

As already mentioned, from the analysis of the dynamics of the actual information when a DMBDoE approach is pursued (Fig. 6 and Fig. 4) it seems that the experiment can be stopped well before its planned duration, while maintaining at the same time a satisfactory parameter estimation.

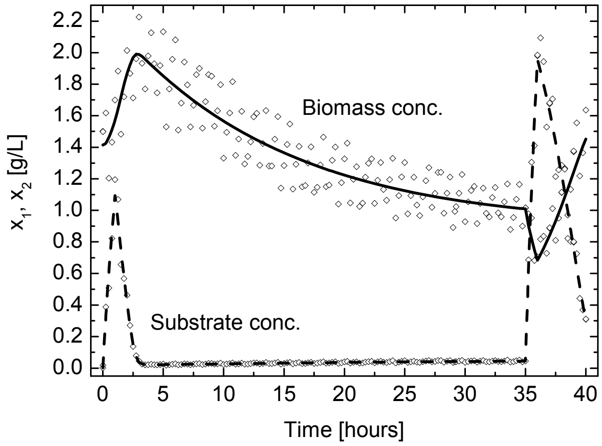

Figure 6. Case B: biomass and substrate concentration profiles as predicted by the model after the parameter identification of the DMBDoE planned experiment; biomass and substrate concentration measurements are indicated by diamonds.

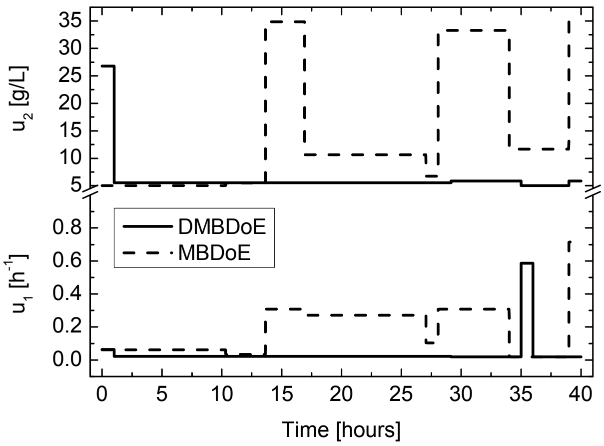

Figure 7. Case B: profiles of the manipulated inputs as optimised by the two different design strategies.

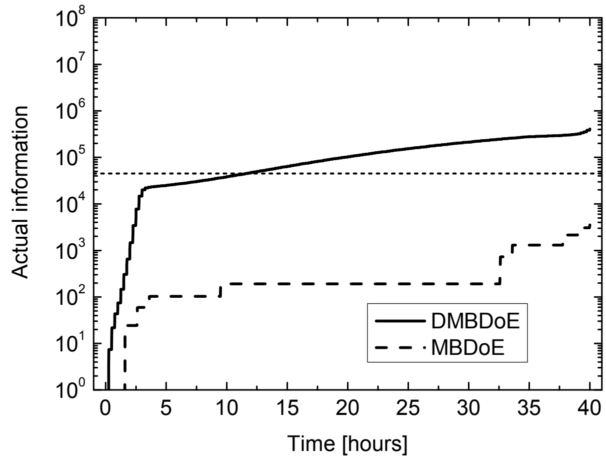

Figure 8. Case B: profiles of actual information for a standard MBDoE and for DMBDoE as given by the summation term of (5); the dotted line represents the A-optimal information limit for a 10% deviation on the final estimate.

Table 3. Case B: comparison of parameter estimations for DMBDoE planned minimal-length experiments (the reference t-value is 1.74 for MBDoE-A2 and 1.66 for DMBDoE-B2 estimation).

	DMBDoE-A2	DMBDoE-B2
Estimate	$[0.3022\ 0.1809\ 0.5359\ 0.0423]^T$	$[0.2872\ 0.1762\ 0.4999\ 0.0457]^T$
Conf. Interval (95%)	[±0.0208 ±0.0050 ±0.0410 ±0.020]	[±0.0643 ±0.0217 ±0.1381 ±0.0192]
t-values	[14.48 35.59 13.33 2.115]	[4.464 8.118 3.619 2.381]
Duration (h)	12.75	14.50

This behaviour is confirmed by performing an additional parameter estimation on two DMBDoE experiments (Table 3) where the trial is stopped as soon as the t-test is satisfied for the entire parametric set.

The results show significant benefits in terms of time saving. For both experiments the precise estimation of θ_4 is critical. When moderately noisy measurements are available (DMBDoE-A2), the approach allows reducing the experiment duration from 40 to 12.75 hours. When only noisy measurements are available (DMBDoE-B2) a slightly longer experiment is required (14.5 h), but still the experiment length can be significantly reduced.

4. FINAL REMARKS

A novel design criterion (DMBDoE), suitable for systems where continuous measurements are available, has been proposed and analysed in this paper. The parametric identification of a nonlinear bioreactor model is significantly improved when DMBDoE is used. The novel design technique allows exploiting different information patterns where the information is maximised since the very beginning of the trial. One additional benefit is that it is possible to reduce the overall duration of the experiment in a substantial way. Future work will assess the effectiveness of the technique to identify the set of parameters of more complex systems where continuous measurement policies can be undertaken.

AKNOWLEDGEMENTS

The financial support of the University of Padova under Progetto di Ateneo 2007 (cod. CPDA074133, "A process systems engineering approach to the development of an artificial pancreas for subjects with type-1 diabetes mellitus") and under Project CPDR095313-2009 ("Towards the development of an artificial pancreas for diabetes mellitus care: optimal model-based design of experiments for parameter identification of physiological models") is gratefully acknowledged.

REFERENCES

Asprey, S. P.; Macchietto, S. (2000). Statistical tools for optimal dynamic model building. *Computers Chem. Eng.*, **24**, 1261-1267.

Dette, H. (1997). E-optimal design for regression models with quantitative factors: a reasonable choice? *Canadian J. Statistics*, **25**, 531-543.

Dias, J. M. L., Pardelha, F., Eusebio, M., Reis, M. A. M., Oliveira, R. (2009). On-line monitoring of PHB production by mixed microbial cultures using respirometry, titrimetry and chemometric modelling. *Process Biochem.*, **44**, 419-427.

Emery, A. F., Blackwell B. F., Dowding K. J. (2002). The relationship between information, sampling rates, and parameter estimation models. *J. Heat Transfer*, **124**, 1192-1199.

Espie, D., S. Macchietto (1989). The optimal design of dynamic experiments. *AIChE J.*, **35**, 223-229.

Franceschini, G., Macchietto S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.*, **63**, 4846-4872.

Galvanin, F., Macchietto S., Bezzo F. (2007). Model-based design of parallel experiments. *Ind. Eng. Chem. Res.*, **46**, 871-882.

Galvanin, F., Barolo M., Macchietto S. and F. Bezzo (2009a). On the optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus. In: *Computer-Aided Chemical Engineering 27*, 10th International Symp. on Process Systems Engineering, Elsevier, Amsterdam (The Netherlands), 183-188.

Galvanin, F., Barolo M. and Bezzo F. (2009b). Online model-based redesign of experiments for parameter estimation in dynamic systems. *Ind. Eng. Chem. Res.*, **48**, 4415-4427.

Mehra, K., R. (1974). Optimal input signals for parameter estimation in dynamic systems: survey and new results. *IEEE Trans. Automat. Contr.*, **19**, 753-768.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. J. Wiley & Sons, New York (U.S.A.).

Titterington, D. M. (1980). Aspects of optimal design in dynamic systems. *Technometrics*, **22**, 287-299.

Tosi, S., Rossi, M., Tamburini, E., Vaccari, G., Amaretti, A., Matteuzzi, D. (2008). Assessment of In-line Near-infrared Spectroscopy for Continuous Monitoring of Fermentation Processes. *Biotechnol. Prog.*, **19**, 1816-1821.

Zhang, Y., Edgar T. F. (2008). PCA combined model-based design of experiments (DOE) criteria for differential and algebraic system parameter identification. *Ind. Eng. Chem. Res.*, **47**, 7772-7783.

Zullo, L. (1991). Computer aided design of experiments. An engineering approach. Ph.D. Thesis, The University of London, U.K.