Control and Cybernetics

vol. **35** (2006) No. 2

A new OSS design based on parameter sensitivity to changes in measurements

by

R. Kalicka[1] **and D. Bochen**[2]

[1] Technical University of Gdańsk
Narutowicza 11/12, 80-952 Gdańsk, Poland
[2] Technical University of Białystok
Wiejska 45D, 15-351 Białystok, Poland

Abstract: As a result of the severe practical and ethical constraints imposed on medical measurements, the parameter estimation procedure designed for diagnosis and therapy is often a difficult problem. When blood sampling provides the data, the number of samples and the observation interval should be minimized. Designing an experiment for parameter estimation requires a step known as quantitative experiment design, usually preceded by a step called qualitative experiment design. The latter answers if a model is identifiable under particular experimental conditions. The former is suitable for the purpose of obtaining the maximum information from the data to be collected. An experiment design is based on the optimization of a suitable criterion formulated with respect to the analyzed variables of the experiment (input shape, sampling schedule). This paper considers sampling schedule design. New criteria for new optimal sampling schedules (OSS) have been formulated on the basis of sensitivity function. These are referred to as S-OSS and RS-OSS designs. The results of optimization for both criteria are compared with the result obtained with a reputable established D-optimal design based on the Fisher information matrix. By showing the results of S-OSS and RS-OSS design we can present the reliability and efficiency of the new criteria in comparison to D-OSS design. Illustrative examples are presented.

Keywords: optimal sampling schedule, sensitivity criterion.

1. Introduction

Advanced mathematical modelling and identification methods are necessary for the investigation of biomedical system kinetics. Numerous limitations are imposed on experimental conditions, the most common of which concern the number of input and output samples. When blood is the medium being sampled, the total amount of plasma that can be drawn from the subject is limited. This

requires that as short a measurement interval and as small a number of samples as possible be established for model identification. This falls within the scope of sampling schedule optimization.

Let us consider compartmental models of non-linear regression (model) functions $y_m(\mathbf{p}, t)$, with parameter vector $\mathbf{p}$ and discrete-time noisy measurements. For single-output models these measurements are:

$$y\left(t_i\right) = y_m\left(\mathbf{p}, t_i\right) + e\left(t_i\right);\ i = 1, \ldots, N;\ \mathbf{p} = [p_1, \ldots, p_n] \tag{1}$$

where N and n are the numbers of measurements and number of model parameters, respectively.

Measurement noise $e(t_i)$ - zero mean, variance $\sigma^2(t_i)$ - is independent of model parameters. It is normally distributed at each t_i and all $e(t_i)$ are uncorrelated with the others. It is assumed that a nominal "true" parameter vector is known on the basis of intuitive experiment with a sufficiently large number of measurements. The aim is to design for $y(t_i)$ a minimal optimal sampling schedule ($MOSS$), which consists of a minimum of measured samples N_{MOSS}, equal to the number of model parameters: $N_{MOSS} = n$.

The Cramer-Rao theorem, also known as information inequality, gives the relationship between the obtainable precision of the parameter estimates $\hat{\mathbf{p}}$ and the amount of information concerning these parameters available in the noisy data $\mathbf{y}(t) = f(\mathbf{p}, t) + e(t)$. The Cramer-Rao inequality gives a lower bound (Walter and Pronzato, 1990) for the asymptotic covariance matrix of any unbiased estimate $\hat{\mathbf{p}}$ of $\mathbf{p}$:

$$\begin{aligned} &\operatorname{cov}\left(\mathbf{p}\right) = E\left[\left(\mathbf{p} - \hat{\mathbf{p}}\right)\left(\mathbf{p} - \hat{\mathbf{p}}\right)^T\right] \geq \mathbf{M}^{-1}\left(\mathbf{p}\right) \\ &\mathbf{M}\left(\mathbf{p}\right) \stackrel{def}{=} E\left\{\left[\frac{\partial \ln F(\mathbf{y}|\mathbf{p})}{\partial \mathbf{p}}\right]^T \cdot \left[\frac{\partial \ln F(\mathbf{y}|\mathbf{p})}{\partial \mathbf{p}}\right]\right\}. \end{aligned} \tag{2}$$

$\mathbf{M}(\mathbf{p})$ is Fisher information matrix, $F(\mathbf{y}|\mathbf{p})$ is conditional probability density and E is the expectation operator. The expression $F(\mathbf{y}|\mathbf{p})$ is a complicated function of $\mathbf{p}$. It simplifies (Godfrey, 1983) when samples of measurement error $e_k = e(t_k)$ at each time point are uncorrelated and have identical normal distribution

$$\begin{aligned} &E\left(\mathbf{e}\right) = [0, 0, \ldots 0] = \mathbf{0} \\ &E\left(e_k^2\right) = \sigma_k^2;\ E\left(e_k, e_r\right) = 0;\ k \neq r;\ k, r = 1, 2, \ldots, N \end{aligned} \tag{3}$$

In most practical applications (Godfrey 1983, Brandt, 1999), when the noise is assumed to be white, so that $\mathbf{R}$ is diagonal

$$\mathbf{R} = \begin{bmatrix} \frac{1}{\sigma_1^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sigma_N^2} \end{bmatrix} = \left[\frac{1}{\sigma_i^2}\right];\ i = 1, 2, \ldots, N \tag{4}$$

one obtains the well-known expression

$$\mathbf{M}(\mathbf{p}) = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \cdot \frac{\partial y_m(\mathbf{p}, t_i)}{\partial \mathbf{p}} \cdot \frac{\partial y_m(\mathbf{p}, t_i)}{\partial \mathbf{p}^T} \ . \tag{5}$$

Classical criteria of optimality are scalar cost functions $\Phi(\mathbf{M})$. Expression (5) is then the starting point for the classical mathematical theory of quantitative experiment design. A general cost function $\Phi(\mathbf{M})$ is defined (Walter and Pronzato, 1990; Zarrop, 1979) as:

$$\begin{aligned} &\Phi(\mathbf{M}) = \left[n^{-1} trace\left(\mathbf{H}\mathbf{M}^{-1}\mathbf{H}^T\right)^k\right]^{1/k} \ \ if \det(\mathbf{M}) \neq 0 \\ &\Phi(\mathbf{M}) = \infty \ if \det(\mathbf{M}) = 0 \end{aligned} \tag{6}$$

where $\mathbf{H}$ is $n \times n$, a non-singular matrix, and $k \geq 0$.

Each of the above criteria delivers the $MOSS$. Three examples are given below of the most popular optimality criteria obtained for $\mathbf{H} = \mathbf{I}_{n\times n}$ and $k = 0, 1, \infty$ (Walter and Pronzato, 1990; Zarrop, 1979):

- D-optimality, $\mathbf{H} = \mathbf{I}_{n\times n};\ k = 0;\ \Phi(\mathbf{M}) = [\det(\mathbf{M})]^{-1}$;
- A-optimality, $\mathbf{H} = \mathbf{I}_{n\times n};\ k = 1;\ \Phi(\mathbf{M}) = n^{-1} trace(\mathbf{M}^{-1})$;
- E-optimality, $\mathbf{H} = \mathbf{I}_{n\times n};\ k = \infty;\ \Phi(\mathbf{M}) = \max(\lambda_{\mathbf{M}^{-1}})$.

where $\lambda_{\mathbf{M}^{-1}}$ is the eigenvalue of $\mathbf{M}^{-1}$, and $\Phi(\mathbf{M})$ is minimized with respect to the sampling schedule.

1.1. Local optimal approximate design

The most frequently used approach, known as local optimal design (Fedorov, 1972), is based on designing an optimal experiment for a given nominal value $\mathbf{p}_0$ of the model parameter. The approach leads to an optimal experiment that depends on $\mathbf{p}_0$. For an approximate design approach the number of measurements to be performed need not be specified. The most commonly used criteria for optimal experiment design are D, A, E-optimality (Feng and Siu, 1997; Cobelli and Ruggeri, 1985; DiStefano, 1982; Kalicka, 1999; Kalicka and Bochen, 1999; Landaw, 1982).

The D, E and A-optimality criteria were implemented and their reliability, efficiency and applicability were analyzed and compared (Kalicka and Bochen, 1999, 2005). The resulting conclusions concerning D, E and A optimal minimal sampling schedules are:

- Different optimization criteria used for the same data give different optimal SS.
- D-optimal design shows the best numerical robustness and its physical interpretation is attractive: the design minimizes the volume of the asymptotic confidence region for the maximum likelihood estimate $\hat{\mathbf{p}}$ of $\mathbf{p}$.

- There are some SS optimizations that deliver, instead of OSS, a non-optimal set of time points which are not the co-ordinates of a cost function extreme value. Our earlier work (Kalicka and Bochen, 2005) shows that the noted lack of convergence to the optimal solution (the faulty solution) for compartmental models and for D-optimal design results from the numerical inconveniences of the searching algorithms. A fraction, a priori unknown, of faulty results obliges one to repeat the optimization procedure with different initial conditions to become convinced of the actual optimality of the solution. Faulty results appear less frequently for D-optimal design than for A and E-optimal designs.

For these reasons we have decided to use D-OSS design in further investigation as a gold standard for comparing features of the newly designed criteria.

2. New OSS design

The above conclusions induced us to look for an alternative to the commonly used methods. The alternative should not show the tendency to yield faulty results. This requires a new direct method of finding the location for the optimal sampling points. Our proposal is the method based on sensitivity analysis (Kalicka and Bochen, 1999; Thomaseth and Cobelli, 1999).

Let us consider compartmental model of regression functions $y_m(\mathbf{p}, t)$ and output measurements $y(t_i)$ (1). A "true" parameter vector $\mathbf{p}_0$ (nominal value) was estimated in the time domain by minimizing the weighted residual sum of squares ($WRSS$) with the objective function OF

$$OF\left(\mathbf{y},\mathbf{p}\right) = WRSS\left(\mathbf{y},\mathbf{p}\right) = \sum_{i=1}^{N} \frac{1}{\sigma^2\left(t_i\right)} \cdot \left[y\left(t_i\right) - y_m\left(t_i,\mathbf{p}\right)\right]^2 \tag{7}$$

where $\mathbf{y} = [y(\mathbf{p}, t_1), \ldots, y(\mathbf{p}, t_N)]^T$ represents the column vector of measurements. The measurement noise variance is assumed to be known as a scale factor estimated from the final $WRSS$ and it is equal to $\sqrt{WRSS}$ (Kalicka, 1999). Consequently, $\frac{1}{\sigma^2(t_i)} = \frac{1}{\sigma^2} = R$ is a scalar factor. Parameter estimate $\hat{\mathbf{p}} = \arg\min OF(\mathbf{y}, \mathbf{p})$ satisfies the condition

$$\Delta_p OF\left(\mathbf{y},\mathbf{p}\right)\big|_{\mathbf{p}=\hat{\mathbf{p}}} = \Delta_p OF\left(\mathbf{y},\hat{\mathbf{p}}\right) = 0. \tag{8}$$

The estimate $\hat{\mathbf{p}}$ is unbiased, with an expected value $E(\hat{\mathbf{p}}) = \mathbf{p}_0$. Sensitivity $\mathbf{S}(t, \mathbf{p})$ of model output with respect to the parameters is

$$\mathbf{S}\left(t,\mathbf{p}\right) = \begin{bmatrix} \Delta_p y_m\left(t_1,\mathbf{p}\right)^T \\ \vdots \\ \Delta_p y_m\left(t_N,\mathbf{p}\right)^T \end{bmatrix}. \tag{9}$$

$\mathbf{S}(t, \mathbf{p})$ represents the change δy_m in model output trajectory with respect to the small variation $\delta \mathbf{p}_0$ in the model parameter:

$$\begin{aligned} &\delta y_m = \mathbf{S}\left(t, \mathbf{p}\right) \cdot \delta \mathbf{p}_0 \\ &\delta y_m = y_m\left(t, \mathbf{p}\right) - y_m\left(t, \mathbf{p_0}\right);\ \delta \mathbf{p_0} = \mathbf{p} - \mathbf{p_0}. \end{aligned} \tag{10}$$

The above results from the Taylor series expansion of $y_m(t, \mathbf{p})$ in the neighborhood of $\mathbf{p}_0$. $\mathbf{S}(t, \mathbf{p})$ does not depend on measurement error and therefore perturbation $\delta \mathbf{y}$ in model output is as follows:

$$\delta \mathbf{y} = \begin{bmatrix} \Delta_p y_m\left(t_1, \mathbf{p}\right)^T \\ \vdots \\ \Delta_p y_m\left(t_N, \mathbf{p}\right)^T \end{bmatrix} \cdot \delta \mathbf{p_0}. \tag{11}$$

Objective function (7) in vector notation with scalar factor R and the minimization condition are:

$$\begin{aligned} &OF\left(\mathbf{y}, \mathbf{p}\right) = \left[\mathbf{y} - \mathbf{y}_m\left(\mathbf{p}\right)\right]^T \cdot R \cdot \left[\mathbf{y} - \mathbf{y}_m\left(\mathbf{p}\right)\right]^T = \min \\ &\tfrac{\partial OF}{\partial \mathbf{p}} = 0 = \left[\mathbf{y} - \mathbf{y}_m\left(\mathbf{p}\right)\right]^T \cdot \tfrac{\partial \mathbf{y}_m(\mathbf{p})}{\partial \mathbf{p}}, \tfrac{\partial \mathbf{y}_m(\mathbf{p})}{\partial \mathbf{p}} = \mathbf{S}_\mathbf{p}. \end{aligned} \tag{12}$$

For $\hat{\mathbf{p}}$ that represents the estimate vector at the minimum point of the objective function the regression function can be expressed as:

$$\mathbf{y}_m\left(\mathbf{p}\right) = \mathbf{y}_m\left(\hat{\mathbf{p}}\right) + \mathbf{S}_{\hat{\mathbf{p}}} \cdot \left(\mathbf{p} - \hat{\mathbf{p}}\right). \tag{13}$$

After substituting (13) into (12)

$$\left[\mathbf{y} - \mathbf{y}_m\left(\hat{\mathbf{p}}\right)\right]^T \cdot \mathbf{S}_{\hat{\mathbf{p}}} = \left(\mathbf{p} - \hat{\mathbf{p}}\right)^T \cdot \mathbf{S}_{\hat{\mathbf{p}}}{}^T \cdot \mathbf{S}_{\hat{\mathbf{p}}} \tag{14}$$

and solving the acquired equation, one obtains the formula

$$\begin{aligned} &\left(\mathbf{p} - \hat{\mathbf{p}}\right) = \left[\mathbf{S}_{\hat{p}}^T \cdot \mathbf{S}_{\hat{p}}\right]^{-1} \cdot \mathbf{S}_{\hat{\mathbf{p}}}^T \cdot \left(\mathbf{y} - \mathbf{y}_m\left(\hat{\mathbf{p}}\right)\right) \\ &\delta \hat{\mathbf{p}} = \left[\mathbf{S}^T \mathbf{S}\right]^{-1} \cdot \mathbf{S}^T \cdot \delta \mathbf{y} = \mathbf{W} \cdot \delta \mathbf{y}. \end{aligned} \tag{15}$$

The variation δy_j in measurements y_j; $j = 1, 2, \ldots, N$, being a result of variation in parameter value, causes variation in parameter estimates $\hat{p}_r$, $r = 1, 2, \ldots, n$. Therefore, instead of $\hat{p_r}$, one obtains $\hat{p_r} + \delta \hat{p_r}$ and

$$\begin{aligned} &\begin{bmatrix} \delta \hat{p}_1 \\ \vdots \\ \delta \hat{p}_n \end{bmatrix} = \begin{bmatrix} w_{11} & \ldots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{n1} & \ldots & w_{nN} \end{bmatrix} \cdot \begin{bmatrix} \delta y_1 \\ \vdots \\ \delta y_N \end{bmatrix} \\ &\delta \hat{p}_r = \sum_{j=1}^{N} w_{rj} \cdot \delta y_j;\ r = 1, \ldots, n. \end{aligned} \tag{16}$$

The entries of matrix $\mathbf{W}_{nxN}\,[w_{rj}];\, r = 1, \ldots, n; j = 1, \ldots, N$ show the ability of individual measurements δy_j to produce an effect on the resulting changes $\delta\hat{p}_r$ in the parameter estimates. The larger $|w_{rj}|$, the more significant the impact of measured δy_j on calculated $\delta\hat{p}_r$. For small $|w_{rj}|$ even quite marked changes in δy_j cause no noticeable change $\delta\hat{p}_r$ in parameter estimates. Therefore, the time points which assure the highest $|w_{rj}|$ values are the points at which a change in model parameter causes the most observable change in the measured output. Hence these time points are candidates for optimal sampling points (OSS). The approach presented is the one of approximate design. The final SS being designed does not fulfil an optimality criterion and depends on starting SS. The matrix $\mathbf{W}$ depends on the sensitivity matrix, which in turn depends on a model regression function calculated for a given value of parameter estimates. The idea consists in improving the initial estimate by means of a new experiment design, which uses the estimate obtained from a previous experiment. The previous estimate has to be chosen very carefully and has to be close to a real parameter value, because the design is optimal for the particular chosen regression function together with particular chosen parameter estimates.

2.1. S-optimal design

Large values of $|w_{rj}|$ are desirable for the effective identification of possible variations in model parameters. This inspires the introduction of S-optimal SS design as the set of time points at which $|w_{rj}|$ reaches the largest value. As values of $|w_{rj}|$ within a chosen measurement protocol differ greatly, it is more convenient to analyze $|w_{rj}/\max(w_{rj})|$ instead of $|w_{rj}|$. The objective function for S-OSS design is

$$OF_r = \max_j \left(|w_{rj}/\max\left(w_{rj}\right)|\right);\; r = 1, \ldots, n;\; j = 1, \ldots, N. \tag{17}$$

These maxima are detected for every r for a given experimental protocol. The time points t_j, for which OF_r reaches the maxima, are candidates for S-OSS. Every individual OF_r provides information on one parameter $p_r;\; r = 1, \ldots n$ and delivers a number of candidates for S-OSS. A reduced S-OSS (where the number of time points is equal to the number of parameters) consists of n optimal time co-ordinates. We assumed additionally that every parameter $p_r;\; r = 1, \ldots, n$ has to have at least one representative in the S-OSS.

2.2. RS-optimal design

Let us consider the relative deviations $\varepsilon_{\hat{p}_r}$ and ε_{y_j}

$$\varepsilon_{\hat{p}_r} = \delta\hat{p}_r/\hat{p}_r;\; \varepsilon_{y_j} = \delta y_j/y_j. \tag{18}$$

These are related to w_{rj}, as follows:

$$\begin{aligned} &\varepsilon_{\hat{p}_r} = \sum_{j=1}^{N} w_{rj} \cdot \delta y_j \cdot \tfrac{y_j}{y_j} \cdot \tfrac{1}{\hat{p}_r} = \sum_{j=1}^{N} v_{rj} \cdot \varepsilon_{y_j} \\ &v_{rj} = w_{rj} \cdot \tfrac{y_j}{\hat{p}_r} \ . \end{aligned} \tag{19}$$

The above indicates the possibility of using relative sensitivity for establishing OSS. This we termed relative S-OSS, or RS-OSS. Objective function for RS-OSS design is

$$\begin{aligned} &OF_r = \max_j \left(|v_{rj} / \max \left(v_{rj} \right)| \right) \\ &r = 1, \ldots n; \ j = 1, \ldots N. \end{aligned} \tag{20}$$

Designing the RS-OSS requires a similar procedure to that described for S-OSS. The maxima of OF_r that are detected and time points t_j, for which OF_r reaches the maxima, are taken as candidates for RS-OSS. Each individual OF_r gives information on one parameter p_r; $r = 1, \ldots n$ and delivers a number of candidates for RS-OSS. Reduced S-OSS (where the number of time points is equal to the number of parameters) consists of n optimal time co-ordinates. Each parameter p_r; $r = 1, \ldots n$ has to have at least one representative in the RS-OSS.

The maxima of w_{rj} and the maxima of v_{rj} have different time co-ordinates. This is because the term y_j in equation (19) changes subsequent values of w_r into v_{rj} in a non-linear way.

The MATLAB implementation of $fmins$ procedure was used to solve the D, S and RS-optimal SS problem. The purpose of $fmins$ is to minimize a function of several variables. The syntax $\mathbf{x} = fmins('\mathbf{fun}', \mathbf{x}_0)$ returns a vector $\mathbf{x}$ (a set of optimal sampling points), which is a local minimizer of $\mathbf{fun}(\mathbf{x})$, with $\mathbf{x}_0$ as a starting point. For different optimization criteria $\mathbf{fun}(\mathbf{x})$ gets the proper form of $\Phi(\mathbf{M})$ or of OF_r.

The results obtained for S and RS criteria were compared with the results obtained for D-optimal SS design.

3. Cases under study

To examine S-OSS and RS-OSS and to compare these to D-optimization two stationary model functions (a) and (b) have been chosen. For model (a) the parameter values and the error imitate the result of a pharmacokinetic experiment (McIntosh and McIntosh, 1980) concerning the distribution of gonadotropine administered in an intravenous injection. Intuitive SS consisted of $N = 23$ samples collected over a time interval of $0 \div 117$ hours. Regression function giving the best fit to data was $y_m = p_1 \exp(-p_2 t) + p_3 \exp(-p_4 t)$ with initial "true" parameter vector $\mathbf{p}_0 = [19 \ 0.02 \ 6 \ 0.5]$ and weighted residual sum of squares $WRSS = 0.341$. The model (b) (Thomaseth and Cobelli, 1999) is described

by an input-output relationship $y_m = \exp(-p_1 t) + \exp(-p_2 t) + \exp(-p_3 t)$ with initial parameter vector $\mathbf{p}_0 = [5\ 4\ 1]$.

Simulated data were used to assess the properties of S and RS-optimal designs and to compare these methods with D-optimal design. Regression functions were used as the basis for the simulation, yielding 1000 sets of data. "Measurements" were simulated as sets of numerical values $y_m(\mathbf{p}_0, t_i); i = 1, \ldots, N$ with measurement error $e(t_i)$ - Gaussian, $N(0, \sqrt{WRSS})$. For model (a) each set of simulated data contained 2401 samples related to subsequent time points in the time interval $t \in (0, 240)$ with step $dt = 0.1$. For case (b) each set of simulated data contained 201 samples over the time interval $t \in (0, 2)$ with step $dt = 0.01$.

4. Results

4.1. Simulation technique

To validate the simulation technique adopted we perform a number of tests. Let us take, as an example, model (a). Simulated data were generated by adding uncertainty, selected randomly from normally distributed population $N(0, \sqrt{WRSS})$ to the exact response of the model calculated for $\mathbf{p_0}$. Next, the parameters were re-estimated for the data generated by minimizing the weighted residual sum of squares ($WRSS$). After 1000 simulation runs we obtained 1000 estimates for each p_i. Mean parameter estimates $\bar{p}_i$, their standard deviations $\sigma_{p_i} = std\ dev_{p_i}$ and the errors Δ are defined as follows:

$$\begin{aligned} \bar{p}_i &= \tfrac{1}{1000} \sum_{r=1}^{1000} p_i^r \\ \sigma_{p_i} &= std\ dev_{p_i} = \sqrt{\tfrac{1}{999} \sum_{i=1}^{1000} (p_i - \bar{p}_i)^2} \\ \Delta &= \tfrac{|p_{i0} - \bar{p}_i|}{p_{i0}} [\%] . \end{aligned} \tag{21}$$

The Kolmogorov-Smirnov test (Filliben, 2006; Krysicki, Bartos and Dyczka, 1986) was used to answer the question of whether the sample (1000 parameter estimates) came from a normal distribution population. The test statistic D has the form:

$$D = \max_{1 \le r \le 1000} \left(F(p_i^r) - \frac{r}{1000}, \frac{r}{1000} - F(p_i^r) \right) \tag{22}$$

where F is the theoretical cumulative normal distribution, $i = 1, .., n$ is the parameter number and $r = 1, .., 1000$ is the simulation number. The Kolmogorov-Smirnov test accepts the normality hypothesis concerning parameter estimates at significance level $\alpha = 0.1$. The test statistic D obtained was 0.0170, 0.0202, 0.0197 and 0.0271 respectively for $p_1 \div p_4$, which is not greater than the critical value (the cutoff value) obtained from a table for $N = 1000$ and equal to 0.0386.

The hypothesis is accepted that the data were taken from a normal distribution with a mean of $\bar{p}_i$ and a standard deviation of σ_{p_i}. We therefore concluded that we have 1000 normally distributed estimates for each p_i .

The results are presented in Table 1. As the table shows, the mean parameter estimates $\bar{p}_i$ are very close to their nominal values. This validates the adopted simulation technique and ensures that the simulation is not the reason for any discrepancy which may appear in the optimization results for S, RS and for the D-optimal design.

Table 1. Mean parameter estimates $\bar{p}_i$, standard deviations σ_{p_i} and $\Delta[\%]$ calculated on the basis of 1000 simulation runs for model (a) with nominal parameter vector $\mathbf{p}_0 = [19\ 0.02\ 6\ 0.5]$.

-	$\bar{p}_i$	σ_{p_i}	$\Delta[\%]$
p_1	18.99197	0.07497	0.042
p_2	0.019987	0.00009	0.065
p_3	6.008938	0.24663	0.148
p_4	0.499735	0.03522	0.053

The Kolmogorov-Smirnov test was also adopted in the later parts of this paper when the normality hypothesis was postulated for scatter in S, RS and D-optimal points t_i and scatter in parameters p_i re-estimated for S, RS and D-optimal SS.

4.2. Model (a) - S and RS-OSS optimal design

The time course of $|w_r/\max(w_r)|$ and of $|v_r/\max(v_r)|$ for the nominal parameter vector $\mathbf{p}_0 = [19\ 0.02\ 6\ 0.5]$ are presented in Figs. 1 and 2, respectively. The four curves are drawn for p_r; $r = 1, 2, 3, 4$ over time interval $0 \div 240$ minutes, with a time step of 0.1. The time co-ordinates of maxima of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ were identified (Table 2) and the levels of these maxima (in brackets) were calculated. In Table 2 the time co-ordinates of the maxima are ordered according to the levels of the maxima, from the largest to the smallest. In this way the maxima are ordered from the most significant to the least significant for parameter estimation purpose. The time co-ordinates of subsequent $max|w_r/\max(w_r)|$ and $max|v_r/\max(v_r)|$ are shifted. The shift is larger for time co-ordinates more distant from $t = 0$.

As can be seen from the table (columns 1 and 3), $|w_{r\,\max}|$ and $|v_{r\,\max}|$ differ greatly within the analyzed measurement protocol: $|w_{3\,\max}| : |w_{2\,\max}| \approx 20$ and $|v_{4\,\max}| : |v_{2\,\max}| \approx 80$. Therefore, it is rational to analyze $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ instead of $|w_r|$ and $|v_r|$. Sampling schedules S-OSS and RS-OSS are selected on the basis of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$, respectively. The steps in choosing optimal time points (similar for both approaches) are:

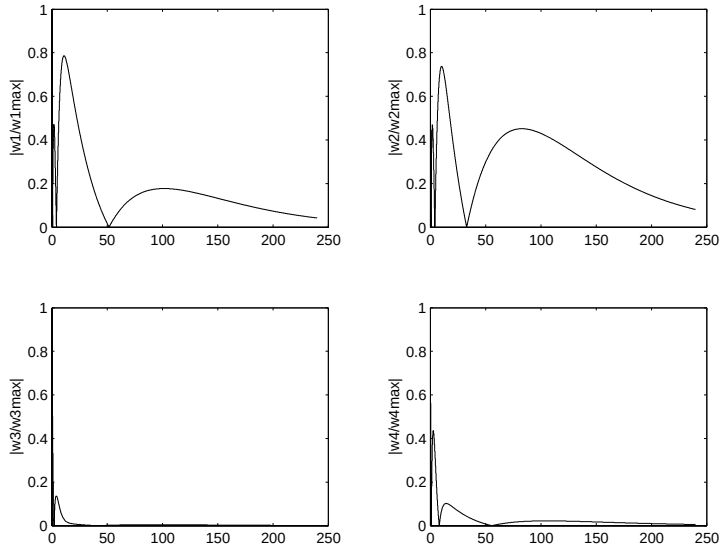


Figure 1. Functions $|w_r/\max(w_r)|$ for parameter p_r, $r = 1, \ldots, 4$ at points $j = 1, \ldots, 2401$ for time $t_j = (j-1)/10 = 0 \div 240$.

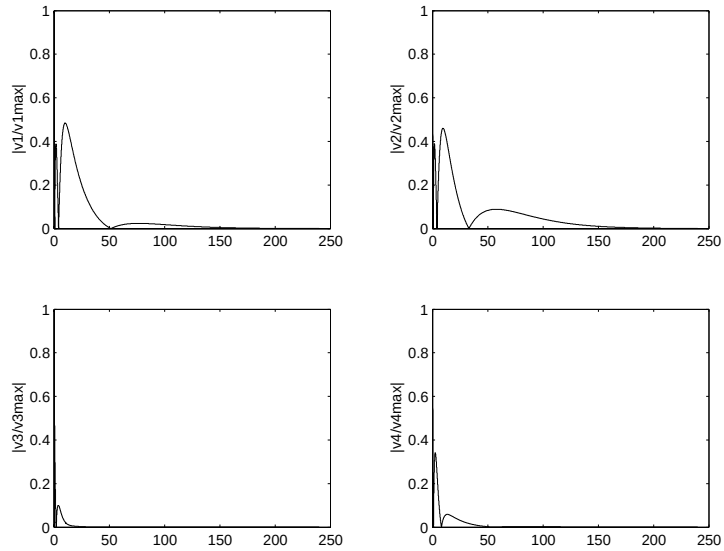


Figure 2. Functions $|v_r/\max(v_r)|$ for parameter p_r, $r = 1, \ldots, 4$ at points $j = 1, \ldots, 2401$ for time $t_j = (j-1)/10 = 0 \div 240$.

Table 2. Co-ordinates of maxima of $|w_r/\max(w_r)|$ and of $|v_r/\max(v_r)|$ for model (a) over time interval $t \in (0, 240)$. The time points in the table are ordered according to the level of the maxima (in brackets), from the largest to the smallest value.

Maxima of $\lvert w_r/\max(w_r)\rvert$	Time co-ordinates of maxima	Maxima of $\lvert v_r/\max(v_r)\rvert$	Time co-ordinates of maxima
With respect to p_1 $\lvert w_1/w_{1\max}\rvert$ $\lvert w_{1\max}\rvert =$ $= 9.8 \cdot 10^{-3}$	$t_1 = 0.00 \to (1.00)$ $t_2 = 11.1 \to (0.86)$ $t_3 = 1.90 \to (0.52)$ $t_4 = 102 \to (0.19)$	With respect to p_1 $\lvert v_1/v_{1\max}\rvert$ $\lvert v_{1\max}\rvert =$ $= 1.3 \cdot 10^{-2}$	$t_1 = 0.00 \to (1.00)$ $t_2 = 10.1 \to (0.53)$ $t_3 = 1.80 \to (0.43)$ $t_4 = 76.8 \to (0.03)$
With respect to p_2 $\lvert w_2/w_{2\max}\rvert$ $\lvert w_{2\max}\rvert =$ $= 8.9 \cdot 10^{-3}$	$t_1 = 0.00 \to (1.00)$ $t_2 = 10.3 \to (0.81)$ $t_3 = 1.90 \to (0.52)$ $t_4 = 83.0 \to (0.49)$	With respect to p_2 $\lvert v_2/v_{2\max}\rvert$ $\lvert v_{2\max}\rvert =$ $= 1.1 \cdot 10^{-2}$	$t_1 = 0.00 \to (1.00)$ $t_2 = 9.90 \to (0.51)$ $t_3 = 1.80 \to (0.43)$ $t_4 = 58.0 \to (0.10)$
With respect to p_3 $\lvert w_3/w_{3\max}\rvert$ $\lvert w_{3\max}\rvert = 0.18$	$t_1 = 0.00 \to (1.00)$ $t_2 = 4.20 \to (0.14)$	With respect to p_3 $\lvert v_3/v_{3\max}\rvert$ $\lvert v_{3\max}\rvert = 0.74$	$t_1 = 0.00 \to (1.00)$ $t_2 = 4.00 \to (0.10)$
With respect to p_4 $\lvert w_4/w_{4\max}\rvert$ $\lvert w_{4\max}\rvert = 0.02$	$t_1 = 0.00 \to (1.00)$ $t_2 = 2.70 \to (0.47)$ $t_3 = 14.4 \to (0.11)$ $t_4 = 106 \to (10^{-3})$	With respect to p_4 $\lvert v_5/v_{5\max}\rvert$ $\lvert v_{4\max}\rvert = 0.88$	$t_1 = 0.00 \to (1.00)$ $t_2 = 2.50 \to (0.37)$ $t_3 = 13.8 \to (0.06)$ $t_4 = 80.9 \to (0.003)$

1. Calculate $|w_r/\max(w_r)|$ ($|v_r/\max(v_r)|$) for the experimental protocol. Find the co-ordinates of $\max|w_r/\max(w_r)|$ ($\max|v_r/\max(v_r)|$) and order them from the most significant to the least significant, as shown in Table 2.
2. Take the best t_1 point and place it in the SS. There are four equally good time points. These are the co-ordinates of the highest maxima for all parameters. The four t_1 time points are identical and for $r = 1, 2, 3, 4$ give one candidate for OSS: $t = 0$. Therefore, at present, S-OSS $= [0, -, -, -]$, (RS-OSS $= [0, -, -, -]$).
3. Take the best t_2 of the largest value of $|w_u/\max(w_u)|$, ($|v_u/\max(v_u)|$). Add it to the previous OSS. Therefore, at present, S-OSS $= [0, t_{2u}, -, -]$, (RS-OSS $= [0, t_{2u}, -, -]$). In the further search do not take into account the already used $|w_u/\max(w_u)|$, ($|v_u/\max(v_u)|$).
4. Take the best t_3 of the largest value of $|w_s/\max(w_s)|$, ($|v_s/\max(v_s)|$). Add it to the previous OSS. The current S-OSS $= [0, t_{2u}, t_{3s}, -]$, (RS-OSS $= [0, t_{2u}, t_{3s}, -]$). In the further search do not take into account the already

used $|w_u/\max(w_u)|$, ($|v_u/\max(v_u)|$) and $|w_s/\max(w_s)|$, ($|v_s/\max(v_s)|$.
5. Take the best t_4 of the largest $|w_t/\max(w_t)|$, ($|v_t/\max(v_t)|$). The final S-OSS $= [0, t_{2u}, t_{3s}, t_{4t}]$ (RS-OSS $= [0, t_{2u}, t_{3s}, t_{4t}]$).

Sensitivities $\max|w_r/\max(w_r)|$ and $\max|v_r/\max(v_r)|$ were calculated and the co-ordinates of the maxima were selected in the programming environment MATLAB, using previously developed software (Kalicka and Bochen, 1999). The results of the individual steps of the algorithm and the final choice for model (a) are presented in Table 3.

S, RS and D-OSS start from the same zero time point. D-OSS ends earlier than RS and S-OSS. In the above example p_1 is represented once in S-OSS and in RS-OSS (the first step), while p_2, p_3 and p_4 are represented twice (steps 2, 3 and 4 respectively).

Table 3. Selection of S-OSS and RS-OSS on the basis of $w_r = |w_r/\max(w_r)|$, and $|v_r/\max(v_r)|$. For comparison, D-OSS is given in the last row of the table.

step	$\lvert w_r/\max(w_r)\rvert$	$\lvert v_r/\max(v_r)\rvert$
1	For $p_r, r = 1,2,3,4$ $[0, \_, \_, \_]$	For $p_r, r = 1,2,3,4$ $[0, \_, \_, \_]$
2	For $p_u, u = 2$ $[0, 11.1, \_, \_]$	$p_u, u = 2$ $[0, 10.1, \_, \_]$
3	For $p_s, s = 3$ $[0, 11.1, 1.9, \_]$	For $p_s, s = 3$ $[0, 10.1, 1.8, \_]$
4	For $p_t, t = 4$ $[0, 11.1, 1.9, 105.8]$	For $p_t, t = 4$ $[0, 10.1, 1.8, 80.9]$
Final choice	S-OSS $[0, 1.9, 11.1, 105.8]$	RS-OSS $[0, 1.8, 10.1, 80.9]$
Gold standard	D-OSS $[0, 1.9, 9.6, 60.3]$	D-OSS $[0, 1.9, 9.6, 60.3]$

4.3. Model (b) - S and RS-OSS optimal design

Time time courses of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ for nominal parameter vector $\mathbf{p}_0 = [5\ 4\ 1]$ are presented in Fig. 3. The time co-ordinates of the maxima of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ and the values of subsequent maxima (in brackets) are presented in Table 4. Similarly, as for case (a), time co-ordinates of subsequent maxima of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ are shifted.

As can be seen from the above, there are two equally good values of t_1 for $|w_r/\max(w_r)|$: $t_1 = 0.10$ (for p_1 and for p_2) and $t_1 = 1.65$ (for p_3). Therefore, the first step for $|w_r/\max(w_r)|$ gives two optimal sampling times. For $|v_r/\max(v_r)|$ all t_1 points are the same: $t_1 = 0.08$.

Table 4. Co-ordinates of the maxima of $|w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$ for model (b) over time interval $t \in (0, 2)$. The time points in the table are ordered according to the values of maxima (in brackets).

Maxima of $\lvert w_r/\max(w_r)\rvert$	Time co-ordinates of maxima	Maximum of $\lvert v_r/\max(v_r)\rvert$	Time co-ordinates of maxima
With respect to p_1 $\lvert w_1/w_{1\max}\rvert$ $w_{1\max} = 2.5$	$t_1 = 0.10 \to (1.00)$ $t_2 = 0.59 \to (0.62)$ $t_3 = 1.87 \to (0.36)$	With respect to p_1 $\lvert v_1/v_{1\max}\rvert$ $v_{1\max} = 1.1$	$t_1 = 0.08 \to (1.00)$ $t_2 = 0.52 \to (0.21)$ $t_3 = 1.54 \to (0.03)$
With respect to p_2 $\lvert w_2/w_{2\max}\rvert$ $w_{2\max} = 1.7$	$t_1 = 0.10 \to (1.00)$ $t_2 = 0.57 \to (0.77)$ $t_3 = 1.84 \to (0.46)$	With respect to p_2 $\lvert v_2/v_{2\max}\rvert$ $v_{2\max} = 0.95$	$t_1 = 0.08 \to (1.00)$ $t_2 = 0.49 \to (0.27)$ $t_3 = 1.52 \to (0.04)$
With respect to p_3 $\lvert w_3/w_{3\max}\rvert$ $w_{3\max} =$ $= 3.8 \cdot 10^{-3}$	$t_1 = 1.65 \to (1.00)$ $t_2 = 0.09 \to (0.65)$ $t_3 = 0.48 \to (0.53)$	With respect to p_3 $\lvert v_3/v_{3\max}\rvert$ $v_{3\max} =$ $= 5.7 \cdot 10^{-3}$	$t_1 = 0.08 \to (1.00)$ $t_2 = 0.43 \to (0.32)$ $t_3 = 1.28 \to (0.16)$

Table 5. Selection of S-OSS and RS-OSS, for model (b), on the basis of $w_r = |w_r/\max(w_r)|$ and $|v_r/\max(v_r)|$. For the purpose of comparison D-OSS is given in the last row of the table.

Step	$\lvert w_r/\max(w_r)\rvert$	$\lvert v_r/\max(v_r)\rvert$
1	For $p_r, r = 1, 2$ [0.10, _, _] For $p_r, r = 3$ [_, 1.65, _]	For $p_r, r = 1, 2, 3$ [0.08, _, _]
2	For $p_u, u = 2$ [_, _, 0.57]	For $p_u, u = 3$ [_, 0.43, _]
3	SS already consists of 3 time points	For $t_s, s = 2$ [_, _, 1.52]
Final Choice	S-OSS $[0.10, 0.57, 1.65]$	RS-OSS $[0.08, 0.43, 1.52]$
Gold Standard	D-OSS $[0.14, 0.51, 1.47]$	D-OSS $[0.14, 0.51, 1.47]$

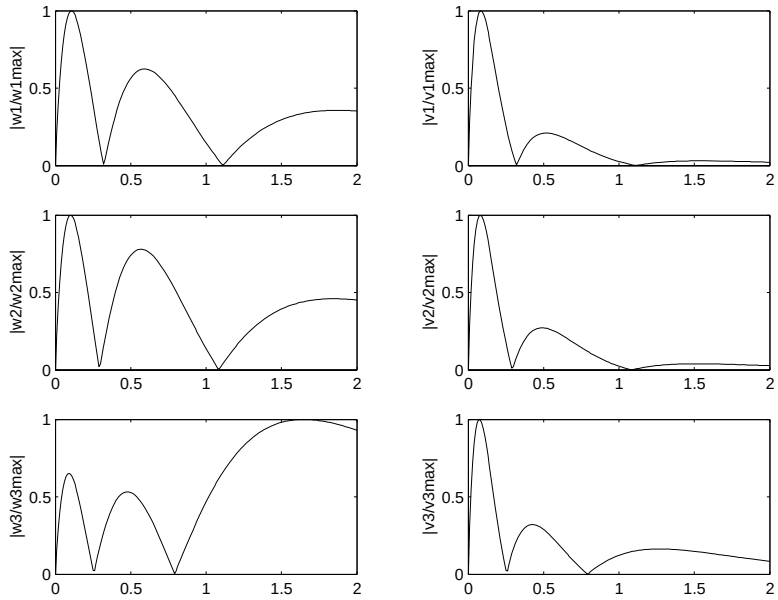


Figure 3. Functions $|v_r/\max(v_r)|$ and $|v_r/\max(v_r)|$ for parameter p_r, $r = 1, 2, 3$ at points $j = 1, \ldots 201$ for time $t_j = (j-1)/100 = 0 \div 2$.

The optimal SS for the three criteria considered do not differ noticeably. D-OSS starts later then the others and ends earlier then RS and S-OSS end. In the above example p_1 and p_3 are represented twice and p_2 is represented once in the S-OSS. In the RS-OSS, p_1 has one representative and p_2 and p_3 have two representatives.

4.4. Comparison of D, S and RS-OSS criteria

More detailed results will be presented on the exemplary model (a), a two-compartmental model with four macro-parameters. In order to compare properties of D, S and RS-OSS we used the simulation technique described in section 4.1. The whole process of simulation and parameter estimation, delivered 1000 simulated model functions and 1000 sets of OSS calculated using S, RS and D-optimal design.

Scatter in t_i was analyzed for the criteria tested. We concluded that t_i are normally distributed (the Kolmogorov-Smirnov test hypothesis of normality was accepted at the significance level $\alpha = 0.01$). Table 6 presents the mean optimal sampling points $\bar{t}_i$ and $\sigma_{t_i} = std\ dev_{t_i}$ and subsequently for S, RS and D-optimal design for model (a).

Table 6. Scatter in $OSS = [t_1\ t_2\ t_3\ t_4]$, mean values $\bar{t}_i = \frac{1}{1000} \cdot \sum_{r=1}^{1000} t_i^r$ standard deviations $\sigma_{t_i} = std\ dev_{t_i}$ and $CV = \sigma_{t_i}/\bar{t}_i[\%]$ for S, RS and D-optimal design, model (a).

S-optimal SS	t_1	t_2	t_3	t_4
$\bar{t}_i$	0.0	1.9	11.1	105.8
$\sigma_{t_i} = std\ dev_{t_i}, CV\%$	0.0	0.13, 6.8%	0.57, 5.1%	0.69, 0.6%
RS-optimal SS	t_1	t_2	t_3	t_4
$\bar{t}_i$	0.0	1.8	10.1	80.9
$\sigma_{t_i} = std\ dev_{t_i}, CV\%$	0.0	0.14, 7.7%	0.56, 5.5%	0.72, 0.9%
D-optimal SS	t_1	t_2	t_3	t_4
$\bar{t}_i$	0.0	1.9	9.6	60.3
$\sigma_{t_i} = std\ dev_{t_i}, CV\%$	0.0	0.1, 5.3%	0.5, 5.2%	0.8 ,1.3%

It should be noted that:

- Different optimization criteria used for the same model function gave different OSS and consequently gave different mean sampling points $\bar{t}_i$.
- S and RS-OSS and D-optimal criteria give $t = 0$ as the first optimal point (while E-optimal design, for instance, does not).
- When compared, scatter in OSS for S, RS and D-OSS are found to be similar. The last optimal sample t_4 has an even smaller coefficient of variation value for S-optimal ($CV = 0.6\%$) and RS-optimal ($CV = 0.9\%$) design than it has for D-optimal design ($CV = 1.3\%$).

Table 7 presents the comparison of parameter estimates obtained on the basis of intuitive sampling schedule ISS (23 samples not uniformly, but intuitively allocated over a time interval of 0-117 hours) and of D, S, and RS-OSS designs. These particular results were obtained for an exemplary simulation (one out of 1000 simulation runs).

The general conclusion is that S, RS and D-OSS give very similar results.

There is a loss in parameter accuracy for every minimal optimal SS (4 samples) in relation to the numerous ISS (23 samples). This is the price paid for decreasing the cost and onerousness of clinical measurements. D, S and RS-OSS give almost the same parameter estimates and coefficients of variation. Therefore, the newly designed criteria S and RS give equally good results as the D criterion as far as parameter estimates and accuracy are concerned. The newly designed criteria have an important advantage over D-optimal design. They do not have a tendency to give a non-optimal solution instead of the optimal one, as happens with D-optimal design.

Other results (Kalicka and Bochen, 1999, 2005) show that there is a fraction of optimization procedures which ends in a faulty solution and not in the optimal

Table 7. Parameter estimates p_i and coefficient of variation $CV\%$ calculated for model (a) with nominal parameter vector $\mathbf{p}_0 = [19\ 0.02\ 6\ 0.5]$.

	p_1 $(CV\%)$	p_2 $(CV\%)$	p_3 $(CV\%)$	p_4 $(CV\%)$	$WRSS$
ISS 23 non optimal samples	19.19 (2.5%)	0.020 (3.6%)	6.075 (9.1%)	0.497 (21%)	0.265
D-OSS 4 optimal samples	18.94 (4.8%)	0.0197 (10%)	6.811 (16%)	0.554 (32%)	0.299
S-OSS 4 optimal samples	18.95 (5.0%)	0.0198 (13%)	6.797 (16%)	0.555 (32%)	0.299
RS-OSS 4 optimal samples	18.95 (4.7%)	0.0198 (11%)	6.800 (15%)	0.559 (31%)	0.299

one according to the D-optimality criterion. For a k-compartmental model, minimal OSS consists of $n = 2 * k$ time points. The objective function is then spread over the multi-dimensional time space. A special procedure was designed (Kalicka and Bochen, 2005) to determine the exact number and location of the objective function extremes in an allowable time domain. The noted lack of convergence to the optimal solution (the faulty solution) for some searching algorithms results from numerical inconveniences and not from the existence of local extremes of the objective functions of compartmental models. The lack of local extremes does not guarantee that the optimization algorithm provides the global extreme. The proportion of faulty solutions for D-optimal design is of several percent. The proposed S and RS-optimal designs find candidates for OSS over the time interval on the basis of the investigation of subsequent w_{rj} and v_{rj}, $r = 1, .., n$ for the maxima. There is therefore a need to repeat the calculation to be convinced of the D-optimality of the solution.

The last column of Table 7 contains the weighted residual sum of squares ($WRSS$). When $WRSS$ is calculated for OSS estimates with respect to OSS time points, it is equal to zero. This does not mean that the estimates are faultless but rather that the model function proposed provides a correct description of the functioning of the system. Therefore, $WRSS$ was calculated with respect to ISS for estimates based on OSS, in other words – with respect to $N = 23$ points of intuitive SS. A slight increase in $WRSS$ for reduced OSS causes an increase in $CV\%$ calculated for parameter estimates based on OSS.

A nominal parameter vector $\mathbf{p}_0 = [19\ 0.02\ 6\ 0.5]$ was used for simulation of 1000 model functions, which gave 1000 sets of OSS calculated using S, RS and D-optimal design. For every reduced $OSS_s = [t_1\ t_2\ t_3\ t_4]$, $s = 1, .., 1000$, the parameters p_i were re-estimated. Altogether there were 3000 simulation sets with 1000 for each criterion. The histograms presented in Fig. 4 show scatter in p_i re-estimated on the basis of 1000 simulations. Assuming a normal distribution

Table 8. For model (a): mean $\bar{p}_i$, standard deviations σ_{p_i} and Δ[%] for parameters $p_1 \div p_4$ obtained for S, RS and D-OSS design on the basis of 1000 simulation runs.

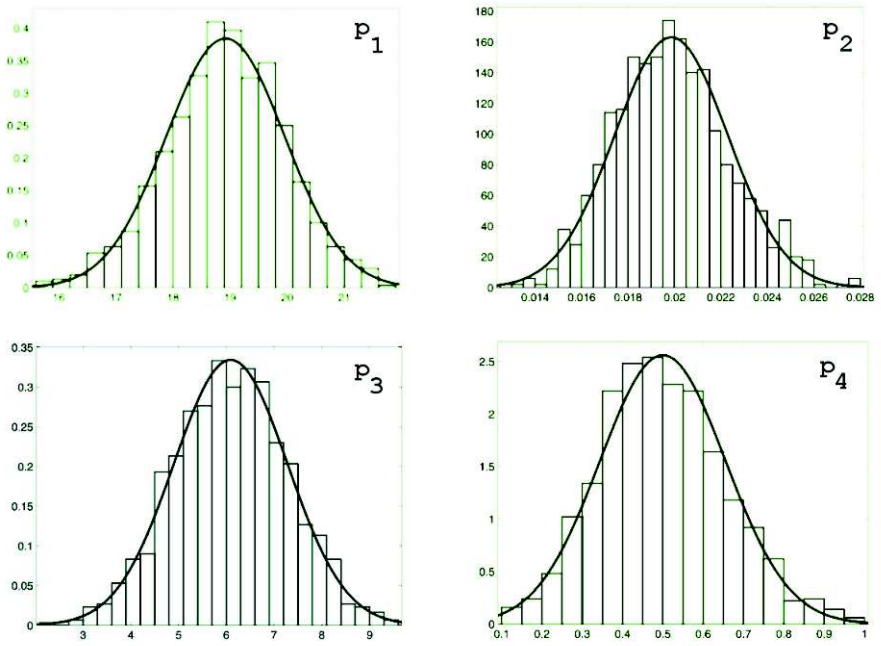
		p_1 (Δ%)	p_2 (Δ%)	p_3 (Δ%)	p_4 (Δ%)
S-OSS	$\bar{p}_i$	**18.92 (0.5%)**	**0.0201 (0.5%)**	**6.096 (1.6%)**	0.5612 (12.2%)
	σ_{p_i}	1.049	0.0028	1.209	0.2792
D-OSS	$\bar{p}_i$	18.79 (1.1%)	0.0198 (1.0%)	6.217 (3.6%)	0.5284 (5.7%)
	σ_{p_i}	1.081	0.0022	1.214	0.2268
RS-OSS	$\bar{p}_i$	18.84 (0.8%)	**0.0199 (0.5%)**	**6.175 (2.9%)**	**0.5248 (4.9%)**
	σ_{p_i}	1.025	0.0023	1.169	0.2098

of p_i (the Kolmogorov-Smirnov test hypothesis of normality was accepted at significance level $\alpha = 0.01$), we have calculated $\bar{p}_i$ and σ_{p_i}, which are shown in Table 8. Additionally, in brackets, deviations are given $\Delta = \frac{|p_{i0} - \bar{p}_i|}{p_{i0}}$ [%] for p_{i0} and $\bar{p}_i, i = 1, 2, 3, 4$. A comparison of $\bar{p}_i$ and σ_{p_i} allows us to frame conclusions concerning the OSS criteria under scrutiny. The smaller the dispersion σ_{p_i} and the closer to the initial parameter value the mean $\bar{p}_i$ (the smaller Δ), the more reliable and efficient the optimization criterion.

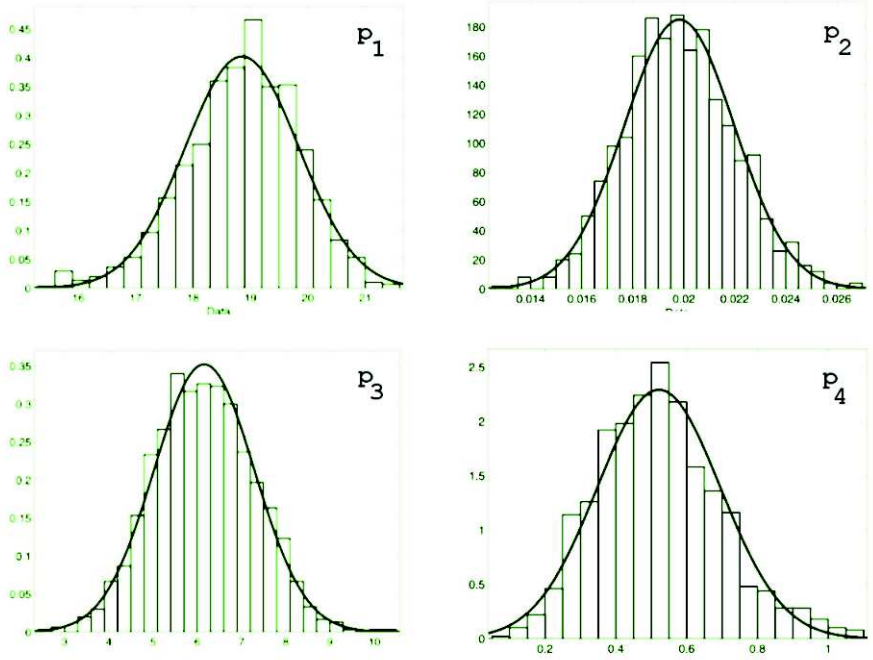
The shaded cells in Table 8 show estimates which are the closest to their nominal values. In the column related to p_1 the most accurate (the smallest σ_{p_i} and Δ[%]) are estimates obtained for S-OSS design. For parameter p_2, the S and RS designs are equally good and are better than D-OSS. For p_3 and for p_4 S-optimal and RS-optimal design respectively turned out to be the best. D-OSS was never indicated as being the best with respect to Δ[%]. This conclusion is in compliance with expectation - S and RS-OSS search for the time points which are most capable of reflecting the actual parameter value.

S and RS-OSS do not give faulty solutions. This is because distribution of maximum values of objective functions over time interval in known exactly on the basis of analysis of $w_{rj}(v_{rj})$. Criteria like that for D-optimal design, based on the Fisher information matrix, search the multidimensional space to determine the extreme point of criterion function. The percentage fraction (in 1000 simulation runs) of faulty, or non-optimal, SS was 1.9% for D-OSS. No solution obtained from the reduced optimal S and RS sampling schedule was faulty.

a) For S-OSS



b) For RS-OSS



c) For D-OSS

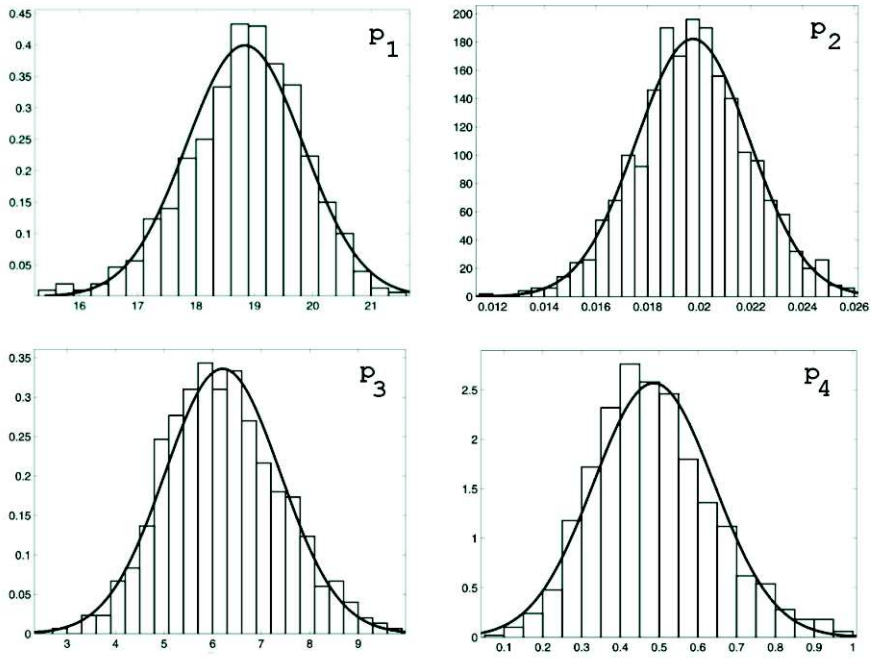


Figure 4. Histograms showing scatter in $p_r,\ r = 1, \ldots, 4$ (model (a)) calculated for 1000 simulation runs. In the background there are normal distribution functions accepted by the Kolmogorov-Smirnov statistical test at significe level $\alpha = 0.01$.

5. Conclusions

An S and RS-optimal sampling schedule design has been presented. These optimization criteria are based on maximization of the parameter estimate sensitivity to changes in measurements. This method allows the location of optimal sampling points to be found in a direct manner. This is a crucial asset in comparison with methods which search for an extreme of objective function (expressed by means of the Fisher information matrix) spread over multi-dimensional space of time. This is the reason for accidental stopping of the algorithm in a faulty solution. The optimal designs presented are free from this inconvenience. The location of maximum values of an objective function over a time interval is known on the basis of investigation of subsequent w_{rj} and v_{rj}, $r = 1, .., n$ for maxima at t_j. Co-ordinates t_j are candidates for OSS. The best n from these form the reduced S-OSS (RS-OSS). The search procedure is free from the in-

conveniences which may occur when the minimum of an objective function is sought (such as for D-OSS) over multi-dimensional parameter space and the procedure may end in a false solution.

The new criteria were compared with the D-OSS approach. Different criteria give different OSS. A simulation technique was used for analyzing the way in which the OSS criteria under consideration influence parameter estimates. The simulation technique adopted was tested to show that it is not the reason for a discrepancy in optimization results. The process of simulation (1000 simulation runs) and parameter estimation gave 1000 sets of parameter estimates for each criterion. The results of simulations show that parameter estimates were normally distributed, which was verified by means of the Kolmogorov-Smirnov goodness-of-fit tests. To validate the simulation technique adopted we performed a set of simulations. The parameter estimates obtained on the basis of numerous (not for reduced minimal) simulated measurements with normal additive noise turned out to be normally distributed at significance level $\alpha = 0.1$, which validates the simulation technique adopted.

Next, the process of simulation and parameter estimation was used to obtain 1000 simulated model functions and 1000 sets of OSS calculated using S, RS and D-optimal design and 1000 of the model parameters re-estimated for the reduced optimal designs. Scatter in optimal t_i and in re-estimated p_i was analyzed for the criteria tested. The hypothesis of normality for t_i and for p_i was accepted at significance level $\alpha = 0.01$. Mean parameter estimates and standard deviations ware calculated, analyzed and compared. The results of 1000 simulation runs for each method show that S and RS-optimization could be considered as an alternative to D-optimization. All the methods have a similar accuracy, while S and RS-optimization never give faulty solutions.

References

Brandt, S. (1999) *Data Analysis. Statistical and Computational Methods for Scientists and Engineers*. Springer Verlag, New York.

Cobelli, C., Ruggeri, A. and Thomaseth, K. (1985) Optimal Input and Sampling Design for Physiological system Identification. Some Theoretical and in vivo Results on a Compartmental model of Glucose Kinetics. IMACS, 161-172.

Cobelli, C., Ruggeri, A., DiStefano, J.J and Landaw, E.M.(1985) Optimal design of multioutput sampling schedules - Software and applications to endocrine-metabolic and pharmacokinetic models. *IEEE Trans. Biomedical Eng.* **32** (4), 249-256.

DiStefano, J.J (1982) Algorithms, software and sequential optimal sampling schedule design for pharmacokinetic and physiologic experiments. *Mathematics Computer Simulation* **24**, 531-534.

Eykhoff, P. (1980) *System Identification and State Estimation*. PWN, Warsaw.

Fedorov, V.V. (1972) *Theory of Optimal Experiments.* Academic Press, New York.

Feng, D., Li, X. and Siu, W.C. (1997) Optimal Sampling Schedule Design for System Modelling Based on Continuous Data Acquisition and Time Accumulated Counts Measurements. A proceedings volume from the IFAC Symposium. Warwick, UK, 23-26 March 1997, 1277-282.

Filliben, J.J. (2006) NIST/SEMATECH e-Handbook of Statistical Methods. www.itl.nist.gov/div898/handbook, 2006, section 1.3.5.16.

Godfrey, K. (1983) *Compartmental Models and their Application.* Academic Press, London.

Kalicka, R. (1994a) Optimal Multicompartmental Sampling Design. Attributes of the Main Algorithms and All-Purpose Software. IFAC Symposium. Galveston Texas USA, March 1994, 273-275.

Kalicka, R. (1994b) Optimal Experiment Design for Identification of Compartmental Model Parameters of Therapeutic Processes. IEEE IMTC, Massachusetts USA, April 1995, 277-281.

Kalicka, R. (1999) Optimal Design and Organisation of Biomedical Experiment. *Measurement* **26** 19-44, Elsevier Science.

Kalicka, R. and Bochen, D. (1999a) Various criteria of sampling schedule optimisation. Are they equivalent? *Conference on Application of Mathematics in Biology and Medicine*, Ustrzyki Górne, 76-81.

Kalicka, R. and Bochen, D. (1999b) Reliability and efficiency of the D, E, S and A-optimal sampling schedule design. *Task Quarterly* **3** (4), Gdańsk, 453-464.

Kalicka, R. and Bochen, D. (1999c) S-optimal sampling schedule for nonstationary nonlinear pharmacokinetic model. *XII Conference Biocybernetics and Biomedical Engeenering*, Warsaw 4-6.XII.1999, 101-105.

Kalicka, R. and Bochen, D. (2005) Properties of D-optimal Sampling Schedule for Compartmental models. *Biocybernetics and Biomedical Engeenering* **25** (1), 23-36.

Krysicki, W., Bartos, J. and Dyczka, W. (1986) *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach* (in Polish). PWN, Warszawa, 123-125.

Landaw, E.M. (1982) Optimal multicompartmental sampling designs for parameter estimation: practical aspects. *Math. Comp. Simul.* **24**, 525-530.

McIntosh, J.E.A. and McIntosh, R.P. (1980) *Mathematical Modelling and Computers in Endocrinology.* Springer-Verlag, New York.

Thomaseth, K. and Cobelli, C. (1997) Parameter information content during model identification experiments. *A proceedings volume from the IFAC Symposium*, Warwick, UK, 23-26 March 1997, 107-112.

Thomaseth, K. and Cobelli, C. (1999) General Sensitivity Functions in Physiological System Identification. *Annals of Biomedical Engineering* **27**, 607-616.

WALTER, E. and PRONZATO L. (1990) Qualitative and quantitative experiment design for phenomenological models - A survey. *Automatica* **26** (2), 195-213.

ZARROP, M.B. (1979) *Optimal Experiment Design for Dynamic System Identification.* Springer Verlag, Berlin Heidelberg New York.