Journal of Power Sources 190 (2009) 48–55

Contents lists available at ScienceDirect

Journal of Power Sources

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

Nonlinear D-optimal design of experiments for polymer–electrolyte–membrane fuel cells

M. Meiler[a,*], D. Andre[a], Á. Pérez[a], O. Schmid[a], E.P. Hofer[b]

[a] Department of MEA and Stack Technology, Daimler AG, Neue Str. 95, D-73230 Kirchheim/Teck, Germany
[b] Institute of Measurement, Control and Microtechnology, University of Ulm, Albert-Einstein-Allee 41, D-89081 Ulm, Germany

ARTICLE INFO

Article history:
Received 10 June 2008
Received in revised form 22 July 2008
Accepted 23 August 2008
Available online 2 September 2008

Keywords:
Polymer–electrolyte–membrane fuel cell
Design of experiments (DoE)
Sequential search

ABSTRACT

Using empirical models, parameters have to be estimated from experimental data. Experimental characterization of fuel cell stacks is an expensive and time-consuming task. Therefore it is very important to choose an experimental design, which maximizes the statistical quality of the resulting information. Box and Lucas (Biometrika 46 (1959)) showed that it is possible to optimize nonlinear experimental designs by the minimization of the covariance matrix of the least squares estimate. The aim of this work is to adopt this general method in order to investigate its ability for application in polymer–electrolyte–membrane fuel cell (PEMFC) characterization. Based on an empirical PEMFC model a D-optimal design criterion has been developed and validated. Numerical methods, evolutionary and heuristic are investigated with respect to fast and robust evaluation of the design criterion. For a given set of experimental data best results are achieved using a heuristic approach, a so-called sequential search. Based on that result an algorithm to obtain an optimal design of experiments (DoE) in a nondeterministic operating area is introduced. The proposed algorithm is able to take into account experimental limitations due to test facilities or examinee. The algorithm further allows to include existing and for reference needed experiments.

1. Introduction

Intelligent energy management is a cost effective key path to realize efficient automotive drive trains. To develop operating strategy in fuel cell drive trains, precise and computational efficient models of all system components, especially the fuel cell stack, are needed.

System identification is an essential step in empirical modelling. Especially, the estimation of unknown parameters is a typical problem in the development of a PEMFC model. The standard method is the analysis of experimental data from measurements. Due to system complexity of a PEMFC the experimental investigation is an expensive and time-consuming task. In most of the cases just a limited operating area is available for the measurements caused by the fuel cell capabilities, physical laws, test equipment, or interactions between these components. To give a simple example for equipment limitations just consider the gas supply to the cell. Mass flow in general is limited to a certain range and therefore the stoichiometry λ (λ = ratio of supplied reactant to chemically needed reactant) is also limited in dependency of the current density. At a low load point using maximum gas flow a higher stoichiometry can be realized than at high load points. Also the fuel cell itself will work properly with low gas pressure at low loads, but not at high loads.

Therefore it is on the one hand important to minimize the experimental work with consideration of these constraints and on the other hand to achieve the best possible parameter estimation. The aim of this article is to show and demonstrate a systematic method for investigation of a fuel cell, which performs both tasks simultaneously. It is known that the quality of the parameter estimation depends on the used measurement points and hence on the design of experiments (DoE). Therefore the issue can also be understood as to find the data set, which delivers the largest information content if experimental data are available or if not to provide an optimized experimental design for a PEMFC stack.

Box and Lucas [1] show in their work a mathematical theory for increasing the quality of least squares based parameter estimation by minimizing the covariance matrix of its solution. Several criteria are published in the literature [2] to evaluate this optimization. In this paper these criteria are compared and the most promising one is applied on a set of 405 test points, to find the set of 50 points which is most suitable for the parameter estimation of an empirical fuel cell voltage model. Moreover, it is investigated whether this

* Corresponding author. Tel.: +49 7021 89 35 20; fax: +49 711 30 52 19 35 20.
E-mail address: markus.meiler@daimler.com (M. Meiler).

0378-7753/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpowsour.2008.08.068

Table 1
Typical factor levels of selected inputs in PEMFC characterization

Input Θ	Symbol	Unit	Factor levels
Current density	i	A/cm^2	10–20
Cathode gas pressure	p_c	bar	3–6
Anode gas pressure	p_a	bar	3–6
Relative humidity cathode gas	rh_c	–	3–6
Relative humidity anode gas	rh_a	–	3–6

optimized data set increases the quality of parameter estimation and therefore the stack model accuracy.

2. Design of experiments

2.1. Design methods

If the operating area is unrestricted and the number of fuel cell inputs are low, a full factorial DoE can be performed. Interactions as well as nonlinearities can be detected if more than three factor levels of each input are applied. But with a high number of inputs or factor levels the experimental effort rises exponentially. In Table 1 the typical number of factor levels is given. A full factorial design of experiments with the maximum values of each input would thus contain 25.920 experimental points.

If a system behaves linear within an unrestricted operation area a fractional factorial DoE can be chosen. That technique reduces the number of experiments compared to full factorial designs but is not able to detect interactions between inputs or nonlinear behaviour. Only a fraction of the edges of the hypercube, defined by the dimension of the operating area, is tested.

A trade-off between effort and ability to detect interactions and nonlinear behaviour gives a so-called central composed design. All edges of the hypercube are used. Additional, centre points and star points are added. These three designs are easy to construct, but all require an unrestricted operating area and are therefore not suited for global fuel cell characterization.

A first approach to design experiments in restricted areas is the so-called Box–Benken design. In this design the number of experiments is fixed and cannot be varied. The experiments are not placed in the edges of the operating area. More details about the mentioned and additional designs can be found in DoE literature, like Refs. [3–5]. More suited to the requirements for PEMFC characterization are optimal DoEs.

2.2. Optimal DoE

Since powerful computing has become affordable, more complex and powerful design criteria have been developed. The group of optimal criteria is able to design nonlinear DoEs with a low number of experiments in arbitrary restricted operating areas without being limited to a fixed number of experiments or factor levels.

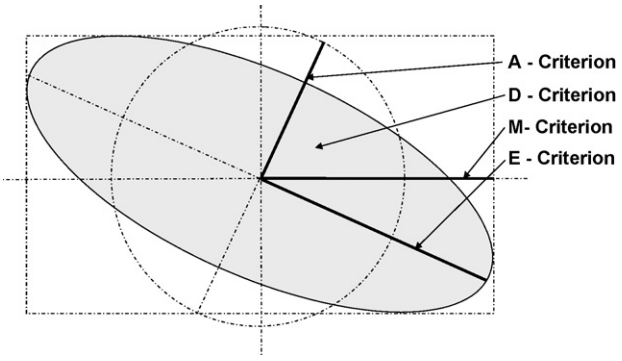

Fig. 1. Graphic interpretation of the optimal DoE criteria for minimizing the covariance matrix.

Table 2
Experimental conditions

State	Symbol	Operating range	
		Cathode	Anode
Pressure	p_c, p_a	1.05–2.5 bar	1.05–2.5 bar
Relative humidity	rh_c, rh_a	0.32–1	0.32–1
Temperature	T	75 °C	
Current density	i	0–2 A/cm^2	

Ref. [1] shows that the covariance matrix COV of the estimated model parameters ξ obtained by a least squares estimation can be approximated by the following equation:

$$\mathrm{COV} = (\mathrm{FIM}' \cdot \mathrm{FIM})^{-1}\sigma^2. \quad (1)$$

with the, in general unknown, variance σ^2 and FIM the Fisher information matrix. The FIM is a Φ-by-ϕ matrix of the partial derivative fim (Eq. (2)) of the θth model parameter ξ_θ for the dth set of experimental conditions Θ_d (Eq. (3)). Here d is the run index of the experimental conditions. It starts with 1 and ends at Φ. Additionally, θ is the index of the model parameters in the range from 1 to ϕ.

$$\mathrm{FIM} = \{\mathrm{fim}_{\theta d}\}, \quad \theta = 1 \ldots \phi, \quad d = 1 \ldots \Phi. \quad (2)$$

$$\mathrm{fim}_{\theta d} = \left[\frac{\partial\, \mathrm{fim}(\Theta, \xi)}{\partial \xi_\theta}\right]_{\theta_d}. \quad (3)$$

The basic idea of all optimal DoEs is the minimization of a quality criterion J related to COV or $(\mathrm{FIM}'\cdot\mathrm{FIM})^{-1}$ respectively. This optimization can be achieved by several design criteria [5].

The trace-criterion (A-criterion) (Eq. (4)) minimizes the trace,

$$J_A = \mathrm{trace}\left[\left(\mathrm{FIM}^{\mathrm{T}} \cdot \mathrm{FIM}\right)^{-1}\right] \quad (4)$$

the determinant-criterion (D-criterion) (Eq. (5)) minimizes the determinant,

$$J_D = \det\left[\left(\mathrm{FIM}^{\mathrm{T}} \cdot \mathrm{FIM}\right)^{-1}\right] \quad (5)$$

the eigenvalue-criterion (E-criterion) (Eq. (6)) minimizes the maximum eigenvalue (EIG),

$$J_E = \max\left\{\mathrm{EIG}\left[\left(\mathrm{FIM}^{\mathrm{T}} \cdot \mathrm{FIM}\right)^{-1}\right]\right\} \quad (6)$$

Table 3
Experimental details of campaigns 1–3 ($rh_{a/c}$ = 0.92)

Campaign					
1		2		3	
Gas pressure					
p_c (bar)	p_a (bar)	p_c (bar)	p_a (bar)	p_c (bar)	p_a (bar)
1.05	1.05	1.05	1.50	1.50	1.05
1.10	1.10	1.10	1.50	1.50	1.10
1.25	1.25	1.25	1.50	1.50	1.25
1.50	1.50	1.50	1.50	1.50	1.50
1.75	1.75	1.75	1.50	1.50	1.75
2.00	2.00	2.00	1.50	1.50	2.00
2.50	2.50	2.50	1.50	1.50	2.50

and the confidence-interval-criterion (M-criterion) (Eq. (7)) minimizes the maximum element of the main diagonal.

$$J_M = \max\left\{\sqrt{\left[\left(\mathrm{FIM}^{\mathrm{T}}\cdot\mathrm{FIM}\right)^{-1}\right]_{\theta\theta}}\right\}, \quad \theta = 1, 2\ldots\phi. \tag{7}$$

Also a geometric interpretation of these criteria is possible, whereas the A-criterion equals the mean axis length and the E-criterion the longest axis of the confidential ellipsoid spanned by the estimated parameters. The M-criterion can be explained by the longest side of an imaginary box around the ellipsoid and the D-criterion can be interpreted as the volume of the confidential ellipsoid. An illustration of these criteria is drawn in Fig. 1.

Table 4
Experimental details of campaigns 4–6 ($p_{c/a}$ = 1.5 bar)

Campaign					
4		5		6	
Gas relative humidity					
rh_c	rh_a	rh_c	rh_a	rh_c	rh_c
0.32	0.32	0.32	0.92	0.92	0.32
0.40	0.40	0.40	0.92		0.40
0.80	0.80	0.80	0.92		0.80
0.92	0.92	0.92	0.92		0.92
1.00	1.00	1.00	0.92	0.92	1.00

3. Experimental data and fuel cell model

3.1. Experimental data

For the validation of the investigated design criteria a set of experiments is used. In total 405 different data points are available, since 33 experimental runs have been carried out varying cathode pressure p_c, anode pressure p_a, cathode gas relative humidity rh_c and anode gas relative humidity rh_a and a complete polarization curve was measured for each run. All measurements were performed under isothermal conditions of 75 °C. To avoid the influence of gas supply to fuel cell voltage, the cell was fed with 30 times more reactants than the cell consumed. An overview of all experimental conditions is given in Table 2.

During the first three campaigns, gas pressures are varied while relative humidity of anode and cathode gas is kept constant at $rh_{a/c}$ = 0.92. During campaigns 4 and 6 both pressures remained at $p_{a/c}$ = 1.5 bar and the relative humidity values at cathode and anode are 0.32, 0.40, 0.80, 0.92, and full saturated conditions with a relative humidity $rh_{a/c}$ = 1.0. In Tables 3 and 4 more details about the experimental conditions are given.

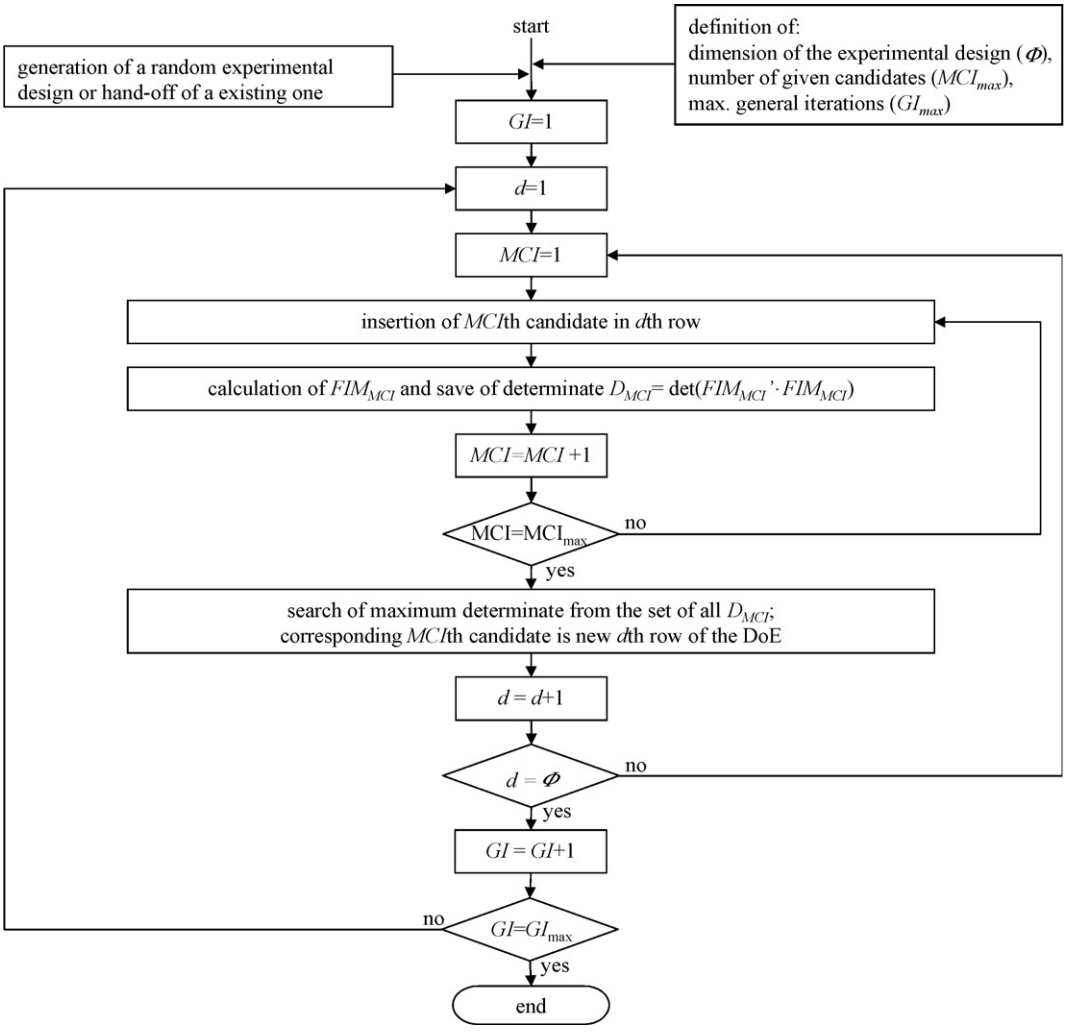

Fig. 2. Flow chart of sequential search algorithm for deterministic operating area.

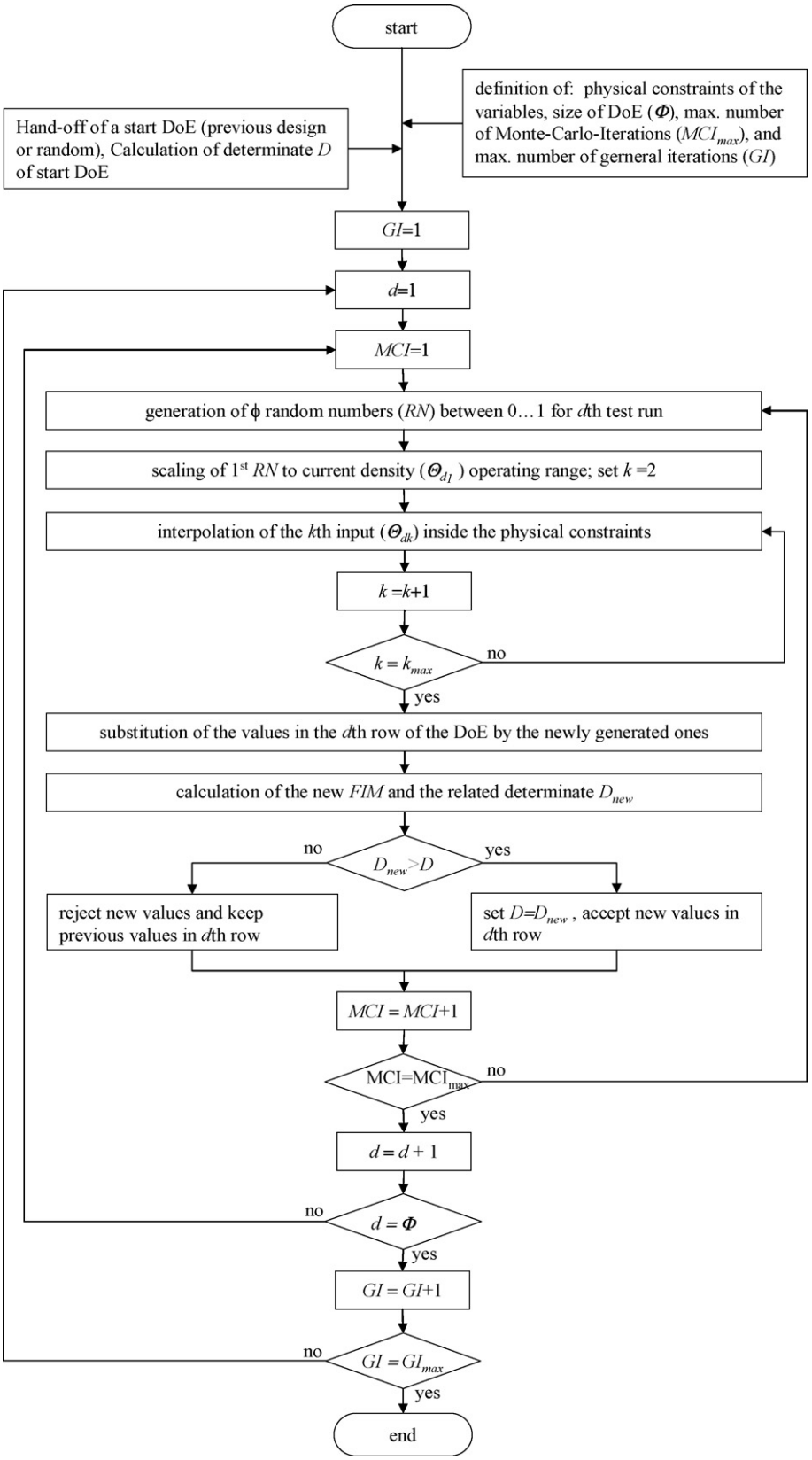

Fig. 3. Flow chart of sequential search algorithm for a restricted nondeterministic operating area.

Table 5
Estimated parameters of the SEES PEMFC model [8]

Parameter	Value	Parameter	Value
ξ_1	0.85620258	ξ_8	0.02675283
ξ_2	−0.05591375	ξ_9	0.15494972
ξ_3	0.0005173	ξ_{10}	0.71394995
ξ_4	0.04775405	ξ_{11}	0.05764422
ξ_5	0.00430726	ξ_{12}	1.39237089
ξ_6	12.5663035	ξ_{13}	4.48394466
ξ_7	−0.02481325	ξ_{14}	0.00126953

More details about experimental work and results have been previously published in Refs. [6,7].

3.2. Fuel cell model

In a previous publication of our team [8] a valid semi-empirical expert system (SEES) fuel cell model was presented. The stationary cell voltage is described by the following equation:

$$U_{\text{cell}} = U_{\text{ocv}} - U_{\text{act}} - U_{\text{ohm}} - U_{\text{dif}} - U_{\text{hum}}. \quad (8)$$

The open circuit voltage U_{ocv} in Eq. (8) is described as a function of hydrogen and air pressure. Both influences are lumped together in a simple equation (Eq. (9)) with just three free parameters:

$$U_{\text{ocv}} = \xi_1 + \xi_2 \ln(p_c) + \xi_3 \ln(p_a). \quad (9)$$

If the cathode pressure rises, the first term in Eq. (10) will decrease leading to lower activation loss U_{act}. The logarithmic dependency of the load i delivers even for a load of zero valid values because of a leakage current parameter ξ_6.

$$U_{\text{act}} = [\xi_4 - \xi_5 \ln(p_c)] \ln(i + \xi_6). \quad (10)$$

The quasi-linear ohmic resistance part of the current–voltage curve shows a dependency on cathode pressure which is described with a linear interaction term of current density and cathode pressure as given in the following equation:

$$U_{ohm} = \xi_7 \cdot i + \xi_8 \cdot i \cdot p_c \quad (11)$$

To model the diffusion loss U_{dif}, Eq. (12) is used.

$$U_{\text{dif}} = \xi_9 {p_c}^{-1} \exp(\xi_{10} i). \quad (12)$$

Eq. (13) calculates the last term of Eq. (8), the loss due to insufficient membrane humidity:

$$U_{\text{hum}} = \xi_{11} \exp(\xi_{12} i)(1.01 - \text{rh}_a)^{\xi_{13}} + \xi_{14} \text{rh}_a. \quad (13)$$

The given PEMFC model uses $k_{\max} = 4$ inputs (load i, cathode pressure p_c, anode pressure p_a, and anode gas relative humidity rh_a) and $\phi = 14$ model parameters. A valid set of these parameters is given in Table 5.

Using the PEMFC model [Eqs. (8)–(13)] in Eqs. (2) and (3) the FIM can be calculated and used for the introduced design criteria [Eqs. (4)–(7)].

4. Optimization algorithm

4.1. Algorithm for deterministic operating areas

From the mathematical point of view, a nonlinear optimization problem has to be solved. However, the design task is restricted regarding limitations of experimental equipment and examinee. Another challenge is the large number of degrees of freedom that can be adjusted to optimize the design criterion. Looking at the used fuel cell model with $k_{\max} = 4$ inputs (Θ) and a wanted DoE with $\Phi = 50$ independent test points, 200 parameters would influence the design criterion. All these parameters have to be investigated to get the optimal experimental design. Most used numerical gradient methods or gradient free optimization algorithms will have problems with disadvantageous start values [9]. However, genetic algorithms are less sensitive due to start values but convergence speed is low [10]. Therefore the optimal DoE is a promising way.

Since the A-, E- and M-design criteria are just approximations of the searched volumes the D-criterion is always the best possible

Table 6
Candidate set for the simple example

#	i (A/cm^2)	p_a (bar)	#	i (A/cm^2)	p_a (bar)	#	i (A/cm^2)	p_a (bar)	#	i (A/cm^2)	p_a (bar)
1	0	1.1	29	0	1.5	57	0	1.9	85	0	2.3
2	0.01	1.1	30	0.01	1.5	58	0.01	1.9	86	0.01	2.3
3	0.05	1.1	31	0.05	1.5	59	0.05	1.9	87	0.05	2.3
4	0.1	1.1	32	0.1	1.5	60	0.1	1.9	88	0.1	2.3
5	0.2	1.1	33	0.2	1.5	61	0.2	1.9	89	0.2	2.3
6	0.4	1.1	34	0.4	1.5	62	0.4	1.9	90	0.4	2.3
7	0.6	1.1	35	0.6	1.5	63	0.6	1.9	91	0.6	2.3
8	0.8	1.1	36	0.8	1.5	64	0.8	1.9	92	0.8	2.3
9	1	1.1	37	1	1.5	65	1	1.9	93	1	2.3
10	1.2	1.1	38	1.2	1.5	66	1.2	1.9	94	1.2	2.3
11	1.4	1.1	39	1.4	1.5	67	1.4	1.9	95	1.4	2.3
12	1.6	1.1	40	1.6	1.5	68	1.6	1.9	96	1.6	2.3
13	1.8	1.1	41	1.8	1.5	69	1.8	1.9	97	1.8	2.3
14	2	1.1	42	2	1.5	70	2	1.9	98	2	2.3
15	0	1.3	43	0	1.5	71	0	2.1	99	0	2.5
16	0.01	1.3	44	0.01	1.7	72	0.01	2.1	100	0.01	2.5
17	0.05	1.3	45	0.05	1.7	73	0.05	2.1	101	0.05	2.5
18	0.1	1.3	46	0.1	1.7	74	0.1	2.1	102	0.1	2.5
19	0.2	1.3	47	0.2	1.7	75	0.2	2.1	103	0.2	2.5
20	0.4	1.3	48	0.4	1.7	76	0.4	2.1	104	0.4	2.5
21	0.6	1.3	49	0.6	1.7	77	0.6	2.1	105	0.6	2.5
22	0.8	1.3	50	0.8	1.7	78	0.8	2.1	106	0.8	2.5
23	1	1.3	51	1	1.7	79	1	2.1	107	1	2.5
24	1.2	1.3	52	1.2	1.7	80	1.2	2.1	108	1.2	2.5
25	1.4	1.3	53	1.4	1.7	81	1.4	2.1	109	1.4	2.5
26	1.6	1.3	54	1.6	1.7	82	1.6	2.1	110	1.6	2.5
27	1.8	1.3	55	1.8	1.7	83	1.8	2.1	111	1.8	2.5
28	2	1.3	56	2	1.7	84	2	2.1	112	2	2.5

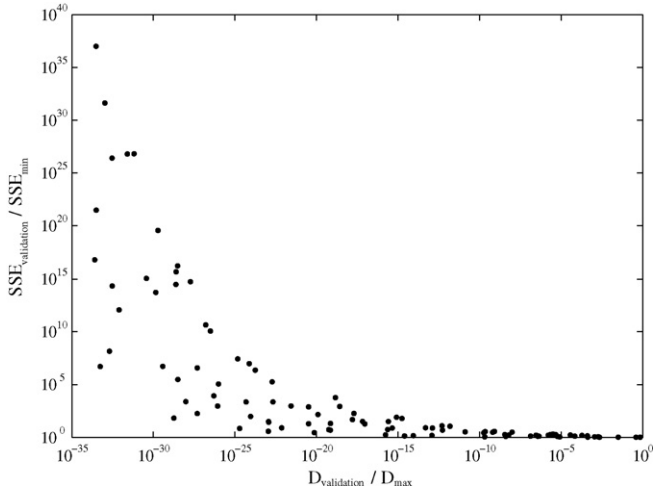

Fig. 4. SSE of the model in dependency of the determinant of the covariance matrix of the used experimental design.

choice. The other criteria are just advisable for very fast optimizations because of a reduced computing time.

If there is only a low hysteresis in fuel cell behaviour, the order of performing the experiments will not influence much the measured voltage and it is possible to change the order of the experimental conditions. That offers the opportunity to optimize the experimental design sequentially. In Fig. 2 the flow chart of a developed sequential search (SS) algorithm which identifies the best Φ experiments of MCI_{max} given candidates is shown. At first, the number of general iterations GI_{max} and a start experimental design, randomly chosen or predetermined, must be defined. Then the first row ($d = 1$) of the start experimental design is substituted by the first candidate. After this step the FIM_1 is determined and the determinant D_1, as in the following equation:

$$D_1 = J_{D_1}{}^{-1} = \det(\mathrm{FIM}_1{}' \cdot \mathrm{FIM}_1). \tag{14}$$

is calculated. Whereas $J_{D_{MCI}}$ (Eq. (5)) needs to be minimized (with MCI as the run index in the range of 1 to MCI_{max}), D_{MCI} must be maximized to find an improved experimental design. The advantage of using D_{MCI} instead of $J_{D_{MCI}}$ is the avoidance of numerical problems, caused by the inverting of a singular matrix, if a disadvantageous experimental design is tested. Both steps are repeated for all candidates. At the end, the candidate that delivered the maximum D_{MCI} is stored and taken at the new first row. Now the algorithm continues with the second row ($d = 2$) and again all candidates are tested

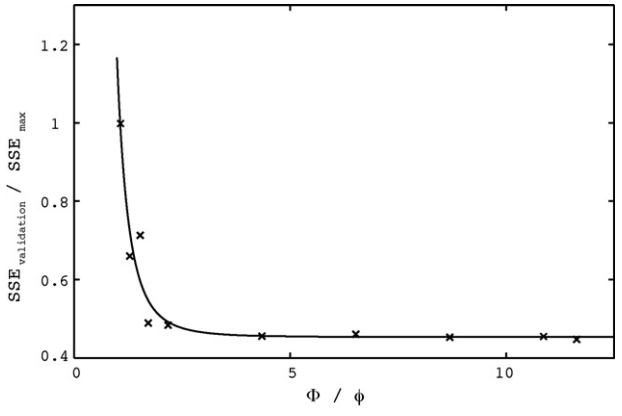

Fig. 5. SSE of the model in dependency of the ratio of experiments (Φ) to the number of parameters (ϕ).

Table 7
Results for maximization of D by using different optimization algorithms at a workstation (DELL Precision 390 Workstation, Intel core 2 duo 1.86 GHz, 8 GB RAM, WinXP64, MatlabR2007b)

	SS	GA	LM
Duration	175 s	9.5 h	1.3 h
D	1.62×10^{-6}	6.53×10^{-8}	3.98×10^{-14}

and the candidate with the largest D_{MCI} is stored. This is repeated until all d experiments are optimized once ($d = \Phi$). After that, this loop is repeated GI_{max}-times.

4.2. Algorithm for restricted nondeterministic operating areas

In many cases no experimental data will be available. The advantage is an improved quality of the parameter estimation since all operating points in the restricted area can be investigated in consideration of maximizing the covariance matrix. The disadvantage is that our sequential search has to be upgraded by the generation of MCI_{max} candidates. Fig. 3 shows the varied flow chart. The only difference to the hitherto used algorithm is that k_{max} random numbers in the range of 0–1 are generated. The current density Θ_{d1} is generated randomly by a Monte-Carlo method and the other inputs Θ_{dk} ($k = 2 \ldots k_{max}$) are interpolated in the nonlinear restricted operating area in dependence of the current density.

The generated candidates are tested and compared in view of the parameter estimation quality and the experimental design is gradual improved equal to the previous algorithm.

5. Simple example using sequential search to create D-optimal DoE

In this section a simple example is given demonstrate how to create a D-optimal DoE for a PEMFC using the sequential search approach.

To keep the example clearly represented, only the influence of anode pressure dependency is considered in this section. The simplified model is set-up by reducing Eqs. (8)–(12) to current density i an anode pressure p_a dependent terms. The resulting equation reads as

$$U_{cell} = \alpha_1 + \alpha_2 \ln(p_a) - \alpha_3 \ln(i + \alpha_4) - \alpha_5 i - \alpha_6 \exp(\alpha_7 i). \tag{15}$$

This simple model take $k_{max} = 2$ inputs into account. The first input (Θ_1) is the current density i, and the second (Θ_{12}) is anode pressure. Further the model includes in total $\phi = 7$ model parameters. Thus, the parameter index θ counts from 1 to 7.

Next the general $\mathrm{fim}_{\theta d}$ is calculated. For an arbitrary set of experimental conditions d, fim includes 7 columns and reads as given in the following equation:

$$\mathrm{fim}_d = \begin{bmatrix} 1 \\ \ln(p_a) \\ -\ln(i + \alpha_4) \\ \dfrac{-\alpha_3}{i + \alpha_4} \\ -i \\ -\exp(\alpha_7 \cdot i) \\ -\alpha_6 \cdot i \cdot \exp(\alpha_7 \cdot i) \end{bmatrix}^T \tag{16}$$

To calculate the design criterion D, as is given in Eq. (14), the initial model parameter must be given to Eq. (16). Each α corresponds to a ξ of the original model, as described in Eqs. (8)–(11), which are given in Table 5, and are calculated with an assumed cathode

Table 8
Calculated determinant by sequential search after each iteration step for different start sets

GI	D					
	0	1	2	3	4	5
Start set A	0	1.45×10^{-6}	1.61×10^{-6}	1.62×10^{-6}	1.62×10^{-6}	1.62×10^{-6}
Start set B	2.62×10^{-18}	1.58×10^{-6}	1.61×10^{-6}	1.62×10^{-6}	1.62×10^{-6}	1.62×10^{-6}
Start set C	2.06×10^{-16}	1.57×10^{-6}	1.61×10^{-6}	1.62×10^{-6}	1.62×10^{-6}	1.62×10^{-6}

pressure of 1.5 bar following equation:

$$\alpha = \begin{bmatrix} \xi_1 + \xi_2 \cdot \ln(1.5) \\ \xi_3 \\ \xi_4 - \xi_5 \cdot \ln(1.5) \\ \xi_6 \\ \xi_7 + \xi_8 \cdot 1.5 \\ \xi_9 \cdot 1.5^{-1} \\ \xi_{10} \end{bmatrix}. \quad (17)$$

Now we assume Ui-curves are experimentally measured from 0 up to 2 A/cm^2 with 14 load points each Ui-curve, and anode pressure is varied in the range from 1.1 to 2.5 bar in 0.2 bar steps.

To design the D-optimal DoE the maximal number of experiments Φ has to be defined. In this example, two times more experiments shall be carried out to identify model parameters well. This leads to $\Phi = 14$. A traditional DoE set-up by a scientist would now measure for every anode pressure a single Ui-curves, which would lead to $8 \times 14 = 112$ single measurements.

These 112 experimental conditions are given in Table 6, and are used as candidate set ($MCI_{max} = 112$) for the sequential search algorithm.

The sequential search algorithm is initialized with the first 14 candidates, which equals a initial design criterion value of $D_{initial} = -1.77 \times 10^{-44}$. This very poor value is related to the fact that there are only initial candidates at $p_a = 1.1$ bar. During first run ($d = 1$) of the algorithm the first entry of initial DoE is sequentially substituted by all 112 candidates, and the resulting design criteria are calculated. At the end candidate #108 gives best design criterion for first entry with a $D_1 = 1.46 \times 10^{-28}$, and therefore, is stored to the DoE. The design criterion improved by 72 orders of magnitude just by replacing the first entry. Best candidate for second entry is found to be candidate #100, with $D_2 = 2.29 \times 10^{-28}$. After complete optimization of all 14 entries in the DoE the sequential search algorithm repeats two times ($GI_{max} = 3$) the complete process. The final design criterion (D_{final}) can be improved to $D_{final} = 1.84 \times 10^{-27}$, and contains the candidates #1, 5, 8, 10, 13, 14, 99, 100, 103, 105, 106, 109, 111, and 122.

6. Experimental validation

6.1. D-criterion

For the validation of the D-optimal design criterion the $\phi = 14$ parameters of the PEMFC model are estimated for different sets of $\Phi = 50$ data points selected from the 405 experimental data points. For the evaluation of the estimated model, the sum of squares due to errors (SSE) between measured U_d and model predicted fuel cell voltage $\hat{U}_d$, as given in the following equation, is used.

$$SSE = \sum_{d=1}^{\phi} \left(U_d - \hat{U}_d\right)^2 \quad (18)$$

The SSE of each validation trail ($SSE_{validation}$) is plotted in Fig. 4 against the achieved determinant ($D_{validation}$) to observe potential correlations. A suitable scaling of the respective optimum (SSE_{min} and D_{max}) is chosen for a better illustration.

Clearly a strong correlation between both states can be seen. Moreover a strong decrease of SSE can be seen with increased determinant. A model with parameters from an experimental design with a small determinant has a significantly bigger SSE than a model with parameters from an experimental design with a large determinant. Since the SSE decreases strongly exponential with an increasing determinant, the D-criterion is for sure an excellent method to evaluate the quality of parameter estimation and was used for all further validations.

6.2. Sequential search

The performances of three different optimization approaches to maximize D were investigated. The proposed sequential search was compared to a Levenberg–Marquadt (LM) algorithm and a genetic algorithm (GA). Further information about both methods can be found in Refs. [9,11,12]. All three algorithms were initialized with the same set of randomly chosen 50 experimental conditions with $D_{initial} = 2.06^{-16}$. In Table 7 the maximum achieved determinant and the duration to get this result is given for each algorithm.

Investigating various configurations of the LM algorithm, only a slight improvement of D could be achieved and 1 h 20 min were needed to calculate the best result. In most of the investigated LM algorithm set-ups even no improvement of the design criterion was achieved. Using a GA an improved maximization by 8 orders of magnitude of D is gained by cost of a very long calculating time. But also the GA depends strongly on the used algorithm options and delivered no results in many cases. Using the sequential search algorithm a 24.8 times greater D in a very fast calculation of less than 3 min was found at a DELL Precision 390 Workstation, using Intel core 2 duo 1.86 GHz P, 8 GB RAM, WinXP64, and MatlabR2007b.

6.3. Robustness of sequential search

Usually there are no good start values available, which mean a good experimental design. Different sets were used for the sequential search of the optimal 50 candidates for investigating the quality of this method in consideration of the start values. The results for three different start sets are displayed in Table 8. Even starting with a very poor experimental design as start set A, the algorithm works properly. The maximum determinant was found for every start set already after three iterations ($GI = 3$) and even the first and second iteration delivered useful determinants. Thus the sequential search is not only a very fast, but also a very robust method to optimize the proposed design criterion.

6.4. Experimental effort

Another issue of this work is the question how many experiments are needed for getting satisfying parameter estimations. Therefore the ratio of the number experiments to the number parameters is plotted against the SSE between model and experimental data in Fig. 5. A significant improvement of the model can be observed till a ratio of four. More experiments deliver just a low improvement and the SSE reaches a steady state which can not be decreased because of inaccuracies of the model itself.

7. Discussion

A constraint in this work is that no dynamic database was investigated. Furthermore, the gained experimental design can be a local optimum and so there is no guarantee that not a higher, unnoticed global optimum exists.

For reducing the computation time the replacement of the determined number of iterations a in Fig. 2 by an additional algorithm stop similar to the abort criterion for numerical algorithm is one of the next steps. If gradient of D does not improve significantly between two loops, no further iterations are needed. The advantage would be an optimized number of iterations and the avoiding of unnecessary steps as seen in Table 8.

Also interesting would be a hybrid approach for optimizing the experimental plan. The sequential insertion of all candidates or the generation of random variables could be replaced by a local search of the optimum row by using a numerical algorithm like for example a LM algorithm.

Problems with the introduced approach could appear, when old stacks with thick membranes are used. These stacks often have a hysteresis of about 20 mV. In this case, the order of experimental conditions is not negligible and the SS may deliver invalid results.

8. Conclusions

It is shown that the model is valid for DoE and prior improvement of the experimental design is a promising way for enhanced parameter estimation. Especially, the D-optimal design approach is suitable to identify model parameters efficiently as Fig. 4 demonstrates. The developed sequential search is compared to usual used methods of optimization and shows a significant lower calculating time as well as a couple of orders of magnitude larger determinant.

The sequential search is moreover recommendable for DoE because of a low number of iterations are needed and the method proves a high robustness against disadvantageous chosen start designs. Already a number of experimental data points, which is as big as four times the number of model parameters, delivers an accuracy model simulation result, further experiments delivers just low improvements and are in the most cases not recommendable in view of increasing experimental effort.

All in all the prior optimization of the experimental design in order to improve the parameter estimation is an intelligent way to save time and thus to reduce the cost of experimental work significantly.

The benefit performing such a DoE is, to use expensive experimental resources much more efficiently. Traditionally, fuel cells are characterized by Ui-curves, as described in Refs. [6,7]. There, the characterization for just 4 inputs causes an effort of 405 single measurements. Using the proposed D-optimal DoE together, with the sequential search, this effort can be reduced by a factor of more than 8 to a total cost of 50 single measurements. With the proposed method, experimental groups are able to characterized fuel cells or fuel cell stacks in a very large operation area with an affordable experimental effort. This helps to speed up global fuel cell modelling tremendously, and provide urgent model-input to fuel cell system concept, design, and engineering work. This contribution helps to provide efficiently high quality models, which are the key-tools in modern drive train development. Reducing engineering costs, and improving fuel cell drive train performance is the major hurdle in commercialization this zero emission technology, and is supported by this work.

Further work is advisable in terms of taking over this general method to other fields of science and so to establish our method as a general valid way for optimizing experimental designs.

Acknowledgements

We thank all colleagues at Daimler Research Centre for their help with the experimental work and fruitful discussions. Further, the help of the Institute of Measurement, Control and Microtechnology at University of Ulm is acknowledged.

References

[1] G.E.P. Box, H.L. Lucas, Biometrika 46 (1959).
[2] S. Körkel, Numerische Methoden für Optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen, Ph.D. Thesis, University of Heidelberg, 2002.
[3] W. Kleppmann, Taschenbuch Versuchsplanung, Carl Hanser Verlag, München, 2006.
[4] H. Petersen, Grundlagen der statistischen Versuchsplanung, Band 2, ecomed, Landsberg/Lech, 1991.
[5] F. Pukelsheim, Optimal Design of Experiments, John Wiley & Sons, Inc., New York, 1993.
[6] C. Heller, Design von PEM-Brennstoffzellen und dessen Auswirkungen auf Wirkungsgrad und Komplexität von Subsystemen, Logos Verlag, Berlin, 2004.
[7] M. Schudy, Einfluss thermo- und hydrodynamischer parameter und des Zelldesigns auf den Energie- und Stofftransport in PEM-Brennstoffzellen, Verlag dissertation.de, 2004.
[8] M. Meiler, E.P. Hofer, M. Schmid, M. Schudy, An empirical stationary fuel cell model using limited experimental data for identification, in: Proceedings of workshop "Neue Technologien für Zell- und Stack-Design", Hamburg, 2007.
[9] W. Alt, Nichtlineare Optimierung, Vieweg, Wiesbaden, 2002.
[10] H. Feltl, Ein Genetischer Algorithmus für das Generalized Assignment Problem, Diploma Thesis, 2003.
[11] K. Weicker, Evolutionäre Algorithmen, Teubner, Wiesbaden, 2007.
[12] S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, T. Murata (Eds.), Evolutionary Multi-Criterion Optimization, Fourth International Conference, EMO 2007 Matsushima, Japan, March 5–8, 2007.