

PUBLICATION [P6]

Stack Temperature Estimation in System
Environment by Utilizing the Design of
Experiments Methodology

In: ECS Transactions 57 (1), pp. 205-214
© 2013 The Electrochemical Society

Reproduced by the permission of ECS – The
Electrochemical Society

Stack Temperature Estimation in System Environment by Utilizing the Design of Experiments Methodology

M. Halinen, A. Pohjoranta, J. Pennanen, and J. Kiviaho

VTT Technical Research Centre of Finland, Espoo, Finland

The behavior of the maximum temperature measured inside a SOFC stack with respect to three independent input variables (stack current, air flow and air inlet temperature) is examined by using a full factorial screening experiment, following the design of experiments methodology. The experiments were carried out with a complete 10 kW_e SOFC system. Multivariate regression models are developed to estimate said temperature and a statistical analysis is carried out on the model parameters.

Introduction

The temperature of and inside a planar solid oxide fuel cell (SOFC) stack is a property of interest for several reasons; the stack temperature affects the stack performance as well as its degradation rate (1,2). Moreover, the stack characteristics do not remain identical over its lifetime due to voltage degradation – with constant system inputs the stack temperature will increase as the stack degrades. Therefore, it is highly desirable to monitor the internal stack temperature for control and diagnosis purposes. The temperature inside an SOFC stack can be measured directly by installing measurement probes inside the stack. However, in commercial systems direct measurement of the temperature can be difficult and costly, and can also decrease the stacks' durability due to the increased complexity. Therefore, estimating the stack temperature indirectly, from other process measurements, is well-motivated. In reality, the stack temperature is certainly not a single temperature value, but an infinite number of temperature values distributed continuously over the stack. For practical purposes, however, a single representative value is chosen, and this is the approach taken in this paper as well.

This paper reports a systematic effort carried out to create a simple numerical estimator for the maximum temperature measured inside a SOFC stack. Furthermore, the test protocol utilized in this work was also used to evaluate which system inputs are most effective for controlling the said temperature.

Multivariable Regression Modeling and the Design of Experiments

Multivariable linear regression (MLR) models are functions of the form

$$\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1} = \mathbf{X}\boldsymbol{\beta}, \text{ where } \mathbf{X} = [x_1 \quad x_2 \quad \dots \quad x_{p-1}]. \quad [1]$$

x_i and \hat{y} are (column) vectors of the measured input and estimated output values, respectively. The parameter vector $\boldsymbol{\beta} = [\beta_1 \quad \beta_2 \quad \dots \quad \beta_{p-1}]^T$ is calculated so to produce

an estimate which is optimal in the least squared error sense. See e.g. (3) for details on MLR modelling.

MLR models can be evaluated based on several criteria, but in this work mostly the adjusted- R^2 [2] and the root-mean-squared-error [3] were utilized – the former for its scalability with a varying number of model parameters and the latter for its ease of interpretation. Their values were computed as follows:

$$R_{adj}^2 = 1 - \frac{(N - 1) \sum_{i=1}^N (\hat{y} - y)^2}{(N - p) \sum_{i=1}^N (\bar{y} - y)^2} \quad [2]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y} - y)^2}{(N - p)}} \quad [3]$$

In [2]-[3] N is the number of data points used, y the measured property value and \bar{y} is the mean of y . p is the number of parameters in the regression model. The closer R_{adj}^2 is to 1 or the smaller $RMSE$ is the better is the model considered to be.

Obtaining rich data that sufficiently covers the domain of interest is the most important step in creating regression models. To this end, a series of experiments were designed by following the design of experiments (DoE) methodology. The DoE methodology is especially suitable to be applied hand-in-hand with multivariable regression modeling. Good coverage of the DoE methodology is found in several textbooks, e.g. (3) and therefore DoE is not discussed here in detail.

Experimental

A full factorial experiment design, with three factors (input variables), each of them obtaining two levels (the minimum and the maximum) and a nominal level, was carried out. The system response, i.e. the maximum temperature of the stack, was obtained by recording the values of several temperature measurements placed inside a SOFC stack and then using the maximum of the measured values. The test plan, i.e. the factors, their levels and the experimental conditions' execution order is given in Table I. The test plan was designed so that the execution time of all experiment conditions (relative to start time of the designed experiments) was fixed in advance. Then, a D-optimal test plan considering the three factors plus time as the model input was calculated by varying the experiment conditions' order. The aim was to maximize the information obtained with the experiment design, also with respect to the stack performance degradation, which increases with time.

A minimum stabilization period of 24 hours was reserved per condition. A longer interval between conditions 4 and 5 is due to the weekend falling amidst the test run and the system was brought to the nominal operating condition (NOC) for this period, enabling thus also the repetition of the NOC.

The experiments were carried out according to the test plan with a complete 10 kW_e SOFC demonstration unit (4). The demo unit was equipped with a planar SOFC stack by

Versa Power Systems (5). The stack had 64 cells with ca. 550 cm² of active area. The designed experiments were started at 145 hours into the test run. Before this, the system was brought to NOC for stabilization after the start-up. All-together the test run lasted for 470 hours, of which the designed experiments (including the first and last NOC) took ca. 300 hours.

TABLE I. Test Plan of the Full Factorial Experiment and Description of Additional Experiment Conditions. The experiment condition numbers refer to figure 2(b).

Data short desc.	Exp. cond. #	Current (x_1) A	Air flow (x_2) l _N /min	Air inlet temperature (x_3) °C	Time from test start (t) h
Nominal	3, 4, 5, 10, 11, 12, 17	200	1063	695	-
DoE	6	205	962	705	145
	7	195	1162	685	169
	8	205	1162	685	193
	9	195	962	705	217
	13	195	1162	705	313
	14	205	962	685	337
	15	195	962	685	361
16	205	1162	705	385	
Part-load	1, 2	150	960	705-695	25-50

A graphical representation of the 2³ full factorial experiment design is given in Figure 1.

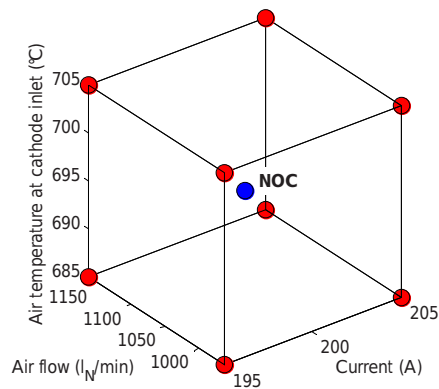


Figure 1. The input variables' domain of variation in the experiment design.

Data Analysis and MLR Modeling

Figure 2 shows the input variables' time-series during the experiment, with the steady-state experiment conditions marked with red circles. The completed experiment provided data from altogether 17 conditions (points 1-17 in Fig. 2(b)): 8 non-nominal experiment conditions (pts. 6-9, 13-16), 7 repetitions of NOC at different time instants (with a minimum of 24 hours between the instants, pts. 3-5, 10-12, 17), and 2 extra

conditions measured during system start-up when the system was operating at part-load (with stack current of 150A, pts. 1-2). The 7 repeated measurements at NOC were carried out to capture system degradation effects and the part-load conditions were used for exploratory purposes, described further below. In Figure 2, the period of the designed experiments is indicated with the dashed vertical lines.

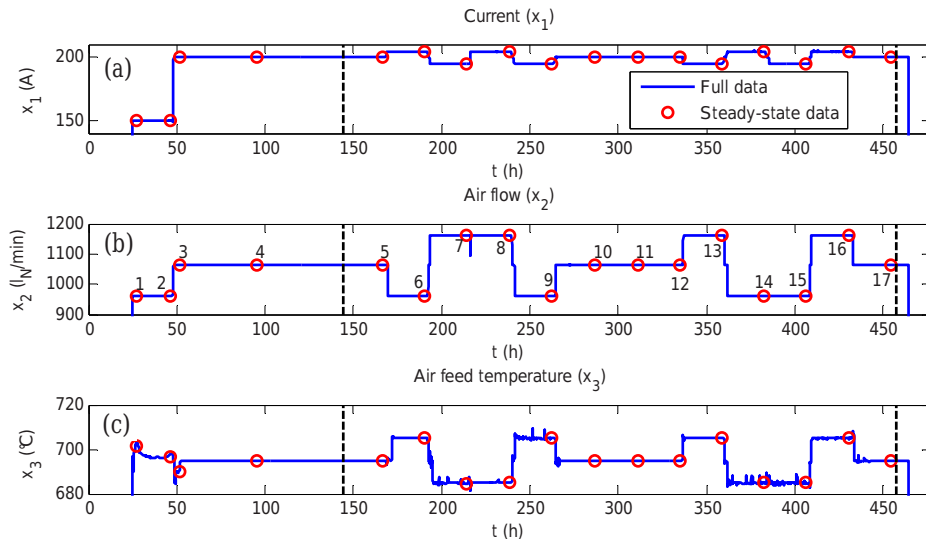


Figure 2. The input variables during the test run. Steady-state conditions are denoted with circles and indexed with a running number in subfigure (b).

Data Analysis

One aim of the data analysis was to resolve the significant input variables (and significant input interactions) with respect to the measured stack temperature. For the purpose of the input variables' significance analysis, the inputs were all scaled to the range $[-1, 1]$ by using equation [4]. The input interactions were then calculated after scaling the inputs. Only the 13 data points from the period of designed experiments (8 non-nominal conditions + 5 repetitions in NOC) were used in this analysis.

$$x_{i,sc} = 2 \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} - 1 \quad [4]$$

For further analysis, the linear regression models [5] and [6], based on all (scaled) input variables and their interactions, as well as on only the input variables were fit to the data. Figure 3 shows the model coefficients of the models per input and input interaction, and the corresponding coefficients' 95 % confidence intervals (CIs). If the confidence interval of a coefficient includes zero, the respective term in the regression model can be concluded insignificant.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 \quad [5]$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad [6]$$

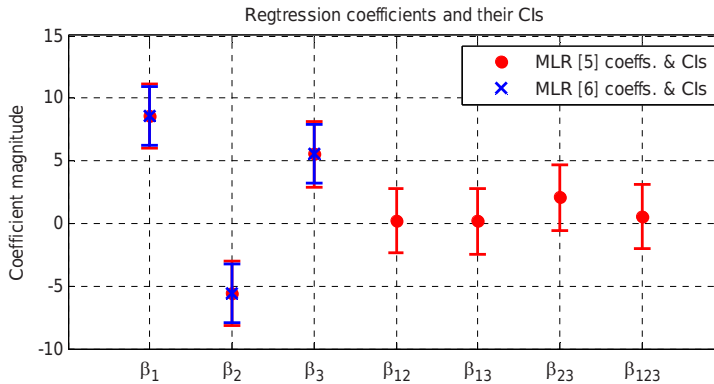


Figure 3. Regression coefficients of the MLR models [5] and [6].

Judging by the confidence intervals of model [5], all of the inputs (x_1, x_2, x_3) clearly have a significant (mean) effect on the output. Notably, the interaction of air inlet temperature and air volumetric flow (x_2x_3) seems like the only interaction term to have some significance. This would be reasonable as that term (more typically denoted $\dot{V}_{gas}T_{gas}$) is directly related to the heat transferred convectively by a gas flow.

The coefficients of model [6] enable assessing the relative significance of the inputs with respect to the output. Notably, $\beta_{1\dots 3}$ of [5] and [6] nearly coincide due to the insignificant impact of the interactions. Clearly, current (x_1) has the strongest relative effect on the output ($\beta_1=8.57$), whereas air flow (x_2) and the air inlet temperature (x_3) have a close to equal, but opposite-signed relative mean effect ($\beta_2=-5.59$, $\beta_3=5.54$). Thus, if considering control of the stack temperature within the examined operating domain, either air inlet temperature or the air flow may be used equally effectively. The air inlet temperature, however, can be manipulated without affecting the parasitic losses of the system and is thus preferable.

Preliminary MLR Modeling

With the data analysis as background, an initial MLR estimate [7] for the stack temperature was created using all three inputs as regressors. Again, only the 13 data points (8 DoE + 5 NOC) described above were used. The R_{adj}^2 and $RMSE$ values for the fitting data are 0.913 and 2.90 °C, respectively. Figure 4 and Figure 5 illustrate the obtained estimate accuracy for both the utilized steady-state data and the full data time-series, respectively. Most of the steady-state estimates lie within the point-wise confidence intervals (Figure 4a) and the estimation error has a zero mean (Figure 4b). However, when looking at the time-series data in Figure 5, problems with the estimator become obvious. The estimate is stationary and differs significantly from the measurement during set-point changes. Also, the estimator fails to capture the measurement trend, due to the stack degradation. Both of these problems are consequences of the MLR model structure.

$$\hat{y} = 141 + 1.71x_1 - 0.0559x_2 + 0.542x_3 \quad [7]$$

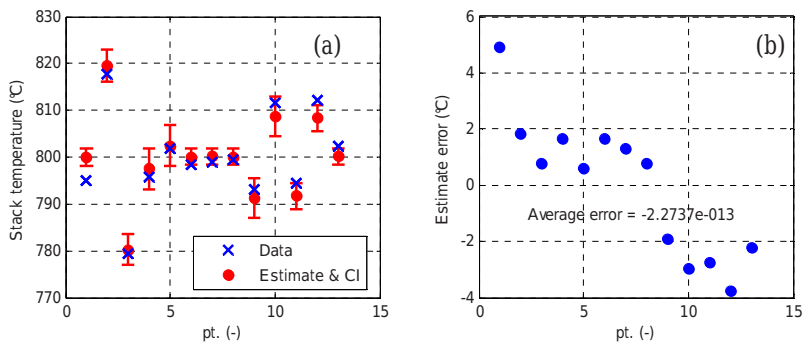


Figure 4. The steady-state data and the corresponding estimate of model [7] compared (a). The estimation error (b).

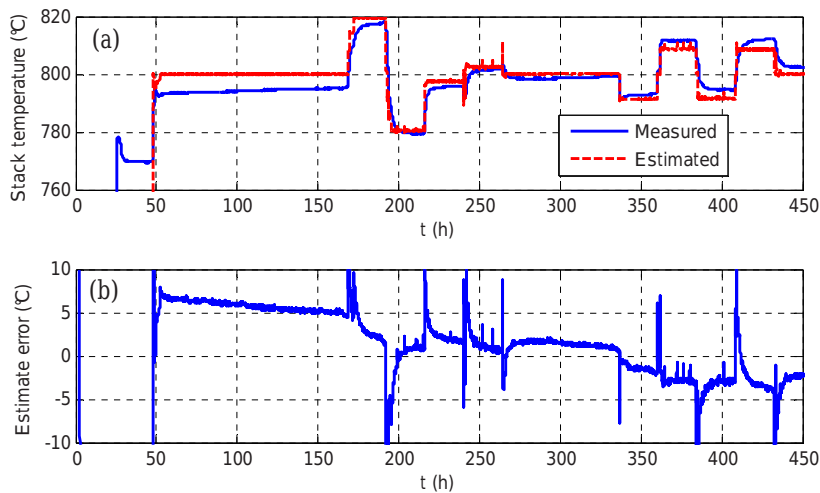


Figure 5. The time-series data and the corresponding estimate of model [7] compared (a). The estimation error (b).

Improved MLR Models

In order to improve the MLR model estimation accuracy on behalf of both the accuracy at stationary points, as well as during transients, two means were found especially useful:

- 1) Including the cathode outlet flow temperature measurement as regressor both (i) improves the estimation accuracy by capturing the measurement trend and (ii) provides a “dynamic input” to the MLR estimator, thus improving the estimate accuracy during transient phases
- 2) Expanding the model fitting data series to include part-load data improves the model accuracy as well as expands its feasible application area.

Cathode Outlet Flow Temperature as Regressor

An MLR model [8] with the cathode outlet flow temperature (x_4) as a regressor variable was created to enable better estimation accuracy during both transient operation as well as in steady-state. The model R_{adj}^2 and $RMSE$ values for the fitting data are 0.991 and 0.946 °C, respectively.

$$\hat{y} = -75.5 - 0.0877x_1 - 0.0265x_2 - 0.290x_3 + 1.51x_4 \quad [8]$$

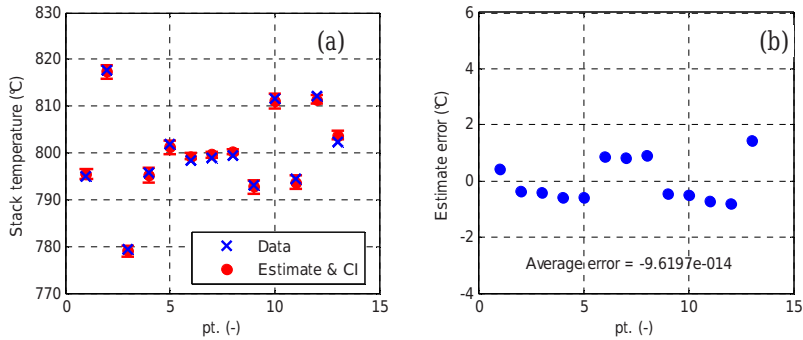


Figure 6. The steady-state data and the corresponding estimate of model [8] compared (a). The estimation error (b).

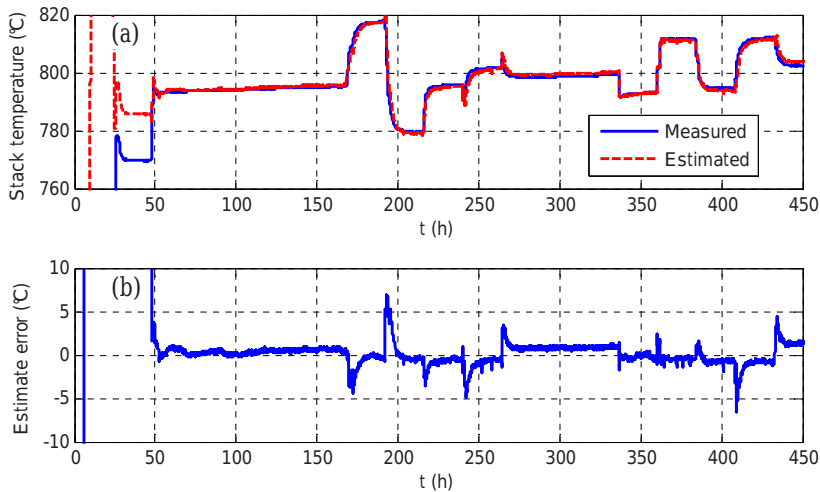


Figure 7. The time-series data and the corresponding estimate of model [8] compared (a). The estimation error (b).

When comparing Figure 4a and 6a, it is clear how introducing x_4 as regressor narrows the estimate confidence intervals (for the fitting data). Also the reduction in the estimation error from Figure 4b to 6b is evident. The most significant improvement in the

estimation accuracy is, however, noticeable in the time-series data, Figure 7a-b. Since x_4 is a stack output variable, it changes according to the stack dynamics and thus provides the MLR estimate with information of the stack dynamics. The downside of including x_4 as a regressor is that the MLR model coefficients lose their physical interpretability (because the regressors are no longer independent of each other).

Although the estimator [8] already provides good estimation accuracy when operating around the nominal operating conditions, the accuracy is poor outside this operating domain. To correct this, the MLR model parameter fitting data was expanded to include part-load operation data, although this data was not a part of the original test plan.

Adding Part-load Data to Regression Coefficient Calculation

The system was operated at part-load (150 A) during 25-50h into the test run and, as seen in e.g. Figure 7a, the stack temperature is significantly lower there than at NOC. Two measurements from this operating period were included in the MLR model parameter calculation data, yielding model [9], with R_{adj}^2 and $RMSE$ for the fitting data 0.994 and 1.06 °C, respectively. The estimation results for model [9] are shown in Figures 8-9.

$$\hat{y} = -33.8 + 0.353x_1 - 0.0327x_2 - 0.110x_3 + 1.17x_4 \quad [9]$$

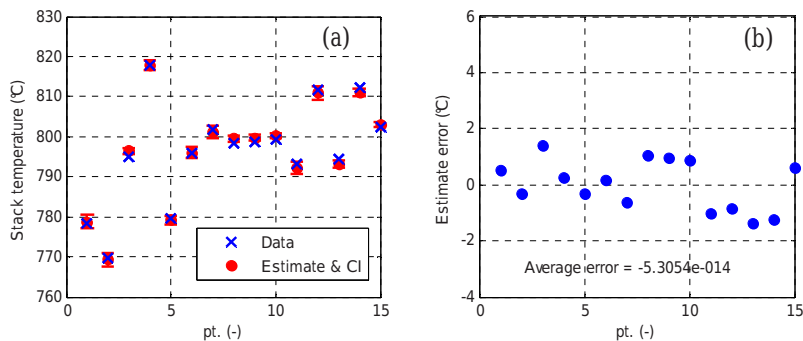


Figure 8. The steady-state data and the corresponding estimate of model [9] compared (a). The estimation error (b).

As is seen in Figure 8 and 9, the model estimation accuracy is still good, though the $RMSE$ value for the fitting data increased compared to [7] and also the error in Figure 9b is visibly greater than in Figure 7b. With the slight sacrifice of estimation accuracy at NOC, the estimator can be seen to have significantly improved over the part-load operating period (25-50h) in the data time-series, Figures 7a and 9a. The confidence intervals of the fitting data estimates still remain narrow, also for the added data points (pts. 1-2).

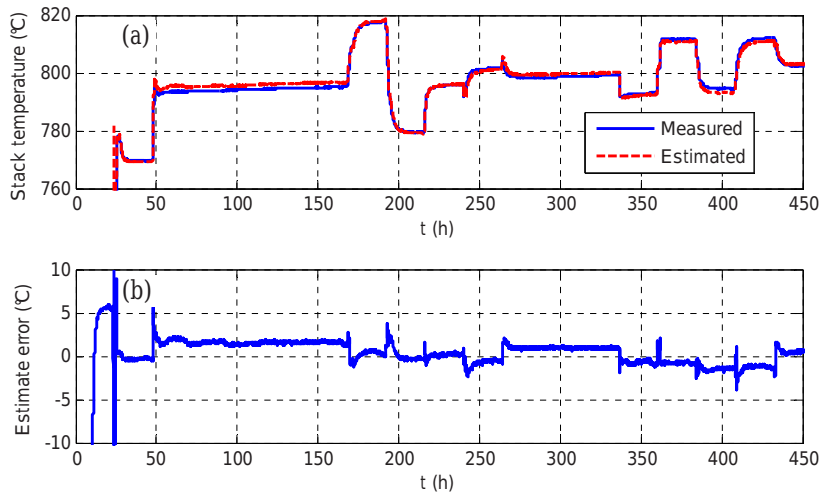


Figure 9. The time-series data and the corresponding estimate of model [9] compared (a). The estimation error (b).

Conclusions

A series of multi-variable linear regression models were created based on experimental data from a complete 10kW SOFC system to estimate the maximum internal temperature measured from a SOFC stack. The data was obtained with a full factorial experimental design, where three input variables (stack current, air flow and air inlet temperature) were manipulated.

The benefits of a systematic, designed, experimental procedure were found to manifest as straightforward regression modelling and in the accuracy of the consequent models. Furthermore, the full factorial experimental design enabled clearly quantifying the significance of the inputs with respect to their impact on the monitored output value. This information is valuable when developing a control strategy for the system.

Because the MLR model itself is only a static mapping of a series of regressor variables to the estimated output it was found beneficial, from a practical viewpoint, to include such variables to the regressors which convey the dynamics of the estimated property. In this case the cathode outlet flow temperature was used as regressor to include the dynamics of the stack temperature. Furthermore, as the designed experiments may cover only a part of the system operating domain, it is beneficial to include data from outside the domain of the designed experiments. In this case, part-load data was used to expand the applicable area of the MLR model.

Acknowledgments

Funding for this study was obtained through the project RealDemo. The Finnish Funding Agency for Technology and Innovation (TEKES) as well as the companies participating in the project are gratefully acknowledged for their financial support. Additionally, A. Pohjoranta would like to thank the gentlemen Fabio Postiglione from the University of Salerno and Angelo Esposito from the European Institute for Energy Research for the many helpful discussions related to the design of experiments and data analysis. M. Pastula from Versa Power Systems is acknowledged for good collaboration and helpful comments. M. Rautanen and M. Kotisaari from VTT are thanked for assisting in the experiments.

References

1. R. Knibbe, A. Hauch, J. Hjelm, S. Ebbesen and M. Mogensen, *Green*, **1**, 141 (2011)
2. A. Nakajo, Z. Wullemin, J. Van herle and D. Favrat, *J. Power Sources*, **193**, 203 (2009)
3. D.C. Montgomery, *Design and Analysis of Experiments*, 8th ed., John Wiley & Sons, Singapore (2009)
4. M. Halinen, M. Rautanen, J. Saarinen, J. Pennanen, A. Pohjoranta, J. Kiviaho, M. Pastula, B. Nuttall, C. Rankin, B. Borglum, *ECS Trans.*, **35**(1), 113 (2011)
5. B. Borglum, E. Tang, M. Pastula, *ECS Trans.*, **35**(1), 63 (2011)