Modelling the production of soluble hydrogenase in Ralstonia eutropha by on-line optimal experimental design

Flavia Neddermeyer * Volker Marhold * Christoph Menzel * Dominik Krämer * Rudibert King

*Chair of Measurement and Control, Technische Universität Berlin, Secr. ER 2-1, Hardenbergstraße 36a, 10623 Berlin, Germany (e-mail: flavia.neddermeyer@tu-berlin.de, rudibert.king@tu-berlin.de).

Abstract: This paper presents a case study on integrating a new model extension describing the production of the enzyme soluble hydrogenase (SH) into an already existing model by means of on-line optimal experimental design (OED). These investigations were done on the autotrophic cultivation system and process model of Ralstonia eutropha, i.e., a cultivation where aeration by air is replaced by a gas consisting of H₂/O₂/CO₂. Prior to the experiment, three different structures of the new model extension are postulated and coarsely identified. Off-line OED with the best model yields initial feeding trajectories. During the experiment, a repetitive model refinement is performed. This consists of identifying the parameters of the new model extension for all three models, selecting the best model and recalculating optimal feeding trajectories. It is shown that on-line OED in the early modelling stage followed by a manual modelling step results in an acceptable quality of the description of SH production.

Keywords: autotrophic cultivation, closed-loop, fed-batch, Fisher-information matrix, Ralstonia eutropha, model-based control, modelling, optimal experimental design, soluble hydrogenase.

1. INTRODUCTION

The aim of this work is to add and validate a model extension describing the production of the enzyme soluble hydrogenase (SH) to an existing autotrophic cultivation model of Ralstonia eutropha (R. e.). On-line optimal experimental design (OED) is used for planning and control. The focus is on the enzyme SH because it reduces expensive cofactors (see Lauterbach et al. (2011)) which are employed in biotransformation processes. Whenever cofactors participate in a reaction, they have to be replaced or regenerated. In situ cofactor regeneration by applicable enzymes such as SH reduces costs (Zhao and van der Donk (2003)). To utilize SH in industrial biotransformations, large amounts are required. R. e. is a bacterium that expresses SH. Therefore, the cultivation of R. e. represents a suitable SH production process. A general process model as well as a cultivation system already exist for this strain. To optimize the expression of SH efficiently by means of model-based control, a model extension describing the production of SH has to be integrated into the existing process model. Hence, this work focuses on the modelling of SH by on-line OED.

Optimal experimental design is a powerful tool to increase the quality of models using comparatively few experiments (Franceschini and Macchietto (2008), Bernaerts et al. (2000)). Especially for time consuming biological cultivations diminishing the number of experiments is beneficial. If the structure of the biological model is known and the parameters are coarsely estimated, OED yields the trajectories for experiments with the highest information content for a subsequent parameter identification. In this case study, the iterative OED cycle is run on-line to improve and discriminate between initial models of low quality during the cultivation by parameter estimation and model selection. On-line parameter estimation and OED using a single model was also investigated in a simulation study by Qian et al. (2014). Schenkendorf and Mangold (2013) studied robust on-line model discrimination in silico by planning the process on-line in a discriminatory manner. In the work presented here, the optimal stimuli for the real cultivation resulting in data with high information content for the subsequent parameter estimation of the priorly selected model are calculated. For this, the overall process model introduced in Sec. 3 is used. This model describes biomass growth but not the production of the target compound SH. Since OED, in contrast to traditional experimental planning, requires a description of SH production prior to the experiment, three structures of possible model extensions are proposed. The initial parameters are identified based on 14 SH measurements in total, derived from three different preliminary cultivations (not shown here). The model with the best cost function value (defined in Sec. 4) is used for an initial calculation of the optimal trajectories prior to the start of the experiment. During the cultivation, sampling and SH analysis, parameter estimation (est.), model selection and OED form a repetitive on-line cycle resulting in a selected model with estimated parameters. Figure 1 illustrates the off-
Fig. 1. On-line OED workflow for SH modelling. The repetitive on-line cycle is marked in grey.

and on-line workflow. In order to utilize this method, the measurements must be present on-line or at least at-line. The priori defined SH model extensions suffice for OED but might be structurally incomplete. Hence, a manual modelling step evaluating the data gained in the on-line OED experiment finishes the workflow.

This paper is organized as follows: After describing the experimental setup in Sec. 2 and the overall process model in Sec. 3, the control scheme for on-line OED is explained in Sec. 4. Section 5 presents different SH descriptions of three models which are used in the on-line OED cultivation. The results of the OED-planned experiment including post-experimental modelling are illustrated in Sec. 6. Finally, a conclusion of this case study is drawn in Sec. 7.

2. MATERIALS AND METHODS

All in all, three fed-batch cultivations were performed. One was run with on-line OED and two cultivations served exclusively for validation. The used strain was Ralstoniaeutropha (H16). The feeding compositions are given in Tab. 1. The initial volume of the cultivation was 10.5 L of defined medium including inoculum.

<table>
<thead>
<tr>
<th>Feeding</th>
<th>Chemical</th>
<th>Conc. [g-L⁻¹]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphate (P)</td>
<td>Na₃H₂PO₄ (2H₂O)</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>KH₂PO₄</td>
<td>13</td>
</tr>
<tr>
<td>Ammonium (N)</td>
<td>NH₄Cl</td>
<td>296</td>
</tr>
<tr>
<td>Iron and trace elements</td>
<td>MgSO₄ (7H₂O)</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>CaCl₂ (2H₂O)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>FeCl₃ (6H₂O)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NiCl₂ (6H₂O)</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>MnCl₂ (2H₂O)</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>H₂BO₃</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>CuCl₂ (2H₂O)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Na₂MoO₄ (2H₂O)</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>CoCl₂ (6H₂O)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

The initial medium concentrations of iron and trace elements were 1/100 of those in the feeding. Ammonium concentration was 1/5 of the feeding and phosphate 1/10 of the feeding concentration. The digital control unit of the explosion-protected 15 L stirred tank reactor with 3 Rushton turbines regulated the temperature at 30°C, the pH at 6.8 and the stirrer speed at 500 rpm. The dissolved gas probes are Mettler Toledo InPro6800 and InPro5000i, the gas sensors are the BlueOxOne Cell and BCP-H2 from BlueSens, the mass flow controllers and the pressure sensor are produced by Bronkhorst. Optical density (OD) was manually measured at 436 nm photometrically.

3. OVERALL PROCESS MODEL

To keep this contribution self-contained, the model presented by Neddermeyer et al. (2015) is summarized here. The evaluation of the states biomass (X), ammonium (N), phosphate (P), polyhydroxybutyrate (PHB), membrane-bound hydrogenase (MBH) and of the dissolved (index l) gases are derived from mass balances. The amount of fed base minus fed acid (BaAc) and reactor liquid phase volume (V₁) are derived from volumetric balances. Calculating the evolution of the states is expressed by reaction rates (μ), yield coefficients (Y), conversion factors (K) and terms describing gas transport (index tr) between the two phases. The model is given by:

\[
\begin{align*}
\dot{\mu}_X &= (Y_{N,X,PHB} + Y_{P,X,PHB} + 1) \cdot \mu_X \cdot \mu_X \cdot m_X + \mu_X \cdot m_X \\
\dot{\mu}_N &= -Y_{N,X,PHB} \cdot \mu_X \cdot \mu_X \cdot m_X - Y_{N,MBH} \cdot \alpha_{MBH} + u_N \cdot \alpha_{feed} \\
\dot{\mu}_P &= -Y_{P,X,PHB} \cdot \mu_X \cdot \mu_X \cdot m_X + u_P \cdot \alpha_{feed} \\
\dot{\mu}_{PHB} &= \mu_{PHB} \cdot \mu_X - \mu_X \cdot \mu_X \\
\dot{\mu}_{MBH} &= \mu_{MBH} \cdot \mu_X \\
\dot{\mu}_{H_{2},1} &= -(Y_{H_{2},X,PHB} \cdot \mu_X \cdot \mu_X + Y_{H_{2},X,PHB}) \cdot \mu_X + \mu_{H_{2}} \\
\dot{\mu}_{CO_{2},1} &= -(Y_{CO_{2},X,PHB} \cdot \mu_X + Y_{CO_{2,PHB} \cdot \mu_{PHB}}) \cdot \mu_X + \mu_{CO_{2}} \\
\dot{\mu}_{O_{2},1} &= -(Y_{O_{2},X,PHB} \cdot \mu_X \cdot \mu_X + Y_{O_{2},X,PHB}) \cdot \mu_X + \mu_{O_{2}} \\
\dot{\mu}_{BaAc} &= (Y_{N,X,PHB} + Y_{P,X,PHB} \cdot \mu_X \cdot \mu_X + Y_{N,MBH} \cdot \alpha_{MBH} - 18 \cdot \frac{Y_{N,MBH}}{Y_{Na}} \cdot \frac{Y_{P,PHB}}{Y_{P,PHB}} + Y_{H_{2}O,PHB} \cdot \mu_X \cdot \mu_X - Y_{H_{2}O,PHB} + 1) \cdot \frac{\mu_{BaAc} \cdot \mu_{CO_{2},1} \cdot 1000 \cdot m_{BaAc}}{K_{BaAc,CO_{2}} \cdot \mu_{CO_{2},1} \cdot Y_{H_{2}O,PHB} \cdot \mu_X \cdot \mu_X - Y_{H_{2}O,PHB} + u_{BaAc}} \cdot \frac{u_{BaAc}}{u_{BaAc}} \\
\dot{\mu}_{I} &= \frac{1}{m_{Phosph}} \cdot Y_{H_{2}O,PHB} \cdot \mu_X \cdot \mu_X + \frac{1}{m_{Phosph}} \cdot Y_{H_{2}O,PHB} \cdot \mu_X \cdot \mu_X - Y_{H_{2}O,PHB} + u_{BaAc} \\
\mu_X &= \mu_{max,X} \cdot \text{MiMe}(c_{N}, k_{N}^{X}) \cdot \text{MiMe}(c_{P}, k_{P}^{X}) \cdot \text{MiMe}(c_{N}^{2}, k_{N}^{2X}) \cdot \text{Spec}_{2}(c_{O_{2},1}, k_{O_{2}}^{X}) \cdot \text{Spec}_{2}(c_{CO_{2},1}, k_{CO_{2}}^{X}) \\
\mu_{PHB} &= \mu_{max,PHB} \cdot \text{Spec}_{3}(c_{N}, k_{N}^{PHB}) \cdot \text{Spec}_{3}(c_{P}, k_{P}^{PHB}) \cdot \text{Spec}_{3}(c_{N}^{2}, k_{N}^{2PHB}) \cdot \text{Spec}_{3}(c_{O_{2},1}, k_{O_{2}}^{PHB}) \cdot \text{Spec}_{3}(c_{CO_{2},1}, k_{CO_{2}}^{PHB}) \cdot \text{Spec}_{3}(c_{MBH}, k_{MBH}^{PHB}) \cdot \text{Spec}_{3}(c_{MBH}, k_{MBH}^{PHB}) \\
\mu_{MBH} &= \mu_{max,MBH} \cdot \text{MiMe}(c_{N}, k_{N}^{X}) \cdot \text{MiMe}(c_{P}, k_{P}^{X}) \cdot \text{MiMe}(c_{N}^{2}, k_{N}^{2X}) \cdot \text{MiMe}(c_{O_{2},1}, k_{O_{2}}^{PHB}) \cdot \text{MiMe}(c_{CO_{2},1}, k_{CO_{2}}^{PHB}) \cdot \text{MiMe}(c_{MBH}, k_{MBH}^{PHB}) \cdot \text{MiMe}(c_{MBH}, k_{MBH}^{PHB}) \\
\end{align*}
\]

For identification, many different measurements are available. Prior to SH measurements, the measurement vector consisted of biomass concentration (y₁), optical density (y₂), gas fractions in the headspace (y₃₋₆), system excess pressure (y₇), inlet gasflows (y₇₋₉), ammonium and phosphate concentrations (y₁₀₋₁₁), dissolved carbon-dioxide and oxygen (y₁₂₋₁₃), volume of base fed minus acid fed.
(y_{14}), specific activity of MBH (y_{15}), PHB concentration (y_{16}) and liquid volume (y_{17})

\[
y_1 = m_X + m_{PHB} / V_l \\
y_2 = K_{OD,X} \cdot \frac{m_X}{V_l} + K_{OD,PHB} \cdot \frac{m_{PHB}}{V_l} \\
y_3 = x_{H_2,v} \\
y_4 = x_{CO_2,v} \\
y_5 = x_{O_2,v} \\
y_6 = \frac{p_{v}}{1000 \, mm} \\
y_7 = \frac{(m_{CO_2, set} + m_{O_2, set}) \cdot R \cdot T \cdot 10^3 \, L}{m} \\
y_8 = \frac{(m_{CO_2, set} + m_{O_2, set}) \cdot R \cdot T \cdot 10^3 \, L}{m} \\
y_9 = \frac{(m_{O_2, set} + m_{CO_2, set}) \cdot R \cdot T \cdot 10^3 \, L}{m} \\
y_{10} = \frac{m_{BA}}{V_l} \\
y_{11} = \frac{m_{Fe}}{100 \, mg} \\
y_{12} = \frac{V_{pH}}{100 \%} = pO_2 \\
y_{13} = V_{BaAc} \\
y_{14} = \alpha_{MBH} \\
y_{15} = \frac{m_{a}}{m} \\
y_{16} = \frac{m_{v}}{m} \\
y_{17} = \frac{1}{V_l}
\]

with \(y_7 - y_9\) depending on environmental pressure (\(P_0\)), a leakage flow (index out) and molar masses (\(M\)). All gaseous (index \(v\)) volume flows (\(q\)) and liquid volume flows (\(u\)) are assembled in the input vector of the system

\[
\mathbf{u}^T = (u_N, u_{Fe}, u_P, q_{H_2,v}, q_{CO_2,v}, q_{O_2,v}, u_{base}, u_{acid}, u_{antifoam}).
\]

4. CLOSED-LOOP CONTROL SCHEME

The organism *Ralstonia eutropha* is cultivated autotrophically in this study, meaning that it assimilates carbon dioxide and gets its energy from oxidising hydrogen. Hence, dissolved gas concentrations determine growth. Since the dissolved gas amounts depend on the gas composition in the headspace, commanded variables are the gas fractions in the headspace \((x_{H_2,v, set}, x_{CO_2,v, set}, x_{O_2,v, set})\), see Fig. 2).

![Closed-loop control scheme for the autotrophic cultivation of R. e.](image)

The desired gas composition and the desired excess pressure (\(\Delta P_{set}\)), which is set constant, are realized by an underlying multiple-input multiple-output PI gas phase controller which calculates the required gas flows (\(q_{H_2,v}, q_{CO_2,v}, q_{O_2,v}\)). Besides set or reference values, the estimated gas flows are needed for gas phase control. Those are passed from a feed-forward disturbance rejection (FFDR) which evaluates the estimated state vector \((\hat{x}_{ext})\) and calculates the simulated gas flow \(q_{gas,v,est} = \sum_j y_j\). A sigma-point Kalman filter (SPFK) estimates the state vector which is not only passed to the FFDR but also serves as the initial vector for on-line OED. Manipulating variables for OED are the gas fractions in the headspace. In contrast to the traditional strategy, during the OED experiment SH has to be analyzed at-line for an on-line parameter identification of three postulated models. Afterwards, the model with the smallest cost function value

\[
S = \sum \frac{(y_{sim} - y_{SH})^2}{\sigma^2}
\]

is selected and used to calculate the optimal gas fraction trajectories, with \(y_{sim}\) and \(y_{SH}\) being the simulated and measured SH activities, respectively. For all on-line calculations, the variance \(\sigma^2\) is regarded to be SH measurement noise which was chosen as \((0.1 \, y_{sim})^2\), but at least \(0.2^2 \, U^2 \, mg_{O_2}\). The variance increases with high measured SH values because the sample has to be diluted prior to measurement which is error-prone. A minimal variance is proposed due to the accuracy of the method.

Parameter identification, model selection and trajectory calculation form the repetitive on-line cycle. Ideally, in the course of the experiment, the calculated trajectories converge and the selected model is consistent. In case the chosen model changes often during the experiment, the models do not differ significantly.

To analyze the parameter uncertainties and for OED the Fisher information matrix \(F\) is calculated according to Walter and Pronzato (1997) as

\[
F = \frac{\partial y_{SH}}{\partial p}^T \cdot \frac{1}{\sigma^2} \cdot \frac{\partial y_{SH}}{\partial p},
\]

where \(y_{SH}\) is the SH measurement and \(p\) the parameter vector of the SH production. It is assumed that the errors between an SH simulation and measurements are mainly caused by the model structure. Thus, for all post-experimental calculations the variance \(\sigma^2\) of the SH measurement is approximated by

\[
\sigma^2 = \frac{1}{n_m - n_p} \cdot \sum (y_{sim} - y_{SH})^2,
\]

with \(n_m\) being the number of measurements and \(n_p\) the number of parameters. Here, the A-criterion is taken as a cost function in the minimization problem,

\[
\Phi = \arg \min \{\text{trace}(F^{-1})\}.
\]

5. SH MODEL EXTENSIONS FOR ON-LINE OED

For OED, the model presented in Sec. 3 has to be expanded prior to the experiment. Descriptions of three different SH model extensions used in on-line OED are given below.

5.1 State and measurement equations

The amount of the enzyme SH (\(m_{SH}\)) depends on the active biomass \(m_X\) and the formation rate \(\mu_{SH}\)

\[
m_{SH} = \mu_{SH} \cdot m_X.
\]
Table 2. List of used kinetic functions $g(c)$ depending on the substrate concentrations $c$ and the constant parameters $k$ that have to be identified.

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Kinetic function $g(c)$</th>
<th>Standardisation $g_{\text{max}}$</th>
<th>Type</th>
<th>Trend curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michaelis Menten</td>
<td>MiMe(c, k)</td>
<td>$\frac{c}{c+c_k}$</td>
<td>-</td>
<td>lim</td>
<td></td>
</tr>
<tr>
<td>Aiba</td>
<td>Ai(c, k)</td>
<td>$e^{-kc}$</td>
<td>-</td>
<td>inh</td>
<td></td>
</tr>
<tr>
<td>Special</td>
<td>Spec$_1$(c, k)</td>
<td>$\frac{g_{\text{max}}e^{-kc}}{1+e^i(c)}$</td>
<td>$1 + \frac{1}{k}$</td>
<td>lim, inh</td>
<td></td>
</tr>
<tr>
<td>Special</td>
<td>Spec$_2$(c, k$_1$ k$_2$)</td>
<td>$\frac{g_{\text{max}}e^{-kc}}{1+e^i(c)}$</td>
<td>$1 + \frac{1}{k} \left( \frac{k_2(k_2-1)}{k_1} \right)$</td>
<td>lim, inh</td>
<td></td>
</tr>
<tr>
<td>Special</td>
<td>Spec$_3$(c$_1$ c$_2$, k$_1$ k$_2$)</td>
<td>max$(e^{-k_1c_1}, e^{-k_2c_2})$</td>
<td>-</td>
<td>inh</td>
<td></td>
</tr>
<tr>
<td>Møller</td>
<td>Mo(c, k$_1$ k$_2$)</td>
<td>$\frac{c^{k_1}}{c^{k_1}+c^{k_2}}$</td>
<td>-</td>
<td>step</td>
<td></td>
</tr>
</tbody>
</table>

For Model 1 and 2, $\mu_{\text{SH}}$ implies exclusively formation. Degradation is neglected as it was deemed not to be beneficial from an energetic point of view. As for MBH, see Sec. 3, the actual mass of SH cannot be measured directly. However, in a certain enzymatic reaction, the activity $A_{\text{SH}}$ of SH can be quantified by the amount of used proteins

$$ y_{\text{SH}} = \frac{10^3 \cdot A_{\text{SH}}}{m_{\text{Pr}}} = a_{\text{SH}} $$

with the preliminary analyzed averaged amount of protein $m_{\text{Pr}} = 0.6 \cdot m_X$. Specific activity ($a_{\text{SH}}$) is assumed to be proportional to the mass of the enzyme

$$ a_{\text{SH}} = K_{\text{SH,a}} \cdot m_{\text{SH}}.$$\hspace{1cm} (8)

Combining (6) and (8) yields

$$ \dot{a}_{\text{SH}} = K_{\text{SH,a}} \cdot \mu_{\text{SH}} \cdot m_X.$$\hspace{1cm} (9)

5.2 Kinetics

Any kind of production rate in chemical and biochemical reactions depends on the availability of the substrates and the exposure to inhibitors. Those biochemical dependencies are described in various kinetic functions. Table 2 shows the kinetic functions that are used here. The Michaelis-Menten (MiMe) kinetic describes growth on a limiting (lim) substrate. An inhibiting (in) dependency on a single substrate is expressed by the Aiba (Ai) kinetic (Bastin and Dochain, 1990). Limiting effects on growth at low concentrations followed by inhibition at high substrate amounts (lim, in) are realized with a special function, e.g., Spec$_2$ taken from Roessner (2014).

Model 1: For the first model it is assumed that high concentrations of dissolved hydrogen and the amount of already existing SH both inhibit the expression of SH resulting in

$$ \mu_{\text{SH}} = \mu_{\text{SH, max}} \cdot \text{Ai}(c_{H_2}, k_{H_2}^{\text{SH}}) \cdot \text{Ai}(a_{\text{SH}}, k_{\text{SH}}^{\text{SH}}). $$\hspace{1cm} (10)

The first kinetic term is chosen in analogy to the MBH production. The second term implies that there is a limit of SH production in a cultivation, e.g., due to an inhibition caused by SH directly or indirectly. Unfortunately, in the experiment conducted, the second term was erroneously implemented as $\text{Ai}(A_{\text{SH}}, k_{\text{SH}}^{\text{SH}})$ using the absolute value $A_{\text{SH}}$ instead of the relative one, $a_{\text{SH}}$. Since the expression rate $\mu_{\text{SH}}$ (10) is inserted into (9), and $K_{\text{SH,a}}$ is also unknown, only the product of $K_{\text{SH,a}}$ and $\mu_{\text{SH, max}}$ can be identified.

Model 2: The second model also postulates an inhibiting effect of dissolved hydrogen and SH on the expression of SH and postulates an optimum for dissolved oxygen

$$ \mu_{\text{SH}} = \mu_{\text{SH, max}} \cdot \text{Ai}(c_{H_2}, k_{H_2}^{\text{SH}}) \cdot \text{Ai}(a_{\text{SH}}, k_{\text{SH}}^{\text{SH}}) \cdot \text{Spec}_2(c_{O_2}, k_{O_2}^{\text{SH}}). $$\hspace{1cm} (11)

Oxygen concentrations beyond the optimum are not beneficial for SH expression accordingly. This kinetic term is chosen because for general growth (see Sec. 3) a similar kinetic yields good results. As in Model 1, the parameters $K_{\text{SH,a}}$ and $\mu_{\text{SH, max}}$ can only be identified together and the second term was implemented as $\text{Ai}(A_{\text{SH}}, k_{\text{SH}}^{\text{SH}})$.

Model 3: The third model assumes a minimal SH amount at high hydrogen concentrations ($m_{\text{SH, min}}$) and a maximal SH amount at low hydrogen concentrations ($m_{\text{SH, max}}$). Due to cellular regulations it is postulated that the cell aims for a certain value of SH between these boundaries depending on the availability of $H_2$ and $O_2$. Naming this value a saturation amount $m_{\text{SH, sat}}$, the kinetic is given by

$$ \mu_{\text{SH}} = \mu_{\text{SH, max}} \cdot (m_{\text{SH, sat}} - m_{\text{SH}}). $$\hspace{1cm} (12)

Multiplying this equation with $K_{\text{SH,a}}$ gives an expression in the specific activities.

$$ K_{\text{SH,a}} \cdot \mu_{\text{SH}} = \mu_{\text{max, SH}} \cdot (a_{\text{SH, sat}} - a_{\text{SH}}). $$\hspace{1cm} (13)
The saturation activity $a_{\text{SH, sat}}$ depends on the dissolved hydrogen and oxygen concentration. Low hydrogen concentrations enhance SH expression, according to an Aiba kinetic with the parameter $k_{\text{H}_2}^{\text{SH}}$

$$a_{\text{SH, sat}} = a_{\text{SH, min}} + (a_{\text{SH, max}} - a_{\text{SH, min}}) \cdot \left( \frac{[\text{H}_2]}{K_{\text{H}_2}^{\text{SH}}} \right).$$  \hspace{1cm} (14)

In contrast to Model 1 and 2, a maximum amount of SH per cell exists ($a_{\text{SH, max}}$) and the amount of produced SH does not affect its formation. Moreover, the activity of SH converges towards the saturation activity and thus can decrease.

6. RESULTS

In the first part of this section, the on-line OED results are presented. The identified parameters are listed in Tab. 3 for Model 2 only. Since none of the initially proposed models were able to describe the validation data to an acceptable level, the on-line OED experiment was interpreted manually leading to an alternative and validated model for SH. The alternative model is presented in the second part of this section. It also includes an extension of the general growth description. This was possible, as on-line OED pushed the process into a region which was not considered before and therefore was not included in the general model.

**On-line OED results:** During on-line OED the optimizer calculated dynamic inputs for the gas fractions of hydrogen and oxygen (see Fig. 3, lower right). This is the case because the gas fractions determine the dissolved gas concentrations which are relevant for the kinetic formulas in the reaction rates of the proposed model extensions. In the experiment, Model 2 was the finally selected model. In the course of the cultivation the model number changed only in the beginning as it was expected since the models differ significantly. Identified parameters are shown in Tab. 3. The identified values of $K_{\text{SH,a}}$, $\mu_{\text{SH,max}}$ and $K_{\text{H}_2}^{\text{SH}}$ hit pre-set constraints. If the model were used for further investigations, the plausibility of the identified parameter values would have to be analyzed. However, Model 2 does not describe the SH measurements of the validation cultivations to a sufficient extent see Fig. 4. As mentioned in Sec. 5 the absolute activity was implemented erroneously in the Aiba kinetic instead of the relative one. Since the absolute activity increases with biomass, at high biomass amounts the SH production rate is always low according to the Aiba kinetics. Vice versa, at low biomass amounts in the early stage of each cultivation, the model predicts high SH production rates contradicting the measurements of the validation experiments. However, the on-line OED experiment supplied a rich data set which was used for manual modelling of SH. In the following, a subsequent alternative model is presented.

**Alternative model:** Due to the initially unknown fact that measured SH increases and decreases over time, the alternative model includes SH formation and degradation

$$\dot{a}_{\text{SH}} = (\mu_{\text{SH}} - \mu_{\text{SH, deg}}) \cdot m_X.$$  \hspace{1cm} (15)

With (8) resulting in

$$\dot{a}_{\text{SH}} = K_{\text{SH,a}} \cdot (\mu_{\text{SH}} - \mu_{\text{SH, deg}}) \cdot m_X.$$  \hspace{1cm} (16)

The degradation

$$\mu_{\text{SH, deg}} = \mu_{\text{SH, deg, max}} \cdot \text{MiMe}(\text{CO}_2, k_{\text{deg, SH}})$$  \hspace{1cm} (17)

is realised through a MiMe kinetic. As only activities are measured, an apparent decrease of SH can be caused by degradation but also by an inactivation or inhibition. Moreover, in contrast to the production of MBH, the
expression of SH seems not to be influenced by H₂ as supposed above. Thus, only a dependence on O₂ is introduced
\[ \mu_{SH} = \mu_{SH,max} \cdot \text{Spec}_2(\text{CO}_2, l, k_3, \text{SH}, \text{O}_2, k_4). \]  
(18)

In Fig. 5, the simulations of the alternative model (first row) are shown which generally describe the experimental results much better than with Model 2.

![Fig. 5. Development of concentration measured (circles) and simulated by the alternative model (black lines) in all three experiments. Additionally, for the OED experiment, the OD is plotted. For Validation exp. 1 the SH simulations were performed using the simulated (black) and the measured (grey) dissolved oxygen as input.](image)

However, there are discrepancies between the SH simulation and measurements in Validation exp. 1. These differences are caused by a slightly wrong description of dissolved oxygen as the overall model uses the gas fractions of the headspace \( (y_1-y_5) \) as inputs and calculates the dissolved gases itself (for further information see Neddermeyer et al. (2015) and Sec. 3). If the directly measured, dissolved gas concentration is taken as input, measurements and simulations are much closer (see Fig. 5, grey). During the on-line OED experiment the cells were continuously exposed to high dissolved oxygen concentrations (see Fig. 3) which was beneficial for the information content of the data. These high oxygen concentrations, however, had effects on the overall process. On the one hand, discrepancies between simulations and measurements at high oxygen concentrations are notable. They occur because the \( y_{13} \)-sensor is calibrated to 20 % oxygen in the gas phase and consequently higher concentrations cannot be measured reliably. On the other hand, high oxygen leads to long term inhibition of growth. For this reason the measured OD is not reflected by the simulations (Fig. 3). Hence, a state \( (I) \) accounting for oxygen inhibition on growth was introduced into the global model of Sec. 3
\[ \dot{i} = Mo(\alpha_1, k_{O_1}, k_{O_3}). \]  
(19)

The inhibiting state \( I \) is constrained to values between 0 and 1 and affects growth. Accordingly, oxygen exposure leads to decreased growth rates
\[ \mu_{X,max} = \mu_{X,max} \cdot (1 - I), \]  
(20)

such that the simulated OD, see Fig. 5, better describes the measurements of this experiment, see Fig. 3, and past conducted cultivations (not shown) than Model 2.

7. CONCLUSION

The aim of this work was to integrate a description of SH production into an existing process model. One experiment controlled by on-line OED and two validation cultivations were run. The experimental data of the on-line OED cultivation contained enough information for modelling. After a manual modelling step, a description for SH was obtained which is in accordance with the cross-validations. By using on-line OED the data information content was increased. Therefore, only one experiment was sufficient to create and identify the SH model extension. Prospectively, the generated model will be used to maximize the SH amount in a controlled cultivation. Future work will also focus on automating the manual modelling step (see Herold and King (2014)) and integrating it into the on-line OED cycle.

REFERENCES