Benchmarking real-time monitoring strategies for ethanol production from lignocellulosic biomass

Pau Cabaneros Lopez[1], Hannah Feldman[1], Miguel Mauricio-Iglesias [2], Helena Junicke[1], Jakob Kjøbsted Huusom[1], Krist V. Gernaey[1]

[1]PROSYS Research Center, Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), Building 229, 2800 Lyngby, Denmark {pacalo, hafe, heljun, jkh, kvg @kt.dtu.dk}

[2]Department of Chemical Engineering. Universidade de Santiago de Compostela. 15782, Santiago de Compostela, Spain. miguel.mauricio@usc.es

Corresponding author: Krist V. Gernaey, kvg@kt.dtu.dk

Journal: Biomass and Bioenergy

Abstract

The goal of this paper is to review and critically assess different methods to monitor key process variables for ethanol production from lignocellulosic biomass. Because cellulose-based biofuels cannot yet compete with non-cellulosic biofuels, process control and optimization are of importance to lower the production costs. This study reviews different monitoring schemes, to indicate what the added value of real-time monitoring is for process control. Furthermore, a comparison is made on different monitoring techniques to measure the off-gas, the concentrations of dissolved components in the inlet to the process, the concentrations of dissolved components in the reactor, and the biomass concentration. Finally, soft sensor techniques and available models are discussed, to give an overview of modeling techniques that analyze data, with the aim of coupling the soft sensor predictions to the control and optimization of cellulose to ethanol fermentation. The paper ends with a discussion of future needs and developments.

Keywords

Real-time monitoring; monitoring devices; fermentation; models; soft sensors; cellulosic ethanol

1 Introduction

The monitoring of bioprocesses in real-time is a widely studied area, as real-time measurements of reactor conditions allow for a higher degree of control and process optimization than off-line monitoring [1]. In large scale biotechnology processes there is usually only a rather limited capability for real-time monitoring of the process due to lack of suitable – and affordable – monitoring techniques. Monitoring applications have been developed mainly for laboratory use [2]. There are many reports on the availability, advantages, and challenges of different monitoring techniques, but large scale monitoring in real-time with advanced sensors is rarely done. This is because there are hardly any investigations on the potential benefits of these

methods [3]. Kiviharju *et al.* [4] compared different monitoring methods based on specific requirements for biomass monitoring, providing a guide to select the appropriate method under specific conditions. A number of papers review specific monitoring devices, in which the devices are described as single entities. For instance, Marison *et al.* [5], gave an extensive review of near infrared spectroscopy (NIR), mid infrared spectroscopy (MIR), Raman spectroscopy, dielectric spectroscopy (DS), and biocalorimetry, and Marose *et al.* [6], studied in situ microscopy, NIR spectroscopy, and fluorescence spectroscopy. Nevertheless, as the performance of these methods for specific monitoring objectives was not compared, it did not provide the reader enough information to support the selection of a given alternative.

Cellulose-based biofuels are produced from biomass mainly consisting of plant material, in which sugars are fixed in structures of cellulose and hemicellulose that are intertwined with lignin [7]. The cost of cellulose-based biofuels production cannot yet compete with non-cellulosic biofuels [8], in which the carbon source comes from relatively simple and easily accessible sources such as corn or sugar beet **(FIGURE 1)**. Non-cellulosic biofuels have been produced for more than two decades [9] and are now a mature technology given the considerable experience gathered operating and building plants for non-cellulosic biofuel production. As a consequence, monitoring is essential in cellulosic biofuel production in order to ensure that the process runs at the optimal process conditions and to compensate the relative lack of process understanding of this technology [10]. One of the goals of monitoring cellulose to ethanol fermentation is to increase the profit associated with the process [11]. An increased profit can be obtained by reaching a high ethanol yield (income increase) and by running the process under non-sterile conditions (cost reduction). However, there is an increased risk of contamination when working under non-sterile conditions, which would decrease the ethanol yield, compared to a sterile process. Monitoring the process to detect contaminations is therefore of importance to be able to stop the process as soon as a contamination is detected and avoid the loss of substrate.

Another challenge in cellulosic ethanol production is overcoming the action of inhibitors. Indeed, inhibition decreases the productivity, which makes the process last longer, and thus increases the costs. Furthermore, the longer the process lasts, the higher the risk of contamination. Monitoring of inhibitory components in the feed and the reactor is therefore needed so that a real-time strategy can be applied to improve the process performance. As the introduction of novel control and optimization techniques has a cost related not only to the equipment and implementation but also to the training of operators, the benefits must be demonstrated and clearly outperform the current process as it is operated. In optimal conditions, one would desire direct measurements of all components of interest – for example, substrates, biomass, inhibitory substances, and product concentration. However, this is not an economically viable option, due to the high costs associated with installing and maintaining the equipment needed to establish the different monitoring techniques.

The contribution of this study is to assess the alternatives for real-time monitoring of fermentation and link them with industrial challenges faced during ethanol production from lignocellulosic biomass. The application of combined techniques for advanced monitoring is covered for the first time. Beyond the review made by Pohlschleidt *et al.,* [12], this study explicitly relates the monitoring equipment and combinations thereof with the specific objectives of the process, in particular for the production of cellulosic ethanol [12]. Furthermore, this study evaluates the potential benefits of the methods with a case study involving the production of cellulosic ethanol. This is a relevant case study, as the complex feed stream containing multiple carbon sources, inhibitors, and particulates needs accurate monitoring to obtain knowledge of the process characteristics **(FIGURE 1**). Furthermore, because the feed stream contains a significant amount of solid particles, the monitoring techniques need to be able to distinguish between relevant and irrelevant compounds.

The organization of this paper is such that section 2 describes the process layout of a cellulosic ethanol fermentation. Section 3 focuses on the added value of monitoring different key process variables versus the requirements for such set-ups, while section 4 discusses different sampling techniques in case of at-line sampling. In section 5, different techniques per monitoring approach, are evaluated based on the previously defined requirements. This section will go more in depth about specific measuring devices to monitor the key process variables. Section 6 and 7 discuss models and soft sensors, which can be used for optimization and control of a fermentation process. Finally, the discussion evaluates the applicability to the case study and discusses an optimal strategy for the monitoring of cellulose to ethanol fermentation.

[FIGURE 1 should be approximately here]

Figure 1: Differences between non-cellulosic and cellulosic ethanol production. Note that both, non-cellulosic and cellulosic ethanol can also be produced from other sources such as sugar beets or wood chips respectively.

2 The Cellulosic Ethanol Fermentation Process

The process to produce cellulosic ethanol typically consists of four consecutive steps: the pretreatment of the lignocellulosic material, the enzymatic hydrolysis of the pretreated material, the fermentation of the hydrolysate and the separation processes (**FIGURE 1**) [13]. In the pretreatment, the lignocellulosic fibers are broken down to smaller pieces, and exposed to increase the hydrolysis rate in the following step. Several methods are available for the pretreatment, most of them including high temperatures and pressures, and pH variations by addition of acid or base. Some conventional pretreatment methods are acid hydrolysis, steam explosion or ammonia treatment [13]. The choice of a specific pretreatment strategy will have an impact on the downstream processing, the hydrolysis and the fermentation steps, and may raise different challenges for the implementation of analytical methods to monitor the fermentation

process, which must be considered. The enzymatic hydrolysis is the step in which the fibers of lignocellulose are enzymatically hydrolyzed to release the monosaccharides. In some cases, the enzymatic hydrolysis and the fermentation are performed simultaneously (simultaneous saccharification and fermentation, SSF) and in some other cases they are done consecutively (separate hydrolysis and fermentation, SHF) [13,14]. The performance of the hydrolysis will also have an impact on the fermentation, as it determines the concentration of fermentable sugars. The system considered in this case study is the fermentation step in a separate hydrolysis and fermentation process.

The fermentation for the production of cellulose-based ethanol usually consists of a batch phase, followed by a fed-batch phase, and finally another batch phase. A fed-batch operation typically starts and ends with a batch phase[15]. In the first batch phase, the cells grow at a maximum growth rate, and the cell density increases significantly. Cell and process characteristics define this growth rate. During the fed-batch phase, a feed stream enters the reactor increasing the volume in the reactor. In the reactor, anaerobic conversion of the substrates to product and biomass takes place with a rate dependent on the microorganisms used and the process characteristics (Figure 2). During the fed-batch phase, the conversion rate is limited by the feed rate and the detoxification of inhibitors. In effect, the admissible feed rate is generally limited by the presence of inhibitors in the feed and cannot proceed faster than the capacity of the micro-organism to detoxify the medium. It is therefore important that the amount of inhibitors is monitored during the fed-batch phase as a means to maximize the feed rate and productivity. During the final batch phase the consumption, production, and growth rates are not controlled. The capacity is defined by the host organism and the process characteristics, such as pH and temperature, and the presence of inhibitory compounds. The main cost components of this fermentation process are the feedstock, utilities, and capital cost [11]. It is therefore desirable to utilize as much of the feedstock as possible for ethanol production, so a high ethanol yield is

147 required. Furthermore, to minimize the utilities cost, a high productivity is desired to minimize the
148 fermentation time.
149
150 For this case study, it is assumed that the carbon source originates from wheat straw, which
151 yields mainly glucose, xylose, furfural, 5-HMF, acetic acid, and lignin after pre-treatment and
152 enzymatic hydrolysis [11]. In the current study, it is assumed that the yeast, which is a
153 genetically engineered strain, can consume glucose and xylose simultaneously [16]. The
154 productivity of the process is mainly dependent on the xylose consumption rate, as this is the
155 rate-limiting step in mixed glucose/xylose fermentation. Furfural is a major inhibitor of yeast [17],
156 and it is therefore important to keep this concentration low throughout the fermentation process.
157 Acetic acid is also a major inhibitor, but the inhibition effect of this compound depends on the pH
158 as only the unionized (neutral) form is inhibitory. This indicates that pH control is of importance
159 for the process. As the fermentation is run without gas sparging, oxygen will be present in the
160 beginning of the fermentation. This is important to monitor, as the presence of oxygen decreases
161 the ethanol yield, as ethanol is produced under anaerobic conditions. It is therefore desired that
162 the oxygen has been consumed before the fed-batch phase starts. The most important variables
163 of cellulose to ethanol fermentation are therefore the carbon sources glucose and xylose, the
164 product ethanol, the inhibitors furfural and acetic acid, carbon dioxide and oxygen, and the pH.
165 These variables can be monitored in real-time by either direct measurement or indirect modeling
166 techniques. The monitored variables can then be used in a model for optimization and control,
167 as shown in Table 1 and Table 2, where respectively the process objectives and different risks
168 and solutions associated with cellulosic ethanol fermentation are shown.
169
170 **Table 1: Monitoring targets to achieve process objectives**

171 [TABLE 1 should be approximately here]
172
173 [FIGURE 2 should be approximately here]
174 **Figure 2: Schematic overview of a fed-batch reactor with a feed rate F_{in}. The components in italics indicate**
175 **the uncertainties in the process. These are the substrate (S), product (P), biomass (X), dissolved oxygen (DO),**

and the possible presence of a contamination. The feed rate F_{in} is known and controlled. The off-gas composition is not known, but as it is an indirect indication of the state of the process, it is not a direct uncertainty in how the fermentation behaves. The volume (V), pH, and temperature (T) are usually monitored and controlled, and therefore not uncertain.

Table 2: Overview of risks associated with cellulose to ethanol fermentation.

[TABLE 2 should be approximately here]

3 Key process variables

In this section, the added value of the monitoring of each key process variable will be evaluated in terms of what a monitoring strategy of different variables can add to the total quantity of process data that can be analyzed. Figure 2 gives an overview of the uncertain elements in cellulose to ethanol fermentation, which are shown in italics. A comparative table of the evaluation results can be found in Table 3, where each monitoring step has been assigned a number of points, depending on how much the measurements contribute to the analysis of the key variables. While Table 1 describes why components are measured, Table 3 describes how they are measured. Temperature and pH, which are standard monitoring techniques, are set at zero points. The other techniques are pointwise compared to the added value of temperature and the pH. The next few paragraphs will focus on how the table and figure are linked, and how the system is graded. The end rankings, which were reviewed by an industrial panel in ØRSTED (Denmark), with plenty of experience in operating a cellulosic ethanol demonstration plant, are a result of combining an extensive literature study, including academic research and published patents, and the authors' experience with monitoring and control. The targets addressed are monitoring the off-gas, the components dissolved in the feed stream, the components dissolved in the reactor, the biomass concentration, and detecting contaminations, such as the occurrence of lactic acid bacteria.

3.1 Temperature and pH

The most basic approaches to monitor a fermentation process are through the temperature and the pH. Most fermentation processes are run at constant pH with a relatively loose control. As carbon dioxide is produced along the fermentation, base is added to keep the pH constant. Under normal circumstances, the addition of base to the reactor at a relatively constant pace would indicate stable growth and ethanol formation. However, an abnormally large addition of base is an indication of a contamination with undesired lactic acid bacteria, as the production of lactic acid substantially acidifies the medium [18]. When a contamination is detected, the most convenient solution is to stop the process, as the substrate represents a major share of the production costs, and a contamination will take valuable carbon source away from ethanol production.

3.2 The off-gas

Measurements of the off-gas give the highest added value as a stand-alone method. It is possible to detect carbon dioxide, oxygen, and ethanol directly in the off-gas, and thus predict the concentrations in the liquid phase. This is usually done by using Henry's law, which is dependent on the process conditions, in particular temperature and, for carbon dioxide, pH. One can also indirectly monitor the growth rate, the total sugar consumption, and detect contaminations through mass balances and growth kinetics [19]. The ethanol concentration can give information on the process yield, while the process rates indicate the productivity of the process. Furthermore, monitoring the oxygen in the off-gas is important, as the presence of oxygen is unwanted in cellulose to ethanol fermentation.

3.3 The off-gas and components dissolved in the inlet

Combining off-gas measurements with measurements of the components dissolved in the inlet can give, additionally, the xylose and glucose concentrations in the liquid, as these can be estimated through mass balances and growth kinetics [20]. The off-gas provides feedback

information about the rate of consumption/growth whereas the inlet measurements give feedforward information about the actual substrate provided. This also allows better estimation of the biomass concentration, as compared to solely off-gas measurements. Another advantage is that the inhibitory components entering the reactor are directly monitored, which has the result that the feed rate can be manipulated, to maintain a low concentration of inhibitors in the reactor during the fed-batch phase.

3.4 The off-gas and components dissolved in the reactor

When combining off-gas measurements with measurements of the components dissolved in the reactor, no predictions are needed to acquire these data, and real-time information of the actual state of the process can be obtained. On the other hand, not having any measurements of the inlet is a disadvantage because characterizing the inlet is important in general control of fermentations, especially when the inlet can be a potential source of disturbances. In this strategy, such disturbances would only be measured inside the reactor. During the fed-batch phase, the only manipulated variables are the feeding rate, the addition of base or to stop the batch and start all over again. While the pH is often maintained within certain bounds (at the expense of using base, which is expensive), the feeding rate can be adjusted to keep the concentration of inhibitors inside the reactor below a threshold. In this regard, the difference between monitoring the components dissolved in the reactor or in the inlet would be that the former would allow to control the feeding rate based on actual measurements, while the latter would depend on the prediction of how fast the cell culture can detoxify the inhibitors. Also, by monitoring compounds dissolved in the reactor it would be possible to directly measure the concentration of lactic acid, which would allow to early detect contaminations by lactic acid bacteria and to stop the batch on time.

3.5 The off-gas and the biomass concentrations

Another option is combining the off-gas measurements with the monitoring of the biomass concentration. This will not yield direct concentrations of glucose, xylose, ethanol, and furfural, but with the right measuring method contaminations could be observed directly. This is the only beneficial aspect of monitoring the biomass concentration instead of the before mentioned monitoring schemes, although as will be described in section 5.3, so far no applications are available that can distinguish cells on-line in industrial scale. The effect of inhibitory compounds can be seen in the biomass activity, but there is no knowledge of the amount of inhibitors that are present. This makes control of especially the feed rate significantly more complex.

3.6 Components dissolved in the inlet and in the reactor

If off-gas measurements are not possible, one could also measure the components in the inlet and in the reactor. This does not change the added value compared to the previous two mentioned methods, but measuring components dissolved in the liquid phase is more complex than off-gas measurements. Section 5 will elaborate in more detail on these differences.

3.7 Addition of multiple monitoring methods

Increasing the number of monitoring methods to three or four increases the added value, as different measurements will add more direct data. However, it should be noted how much additional monitoring approaches contribute to the total amount of information obtained from combining hardware and software sensors, as soft sensors are often capable of analyzing what is going on in the reactor from less complex measuring methods, such as the off-gas composition measurement. Hardware sensors should be better capable of giving accurate information on the current reactor state. However, this is only true if the sensors can measure all components of interest, are accurate and not subjected to interference. Furthermore, fast response times are beneficial for fast control, but the techniques should not be too expensive. Biomass concentration measurement techniques should be able to detect contaminations and

distinguish between viable and non-viable cells to be of any extra value. Most techniques will need soft sensors for calibration and to convert the measured data into valuable information. The complexity of the calibration, maintenance and data analysis differs per technique, and this can be of importance when considering that factories are often built in remote areas, where expert knowledge will not always be available at all times. These considerations will be taken into account in section 5, where equipment is discussed.

Table 3: The added value of (combinations of) different monitoring strategies of key process variables. -: does not monitor, +: monitors indirectly, ++: monitors indirectly through different models, +++: monitors directly. Each plus counts as one point, while the points of the standard setup (pH and temperature measurements) are subtracted from the total amount of gained points for each monitoring strategy.
[TABLE 3 should be approximately here]

4 Sampling

Real-time measurements can be performed either in-line, on-line or at-line (**¡Error! No se encuentra el origen de la referencia.**) [21]. With in-line monitoring, the measurements are performed directly inside the reactor without removing or diverting the sample from the process stream. On-line and at-line measurements, in contrast, take place outside the reactor. While the sample is diverted and may be returned to the reactor (e.g. analysis through a flow cell) for on-line measurements, the sample is removed when performing at-line measurements. In order to maintain real-time measurements, on/at-line methods need to be automated for industrial applications. There is, therefore, a need for a reliable sampling technique, connected to one or multiple pre-treatment devices, and subsequently the measuring device. The pre-treatment devices often include filtration units to remove the suspended solid particles and flow systems to prepare the samples (e.g., to dilute or stain them). A promising automated pre-treatment method is cross-flow filtration, where a constant flow through a hollow fiber keeps solid particles from clogging the membrane [22,23]. This method has been used by Meschke *et al.* [22] in combination with high-performance liquid chromatography (HPLC), and by Rocha and Ferreira [23] with an amperometric biosensor. Also, the wastewater treatment sector applies cross-flow

filtration in order to remove particles from the water or retain biomass in the reactor [24,25]. Another type of automated sampling techniques which is being developed is applied in the BioScope [26]. The BioScope can be used for experimental research of microbial kinetics in a fermentation, in which rapid sampling is desired. However, so far, this technique is developed for experimental research, and not for industrial applications. Automated sampling devices combined with a sample preparation system have also been described for the application of flow cytometry [27]. A general challenge for an automated sampling system is that sterility in the reactor needs to be maintained. However, for the case of cellulosic ethanol production, this is not an issue as the reactor is operated under non-sterile conditions.

[FIGURE 3 SHOULD BE APPOXIMATELY HERE]

Figure 3. Conceptual approaches to real-time monitoring according to the guidance for industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance [21] . PAT: process analytical technology.

5 Sensors

This section will evaluate different measuring techniques for the monitoring approaches discussed in section 3.

5.1 The off-gas analyzer

Previously, an extensive review has been written on different methods to measure the off-gas composition, which is considered as a continuous measurement [19]. For this study it was chosen to only focus on techniques that can measure all gas components of interest, as combining different gas monitoring methods in tandem will be more expensive [3]. The only two techniques that can measure all components of interest, carbon dioxide, oxygen, and ethanol, are electronic noses and mass spectrometry, as these methods are capable of measuring a broad spectrum of volatile components in the off-gas. The electronic nose works as a semiconductor, where the resistance of sensors changes when exposed to volatile organic compounds (VOC) or gases. An electronic nose consists of multiple sensors with high

sensitivity, but a slow response time. Furthermore, pattern recognition algorithms are needed to analyze the obtained data. Another issue is that background gases such as water vapor can interfere with the measurements [28]. A solution for this has been proposed, where samples were dehydrated before injection into the electronic nose [29]. A significant amount of research is being conducted on the electronic nose, but the main applications are in food technology, and most of the applications are still performed on lab scale. Mass spectrometry (MS) on the other hand is a well-established method [30] capable to quantify a broad range of substances with high accuracy, typically from 100% to a few parts per million. One can choose for quadrupole MS, which is the cheaper option, or magnetic MS, which is more expensive, but also more stable and offers a higher resolution.

5.2 Components dissolved in the liquid

In Table 5, the different techniques to monitor components dissolved in the liquid for both the inlet and the reactor are compared. The techniques evaluated are in-line, on/at-line near-infrared spectroscopy (NIR), mid-infrared spectroscopy (MIR), Raman spectroscopy, UV-Vis spectroscopy, biosensors, and HPLC. Fluorescence spectroscopy is not considered in this section because the key components dissolved in the liquid (i.e. glucose, xylose, ethanol, acetic acid, lactic acid, furfural and HMF) are not fluorescent. The evaluation is based on the following eight requirements of measured components: sensitivity, accuracy, drift, calibration and data analysis, sample preparation, response time, industrial availability, and costs. In order to compare the potential of each technique, a scoring matrix is introduced which is made considering each of the previous criteria. The scoring matrix aims at reflecting the applicability and complexity of each method to provide a better understanding of the possibilities of each technique. The requirements were based on previous literature [12] and discussions with industry. An example of how the scoring matrix is done is provided for the first criterion

(measured components) in Table 4. A detailed explanation of the development of the scoring matrix for the remaining seven criteria is provided in the supplementary material.

Table 4. Scoring matrix to evaluate the capabilities of the different methods to monitor the key compounds of the cellulose to ethanol fermentation. A method capable of monitoring all the relevant compounds would receive a score of 3, while a method unable to monitor any of the compounds would receive a score of 0.

[Table 4 should be approximately here]

5.2.1 Vibrational spectroscopy

Vibrational spectroscopy (UV-Vis, NIR, MIR and Raman spectroscopy) is a group of analytical techniques that allow a fast detection of several compounds directly from the fermentation media without the need for sample preparation. The primary challenge for the application of vibrational spectroscopy to monitor cellulose to ethanol fermentation is the high content of suspended solid particles derived from lignin and biomass. These particles interfere with the light, reflecting and scattering it. This limits the implementation of vibrational transmission spectroscopy to on-line or at-line modes only, where a filtration unit is added before the spectroscopic analysis [31]. In contrast, reflectance vibrational spectroscopy (mainly attenuated total reflectance (ATR) and diffuse reflectance [32]), and backscattered Raman spectroscopy do not depend on the light transmitted through the media but on the light reflected or backscattered by the media, making these methods more suited for in-line monitoring cellulose to ethanol fermentations [33–35]. Despite the advantages of reflectance and backscattered spectroscopy, the interference between the particles and the light still entails extensive data pre-treatments and

382 results in lower accuracy and sensitivity [35,36]. For this reason, vibrational spectroscopy
383 methods performed better in the evaluation for on-line or at-line modes than for in-line modes.
384
385 Among the different vibrational spectroscopy techniques, near-infrared (NIR) spectroscopy is the
386 most mature and well-established method [14,35,37], and it has been applied to monitor a wide
387 variety of fermentations [33,38–40] including cellulose to ethanol processes [34,41,42]. Pinto *et*
388 *al.* [41] used at-line transmission NIR to monitor the concentration of glucose and ethanol during
389 cellulose to ethanol fermentation at lab-scale. Despite filtering the samples before analysis, the
390 high interference of NIR with water and the highly overlapped spectra resulted in high prediction
391 errors (6.60 g/L and 3.02 g/L for glucose and ethanol respectively). In another study, Sundvall *et*
392 *al.* [42] used an on-line NIR probe (score of 11) in a demonstration-scale cellulose-to-ethanol
393 plant (EPAB/SEKAB E-Technology, Sweden) to monitor the concentration of total sugars,
394 glucose, ethanol, and suspended solids. Despite the good correlation between the off-line and
395 the on-line samples, the reported concentration ranges were quite high (17-30 g/L and 2-40 g/L
396 for glucose and ethanol respectively) and more sensitive measurements would be needed for
397 accurate monitoring of the fermentation. Austin *et al.* [34] monitored the concentration of total
398 sugars, glucose, and ethanol in a 23 m^3 reactor using an in-line diffuse reflection probe (score of
399 11). The measurements were noisy due to the high concentration of solid particles but in
400 accordance with the off-line measured samples, giving valuable qualitative information about the
401 process endpoint. In general, NIR spectroscopy has the advantage of being a robust method
402 that can be implemented in, on or at-line, and requiring very little or no sample preparation.
403 Although it is not as sensitive and accurate as other techniques, NIR delivers qualitative
404 information that can increase the process knowledge. For these reasons, on-line and in-line NIR
405 is given a score of 11.
406

Mid-infrared spectroscopy (MIR) offers a higher accuracy and a larger number of variables to be analyzed compared to NIR [5,34,43]. Several implementations of MIR in cellulose to ethanol bioprocesses are reported in the literature [31,34,44]. Juhl *et al.* [31,44] used an at-line transmission system to monitor the concentration of glucose, lactic acid, glycerol, acetic acid, and ethanol. The samples were filtered prior to analysis in order to avoid the interactions with the solid particles. The predictions with MIR had a significant lower root mean square error of prediction (RMSEP) when compared to the ones obtained with NIR in a similar set-up (e.g., the RMSEP for glucose was 0.12% for MIR and 0.26% for NIR). In another study, Austin *et al.* [34] used an in-line attenuated total reflectance MIR (ATR-MIR) probe (score of 13) to monitor the glucose, xylose, lactic acid, acetic acid and ethanol concentration in a 23 m^3 reactor. Their results were directly compared with in-line diffuse reflectance NIR and showed that ATR-MIR had a significantly higher accuracy than NIR [34], allowing a better understanding of the dynamics of the fermentation. On-line ATR-MIR (score of 12) has also been applied to the hydrolysis step of starch-based ethanol production and brewing processes, which present similar challenges as cellulose-based ethanol production regarding suspended solid particles [45,46]. ATR-MIR has a shallow penetration depth in the sample media, making it more robust in media with suspended particles than transmission MIR. The main disadvantages of ATR-MIR are the fouling on the surface of the ATR crystal [31] and the high costs associated with the optical fibers required to transmit the signal. ATR-MIR scores higher than NIR spectroscopy (12 and 13 for on-line and in-line respectively) due to the higher sensitivity and accuracy, and due to the potential to measure lactic acid, a crucial compound to detect contaminations.

Raman spectroscopy is an attractive method foremost because there is, unlike for NIR and MIR, no water interference. Additionally, Raman spectra are better resolved and require less modeling efforts than NIR and MIR [47,48]. However, the Raman signal is relatively weak and attenuated mainly by the suspended solid particles and by the background fluorescence emitted by lignin

[48], altogether, limiting its potential for in-line monitoring. Ewanick *et al.* [49,50] used on-line Raman spectroscopy (score of 12) to measure the concentrations of glucose and ethanol in a lab-scale cellulose to ethanol fermentation (1.3 L). In order to avoid the interference with suspended solid particles, the fermentation medium was filtered prior to the fermentation. The concentration of glucose and ethanol were monitored with a prediction error of 1 g/L. Also at lab-scale, Iversen *et al.* monitored the concentration of glucose, ethanol and acetic acid using in-line Raman spectroscopy (score of 12) [47,48]. To account for the reduction of fluorescence caused by the suspended solid particles, Iversen *et al.* included an internal standard as a correction factor [51]. Despite the efforts to minimize the effect of the solid particles, their research showed that accuracy of Raman spectroscopy improves when lignin particles are removed before the measurement, which on full-scale could be achieved by using an automated sample port in combination with a filtration or sedimentation step. In spite of the potential of Raman spectroscopy as analytical technique, the expensive material and the lack of relevant industrial implementation lead to suggest a final score of 12 for both, on-line and in-line Raman spectroscopy.

UV-Vis spectroscopy is often not considered as a method for real-time monitoring of fermentations because it cannot detect many key compounds (e.g. glucose or ethanol) and because the light scattering caused by the suspended solid particles dominates the absorption process [35,52]. However, in the context of cellulose to ethanol fermentation, the technique gains special relevance because many of the inhibitors present in lignocellulosic hydrolysate including furfural, HMF or acetic acid absorb in this region [53]. Pinto *et al.* [53] used at-line UV-Vis spectroscopy to quantify the concentration of furfural and HMF from filtered samples, attaining a high sensitivity and low prediction errors (RMSEP of 0.375 g/L and 0.041 g/L for furfural and HMF respectively). UV-Vis is a useful method to quickly detect inhibitory compounds

458 and lactic acid (useful to detect contaminations), in an inexpensive manner. For this reason, UV-
459 Vis gets an overall score of 10.
460

461 ### 5.2.2 Biosensors

462 Biosensors (total score of 10) in general use enzymatic reactions to monitor concentrations of
463 specific components [54]. The way the reactions are monitored differs per type of biosensor. The
464 most widely known biosensor is the amperometric glucose sensor, which is used by diabetes
465 patients to measure glucose levels in the blood [55]. In the biosensor glucose oxidase converts
466 glucose to hydrogen peroxide (H_2O_2), which reacts with specific compounds in the sensor and
467 generates, in the case of an amperometric biosensor, a current, which is measured. Ethanol can
468 be monitored by the same principle through the use of alcohol dehydrogenase [56]. The
469 measurement of xylose can be monitored simultaneously with glucose by the YSI 2700 SELECT
470 probe (YSI Life Sciences, Yellow Springs, Ohio, USA), but sample filtration and dilution are
471 required. Concentration ranges of 0.05 g/L – 9 g/L and 0.5 g/L – 30 g/L were reported for
472 glucose and xylose, respectively (YSI Life Sciences, 2008). Amperometric sensors for the
473 detection of lactic acid have been developed and applied to monitor malolactic fermentations
474 [58]. This can be used to detect lactic acid bacteria. The measurements are fast, sensitive, and
475 have a high selectivity. However, the sensors have limited long term stability and drift is
476 encountered [3]. This happens in the time range of days to months, depending on the sensor
477 [56]. Electrochemical sensors to monitor the concentration of acetic acid in fermentations have
478 also been described in the literature [59]. There are also no reports on the measurement of
479 furfural through biosensors, but as the sensors work enzymatically, this should theoretically be
480 possible. The technique has not yet been applied on industrial scale, which forms an indication
481 that there is still considerable development work needed.

5.2.3 High-performance liquid chromatography (HPLC)

The most widely used and known method of measuring specific components is HPLC, which is commonly used as reference measurement to calibrate other monitoring methods. For at-line applications, a flow injection system that withdraws, filters and prepares the sample is required so that only particle-free liquid is analyzed by the HPLC [2]. This adds complexity to the set-up and increases its costs and operational time. Furthermore, the HPLC columns need to be washed regularly to guarantee that one obtains reliable results. In order to reduce the complexity of the set-up, it is desired to use a single chromatographic column able to analyze as many relevant compounds as possible. In the context of cellulosic ethanol production, the simultaneous quantification of sugars (glucose and xylose), ethanol, acetic acid and common inhibitors (HMF and furfural) is challenging and slow due to their different chemical properties and concentration ranges [60]. The simultaneous quantification of the previously mentioned compounds has only been reported using an Aminex HPX-87H column and requires between 40 to 55 minutes for one analysis depending on the mobile phase [60–62]. Faster analysis (up to 15 minutes) would be achieved using different columns, but it would increase the costs of the set-up and the complexity of the operation [5,22,61,63]. At-line HPLC gets a high score as an analytical tool (as it can measure all relevant compounds with high sensitivity and accuracy, and a small drift), but it is somewhat challenging to automate, requires sample preparation and has a slow response time (total score of 12).

Table 5: Overview of all the techniques discussed to monitor components in the liquid phase. Scores from 0 to +++ are given for each criterion, 0 indicating a negative effect and +++ indicating a positive effect. The costs are evaluated with scores from --- to 0, --- indicating more costly and 0 less costly. A thorough description of the scoring system is provided in the supplementary material.

[TABLE 5 should be approximately here]

5.3 The biomass

Monitoring biomass in cellulose to ethanol fermentation is a significant challenge foremost because the conventional methods used in other fermentation processes (e.g., optical density probes or infrared spectroscopy) fail at differentiating cellular biomass from the suspended solid particles and therefore are not suitable for lignocellulosic ethanol fermentations [6]. Moreover, standard methods to assess cell culture viability (e.g., methylene blue test) cannot be applied due to the dark color of the media [64]. In biomass monitoring, unlike in methods to monitor compounds in the liquid, the samples cannot be filtered prior to analysis because that would also remove the cells. In this section, different methods to monitor the biomass concentration are discussed and evaluated regarding their ability to differentiate between biomass and solid particles, to assess the cell culture viability, to detect contaminations, sample preparation, calibration, and data analysis, industrial availability and costs. An overview of the evaluation can be found in Table 6, and a detailed explanation of the scoring system is provided in the supplementary material.

5.3.1 Multi-wavelength fluorescence spectroscopy

Fluorescence spectroscopy (total score of 7) can monitor biological compounds such as NADH, tryptophan, and riboflavin [6]. These compounds are closely related to the generation of cells and can be used as indirect measurements of biomass [65–68]. Multi-wavelength fluorescence spectroscopy produces three-dimensional data sets (time, excitation spectra and emission spectra) which are analyzed using advanced chemometric methods (typically using parallel factor analysis, PARAFAC [65–67]). By using these models, it is possible to resolve the pure spectra of each fluorophore from the mixture, making multi-wavelength fluorescence more robust to changes in the composition of the media and to the background fluorescence emitted

533 by lignin [65,66,69]. In addition, similarly to other spectroscopic techniques, fluorescence
534 spectroscopy is also affected by the high content of suspended solid particles. Multi-wavelength
535 fluorescence has previously been used to monitor ethanol fermentations at lab-scale, but there
536 are no reports of utilizing fluorescence spectroscopy for cellulose-based ethanol production. The
537 BioView fluorescence spectrometer (Delta, Hørsholm, Denmark) claims to be applicable in
538 industrial settings [70] but has to our knowledge not been used for the monitoring of ethanol
539 production from lignocellulosic biomass at pilot or even larger scale.
540

541 **5.3.2 Biocalorimetry**

542 A biocalorimeter (score of 8) monitors biomass growth based on the metabolic heat, which is
543 calculated from all the heat flows concerning the reactor [71]. The main advantage of this
544 technique is that the equipment needed, mainly temperature probes and flow meters, is cheap
545 [72]. A direct relation was even found between the consumption of cooling water and the
546 metabolic heat generation in an industrial-sized bioreactor of 100 m^3, where the biomass
547 concentration could be estimated more accurately using the cooling water consumption data
548 than from elemental and electron balances [72]. In fact, as the scale of the reactor increases,
549 smaller influences such as heat loss to the environment and noise become less significant. This
550 method monitors the biomass concentration indirectly through heat balances, in a similar way as
551 it can be monitored through a carbon balance, although no distinction between cell types can be
552 made. The initial biomass concentration needs to be known to estimate the concentration over
553 time from the metabolic activity. Response times are between 1 and 2 minutes [5].

554 **5.3.3 Flow cytometry**

555 Flow cytometry (score of 8) is an at-line method to characterize and count cells through light
556 scattering and fluorescence [73]. It can monitor the biomass concentration accurately and allows
557 to distinguish between viable cells, non-viable cells, and other types of biomass [74]. Flow

cytometry is expensive, but it has been applied on large scale and many different devices are available [73]. In addition, several approaches have been developed to automate the sampling procedure, dilution, and staining of the cells via flow injection systems, thereby reducing the required labor and allowing the design of control strategies based on the physiological properties of the cell culture [75–79]. The main drawback of flow cytometry in cellulose to ethanol fermentations are the suspended solid particles, which cannot be filtered and can only be differentiated from the biomass via expensive fluorescent stains and not via light scattering. Apart from that, accuracies have been reported to be good enough up to a concentration of $2 \cdot 10^6$ cells/mL, which means that the samples will need to be diluted. As dilutions also increase the measurement error, it was observed that flow cytometry can only work well with a total concentration of up to $30 \cdot 10^6$ cells/mL [75]. The dilution steps will also increase the time needed for sample preparation. Sampling results can be obtained every 15 minutes [3].

5.3.4 Dielectric spectroscopy

Dielectric spectroscopy (score of 12), the most advantageous technique according to this study (Table 6), can monitor viable cells in-line by using an electric field at different frequencies to characterize the capacitance and conductivity of the system. The applied electric field induces the polarization of viable cells only [80,81], and this is reflected in the capacitance of the system [82]. Since polarization is only induced in viable cells, this method has no interference with gas bubbles and dead cells [83]. Dielectric spectroscopy has been applied to monitor cell viability in different fermentations with concentration ranges reported to be between 0 g/L and 200 g/L [4,5]. Furthermore, this technique has also been applied to control fermentations based on the specific growth rate [84]. Bryant *et al.,* [80] applied dielectric spectroscopy to monitor the hydrolysis of pretreated lignocellulose in a simultaneous saccharification and fermentation (SSF) process. Wang *et al.* [64] combined dielectric spectroscopy with multivariate analysis to measure

583 the viability of yeast during a fed-batch SSF. Despite the positive results, the method requires
584 extensive calibration to account for the different process parameters that affect dielectric
585 spectroscopy (e.g., suspended solids, ethanol concentration or conductivity of the media).
586 Another advantage of this technique is that it is available for industrial use, as industrial brewing
587 processes already apply dielectric spectroscopy [5].
588

589 **5.3.5 Microscopy and image analysis**

590 Microscopy combined with image analysis (score of 11) is an automatic cell counting method
591 based on the identification of individual cells from pictures taken with microscopy from
592 fermentation samples [85]. It was developed 30 years ago in the brewing industry, and it has
593 significantly developed with the recent advances in machine learning and improvements in
594 detection sensors (i.e., charge coupled devices) [85,86]. Image analysis has also been used to
595 correlate several features (e.g., cell size or cell volume) with cell viability. Donnelly *et al.* [87]
596 developed a method to predict the viability of cell cultures with the cell volume distribution and
597 used it to calculate the pitch size in industrial fermentations. Belini *et al.* [88] used in-line
598 microscopy combined with image analysis to monitor yeast growth in a lab-scale molasses-to-
599 ethanol fermentation. By using classification algorithms, they were able to differentiate between
600 yeast cells and other solid compounds present in the fermentation media (e.g., plant fibers,
601 sugar crystals or gas bubbles). If the resolution of the microscope is high enough, this method
602 can also be used to detect microbial contaminations, as suggested by Belini *et al.* [86].
603

Table 6: Overview of all the techniques discussed to monitor the biomass concentration. Scores from 0 to +++ are given for each criterion, 0 indicating a negative effect and +++ indicating a positive effect. The costs are evaluated with scores from --- to 0, --- indicating more costly and 0 less costly. A thorough description of the scoring system is provided in the supplementary material.

[TABLE 6 Should be approximately here]

6 Previous modelling efforts

The previous sections evaluated what measurements add to the extent of knowledge of cellulose to ethanol processes, and what measurement equipment is actually available for full-scale bioreactors. Models will be needed to predict the yield and productivity from the available data. The use of models is beneficial to control the process and optimize at specific points, such as the feed rate. Furthermore, it is important to model the variables that are considered as risks in Table 2, namely if there is contamination, inhibition, or presence of oxygen. These risks can be monitored directly through measurements, as described previously or indirectly through modelling. This section will look into the available models that take into account the measurements that were previously shown to be important to monitor the yield, productivity, and risks. A list of the models that have been evaluated can be found in Table 7. The models evaluated in this study are all unstructured models with simplified kinetic expressions (containing only substrate, product, and biomass), as structured models, containing synthesis rates of enzyme and intracellular metabolite production are considered too complex for routine daily use in a production environment. An interesting observation is that only one of the models described takes carbon dioxide in the form of total inorganic carbon into account [89], while the monitoring of this compound in the off-gas can relate significantly to the process characteristics. However, as cellulose to ethanol fermentation is not aerated or sparged, it is relatively difficult to monitor the gas flow rate out of the reactor and relate it to the dissolved CO_2 concentration. Therefore, it would be necessary to compare it with previous fermentations, and generate a relation based on experience. All evaluated models contain inhibition functions, often with Monod type kinetics. All studied models take product inhibition into account. Substrate inhibition and furfural inhibition,

631 which was previously mentioned to be a strong inhibitor (see section 1), are also often modelled.
632 In fact, Navarro *et al.* [90] only used furfural as state variable to describe the process. Monitoring
633 the inhibitory compounds is important in a cellulose to ethanol fermentation, as the amount of
634 inhibitory compounds in the reactor can be controlled through the feed rate. With the exception
635 of the model published by Navarro *et al.* [90], all models contain at least the substrate and
636 product as state variables, while the cell biomass is often present. These state variables are
637 important to model the yield and productivity of the fermentation. Furthermore, sudden changes
638 in yield and productivity can indicate the presence of inhibitory compounds or a contamination.
639 In the case of Hanly and Henson [91], Palmqvist *et al.* [92], and Mauricio-Iglesias *et al.* [89],
640 other major components present in the reactor are also included. In general, the more
641 components are added in a model, the more accurate balance equations can be applied, and
642 the more time will be spent on model development as well. Balance equations use relationships
643 that are derived from theory or experiments to estimate states from measurements [93].

644
645 **Table 7: Overview of the process models researched in this study.**

646
647
648 [TABLE 7 should be approximately here]

649 ## 7 Soft sensors

650 Soft sensors are important for data analysis, process control, and process optimization. Data
651 driven soft sensors are used to calibrate and interpret the data from measuring devices
652 (hardware sensors), and to perform fault detection, from which deviating activity in the system
653 can be found [101]. The most used soft sensors for this purpose are based on principal
654 component analysis (PCA) decomposition and partial least squares (PLS) regression [35–
655 37,67,101,102], which are applicable to linear relationships. For non-linear relationships, artificial
656 neural networks (ANN) are often used. A challenge of ANN's is that they tend to get stuck in
657 local minima [101]. For this reason ANN's need a significant amount of calibration data and

tuning [103]. Soft sensors based on chemometrics, PCA for exploratory analysis and PLS regression are a mature technology and currently the most frequently applied tools in industry for monitoring fermentation processes. Furthermore, these soft sensors comply with the process analytical technology (PAT) initiative by the American Food and Drug Administration. These methods are very efficient for quality surveillance in order to detect if a particular process is following the intended production recipe. Hence, these tools provide insight into the current behavior of the principal components, but do not provide information which can be directly coupled with a first principles process model in order to predict or optimize future behavior.

Model-driven soft sensors on the other hand are applied to estimate variables from other monitored variables, to work as a backup for when hardware sensors fail, and to perform fault detection. The model-based soft sensors rely on first principles process models (balance equations for mass and energy as well as constitutive equations for e.g. reactions and transport) and on an algorithm that reconciles the available measurements with predictions by the model. This is also known as a filter or a state observer. Examples of such algorithms are Luenberger or Kalman filters or asymptotic observers [104,105].

Soft sensor technology has been utilized in the bulk chemical industry for decades but industrial applications in the biochemical industry are recent and under development [106,107]. The reasons for later utilization in e.g. fermentations can be several, among others, process-model mismatch, nonlinear dynamics, noisy measurements and that the development of state estimators of sufficient quality is troublesome for many industrial fermentation processes. Much of the research in state estimation focuses on ensuring the long-term (asymptotic) convergence of the developed algorithms. However, as the biochemical industry is dominated by batch and fed-batch processes (time limited), the ability of many popular state estimators to monitor bioprocesses is somewhat limited [104]. Furthermore, the instrumentation can be insufficient in

order to have enough information available for the estimation. In industrial fermentation applications, spectroscopic methods dominate to a high degree, and these are not as straightforward to couple to the estimation scheme as direct measurements of e.g. temperature, pressure or pH, as is the case in classic chemical processes.

According to Luttmann *et al.* [108] soft sensors are mainly applied to determine the rate of oxygen consumption and carbon dioxide production, as well as the relationship between the two, the respiratory quotient (RQ) [109], but the number of applications at industrial scale is low. Furthermore, the RQ is not applicable to cellulose to ethanol fermentation, as there is no oxygen consumption. Mauricio-Iglesias *et al.* [89] explored the use of the continuous-discrete extended Kalman filter to estimate biomass, furfural and acetic acid by measuring glucose, xylose, ethanol and pH. The *in silico* results were promising as the estimation was reasonably good even in conditions of simulated contamination by lactic acid bacteria. So, to our opinion this is certainly a route that could be exploited further, for example for more standardized comparison of sensors, monitoring and control strategies *in silico*. Here, inspiration can be found in the wastewater treatment field, where benchmarking efforts aiming at *in silico* comparison of control strategies have been ongoing for almost 20 years now [110].

8 Discussion

This paper aimed to identify key variables to monitor in cellulose to ethanol fermentation. As cellulosic ethanol cannot yet compete with non-cellulosic ethanol regarding process economy, it is important to reduce the costs, which are mainly associated with utilities, substrate, biomass, and capital costs. Hence, an increase in profit can be achieved by increasing the yield and productivity as well as by running the fermentation in non-sterile conditions. However, to reach these objectives and to maintain the highest possible yield and productivity, monitoring and control are needed.

The current real-time monitoring methods used in the non-cellulosic ethanol industry (as in many other low-value, high-volume processes) consist of secondary measurements such as pH, turbidity, CO_2 in the offgas or temperature [111]. Although these measurements provide valuable information about the process, they do not directly relate to the state of the system, making it challenging to establish advanced control strategies. Similar to fermentation processes for non-cellulosic ethanol production, cellulosic ethanol fermentations are subject to fluctuations in the substrate composition that change the dynamics of the fermentation. Therefore, these processes would benefit from more advanced monitoring methods that can generate data that can be used for adjusting the operation of the process. When compared with non-cellulosic ethanol production processes, cellulose-based ethanol production is a more complicated process involving more phenomena such as inhibition, or a mixed substrate. In consequence, the monitoring methods typically used for the production of non-cellulosic ethanol fail in cellulosic ethanol production processes at providing real-time information, which would otherwise be useful for implementing control strategies. Additionally, advanced monitoring methods are required to improve the performance of cellulosic ethanol fermentations.

Models are needed to control and optimize the process. For reliable and accurate models, measurements are necessary. In the reactor, fast response times are also desired, as the process characteristics will constantly change. As the response times needed differ per process, it would be of value to investigate the actual response times needed in different processes. Automatic controllers will also need real-time measurements as input. However, real-time monitoring of cellulosic ethanol fermentation is complex and troublesome due to the presence of suspended solid particles and the complexity of the fermentation matrix, while mixed substrate consumption and the presence of inhibitory compounds will further increase the complexity of the model. The choice of a suitable monitoring strategy depends on the model and the specific

equipment requirements. Quantitative data (e.g., on accuracy, costs or concentration ranges) is desired for making objective decisions for control and optimization, but also to support and justify the choice of specific equipment. The collection of quantitative data is somewhat troublesome, as data from different sources either contradicted one another, as this could be dependent on the manufacturer and the specific reactor conditions, or was not available at all. The most reliable option, but also the most expensive and time-consuming one, is to test measurement equipment under practical conditions on a cellulose to ethanol fermentation plant and to make the results available to a broader public. It is not very realistic to assume that one organization can perform such tests alone. Therefore, it would be obvious to set up a consortium of stakeholders such that the test work – and the costs related to it – can be shared. It should also be in the interest of the measurement equipment manufacturers if an objective evaluation of the potential of the different measurement techniques would be available.

In Section 3 it was determined that the off-gas is the easiest to monitor in real-time because it avoids the interferences with the suspended solid particles. Also, off-gas analyzers that can detect oxygen, carbon dioxide or ethanol are often available in the industry. With this information, the controller can increase or decrease the batch times, and adjust the feeding rate based on the productivity of the fermentation. In section 5.1 it was evaluated that magnetic mass spectrometers are the most advantageous because they can evaluate a broad range of substances in a wide range of concentrations. Although the off-gas can give insight into the reactor characteristics, the evaluation of several models showed that the gas components such as carbon dioxide are hardly considered, while the ethanol stripping is not considered at all. Modeling the carbon dioxide concentration could potentially be useful in detecting uncommon behaviors in the system, as a deviation from mass balances might indicate that something is wrong in the process. However, it was shown that most models mainly consider substrates, products, biomass, and inhibitors, which can only be monitored in the liquid phase, and

761 predictions based on off-gas only would not be as accurate. Contaminations by lactic acid
762 bacteria can also be potentially monitored through mass balances and kinetics, but this option
763 has not been thoroughly explored yet.
764
765 Monitoring the compounds dissolved in the liquid phase allows measuring the concentration of
766 substrates, products, and inhibitors directly, giving a more clear picture of the actual state of the
767 system. This information permits a better estimation of the biomass concentration and a control
768 of the fermentation time and the feeding rate based on the actual concentrations of substrates
769 and inhibitors. The main challenges are the interference with the suspended solid particles and
770 the complex fermentation matrix of cellulose-to-ethanol fermentations. In this context, the choice
771 of a monitoring method for the compounds in the liquid phase is not obvious and becomes a
772 trade-off between the quality of the measured data, the speed of the analysis and the ease of
773 the operation. On the one side of the spectrum, HPLC (score 12) is an excellent and well-known
774 analytical tool with very high sensitivity and accuracy, but somewhat slow and complex to use. In
775 addition to measuring substrates, products, and inhibitors, HPLC can measure the concentration
776 of lactic acid, allowing the direct detection of contaminations by the LAB. On the other side of the
777 spectrum, different in-line spectroscopies are easy to implement and have a high measuring
778 frequency, but the measurements are noisy and less accurate. The accuracy of the
779 spectroscopic methods improves when a filtration unit is added before the analysis, but this also
780 increases the complexity of the operation. Among the different spectroscopic methods, in-line
781 ATM-MIR is evaluated with the highest score (total score of 13) because it can measure the
782 substrates, products, and lactic acid and it has been tested in demonstration scale cellulosic-
783 ethanol fermentation. UV-Vis spectroscopy (score of 11) is also an interesting option as a fast
784 on-line method to measure the concentration of inhibitors in the inlet or in the reactor.
785 Biosensors (score of 10) obtained the lowest score, as they are sensitive and accurate methods
786 to measure with high frequency the concentrations of glucose, xylose, ethanol or lactic acid, but

787 they cannot be implemented in-line and require clear and diluted samples. The main challenges
788 are that the sensors have limited long-term stability and will encounter drift, while there are also
789 no furfural or 5-HMF biosensors available yet.
790
791 For biomass monitoring (section 5.3) dielectric spectroscopy was the most beneficial (total score
792 of 12, Table 6) since it can differentiate cells from other suspended solid particles, it is able to
793 detect viable cells, and it has been shown to work on lab-scale in cultures with lignocellulosic
794 material. Although contaminations cannot be detected with this method, this study has shown
795 other indirect methods to detect contaminations, such as the observation of a sudden increase in
796 base addition to indicate lactic acid production from lactic acid bacteria. Unlike dielectric
797 spectroscopy, flow cytometry (score of 8) can directly detect contaminations by lactic acid
798 bacteria. However, flow cytometry is an expensive technique difficult to implement for on/at-line
799 monitoring. 2D fluorescence and bio-calorimetry (scores of 7 and 8 respectively) are indirect
800 methods to measure biomass, but they cannot detect contaminations. Finally, microscopy and
801 image analysis (score of 11) appears as a method with the potential to measure biomass since it
802 can differentiate cells from particles, viable and non-viable cells and also contamination.
803 However, this method still needs further development.
804
805 When deciding on extending the monitoring scheme, one should first gain insight into what
806 strategies will be the most useful for control and optimization. This will depend on how the
807 process is modeled, but also on the type of process and the specific conditions applied. Off-gas
808 measurements by mass spectrometry were found to be the most important in cellulosic ethanol
809 fermentation, followed by the addition of the monitoring of the inlet. If it is assumed that the inlet
810 composition is not dynamic, a delay in measurements is not an issue at all. HPLC is therefore
811 suitable and reliable to monitor the inlet under this assumption. These two measurement
812 techniques combined with kinetic models can generate data needed for control. Monitoring

dissolved components and biomass in the reactor is of importance for fault detection and optimization, as this will need accurate data on the state of the reactor. A simulation study [89] including the addition of in situ measurements to estimate state variables, showed that the prediction error decreased when the reactor holdup, substrates, product, and pH were monitored with a sampling interval of 240 minutes. Interestingly, when excluding the pH from these measurements, the prediction error increased. Although total inorganic carbon was a state variable in this study, no off-gas monitoring was performed. It is recommended that a similar study is performed when a monitoring scheme is considered, to give a better insight into the added value of a specific monitoring scheme linked with a specific model.

Considering that cellulosic ethanol production processes have now reached a stage of maturity which allows operating a process at demonstration scale or even full-scale, it would be obvious to allocate some more resources to investigating the potential of further improving the operation of such installations by adding more on-line monitoring and control. In order to reach a situation where real-time control is put in operation on the basis of on-line measured data, our suggestion is, therefore, to focus on a detailed evaluation of the most promising monitoring methods that have been highlighted in this manuscript. As mentioned before, an *in-silico* approach could be useful here, inspired by the work on benchmarking of control strategies that has been done in the wastewater field [110].

9 Conclusion

Cellulose to ethanol fermentation is a complex process that is often operated far from its optimal conditions. In consequence, the implementation of advanced monitoring and control strategies is necessary to improve the process efficiency compared to non-cellulosic ethanol production processes.

838 Lignocellulosic waste includes a wide variety of materials ranging from wood chips to different
839 kinds of straw. These materials have very different properties and compositions, and affect the
840 fermentation differently. Likewise, the influence of the available process alternatives must be
841 carefully considered before deciding on the most adequate monitoring and control system. In
842 this review, different monitoring schemes and methods for cellulosic ethanol fermentation have
843 been reviewed. The fermentation of wheat straw hydrolysate in an SHF process was used as a
844 case study. However, the challenges described for this case study (e.g., high concentration of
845 suspended solids, the complex fermentation matrix or the presence of inhibitors) are common to
846 other substrates or process configurations.
847
848 The risk of contamination by lactic acid bacteria, the inhibition by furfural and acetic acid and the
849 presence of oxygen in the fermenter were identified as the major threats for the cellulose to
850 ethanol fermentation. Among the different monitoring schemes reviewed in this article, it was
851 found that monitoring the off-gas, the inlet, and the liquid phase of the reactor would add
852 significant value to the currently used monitoring methods (i.e., pH and temperature). Among all
853 the methods available to monitor off-gas, only electronic noses and mass spectrometry are
854 considered in this review as the two techniques able to simultaneously detect all the compounds
855 of interest (glucose, xylose and ethanol). Despite the significant amount of research done in
856 electronic noses, mass spectrometry is a more mature and implemented technology. To monitor
857 the inlet and the liquid phase in the reactor, in-line ATR-MID spectroscopy was deemed as the
858 most advantageous technique because it is able detect simultaneously most of the compounds
859 of interest, it does not require sample preparation and it is not affected too much by the high
860 concentrations of suspended solids. Monitoring the biomass was also found to be valuable. The
861 most suited analytical instrument for real-time monitoring of the biomass is dielectric
862 spectroscopy. However, the developments in microscopy and in image analysis make the
863 technology attractive, especially for its potential to detect contaminations. It was found that quite

some quantitative data on measuring devices is missing in the literature and that the available data can vary considerably depending on the manufacturer of a device, and on the reactor conditions. Research on the objective comparison of different devices in specific case studies or applications would be of interest, especially to companies aiming at selecting a device for a specific application.

Another important step is to investigate in more detail how the monitoring can contribute specifically to the control and optimization of industrial applications, and the most viable option there seems to use an in-silico approach to save on costs.

Acknowledgments

This work was partially financed by the European Regional Development Fund (ERDF) and Region Zealand (Denmark) through the BIOPRO-SMV project. Furthermore, the work received funding from Innovation Fund Denmark (BIOPRO2 strategic research center, project number 4105-00020B). We would like to acknowledge critical comments by Jesper Bryde-Jacobsen (BIOPRO), and Laila Thirup, Michael Elleskov, Pia Jørgensen, Flemming Mathiesen and Remus Mihail Prunescu from Ørsted. This project has also been supported partially by the EUDP project 'Demonstration of 2G ethanol in full scale, MEC' (Jr. no. 64015-0642). Finally, we wish to acknowledge the support obtained from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 713683 (COFUNDfellowsDTU) and from the Danish Council for Independent Research in the frame of the DFF FTP research project GREENLOGIC (grant agreement number 7017-00175A). Miguel Mauricio-Iglesias belongs to the Galician Competitive Research Group GRC2013-032 and the CRETUS strategic partnership (AGRUP2015/02), co-funded by FEDER (EU).

10 References

[1] C. Svendsen, T. Skov, F.W.J. van den Berg, Monitoring fermentation processes using in-process measurements of different orders, J. Chem. Technol. Biotechnol. 90 (2015) 244–254. doi:10.1002/jctb.4483.

[2] K. Schügerl, Progress in monitoring, modeling and control of bioprocesses during the last 20 years., J. Biotechnol. 85 (2001) 149–173. doi:10.1016/S0168-1656(00)00361-8.

[3] B. Sonnleitner, Automated Measurement and Monitoring of Bioprocesses: Key Elements of the M3C Strategy, Adv. Biochem. Eng. Biotechnol. 132 (2013) 1–33. doi:10.1007/10.

[4] K. Kiviharju, K. Salonen, U. Moilanen, T. Eerikäinen, Biomass measurement online: The performance of in situ measurements and software sensors, J. Ind. Microbiol. Biotechnol. 35 (2008) 657–665. doi:10.1007/s10295-008-0346-5.

[5] I. Marison, S. Hennessy, R. Foley, M. Schuler, S. Sivaprakasam, B. Freeland, The Choice of Suitable Online Analytical Techniques and Data Processing for Monitoring of Bioprocesses, Adv. Biochem. Eng. Biotechnol. 132 (2013) 249–280. doi:10.1007/10.

[6] S. Marose, C. Lindemann, R. Ulber, T. Scheper, Optical sensor systems for bioprocess monitoring, Trends in Biotec. 17 (1999) 30–34. doi:10.1007/s00216-003-1930-1.

[7] K.L. Kadam, E.C. Rydholm, J.D. McMillan, Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass., Biotechnol. Prog. 20 (2004) 698–705. doi:10.1021/bp034316x.

[8] S. Macrelli, J. Mogensen, G. Zacchi, Techno-economic evaluation of 2nd generation bioethanol production from sugar cane bagasse and leaves integrated with the sugar-based ethanol process, Biotechnol. Biofuels. 5 (2012) 22. doi:10.1186/1754-6834-5-22.

[9] W.E. Tyner, The US Ethanol and Biofuels Boom: Its Origins, Current Status, and Future Prospects, Bioscience. 58 (2008) 646. doi:10.1641/B580718.

913 [10] J.D. Stephen, W.E. Mabee, J.N. Saddler, Will second-generation ethanol be able to
914 compete with first-generation ethanol? Opportunities for cost reduction, Biofuels, Bioprod.
915 Biorefining. 6 (2012) 159–176. doi:10.1002/bbb.331.
916 [11] J. Larsen, M.Ø. Haven, L. Thirup, Inbicon makes lignocellulosic ethanol a commercial
917 reality, Biomass and Bioenergy. 46 (2012) 36–45. doi:10.1016/j.biombioe.2012.03.033.
918 [12] M. Pohlschleidt, S. Charaniya, C. Bork, M. Jenzsch, T.L. Noetzel, A. Luebbert, Bioprocess
919 and Fermentation Monitoring, in: Encycl. Ind. Biotechnol., 2013: pp. 1471–1491.
920 [13] C.M. Drapcho, N.P. Nhuan, T.H. Walker, Biofuels Engineering Process Technology,
921 2008. doi:10.1036/0071487492.
922 [14] A.E. Lantz, K. V. Gernaey, C.J. Franzen, L. Olsson, Online monitoring of fermentation
923 processes in lignocelluloses-to-bioalcohol production, in: K. Waldron (Ed.), Bioalcohol
924 Prod. Biochem. Convers. Lignocellul. Biomass., Woodhead Publishing LTD., Oxford,
925 2010: pp. 315–39.
926 [15] H.S. Shin, H.C. Lim, Cell-mass maximization in fed-batch cultures, Bioprocess Biosyst.
927 Eng. 29 (2006) 335–347. doi:10.1007/s00449-006-0082-z.
928 [16] M.S. Krishnan, N.W. Ho, G.T. Tsao, Fermentation kinetics of ethanol production from
929 glucose and xylose by recombinant Saccharomyces 1400(pLNH33)., Appl. Biochem.
930 Biotechnol. 77–79 (1999) 373–388. doi:10.1385/ABAB:78:1-3:373.
931 [17] E. Palmqvist, B. Hahn-Hagerdal, Fermentation of lignocellulosic hydrolyzates. II: inhibitors
932 and mechanisms of inhibition, Bioresour. Technol. 74 (2000) 25–33.
933 [18] P. Oliva-Neto, F. Yokoya, Evaluation of bacterial contamination in a fed-batch alcoholic
934 fermentation process, World J. Microbiol. Biotechnol. 10 (1994) 697–699.
935 doi:10.1007/BF00327963.
936 [19] D. Pollard, J. Christensen, Vent Gas Analysis, in: M.C. Flickinger (Ed.), Encycl. Ind.

Biotechnol. Bioprocess, Biosep. Cell Technol., John Wiley & Sons, Inc., 2010: pp. 1–15.

[20] N. Leksawasdi, E.L. Joachimsthal, P.L. Rogers, Mathematical modelling of ethanol production from glucose / xylose mixtures by recombinant Zymomonas mobilis, Biotechnol. Lett. 23 (2001) 1087–1093.

[21] U.D. of H. and H. FDA, Guidance for Industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacuring, and Quality Assurance, 2004. doi:http://www.fda.gov/CDER/guidance/6419fnl.pdf.

[22] J. Meschke, H. Bennemann, H. Herbst, S. Dormeier, D.C. Hempel, On-line HPLC-measurement and control of substrate in a continuously operated biological tankreactor, Bioprocess Eng. 3 (1988) 151–157.

[23] I. Rocha, E.C. Ferreira, On-line simultaneous monitoring of glucose and acetate with FIA during high cell density fermentation of recombinant E. coli, Anal. Chim. Acta. 462 (2002) 293–304. doi:10.1016/S0003-2670(02)00347-1.

[24] E.B. Muller, A.H. Stouthamer, H.W. Verseveld, D.H. Eikelboom, Aerobic domestic waste water treatment in a pilot plant with complete sludge retention by cross flow filtration, Water Res. 29 (1995) 1179–1189.

[25] R. Bai, H.F. Leow, Microfiltration of activated sludge wastewater-The effect of system operation parameters, Sep. Purif. Technol. 29 (2002) 189–198. doi:10.1016/S1383-5866(02)00075-8.

[26] D. Visser, G. a. Van Zuylen, J.C. Van Dam, A. Oudshoorn, M.R. Eman, C. Ras, W.M. Van Gulik, J. Frank, G.W.K. Van Dedem, J.J. Heijnen, Rapid sampling for analysis of in vivo kinetics using the BioScope: A system for continuous-pulse experiments, Biotechnol. Bioeng. 79 (2002) 674–681. doi:10.1002/bit.10328.

[27] J. Kacmar, F. Srienc, Dynamics of single cell property distributions in Chinese hamster

961 ovary cell cultures monitored and controlled with automated flow cytometry, J. Biotechnol.
962 120 (2005) 410–420. doi:10.1016/j.jbiotec.2005.06.031.

963 [28] F. Röck, N. Barsan, U. Weimar, Electronic nose: current status and future trends, Chem.
964 Rev. 108 (2008) 705–725.

965 [29] M. Calderon-Santoyo, P.U. Bautista-rosales, G. Luna-solano, C. Ghommidh, J.A.
966 Ragazzo-sanchez, Monitoring of Lactic Fermentation with a Coupling Electronic Nose and
967 Gas Chromatography, Eng. 2013 (2013) 13–19.

968 [30] E. de Hoffmann, Mass Spectrometry, in: Kirk-Othmer Encycl. Chem. Technol., John Wiley
969 & Sons, Inc., 2005.

970 [31] H. Juhl, O.W. Hansen, Method and device for monitoring of bioalcohol liquor prduction.,
971 WO 2009/121423 A1, 2009.

972 [32] Z.M. Khoshhesab, Infrared Spectroscopy - Materials Science, Engineering and
973 Technology, Infrared Spectrosc. - Mater. Sci. Eng. Technol. (2012) 234–244.
974 doi:10.5772/37180.

975 [33] N. Petersen, P. Ödman, A.E. Cervera Padrell, S. Stocks, A.E. Lantz, K. V. Gernaey, In
976 situ near infrared spectroscopy for analyte-specific monitoring of glucose and ammonium
977 in Streptomyces coelicolor fermentations, Biotechnol. Prog. 26 (2010) 263–271.
978 doi:10.1002/btpr.288.

979 [34] G. Austin, E.J. Becker, C. Beckstrom, G. Djordjevic, I. Dobson, H. Mason, Method NIR
980 and MIR for ethanol and other compounds.pdf, WO 2015/095255 A1, 2015.

981 [35] N.D. Lourenço, J. a. Lopes, C.F. Almeida, M.C. Sarraguça, H.M. Pinheiro, Bioreactor
982 monitoring with spectroscopy and chemometrics: A review, Anal. Bioanal. Chem. 404
983 (2012) 1211–1237. doi:10.1007/s00216-012-6073-9.

984 [36] A.E. Cervera, N. Petersen, A.E. Lantz, A. Larsen, K. V. Gernaey, Application of near-

infrared spectroscopy for monitoring and control of cell culture and fermentation, Biotechnol. Prog. 25 (2009) 1561–1581. doi:10.1002/btpr.280.

[37] M. Scarff, S.A. Arnold, L.M. Harvey, B. McNeil, Near infrared spectroscopy for bioprocess monitoring and control: current status and future trends., Crit. Rev. Biotechnol. 26 (2006) 17–39. doi:10.1080/07388550500513677.

[38] J. Crowley, S.A. Arnold, N. Wood, L.M. Harvey, B. McNeil, Monitoring a high cell density recombinant Pichia pastoris fed-batch bioprocess using transmission and reflectance near infrared spectroscopy, Enzyme Microb. Technol. 36 (2005) 621–628. doi:10.1016/j.enzmictec.2003.12.016.

[39] M. Blanco, A.C. Peinado, J. Mas, Analytical monitoring of alcoholic fermentation using NIR spectroscopy, Biotechnol. Bioeng. 88 (2004) 536–542. doi:10.1002/bit.20214.

[40] P. Luoma, A. Golabgir, M. Brandstetter, J. Kasberger, C. Herwig, Workflow for multi-analyte bioprocess monitoring demonstrated on inline NIR spectroscopy of P. chrysogenum fermentation, Anal. Bioanal. Chem. 409 (2017) 797–805. doi:10.1007/s00216-016-9918-9.

[41] A.S.S. Pinto, S.C. Pereira, M.P.A. Ribeiro, C.S. Farinas, Monitoring of the cellulosic ethanol fermentation process by near-infrared spectroscopy, Bioresour. Technol. 203 (2016) 334–340. doi:10.1016/j.biortech.2015.12.069.

[42] E. Sundvall, T. Der Meulen, NIR measurements in production of a target chemical from cellulose, WO 2012/066042 A1, 2012.

[43] S. Tosi, M. Rossi, E. Tamburini, G. Vaccari, A. Amaretti, D. Matteuzzi, Assessment of In-Line Near-Infrared Spectroscopy for Continuous Monitoring of Fermentation Processes, Biotechnol. Prog. 19 (2003) 1816–1821. doi:10.1021/bp034101n.

[44] H. Juhl, P.W. Hansen, Infrared monitoring of bioalcohol production, WO 2009/121416 A1,

2009.

[45] E. Hoffmann Petersen, A.K. Kunov-Kruse, Method for online monitoring of mashing processes using infrared spectroscopy, WO 2015/155353 A1, 2015.

[46] M. Mosher, In-line detection of chemical compounds in beer., Wo 2017/218039 A1, 2017.

[47] J.A. Iversen, R.W. Berg, B.K. Ahring, Quantitative monitoring of yeast fermentation using Raman spectroscopy, Anal. Bioanal. Chem. 406 (2014) 4911–4919. doi:10.1007/s00216-014-7897-2.

[48] J.A. Iversen, B.K. Ahring, Monitoring lignocellulosic bioethanol production processes using Raman spectroscopy, Bioresour. Technol. 172 (2014) 112–120. doi:10.1016/j.biortech.2014.08.068.

[49] S.M. Ewanick, W.J. Thompson, B.J. Marquardt, R. Bura, Real-time understanding of lignocellulosic bioethanol fermentation by Raman spectroscopy, Biotechnol. Biofuels. 6 (2013) 1–8. doi:10.1186/1754-6834-6-28.

[50] S.M. Ewanick, E. Schmitt, R. Gustafson, R. Bura, Use Raman Spectroscopy for continuous monitoring and control of lignocellulosic biorefinery processe., Pure Appl. Chem. 86 (2014) 867–879. doi:10.1515/pac-2013-1022.

[51] A. Picard, I. Daniel, G. Montagnac, P. Oger, In situ monitoring by quantitative Raman spectroscopy of alcoholic fermentation by Saccharomyces cerevisiae under high pressure, Extremophiles. 11 (2007) 445–452. doi:10.1007/s00792-006-0054-x.

[52] J. Roberts, The Use of UV-Vis Spectroscopy in Bioprocess and Fermentation Monitoring, Fermentation. 4 (2018) 18. doi:10.3390/fermentation4010018.

[53] A.S.S. Pinto, M.P.A. Ribeiro, C.S. Farinas, Fast spectroscopic monitoring of inhibitors in the 2G ethanol process, Bioresour. Technol. 250 (2018) 148–154. doi:10.1016/j.biortech.2017.11.033.

[54] K. Hantelmann, M. Kollecker, D. Hüll, B. Hitzmann, T. Scheper, Two-dimensional fluorescence spectroscopy: A novel approach for controlling fed-batch cultivations, J. Biotechnol. 121 (2006) 410–417. doi:10.1016/j.jbiotec.2005.07.016.

[55] S. Borgmann, A. Schulte, S. Neugebauer, W. Schuhmann, Amperometric biosensors, in: R.C. Alkire, D.M. Kolb, J. Lipkowski (Eds.), Adv. Electrochem. Sci. Eng., WILEY-VCH Verlag GmbH C& Co. KGaA, Weinheim, 2011: pp. 1–83. doi:10.1016/0168-1656(90)90029-B.

[56] S. Piermarini, G. Volpe, M. Esti, M. Simonetti, G. Palleschi, Real time monitoring of alcoholic fermentation with low-cost amperometric biosensors, Food Chem. 127 (2011) 749–754. doi:10.1016/j.foodchem.2011.01.008.

[57] YSI Life Sciences, Application note - Rapid Measurement of Xylose & Glucose Monitoring Corn Stover Fermentation in Bioethanol Production, Yellow Springs, Ohio, USA, 2008.

[58] A. Sannini, D. Albanese, F. Malvano, A. Crescitelli, M. Di Matteo, An amperometric biosensor for the determination of lactic acid during malolactic fermentation, Chem. Eng. Trans. 44 (2015) 283–288. doi:10.3303/CET1544048.

[59] A.L. Ndiaye, S. Delile, J. Brunet, C. Varenne, A. Pauly, Electrochemical sensors based on screen-printed electrodes: The use of phthalocyanine derivatives for application in VFA Detection, Biosensors. 6 (2016). doi:10.3390/bios6030046.

[60] J.P. Yuan, F. Chen, Simultaneous separation and determination of sugars, ascorbic acid and furanic compounds by HPLC - Dual detection, Food Chem. 64 (1999) 423–427. doi:10.1016/S0308-8146(98)00091-0.

[61] M. Soleimani, L. Tabil, Simultaneous Quantification of Carbohydrates, Alcohols, and Toxic Components in a Bio-Based Medium Using Dual-Detection HPLC Analysis, Am. J. Anal. Chem. 04 (2013) 265–272. doi:10.4236/ajac.2013.45033.

[62] X. Liu, N. Ai, H. Zhang, M. Lu, D. Ji, F. Yu, J. Ji, Quantification of glucose, xylose, arabinose, furfural, and HMF in corncob hydrolysate by HPLC-PDA-ELSD, Carbohydr. Res. 353 (2012) 111–114. doi:10.1016/j.carres.2012.03.029.

[63] C.J. Scarlata, D.A. Hyman, Development and validation of a fast high pressure liquid chromatography method for the analysis of lignocellulosic biomass hydrolysis and fermentation products, J. Chromatogr. A. 1217 (2010) 2082–2087. doi:10.1016/j.chroma.2010.01.061.

[64] R. Wang, Bioprocess development for biochemical conversion of lignocellulose, Chalmers University of Technology, 2017.

[65] A. Surribas, J.M. Amigo, J. Coello, J.L. Montesinos, F. Valero, S. Maspoch, Parallel factor analysis combined with PLS regression applied to the on-line monitoring of Pichia pastoris cultures, Anal. Bioanal. Chem. 385 (2006) 1281–1288. doi:10.1007/s00216-006-0355-z.

[66] J.M. Amigo, A. Surribas, J. Coello, J.L. Montesinos, S. Maspoch, F. Valero, On-line parallel factor analysis. A step forward in the monitoring of bioprocesses in real time, Chemom. Intell. Lab. Syst. 92 (2008) 44–52. doi:10.1016/j.chemolab.2007.12.001.

[67] P. Ödman, C.L. Johansen, L. Olsson, K. V. Gernaey, A.E. Lantz, On-line estimation of biomass, glucose and ethanol in Saccharomyces cerevisiae cultivations using in-situ multi-wavelength fluorescence and software sensors, J. Biotechnol. 144 (2009) 102–112. doi:10.1016/j.jbiotec.2009.08.018.

[68] Martin B. Haack, A.E. Lantz, P.P. Mortensen, L. Olsson, Chemometric Analysis of In-Line Multi-Wavelength Fluorescence Measurements Obtained During Cultivations With a Lipase Producing Aspergillus oryzae Strain, Biotechnol. Bio. 96 (2006) 904–913. doi:10.1002/bit.21170.

[69] J.S. Lupoi, S. Singh, B. a. Simmons, R.J. Henry, Assessment of Lignocellulosic Biomass

Using Analytical Spectroscopy: An Evolution to High-Throughput Techniques, Bioenergy Res. 7 (2014) 1–23. doi:10.1007/s12155-013-9352-1.

[70] S. Marose, C. Lindemann, T. Scheper, Two-dimensional fluorescence spectroscopy: A new tool for on-line bioprocess monitoring, Biotechnol. Prog. 14 (1998) 63–74. doi:10.1021/bp970124o.

[71] D. Voisard, P. Pugeaud, a. R. Kumar, K. Jenny, K. Jayaraman, I.W. Marison, U. Von Stockar, Development of a large-scale biocalorimeter to monitor and control bioprocesses, Biotechnol. Bioeng. 80 (2002) 125–138. doi:10.1002/bit.10351.

[72] M. Türker, Development of biocalorimetry as a technique for process monitoring and control in technical scale fermentations, Thermochim. Acta. 419 (2004) 73–81. doi:10.1016/j.tca.2004.01.036.

[73] M. Díaz, M. Herrero, L. a. García, C. Quirós, Application of flow cytometry to industrial microbial bioprocesses, Biochem. Eng. J. 48 (2010) 385–407. doi:10.1016/j.bej.2009.07.013.

[74] T.L. Da Silva, J.C. Roseiro, A. Reis, Applications and perspectives of multi-parameter flow cytometry to microbial biofuels production processes, Trends Biotechnol. 30 (2012) 225–231. doi:10.1016/j.tibtech.2011.11.0053.

[75] N.R. Abu-Absi, A. Zamamiri, J. Kacmar, S.J. Balogh, F. Srienc, Automated flow cytometry for acquisition of time-dependent population data., Cytometry. A. 51 (2003) 87–96. doi:10.1002/cyto.a.10016.

[76] J. Kacmar, A. Gilbert, J. Cockrell, F. Srienc, The cytostat: A new way to study cell physiology in a precisely defined environment, J. Biotechnol. 126 (2006) 163–172. doi:10.1016/j.jbiotec.2006.04.015.

[77] A. Brognaux, S. Han, S.J. Sørensen, F. Lebeau, P. Thonart, F. Delvigne, A low-cost,

multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors, Microb. Cell Fact. 12 (2013). doi:10.1186/1475-2859-12-100.

[78] F. Delvigne, J. Baert, S. Gofflot, A. Lejeune, S. Telek, T. Johanson, A.E. Lantz, Dynamic single-cell analysis of Saccharomyces cerevisiae under process perturbation: Comparison of different methods for monitoring the intensity of population heterogeneity, J. Chem. Technol. Biotechnol. 90 (2015) 314–323. doi:10.1002/jctb.4430.

[79] J. Baert, R. Kinet, A. Brognaux, A. Delepierre, S. Telek, S.J. Sørensen, L. Riber, P. Fickers, F. Delvigne, Phenotypic variability in bioprocessing conditions can be tracked on the basis of on-line flow cytometry and fits to a scaling law, Biotechnol. J. 10 (2015) 1316–1325. doi:10.1002/biot.201400537.

[80] D.N. Bryant, S.M. Morris, D. Leemans, S. a. Fish, S. Taylor, J. Carvell, R.W. Todd, D. Logan, M. Lee, N. Garcia, A. Ellis, J. a. Gallagher, Modelling real-time simultaneous saccharification and fermentation of lignocellulosic biomass and organic acid accumulation using dielectric spectroscopy, Bioresour. Technol. 102 (2011) 9675–9682. doi:10.1016/j.biortech.2011.07.084.

[81] J.-P. Pitkänen, H. Turkia, H. Sirén, A. Rissanen, Literature review of on-line bioprocess monitoring Report, (2010) Deliverable D.8.3. http://www.vtt.fi/sites/nanobe/nanobe_publications.jsp.

[82] G.D. Austin, R.W.J. Watson, T. D´amore, Studies of On-Line Viable Yeast Biomass with a Capacitance Biomass Monitor, Biotechnol. Bioeng. 43 (1994) 337–341. doi:10.1016/0022-3999(86)90018-8.

[83] C. Justice, a. Brix, D. Freimark, M. Kraume, P. Pfromm, B. Eichenmueller, P. Czermak, Process control in cell culture technology using dielectric spectroscopy, Biotechnol. Adv. 29 (2011) 391–401. doi:10.1016/j.biotechadv.2011.03.002.

[84] L. Habegger, K. Rodrigues Crespo, M. Dabros, Preventing Overflow Metabolism in Crabtree-Positive Microorganisms through On-Line Monitoring and Control of Fed-Batch Fermentations, Fermentation. 4 (2018) 79. doi:10.3390/fermentation4030079.

[85] T. Höpfner, A. Bluma, G. Rudolph, P. Lindner, T. Scheper, A review of non-invasive optical-based image analysis systems for continuous bioprocess monitoring, Bioprocess Biosyst. Eng. 33 (2010) 247–256. doi:10.1007/s00449-009-0319-8.

[86] V.L. Belini, P. Wiedemann, H. Suhr, In situ microscopy: A perspective for industrial bioethanol production monitoring, J. Microbiol. Methods. 93 (2013) 224–232. doi:10.1016/j.mimet.2013.03.009.

[87] D. Donnelly, G. Cahill, Methods of calculating amounts of yeast for use in a process, GB2341610A, 2000.

[88] V.L. Belini, G.A.P. Caurin, P. Wiedemann, H. Suhr, Yeast fermentation of sugarcane for ethanol production: Can it be monitored by using in situ microscopy?, Brazilian J. Chem. Eng. 34 (2017) 949–959. doi:10.1590/0104-6632.2017034420160162.

[89] M. Mauricio-Iglesias, K. V. Gernaey, J.K. Huusom, State Estimation in Fermentation of Lignocellulosic Ethanol. Focus on the Use of Ph Measurements, in: 12th Int. Symp. Process Syst. Eng. 25th Eur. Symp. Comput. Aided Process Eng. Copenhagen, Denmark., Copenhagen, Denmark, 2015.

[90] A.R. Navarro, Effects of furfural on ethanol fermentation bySaccharomyces cerevisiae: Mathematical models, Curr. Microbiol. 29 (1994) 87–90. doi:10.1007/BF01575753.

[91] T.J. Hanly, M. a. Henson, Dynamic model-based analysis of furfural and HMF detoxification by pure and mixed batch cultures of S. cerevisiae and S. stipitis, Biotechnol. Bioeng. 111 (2014) 272–284. doi:10.1002/bit.25101.

[92] E. Palmqvist, J. Almeida, B. Hahn-Hagerdal, Influence of furfural on anerobic glycolytic

kinetics of Saccharomyces cerevisiae in batch culture., Biotechnol. Bioeng. 62 (1999) 447–454.

[93] C. Venkateswarlu, Advances in monitoring and state estimation of bioreactors, J. Sci. Ind. Res. (India). 63 (2004) 491–498.

[94] J. Zhang, X. Shao, O. V. Townsend, L.R. Lynd, Simultaneous saccharification and co-fermentation of paper sludge to ethanol by Saccharomyces cerevisiae RWB222 - Part I: Kinetic modeling and parameters, Biotechnol. Bioeng. 104 (2009) 920–931. doi:10.1002/bit.22464.

[95] J.H.. Luong, Ethanol Inhibition in Alcohol Ferment, Biotechnol. Bioeng. XXVII (1985) 280–285.

[96] M. Starzak, L. Kryzstek, L. Nowicki, H. Michalski, Macroapproach kinetics of ethanol fermentation by Saccharomyces cerevisiae: experimental studies and mathematical modelling, Chem. Eng. J. Biochem. Eng. J. 54 (1994) 221–240.

[97] M. Phisalaphong, N. Srirattana, W. Tanthapanichakoon, Mathematical modeling to investigate temperature effect on kinetic parameters of ethanol fermentation, Biochem. Eng. J. 28 (2006) 36–43. doi:10.1016/j.bej.2005.08.039.

[98] D. Pinelli, Y.R. a González-Vara, D. Matteuzzi, F. Magelli, Assessment of kinetic models for the production of L- and D-lactic acid isomers by Lactobacillus casei DMS 20011 and Lactobacillus coryniformis DMS 20004 in continuous fermentation, J. Ferment. Bioeng. 83 (1997) 209–212. doi:10.1016/S0922-338X(97)83586-6.

[99] A. Athmanathan., M. Sedlak, N.W.Y. Ho, N.S. Mosier, Effect of product inhibition on xylose fermentation to ethanol by saccharomyces cerevisae 424A (LNH-ST), Biol. Eng. 3 (2010) 111–124.

[100] R. Wang, R. Koppram, L. Olsson, C.J. Franzén, Kinetic modeling of multi-feed

simultaneous saccharification and co-fermentation of pretreated birch to ethanol, Bioresour. Technol. 172 (2014) 303–311. doi:10.1016/j.biortech.2014.09.028.

[101] P. Kadlec, B. Gabrys, S. Strandt, Data-driven Soft Sensors in the process industry, Comput. Chem. Eng. 33 (2009) 795–814. doi:10.1016/j.compchemeng.2008.12.012.

[102] P. Roychoudhury, L.M. Harvey, B. McNeil, The potential of mid infrared spectroscopy (MIRS) for real time bioprocess monitoring, Anal. Chim. Acta. 571 (2006) 159–166. doi:10.1016/j.aca.2006.04.086.

[103] C.F. Mandenius, Design of monitoring and sensor systems for bioprocesses by biomechatronic methodology, Chem. Eng. Technol. 35 (2012) 1412–1420. doi:10.1002/ceat.201100553.

[104] D. Dochain, State and parameter estimation in chemical and biochemical processes: A tutorial, J. Process Control. 13 (2003) 801–818. doi:10.1016/S0959-1524(03)00026-X.

[105] a. C. a Veloso, I. Rocha, E.C. Ferreira, Monitoring of fed-batch E. coli fermentations with software sensors, Bioprocess Biosyst. Eng. 32 (2009) 381–388. doi:10.1007/s00449-008-0257-x.

[106] L. Olsson, U. Schulze, J. Nielsen, On-line bioprocess monitoring - An academic discipline or an industrial tool?, TrAC - Trends Anal. Chem. 17 (1998) 88–95. doi:10.1016/S0165-9936(97)00125-8.

[107] C. Komives, R.S. Parker, Bioreactor state estimation and control, Curr. Opin. Biotechnol. 14 (2003) 468–474. doi:10.1016/j.copbio.2003.09.001.

[108] R. Luttmann, D.G. Bracewell, G. Cornelissen, K. V. Gernaey, J. Glassey, V.C. Hass, C. Kaiser, C. Preusse, G. Striedner, C.F. Mandenius, Soft sensors in bioprocessing: A status report and recommendations, Biotechnol. J. 7 (2012) 1040–1048. doi:10.1002/biot.201100506.

[109] H. Sundström, S.O. Enfors, Software sensors for fermentation processes, Bioprocess Biosyst. Eng. 31 (2008) 145–152. doi:10.1007/s00449-007-0157-5.

[110] K.V. Gernaey, U. Jeppssson, P.A. Vanrolleghem, J.B. Copp, Benchmarking of Control Strategies for Wastewater Treatment Plants e, IWA Publishing, London, UK, 2014.

[111] W.E. Herrera, R. Maciel Filho, Development of a Monitoring Hybrid System for Bioethanol Production, Icheap-11 11th Int. Conf. Chem. Process Eng. Pts 1-4. 32 (2013) 943–948. doi:10.3303/CET1332158.

11 Supplementary material

11.1 Scoring method for the evaluation of the discussed methods to monitor the dissolved components

All methods were evaluated based on the following eight criteria: measured compounds, sensitivity, accuracy, drift, calibration and data analysis, sample preparation, response time, industrial implementation and costs.

The scores for measured compounds were based on the capabilities of each method to monitor key compounds of the cellulose to ethanol fermentation (Table S 1). A method capable of monitoring all the relevant compounds would receive a score of 3, whilst a method able to monitor none of the compounds would receive a score of 0. Methods able to monitor glucose

Table S 1. Scores given based on the capabilities to measure relevant compounds in the liquid phase.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors*	At-line HPLC
Glucose	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Xylose	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Ethanol	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Acetic acid	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lactic acid	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Furfural	No	No	No	No	No	No	Yes	No	Yes
HMF	No	No	No	No	No	No	Yes	No	Yes
Total score	**1**	**1**	**2**	**2**	**2**	**2**	**1**	**2**	**3**

Accuracy and sensitivity are evaluated based on the values found in the literature and discussed in Section 0 (Table S 2).

Table S 2. Scores given based on the sensitivity and accuracy of each method.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
Sensitivity	**1**	**1**	**2**	**2**	**2**	**2**	**1**	**2**	**3**
Accuracy	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**2**	**3**

Drift is evaluated based on the deviation of the measurements over time. All methods start with a maximum score of 3. Long-term deviations result in the subtraction of 1 point. Drift between and within batches results in the subtraction of 1 and 2 points, respectively (Table S 3).

Table S 3. Scores given based on the basis of information collected about drift of each method.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
Within batch	No	No	No	No	No	No	No	Yes	No
Between batches	Yes	Yes	Yes	Yes	No	No	No	No	No
Long-term deviations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drift	**1**	**1**	**1**	**1**	**2**	**2**	**2**	**0**	**2**

Scores related to the calibration and data analysis are based on two criteria: the complexity of calibration methods and the pre-processing requirements of each type of data. Univariate methods are the simplest ones and receive a score of 3, multivariate methods receive a score of 2 and multiway methods a score of 1. Preprocessing requirements are classified into P1 (including basic pre-processing techniques such as base-line correction or mean centering) and P2 (including P1 and additional methods to correct for other disturbances). A method requiring a pre-processing of type P1 or P2 would receive -1 or -2 points in their final scores, respectively (Table S 4).

Table S 4. Scores given to each method according to the required calibration methods and data analysis. P1 includes basic pre-processing techniques such as base-line correction or mean centering. P2 includes P1 and additional methods to correct for other disturbances.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
Univariate	No	No	No	No	No	No	No	Yes	Yes
Multivariate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Multiway	No	No	No	No	No	No	No	No	No
Pre-process	P1	P2	P1	P2	P1	P2	P1	P1	P1
Total score	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**2**	**2**

The sample preparation is evaluated based on the number of steps required prior to analysis. A method requiring no sample preparation (in-line methods) would receive a score of 3, whilst methods requiring 1, 2, or 3 steps, would receive a score of 2, 1 or 0, respectively (Table S 5).

Table S 5. Scores assigned to each method according to the sampling preparation requirements.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
Filtration	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes
Dilution	No	No	No	No	No	No	No	Yes	Yes
Derivation[1]	No	No	No	No	No	No	No	No	Yes
Total score	**2**	**3**	**2**	**3**	**2**	**3**	**2**	**1**	**0**

[1] Derivation may include sample staining, or

The sampling frequency is divided into methods able to deliver almost real-time information (< 5 min), which receive a score of 3, methods with a delay of less than one hour (receiving a score between 2 if they need less than 20 minutes and 1 if they need more) and methods with a delay greater than one hour (receiving a score of 0) (Table S 6)

Table S 6. Scores given to each method according to sample frequency.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
< 5 min	No	Yes	No	Yes	No	Yes	No	No	No
< 1 hour	Yes	No	Yes	No	Yes	No	Yes	Yes	No
> 1 hour	No	No	No	No	No	No	No	No	Yes
Total score	**2**	**3**	**2**	**3**	**2**	**3**	**2**	**1**	**0**

The evaluation of industrial implementation has been based on an extensive review of papers and patents. Industrial implementation refers to any fermentation process and it is not limited to cellulose to ethanol fermentations. Methods not implemented at industrial scale or that are rarely used would receive 0 and 1 point respectively, and methods commonly used at industrial scale would receive 2 points. Methods tested in large scale cellulose to ethanol fermentations would receive an additional point (Table S 7).

The scores regarding costs are divided into operational and investment costs and they are compared relatively to each other. A score of -3 is given to the most expensive equipment and a score of 0 is given to the cheapest one. The final score results from the rounded up average between the operational and the investment costs (Table S 8).

Table S 7. Industrial implementation.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
None	-	-	-	-	-	+	+	+	-
Rarely used	+	+	+	+	+	-	-	-	+
Commonly used	-	-	-	-	-	-	-	-	-
Tested in large scale 2G ethanol	+	+	+	+	-	-	-	-	-

Total score	**2**	**2**	**2**	**2**	**1**	**0**	**0**	**0**	**1**

Table S 8. Scores of each method related to the investment and operation costs.

	On-line NIR	In-line NIR	On-line MIR	In-line MIR	On-line Raman	In-line Raman	On-line UV-Vis	Biosensors	At-line HPLC
Operation	-1	0	-1	0	-1	0	-1	-1	-2
Investment	-2	-2	-3	-3	-3	-3	-1	0	-2
Total score	**-1**	**-1**	**-2**	**-1**	**-2**	**-1**	**-1**	**0**	**-2**

11.2 Scoring method used to evaluate the discussed methods to monitor biomass according to the different evaluation criteria

Each method is given 3 points if they are able to detect the corresponding feature (cells/particles, viable/dead or contaminations. The final score is obtained from the sum of each individual score.

Table S 9. Scores based on the capabilities to differentiate cells and solid particles, to assess the viability of the cell culture and to detect contaminations.

	OD probes	IR Spectroscopy	2D fluorescence	Bio-calorimetry	Flow cytometry	Dielectric spectroscopy	Image analysis
Cells/particles	0	0	3	3	3	3	3
Viable/dead	0	0	0	0	3	3	3
Contaminations	0	0	0	0	3	0	3
Total score	**0**	**0**	**3**	**3**	**9**	**6**	**9**

The sample preparation is evaluated based on the number of steps required prior to the analysis. A method requiring no sample preparation (in-line methods) would receive a score of 3, whilst methods requiring dilution, derivation or both, will receive between 0 and 2 points (Table S 10).

Table S 10. Scores given to each method according to the sample preparation requirements.

	OD probes	IR Spectroscopy	2D fluorescence	Bio-calorimetry	Flow cytometry	Dielectric spectroscopy	Image analysis
Dilution	Yes	No	No	No	Yes	No	Yes
Derivation	No	No	No	No	Yes	No	No
Total score	**2**	**3**	**3**	**3**	**0**	**3**	**2**

Scores related to the calibration and data analysis are based on two criteria: the complexity of calibration methods and the pre-processing requirements of each type of data. Univariate methods are the simplest ones and receive a score of 3, multivariate receive a score of 2 and multiway methods and non-linear machine learning a score of 1. Preprocessing requirements are classified into P1 (including basic pre-processing techniques such as base-line correction or mean centering) and P2 (including P1 and additional methods to correct for other disturbances). A method requiring a pre-processing of type P1 or P2 would receive -1 or -2 points in their final scores, respectively (Table S 11).

Table S 11. Scores given to each method according to the calibration and data analysis requirements. P1 includes basic pre-processing techniques such as base-line correction or mean centering. P2 includes P1 and additional methods to correct for other disturbances.

	OD probes	IR Spectroscopy	2D fluorescence	Bio-calorimetry	Flow cytometry	Dielectric spectroscopy	Image analysis
Univariate	Yes	No	No	Yes	Yes	No	No
Multivariate	No	Yes	No	No	No	Yes	No
Multiway/Non-linear machine learning	No	No	Yes	No	No	No	Yes
Pretreatment	No	P1	No	No	P1	P1	P1
Total score	**3**	**1**	**1**	**3**	**2**	**1**	**0**

The evaluation of industrial implementation has been based on an extensive review of papers and patents. Industrial implementation refers to any fermentation process and it is not limited to

cellulose to ethanol fermentations. Methods not implemented at industrial scale or that are rarely used would receive 0 and 1 point respectively, and methods commonly used at industrial scale would receive 2 points. Methods tested in large scale cellulose to ethanol fermentations would receive an additional point (Table S 12).

The scores regarding costs are divided into operational and investment costs and they are compared relative to each other. A score of -3 is given to the most expensive equipment and a score of 0 is given to the cheapest one. The final score results from the rounded up average between the operational and the investment costs (Table S 13).

Table S 12. Scores given to each method according to the industrial availability.

	OD probes	IR Spectroscopy	2D fluorescence	Bio-calorimetry	Flow cytometry	Dielectric spectroscopy	Image analysis
None	-	-	-	+	+	-	-
Rarely used	-	+	+	-	-	+	+
Commonly used	+	-	-	-	-	-	-
Tested in large scale cellulose-to ethanol	-	-	-	-	-	+	-
Total score	**2**	**1**	**1**	**0**	**0**	**2**	**1**

Table S 13. Scores given to each method according to operational and investment costs.

	OD probes	IR Spectroscopy	2D fluorescence	Bio-calorimetry	Flow cytometry	Dielectric spectroscopy	Image analysis
Operation	0	0	0	0	-3	0	-1
Investment	-1	-2	-2	-2	-3	-1	-1
Total score	**0**	**-1**	**-1**	**-1**	**-3**	**0**	**-1**

1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356