



# Article Off-Gas-Based Soft Sensor for Real-Time Monitoring of Biomass and Metabolism in Chinese Hamster Ovary Cell Continuous Processes in Single-Use Bioreactors

Tobias Wallocha \* D and Oliver Popp

Roche Diagnostics GmbH, Pharma Research and Early Development, 82377 Penzberg, Germany; oliver.popp@roche.com

\* Correspondence: tobias.wallocha@roche.com; Tel.: +49-8856-6013298

Abstract: In mammalian cell culture, especially in pharmaceutical manufacturing and research, biomass and metabolic monitoring are mandatory for various cell culture process steps to develop and, finally, control bioprocesses. As a common measure for biomass, the viable cell density (VCD) or the viable cell volume (VCV) is widely used. This study highlights, for the first time, the advantages of using VCV instead of VCD as a biomass depiction in combination with an oxygen-uptake- rate (OUR)-based soft sensor for real-time biomass estimation and process control in single-use bioreactor (SUBs) continuous processes with Chinese hamster ovary (CHO) cell lines. We investigated a series of 14 technically similar continuous SUB processes, where the same process conditions but different expressing CHO cell lines were used, with respect to biomass growth and oxygen demand to calibrate our model. In addition, we analyzed the key metabolism of the CHO cells in SUB perfusion processes by exometabolomic approaches, highlighting the importance of cell-specific substrate and metabolite consumption and production rate qS analysis to identify distinct metabolic phases. Cellspecific rates for classical mammalian cell culture key substrates and metabolites in CHO perfusion processes showed a good correlation to qOUR, yet, unexpectedly, not for qGluc. Here, we present the soft-sensoring methodology we developed for qPyr to allow for the real-time approximation of cellular metabolism and usage for subsequent, in-depth process monitoring, characterization and optimization.

**Keywords:** process analytical technologies (PAT); off-gas analytic; real-time monitoring; viable cell biomass; perfusion process; continuous process; single-use bioreactor (SUB); oxygen uptake rate (OUR); soft sensor

# 1. Introduction

Chinese hamster ovary cells (CHO) represent the backbone of commercial and research expression hosts used for the manufacture of monoclonal antibodies (mAB) for therapeutic purposes. In the last decades, a variety of different production strategies and processes have been developed to ensure high yields and product quality as well as operational efficiency and reproducibility [1]. In order to cope with these demands, a smart synergy of flexible, state-of-the-art bioreactor systems and up-to-date process analytical techniques (PAT) is essential and highly recommended by the FDA [2]. Single-use bioreactor (SUB) systems have gained tremendous interest in the biopharmaceutical industry as these bioreactors and their peripheral solutions can remarkably elevate the efficiency and, of course, the flexibility of modern manufacturing processes [3]. SUB systems are currently an often used alternative and cover a broad range of different applications despite lacking many conventional engineering parameters or availability of compatible single-use sensors [4,5]. In the context of modern bioprocess development and optimization, the online monitoring of crucial key performance indicators (KPIs) is necessary to ensure high process and product quality [6]. Therefore, monitoring sensors that can be applied to SUB systems and



Citation: Wallocha, T.; Popp, O. Off-Gas-Based Soft Sensor for Real-Time Monitoring of Biomass and Metabolism in Chinese Hamster Ovary Cell Continuous Processes in Single-Use Bioreactors. *Processes* 2021, 9, 2073. https://doi.org/10.3390/ pr9112073

Academic Editor: Philippe Bogaerts

Received: 21 October 2021 Accepted: 16 November 2021 Published: 19 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). deliver high-grade, real-time information is a significant need, especially for one of the most valuable indicators, the biomass concentration [7,8]. Even if common sensor systems, e.g., hard-type sensors that have been adjusted to match single-use bioreactor requirements, are available [9], there is still a challenge to measure the biomass concentration online [10]. As hard-type sensor probes need to fulfill a variety of prerequisites, offline-based methods, such as image analysis with trypan blue staining and subsequent cell counting, lack limited amount of samples, are time consuming and require sample taking [11]. An appropriate alternative for online measuring of biomass is soft sensors. Soft sensors constitute the interoperation of hard-type sensors that can be implemented, or peripherally attached, to the bioreactor system in numerous ways using a software-based model as estimator [12,13]. Utilization of different soft-sensor approaches for online determination of biomass has been reported in many valuable contributions focusing on radio frequency impedance, Raman spectroscopy or off-gas analytic techniques [10,11,14–20]. Online biomass estimation via off-gas measurement generally relies on the oxygen uptake rate (OUR) as this variable is one of the best indicators for cell physiological activity and correlates very well with metabolic turnover rates and the concentration of viable biomass.

The knowledge of the OUR as a metabolic marker allows a deeper understanding of intrinsic physiological performance of the biomass and can be merged with other process variables to create meaningful new information in the sense of soft sensing [17,21–25]. The online measure of the OUR can be achieved by the global mass balance (GMB) approach which is easy to implement in SUB systems and enables a disturbance-free measure of the OUR. In perfusion and continuous processes, process interventions or other perturbations that might affect the steady-state mode are undesirable [24]. Only the knowledge of volumetric flow, composition of the gas in the inlet and outlet and the fermenter volume is required, constituting an advantage when compared to liquid-phase measurements, such as the dynamic method [23]. Perfusion processes require an optimized perfusion strategy to allow VCD and cellular biomass generation and continuous product formation and, at the same time, avoid unintentionally high economical demands of perfusion media. The exchange of unconditioned media is highly dependent upon the applied perfusion rate D, which ensures a sufficient supply of substrates to keep, in the best case, the substrate concentrations in a steady state. In parallel, the constant removal of conditioned media preserves the level of known metabolites, such as lactate, ammonia and metabolic and abiotic break-down products of amino acids, nucleotides and lipoid metabolism [26], as well as unknown cytotoxic/cytostatic metabolites at uncritical levels. It has been shown previously that the cell-specific nutrient uptake rates in perfused cell cultures correlated with the cell-specific growth rate for hybridomas and recombinant CHO cell lines [27]. This knowledge strengthens the goal to analyze and use cell- and biomass-specific rates and further metabolic KPIs for process development and optimization. In this work, we present a real-time off-gas-based biomass soft sensor that can be applied for the perfusion-based biomass growth phase of continuous processes. We used a proof-of-concept data set of 14 similar SUB continuous fermentation processes including 14 different mAB expressing CHO cell lines for model calibration. We aimed for this heterologous set-up of cell lines in order to cover a broad spectrum of different metabolically active cells as the oxygen demand between cell lines can differ significantly [24]. The soft sensor consists of two different models to predict the biomass in terms of viable cell density (VCD) and viable cell volume concentration (VCV) using a multilinear regression approach. The OUR was measured as a major input variable for both using the GBM technique and observed noise distortions were minimized by data preprocessing to improve model accuracy. Model prediction quality assessment was done by RMSE as well as MAPE and MdAPE calculation, enabling an in-depth analysis of errors and their distribution. Real-time biomass prediction was then applied on three different and unknown cell lines to the prediction models by utilization of two moving average methods to remove the OUR signal noise. Furthermore, we elaborate on the question of how biomass can be described most properly in modern bioprocesses. VCD, as a commonly used measure, lacks the information of the cellular volume and refers solely to the number of cells. Our utilization of VCV as a measure for biomass delivers more information, taking into account cell volume, which can lead to more precise correlations with the OUR [14]. In addition, cell size can have a direct impact on oxygen demand, leading to higher oxygen requirements from larger cells compared to smaller cells, reflecting a positive correlation between OUR and cell size [22]. We also highlight the advantage of the shown off-gas-based biomass soft sensor in SUB continuous processes and illustrate how the biomass can be described best when VCD or VCV are applied as descriptive measures. Process variables, such as specific oxygen uptake rate per single cell or per viable cell volume, can raise a different picture when the volume per cell is not constant.

# 2. Materials and Methods

# 2.1. Cell Lines

For this study, 17 different in-house-generated Chinese hamster ovary clonal cell lines (CHO-K1), engineered to produce 14 different mAbs, were used (Table 1). All clones were cultivated using a proprietary, chemically-defined (CD), serum-free, in-house base and perfusion medium. In general, cells can also be cultivated in commercially chemical-defined media, such as CD-CHO (Thermo Fisher Scientific, Waltham, MA, USA).

**Table 1.** Overview of all processes with their relation to the expressed antibody and clone that were used to create the biomass prediction models and for validation.

No.	Process	Expressed Antibody	Clone	Purpose
1	P-01	A-01	C-01	Training
2	P-02	A-02	C-02	Training
3	P-03	A-03	C-03	Training
4	P-04	A-04	C-04	Training
5	P-05	A-01	C-05	Training
6	P-06	A-05	C-06	Training
7	P-07	A-06	C-07	Training
8	P-08	A-07	C-08	Training
9	P-09	A-05	C-09	Training
10	P-10	A-08	C-10	Training
11	P-11	A-09	C-11	Training
12	P-12	A-05	C-12	Training
13	P-13	A-10	C-13	Training
14	P-14	A-11	C-14	Training
15	P-15	A-12	C-15	Validation
16	P-16	A-13	C-16	Validation
17	P-17	A-14	C-17	Validation

# 2.2. Cell Cultivation and in Process Control

Cells were thawed in a shake flask and maintained in a humidified shaking incubator (Multitron Cell, Infors AG, Headoffice, Switzerland) at 36.5 °C under 7% (v/v) carbon dioxide (CO<sub>2</sub>), applying a constant shaking rate and relative humidity of 70%. Cell passage took place every 3–4 days for scale-up purposes. After 10 days, cells were transferred into a wave-mixed SUB (Biostat<sup>®</sup> RM, Sartorius Stedim Biotech GmbH, Göttingen, Germany) for a 4 day long inoculation train as the batch phase for further cell propagation. During this step, the temperature was controlled at 36.5 °C and device-internal optical probes were used to control pH at 7.00 and dissolve oxygen (DO) to 30% saturation by gassing with a mixture of process air, nitrogen (N<sub>2</sub>), carbon dioxide (CO<sub>2</sub>) and pure oxygen (O<sub>2</sub>). Rocking motion was held constant at 15 rocks/min at an angle of 9 degrees.

After the inoculation train process was finished, cells were transferred to a stirred SUB (HyPerforma<sup>™</sup> SUB, Thermo Fisher Scientific, Waltham, MA, USA) with an appropriate seeding cell density depending on the doubling time in the preceding inoculation train process. Temperature was controlled to 36.5 °C and stirrer speed was set constant at 140 rpm. DO concentration was measured using a stainless steel optical probe (VisiFerm<sup>™</sup>

DO ECS 225, Hamilton, Switzerland) and controlled to 30% saturation analogous to the wave-mixed process. pH monitoring was done by pH probe (Inpro<sup>®</sup> 3253/225/PT100, Mettler Toledo, Columbus, OH, USA) and regulated by  $CO_2$  gassing and 1 M sodium carbonate addition to pH 7.00.

Cell retention was enabled using a hollow fiber module (KrosFlo<sup>®</sup> MBT<sup>®</sup>, Repligen, Waltham, MA, USA) with 0.2  $\mu$ m pore size. Perfusion was started 24 h after inoculation and modified stepwise according to the following protocol: 24 h after inoculation, perfusion mode was started with normalized fermentation volumes per day (vvd<sub>n</sub>) of 1 and increased to 2 vvd<sub>n</sub> after another 24 h. Further increase was done after 72 h to a vvd<sub>n</sub> of 3 until the last raise was applied 120 h after inoculation up to a vvd<sub>n</sub> of 4.55 towards process end. During perfusion mode, the fermentation volume was kept constant by weight-controlled addition of fresh perfusion media and no cell bleed took place. After the dynamic state with altering perfusion rates, the steady-state process would start with a constant normalized perfusion rate of 4.55 vvd<sub>n</sub> with parallel cell bleed to keep biomass concentration and product titer stable while product yield rises (Figure 1A). The biomass soft sensor presented in this work is proposed for automatically controlling cell bleeding during the steady-state perfusion process.

Offline samples were drawn at least once per day using a sterile syringe (Omnifix Luer Lock Solo, B. Braun Melsungen AG, München, Germany) and aliquoted for further analytic purposes.

## 2.3. Oxygen Balancing and OUR Calculation

To quantify the OUR, the well-known global mass balance approach is used as shown in Equation (1):

$$\frac{dc_{O2}}{dt} = OTR(t) - OUR(t)$$
(1)

The rate of oxygen that is transferred from the gas phase into the liquid phase (OTR) is influenced by several factors, such as the fluid-side oxygen mass transfer coefficient ( $k_La$ ), maximum possible oxygen solubility ( $c_{o2}^*$ ) and the current dissolved oxygen concentration ( $c_{O2,L}$ ). Therefore, Equation (1) can be written as:

$$\frac{dc_{O2}}{dt} = k_L a(t) \cdot (c_{O2}^*(t) - c_{O2,L}(t)) - OUR(t)$$
(2)

Since the oxygen saturation during all fermentation processes in this work was controlled to 30%, steady-state conditions can be assumed, thus OTR equals OUR, leading the temporal change of soluble oxygen concentration to 0. Within this condition, the OUR calculation is possible using a sensitive off-gas analyzer and the further application of a mass balance approach that is based on the mass of oxygen that enters (O<sub>2/in</sub>) and leaves (O<sub>2/out</sub>) the bioreactor system [5]. Since the temperature of the gas mixture (consisting of air, N<sub>2</sub>, CO<sub>2</sub>, O<sub>2</sub>) that flows into the system is known, the oxygen mass intake can be calculated using the ideal gas law (R = gas constant, M<sub>O2</sub> = molar mass of oxygen). The same principle applies to the calculation of the oxygen mass leaving the system where the measured oxygen volume fraction and the gas flow rate are used. As all fermentation processes were operated without any overpressure, the inlet gassing flow rate equals the outlet flow rate (F<sub>in</sub> = F<sub>out</sub>). Furthermore, the gas inlet and outlet temperature is presumed to be equal due to the length of the tubing (T<sub>in</sub> = T<sub>out</sub>). Finally, the current liquid fermentation volume (V<sub>L</sub>) and the ambient pressure (P<sub>amb</sub>) are needed to calculate the OUR as shown in Equation (3):

$$OUR(t) = (O_{2,in}(t) - O_{2,out}(t)) \cdot \frac{1}{V_L(t)} = \frac{(V_{O2,in}(t) - V_{O2,out}(t)) \cdot M_{O2} \cdot p_{amb}}{R \cdot T \cdot V_L(t)}$$
(3)

For the graphical representation in this study, the volumetric OUR was normalized to the maximum value in the training data set.

Α

Biomass / Product Titer /

**Product Yield** 

0

**Dynamic State** 

D ≠ const.

Bleed < D



Process time



Figure 1. Schematic overview of (A) continuous process consisting of perfusion-based dynamic state (red marked area) and cell-bleed-based steady-state phase and (B) the off-gas measurement set-up for continuous processes in single-use bioreactors.

#### 2.4. Off-Gas Measurement Set-Up

To perform off-gas analytic measurements in single-use bioreactors, a bypass solution was applied. The gas mixture enters the SUB at the bottom via a microsparger or open pipe. After the gas leaves the bioreactor, it passes heated filters and is then transported into a bottle that serves as a divider as well as a condensate trap. Here the gas stream is separated into two parts where the major part leaves the system via main exhaust whereas a smaller portion is distributed by a self-made gas manifold chamber and actively drawn using a membrane pump (Laboport<sup>®</sup> N96, KNF Neuberger GmbH, Freiburg, Germany) to the gas analyzer (DasGip<sup>®</sup> GA4, Eppendorf AG, Hamburg, Germany). The gas manifold and the multi-channel gas analyzer allow simultaneous off-gas measurements on up to four fermentation systems in parallel. Therefore, no multiplexing or flushing steps were necessary. We used gas-tight Teflon tubing for the whole transportation of the gas stream after it leaves the heated sterile filters as silicone tubing tends to be permeable for gases (Figure 1B). The gas analyzer was two-point calibrated before each process with air and a defined gas mixture (Linde AG, Höllriegelskreuth, Germany) containing 10% CO<sub>2</sub> and 2% O<sub>2</sub>. Unused gas analyzer channels were flushed with humidified air to preserve sensor lifetime.

#### 2.5. Data Collection and Preprocessing

All used offline and online data points for model generation were taken from perfusionbased biomass growth phases of continuous processes (P-01 to P-14). These processes were carried out in two identical SUB fermentation systems. Processes were technically identical in terms of used media, cultivation conditions and process operating conditions, as described above. Table 1 provides an overview on the respective cell lines, expressed antibody and data used for model creation and validation. In order to build a biomass prediction model, the collected OUR data gathered from off-gas analytics needed to be preprocessed to remove signal noise and measurement distortions.

All OUR raw data were fitted by higher order polynomials using corresponding regressions. The degree of the applied polynomials was chosen by the highest correlation coefficient ( $R^2$ ) in order to select the most descriptive regression model for each process. Since the OUR is the main factor affecting model quality, this pretreatment step was mandatory for achieving high model prediction accuracy.

#### 2.6. Model Generation and Assessment

The prediction model was built and evaluated using the statistical software JMP<sup>®</sup> 15.2.0 (SAS Institute, Cary, NC, USA). For the modeling procedure, treated OUR variables (OUR<sub>fitted</sub>), offline VCV and VCD and process time were used to fit a multilinear regression model. To address model performance, the root mean square error (RMSE) with Bessel's correction was calculated as shown in Equation (4):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_i')^2}{n - 1}}$$
(4)

In order to compare the performance of both prediction models, the mean absolute percentage error (MAPE) was calculated (Equation (5)):

MAPE = 
$$\frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - y_i'}{y_i} \right|$$
 (5)

In both equations,  $y_i$  represents the observed values,  $y_i'$  the corresponding predicted values and n the number of fitted points in total. Since the RMSE and MAPE are based on averages, outliers can negatively distort their predication [28]. Therefore, the MAPE was also calculated using the median of absolute percentage errors (MdAPE) to paint a more robust view on model accuracy.

#### 2.7. Real-Time Prediction and Validation

The models were implemented in data visualization and analyzer software SEEQ (SEEQ Corporation, Seattle, WA, USA) for calculation of VCV and VCD predictions in real time. Therefore, three new data sets from technical replicate processes (P-15, 16 and

17) with unknown cell lines to both models (C-15, 16 and 17) were used to validate the prediction models. Because online OUR raw data have low signal-to-noise ratios, as described above, two moving average smoothing algorithms were applied to assess their impact on final prediction accuracy. The used algorithm was either a Savitzky–Golay (SG) or a locally estimated scatterplot smoothing (LOESS) algorithm with equivalent analytical design regarding the investigation and sample output time range. Through the described analysis, real-time signal cleansing of OUR raw data was possible, leading to more accurate predictions. Figure 2 gives an overview of the performed model generation and validation workflow.



Figure 2. Workflow overview for biomass model generation and consequent validation.

#### 2.8. Off-Line Measurements

All cell physiological measures, such as viable cell density (VCD), average cell diameter (ACD), average volume per cell (AVC) and culture viability, were determined using an automatic cell counting device (Cedex HiRes<sup>®</sup>, Roche Diagnostics, Mannheim, Germany). In this work, the shape of a cell is assumed to be spherical, hence the AVC is calculated as follows:

$$AVC = \frac{4}{3}\pi \left(\frac{ACD}{2}\right)^3 \tag{6}$$

All samples were immediately processed after the sample was taken, as described in Section 2.3. Accordingly, 300  $\mu$ L of the cell containing sample was transferred into a Cedex HiRes sample cup and measured directly to avoid long-term storage. Depending on VCD concentration, the sample was diluted properly using 3% (m/v) Pluronic F68 dissolved in PBS. Furthermore, the measured VCD was used to calculate VCV according to Equation (7):

$$VCV = VCD \cdot AVC \tag{7}$$

For the graphical representation in this study, the biomass measures VCD and VCV were normalized to the maximum value in the training data set.

A biochemical analyzer (Cedex Bio HT<sup>®</sup>, Roche Diagnostics, Mannheim, Germany) was used to determine the metabolites glucose, lactate, pyruvate and ammonium. Therefore, cell suspension was centrifuged (Heraeus Multifuge 1S-R, Thermo Fisher Scientific, Waltham, MA, USA) at  $3500 \times g$  for 10 min. The cell pellet was discarded, and the supernatant was used subsequently.

Amino acid analysis was performed using an in-house LC-MS (Ultivo Triple Quadrupole LC/MS System, Agilent Technologies Inc., Santa Clara, CA, USA) procedure with stable isotope-labeled internal standards for calibration.

# 2.9. Cell-Specific Substrate and Metabolite Consumption and Production Rate, Product Formation Rate and Yield Calculation

The cell-specific substrate consumption and metabolite production rates in the dynamic state of continuous process were calculated, as recently described by Bausch et al., [29] according to the following balancing Equation (8):

$$\frac{\mathrm{dS}}{\mathrm{dt}} = \mathrm{D}(\mathrm{S}_{\mathrm{in}} - \mathrm{S}) + \mathrm{q}_{\mathrm{S}}\mathrm{X} \tag{8}$$

where S represents the molar concentration of the substrate or metabolite, D is the perfusion rate,  $S_{in}$  is the substrate molar concentration in the perfusion medium, X is the cell number and qS is the molar cell-specific substrate/metabolite production rate. In a simplified approach neglecting abiotic degradation of instable compounds, the cell-specific qS at discrete process time points, i, are calculated as described in Equation (9):

$$qS_{i} = \frac{1}{X_{i}} \left( \frac{(S_{i} - S_{i-1})}{(t_{i} - t_{i-1})} - D(S_{in} - S_{i}) \right)$$
(9)

A negative and positive value for qS represent consumption and production of a compound, respectively. The product formation rate qP can be calculated analogous to Equation (9). The metabolic yield coefficients  $Y_{Lac/Glc}$  and  $Y_{NH4/Gln}$  for the assessment of the metabolic state and efficiency were calculated as follows in Equations (10) and (11), respectively:

$$Y_{\text{Lac},i/\text{Glc},i} = \frac{q_{\text{Lac},i}}{q_{\text{Glc},i}}$$
(10)

$$Y_{\rm NH4,i/Gln,i} = \frac{q_{\rm NH4,i}}{q_{\rm Gln,i}}$$
(11)

#### 3. Results and Discussion

#### 3.1. Online Parameter Evaluation and Preprocessing

To monitor the biomass formation of 14 different CHO cell lines expressing different target proteins, we used an at-line-based viable cell density assessment, as described above. Although all CHO lines were derived from the same native CHO host cell line, we detected, as shown before by others, significant, process time-dependent differences in cell growth characteristics, such as viable cell density formation and timing for cell doubling, as well as in volume per cell among all tested clones (Figure 3A,C,E). Usually, biomass formation analysis is performed only once per day, resulting in an erroneous, discrete monitoring of this critical KPI, in conflict with the continuous use of this variable for dynamic calculations of important bioprocess key performance indicators, such as the cell-specific product formation rate and feedback control process strategies. Continuous assessment of cell biomass formation is a prerequisite for efficient bioprocess development and economic target protein production.

For the biomass soft-sensor model, general assumptions were made. Process variations occur from: Media lot-to-lot differences, pH and DO probe behavior, mass flow divergences, off-gas sensor response time and the metabolic performance of the used clone. These variations can influence the oxygen transfer and/or its solubility and, therefore, the oxygen level

becomes a sum parameter. Specifically, the different metabolic behavior of the clones (that leads to different controller responses and correction agent additions) might have the greatest impact on the fermentation broth and its physicochemical characteristics in terms of oxygen transfer. In addition, ambient conditions may vary during the course of a fermentation that affect the off-gas measurement. Residence time of the off-gas in the headspace as well as in the tubing and condensate trap bottles may further influence the proper calculation of the OUR [30]. Differences in the gas inlet and outlet temperature also have an impact on measured volume fractions, especially in cases where an off-gas cooler is used [16]. Since we did not use any off-gas cooling, we assumed the inlet and outlet gas temperature difference to be negligible. It has been shown that correction functions or description models generated from perturbation experiments can be applied to enhance accuracy of off-gas measurements for OUR calculations [30–32]. These approaches require considerable effort and profound knowledge about the characteristics of  $O_2$  transport kinetics within the whole system. However, we took none of the mentioned factors into consideration as we wanted to create a robust and relatively simple model that allows for easy implementation and good prediction quality in contrast to alternative, soft-sensing approaches. As described most recently by Tuveri et al., precise estimation of bioprocess variables such as biomass can be realized by comprehensive yet more complex approaches, such as the utilization of Kalman filters [33]. However, in our study, we were aiming for a biomass soft-sensor model that can handle the metabolic diversity and its effects on process properties that are caused by the varying clonal behavior of not yet in-depth, characterized CHO cell lines.

All available online variables were investigated regarding their ability to predict the biomass in terms of VCD and VCV. We found the OUR and process time (PT) to be the most predictive variables using a JMP predictor screening algorithm, which confirmed the known, high correlation of cellular biomass and respective volumetric OUR in cell culture (Figure 3B). The assembled OUR data (OUR<sub>raw</sub>) showed a low signal-to-noise ratio at the beginning of all processes up to several process days. The ratio was heavily influenced by low biomass concentrations as well as DO and pH controller response. Once the biomass reached a critical level accompanied by higher oxygen demand, the measured OUR signal became stable (Figure 4A).

Due to the use of different cell lines with diverse growth behavior as part of the data set, this condition differed clearly with respect to process time (Figure 4B) and mainly influenced the choice of the used higher order polynomials to describe the data set of each process as accurate as possible. We found polynomials of the third to fifth degree to fit best to the observed OUR raw data sets. Table 2 provides an overview of used polynomial degree and their related R<sup>2</sup>.

Process	Polynomial Degree	<b>R</b> <sup>2</sup>
P-01	5	0.863
P-02	4	0.983
P-03	4	0.878
P-04	4	0.888
P-05	4	0.712
P-06	4	0.971
P-07	4	0.991
P-08	3	0.892
P-09	3	0.982
P-10	3	0.977
P-11	3	0.918
P-12	4	0.956
P-13	4	0.977
P-14	4	0.954

Table 2. Overview of applied polynomial degree to describe the OUR raw data.



**Figure 3.** Time course of (**A**) normalized viable cell density (VCD), (**B**) normalized viable cell volume (VCV) vs. normalized volumetric oxygen uptake rate (OUR), (**C**) cell volume, (**D**) cell-specific OUR, (**E**) cell doubling time and (**F**) cell-specific OUR vs. normalized VCV of 14 different CHO cell lines (training data set, see Table 1) expressing different target proteins in a seven-day perfusion process. Black arrows and blue dotted lines show the perfusion rate protocol with respective normalized perfusion rate (in volume media per volume fermenter and day, vvd<sub>n</sub>) and timing strategy. The black lines represent the fit among all tested clones and runs and the grey area highlights the confidence of the fit with  $\alpha = 0.05$ .



**Figure 4.** (**A**) Example of normalized OUR raw data fit from process P-04 with high signal–noise ratio in the first 100 h and following stable signal towards end of fermentation. A polynomial fit of fourth grade was used to describe the OUR with an  $R^2$  of 0.89. (**B**) Example of normalized OUR raw data fit from process P-09 with high signal–noise ratio in the first 85 h and also towards end of fermentation. A polynomial fit of fourth grade was used to describe the OUR with an  $R^2$  of 0.95.

#### 3.2. Biomass Model Generation and Assessment

Two descriptive models were built using the preprocessed OUR and process time (PT) as input variables to predict the VCV and the VCD, respectively. Both regression models allow a good description of the biomass for each variable (Figure 5A,B). The VCV model has a normalized prediction error of RMSE = 0.0339, whereas the VCD model reaches 0.0469. Referring to relative model performance evaluation, the accuracy for VCV prediction was calculated as MAPE<sub>VCV</sub> = 31.79% and MdAPE<sub>VCV</sub> = 13.19%. Lower forecast performance values were obtained from the VCD model with MAPE<sub>VCD</sub> = 56.59% and  $MdAPE_{VCD} = 19.78\%$ . The differences between MAPE- and MdAPE-derived values can be explained by the nature of the observed errors and their distribution during the fermentations that were used to create these models. Despite the similarity from the observed residuals to the normal distribution (Shapiro-Wilk for VCV residuals is 0.89 and 0.82 for VCD), it is noticeable that, within both models, the difference between actual and predicted values begins to scatter with progressing process time and biomass concentration (Figure 5C,D). Small dimension residuals were observed up to 60-80 h after the process start and were highest towards the end of processes. However, the lack of prediction performance of both models is located at the beginning of the processes, as the magnitude of absolute percentage errors (APE) reveals (Figure 5E,F). Both prediction models show comparable behavior regarding the APE distribution but, significantly, lower APE magnitudes were found from the VCV model. This local APE density is mainly influenced by the high signal noise produced by the OUR raw data combined with comparably low biomass concentrations and, therefore, low oxygen demands. Despite the fact that the OUR data is preprocessed as described, the impact of the low signal-to-noise ratio heavily reduces the accuracy of both models. Furthermore, this is the leading cause for the described differences between MAPE and MdAPE values as the median is not as affected as the mean is by high APE occurrence, as mentioned. Additionally, a likely explanation for an increase in scattering residuals might be related to the necessary dilution to stay within the manufacturer's specifications and calibration ranges for the Cedex HiRes<sup>®</sup> cell density assay.



**Figure 5.** (**A**) Soft-sensor model for VCV prediction with model assessment RMSE and  $\mathbb{R}^2$ . Black dots represent normalized values, red line describes the found model with prediction confidence  $\alpha = 0.05$ . (**B**) Soft-sensor model for VCD prediction with model assessment RMSE and  $\mathbb{R}^2$ . Black dots represent normalized values, red line describes the found model with prediction confidence of  $\alpha = 0.05$ . (**C**) VCV normalized residuals plotted against normalized predicted values with process time indication. (**D**) VCD normalized residuals plotted against normalized predicted VCV values with process time indication. (**E**) VCV-model-derived absolute percentage errors plotted against normalized predicted VCV values with process time indication. (**F**) VCD-model-derived absolute percentage errors plotted against normalized predicted VCD values with process time indication.

Even if a high cell concentration in a given sample might decrease the measurement error, the probability is increased when covering a more characteristic amount of cells in the analyzed sample because pre-dilution procedures are prone to cause unintended consequential errors [34,35]. Therefore, we consider the user-dependent and manually applied dilution step as the root cause for the observed residual increase during the course of the processes. The second input variable, process time, is further expected to represent an indirect measure of biomass growth rate.

The estimation functions are listed below:

......

VCV <sub>Predicted</sub> ,Normalized	
$= -0.0907 + 0.425 \cdot \text{OUR}_{\text{Normalized}} + 0.00194 \cdot \text{PT}$	(12)
$+(PT - 74.736) \cdot ((OUR_{Normalized} - 0.207) \cdot 0.00571)$	

$$VCD_{Predicted,Normalized} = -0.0994 + 0.362 \cdot OUR_{Normalized} + 0.00243 \cdot PT + (PT - 74.736) \cdot ((OUR_{Normalized} - 0.207) \cdot 0.00554)$$
(13)

#### 3.3. Real-Time Prediction and Quality of Online OUR Monitoring

Using the identified models as a biomass soft sensor under real-time circumstances was considered as the chosen path of validation in this work. The estimator equations were implemented in SEEQ to perform online biomass prediction of dynamic state for the continuous processes (P-15, P-16 and P-17) with unknown cell lines (C-15, C-16 and C-17) to both models. The processes were executed in the same manner as described above, hence they are technical replicates, such as the processes P-01–P-14. In order to remove the signal noise from the calculated OUR, two signal-smoothing algorithms were applied in real time. As Bassey et al. [36] found the Savitzky–Golay (SG) filter algorithm to be well suited for gas-sensor-derived signal smoothing, we also applied the SG filter to remove signal distortions from the OUR signal. In addition, we tested a locally estimated scatterplot smoothing (LOESS)-based algorithms represent moving average functions that investigate a filter time window of 25 min with a permanent output frame of 30 s.

Using this approach, the SG applies a polynomial regression of first order, whereas the LOESS filter uses the best-fit line, which can either be a linear or a higher polynomial function. Due to the growth rate of animal cells of about 24 h, we consider the filter time delay to be negligible. Both soft-sensor models can predict the biomass in terms of VCV and VCD with good prediction accuracy regardless of whether the real-time OUR smoothing was done with the SG or LOESS algorithm (Figure 6A–D). Nevertheless, referring to model assessment parameters, the VCV model shows a significantly higher goodness of fit in each case (Tables 3 and 4). Calculated MAPE values are half the magnitude from the VCV model (MAPE<sub>VCVLOESS/SG</sub>  $\approx$  14%) compared to VCD-model-derived MAPE (MAPE<sub>VCD,LOESS/SG</sub>  $\approx$  33%). Therefore, the VCV model is leading to predictions that are more precise on average. Beyond that, the difference between MAPE and MdAPE values is still noticeable in a comparable period after process start (Figure 6A,B) as the same root cause of a high signal-to-noise ratio creates a high local APE density. However, MdAPE values between both models are quite comparable and are in the range of  $MdAPE_{VCD,LOESS/SG} \approx 8\%$  and, for the VCV model,  $MdAPE_{VCV,LOESS} = 6.6\%$  and  $MdAPE_{VCV,SG} = 8.3\%$ . Half of the prediction errors are located above and below these values and, in reference to the calculated average prediction errors, the VCV model has the best prediction performance validated on the novel cell lines C-15, 16 and 17.

In contrast to offline-based measurements which usually consists of only one or a few measurement points per day, the prediction provides a continuous description of biomass during the processes, filling in the gaps between offline-derived measurements (Figure 6E,F). These real-time predictions can be further utilized to calculate other meaningful process variables in a soft-sensing manner, such as production or consumption rates

Α

0.9

0.8

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0.0

0.9

0.8

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0.0

0.9

0.8

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0.0

0 20 40 60 80

Process time [h]

Ε

VCV Actual (Normalized)

VCD Actual (Normalized)

С

VCV Actual (Normalized)



(see Section 3.5). Additionally, high quality online biomass forecasts enable a verification of erroneous offline-based readings, revealing possible measurement errors.

Figure 6. (A) Normalized measured VCV vs. real-time normalized predicted VCV data using the LOESS filter algorithm with model assessment for all three processes represented by colored dots. (B) Normalized measured VCV vs. real-time normalized predicted VCV data using the SG filter algorithm with model assessment for all three processes represented by colored dots. (C) Normalized measured VCD vs. real-time normalized predicted VCD data using the LOESS filter algorithm with model assessment for all three processes represented by colored dots. (D) Normalized measured VCD vs. real-time normalized predicted VCD data using the SG filter algorithm with model assessment for all three processes represented by colored dots. (E) Exemplary normalized predicted VCV<sub>LOESS/SG</sub> values (orange and gray lines) and actual normalized VCV values from process P-15 with clone C-15 represented by blue dots. Error bars describe an assumed 11% error for all VCV measurements. (F) Exemplary normalized predicted VCD<sub>LOESS/SG</sub> values (orange and gray lines) and actual normalized VCD values from process P-15 with clone C-15 represented by blue dots. Error bars describe an assumed 10% error for all VCD measurements. Biomass measures VCV and VCD are normalized to the maximum value in the training data set (P-01-P-14).

100 120 140 160 180

0.2

0.1

0.0

0 20 40

60 80 100 120

Process time [h]

140 160 180

Process	R <sup>2</sup>		RMSE (Normalized)		<b>MAPE</b> [%]		MdAPE [%]	
	SG	LOESS	SG	LOESS	SG	LOESS	SG	LOESS
P-15	0.999	0.998	0.017	0.015	18.13	18.15	4.93	3.85
P-16	0.974	0.980	0.057	0.052	14.94	14.61	14.35	13.77
P-17	0.995	0.994	0.025	0.025	7.99	8.23	6.61	6.60

Table 3. Overview of measured VCV and real-time predicted VCV data using SG or LOESS algorithm.

Table 4. Overview of measured VCD and real-time predicted VCD data using SG or LOESS algorithm.

Process	R <sup>2</sup>		RMSE (Normalized)		<b>MAPE</b> [%]		MdAPE [%]	
	SG	LOESS	SG	LOESS	SG	LOESS	SG	LOESS
P-15	0.997	0.995	0.035	0.033	33.09	32.57	9.85	8.29
P-16	0.965	0.971	0.072	0.068	16.75	15.21	19.43	11.16
P-17	0.996	0.995	0.025	0.031	48.81	49.71	6.79	7.99

Since the increase in biomass is always accompanied by a growth in cell number and cell volume, we consider a description of the biomass solely by cell number in terms of VCD to be insufficient. VCD is a coarse measure of the viable biomass, because even small changes in mean cell diameter result in large differences in cell volume [37]. An analysis of cell size, especially its distribution during fermentation process time, can deliver valuable information that cannot be seen by only looking at cell numbers. Besides the fact that trypan blue-based automatic cell counting enables a differentiation in viable and nonviable cells, numerous publications can be found that highlight the advantages and also the necessity of cell size in terms of cell volume measurements [37–41]. Mammalian cell volume differs not only between cell lines but also during an ongoing process, which leads to changing biomass in terms of volume and cellular mass itself. In addition, the process mode, growth conditions and other parameters can influence cell size. For example, larger cells tend to consume more oxygen than smaller cells, and the rapid adaptability of cells to process conditions such as osmolality, where a rise results in cell size increase, underlines the advantages of having cell size measured [22]. All factors support our preference for more accurate correlations for a VCV-based biomass description. Furthermore, it has been demonstrated that packed cell volume measurements can reach errors below 5%, whereas standard trypan blue cell counting techniques still struggle with errors up to 15% [42].

# 3.4. Biomass-Specific Oxygen Demand and Key Metabolism Analysis

The metabolism of CHO cell lines during classical batch and fed-batch cultivation is highly dynamic, and metabolic steady-state descriptions can be used to analyze the coherences by mechanistic modeling approaches [43]. These significant metabolic changes originate from alterations in the dynamic cell environmental media matrix composition, such as substrate and cofactor consumption, (toxic) metabolite production and shifts in chemicophysical parameters, such as medium osmolality, buffer capacity and redox potential [44–47]. Perfusion cell cultivation processes can be used to overcome these media matrix variations by an optimized constant replacement of conditioned media with fresh media and by using parameters such as the cell-specific perfusion rate (CSPR). Nevertheless, the optimization and analysis of CSPR was not the goal of this study.

We analyzed the biomass-specific OUR and metabolism of key substrates and metabolites of tested CHO cell lines in SUB perfusion processes in more detail. As shown previously [14,48], the volumetric OUR of tested CHO cell lines followed the previously described cell density kinetics during cell cultivation (Figure 3B). The observed cell density formation consequently showed differences among each tested CHO cell line, and the final volumetric OUR of the perfusion processes showed a constant increase over process time. At the end of the cultivation, the volumetric OUR showed a broad variation between all tested CHO cell lines and, for C-05 and C-13, up to more than 100% more than the respective variance observed for viable cell densities (Figure 3B). The cell-specific OUR (qOUR), however, showed an initial slight increase followed by highly homogenous qOUR for all tested clones and plateaued on a stable level of approximately 41.7 amol cell<sup>-1</sup> s<sup>-1</sup> from day 5 until the end of the perfusion process at day 7 (Figure 3D). The observed level of qOUR fits well with previously reported qOURs for CHO suspension cells [24,27,48–50]. Plotting the viable cell volume of the process (VCV) vs. qOUR revealed very high qOUR and, subsequently, a fast decrease of qOUR in the beginning of the perfusion cultivation where low biomass was available, followed by a stable plateauing of qOUR (Figure 3F). Both observations, the initial increase in qOUR followed by a stabilization at a lower qOUR level at the later cell cultivation phases and the higher biomass levels, confirm previously reported trends for CHO cells in perfusion cultivations [27,51]. The early qOUR peaks were attributed to an initial metabolic acclimation phase when cells were seeded into an unconditioned media with high substrate concentrations and the cultivation conditions at start of cell culturing.

To understand the reason for this shift in early and late qOUR kinetics, we analyzed the concentration and consumption/production rates of glucose, lactate, glutamine and ammonia as key substrates in mammalian cell cultures. Significant changes in volumetric glucose and glutamine substrate availability, as well as lactate and ammonium byproduct levels, were observed by using the applied perfusion process strategy. Both glucose and glutamine levels dropped during the course of the perfusion process, with an earlier decline in glutamine, which may be due to additional abiotic degradation (Figure 7A). Both byproducts, lactate and ammonium showed an initial increase followed by an intermediate plateau phase between day 3 and 5 and a final metabolic inverse shift with a decreasing level of lactate and, subsequently, an increase in ammonium from day 5 until the end of perfusion fermentation at day 7 (Figure 7A). The analysis of the cell-specific rates of these substrates and metabolites emphasizes the metabolic shift at day 5 with a stagnation in low levels of cell-specific glutamine consumption qGln and lactate formation rates qLac (Figure 7B).

The yield coefficients  $Y_{Lac/Glc}$  and  $Y_{NH4/Gln}$  are characteristic bioprocess key parameters (KPI) for assessing the metabolic status of cellular systems and the utilized pathways for energy production. By applying these parameters, we temporally analyzed the yield coefficient  $Y_{Lac/Glc}$  and  $Y_{NH4/Gln}$  along the perfusion process time. Through our analysis, we identified three distinct metabolic phases: (i) from day 0 to day 3, a phase of high anaerobic lactate production and glutaminolysis-driven ammonium formation with a clone-dependent  $Y_{Lac/Glc}$  of 1–2 mol/mol and  $Y_{NH4/Gln}$  of 0.5–3.6 mol/mol, (ii) from day 3 to day 5, a metabolic transition phase switching to aerobic metabolism and low glutaminolysis activity and (iii) from day 5 to day 7, an almost complete aerobic phase with practically no lactate production and clone-dependent increasing glutaminolysis again ( $Y_{Lac/Glc}$  of 0.03–0.64 mol/mol,  $Y_{NH4/Gln}$  of 0.7–1.7 mol/mol) (Figure 7C,D).

The yield analysis by  $Y_{Lac/Glc}$  and  $Y_{NH4/Gln}$  suggests an alternative reason for the observed metabolic switch rather than substrate limitation since glucose and glutamine are available in the fermentation media matrix in high amounts during the whole perfusion process (Figure 7A). The limitation of pyruvate during the perfusion process was identified as a putative reason for the metabolic switch. The slight increase of cell-specific glucose consumption qGluc and of the glutaminolysis and ammonium formation qS<sub>NH4</sub> from day 5 onward correlates with the limitation of pyruvate (Figure 8A) and stagnation of cell-specific pyruvate consumption rate qPyr (Figure 8B). In general, pyruvate is an important alternative, energy-generating carbon source for fast proliferating mammalian cell lines and for reducing cell growth-inhibiting ammonium production in cell cultures [52]. Analysis of qPyr vs. the available global pyruvate concentrations in the culture suggests a concentration-dependent shift of qPyr at levels lower than 2 mM, which correlates with the

increase of cell-specific ammonium formation with the drop in cell doubling time (data not shown).

In principle, the accumulations of cytostatic/toxic metabolic byproducts, other than lactate and ammonium, originating from the amino acid break-down metabolism in CHO fermentation processes are well characterized triggers which induce decreased biomass formation and increased cell doubling time [53]. In our study, however, we focused on the classical cell culture substrates and metabolites yet encouraged the analysis of these amino acid break-down products in the future to allow for optimized perfusion process designs with efficient depletion of known and unknown cytostatic/toxic metabolic byproducts.



**Figure 7.** Metabolic analysis of the dynamic state of continuous CHO SUB processes. Kinetic of key substrates glucose and glutamine and metabolites lactate and ammonium concentrations (**A**,**B**) cell-specific rates. Time-resolved analysis of the cell-specific rate-based yield coefficients (**C**)  $Y_{Lac/Glc}$  and (**D**)  $Y_{NH4/Gln}$ . The colored dots represent the tested 14 clones and the black, blue, red and green lines represent the fit of Gln concentration or cell-specific Gln consumption/production rate qGln, NH<sub>4</sub><sup>+</sup> concentration or cell-specific NH<sub>4</sub><sup>+</sup> consumption/production rate qNH4, glucose concentration or cell-specific glucose consumption rate and lactate concentration or cell-specific lactate consumption/production rate, respectively. The black, blue, red and green areas highlight the confidence of the fits with  $\alpha = 0.05$ . Black arrows and blue dotted lines show the perfusion rate protocol with respective normalized perfusion rate (in volume media per volume fermenter and day, vvd<sub>n</sub>) and timing strategy.



**Figure 8.** Real-time prediction of qPyr in dynamic state of continuous CHO SUB processes. (**A**) Analysis of pyruvate concentration over perfusion process time. The colored dots represent the tested 14 clones from the training data set and the black line represents the fit with  $\alpha = 0.05$ . (**B**) qPyr cell-specific consumption/production rates. The black area highlights the confidence of the fit with = 0.05. (**C**) Actual vs. predicted plot of a logistic regression model for qPyr for all tested clones from the training data set C-1 to C-14 (black dots) with regression model prediction (red line) and mean of all tested qPyr (blue line). (**D**) Online prediction of qPyr for a model validation perfusion process with C-15 (grey line) with qPyr actuals (blue dots). Error bars describe an assumed 10% error for actual qPyr values.

Mammalian amino acid metabolism is highly dependent upon the availability of bioavailable oxygen as an electron acceptor to allow for an indirect regeneration of redox equivalents NAD<sup>+</sup> and FAD in the tricarbon cycle (TCA), which are finally needed for the oxidative phosphorylation and energy production in cells [54]. Since there is no report that describes the correlation of specific amino acid and key metabolite consumption/production rates qS with qOUR of CHO cell lines in SUB continuous processes, we aimed to analyze this important investigation in our experimental set-up. Unexpectedly, cell-specific qGluc showed no correlation to qOUR (R<sup>2</sup>: 0.007, RMSE: 0.31 pmol cell<sup>-1</sup> d<sup>-1</sup>) yet the following important metabolic rates of key substrates and metabolites revealed a sound correlation: qGln (R<sup>2</sup>: 0.389, RMSE: 0.13 pmol cell<sup>-1</sup> d<sup>-1</sup>), qAla (R<sup>2</sup>: 0.540, RMSE: 0.06 pmol cell<sup>-1</sup> d<sup>-1</sup>), qPyr (R<sup>2</sup>: 0.521, RMSE: 0.23 pmol cell<sup>-1</sup> d<sup>-1</sup>), qLac (R<sup>2</sup>: 0.324, RMSE:

0.32 pmol cell<sup>-1</sup> d<sup>-1</sup>) and qNH4 (R<sup>2</sup>: 0.741, RMSE: 0.08 pmol cell<sup>-1</sup> d<sup>-1</sup>) (Figure 9A). In addition, the cell-specific product formation rate qP revealed no correlation to the cell biomass-specific OUR (R<sup>2</sup>: 0.062, RMSE: 4.42 pg cell<sup>-1</sup> d<sup>-1</sup>) (Figure 9B).



**Figure 9.** Correlation of **(A)** cell-specific substrate and metabolite formation/consumption rates and **(B)** product formation rate. The colored dots represent the tested 14 clones and for **(A)** the black, blue, red, green, violet and brown lines represent the fit of qGluc, qGln, qAla, qPyr, qLac and qNH4 cell-specific consumption/production rates and for **(B)** the black line represent the fit of qP. The dark black, blue, red, green, violet and brown areas highlight the confidence of the fits with  $\alpha$ = 0.05 and light-colored areas the respective confidences of the predictions.

# 3.5. Online Prediction of Cellular Metabolic Rates

As shown in the previous section, the calculations and analyses of biomass-specific substrate consumption and metabolite production rates, qS, are mandatory to identify distinct cell metabolic phases, which can be preferably used to optimize perfusion media and rates for an efficient continuous cultivation of CHO cell lines. Solely monitoring global substrate and metabolite concentrations is not sufficient to allow for an equivalent characterization of cell cultivation processes, such as the described CHO perfusion process in SUBs.

As a proof of concept, we developed a soft-sensor-based real-time prediction of the cellspecific pyruvate consumption/production rate qPyr by using available real-time estimates for qOUR and biomass measures, as described before. The importance of an immediate estimation of the metabolic pyruvate flux into the cell is justified by its central role in the direct and indirect control of the cellular energy metabolism. Pyruvate is funneled into the TCA by the pyruvate dehydrogenase complex and/or by the anaplerotic reaction regulated by the pyruvate carboxylase [55,56]. Therefore, monitoring coupled with tailored control of qPyr is generally envisioned to improve the cellular energy state and avoid the lactate accumulation in cell culture fermentation processes.

qPyr correlated well with discrete cell-specific qOUR values ( $R^2$  of 0.521, RMSE of 0.23) by using the discrete qPyr and real-time predicted qOUR data of the 14 different CHO cell lines and perfusion processes (Figure 9), suggesting the possibility to directly use this important information on the respiratory metabolism for a soft-sensoring approach for real-time qPyr prediction. As a proof of concept, a suitable logistic multiregression model was generated for the generalized, sigmoid qPyr time course by simply using the available online OUR data and the predicted, SG-smoothed VCD, VCV and cell- and biomass-specific qOUR values ( $R^2$  of 0.8, RMSE of 0.0334 pmol cell<sup>-1</sup> d<sup>-1</sup>) (Figure 8C). The used estimation functions for qPyr prediction are shown in following Equation (14):

$$qPyr_{Predicted} = Logist(1.046 \cdot 10^{12} + 67.365 \cdot Logist(OUR) - 0.259)$$
$$\cdot Logist(VCD_{Predicted}) + 1.87 \cdot 10^{14} \cdot qOUR_{cellvolumePredicted} -2.092 \cdot 10^{12} \cdot qOUR_{cellPredicted})$$
(14)

We validated the prediction estimation model for qPyr by using a validation data set with CHO clone C-15 and perfusion process P-15 (Table 1). By this, the real-time prediction and discrete actuals for qPyr showed a technically relevant, good correlation in this validation data set (Figure 8D). The reason for the observed offset likely originates from the erroneous discrete qPyr measurement and respective error propagation by the calculation and/or by the cell-specific metabolic nature, often described for CHO cells with a high genetic plasticity [57]. The first 24 h of the prediction were not used due to the high noise in the OUR signal due to reasons described before. In general, more elaborated non-linear modeling approaches, such as decision trees and artificial neuronal nets, may also be used in the future for an increased precise estimation of cell-specific rates such as qPyr. Regardless, using these powerful modeling approaches requires large, annotated data sets that can be technically realized simply over a longer period of time.

# 4. Conclusions

In this work, we present, for the first time, an off-gas-based soft sensor for real-time biomass prediction in SUB continuous processes with CHO cell lines. The 14 different CHO cell lines that were used to build the soft-sensor models cover a variety of phenotypically different CHO cell lines. Given the diversity of our training data set, we expect the resulting models to be applicable to a broad range of CHO cell lines. This application is underlined by a high prediction accuracy achieved by the models on the bioprocesses of three novel CHO cell lines which were previously unknown to both models. The detailed analysis of both the model residuals as well as the absolute percentage errors disclosed some weaknesses that are primarily process related. The noisy OUR raw signal that was observed during the onset of all cell cultivation processes is caused by the pH controller response leading to very high prediction errors for up to 80 h after the processes were started. Optimization of pH controller settings and strategies or using more basic pH set points could overcome these technical challenges (data not shown). In addition, a split into different forecast models where altered pH controller interferences are present could lead to lower prediction errors. In addition, alternative yet computational and model-calibration-intensive forecast approaches such as Kalman filtering could significantly increase the prediction quality and should be considered for further, more elaborated closed-loop variable predictions and process control strategies [33].

Our data also demonstrated that higher model accuracy was established when VCV instead of VCD was used as biomass depiction. This strengthens our strong belief in a paradigm change regarding biomass description in modern bioprocesses. VCD should no longer be the leading, or the only, measurement looked at when it comes to biomass determination. The cell size or volume, its distribution over time and, of course, the VCV should be used by default to accurately describe the biomass and all derived metabolic variables, such as mAB, lactate production rate, or glucose/oxygen consumption rates. Conclusions, based only on cell density measurements, can lead to wrong assumptions, calculations or other unforeseen misinterpretations, generating a fragmented picture of the biomass [38,40]. As modern bioprocesses can be highly complex and dynamic, the biomass and cellular metabolism analysis should be as comprehensive as possible to generate a comparable and reproducible data basis. Furthermore, the utilization of an off-gas-based soft sensor is easy to implement in SUB systems as well as in common stainless steel plants. For this purpose, the installation of any hard-type probes inside the bioreactor is not necessary and does not increase handling or decrease safety and therefore prevents possible contamination risks. The fundamental correlation of biomass growth and increasing oxygen demand can be used, optimized and extended to generate profound real-time knowledge on diverse bioprocess variables such as the shown biomass and metabolic nutrient rate soft sensor. Moreover, off-gas analysis can be used to determine the true bioreactor pH without any sampling or as non-invasive method for online  $pCO_2$  monitoring, which underlines the flexibility and outstanding character of having an off-gas analyzer implemented and running [58].

**Author Contributions:** T.W. and O.P. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This work is exclusively funded by F. Hoffman-La Roche AG.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** Restrictions apply to the availability to these data; therefore, data sharing is not possible.

Acknowledgments: The authors thank the whole Bioprocess Research Team for their support. The authors thank Frederik Schröter, Ulrike Vollertsen and Katja Montan for the measurements of amino acids using the described HT-MS method. In addition, we would like to thank Tobias Grosskopf for critical reading. Furthermore, we thank Michalea Poth for the helpful discussion and Kathryn Perez for proofreading.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- 1. Aehle, M.; Kuprijanov, A.; Schaepe, S.; Simutis, R.; Lübbert, A. Increasing batch-to-batch reproducibility of CHO cultures by robust open-loop control. *Cytotechnology* **2011**, *63*, 41–47. [CrossRef] [PubMed]
- FDA. Guidance for Industry: PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance; FDA: Silver Spring, MD, USA, 2004.
- 3. Lopes, A.G. Single-use in the biopharmaceutical industry: A review of current technology impact, challenges and limitations. *Food Bioprod. Process.* **2015**, *93*, 98–114. [CrossRef]
- 4. Löffelholz, C.; Husemann, U.; Greller, G.; Meusel, W.; Kauling, J.; Ay, P.; Kraume, M.; Eibl, R.; Eibl, D. Bioengineering Parameters for Single-Use Bioreactors: Overview and Evaluation of Suitable Methods. *Chem. Ing. Tech.* **2013**, *85*, 40–56. [CrossRef]
- 5. Junne, S.; Neubauer, P. How scalable and suitable are single-use bioreactors? Curr. Opin. Biotechnol. 2018, 53, 240–247. [CrossRef]
- 6. Mandenius, C.-F. Measurement Technologies for Upstream and Downstream Bioprocessing. *Processes* **2021**, *9*, 143. [CrossRef]
- Gnoth, S.; Jenzsch, M.; Simutis, R.; Lübbert, A. Process Analytical Technology (PAT): Batch-to-batch reproducibility of fermentation processes by robust process operational design and control. *J. Biotechnol.* 2007, 132, 180–186. [CrossRef]
- Justice, C.; Brix, A.; Freimark, D.; Kraume, M.; Pfromm, P.; Eichenmueller, B.; Czermak, P. Process control in cell culture technology using dielectric spectroscopy. *Biotechnol. Adv.* 2011, 29, 391–401. [CrossRef]
- 9. Busse, C.; Biechele, P.; De Vries, I.; Reardon, K.F.; Solle, D.; Scheper, T. Sensors for disposable bioreactors. *Eng. Life Sci.* 2017, 17, 940–952. [CrossRef]

- 10. Steinwandter, V.; Zahel, T.; Sagmeister, P.; Herwig, C. Propagation of measurement accuracy to biomass soft-sensor estimation and control quality. *Anal. Bioanal. Chem.* **2016**, 409, 693–706. [CrossRef]
- 11. Kiviharju, K.; Salonen, K.; Moilanen, U.; Eerikäinen, T. Biomass measurement online: The performance of in situ measurements and software sensors. J. Ind. Microbiol. Biotechnol. 2008, 35, 657–665. [CrossRef]
- 12. Mandenius, C.-F.; Gustavsson, R. Mini-review: Soft sensors as means for PAT in the manufacture of bio-therapeutics. J. Chem. *Technol. Biotechnol.* 2015, 90, 215–227. [CrossRef]
- 13. Luttmann, R.; Bracewell, D.G.; Cornelissen, G.; Gernaey, K.V.; Glassey, J.; Hass, V.C.; Kaiser, C.; Preusse, C.; Striedner, G.; Mandenius, C.-F. Soft sensors in bioprocessing: A status report and recommendations. *Biotechnol. J.* **2012**, *7*, 1040–1048. [CrossRef]
- Pappenreiter, M.; Sissolak, B.; Sommeregger, W.; Striedner, G. Oxygen Uptake Rate Soft-Sensing via Dynamic k L a Computation: Cell Volume and Metabolic Transition Prediction in Mammalian Bioprocesses. *Front. Bioeng. Biotechnol.* 2019, 7, 195. [CrossRef] [PubMed]
- Metze, S.; Ruhl, S.; Greller, G.; Grimm, C.; Scholz, J. Monitoring online biomass with a capacitance sensor during scale-up of industrially relevant CHO cell culture fed-batch processes in single-use bioreactors. *Bioprocess Biosyst. Eng.* 2020, 43, 193–205. [CrossRef] [PubMed]
- 16. Aehle, M.; Kuprijanov, A.; Schaepe, S.; Simutis, R.; Lübbert, A. Simplified off-gas analyses in animal cell cultures for process monitoring and control purposes. *Biotechnol. Lett.* **2011**, *33*, 2103–2110. [CrossRef]
- 17. Zeng, A.-P.; Byun, T.-G.; Deckwer, W.-D. On-line estimation of viable biomass of a microaerobic culture using exit gas analysis. *Biotechnol. Tech.* **1991**, *5*, 247–250. [CrossRef]
- 18. Dorresteijn, R.C.; Numan, K.H.; De Gooijer, C.D.; Tramper, J.; Beuvery, E.C. On-line estimation of the biomass activity during animal-cell cultivations. *Biotechnol. Bioeng.* **1996**, *51*, 206–214. [CrossRef]
- Goldrick, S.; Umprecht, A.; Tang, A.; Zakrzewski, R.; Cheeks, M.; Turner, R.; Charles, A.; Les, K.; Hulley, M.; Spencer, C.; et al. High-Throughput Raman Spectroscopy Combined with Innovate Data Analysis Workflow to Enhance Biopharmaceutical Process Development. *Processes* 2020, *8*, 1179. [CrossRef]
- 20. Downey, B.J.; Graham, L.J.; Breit, J.F.; Glutting, N.K. A novel approach for using dielectric spectroscopy to predict viable cell volume (VCV) in early process development. *Biotechnol. Prog.* **2014**, *30*, 479–487. [CrossRef]
- 21. Yoon, S.-J.; Konstantinov, K.B. Continuous, real-time monitoring of the oxygen uptake rate (OUR) in animal cell bioreactors. *Biotechnol. Bioeng.* **1994**, *44*, 983–990. [CrossRef]
- 22. Wagner, B.A.; Venkataraman, S.; Buettner, G.R. The Rate of Oxygen Utilization by Cells. *Free Radic. Biol. Med.* 2011, 51, 700–712. [CrossRef]
- 23. Ruffieux, P.-A.; von Stockar, U.; Marison, I.W. Measurement of volumetric (OUR) and determination of specific (qO2) oxygen uptake rates in animal cell cultures. *J. Biotechnol.* **1998**, *63*, 85–95. [CrossRef]
- Goudar, C.T.; Piret, J.M.; Konstantinov, K.B. Estimating cell specific oxygen uptake and carbon dioxide production rates for mammalian cells in perfusion culture. *Biotechnol. Prog.* 2011, 27, 1347–1357. [CrossRef]
- Fleischaker, R.J.; Sinskey, A.J., Jr. Oxygen demand and supply in cell culture. *Appl. Microbiol. Biotechnol.* 1981, 12, 193–197. [CrossRef]
- 26. Pereira, S.; Kildegaard, H.F.; Andersen, M.R. Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients. *Biotechnol. J.* **2018**, *13*, e1700499. [CrossRef] [PubMed]
- 27. Heidemann, R.; Lütkemeyer, D.; Büntemeyer, H.; Lehmann, J. Effects of dissolved oxygen levels and the role of extra- and intracellular amino acid concentrations upon the metabolism of mammalian cell lines during batch and continuous cultures. *Cytotechnology* **1998**, *26*, 185–197. [CrossRef] [PubMed]
- 28. De Jonge, L.; Heijnen, J.; van Gulik, W. Reconstruction of the oxygen uptake and carbon dioxide evolution rates of microbial cultures at near-neutral pH during highly dynamic conditions. *Biochem. Eng. J.* **2014**, *83*, 42–54. [CrossRef]
- 29. Moreno, J.J.M.; Pol, A.P.; Abad, A.S. Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema* **2013**, 25, 500–506. [CrossRef]
- Bausch, M.; Schultheiss, C.; Sieck, J.B. Recommendations for Comparison of Productivity Between Fed-Batch and Perfusion Processes. *Biotechnol. J.* 2019, 14, e1700721. [CrossRef]
- 31. Wu, L.; Lange, H.C.; Van Gulik, W.M.; Heijnen, J.J. Determination of in vivo oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. *Biotechnol. Bioeng.* **2002**, *81*, 448–458. [CrossRef]
- 32. Bloemen, H.H.J.; Wu, L.; van Gulik, W.M.; Heijnen, J.J.; Verhaegen, M.H.G. Reconstruction of the O<sub>2</sub> uptake rate and CO<sub>2</sub> evolution rate on a time scale of seconds. *AIChE J.* **2003**, *49*, 1895–1908. [CrossRef]
- Tuveri, A.; Pérez-García, F.; Lira-Parada, P.A.; Imsland, L.; Bar, N. Sensor fusion based on Extended and Unscented Kalman Filter for bioprocess monitoring. J. Process. Control 2021, 106, 195–207. [CrossRef]
- 34. Sarkar, S.; Lund, S.P.; Vyzasatya, R.; Vanguri, P.; Elliott, J.T.; Plant, A.L.; Lin-Gibson, S. Evaluating the quality of a cell counting measurement process via a dilution series experimental design. *Cytotherapy* **2017**, *19*, 1509–1521. [CrossRef] [PubMed]
- Cadena-Herrera, D.; Lara, J.E.E.-D.; Ramírez-Ibañez, N.D.; Lopez-Morales, C.A.; Pérez, N.O.; Flores-Ortiz, L.F.; Medina-Rivero, E. Validation of three viable-cell counting methods: Manual, semi-automated, and automated. *Biotechnol. Rep.* 2015, 7, 9–16. [CrossRef] [PubMed]
- 36. Bassey, E.; Whalley, J.; Sallis, P. An Evaluation of Smoothing Filters for Gas Sensor Signal Cleaning. In Proceedings of the Fourth International Conference on Advanced Communications and Computation, Paris, France, 8–9 February 2014; pp. 19–23.

- Sonderhoff, S.A.; Kilburn, D.G.; Piret, J.M. Analysis of mammalian viable cell biomass based on cellular ATP. *Biotechnol. Bioeng.* 1992, 39, 859–864. [CrossRef]
- Pan, X.; Dalm, C.; Wijffels, R.H.; Martens, D.E. Metabolic characterization of a CHO cell size increase phase in fed-batch cultures. *Appl. Microbiol. Biotechnol.* 2017, 101, 8101–8113. [CrossRef]
- 39. Nielsen, L.K.; Reid, S.; Greenfield, P.F. Cell cycle model to describe animal cell size variation and lag between cell number and biomass dynamics. *Biotechnol. Bioeng.* **1997**, *56*, 372–379. [CrossRef]
- 40. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* 2013, 35, 1050–1055. [CrossRef]
- 41. Frame, K.K.; Hu, W.-S. Cell volume measurement as an estimation of mammalian cell biomass. *Biotechnol. Bioeng.* **1990**, *36*, 191–197. [CrossRef]
- 42. Stettler, M.; Jaccard, N.; Hacker, D.; De Jesus, M.; Wurm, F.M.; Jordan, M. New disposable tubes for rapid and precise biomass assessment for suspension cultures of mammalian cells. *Biotechnol. Bioeng.* 2006, 95, 1228–1233. [CrossRef]
- Popp, O.; Müller, D.; Didzus, K.; Paul, W.; Lipsmeier, F.; Kirchner, F.; Niklas, J.; Mauch, K.; Beaucamp, N. A hybrid approach identifies metabolic signatures of high-producers for chinese hamster ovary clone selection and process optimization. *Biotechnol. Bioeng.* 2016, 113, 2005–2019. [CrossRef]
- 44. Yuk, I.H.; Zhang, J.D.; Ebeling, M.; Berrera, M.; Gomez, N.; Werz, S.; Meiringer, C.; Shao, Z.; Swanberg, J.C.; Lee, K.H.; et al. Effects of copper on CHO cells: Insights from gene expression analyses. *Biotechnol. Prog.* **2014**, *30*, 429–442. [CrossRef]
- 45. Takagi, M.; Hayashi, H.; Yoshida, T. The effect of osmolarity on metabolism and morphology in adhesion and suspension chinese hamster ovary cells producing tissue plasminogen activator. *Cytotechnology* **2000**, *32*, 171–179. [CrossRef] [PubMed]
- 46. Park, J.H.; Noh, S.M.; Woo, J.R.; Kim, J.W.; Lee, G.M. Valeric acid induces cell cycle arrest at G1 phase in CHO cell cultures and improves recombinant antibody productivity. *Biotechnol. J.* **2016**, *11*, 487–496. [CrossRef] [PubMed]
- 47. Chevallier, V.; Andersen, M.R.; Malphettes, L. Oxidative stress-alleviating strategies to improve recombinant protein production in CHO cells. *Biotechnol. Bioeng.* 2020, 117, 1172–1186. [CrossRef]
- 48. Deshpande, R.R.; Heinzle, E. On-line oxygen uptake rate and culture viability measurement of animal cell culture using microplates with integrated oxygen sensors. *Biotechnol. Lett.* **2004**, *26*, 763–767. [CrossRef]
- 49. Jorjani, P.; Ozturk, S.S. Effects of cell density and temperature on oxygen consumption rate for different mammalian cell lines. *Biotechnol. Bioeng.* **1999**, *64*, 349–356. [CrossRef]
- 50. Gray, D.R.; Chen, S.; Howarth, W.; Inlow, D.; Maiorella, B.L. CO<sub>2</sub> in large-scale and high-density CHO cell perfusion culture. *Cytotechnology* **1996**, 22, 65–78. [CrossRef] [PubMed]
- Super, A.; Jaccard, N.; Marques, M.P.C.; Macown, R.J.; Griffin, L.D.; Veraitch, F.S.; Szita, N. Real-time monitoring of specific oxygen uptake rates of embryonic stem cells in a microfluidic cell culture device. *Biotechnol. J.* 2016, *11*, 1179–1189. [CrossRef] [PubMed]
- 52. Genzel, Y.; Ritter, J.B.; König, S.; Alt, R.; Reichl, U. Substitution of Glutamine by Pyruvate to Reduce Ammonia Formation and Growth Inhibition of Mammalian Cells. *Biotechnol. Prog.* 2008, 21, 58–69. [CrossRef]
- Mulukutla, B.C.; Mitchell, J.; Geoffroy, P.; Harrington, C.; Krishnan, M.; Kalomeris, T.; Morris, C.; Zhang, L.; Pegman, P.; Hiller, G.W. Metabolic engineering of Chinese hamster ovary cells towards reduced biosynthesis and accumulation of novel growth inhibitors in fed-batch cultures. *Metab. Eng.* 2019, 54, 54–68. [CrossRef] [PubMed]
- 54. Jeremy, M.; Berg, J.L.T. The Citric Acid Cycle. Biochemistry, 5th ed.; W H Freeman: New York, NY, USA, 2002.
- 55. Möller, J.; Bhat, K.; Guhl, L.; Pörtner, R.; Jandt, U.; Zeng, A. Regulation of pyruvate dehydrogenase complex related to lactate switch in CHO cells. *Eng. Life Sci.* 2021, 21, 100–114. [CrossRef] [PubMed]
- 56. Gupta, S.K.; Srivastava, S.K.; Sharma, A.; Nalage, V.H.H.; Salvi, D.; Kushwaha, H.; Chitnis, N.B.; Shukla, P. Metabolic engineering of CHO cells for the development of a robust protein production platform. *PLoS ONE* **2017**, *12*, e0181455. [CrossRef] [PubMed]
- 57. Dhiman, H.; Gerstl, M.P.; Ruckerbauer, D.; Hanscho, M.; Himmelbauer, H.; Clarke, C.; Barron, N.; Zanghellini, J.; Borth, N. Genetic and Epigenetic Variation across Genes Involved in Energy Metabolism and Mitochondria of Chinese Hamster Ovary Cell Lines. *Biotechnol. J.* **2019**, *14*, 1800681. [CrossRef] [PubMed]
- 58. Klinger, C.; Trinkaus, V.; Wallocha, T. Novel Carbon Dioxide-Based Method for Accurate Determination of pH and pCO2 in Mammalian Cell Culture Processes. *Processes* **2020**, *8*, 520. [CrossRef]