Computers and Chemical Engineering 168 (2022) 108059

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Optimising bioreactor processes with in-situ product removal using mathematical programming: A case study for propionate production

Lucas Van der Hauwaert [a,*], Alberte Regueira [a,c,d], Ludwig Selder [b], An-Ping Zeng [b], Miguel Mauricio-Iglesias [a]

[a] CRETUS, Department of Chemical Engineering. Universidade de Santiago de Compostela, Spain. Rúa Lope Gómez de Marzoa, s/n. 15782, Spain
[b] Institute of Bioprocess- and Biosystems Engineering, Hamburg University of Technology, Denickestr. 15, 21073 Hamburg, Germany
[c] Center for Microbial Ecology and Technology (CMET), Ghent University, Coupure links 653, B-9000 Ghent, Belgium
[d] Center for Advanced Process Technology for Urban Resource recovery (CAPTURE) (www.capture-resources.be), Belgium

ARTICLE INFO

Keywords:
Kinetic modelling
Product inhibition
Downstream processing
Design under uncertainty
Biotechnology
Operational design

ABSTRACT

Designing and operating bioreactors with in-situ product removal (ISPR) can be challenging, particularly in discontinuous systems, where the ISPR and substrate feeding need to be effectively scheduled. Mathematical models can help assess different scheduling regimes in the fermentation medium and provide a means to optimise the process. Focusing on a propionate production case study, a model of a co-culture batch fermentation with electrodialysis (the ISPR system), was developed. Using this model, the product yield and/or the productivity were maximised by 1) single objective optimisation maximising the product yield (0.49 $g_{propionate}$ /$g_{glucose}$) or productivity (0.75 $g_{propionate}$/L/h), 2) multi objective optimisation to pursue trade-off solutions between the yield and productivity and 3) a stochastic optimisation maximising the productivity robustly (0.64 $g_{propionate}$/L/h) to account for uncertainties associated to the model parameters. With this contribution it is demonstrated that, through mathematical models, ISPR can be implemented and adapted to the user's objectives.

1. Introduction

Many biotechnological processes can still not compete economically with traditional production methods. Currently, the economic feasibility of many fermentation pathways is limited due to diluted concentrations in the broth leading to high downstream costs (Woodley et al., 2008). On top of this, many fermentation processes have limited yields and limited productivities (Zeng, 2019). A key problem associated with these limitations is the inhibition of the microbial community by either the substrate or product formation (Santos et al., 2021).

To alleviate the problem of product inhibition, in situ product removal (ISPR) can be applied during the fermentation process. The goal of this technique is to separate the products, during the fermentation process, before their concentration becomes inhibitory to the microorganisms, increasing the potential yield, productivity and titre (López-Garzón and Straathof, 2014). Additionally, it can be seen as the first step of downstream processing, reducing the number of processes needed for the extraction of the valued products (Woodley et al., 2008).

Unfortunately, designing processes with ISPR still remains a challenging task, because the way ISPR is implemented can have critical effects on the outcome of these reactors. ISPR can either be installed directly inside the reactor (in-situ) or a stream of the fermentation broth can be directed out of the reactor towards the ISPR unit (ex-situ). The configuration depends on the properties of the product, the number of phases involved and the type of recovery process employed (Buque-Taboada et al., 2006). Regarding the operation of the ISPR, it faces challenges such as: 1) higher risks for contamination of the microbial community due to longer operational times which can result in unexpected process oscillations (e.g., a slower product formation due to a slower growth rate of the biomass). In the case of ex-situ recovery processes the risk of contamination is also substantial; 2) large energy consumptions of the extraction procedures and; 3) attaining a maximum product recovery as many ISPR techniques do not completely extract the targeted product (Van Hecke et al., 2014; Woodley et al., 2008).

To facilitate the design or process optimisation of ISPR systems, mathematical modelling is an invaluable tool, minimizing the need for experiments and avoiding costly trial-and-error approaches (Van Hecke et al., 2014). Several works have already created and used models to optimise various reactor systems implementing different ISPR techniques (Table 1). For example, Byun et al. (2020) modelled and optimised a continuous acetone-butanol-ethanol (ABE) fermentation

* Corresponding author.
E-mail address: lucas.vanderhauwaert@usc.es (L. Van der Hauwaert).

https://doi.org/10.1016/j.compchemeng.2022.108059
Received 23 February 2022; Received in revised form 26 October 2022; Accepted 1 November 2022
Available online 6 November 2022

Nomenclature

Parameters	*Units*
km_{ba}	Maximum substrate uptake rate of *Bacillus* $gCOD_{glu}$ $gCOD_x^{-1}$ h^{-1}
km_{ve}	Maximum substrate uptake rate of *Veillonella* $gCOD_{lac}$ $gCOD_x^{-1}$ h^{-1}
km_{ba_yes}	Maximum yeast extract uptake rate of *Bacillus* $gCOD_{ye}$ $gCOD_x^{-1}$ h^{-1}
km_{ve_ye}	Maximum yeast extract uptake rate of *Veillonella* $gCOD_{ye}$ $gCOD_x^{-1}$ h^{-1}
Ks_{glu}	Affinity constant glucose uptake gCOD L^{-1}
Ks_{lac}	Affinity constant lactate uptake gCOD L^{-1}
Ks_{yeb}	Affinity constant yeast extract uptake *Bacillus* gCOD L^{-1}
Ks_{yev}	Affinity constant yeast extract uptake *Veillonella* gCOD L^{-1}
$Kdec_{ve}$	Decay rate coefficient of *Veillonella* h^{-1}
Y_{glu}	Yield of glucose to biomass *Bacillus* $gCOD_X$ $gCOD_{glu}^{-1}$
Y_{lac}	Yield of lactate to biomass of *Veillonella* $gCOD_X$ $gCOD_{lac}^{-1}$
Y_{yeb}	Yield of yeast extract to biomass of *Bacillus* $gCOD_X$ $gCOD_{ye}^{-1}$
Y_{yev}	Yield of yeast extract to biomass of *Veillonella* $gCOD_X$ $gCOD_{ye}^{-1}$
KI_{lac}	Lactate inhibition constant gCOD L^{-1}
KI_{pro}	Propionate inhibition constant gCOD L^{-1}
KI_{ace}	Acetate inhibition constant gCOD L^{-1}
cf_1	Conversion factor: gCOD lactate to gCOD propionate $gCOD_{pro}$ $gCOD_{lac}^{-1}$
cf_2	Conversion factor: gCOD lactate to gCOD acetate $gCOD_{ace}$ $gCOD_{lac}^{-1}$
Inh	Inhibition term of propionate and lactate on *V. criceti* (−)

Abbreviations

COD	Chemical oxygen demand
ISPR	In situ product removal
REED	Reverse enhanced electro dialysis

process with continuous ex situ recovery by adsorption. With the obtained model the authors try to maximise productivity, yield, and minimise the sugar loss. Similarly Lin and Wang (2008) and Z. Shi et al. (2005) also tried to maximise the productivity with the use of optimisation algorithms applied to the reactor models. Note that in almost all of these examples the extraction procedure is continuous. In other words, by modelling a system that incorporates ISPR, it is possible to find optimal strategies that maximise certain operational performance indexes (e.g., product yield, productivity or titre).

Focusing on attaining a maximum product recovery, the operational design needs to be carefully chosen, certainly for systems implementing ISPR discontinuously. For example, how long should an extraction cycle last, how many extraction cycles should there be, when is the best moment to activate the extraction process to avoid inhibition phenomena, when should feeding be introduced with respect to the extraction cycles so a maximum amount of product can be extracted, etc. Answers to these questions are scarce in literature and deserve further attention.

Experimental works that implement ISPR discontinuously are that of Wang et al. (2012) and Liu et al. (2021). The authors always choose to implement the extraction cycles at the same time intervals or when a certain amount of product is produced. However different methods of scheduling these extraction cycles are not explored. In this work it is hypothesised that the scheduling of the extraction cycles can have a large effect on the performance of the reactor.

In this context, the goal of this work is to optimise the experimental system of Selder et al. (2020) to find operational strategies using a mathematical model. The experimental system is a co-culture fermentation in a batch reactor set up, featuring electrodialysis as ISPR. This reactor produces propionate from glucose, with lactate as an intermediate product, where feeding and ISPR are applied discontinuously. From the solutions of the optimisation problems, suitable operational designs relating to the scheduling of the feed and extraction cycles can be proposed, that maximise the productivity and/or the yield.

To find efficient scheduling practices for this system the following workflow is proposed: 1) develop a robust mechanistic model describing the fermentation 2) validate the model by comparing a simulation of an experiment to independent experimental data and 3) define corresponding optimisation problems depending on the needs of the operator. An approach to account for model and parametric uncertainty when defining an optimisation problem, is also showcased. This approach can be used for situations where a robust implementation (i.e., an operation with little associated risk) is needed to obtain a desired objective. A detailed description of this workflow is included in the methodology so that it can be adapted to other ISPR featuring bioprocesses.

This case study is particularly challenging because: 1) The model needs to describe a co-culture fermentation instead of a pure culture, 2)

Table 1
Literature overview of previous models and experimental systems incorporating In-situ product removal to produce fermentative products.

Target compound	Extraction technique	Productivity (g/L/h)	Description	Source
Lactate	Absorption column (continuous)	4.35	ISPR used for pH control not to reduce inhibition. Model was used to optimise of the final lactate concentration by varying the size of the cell immobilizing cubes and the absorption loading	Sun et al. (1999)
Butanol	Flash extraction (continuous)	26.0	Optimising the concentration feed, dilution rate and number of staged to find trade-offs between productivity, energy requirements and product purity	Z. Shi et al. (2005)
Lactate	Extraction with organic solvent (continuous)	120.0	Multi-objective optimisation using Fuzzy Optimisation. Optimising the productivity, the overall glucose conversion and the reactor yield. The operational conditions that where varied: the feed concentrations, the bleed ratio, the fermenter volume ratio, and the flow rate ratio.	Lin and Wang (2008)
Phenylethanol / Phenylethylacetate	Membrane extraction (continuous)	0.25	No optimisation applied. A model created incorporating ISPR	Adler et al. (2011)
Propionate	Anion Exchanger (discontinuous)	0.37	No model made. Comparison of Anion Exchanger-Based (the ISPR) in direct and indirect contact of cells in a batch fermentation.	Wang et al. (2012)
Acetate, Butanol, Ethanol	Absorption column (continuous)	3.75	Optimising productivity using Multi-objective optimisation. The feeding rate, feed concentration and circulation rate are considered as the operating variables. They are optimised with respect to the objectives: maximising the productivity, maximising the product yield, or minimising the substrate loss.	Byun et al. (2020)
ε-Poly-L-lysine	Cationic ion-exchange resin (discontinuous/ continuous)	0.39	Experimental system comparing discontinuous and continuous ISPR. No model made	Liu et al. (2021)

unlike most ISPR systems, feeding and extraction are applied discontinuously (as seen in Table 1), potentially extracting the substrate and 3) depending on the objective of the operator there are different ways to operate the system.

With this in mind the paper is organised as followed: First the methods for model development, including calibration and validation, are described. Then, the methods for optimisation are covered, featuring different levels of complexity depending on the user's objectives and requirements of solution robustness. Finally, (advanced) monitoring techniques are identified as a means to improve the real-time operation of ISPR.

2. Methods

To optimise the scheduling of the extraction cycles and feeding for the experimental system of Selder et al. (2020), a model that accurately simulates the system was developed using experimental data. The code generated to create this model, as well as the implementation of the optimisation problems can be found on the GitHub repository: https://github.com/llvdrhau/Co-culture_ISPR_Optimisation-. This experimental system is particularly interesting to model as the lactate platform is becoming increasingly more relevant, holding the potential to replace traditional petrochemical processes (Shahab et al., 2020).

2.1. Experimental system to model

In this system, the main goal is to produce propionate in a co-culture fermentation, enhanced by the implementation of ISPR. The batch reactor uses glucose as substrate and yeast extract as mineral nutrients, where glucose was sporadically added by spikes during the process. The two bacterial species used in this co-culture were *Bacillus coagulans*, which converts glucose into lactate, and *Veillonella criceti*, which ferments lactate into propionate and acetate in a 2:1 molar ratio (Fig. 1). In the beginning the reactor is only inoculated with *B. coagulans* and after 12 h *V. criceti* is added. Additionally, yeast extract was found to be consumed by both bacteria and was therefore incorporated in the model as an additional substrate. The bottleneck of this process lies with *V. criceti* which is sensitive to substrate (i.e., lactate) and product inhibition (Sabra et al., 2013). To tackle the problem of product inhibition ISPR was applied by means of reverse electro-enhanced dialysis (REED), which selectively removes monovalent organic acid ions (i.e., lactate, acetate and propionate) with an electrical field as driving force to an accumulation reservoir. The REED unit has a membrane area of 64 dm^2 with a current density of 400 A m^{-2} where the polarity of the membrane is reversed every 60 s. In the accumulation reservoir (volume of 1 L) NaOH is recirculated at a concentration of 0.3 M. In the reactor (volume of 1.2 L) the pH was controlled by a 6 M KOH and a 2.5 M H_2SO_4 solution at a value of 6.2. H_2SO_4 was only used during the electrodialysis due to the addition of OH^- to the fermenter during the extraction. REED was not used continuously throughout the process but was instead applied discontinuously during the fermentation in cycles of 2 h. Longer extraction cycles were not considered as lactate can also be removed during the extraction cycles. This could put *V. criceti* in a potentially substrate-deficient state and severely inhibit its growth.

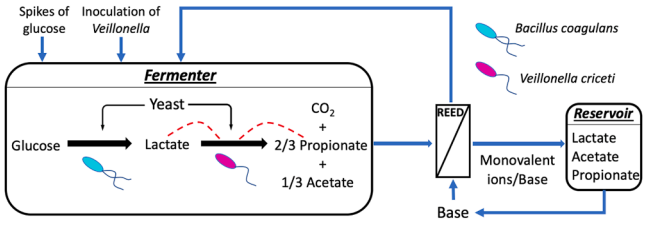

Fig. 1. Schematic representation of the co-culture fermentation in the work of Selder et al. (2020). - - - Red represents the reported inhibition on *V. criceti*.

2.2. Mathematical model

The mathematical model consists of mass balances of the different components (i.e., glucose, yeast, lactate, propionate, acetate, biomass of *V. criceti,* biomass of *B. coagulans* and inert dead biomass) in a discontinuous reactor (Eq. (1)-(2)). The system is considered as a perfectly mixed discontinuous fed-batch reactor where no volume leaves the reactor at a constant pH of 6.2. The changes in volume are due to feeding in the form of glucose spikes and the inoculation of the *V. criceti*. The model does not consider the changes in volume because of foaming, CO_2 solubilization, or the extraction of volume due to sampling (which was only 2 ml per sample). The model also does not consider water transport across the REED unit due to electro-osmosis, as this did not occur during

Table 2
Stoichiometric matrix and the corresponding kinetic expressions. In the matrix Y_{glu} in $gCOD_{glu}/gCOD_x$, Y_{yeb} in $gCOD_{ye}/gCOD_x$, Y_{yev} in $gCOD_{ye}/gCOD_x$ and Y_{lac} in $gCOD_{ye}/gCOD_x$ are the yields of the different substrates to the biomass of *B. coagulans* or *V. criceti*. Parameters cf_1 and cf_2 are parameters to convert gCOD of lactate to gCOD of propionate (0.77 $gCOD_{prop}/gCOD_{lac}$) and acetate (0.23 $gCOD_{prop}/gCOD_{lac}$) respectively and were calculated by stoichiometric relationships. In the equations S_{glu}, S_{lac}, S_{ace}, S_{ye}, X_{ba} and X_{ve} are the concentrations, in gCOD/L, of glucose, lactate, acetate, yeast extract, biomass of *B. coagulans* and biomass of *V. criceti* respectively. The maximum substrate uptake rate of *B. coagulans* and *V. criceti* are represented by km_{ba} ($gCOD_{glu}/gCOD_x/$ h) and km_{ve} ($gCOD_{lac}/gCOD_x$/h) respectively. The maximum yeast uptake of *B. coagulans* and *V. criceti* are represented by km_{ba_ye} and km_{ve_ye} ($gCOD_{ye}/gCOD_x$/h) respectively. Ks_{glu}, Ks_{lac}, Ks_{Yeb} and Ks_{Yev} in gCOD/L are the affinity constants for the glucose uptake by *B. coagulans*, lactate uptake by *V. criceti*, yeast uptake by *B. coagulans* and yeast uptake by *V. criceti* respectively. Finally, $kdec_{ve}$ (h^{-1}) is the decay coefficient of *V. criceti* and Inh_i the inhibition coefficient.

Process	Compounds Glucose (gCOD/L)	Lactate (gCOD/L)	Yeast (gCOD/L)	Propionate (gCOD/L)	Acetate (gCOD/L)	Biomass B. (gCOD/L)	Biomass V. (gCOD/L)	Inert mass (gCOD/L)	Kinetic equations
1. Glucose uptake	−1	($1-Y_{glu}$)	0	0	0	Y_{glu}	0	0	$r_{glu} = km_{ba}\left(\frac{S_{glu}}{Ks_{glu}+S_{glu}}\right)X_{ba}$
2. Yeast extract consumption *B. coagulans*	0	($1-Y_{yeb}$)	−1	0	0	Y_{yeb}	0	0	$r_{yeB} = km_{ba\_ye}\left(\frac{S_{Ye}}{Ks_{Yeb}+S_{Ye}}\right)X_{ba}$
3. Yeast extract consumption *V. criceti*	0	0	−1	($1-Y_{yev}$)	0	0	Y_{yev}	0	$r_{yeV} = km_{ve}\left(\frac{S_{Ye}}{Ks_{Yev}+S_{Ye}}\right)Inh_i \cdot X_{ve}$
4. Lactate uptake	0	−1	0	($1-Y_{lac}$) cf_1	($1-Y_{lac}$) cf_2	0	Y_{lac}	0	$r_{lac} = km_{ve\_ye}\left(\frac{S_{lac}}{Ks_{lac}+S_{lac}}\right)Ihn_i \cdot X_{ve}$
5. Decay of *V. criceti*	0	0	0	0	0	0	−1	1	$r_{dec} = kdec_{ve} \cdot X_{ve}$

the experiments. Considering the density of the feed and reactor holdup as constant, the mass balances of the different components can be defined by the following equations:

$$\frac{dC(t)}{dt} = D(t)\left(C^{spike} - C(t)\right) + R(t) - T(t) \quad (1)$$

Where $C(t)$ is the concentration of the different compounds in the reactor in gCOD/L at time t (ranging from 0 to 63 h), C^{spike} is the concentration of the spikes in gCOD/L and $T(t)$ is the transport flow across the REED extraction membrane in gCOD/L/h. $D(t)$ represents the dilution rate (h^{-1}) and is zero except when a spike is added to the fermentation reactor and is represented by the following equation:

$$D(t) = \frac{F(t)}{V(t)} \; with \; V(t) = V(0) + \int_0^t F(t')dt' \quad (2)$$

Where $V(t)$ is the reactor volume in L and $F(t)$ is feed flowrate in L/h). The initial values of the volume, glucose concentration and bacterial concentration, were obtained from the experimental set up of Selder et al. (2020) (1.2 L, 15 gCOD/L and 0.035 gCOD/L, respectively). All other states are zero at $t = 0$."

The reaction rate, $R(t)$ (Eq.(1)), is calculated by considering the following 5 processes: 1) glucose uptake by *B. coagulans* 2) uptake of yeast extract by *B. coagulans* 3) uptake of yeast extract by *V. criceti* 4) lactate uptake by *V. criceti* and 5) the decay of *V. criceti* (Table 2). These processes can be formulated as equations in the form of Monod kinetics, including inhibition terms for the processes related with *V. criceti*. In cases where the experimental mass balances could not explain the evolution of the compounds, the model included the consumption of yeast extract as represented by process 2) & 3). The decay of *B. coagulans* was not considered because there was not enough data to accurately determine the process. The kinetic expressions describing the reaction rates and the reactions stoichiometry can be found in Table 2. By multiplying the kinetic equations as an array of equations with the stoichiometric matrix, the reaction term R in Eq. (1) can be found for each compound.

In the model the spike and transport events are activated/deactivated by switches in the form of a pseudo-step function taking a value of 1 or 0 at the appropriate times. The pseudo-step function is a hyperbolic tangent function, which is a smooth function making it derivable for the ODE solver. To benchmark the different strategies and make them comparable to the experimental work, all the extraction intervals are equal and last 2 h. Note that this constraint on the extraction intervals would not be necessary if a more accurate representation of the decay of *V. criceti* could be achieved. If this were the case, the optimisation problem could also optimise the length of the intervals where the decay of *V. criceti* does not impact the system dramatically. Unfortunately, to the best of our knowledge there are no works that have estimated the decay of *V. criceti* and new costly experiments would have to be carried out to be able to estimate this parameter. Hence the same 2 h extraction cycles were taken similar to the experiments of Selder et al. (2020). During the extraction cycles, if the concentrations of the extracted compounds were smaller than 0.1 gCOD/L the extraction process is phase out for that specific compound. Otherwise, the extraction rates were set to fixed values representing the average extraction rates that were seen in the data.

The mass balances of each component were implemented in MATLAB (version r2021a) and solved with built-in numerical solvers of MATLAB (*ode45*). The model uses gCOD to be able to accurately determine electron balances. However, for the remainder of this work gCOD will be converted to grams so convenient comparisons can be made (see supplementary materials for the conversion table).

2.3. Inhibition term

In order to include the experimentally observed inhibition of *V. criceti* in the model, the kinetic equations of Table 2 include an inhibition term (Inh_i). Four potential expressions (Eq. (3)-(6)) were considered to describe the experienced inhibition which was either attributed to the combined or individual effects of lactate, propionate or acetate.

$$Inh_1 = \left(\frac{1}{1 + (S_{lac}/KI_{lac})}\right)\left(\frac{1}{1 + (S_{pro}/KI_{pro})}\right)\left(\frac{1}{1 + (S_{ace}/KI_{ace})}\right) \quad (3)$$

$$Inh_2 = \left(\frac{1}{1 + (S_{lac}/KI_{lac})}\right)\left(\frac{1}{1 + (S_{pro}/KI_{pro})}\right) \quad (4)$$

$$Inh_3 = \left(\frac{1}{1 + (S_{lac}/KI_{lac})}\right) \quad (5)$$

$$Inh_4 = \left(\frac{1}{1 + (S_{pro}/KI_{pro})}\right) \quad (6)$$

In these expressions KI_{lac}, KI_{pro} and KI_{ace} are the inhibition constants for lactate, propionate and acetate respectively. To determine the exact mechanism of inhibition experienced by *V. criceti*, the different mathematical expressions were tested during model calibration based on the best experimental data fit.

2.4. Model calibration and validation

To calibrate the parameter values of the kinetic equations, the following methodology was applied: first the parameters related to the two monocultures were calibrated with two datasets each, from batch experiments using individual monocultures (datasets 1.1 and 1.2 for *B. coagulans* and dataset 2.1 and 2.2 for *V. criceti*). The kinetic parameters for *V. criceti*, were calibrated using data from batch experiments involving glucose spikes. These datasets were also used to determine the most appropriate inhibition mechanisms and its related parameters (Eq. (3)-(6)). To calibrate the model to the data, Levenberg-Marquardt and trust-region-reflective methods were applied using a built-in solver *lsqnonlin* from MATLAB (Coleman and Li, 1996; Marquardt, 1963).

Similar to the works of Frutiger et al. (2016), robust parameters were obtained by applying the bootstrap method. Briefly explained, this method consists of the following steps: 1) A reference parameter estimation is carried out to estimate reference residuals (i.e., the error between the experimental and simulation data points); 2) assuming that these residuals have a uniform chance of occurring at each point, the reference residuals can be used to randomly create a new "artificial" set of data points (where the newly generated residuals have the same distribution profile as the previously mentioned residuals in step 1); 3) a new parameter estimation based on the new artificial dataset is carried out; and 4) Steps 2 & 3 are repeated for 500 iterations to find the distribution of the parameters. The parameter values used in the model are the mean values of these distributions (Efron, 1992).

For the parameter estimation, all the half saturation constants and the yield of the consumed yeast extract for both bacterial species, were given fixed values (Table 2) as available experimental data did not provide enough information to estimate them. This approach has a negligible impact on the predictive power of the simulations as substrate limiting conditions (were the half saturation constants become relevant) are barely explored. Additionally, identifiability issues with other parameters and overfitting of the parameters to the data, is avoided (Flotats et al., 2006; Seeliger et al., 2002). As overfitting is avoided, the residuals used during the bootstrap method are not underestimated and hence do not bias the parameter estimation.

To check and visualise if the simulations of the monoculture experiments sufficiently reproduce the experimental data to which the parameters are calibrated to (within a 95% confidence interval), a Monte Carlo uncertainty analysis was performed. This analysis performs simulations using different values of the estimated parameters sampled

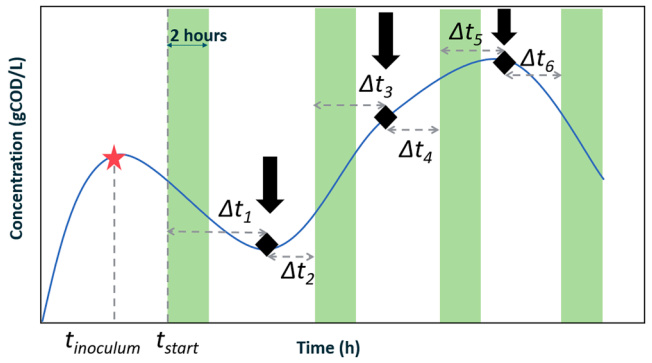

Fig. 2. Depiction of decision variables (vector z) in a typical operation: inoculation of *V. criceti* (★), extraction cycles (■) and glucose spikes (↓). The concentration of the spike is determined by C_{glu}^{spike}.

from the distributions obtained from the bootstrap method. In this case, sampling was performed by using Latin hypercube sampling (Helton and Davis, 2003). Because the model has parameters that are correlated (e.g., $kdec_{ve}$ and Y_{lac}), the sampling was carried out considering these correlations according to the method of Iman and Conover (1982). These simulations result in a distribution of simulation outputs that can be treated statistically and be visualised in a graph reflecting the uncertainty of the model (McClarren, 2018).

To quantitatively evaluate the fit of all the model predictions, the normalised root-mean-square error (NRMSE) of the model outputs (i.e., compound concentrations) were determined, calculated by Eq. (7). The NRMSE serves to aggregate the magnitudes of the errors in predictions for the various components into a single measure of predictive power (Hyndman and Koehler, 2006).

$$NRMSE = \sqrt{\frac{\sum_{t=1}^{n}\left(y_t^{exp} - y_t^{sim}\right)^2}{N}} \cdot nf^{-1} \tag{7}$$

In Eq. (7), N is the amount of recorded data, y_t^{exp} the experimental data at time t, y_t^{sim} the simulation output at time t and nf is a normalisation factor being the difference between the maximum and minimum value of the experimental data for each component. Once satisfactory parameters were found from the separate monoculture experiments, they were then combined into one cohesive co-culture model. To validate this model a simulation replicating a co-culture experiment with extraction cycles was compared to the experimental data.

2.5. Optimisation of the co-culture model

To find strategic solutions to operate the system, different optimisation problems and their objectives need to be defined. Various objective functions can be chosen and, in this case, the yield, the productivity or both at the same time were chosen as performance indexes. These performance indexes are of particular interest because of their importance to the economic performance of the process. The yield is an especially important metric when the substrate needs to be utilised in the most efficient way possible. This efficiency is necessary when either the obtained product or the used substrate is a very valuable resource. On the other hand, productivity is also an important parameter to maximise, certainly in cases where the obtained product is sold cheaply. In these cases, large bulks of product need to be sold regularly to ensure a healthy profit margin with respect to the capital investments. For this process the yield and productivity, are calculated as followed:

$$f_{yield}(z) = \frac{m_{pro}}{m_{glu}} = \frac{C_{pro}^{ext} \cdot V_{res}}{C_{glu,0} \cdot V_{reac} + V_{spikes} \cdot C_{glu}^{spike}} \tag{8}$$

$$f_{prod}(z) = \frac{m_{pro}}{t_{end}} = \frac{C_{pro}^{ext} \cdot V_{res}}{t_{end}} \tag{9}$$

The yield ($f_{yield}(z)$, in $g_{propionate}/g_{glucose}$) is calculated by dividing the mass of extracted propionate (m_{pro} in g) by the mass of glucose added during the entire process (m_{glu} in g) (Eq. (8)). m_{glu} is determined by adding the initial glucose mass in the rector and the glucose mass added by the spikes. To find the mass of extracted propionate the concentration in the reservoir (C_{pro}^{ext} in g/L), and the volume of the reservoir (V_{res} in L) must be known. The initial mass of glucose is calculated as the product of the initial volume of the reactor (V_{reac} in L) and the initial concentration of glucose ($C_{glu,0}$ in g/L). Finally, the mass of glucose added by the spikes is determined by multiplying the total added volume of the spikes (V_{spikes} in L) by the concentration of the glucose spikes (C_{glu}^{spike} in g/L). The productivity ($f_{prod}(z)$, in $g_{propionate}/L/h$) is found by calculating the mass of extracted propionate and dividing it by the operational time of the reactor (t_{end} in h) (Eq. (9)).

2.5.1. Decision variables

The objective functions are ultimately determined by the decision variables z which relate to the operational design of the reactor. For this process 9 decision variables were chosen as represented by the vector $z = [C_{glu}^{spike}, t_{inoculum}, t_{start}, \Delta t_1, \Delta t_2, \Delta t_3, \Delta t_4, \Delta t_5, \Delta t_6]$. Where C_{glu}^{spike} is the concentration of the glucose spikes, $t_{inoculum}$ the time at which *V. criceti* is inoculated, t_{start} the time at which the first extraction cycle is activated and Δt_1 to Δt_6 are time intervals between the spikes and the start of an extraction cycle. To make the operation comparable to that of the experimental system performed by Selder et al. (2020), 6 extraction intervals where chosen. 6 intervals were chosen because it is not known if more extraction cycles would cause fouling and severely limit the extraction rates. For each spike the same concentration (as determined by C_{glu}^{spike}) and volume (24 ml) is introduced to the system. Fig. 2 illustrates the decision variables in relation to a typical operation of the system. Additionally, in this system the spikes are always introduced between extraction cycles where the interval between the start of the extraction cycle and the following spike are longer than 2 h i.e., the time it takes for 1 extraction cycle.

2.5.2. Single objective optimisation

If only one objective for the process is to be maximised, a single objective optimisation problem is used (Eq. (10)-(11)):

$$\max_{z \in \Omega} f(z) \tag{10}$$

$$\Omega = \{z : h(x) = 0,\ g(x) \le 0,\ a \le z \le b\} \tag{11}$$

Where $f(z)$ is the objective function and Ω represents the feasible space of the decision variables z. The feasible space is defined by the (non-)linear equality constraints $h(z)$, the (non-)linear inequality constraints $g(z)$ and the lower and upper variable bounds a and b, respectively. The equality and inequality constraints are mainly put in place to ensure a minimum yield or productivity is achieved. The yield must be larger than 0.63 gCOD/gCOD in the case of optimising the productivity and the productivity must be larger than 0.60 gCOD/L/h in the case of optimising the yield. The bounds of the decision variables make sure that the spikes are always introduce between the extraction cycles.

To solve these single objective problems the built-in function of MATLAB *patternsearch* was used. This algorithm uses derivative-free methods, called generalised pattern search according to the works of Conn et al. (2009). As the optimisation problem is nonconvex, this algorithm was chosen pragmatically given its quick computational speed (compared to other derivative-free methods) while avoiding the numerical issues of derivative-based algorithms.

Table 3
Used parameters for the co-culture model with their respective confidence intervals (α = 0.05), units and source of the used parameters. The maximum consumption rates of the two substrates (glucose and yeast extract or lactate and yeast extract) were considered to be equal of each bacterial species (i.e., $km_{ba} = km_{ba\_ye}$, and $km_{ve} = km_{ve\_ye}$).

Abbreviation	Value	CI	Units	Source
km_{ba}	3.41	[3.22, 3.84]	$gCOD_{glu}\ gCOD_x^{-1}\ h^{-1}$	Calibrated
km_{ba_ye}	3.41	[3.22, 3.84]	$gCOD_{ye}\ gCOD_x^{-1}\ h^{-1}$	Calibrated
km_{ve}	10.33	[9.11, 11.73]	$gCOD_{lac}\ gCOD_x^{-1}\ h^{-1}$	Calibrated
km_{ve_ye}	10.33	[9.11, 11.73]	$gCOD_{ye}\ gCOD_x^{-1}\ h^{-1}$	Calibrated
KS_{glu}	0.10	/	$gCOD\ L^{-1}$	Assumed
KS_{lac}	0.27	/	$gCOD\ L^{-1}$	Assumed
KS_{yeb}	0.10	/	$gCOD\ L^{-1}$	Assumed
KS_{yev}	0.50	/	$gCOD\ L^{-1}$	Assumed
$Kdec_{ve}$	0.04	[0.04, 0.09]	h^{-1}	Calibrated
Y_{glu}	0.07	[0.07, 0.08]	$gCOD_X\ gCOD_{glu}^{-1}$	Calibrated
Y_{lac}	0.03	[0.02, 0.04]	$gCOD_X\ gCOD_{lac}^{-1}$	Calibrated
Y_{yeb}	0.07	/	$gCOD_X\ gCOD_{ye}^{-1}$	Assumed
Y_{yev}	0.03	/	$gCOD_X\ gCOD_{ye}^{-1}$	Assumed
KI_{lac}	10.04	[8.42, 12.01]	$gCOD\ L^{-1}$	Calibrated
KI_{pro}	7.34	[5.33, 10.67]	$gCOD\ L^{-1}$	Calibrated
cf_1	0.78	/	$gCOD_{pro}\ gCOD_{lac}^{-1}$	Calculated
cf_2	0.22	/	$gCOD_{ace}\ gCOD_{lac}^{-1}$	Calculated

2.5.3. Multi-objective optimisation

Strategies can also be found using multi-objective optimisation if multiple targets are desired to be optimised simultaneously (Eq. (12)-(13)). This optimisation technique is especially interesting to apply if the various objectives are conflictual (Custódio et al., 2011). In this case, the solution is not a single optimal result but rather a whole set of solutions called the nondominated set (also known as the Pareto set) representing the potential compromises amongst the objectives (Diwekar, 2020). This optimisation problem was carried out in MATLAB using the built-in function *paretosearch*. This algorithm uses the aforementioned generalised pattern search but instead of updating a single point per iteration, it updates an iterate list of nondominated points (i.e., the points that have the best rank and are closest to the Pareto front) (Custódio et al., 2011). The multi-objective optimisation can be defined as followed:

$$\max_{z \in \Omega} F(z) = (f_1(z), f_2(z), \ \ldots, \ f_n(z)) \tag{12}$$

$$\Omega = \{z : \ h(x) = 0, \ g(x) \leq 0, \ a \leq z \leq b\} \tag{13}$$

Where $F(z)$ represents the vector of objective functions $(f_1(z).f_2(z), \ \ldots, \ f_n(z))$.

2.5.4. Stochastic optimisation

If the strategy is sensitive to uncertain parameters, a more robust strategy can be found by stochastic optimisation. A strategy can be considered as not robust if the implementation of it is not always guaranteed to produce the desired result. This can be due to variations in the process variables such as the concentration of the initial inoculated bacteria, or uncertainty of the model parameters (e.g., the maximum substrate consumption rate of the microbial species). To design a process which is more robust against these uncertainties, a stochastic optimization problem can be solved. The type of algorithm used to solve this

Table 4
Normalised root mean squared errors of the mono-culture models after the bootstrap analysis.

Dataset	NRMSE of components				
	Glucose	Lactate	Propionate	Acetate	Biomass
Mono-culture *B. coagulans**	0.13	0.10	/	/	0.08
Mono-culture *V. criceti* *	/	0.13	0.08	0.13	0.15
Co-culture fermentation**	0.23	0.32	0.18	0.17	0.15

/ Indicates that this compound did not participate in a particular dataset.
* Indicates that the data set was used for parameter calibration
** Indicates the data was used for validation

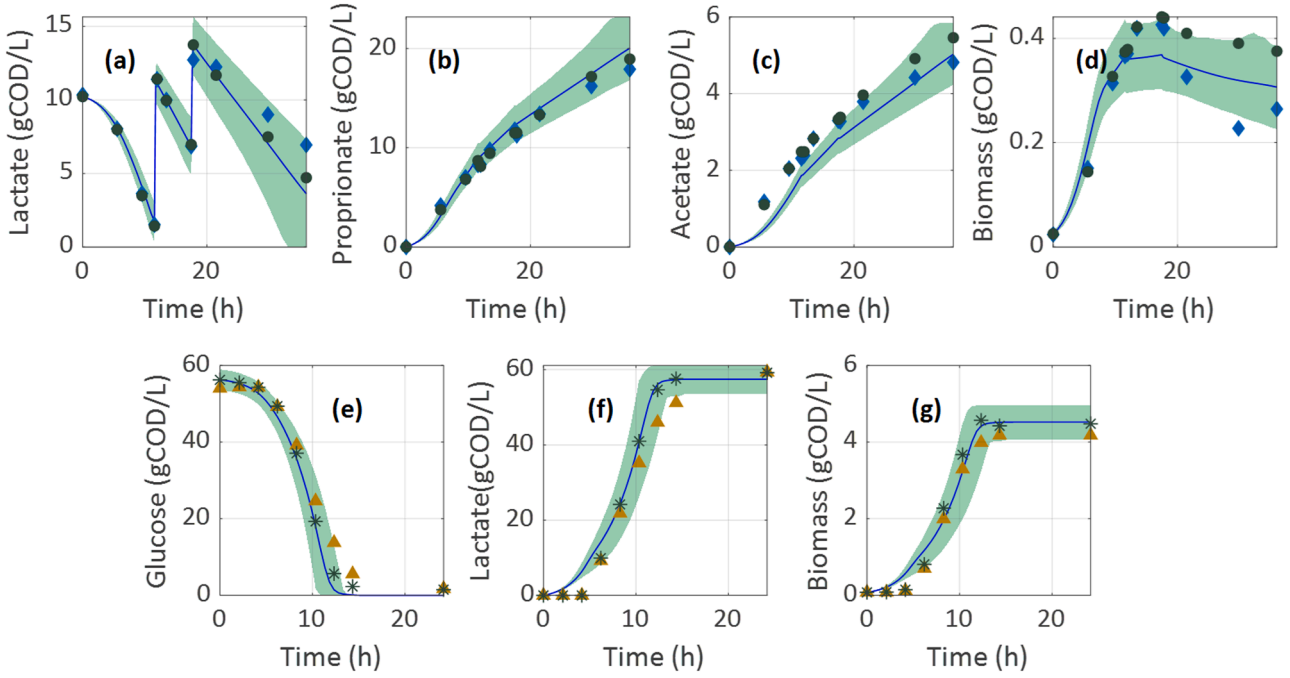

Fig. 3. Plots used to evaluate the fit of the calibrated parameters to the monoculture data. On plots (a) to (d) the concentration profiles of lactate, propionate, acetate and biomass of *V. criceti* plotted with data form dataset 2.1 (♦) and 2.2 (●) from the monoculture experiments with *V. criceti*. On plots (e) to (g) the concentration profiles of glucose, lactate and biomass of *B. coagulans* plotted with data form dataset 1.1 (▲) and 1.2 (*) from the monoculture experiments with *B. coagulans*. The green shading represents the 95% confidence interval of the simulation and were estimated by error propagation of the parameter uncertainties following a Monte Carlo procedure.

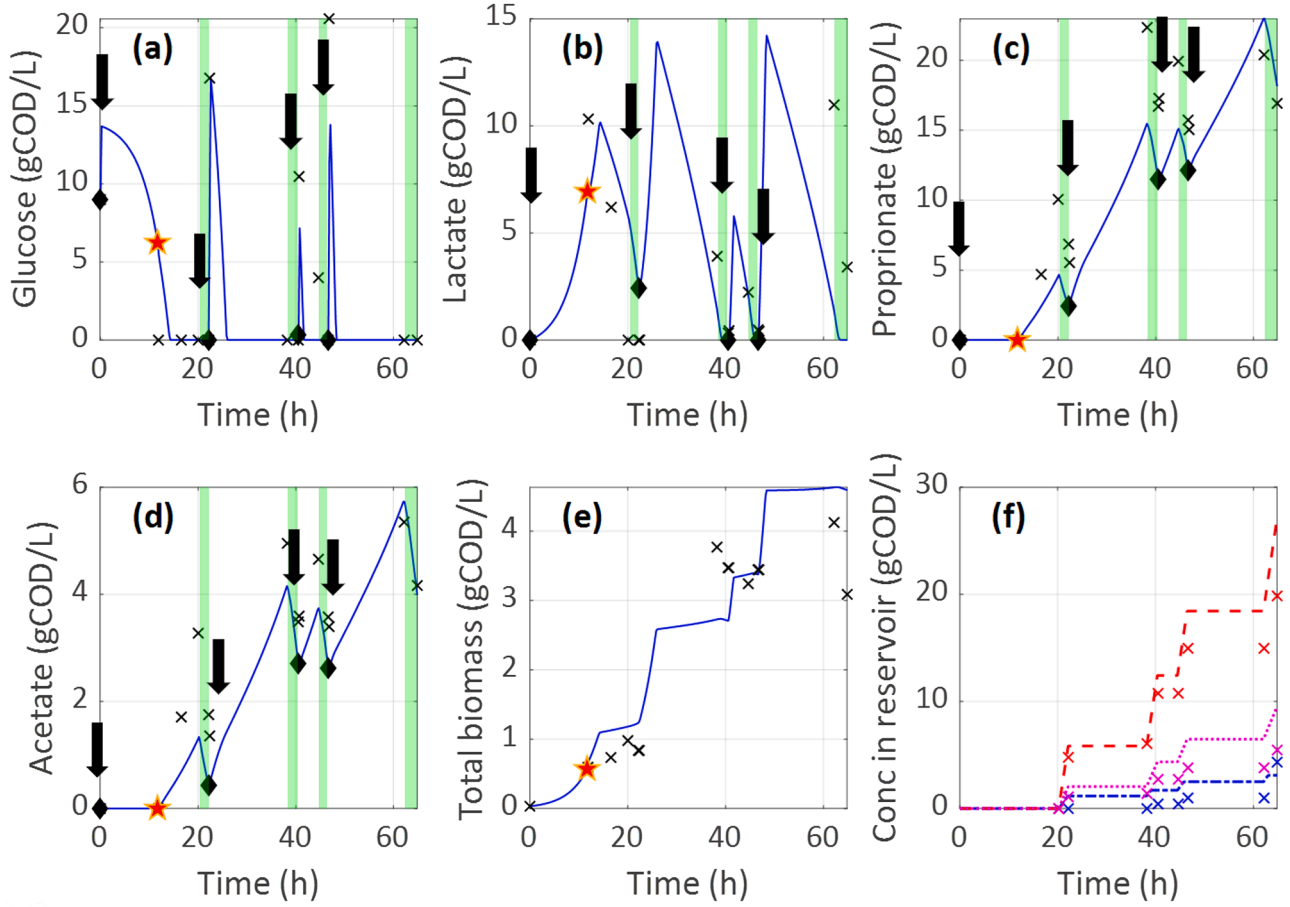

Fig. 4. Validation of the co-culture model on an independent dataset. Plots of glucose (a), lactate (b), propionate (c), acetate (d) and biomass (e) in the reactor can be seen. Data (x) glucose spikes (↓) extraction cycles (■). In subplot (f) the extracted compounds and the corresponding experimental data from the accumulation reservoir are shown, propionate (---), acetate (•••) and lactate (•-•-).

optimisation problem was the "Here and now" algorithm. In this algorithm, the objective function and constraints are expressed in terms of a probabilistic representation e.g., average value, variance or percentiles of a distribution. The lower the percentile chosen as the objective distribution, the more conservative and robust the attained solution is. This probabilistic representation is the result of a Monte Carlo simulation where the parameters are sampled as explained in Section 2.5 from the distributions obtained in the bootstrap method (Diwekar and Rubin, 1991). The optimisation problem is defined as followed:

$$\max_{z\in\Omega} F(z,\ u)\ =\ P_1(f(z,\ u)) \tag{14}$$

$$\Omega = \{z : P_2(h(x,u)) = 0, P_3(g(z),u) \geq 0,\quad a \leq z \leq b\} \tag{15}$$

where u is the vector of uncertain parameters sampled from the distributions attained by the bootstrap analysis and P_i represents the probabilistic representation (e.g., the tenth percentile from the distribution of $f(z, u)$) of the objective and the constraint functions.

3. Results and discussion

3.1. Model calibration and validation

The first stage in calibrating the biochemical model was achieving a right representation for the inhibition of *V. criceti*. For this, model predictions with different inhibition equations (Eq. (3)-(6)) were visually compared to the experimental data. The plots to compare the various simulations using different inhibition equations can be found in the supplementary materials (S1). When only considering lactate as the inhibiting compound (Eq. (5)), the simulation results displayed a clear mismatch to the experimental data. Considering all organic acids or only propionate as the inhibiting compounds (Eq. (3) & (6), respectively) resulted in a much better representation of the data. However, improvements concerning the lactate consumption could still be made. The inhibition equation considering lactate and propionate as inhibiting compounds (Eq. (4)), clearly displayed the best fit and was therefore chosen to describe the inhibition of *V. criceti* in the model. The compounds used in Eq. (4) to describe the Inhibition of *V. criceti*, are also hypothesised to be the cause of inhibition in the experimental work of Sabra et al. (2013), further confirming that the correct inhibition equation was chosen.

Robust parameters from the monoculture datasets were found by calibration with experimental data, using the Levenberg-Marquardt minimisation algorithm embedded in a bootstrap method (Table 3). The distributions of the parameters obtained by the bootstrap method can be found in the supplementary materials (S2).

Comparing the obtained parameters to literature data it can be seen that the yield for *B. coagulans* (0.05 $gCOD_X/gCOD_{glu}$) and a similar *Veillonella* species (0.03 $gCOD_X/gCOD_{lac}$) are comparable to the yields used in the model (0.07 $gCOD_X/gCOD_{glu}$ and 0.03 $gCOD_X/gCOD_{lac}$ respectively) (Seeliger et al., 2002; Glaser and Venus 2018). Other parameters such as the maximum growth rate for *B. coagulans* are considerably different. For example, in the works of Hidaka et al. (2010) a maximum growth rate of 12.8 $gCOD_{glu}/gCOD_x/h$ was found while a growth rate of 3.41 $gCOD_{glu}/gCOD_x/h$ was used in the model. The equation used to describe the lactate consumption in Hidaka et al.

Table 5
Operational design variables for the base case and the 4 strategies with their respective yields and productivities.

Strategy	Optimisation objective	Optimal Decision variables values								Optimisation outcome	
		Spike concentration (g/L)	Spike *Veillonella* (h)	1st extraction (h)	1st spike (h)	2nd extraction (h)	2nd spike (h)	3rd extraction (h)	3rd spike (h)	4th extraction (h)	Yield ($g_{propionate}$ /$g_{glucose}$)	Productivity ($g_{propionate}$ /L/h)
Base case	/	1000	11.8	20.2–22.2	22.4	38.2–40.2	40.7	44.6–46.6	46.8	62.2–64.2	0.23	0.27
Strategy 1	Yield	158.6	4.5	18.5 –20.5	20.6	23.3–25.3	25.3	27.3–29.3	29.3	31.1–33.1	0.49	0.40
Strategy 2	Productivity non-robust	204.4	6.7	10.2–12.5	12.6	12.3–14.3	14.4	14.5–16.5	16.6	16.7–18.7	0.45	0.75
Strategy 3	Productivity & yield	178.6	5.6	10.4–12.4	12.5	16.6–18.6	18.7	20.0–22.0	22.1	22.9 –24.9	0.47	0.56
Strategy 4	Productivity robust	195.9	7.5	12.3–14.3	15.0	15.3–17.3	17.5	17.7–19.7	19.7	19.9–21.9	0.47	0.64

(2010) is considerably more complex and identifiability issues between other parameters might explain the different obtained values (Alberton et al., 2013).

To evaluate the goodness of fit, plots (Fig. 3) to compare the simulation and experimental data were made and the NRMSE was determined (Table 4). Visually, from Fig. 3, it can be seen that in most cases the experimental data points fall in the green shading, representing the 95% confidence interval of the model solution. Model mismatch is only significant for the acetate profile (Fig. 3c), indicating that the model cannot describe acetate production with the same level of confidence. However, considering that acetate is not of primary interest in our evaluation of the system, this flaw in the model is not of much concern. In Table 4 it can also be seen that the NRMSE of the calibration datasets are all quite low giving us confidence in the obtained model parameters.

To have a model able of describing the behaviour of the co-culture system, the kinetic parameters estimated for each individual strain were combined. The complete model was then validated on an independent dataset (Fig. 4). Comparing the RSME of the validation and calibration datasets (Table 4), it can be seen that the values are in the same order of magnitude, indicating that overfitting is not significant. This further confirms our confidence in the obtained parameters from the experimental data. Visually, from Fig. 4, it can be seen that for the glucose profile, the two last spikes of glucose don't reach the datapoints. For the lactate profile the simulation clearly follows the data up until the last 2 data points, while for acetate and propionate the simulation slightly underestimates the concentrations in the beginning of the simulation. Finally, for the biomass concentration and the components in the reservoir the overall trend is followed, although the simulation becomes less accurate towards the end of the simulation. These deviations could be attributed to 2 factors. The first is the fact that the extraction rate can vary in reality. A more realistic and accurate approach to model the extraction rates would be to consider the ions and base concentrations on both sides of the membrane, which oscillate with time. In this way the Donnan exclusion effect could be partially accounted for. I.e., as more of the organic acids are accumulated in the reservoir, the charge gradient over the membrane weakens and a concentration gradient starts pushing the organic acids from the highly concentrated reservoir, back to the reactor with (potentially) lower concentrations. This effect reduces the possible extraction rate over time that can otherwise be obtained (Yaroshchuk, 2000). Considering a variable extraction rate is especially interesting in situations where the concentration of lactate is not low or moderately low ($>$ 4 g/L). As can be seen in Fig. 4 and in most of the optimised simulations (see Section 3.2), the lactate concentrations in the reactor during extraction, are mostly very low. As a result, the assumption that the lactate removal rate is constant and low is almost always valid. For conditions where this is not the case a variable extraction rate, which is dependent on the concentration of organic acids, would be preferred. Additionally such a model would be particularly appropriate for analysing membrane limitations, process control, monitoring or the REED design like e.g., in the works of Prado-Rubio et al. (2011). However, the pragmatic approach followed here (i.e., a constant extraction rate, Eq. (1)) was assessed as sufficient for a first approach to the operational design of the fermentation process (i.e., the scheduling of extraction and feeding cycles). Future works should focus on modelling variable extraction rates for cases where a more detailed model is required e.g., for process control.

Another reason for model mismatch is that the COD (or electron) balance of the experimental data did not always comply during the whole experiment resulting in data points that, in all likelihood, have a significant error. Despite this, it can be claimed that the proposed model is able to describe the experimental process satisfactorily and with robustness.

The simulation, resulting in Fig. 4, obtains a productivity and yield of 0.27 $g_{propionate}$ /L/h and 0.23 $g_{propionate}/g_{glucose}$ respectively (Table 5). This simulation can be seen as the base case, roughly representing the experimental setup of Selder et al. (2020). The base case can be used as a

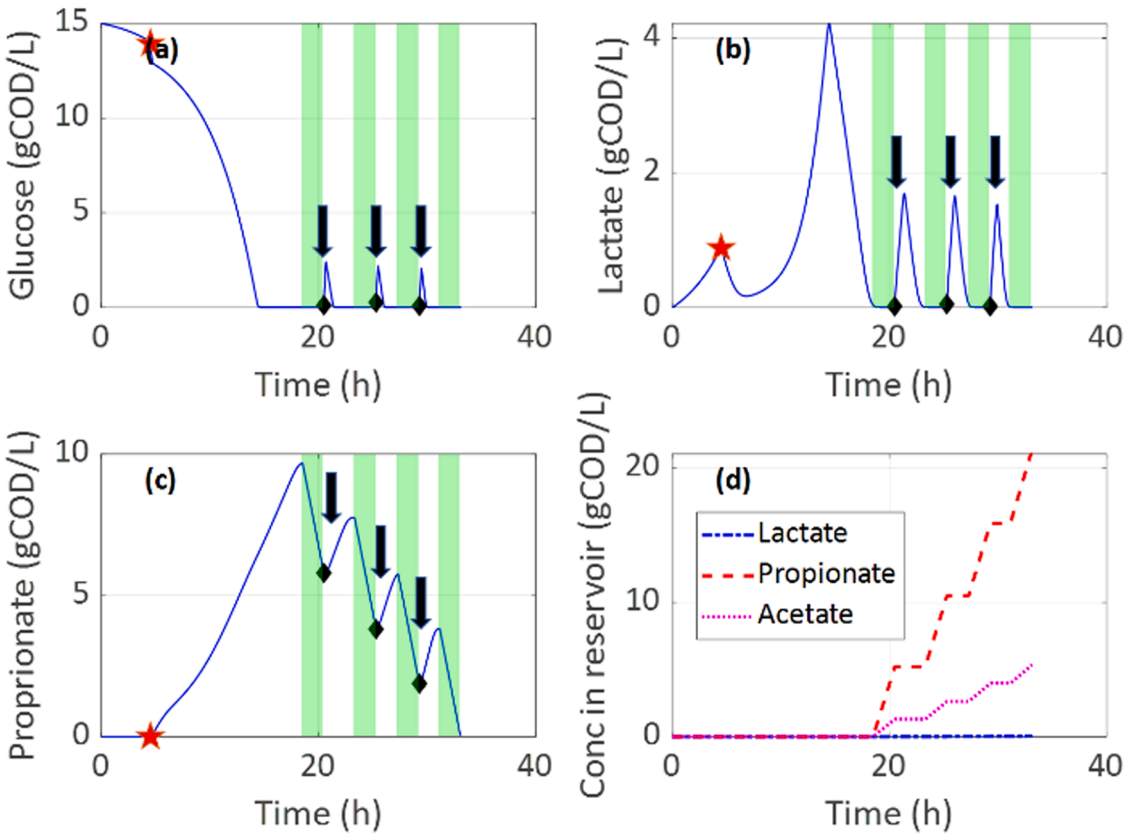

Fig. 5. Simulation of strategy 1 from the single objective optimisation optimising the yield. On plots (a) to (c) the concentration of glucose, lactate, and propionate can be seen. The inoculation of *V. criceti* is represented by ★, extraction cycles by ■ and glucose spikes by ↓. Plot (d) is the concentration of the components in the reservoir.

frame of reference to compare the results of the various optimisation techniques.

3.2. Optimisation problems

With the obtained model, three optimisation problems were applied: a single objective optimisation, a multi-objective optimisation and finally a stochastic optimisation problem. These optimisation problems lead to four different strategies serving different proposes and with different levels of robustness. The operational design (i.e., the concentration of the glucose spikes, the time when *V. criceti* is inoculated and the scheduling of the spikes and extraction cycles) and the obtained yield and productivity for each strategy can be found in Table 5.

3.2.1. Single objective optimisation: optimising the product yield

A single objective optimisation was used to maximise the yield of the system resulting in strategy 1. This strategy is characterised by starting all the extraction cycles when the substrate and intermediates (glucose and lactate) are consumed (Fig. 5a-b). Each time an extraction cycle ends a spike of substrate (glucose) is immediately added to the system. In the first few hours of this simulation the maximum concentration of lactate is reached, which is just over 4 $gCOD_{lactate}/L$, and the maximum propionate concentration is reached, which is just under 10 $gCOD_{propionate}/L$ (Fig. 5b-c). Because of these concentrations *V. criceti* is affected by inhibition, primarily caused by the high propionate concentration, prolonging the time it takes to deplete lactate. However, for each spike, slightly less time is needed to deplete lactate because of the removal of the inhibiting propionate and the increasing concentration of *V. criceti* biomass.

This method of scheduling the spikes and extractions (Table 5, Fig. 5) leads to a very high yield of 0.49 $g_{propionate}/g_{glucose}$, which represents 96% of the theoretical maximum of the system (0.51 $g_{propionate}/g_{glucose}$). The productivity reached for this strategy is 0.40 $g_{propionate}/L/h$.

The biggest advantage of this strategy is that it can be made robust and easy to implement: the strategy relies on the fact that lactate is depleted during each extraction cycle. So, to make this strategy robust/conservative, more time can be taken before activating the extraction cycles after a spike to ensure all the substrate is depleted. In other words, this strategy is best implemented when complete substrate utilisation is prioritised and a simple operational procedure is desired.

The main drawback of this strategy is its very long operational time, which leads to rather low productivities (0.40 $g_{propionate}/L/h$). According to Rodriguez et al. (2014), a general goal to reach in terms of productivity would be 1 g/L/h, to keep capital costs manageable for a commodity chemical such as propionate. In this work the authors based their economic analysis on a pure culture batch reactor with glucose as substrate (including the cost of up and down stream processing) without ISPR. Thus, a cost which Rodriguez et al. (2014) does not take into account, is the cost of energy used during REED extractions. However, since the extraction cycles are considered as a part of the downstream processing, it can be said that a similar productivity needs to be achieved.

To get closer to this goal a single optimisation problem maximising the productivity, was implemented resulting in strategy 2. Strategy 2 almost doubles the productivity (0.75 $g_{propionate}/L/h$) compared to the previous solution but also has a yield that is roughly 10% lower than strategy 1 (0.45 $g_{propionate}/g_{glucose}$) representing 88% of the theoretical maximum of the system). The simulation resulting in this solution starts the first extraction cycle very early and schedules the following extraction cycles and spikes very close to each other (Table 5, Fig. 6). As a result of the quick succession of extraction cycles, the concentration of lactate and propionate in the reactor stays beneath 2 $gCOD_{lactate}/L$ and 3

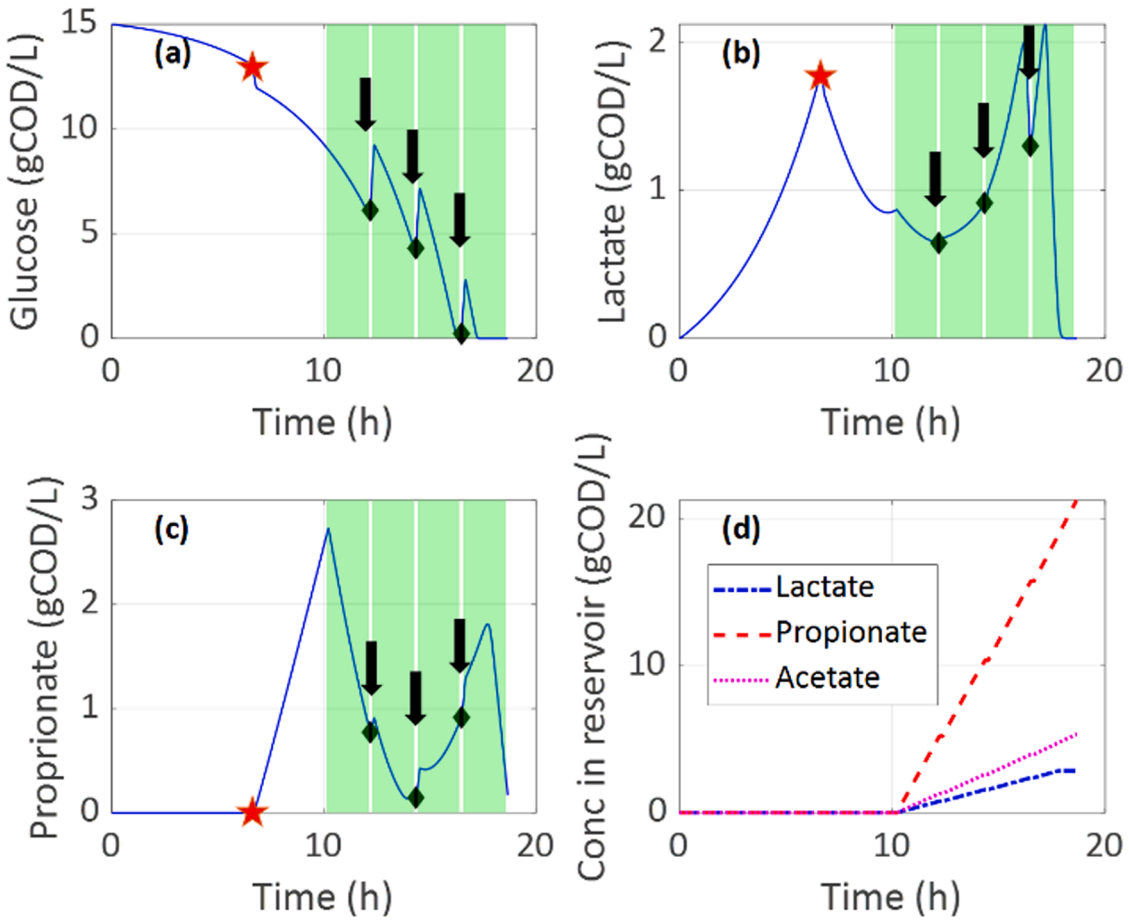

Fig. 6. Simulation of strategy 2 from the multi-objective optimisation optimising the productivity. On plots (a) to (c) the concentration of glucose, lactate, and propionate can be seen. The inoculation of *V. criceti* is represented by ★, extraction cycles by ■ and glucose spikes by ↓. Plot (d) is the concentration of the components in the reservoir.

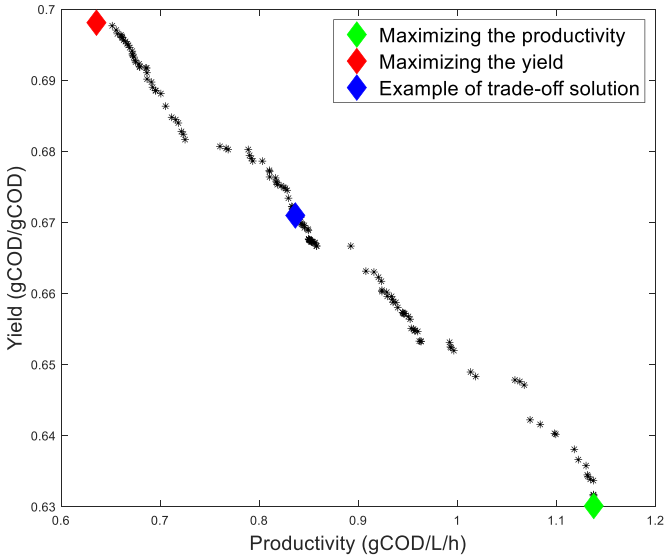

Fig. 7. Obtained Pareto front from the multi-objective optimisation, with strategy 1,2 and 3 indicated by ♦, ♦ and ♦ respectively.

$gCOD_{propionate}/L$ respectively (Fig. 6b-c). At these concentrations the effects of inhibition are not as significant compared to the other strategies, allowing for a quicker production of propionate.

This optimisation problem still does not reach the goal of 1 g/L/h. To do so, a parallel extraction unit or more extraction cycles need be implemented as the system is limited by the REED extraction rate. However, this comes with a significant increase in energy consumption. According to Selder et al. (2020) the energy requirement was 0.03 kWh/$g_{propionate}$. Assuming the energy requirements increase linearly, an extra 0.14 kWh would need to be applied to reach the goal of 1 $g_{propionate}$/L/h representing a 22% increase in energy consumption.

3.2.2. *Multi-objective optimisation: optimising the productivity and yield simultaneously*

In many cases, more than 1 performance index is important to a certain system. As previously identified the yield and productivity are both interesting objectives to improve for this system. To simultaneously maximise these performance indexes a multi-objective optimisation can be used. This optimisation is interesting in this context because solving to maximise the yield and productivity at the same time is conflictual. Either the reactor is operated with a quick turnover time to prioritise the productivity, which leaves part of the substrate unconsumed, or the reactor operation time is long to maximise substrate consumption, sacrificing the productivity. The multi-objective optimisation thus leads to the nondominated set of solutions representing the potential compromises amongst the objectives (i.e., yield and productivity).

Strategy 1 (with the highest yield in the nondominated set) and strategy 2 (with the highest productivity in the nondominated set) are solutions that can be seen at the two extreme ends of the nondominated set of solutions and can be visualised in the form of the Pareto front (Fig. 7). These strategies can be considered as the cases where no

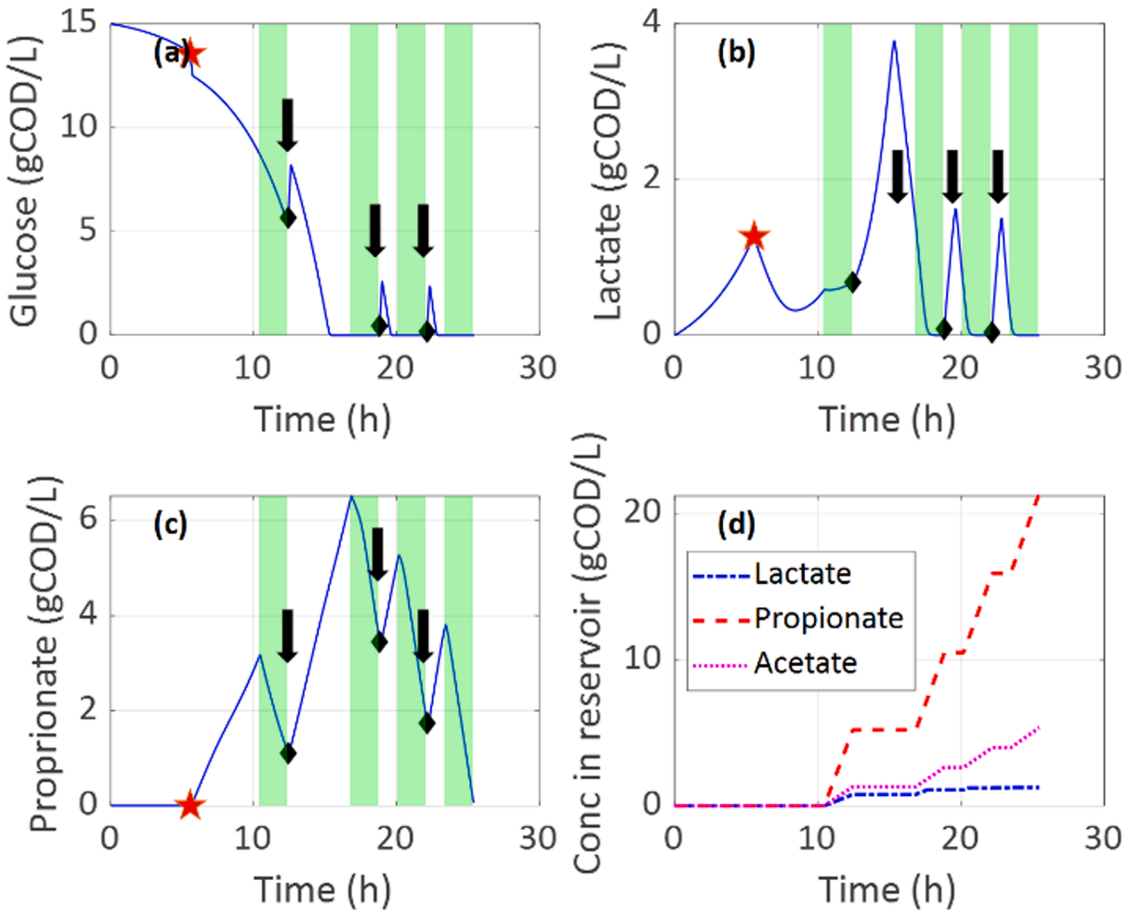

Fig. 8. Simulation of strategy 3 from the multi-objective optimisation with on plots (a) to (c) the concentration of glucose, lactate and propionate in the reactor, respectively. The inoculation of *V. criceti* is represented by ★, extraction cycles by ■ and glucose spikes by ↓. Plot (d) is the concentration of the components in the reservoir.

compromises are made and can thus also be found using the afore mentioned single objective optimisation (Section 3.2.1.).

Strategy 3, a solution found in the middle of the Pareto front (blue-diamond Fig. 7), illustrates a compromise solution (Table 5, Fig. 8) between the yield (0.47 $g_{propionate}/g_{glucose,}$ representing 92% of the theoretical maximum of the system) and the productivity (0.56 $g_{propionate}$/L/h). In this strategy, the first extraction cycle is activated at hour 10.4 without waiting for glucose and lactate to deplete. For the following extraction cycles, lactate is almost completely depleted and the spikes of glucose are always introduced immediately after the extraction cycles. The concentrations reached by lactate and propionate in this strategy fall slightly under 4 $gCOD_{lactate}$/L and slightly over 6 $gCOD_{propionate}$/L (Fig. 7b-c). The effects of inhibition are thus not so prominent as in strategy 1 but still likely influence the rate of substrate consumption (i.e., lactate), meaning that the following extraction cycles are slightly delayed by the inhibition.

Strategy 3 can be considered as a hybrid strategy, employing operational characteristics from strategy 1 (maximising the yield) as well as strategy 2 (maximising the productivity). The first extraction cycle starts early on, without waiting for substrate to be consumed, like strategy 2. The following cycles are spaced more apart like strategy 1, giving time for the intermediate substrate (lactate) to be consumed, thereby utilising it as efficiently as possible. With this hybrid approach, an improved productivity can be achieved that still has very high yields. There are many possible solutions from the Pareto front and it is up to the operator to choose the marginal weights given to each of the objectives. These weights should be chosen after performing a careful economic evaluation of the system.

An important fact to note for strategy 2 and 3 is that due to the uncertainty in the model and the inherent variability of biological systems (e.g., variations in initial bacterial concentration or slight changes in consumption rates), propionate might actually not be present in the reactor when the first extraction cycle starts. Only the intermediate (i.e., lactate) would then be extracted in this case, resulting in a significant loss of yield and productivity. As a result, it can be stated that these solutions cannot be considered as robust strategies. Intuitively it can be said that other solutions from the Pareto set favouring the productivity, are also not robust because they lead to a fast succession of spikes and extraction cycles. These tight strategies increase the likelihood of only extracting lactate due to the inherent model uncertainty. Therefore, care must be taken not to give too much weight to the productivity in a multi-objective optimisation if a robust strategy is pursued amongst the Pareto set.

3.2.3. *Stochastic optimisation: a robust strategy to maximise productivity*

To find a strategy that maximises the productivity and ensures a robust implementation at the same time, a stochastic optimisation problem was solved. To ensure a more conservative and robust strategy, a low percentile of the objective function distribution (in this case the tenth percentile, P10) was selected as the optimisation criterion, resulting in strategy 4 (Table 5, Fig. 9). Based on this selection, implementing this strategy would lead to a 90% probability of obtaining a productivity higher than the optimisation criterion, hence a risk-averse strategy. Compared to a risk-taking approach i.e., optimising the 90th percentile of the objective function, the productivity only increased by 4%. This increase in productivity it not worth the extra associate risk (i.

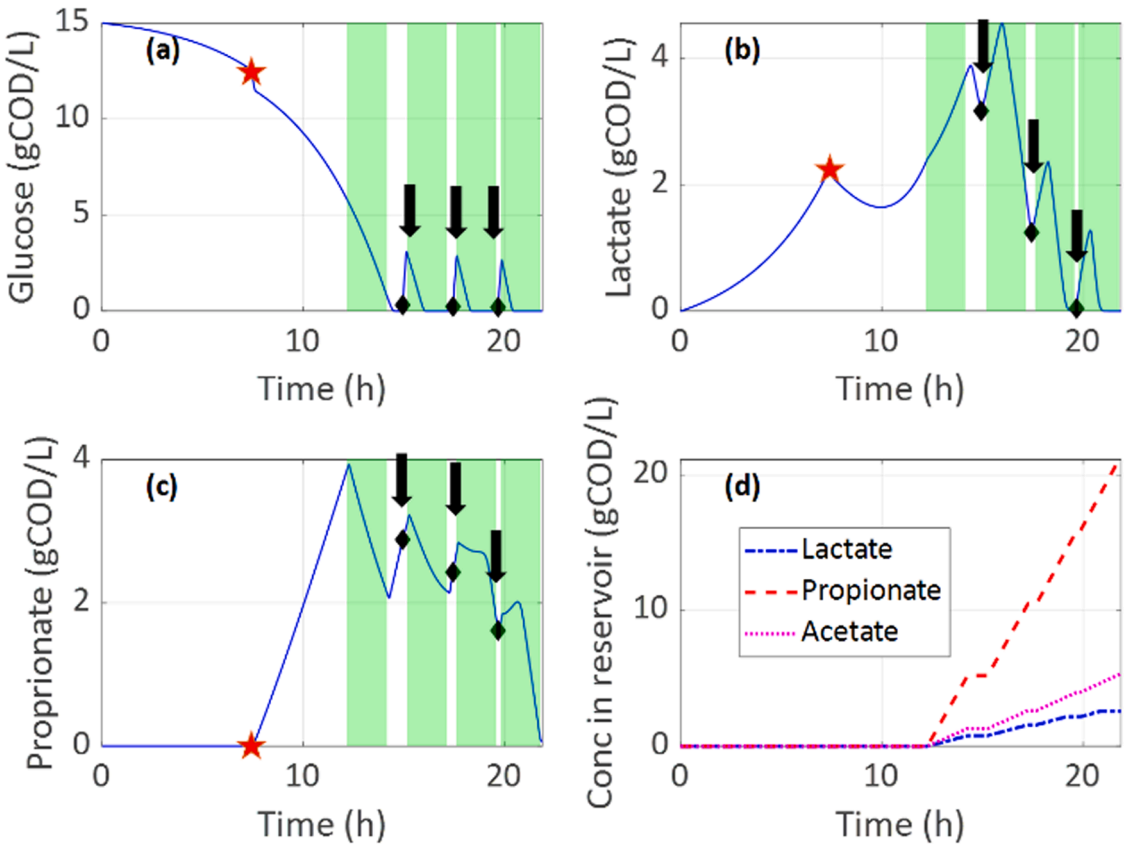

Fig. 9. Simulation of strategy 4 from the stochastic optimisation solution with on plots (a) to (c) the concentration of glucose, lactate and propionate in the reactor. The inoculation of *V. criceti* is represented by ★, extraction cycles by ■ and glucose spikes by ↓. Plot (d) is the concentration of the components in the reservoir.

e., only a 10% probability of obtaining a productivity higher than the optimisation criterion) which is why the more conservative result is the preferred solution. This robust strategy results in a productivity of 0.64 $g_{propionate}$/h/L and a yield of 0.47 $g_{propionate}$/ $g_{glucose}$. As with strategy 2, the concentration of lactate and propionate are low in this strategy and the effect of inhibition are smaller compared to other strategies, resulting once again in a faster production of propionate (Fig. 8b-c).

Compared to strategy 2, the first extraction cycle in this strategy takes place at hour 2. The rationale behind this is to give more time for the propionate concentration to increase, making it more likely that propionate is present in the reactor during the first extraction cycle. Additionally, the following extraction cycles and spikes are organised in a way that propionate has more time to be produced compared to strategy 2, where cycles are densely packed together. This can be observed by the fact that, in the beginning of the simulation the cycles are more spaced out, while at the end of the simulation the extraction cycles follow each other up more quickly. The reason for this is that lactate is more likely to already be converted to propionate due to the higher biomass concentrations at the end of the simulation. Another difference to the previous strategies is that the spikes of glucose are introduced right before extraction cycles take place.

This strategy has a productivity which is 0.11 $g_{propionate}$ /h/L lower than strategy 2 but has the distinct advantage of being robust to implement. Both strategies still do not reach the goal of 1 g/L/h but compared to the base case, significant improvements have been made where the productivity is at least doubled.

3.3. Integration of monitoring in optimal ISPR operations

Real-time data in the form of online measurements could be an important aid to implement the envisioned strategies and decrease the uncertainty of the model predictions by continuous calibration and optimisation. However, real-time monitoring of volatile fatty acids, biomass or other metabolites present in this system is not straightforward. Nevertheless, the use of soft-sensors, or data-driven monitoring systems can help obtain real-time information of the reactor broth thereby aiding the implementation of the strategies. Furthermore, whenever a mathematical model is available, state estimators, such as nonlinear versions of the Kalman filter (extended Kalman filter, unscented Kalman filter) can provide better monitoring and decrease the impact of noisy measurements (Lopez et al., 2020; Mohd Ali et al., 2015).

For the obtained strategies of the case study online monitoring could be implemented in the following ways: the key to implementing strategy 1 is to activate the extraction cycles when lactate is depleted. By monitoring when the lactate is almost consumed, the extraction cycles can be activated automatically with control systems ensuring complete substrate utilisation. For the strategies maximising the productivity (strategy 2 & 4), monitoring the propionate concentration can be very valuable. In these strategies the first extraction cycle can be activated when the concentration of propionate reaches a certain threshold, thereby guaranteeing that propionate is present in the reactor during extraction. This would make strategy 2, which was deemed non-robust, a lot more reliable to implement.

4. Conclusions

A modelling and optimisation-based approach to find various scheduling strategies (relating to the ISPR and feeding) for the experimental system of Selder et al. (2020), was successfully showcased. To accomplish this, a model was developed to describe the system producing propionate using a co-culture fermentation and ISPR. The parameters of this model were calibrated using datasets from monoculture experiments and successfully validated on an independent dataset that

described the entire process. Optimisation problems were then successfully implemented to this model to find several operational strategies. Single objective optimisation was used to find an optimal strategy that maximises the yield or productivity. With multi objective optimisation a whole set of trade-off solutions were found between these two performance indexes. Finally, stochastic optimisation was used to find robust strategies for the cases where model mismatch or inaccuracy in implementation could take place. In all the found strategies the yield and productivity were consistently improved upon compared to the base case and provided an operational rationale that can be adapted, by the ISPR designer. For future studies, the implementations of these strategies should be verified and with the collected data the model can also be further fine-tuned.

CRediT authorship contribution statement

Lucas Van der Hauwaert: Writing – original draft, Software, Data curation. **Alberte Regueira:** Conceptualization, Software, Writing – review & editing. **Ludwig Selder:** Data curation, Resources. **An-Ping Zeng:** Conceptualization, Resources. **Miguel Mauricio-Iglesias:** Supervision, Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The datasets and code generated and/or analysed during the current study are available on a GitHub repository: https://github.com/llvdrhau/Co-culture_ISPR_Optimisation- .

Acknowledgements

This work was supported by project ALQUIMIA (PID2019-110993RJ-I00) funded by the Agencia Estatal de Investigación Alquimia: Proyecto de I- D-i Programa Retos de la sociedad modalidad Jovenes investigadores convocatoria. A. Regueira would like to acknowledge the support of the Xunta de Galicia through a postdoctoral fellowship (ED481B-2021-012). The authors belong to the Galician Competitive Research Group ED431C-2021/37, cofounded by ERDF (EU).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2022.108059.

References

Adler, P., Hugen, T., Wiewiora, M., Kunz, B., 2011. Modeling of an integrated fermentation/membrane extraction process for the production of 2-phenylethanol and 2-phenylethylacetate. Enzyme Microb. Technol. 48, 285–292. https://doi.org/10.1016/j.enzmictec.2010.12.003.

Alberton, K.P.F., Alberton, A.L., Di Maggio, J.A., Díaz, M.S., Secchi, A.R., 2013. Accelerating the parameters identifiability procedure: set by set selection. Comput. Chem. Eng. 55, 181–197. https://doi.org/10.1016/j.compchemeng.2013.04.014.

Buque-Taboada, E.M., Straathof, A.J.J., Heijnen, J.J., van der Wielen, L.A.M., 2006. In situ product recovery (ISPR) by crystallization: basic principles, design, and potential applications in whole-cell biocatalysis. Appl. Microbiol. Biotechnol. 71, 1–12. https://doi.org/10.1007/s00253-006-0378-6.

Byun, H.-E., Kim, B., Lim, J., Lee, J.H., 2020a. Multi-objective optimization of operation of lignocellulosic acetone-butanol-ethanol fermentation with ex situ butanol recovery (ESBR). Comput. Chem. Eng. 140, 106915 https://doi.org/10.1016/j.compchemeng.2020.106915.

Byun, H.-E., Kim, B., Lim, J., Lee, J.H., 2020b. Multi-objective optimization of operation of lignocellulosic acetone-butanol-ethanol fermentation with ex situ butanol recovery (ESBR). Comput. Chem. Eng. 140, 106915 https://doi.org/10.1016/j.compchemeng.2020.106915.

Coleman, T.F., Li, Y., 1996. An interior trust region approach for nonlinear minimization subject to bounds. SIAM J. Optim. 6, 418–445. https://doi.org/10.1137/0806023.

Conn, A.R., Scheinberg, K., Vicente, L.N., 2009. Introduction to derivative-free optimization. SIAM J. Optim. https://doi.org/10.1137/1.9780898718768.

Custódio, A.L., Madeira, J.F.A., Vaz, A.I.F., Vicente, L.N., 2011. Direct multisearch for multiobjective optimization. SIAM J. Optim. 21, 1109–1140. https://doi.org/10.1137/10079731X.

Diwekar, U.M., 2020. Introduction to Applied Optimization, Springer Nature, Springer Optimization and Its Applications. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-55404-0.

Diwekar, U.M., Rubin, E.S., 1991. Stochastic modeling of chemical processes. Comput. Chem. Eng. 15, 105–114. https://doi.org/10.1016/0098-1354(91)87009-X.

Efron, B., 1992. Bootstrap Methods: another Look at the Jackknife. Ann Stat. 569–593. https://doi.org/10.1007/978-1-4612-4380-9_41.

Flotats, X., Palatsi, J., Ahring, B.K., Angelidaki, I., 2006. Identifiability study of the proteins degradation model, based on ADM1, using simultaneous batch experiments. Water Sci. Technol. 54, 31–39. https://doi.org/10.2166/wst.2006.523.

Frutiger, J., Marcarie, C., Abildskov, J., Sin, G., 2016. A comprehensive methodology for development, parameter estimation, and uncertainty analysis of group contribution based property models—an application to the heat of combustion. J. Chem. Eng. Data 61, 602–613. https://doi.org/10.1021/acs.jced.5b00750.

Glaser, R., Venus, J., 2018. Co-fermentation of the main sugar types from a beechwood organosolv hydrolysate by several strains of *Bacillus coagulans* results in effective lactic acid production. Biotechnol. Rep. 18 https://doi.org/10.1016/j.btre.2018.e00245.

Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Saf. 81, 23–69. https://doi.org/10.1016/S0951-8320(03)00058-9.

Hidaka, T., Horie, T., Akao, S., Tsuno, H., 2010. Kinetic model of thermophilic L-lactate fermentation by *Bacillus coagulans* combined with real-time PCR quantification. Water Res. 44, 2554–2562. https://doi.org/10.1016/j.watres.2010.01.007.

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22, 679–688. https://doi.org/10.1016/J.IJFORECAST.2006.03.001.

Iman, R.L., Conover, W.J., 1982. A distribution-free approach to inducing rank correlation among input variables. Commun. Stat. - Simul. Comput. 11, 311–334. https://doi.org/10.1080/03610918208812265.

Lin, H.-.T., Wang, F.-.S., 2008a. Fuzzy optimization of extractive fermentation processes including cell recycle for lactic acid production. Chem. Eng. Technol. 31, 249–257. https://doi.org/10.1002/ceat.200700396.

Lin, H.-.T., Wang, F.-.S., 2008b. Fuzzy optimization of extractive fermentation processes including cell recycle for lactic acid production. Chem. Eng. Technol. 31, 249–257. https://doi.org/10.1002/ceat.200700396.

Liu, S.-.R., Yang, X.-.J., Sun, D.-.F., 2021. Enhanced production of ε-poly-L-lysine by immobilized *Streptomyces ahygroscopicus* through repeated-batch or fed-batch fermentation with in situ product removal. Bioprocess Biosyst. Eng. 44, 2109–2120. https://doi.org/10.1007/s00449-021-02587-7.

López-Garzón, C.S., Straathof, A.J.J., 2014. Recovery of carboxylic acids produced by fermentation. Biotechnol. Adv. 32, 873–904. https://doi.org/10.1016/j.biotechadv.2014.04.002.

Lopez, P.C., Udugama, I.A., Thomsen, S.T., Roslander, C., Junicke, H., Mauricio-Iglesias, M., Gernaey, K.V., 2020. Towards a digital twin: a hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation. Biofuel. Bioprod. Biorefining 14, 1046–1060. https://doi.org/10.1002/bbb.2108.

Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11, 431–441. https://doi.org/10.1137/0111030.

McClarren, R.G., 2018. Introduction to Monte Carlo methods. Comput. Nucl. Eng. Radiol. Sci. Using Python 381–406. https://doi.org/10.1016/b978-0-12-812253-2.00024-8.

Mohd Ali, J., Ha Hoang, N., Hussain, M.A., Dochain, D., 2015. Review and classification of recent observers applied in chemical process systems. Comput. Chem. Eng. 76, 27–41. https://doi.org/10.1016/j.compchemeng.2015.01.019.

Prado-Rubio, O.A., Jørgensen, S.B., Jonsson, G., 2011. Reverse electro-enhanced dialysis for lactate recovery from a fermentation broth. J. Memb. Sci. 374, 20–32. https://doi.org/10.1016/j.memsci.2011.03.007.

Rodriguez, B.A., Stowers, C.C., Pham, V., Cox, B.M., 2014. The production of propionic acid, propanol and propylene via sugar fermentation: an industrial perspective on the progress, technical challenges and future outlook. Green Chem. 16, 1066–1076. https://doi.org/10.1039/C3GC42000K.

Sabra, W., Dietz, D., Zeng, A.P., 2013. Substrate-limited co-culture for efficient production of propionic acid from flour hydrolysate. Appl. Microbiol. Biotechnol. 97, 5771–5777. https://doi.org/10.1007/s00253-013-4913-y.

Santos, A.G., Albuquerque, T.L., Ribeiro, B.D., Coelho, M.A.Z., 2021. In situ product recovery techniques aiming to obtain biotechnological products: a glance to current knowledge. Biotechnol. Appl. Biochem. 68, 1044–1057. https://doi.org/10.1002/bab.2024.

Seeliger, S., Janssen, P.H., Schink, B., 2002. Energetics and kinetics of lactate fermentation to acetate and propionate via methylmalonyl-CoA or acrylyl-CoA. FEMS Microbiol. Lett. 211, 65–70. https://doi.org/10.1111/j.1574-6968.2002.tb11204.x.

Selder, L., Sabra, W., Jürgensen, N., Lakshmanan, A., Zeng, A.P., 2020. Co-cultures with integrated in situ product removal for lactate-based propionic acid production. Bioprocess Biosyst. Eng. 43, 1027–1035. https://doi.org/10.1007/s00449-020-02300-0.

Shahab, R.L., Brethauer, S., Davey, M.P., Smith, A.G., Vignolini, S., Luterbacher, J.S., Studer, M.H., 2020. A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose. Science (80-.) 369. https://doi.org/10.1126/science.abb1214.

Shi, Z., Zhang, C., Chen, J., Mao, Z., 2005a. Performance evaluation of acetone–butanol continuous flash extractive fermentation process. Bioprocess Biosyst. Eng. 27, 175–183. https://doi.org/10.1007/s00449-004-0396-7.

Shi, Z., Zhang, C., Chen, J., Mao, Z., 2005b. Performance evaluation of acetone–butanol continuous flash extractive fermentation process. Bioprocess Biosyst. Eng. 27, 175–183. https://doi.org/10.1007/s00449-004-0396-7.

Sun, Y., Li, Y.-L., Bai, S., Hu, Z.-D., 1999. Modeling and simulation of an in situ product removal process for lactic acid production in an airlift bioreactor. Ind. Eng. Chem. Res. 38, 3290–3295. https://doi.org/10.1021/ie990090k.

Van Hecke, W., Kaur, G., De Wever, H., 2014. Advances in in-situ product recovery (ISPR) in whole cell biotechnology during the last decade. Biotechnol. Adv. 32, 1245–1255. https://doi.org/10.1016/j.biotechadv.2014.07.003.

Wang, P., Wang, Y., Su, Z., 2012. Microbial production of propionic acid with *Propionibacterium freudenreichii* using an anion exchanger-based in situ product recovery (ISPR) process with direct and indirect contact of cells. Appl. Biochem. Biotechnol. 166, 974–986. https://doi.org/10.1007/s12010-011-9485-7.

Woodley, J.M., Bisschops, M., Straathof, A.J.J., Ottens, M., 2008. Future directions for in-situ product removal (ISPR). J. Chem. Technol. Biotechnol. 83, 121–123. https://doi.org/10.1002/jctb.1790.

Yaroshchuk, A.E., 2000. Dielectric exclusion of ions from membranes. Adv. Colloid Interface Sci. 85, 193–230. https://doi.org/10.1016/S0001-8686(99)00021-4.

Zeng, A.-P., 2019. New bioproduction systems for chemicals and fuels: needs and new development. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2019.01.003.