Biochemical Engineering Journal 151 (2019) 107325

Contents lists available at ScienceDirect

Biochemical Engineering Journal

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

Regular article

Inference of dynamic macroscopic models of cell metabolism based on elementary flux modes analysis

Thomas Abbate[a], Sofia Fernandes de Sousa[a], Laurent Dewasme[a], Georges Bastin[b], Alain Vande Wouwer[a,*]

[a] Automatic Control Laboratory, University of Mons, 31 Boulevard Dolez, 7000 Mons, Belgium
[b] Université catholique de Louvain, ICTEAM, Department of Mathematical Engineering, av. G. Lemaitre 4, B1348 Louvain-La-Neuve, Belgium

HIGHLIGHTS

- A simple procedure to infer macroscopic models from metabolic networks is proposed.
- Macroreactions can be selected among the elementary flux modes (EFMs).
- A linear optimization problem is formulated to select the best candidate EFMs.
- Nonlinear parameter estimation is used to fit kinetic models.
- Data from cultures of hybridoma cell line HB58 illustrate the approach.

ARTICLE INFO

Keywords:
Mathematical modeling
Parameter estimation
Model reduction
Biotechnology
Mammalian cells

ABSTRACT

This study aims at providing a methodology based on both data- and knowledge-driven approaches to build dynamic macroscopic models of cell cultures. This methodology proceeds in three steps. A principal component analysis is first applied in order to determine the minimum number of macroreactions necessary to faithfully describe the available data. These reactions are then selected among the elementary flux modes associated with a chosen metabolic network through the definition of an original linear programming problem. Kinetic laws are finally identified so as to reproduce the measurement data. The proposed methodology is illustrated using four different perfusion cultures of hybridoma cells and demonstrates a good capacity to select macroreactions capable of reproducing well the complex experimental data, even with the use of simple kinetic laws and without re-identifying the stoichiometry.

1. Introduction

Mammalian cell cultures are a representative source of a number of biopharmaceutical products, including monoclonal antibodies [1,2], viral vaccines [3], and hormones [4]. Predicting the behavior of mammalian cells during cell culture processes under different culture conditions is highly desirable for both industrial and scientific reasons [5]. The efficiency of these processes can be further increased by model-based optimization and control strategies. Therefore, a reliable kinetic model for the cell metabolism is required to identify the parameters which have a significant impact on cell viability and on protein production and to understand their effects on the cellular phenotype.

Kinetic models can be categorized as structured or unstructured models. Structured models correspond to a relatively detailed representation of the cells, including the interaction between intracellular components [6]. They attempt to explicitly describe the physiological state of the cells, its composition or its regulatory adaptation to the environmental changes, resulting in a high number of equations which are difficult to handle. Instead, unstructured models do not consider the intracellular activity of the cells. They are regarded as a black-box, where substrates are consumed and converted into products. These models are known to describe the evolution of extracellular culture variables, such as biomass, substrates and products by a small set of macroscopic reactions connecting the initial components to the final products. They are simpler than structured models, thus making the model easier to identify, to use for optimization, or design of on-line algorithms for process monitoring and control.

Despite the differences between structured and unstructured

* Corresponding author.
E-mail addresses: thomas.abbate@umona.ac.be (T. Abbate), sofia.afonsofernandes@umons.ac.be (S. Fernandes de Sousa), laurent.dewasme@umons.ac.be (L. Dewasme), Georges.Bastin@uclouvain.be (G. Bastin), alain.vandewouwer@umons.ac.be (A. Vande Wouwer).

https://doi.org/10.1016/j.bej.2019.107325
Received 17 March 2019; Received in revised form 4 July 2019; Accepted 30 July 2019
Available online 12 August 2019
1369-703X/

models, recent studies have highlighted the strong link between both modeling approaches. Several methods have been devised in order to deduce macroscopic reaction schemes from metabolic networks of intracellular reaction pathways thanks to model reduction procedures [7–14].

In this study, a simple procedure to infer macroscopic bioreaction models from metabolic networks is proposed. As in previous reports, the concept of elementary flux modes (EFMs) is used to translate the metabolic network into macroscopic bioreactions linking extracellular substrates to products. However, even considering a small metabolic network, the number of EFMs can be quite large, and a selection of specific EFMs as candidate macro-reactions for the model is needed [13–15].

In this connection, several approaches have been investigated. In [9], a method based upon dynamic metabolic flux analysis (DMFA) is used to find an optimal selection of EFMs based on concentration measurements instead of reaction rates. In [11], the metabolic network is simplified by eliminating insignificant fluxes. Moreover, Hybrid Cybernetic Models (HCM) [16] have been used to reduce the number of EFMs by projection of the modes into the yield space. Another approach is the class of Lumped Hybrid Cybernetic Models (L-HCM) [17], which consist in grouping EFMs into clusters and creating an average EFM for each cluster. Finally, Flux Balance Analysis (FBA) and its extension Dynamic Flux Balance Analysis (DFBA) can also be seen as a method to reduce the number of EFMs using optimization [7]. Indeed, a solution of FBA corresponds to a positive linear combination of EFMs and the solution for any optimal product/substrate ratio always coincide with an elementary mode [18].

In this study, a linear optimization problem is formulated to select the best candidate EFMs that describe the macroscopic bioreactions. This strategy proceeds in three steps:

- first, a data-driven approach is exploited, where principal component analysis is used to infer the minimum number of bioreactions that are required to explain the observed experimental data (this approach follows [19]);
- next, the knowledge of the metabolic network is exploited to select a corresponding number of EFMs that are *a priori* the best candidates to describe the bioreaction network;
- finally, simple kinetic laws, e.g., Monod and inhibition factors, are formulated and nonlinear parameter estimation is used to fit the model to the experimental data. At that stage, the range of operation of the bioreactor (range of concentrations of the main components) and the dynamical changes within this range are determinant in the possibility to estimate kinetic parameters and to possibly model additional phenomena, such as inhibition.

This paper is organized as follows. The next section presents the methodology for determining the minimum number of macroreactions and the most representative EFMs. Information on the experimental databank, i.e. data collected from cultures of hybridoma cell line HB58 in 2 L bioreactor operated in perfusion mode is provided in Section 3. The considered metabolic reaction network is introduced in Section 4. Then, the results of the experimental application are discussed and a dynamic macroscopic model is developed and validated in Section 5. Conclusions are finally drawn in Section 6.

2. Determining minimal sets of macroscopic bioreactions

Based on the method originally developed in [19], this section proposes a methodology to determine a minimal set of macroscopic bioreactions which is consistent with the available experimental data.

The general mass-balance model equations are given by:

$$\frac{dC_m}{dt} = K\varphi + D(C_{m_{in}} - C_m) - Q_m \tag{1}$$

where C_m is the vector of concentrations of the n_m measured species in the culture medium. The full rank (equal to the number of reactions n_r) matrix $K \in \Re^{(n_m \times n_r)}$ represents the stoichiometry of the reaction network. $\varphi \in \Re^{n_r}$ is the vector of reaction rates and D is the dilution rate. The term $C_{m_{in}}$ represents the inflow concentrations. The gaseous flow rate Q_m is the exchange of matter in gaseous form between the surrounding and the reaction medium. At each measurement time t_j, the vector of (measurable) biological rates of consumption/production of each species can thus be defined as:

$$u(t_j) = \frac{dC_m(t_j)}{dt} - D(t_j)(C_{m_{in}}(t_j) - C_m(t_j)) + Q_m(t_j) \tag{2}$$

In general, the computation of flux vectors can be achieved following two approaches, i.e., a rate-based approach where the differentiation of a filtered version of the concentration measurements is achieved (such as for instance in [20,21]), or an extent-based approach [22,23] which makes use of integration rather than differentiation but requires the prior knowledge of the stoichiometry. In this study, as we do not know a priori the structure of the network of macroreactions, but would like to deduce it from the selected EFMs, we cannot apply this latter approach, and we therefore resort to the rate-based approach. In particular, smoothing B-splines are available in Matlab through the function *spaps* and can be applied to the noisy concentration measurements prior to differentiation using the function *fnder* in Matlab. Eq. (1) therefore leads to:

$$u(t_j) = K\varphi(t_j). \tag{3}$$

This equation can be written in matrix form as follows:

$$U = KW \tag{4}$$

where matrix $U^{n_m \times N}$ is made of the vectors $u(t_j)$ at N ($> n_m$) measurement times t_1, …, t_N. In the same way, the associated matrix of reaction rates W is made of N $\varphi(t_j)$ vectors. We assume that the reaction rates are linearly independent (none of the reaction rates can be written as a linear combination of the others, otherwise the model has to be reformulated and some rates eliminated) and that the experimental conditions provide data that are sufficiently informative to have a full rank matrix W (i.e., the time evolution of the rates is such that none of the lines of W can be written as a linear combination of the others). Then, the following classical result can be used (see [19]):

Property 1. *If matrix K has rank n_r, and W has full rank, then the $n_m \times n_m$ matrix $M = UU^T = KWW^TK^T$ has rank n_r. Since it is a symmetric matrix, it can be written:*

$M = P^T \Sigma P$ *where P is an orthogonal matrix ($P^TP = I$) and*

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_{n_r} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

with $\sigma_{i-1} \geq \sigma_i > 0$ for $i \in 2, \dots, n_r$. The number of macroscopic reactions corresponds to the number of non-zero eigenvalues of UU^T.

In practice, due to data sampling and measurement noise, there will be no zero eigenvalues, and the question on how to determine the number of eigenvalues that must be taken into account in order to reasonably approximate the experimental data is answered in [24] by selecting the n_r first principal axes, which represent a total variance larger than a fixed threshold. To avoid conditioning problems, the components $u_k(t_j)$ of the vector $u(t_j)$ are normalised as follows:

$$\tilde{u}_k(t_j) = \frac{u_k(t_j)}{\|u(t_j)\|}, \quad k = 1, \dots, n_m \tag{5}$$

for the computation of the matrix UU^T.

Having determined the number n_r of macro-reactions, the next

challenge is to select an appropriate subset of n_r reactions among the n_e elementary flux modes. In principle, the number of possible combinations is given by

$$C(n_e, n_r) = \frac{n_e!}{n_r!(n_e - n_r)!} \tag{6}$$

The number of combinations therefore increases with the size of the initial metabolic network and the associated number of EFMs. However, not all the combinations considered in Eq. (6) are meaningful, and only the combinations of EFMs connecting all the measured substrates to the measured products should be considered. This number is obviously smaller but cannot be determined *a priori*.

Due to measurement errors and smoothing spline approximations, it may happen that the sign of small fluxes in Eq. (3) becomes uncertain. Therefore, in order to cope with the situation where the fluxes $\varphi(t)$ are known to be of constant sign along the data of the experiment, a linear programming problem is formulated. This problem aims at finding the smallest tolerance corridor ϵ making Eq. (3) feasible with non-negative fluxes:

$$\forall t_j, j = 1, \ldots N;$$

$$\text{Let} \begin{cases} \varepsilon(t_j) = [\varepsilon_1(t_j), \ldots, \varepsilon_{n_m}(t_j)]^T \\ \varepsilon(t_j) \geq 0 \\ \varphi(t_j) = [\varphi_1(t_j), \ldots, \varphi_{n_r}(t_j)]^T \\ \varphi(t_j) \geq 0 \end{cases}$$

$$\min_{\varepsilon(t_j), \phi(t_j)} \sum_{k=1}^{n_m} \varepsilon_k(t_j)$$

$$s.t. \begin{cases} K\varphi(t_j) \leq u(t_j) + \Gamma\varepsilon(t_j) \\ K\varphi(t_j) \geq u(t_j) - \Gamma\varepsilon(t_j) \end{cases} \tag{7}$$

where

- $\varepsilon(t_j) \in \Re^{n_m}$ is a positive vector of minimal tolerances that allow computing an approximate solution to the system of equations $K\varphi(t_j) \approx u(t_j)$ (this system is overdetermined when $n_r < n_m$).
- $\epsilon(t_j) \geq 0$ and $\varphi(t_j) \geq 0$ impose positivity constraints on the solutions,
- $K \in \Re^{(n_m \times n_r)}$ is the stoichiometry matrix of the selected set of macroreactions (among the EFMs),
- the scaling matrix Γ normalizes the errors since the various external fluxes may evolve on different scales, and is defined as follows:

$$\Gamma = \begin{pmatrix} \gamma_1 & 0 & \ldots & 0 \\ 0 & \gamma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \gamma_{n_m} \end{pmatrix}, \tag{8}$$

with

$$\gamma_k = \max_{tj} |u_k(t_j)|, \quad k = 1, \ldots, n_m. \tag{9}$$

In system (7), the unknown variables are thus the tolerances $\epsilon(t_j)$ and the fluxes $\varphi(t_j)$, while the matrix K is a *a priori* information on the stoichiometry (according to the assumption that a set of EFMs is capable of representing the data) and the vectors $u(t_j)$ are the measurement information. The linear programming problem returns the tolerances $\epsilon(t_j)$ and the fluxes $\phi(t_j)$ so that the following inequalities are satisfied:

$$u(t_j) - \varepsilon(t_j) \leq K\phi(t_j) \leq u(t_j) + \varepsilon(t_j) \tag{10}$$

with the tightest corridor of error.

The aim of the overall procedure is to select the most appropriate set of EFMs and corresponding stoichiometry matrix K. A guide in this selection is the capacity of the model associated with K to accurately describe the external data as measured by the indicator:

$$R = \sum_{j=1}^{N} \sum_{k=1}^{n_m} \epsilon_k(t_j) \tag{11}$$

Obviously the model capacity to predict the data is reflected in a decrease of the quantity (11). The several model-candidates (stemming from the combinations of all the subsets of n_r EFMs taken out of the global set of n_e EFMs) can be selected accordingly.

The admissible solutions $\varphi(t_j)$ give the evolution of the fluxes along the EFMs. Further, these kinetics could be modelled using specific kinetic laws (Michaelis–Menten factors, inhibition factors, etc) whose parameters could be estimated using the signals $\varphi(t_j)$ and standard parameter identification techniques.

In Section 5, this methodology is applied to a small metabolic network proposed in [25] and experimental data from perfused hybridoma cell cultures.

3. Material and methods

3.1. Cell line and media

This study is illustrated with experimental data from cultures of hybridoma cell line HB58 (ATCC), which produces antibodies type IgG1, anti-CD54, specific for mouse kappa light chain. These experiments have been performed at the State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology (ECUST), Shanghai [12]. Serum-free medium chemically defined with 1:1 mixture of DMEM and F12 (Gibco) was used and supplemented with 10 mg of bovine insulin, 10 mg of transferring-selenite (Fe-saturated), 500 µmol of ethanolamine and other property additives. Glucose, glutamine and other amino acids were added to the cultures. Table 1 describes the operating conditions of the cultures and the glucose, glutamine and alanine concentrations in the feed stream. It must be stressed that the experiments are designed to cover a wide operating range and capture information about glucose and glutamine saturation and limitation levels.

3.2. Bioreactor operation mode and analysis methods

Cultures were conducted in a 2-L stirred bioreactor (B. Braun Biostat B-DCU) and were settled in a working volume of 1.8 L during the whole cultures. All cultures started in batch mode and were inoculated to reach an initial concentration between 0.2 and 0.5×10^9 cells/L. Temperature was kept at 36.8 °C; the gases air, CO_2, O_2 and N_2 were mixed to maintain DO at 40% air saturation and bicarbonate solution (0.75 mol/L Na_2CO_3 and 0.5 mol/L $NaHCO_3$) was used for pH control around 7.0 ± 0.2. Data acquisition and process control were performed using the supervisory software MFCS/Win 3.0.

Perfusion phase started after 36–48 h of batch culture with a constant dilution rate of 0.0197 h^{-1}. Cells were retained by a spin-filter (20 µm) and the stirring speed was kept at 200 rpm.

Cells were counted with hemocytometer using the trypan blue exclusion method. Glucose, lactate and ammonia concentrations were determined using YSI 7100 biochemical analyzer (Yellow Springs Instruments). The amino acids were analyzed by reverse-phase high

Table 1
Operating conditions for both cultures.

Units	Initial viable cells [10^9 cells/L]	Feed stream			Duration of batch [h]
		Glc_{in} [mM]	Gln_{in} [mM]	Ala_{in} [mM]	
Cult.1	0.19	11	5	0.5	54
Cult.2	0.23	15	11.5	0.1	56.5
Cult.3	0.36	28	4.0	0.3	48
Cult.4	0.36	28	9.5	0.05	44

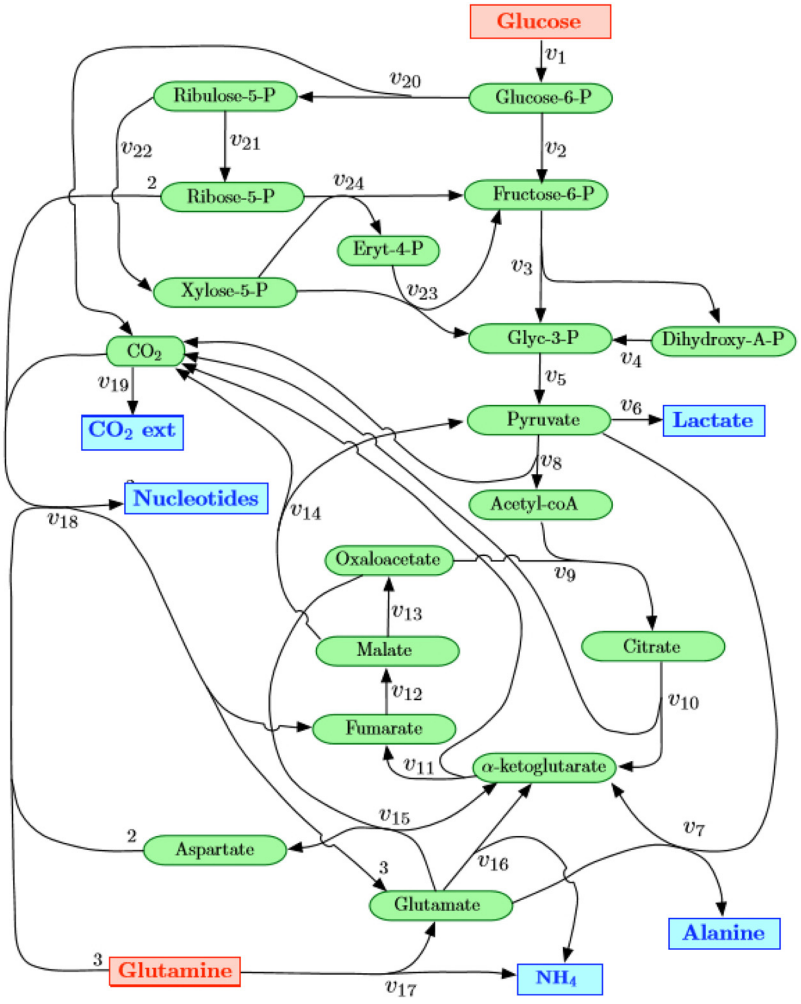


Fig. 1. Metabolic network for mammalian cell central metabolism.

performance liquid chromatography (HPLC) with a UV–vis detector.

4. Metabolic network presentation

A small metabolic network is used to represent the central metabolism of hybridoma cells (see Fig. 1). This metabolic network describes only a part of the metabolism concerned with the utilization of the main energetic nutrients (glucose and glutamine), the secretion of the main extracellular products (lactate, ammonia, alanine and CO_2) and the production of the intracellular nucleotides.

This network is quite small in comparison with more detailed networks (e.g. [26]) but this case study is used to develop and illustrate the concepts in a clear, traceable way. This small-size example is therefore more to be seen as a "proof-of-concept".

From this metabolic network, eleven elementary flux modes are obtained (see [25] for details). They induce the ten macroscopic input/output bioreactions given in Table 2 (note that EFMs e_9 and e_{10} give the same input/output reaction). In the sequel of this study, the macroreactions of Table 2 will be labelled with the notations e_i of the corresponding EFMs. The stoichiometry matrix corresponding to these ten macroreactions is given in Table 3.

5. Application to hybridoma cell cultures in perfusion

5.1. Selection of the candidate macroreactions

Our case study includes the experimental data of four different

Table 2
Macroscopic reactions.

EFM	Macroscopic reaction
e_1	Glc → 2 Lac
e_2	2 Glc + 3 Gln → Ala + Nucleotide + 9 CO_2
e_3	Gln → Lac + 2 N + 2 CO_2
e_4	Gln → 2 N + 5 CO_2
e_5	Gln → Ala + N + 2 CO_2
e_6	2 Glc + 3 Gln → Lac + Ala + Nucleotide + 6 CO_2
e_7	3 Glc → 5 Lac + 3 CO_2
e_8	2 Glc + 3 Gln → 2 Lac + N + Nucleotide + 6 CO_2
e_9, e_{10}	Glc → 6 CO_2
e_{11}	2 Glc + 3 Gln → N + Nucleotide + 12 CO_2

Table 3
Stoichiometry of the macroreactions of Table 2.

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	$e_{9/10}$	e_{11}
Glc	−1	−2	0	0	0	−2	−3	−2	−1	−2
Lac	2	0	1	0	0	1	5	2	0	0
Gln	0	−3	−1	−1	−1	−3	0	−3	0	−3
N	0	0	2	2	1	0	0	1	0	1
Ala	0	1	0	0	1	1	0	0	0	0
Nuc	0	1	0	0	0	1	0	1	0	1
CO_2	0	9	2	5	2	6	3	6	6	12

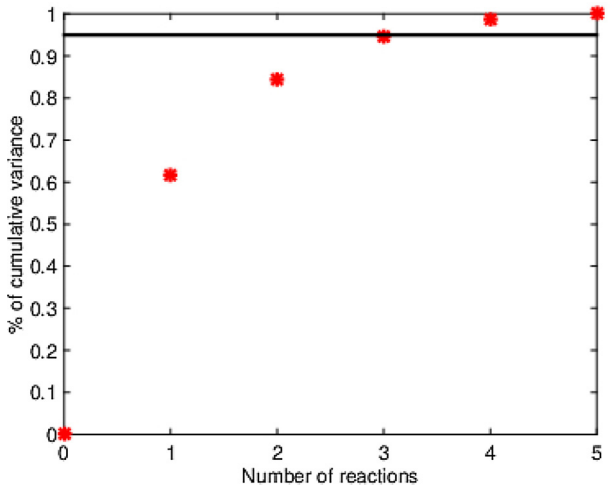


Fig. 2. PCA results giving the total variance explained with respect to the number of reactions.

perfused cell cultures, from which on-line measurements of glucose, glutamine, lactate, ammonia and alanine are available. The vectors are computed by fitting smoothing B-splines to the data and computing their derivatives:

$$u(t_j) = \begin{bmatrix} \left.\frac{\mathrm{dGlc}}{\mathrm{dt}}\right|_{t_j} - D(t_j)(\mathrm{Glc_{in}}(t_j) - \mathrm{Glc}(t_j)) \\ \left.\frac{\mathrm{dLac}}{\mathrm{dt}}\right|_{t_j} + D(t_j)\mathrm{Lac}(t_j) \\ \left.\frac{\mathrm{dGln}}{\mathrm{dt}}\right|_{t_j} - D(t_j)(\mathrm{Gln_{in}}(t_j) - \mathrm{Gln}(t_j)) \\ \left.\frac{\mathrm{dN}}{\mathrm{dt}}\right|_{t_j} + D(t_j)N(t_j) \\ \left.\frac{\mathrm{dAla}}{\mathrm{dt}}\right|_{t_j} - D(t_j)(\mathrm{Ala_{in}}(t_j) - \mathrm{Ala}(t_j)) \end{bmatrix} \quad (12)$$

Then, using the normalization (5), the eigenvalues of the matrix UU^T are computed and the principal component analysis (PCA) indicates, as shown in Fig. 2, that 3 reactions are sufficient to explain 95% of the observed variance, 4 reactions explain 99% of the total variance and 5 reactions obviously allow explaining 100%. In this application, we know by assumption that, whatever the selected macroreactions in Table 2, the fluxes $\varphi(t_j)$ must be positive. The numerical results of LP (7) show that to cope with slight potential deviations in this positivity constraint, some of the flux values have to be saturated to zero on some periods of time. The consequence of these punctual saturations is visible in Table 4, which gathers the residual cost R (Eq. (11)) of the best candidate-combinations of 2–10 macroreactions connecting the substrates to the products, i.e., which allow describing the evolution of all the measured species. It shows that the minimal residual cost decreases with the number of EFMs up to 7, with no significant further improvements with more EFMs. The residual cost for 5 EFMs is relatively small but not zero, and the residual cost does not vanish even when including all the EFMs in the model. This counter-intuitive result is indeed due to the positivity constraint imposed on the solutions $\varphi(t_j)$ of system (7) which transform a problem of linear algebra into a convex analysis. When removing the positivity constraint, the results are in agreement with the PCA outcome and the total data variance is explained by all combinations of 5 elementary fluxes leading to K matrices of rank 5 (independent reactions). In contrast, when the positivity constraint is active, a network of more than 5 EFMs can provide a better representation.

Table 4
Best success combinations of varying numbers of macroreactions passing test (7), where R (11) represents the sum of minimal errors necessary for LP (7) to allow feasible solutions. Note that R has no unit.

# reactions	Macroreactions combination	R
2	e_5, e_8	76.0
3	e_1, e_5, e_{11}	23.9
4	e_1, e_4, e_5, e_{11}	12.6
5	e_1, e_3, e_5, e_8, e_9	7.2
6	e_1, e_3, e_5, e_6, e_9, e_{11}	6.5
7	e_1, e_3, e_5, e_6, e_8, e_9, e_{11}	6.37
8	e_1, e_3, e_4, e_5, e_6, e_8, e_9, e_{11}	6.37
9	e_1, e_2, e_3, e_4, e_5, e_6, e_8, e_9, e_{11}	6.37
10	e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{11}	6.37

This interesting observation induces that the several candidate-combinations corresponding to a specific number of macro-reactions usually do not lead to the same level of accuracy in representing the data. LP (7) therefore constitutes a useful tool to discriminate among candidates, even for combinations of more than 5 reactions. This is illustrated in Fig. 3 which plots the time evolution of the concentrations associated with the solutions obtained from the computation of (7) for all the combinations of 5 EFMs covering the network. Due to the large number of curves to display, the same colors are used several times to plot the curves, which makes difficult to distinguish individual fluxes on the graph. Some important information can however be extracted: (a) when only a few curves are visible, this means that many combinations of 5 EFMs lead to almost identical results for the considered concentration evolution; (b) many EFM combinations are not suited to describe accurately the available data. In fact, among the possible combinations, only 4 seem particularly well suited to reproduce the data, which are presented in Table 5.

The corresponding kinetics, inferred from the computation of (7), have almost identical shapes. As shown in Fig. 4a, which presents the kinetics of the second of these combinations, the complex kinetics shapes might be difficult to reproduce using products of Michaelis–Menten laws or inhibition factors, as it is classically achieved in biological system modelling. In the sequel of this study, we will investigate in more details how well these standard kinetic laws can reproduce the experimental data using parameter identification techniques. Before embarking in this identification work, a first very global approach is taken, which consists in smoothing the raw kinetic signals using B-splines, to integrate them, and to compare the results with the measurements. It is interesting to observe that the concentration evolutions obtained by integration of the smoothed kinetic signals for the four model-candidates are similar and satisfactorily fit the experimental data (Fig. 4b), despite the loss of information associated with the smoothing approximation. This observation supports the possibility to describe the kinetic signals using simple kinetic structures. However, the reproduction of alanine in dataset 3 is less satisfactory, whatever the number of macroreactions considered. We can only conjecture about this observation: (a) the level of alanine is dataset 3 is lower than in the other experiments and might be close to the detectability levels of the HPLC technique, (b) alanine is provided in the feed stream and quantification errors as well as small perturbations can have a direct impact on the concentrations in the culture medium, (c) alanine consumption could occur in some phases of the culture, but as all fluxes are assumed irreversible, this phenomenon is not modelled here (we have not pursued this route, which seems unlikely, because of the limited information in dataset 3 and the lack of further evidence).

Any of the combinations of Table 4 might thus be considered for model identification. As the results of PCA indicate that 3 reactions explain 95% of the variance, it is proposed to also investigate models with 4 and 3 reactions. Combination 1-3-5-9-11 (i.e. combination 2 of Table 5) seems to be a good base to achieve this objective since the residual cost of combinations 1-3-5-11 and 1-5-11 are the second and first best combinations of 4 and 3 elementary fluxes, respectively.

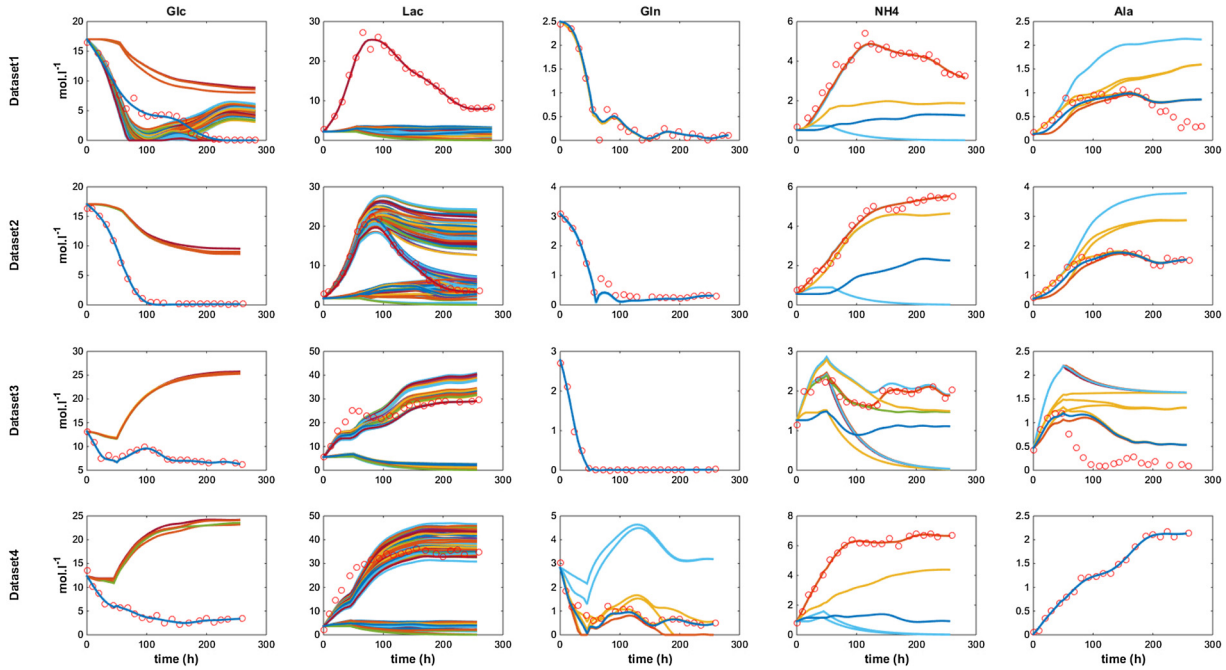


Fig. 3. Time evolution of the concentrations obtained by integration of the solutions of LP (7) for all combinations of 5 EFMs covering the whole network.

Table 5
Best combinations of 5 elementary fluxes selected as candidates to build a macroscopic model.

Combination number	Macroreaction combinations	J
1	e_1, e_3, e_5, e_8, e_9	7.2
2	e_1, e_3, e_5, e_9, e_{11}	7.5
3	e_1, e_4, e_5, e_8, e_9	7.6
4	e_1, e_4, e_5, e_9, e_{11}	8.0

The mass balance ordinary differential equations, for which kinetic evolution have now to be postulated, are given by:

$$\frac{\mathrm{dGlc}}{\mathrm{dt}} = -\varphi_1 X - \varphi_9 X - 2\,\varphi_{11} + D(\mathrm{Glc_{in}} - \mathrm{Glc}) \tag{13}$$

$$\frac{\mathrm{dLac}}{\mathrm{dt}} = 2\,\varphi_1 X - \varphi_3 X - \mathrm{D\,Lac} \tag{14}$$

$$\frac{\mathrm{dGln}}{\mathrm{dt}} = -\varphi_3 X - \varphi_5 X - 3\,\varphi_{11} + D(\mathrm{Gln_{in}} - \mathrm{Gln}) \tag{15}$$

$$\frac{\mathrm{dN}}{\mathrm{dt}} = 2\,\varphi_3 X + \varphi_5 X + \varphi_{11} X - \mathrm{DN} \tag{16}$$

$$\frac{\mathrm{dAla}}{\mathrm{dt}} = \varphi_5 X + D(\mathrm{Ala_{in}} - \mathrm{Ala}) \tag{17}$$

where Glc_{in}, Gln_{in} and Ala_{in} are the concentrations of glucose, glutamine and alanine in the feed stream and D is the dilution rate. The parameter identification procedure is described in the following section.

5.2. Nonlinear parameter estimation

The model identification performed in this work is achieved using classical nonlinear parameter estimation techniques [27]. Since the stoichiometric coefficients of the macroreactions are provided by the elementary fluxes modes, only the kinetic parameters are estimated by minimizing a weighted nonlinear least-square criterion (sum of squared differences between model predictions and experimental measurements) of the form:

$$\min_{\hat{\theta}} \quad J = \sum_{k=1}^{N} (y_k - y_{m,k}(\hat{\theta}))S^{-1}(y_k - y_{m,k}(\hat{\theta}))^T \tag{18}$$

where $\hat{\theta}$ is the set of parameters to identify, J is the value of the cost function, y_k are the measurements, $y_{m,k}$ are the model predictions, S is a scaling matrix whose diagonal elements are the squares of the maximum measurement values of each component concentration, and N is the number of sampling times. Biomass is considered as a known variable, whose values are obtained by smoothing B-splines.

Once an estimation $\hat{\theta}^*$ of the hypothetical true parameters θ^* is found, an *a posteriori* estimation of the measurement error variance $\hat{\sigma}^2$ is obtained by:

$$\hat{\sigma}^2 = \frac{J}{N - P} \tag{19}$$

where P is the number of parameters (dimension of the vector θ^*). This relation is then used to build the *a posteriori* covariance matrix of the measurement errors $\hat{\Sigma}$:

$$\hat{\Sigma} = \hat{\sigma}^2 S \tag{20}$$

which can be used to evaluate the Fisher information matrix (FIM):

$$\mathrm{FIM} = \sum_{k=1}^{N} Y_{\hat{\theta},k} \hat{\Sigma}^{-1} Y_{\hat{\theta},k}^T \tag{21}$$

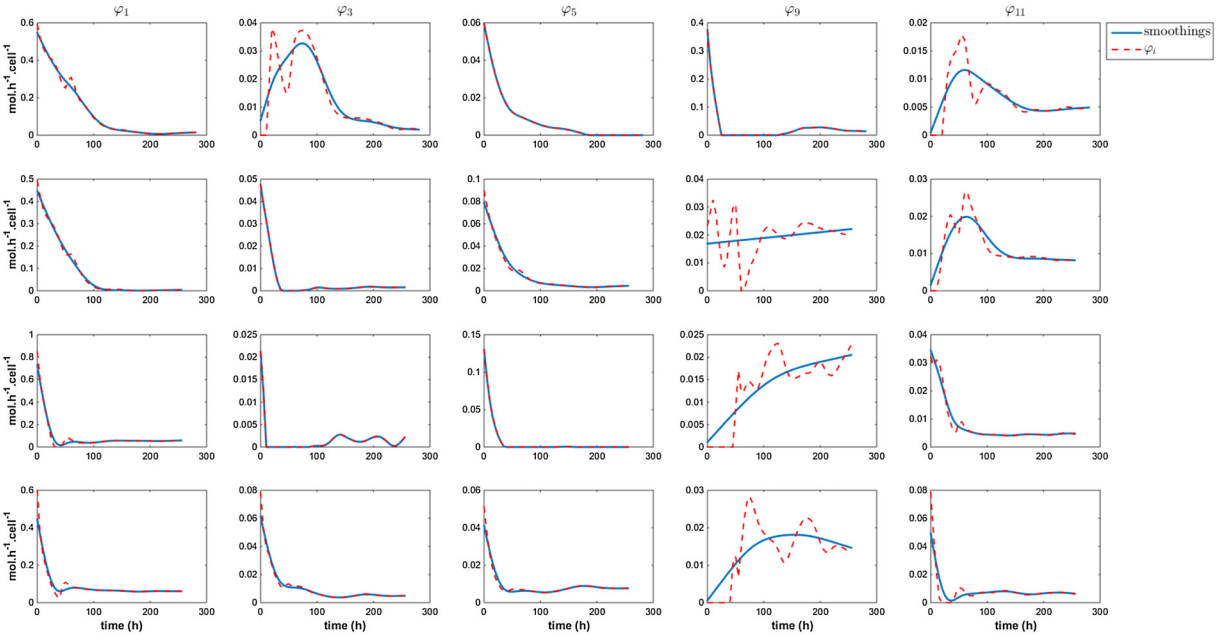
where $Y_{\theta,k}$ is the sensitivity matrix of the model with respect to the parameters at time k. These sensitivities are computed through integration of the parametric sensitivity equations:

$$\frac{\mathrm{dY_\theta}}{\mathrm{dt}} = \frac{\partial f}{\partial Y} Y_\theta + \frac{\partial f}{\partial \theta} \tag{22}$$

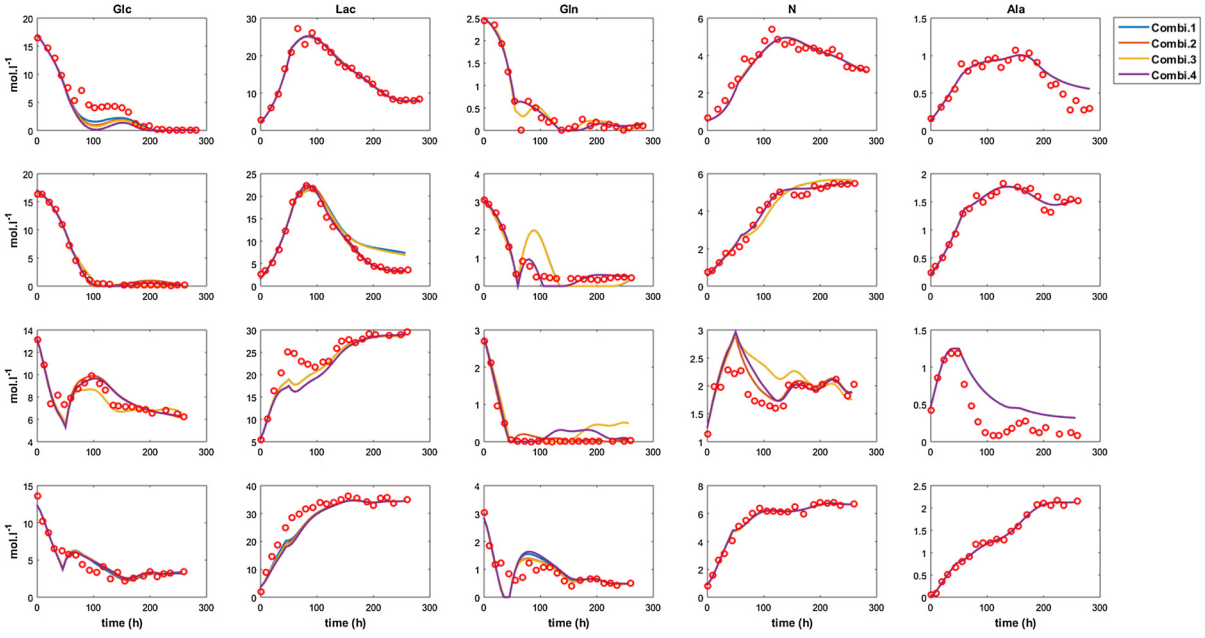
A lower bound of the parameter estimation error covariance matrix is then given by:

$$Q = \mathrm{FIM}^{-1} \tag{23}$$

It must be stressed that due to possible measurement errors affecting the initial concentrations of the experiments, the latter are introduced in the identification problem as unknown parameters. This allows alleviating the bias that could introduce wrong initial concentrations.



(a)



(b)

Fig. 4. Smoothing splines of the fluxes φ for combination 2 (a) and time evolution of the concentration obtained by integration of the smoothed φ signals related to the 4 best combinations of 5 EFMs (b).

6. Results and discussion

In order to start the identification with a model with several degrees of freedom, all kinetics are first set to the same general structure:

$$\varphi_i = \mu_{\max,i} \frac{\text{Glc}}{\text{Glc} + K_{\text{Glc}_i}} \frac{\text{Gln}}{\text{Gln} + K_{\text{Gln}_i}} \frac{K_{\text{Lac}_i}}{\text{Lac} + K_{\text{Lac}_i}} \frac{K_{N_i}}{N + K_{N_i}} \quad i = 1, \ldots, 5 \tag{24}$$

where Glc, Lac, Gln and N stands for the concentrations of glucose, lactate, glutamine and ammonium, respectively, K_{Glc} and K_{Gln} are the half-saturation constants of Michaelis–Menten laws with respect to glucose and glutamine, and K_{lac} and K_N are inhibition constants associated with lactate and ammonium. This kinetic structure can subsequently be adapted depending on the parameter identification results as well as parameter sensitivity analysis.

Moreover, since nonlinear parameter estimation problems are subject to local minima, a multi-start strategy is developed in order to ensure the exploration of the solution space and the identification of the

Table 6
Identifcation results for model 1, 2, 3 and 4. Admissible intervals (A.I.) are set to focus research within the solution space area where the parameters are likely to be found. Results presents the minimum and maximum parameter values of the identifications with residual costs below the defined cut-off threshold as well as the estimated optimal solution $\hat{\theta}^*$ and the related coefficient of variation C.V..

Models			Model 1				Model 2				Model 3				Model 4			
Macroreactions involved			1-3-5-9-11				1-3-5-11				1-5-11				1-3-5-9-11			
Datasets for direct validation			1-2-3-4				1-2-3-4				1-2-3-4				1-2-3			
Number of multistarts			200				200				200				200			
Mutlistart cut-off			$J < 10$				$J < 14$				$J < 18$				$J < 9$			
Remaining identifications			33				64				80				103			
Param.	Units	A.I.	min $\hat{\theta}$	$\hat{\theta}^*$	C.V.	max $\hat{\theta}$	min $\hat{\theta}$	$\hat{\theta}^*$	C.V.	max $\hat{\theta}$	min $\hat{\theta}$	$\hat{\theta}^*$	C.V.	max $\hat{\theta}$	min $\hat{\theta}$	$\hat{\theta}^*$	C.V.	max $\hat{\theta}$
$\mu_{max,1}$	h^{-1}	[0.1–1]	0.080	0.468	2%	0.959	0.169	0.634	2%	0.975	0.231	0.800	1%	1.000	0.025	0.479	2%	0.994
$K_{Glc,1}$	mM	[0.1–30]	1.429	8.112	16%	29.011	5.796	25.961	11%	29.999	5.746	23.912	9%	29.985	1.012	9.498	15%	22.507
$K_{Gln,1}$	mM	[0.01–3]	0.050	0.077	35%	1.856	0.015	0.026	42%	0.153	0.028	0.069	18%	0.195	0.026	0.075	36%	1.871
$\mu_{max,3}$	h^{-1}	[0.01–0.2]	0.010	0.032	51%	0.160	0.014	0.044	41%	0.197	–	–	–	–	0.010	0.030	43%	0.099
$K_{Glc,3}$	mM	[0.1–10]	0.092	0.230	50%	9.615	0.000	5.289	81%	9.631	–	–	–	–	0.157	1.000	63%	9.990
$K_{Gln,3}$	mM	[0.01–3]	0.047	0.063	46%	1.930	0.029	0.070	59%	0.870	–	–	–	–	0.043	0.152	45%	1.902
$K_{N,3}$	mM	[1–10]	0.174	1.832	74%	9.931	0.120	3.153	98%	9.986	–	–	–	–	1.084	8.233	184%	9.999
$\mu_{max,5}$	h^{-1}	[0.01–0.2]	0.010	0.051	47%	0.099	0.014	0.043	39%	0.069	0.023	0.063	48%	0.129	0.011	0.066	59%	0.140
$K_{Gln,5}$	mM	[0.01–3]	0.234	2.850	64%	6.540	0.360	1.650	63%	3.000	0.801	2.920	64%	6.347	0.048	3.785	80%	9.877
$\mu_{max,9}$	h^{-1}	[0.01–0.2]	0.012	0.022	7%	0.073	–	–	–	–	–	–	–	–	0.011	0.020	7%	0.098
$\mu_{max,11}$	h^{-1}	[0.01–0.2]	0.025	0.066	12%	0.126	0.028	0.048	11%	0.101	0.039	0.103	23%	0.200	0.011	0.052	16%	0.099
$K_{Gln,11}$	mM	[0.01–3]	0.048	0.071	33%	1.990	0.020	0.035	36%	0.184	0.035	0.089	12%	0.301	0.038	0.094	32%	1.909
$K_{N,11}$	mM	[1–10]	0.867	1.748	19%	9.532	1.283	2.642	30%	7.883	0.969	1.242	36%	5.395	1.020	2.345	20%	9.771
$K_{Lac,11}$	mM	[10–100]	12.850	43.886	32%	97.940	12.970	47.502	61%	58.716	17.473	31.168	66%	58.314	11.418	82.418	69%	99.998
Residual cost				$J(\hat{\theta}^*) = 8.7$				$J(\hat{\theta}^*) = 12.3$				$J(\hat{\theta}^*) = 16.3$				$J(\hat{\theta}^*) = 6.8$		
Datasets			set 1	set 2	set 3	set 4	set 1	set 2	set 3	set 4	set 1	set 2	set 3	set 4	set 1	set 2	set 3	set 4
RMSE (% of J)		Glc	3%	0%	1%	3%	2%	1%	4%	1%	2%	0%	2%	0%	3%	0%	1%	–
		Lac	4%	1%	2%	1%	4%	20%	2%	6%	3%	12%	1%	3%	5%	1%	2%	–
		Gln	2%	3%	1%	7%	1%	2%	0%	7%	1%	1%	0%	4%	3%	3%	1%	–
		N	13%	5%	6%	3%	13%	3%	2%	2%	23%	1%	1%	21%	17%	6%	6%	–
		Ala	8%	12%	22%	4%	6%	7%	15%	2%	5%	5%	13%	2%	12%	11%	29%	–

optimal parameters θ^*. This strategy consists in performing repeated optimization runs starting from initial guesses θ_0 randomly chosen among predefined admissible intervals (A.I.) in which the parameters are likely to be found.

Since the purpose of the present work is to test the capacity of system (7) to select sets of macroreactions able to reproduce the available data, the identification procedure is first achieved using all datasets at hand for identification and direct validation.

The application of the first round of optimization runs yields a first model (model 1) where the kinetics have the following structure:

$$\varphi_1 = \mu_{\max,1} \frac{\text{Glc}}{\text{Glc} + K_{\text{Glc},1}} \frac{\text{Gln}}{\text{Gln} + K_{\text{Gln},1}} \tag{25}$$

$$\varphi_3 = \mu_{\max,3} \frac{\text{Glc}}{\text{Glc} + K_{\text{Glc},3}} \frac{\text{Gln}}{\text{Gln} + K_{\text{Gln},3}} \frac{K_{N,3}}{N + K_{N,3}} \tag{26}$$

$$\varphi_5 = \mu_{\max,5} \frac{\text{Gln}}{\text{Gln} + K_{\text{Gln},5}} \tag{27}$$

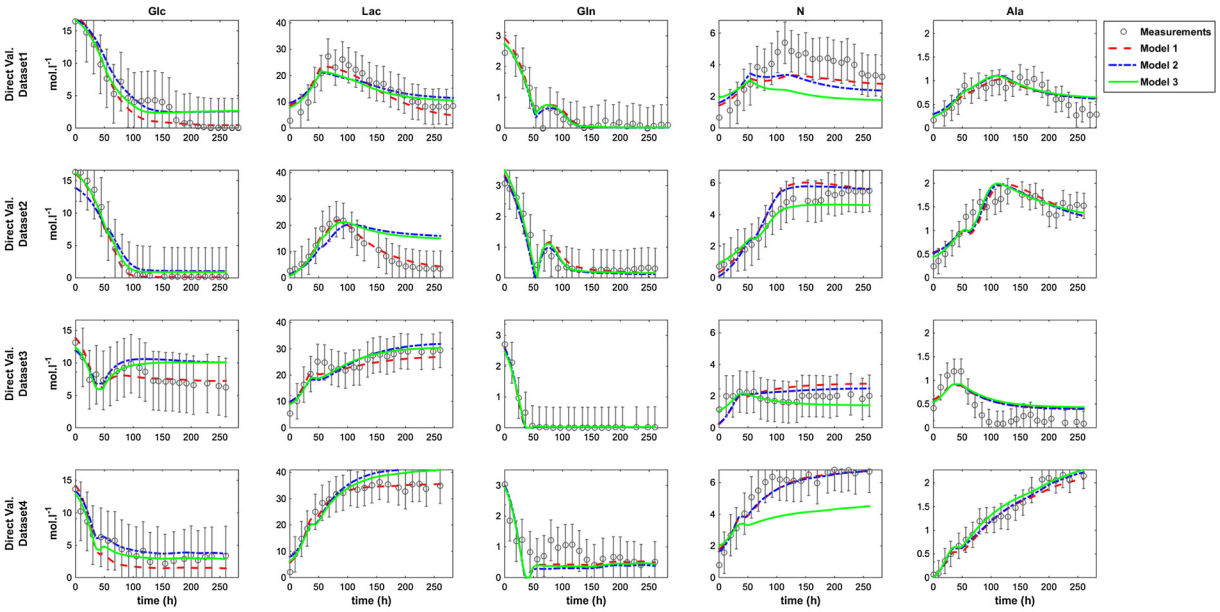
$$\varphi_9 = \mu_{\max,9} \frac{\text{Glc}}{\text{Glc} + 0.01} \tag{28}$$

$$\varphi_{11} = \mu_{\max,11} \frac{\text{Glc}}{\text{Glc} + 0.01} \frac{\text{Gln}}{\text{Gln} + K_{\text{Gln},11}} \frac{K_{N,11}}{N + K_{N,11}} \frac{K_{\text{Lac},11}}{\text{Lac} + K_{\text{Lac},11}} \tag{29}$$

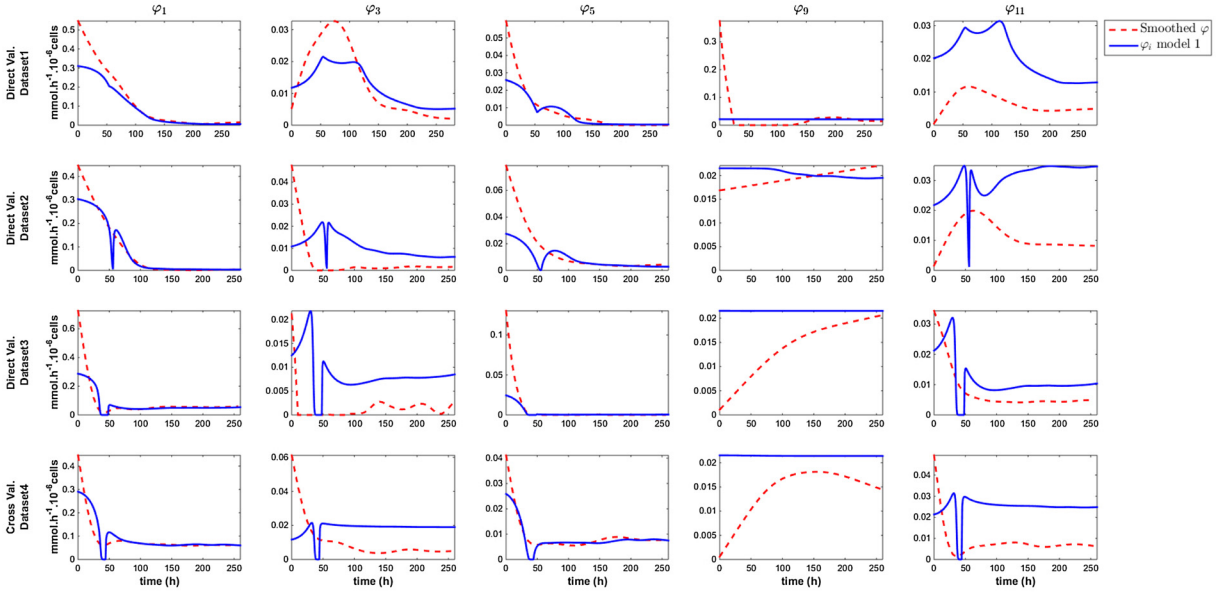
where arbitrary small half-saturation coefficients (0.01 mM) have been introduced in the Michaelis–Menten factors for glucose activation in φ_9 and φ_{11} to ensure the positivity of the model prediction. Indeed, these parameters cannot be estimated precisely from the available data, and have to be fixed in order to resolve identifiability issues. It is notorious that half-saturation coefficients can be delicate to estimate and require enough data in specific concentration ranges (actually data points where the glucose concentration values are relatively close to the half-saturation coefficients). On the other hand, the positivity of the predicted concentrations is an essential feature of any candidate model, which explains why these coefficients cannot be simply eliminated from the model. It is also interesting to note that Michaelis–Menten factors for glutamine and glucose appear in φ_1 and φ_3 respectively, while these components are not consumed in the corresponding reactions. Glucose and glutamine are indeed well-known to be interconnected in the cellular metabolism through the TCA cycle and the related anaplerotic reactions [28–31], which could explain the latter observation.

Identification results are provided in Table 6 which gathers the estimation results of the multistart strategy with a residual cost function J below specified cut-off values. The minimum and maximum parameter values are listed alongside with the best estimate $\hat{\theta}^*$ (leading to the lowest residual cost). Coefficients of variation (standard deviations expressed in percentage of the parameter values) are also evaluated based on the Fisher Information Matrix. The normalized residual mean squared error (RMSE), scaled with the square of the maximum measured values, is also provided. This information is given per dataset and per variable and expressed in % of J (obviously, the sum of these relative errors is 100%).

From 200 optimization runs, 33 present a residual cost lower than 10 and are considered as leading to interesting local minima (potentially in the neighbourhood of the global minimum), while others are rejected. The precision of the best estimate $\hat{\theta}^*$ is satisfactory since the coefficients of variations are mostly below 50%. The model predictions associated with $\hat{\theta}^*$ are given in Fig. 5a where the error bars represent the *a posteriori* 95% confidence intervals on the measurements. Fig. 5b shows the comparison between the fluxes obtained after identification of the kinetic structures and the smoothing of the fluxes obtained by applying LP (7) with the combination 1-3-5-9-11. Clearly, the evolution of the concentrations obtained with model 1 are in good agreement with the data sets. Although the fluxes predicted by the identified model do not exactly match the smoothed fluxes, the deviations are not alarming. In this study, our aim is mostly to show that simple kinetic structures can reproduce the data satisfactorily. However, we have to keep in mind that the identification of kinetic laws is always a difficult task requiring informative data in specific concentration ranges. A thorough study might require experiment design, new experimental



(a)



(b)

Fig. 5. Direct validation of models 1, 2 and 3 (a) and comparison of the predicted fluxes with the smoothing results of Fig. 4a (b).

data collection, and identification of modified kinetic laws. These refinements are considered out of the scope of the present study, whose focus is on the methodology (supported by available data).

It is also interesting to note than 22% of the residual cost J is due to alanine measurements in dataset 3. Yet, as mentioned in Section 5.1, the alanine evolution could be delicate to reproduce given the positivity constraints imposed on the fluxes and the possible errors in the measured feed concentrations.

To further test the capacity of the proposed method to select appropriate model candidates, models 2 and 3, which are described by combinations 1-3-5-11 and 1-5-11, respectively, are now identified. Even though the residual costs increase with the model reduction, the measurement prediction remain acceptable (Fig. 5a). The apparent precision of the parameter estimates also increases as a side effect of the reduction of the number of parameters to estimate from a given amount of data information. These results are remarkable, as identification only targets reaction kinetics, while the stoichiometry is imposed by the selection of EFMs.

The next step consists in challenging the capacity of the model to predict unseen experiments, and to this end, extra identification runs are conducted using only a subset of the available data sets, while keeping the rest for cross-validation.

As mentioned in Section 3.1, the experiments were designed to highlight the limitation and saturation associated with glucose and glutamine levels. Experiments 2 and 3 point out limitations with respect to glucose or glutamine, while experiments 1 is expected to show both.

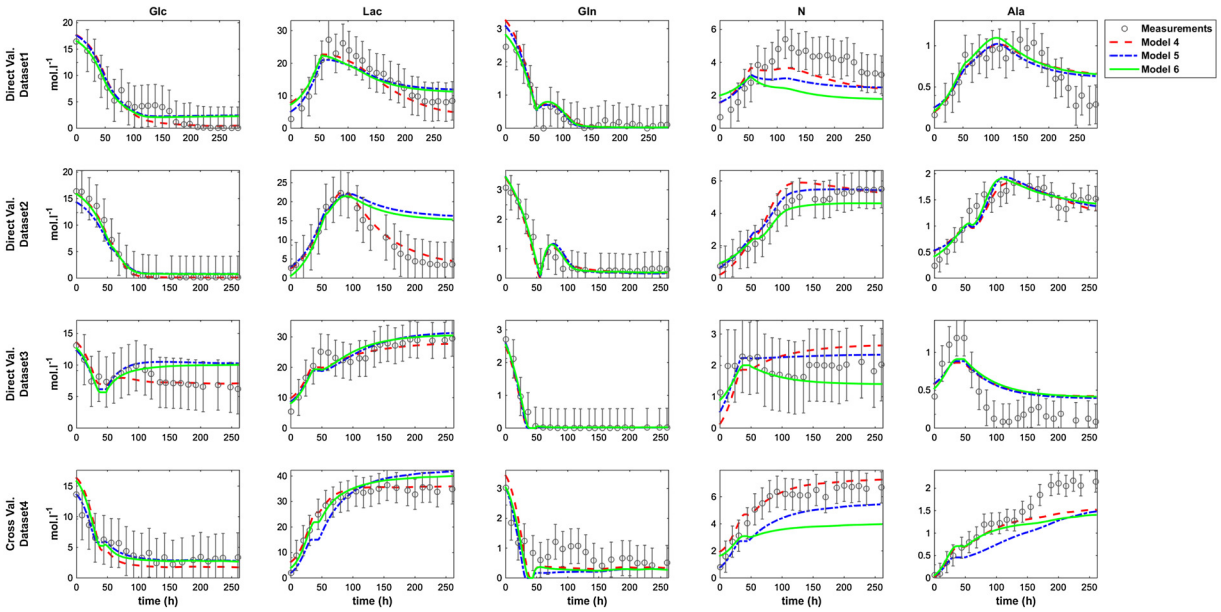


Fig. 6. Direct validation of models 4, 5 and 6.

In contrast, experiment 4 aims at avoiding these limitations. In consequence, extra identification runs are performed using the 3 datasets where limitations are foreseen (data sets 1, 2 and 3) whereas data set 4 is kept for cross-validation. Models 4, 5 and 6 are basically the same as model 1, 2 and 3, respectively, but they are identified based on only 3 data sets. The results are shown in Fig. 6 which confirms that satisfactory results can be obtained using 5 reactions (model 4) both in direct and cross-validation (although the alanine behaviour is not very well reproduced for reasons already mentioned). Other combination of 3 datasets have been tested but none of them leads satisfactory cross-validation results (not shown in this work), which stresses the importance of the information content of the data sets to achieve model selection and identification. As shown in Table 6, the parameter estimates of model 4 are close to the ones of model 1, which supports the robustness of the approach.

7. Conclusions

In this study, a simple procedure to deduce a macroscopic model of cell cultures is developed based on the concept of elementary flux modes. This procedure relies on (1) a data-driven approach where a principal component analysis is used to determine the minimum number of bioreactions required to explain the observed experimental data, (2) the selection of appropriate elementary flux modes based on the solution of a linear programming problem and (3) the identification of the model kinetic parameters using a nonlinear weighted least square criterion.

The positivity constraints imposed on the solutions generally induce a larger number of reactions to explain the data than the minimum number suggested by PCA, and the models based on the selected combinations of elementary flux modes provide a satisfactory prediction of the concentration evolution, even when using simple kinetic structures.

The apparent main limitation of the present study is the use of a very simple metabolic network resulting in only eleven elementary flux modes. The extension of the method to a larger network would imply a significant increase in the computational load to explore the possible combinations among a much larger set of EFMs. However, as the main purpose of the approach is to deduce low-order dynamic models of cell cultures, that could be used for optimization or control, one could question the interest of initiating the procedure with large, detailed, networks.

In its present form, the macroscopic model does not include a biomass production reaction and the cell concentrations are directly deduced from smoothing-splines applied to the experimental measurements. As biomass is nowadays a signal that is relatively easy to measure on-line, the proposed model could be efficiently used for process monitoring and control, the biomass signal becoming a (possibly filtered) dynamic input signal to the model. The construction of a dynamic simulator might seems impossible at first, as it would require a predictor of the biomass. However, this could be achieved in either of two ways: (a) simply consider that biomass is in exponential growth using the estimated growth rate (this parameter is estimated in the procedure) (b) introduce a new reaction in the metabolic network for the production of biomass (assuming a biomass composition).

References

[1] J. Niklas, E. Heinzle, Metabolic flux analysis in systems biology of mammalian cells, Genomics and Systems Biology of Mammalian Cell Culture (2012) 109–132.
[2] F. Sidoli, A. Mantalaris, S. Asprey, Modelling of mammalian cells and cell culture processes, Cytotechnology 44 (1–2) (2004) 27–46.
[3] D. Vester, E. Rapp, S. Kluge, Y. Genzel, U. Reichl, Virus–host cell interactions in vaccine production cell lines infected with different human influenza a virus variants: a proteomic approach, J. Proteomics 73 (9) (2010) 1656–1669.
[4] T. Nottorf, W. Höra, H. Büntemeyer, S. Siwiora-Brenke, A. Loa, J. Lehmann, Production of human growth hormone in a mammalian cell high density perfusion process, Cell Technol. Cell Prod. (2007) 789–793.
[5] D.B. Kell, J.D. Knowles, The role of modeling in systems biology, System Modeling in Cellular Biology: From Concepts to Nuts and Bolts, (2006), pp. 3–18.
[6] A. Zeng, J.-x. Bi, Cell culture kinetics and modeling, Biotechnol. Bioprocess. Ser. 30 (2005) 299.
[7] C. Baroukh, O. Bernard, Metabolic modeling of *C. sorokiniana* diauxic heterotrophic growth, IFAC-PapersOnLine 49 (26) (2016) 330–335.
[8] J.E. Haag, A. Vande Wouwer, P. Bogaerts, Dynamic modeling of complex biological systems: a link between metabolic and macroscopic description, Math. Biosci. 193 (1) (2005) 25–49.
[9] L. Hebing, T. Neymann, T. Thüte, A. Jockwer, S. Engell, Efficient generation of models of fed-batch fermentations for process design and control, IFAC-PapersOnLine 49 (7) (2016) 621–626.
[10] X. Hulhoven, A. Vande Wouwer, P. Bogaerts, On a systematic procedure for the predetermination of macroscopic reaction schemes, Bioprocess Biosyst. Eng. 27 (5) (2005) 283–291.
[11] S. Naderi, M. Meshram, C. Wei, B. McConkey, B. Ingalls, H. Budman, J. Scharer, Metabolic flux and nutrient uptake modeling of normal and apoptotic CHO cells, IFAC Proc. Vol. 43 (6) (2010) 395–400.
[12] H. Niu, Z. Amribt, P. Fickers, W. Tan, P. Bogaerts, Metabolic pathway analysis and reduction for mammalian cell cultures. Towards macroscopic modeling, Chem. Eng. Sci. 102 (2013) 461–473.

[13] A. Provost, G. Bastin, Dynamic metabolic modelling under the balanced growth condition, J. Process Control 14 (7) (2004) 717–728.
[14] F. Zamorano, A. Vande Wouwer, R.M. Jungers, G. Bastin, Dynamic metabolic models of CHO cell cultures through minimal sets of elementary flux modes, J. Biotechnol. 164 (3) (2013) 409–422.
[15] Z.I. Soons, E.C. Ferreira, I. Rocha, Selection of elementary modes for bioprocess control, IFAC Proc. Vol. 43 (6) (2010) 156–161.
[16] H.-S. Song, D. Ramkrishna, Reduction of a set of elementary modes using yield analysis, Biotechnol. Bioeng. 102 (2) (2009) 554–568.
[17] H.-S. Song, D. Ramkrishna, G.E. Pinchuk, A.S. Beliaev, A.E. Konopka, J.K. Fredrickson, Dynamic modeling of aerobic growth of *Shewanella oneidensis* predicting triauxic growth, flux distributions, and energy requirement for growth, Metab. Eng. 15 (2013) 25–33.
[18] S. Schuster, C. Hilgetag, On elementary flux modes in biochemical reaction systems at steady state, J. Biol. Syst. 2 (2) (1994) 165–182.
[19] O. Bernard, G. Bastin, Identification of reaction networks for bioprocesses: determination of a partially unknown pseudo-stoichiometric matrix, Bioprocess Biosyst. Eng. 27 (5) (2005) 293–301.
[20] A.V. Wouwer, C. Renotte, P. Bogaerts, Biological reaction modeling using radial basis function networks, Comput. Chem. Eng. 28 (11) (2004) 2157–2164.
[21] M. Brendel, D. Bonvin, W. Marquardt, Incremental identification of kinetic models for homogeneous reaction systems, Chem. Eng. Sci. 61 (16) (2006) 5404–5420.
[22] N. Bhatt, N. Kerimoglu, M. Amrhein, W. Marquardt, D. Bonvin, Incremental identification of reaction systems – a comparison between rate-based and extent-based approaches, Chem. Eng. Sci. 83 (2012) 24–38.
[23] Y. Liu, R. Gunawan, Parameter estimation of dynamic biological network models using integrated fluxes, BMC Syst. Biol. 8 (1) (2014) 127.
[24] R.A. Johnson, D.W. Wichern, et al., Applied Multivariate Statistical Analysis, Vol. 5, Prentice hall Upper Saddle River, NJ, 2002.
[25] G. Bastin, Quantitative analysis of metabolic networks and design of minimal bioreaction models, Revue Afr. Rech. Inform. Math. Appl. 9 (1) (2008) 41–55.
[26] F. Zamorano, A. Vande Wouwer, G. Bastin, A detailed metabolic flux analysis of an underdetermined network of CHO cells, J. Biotechnol. 150 (2010) 497–508.
[27] E. Walter, L. Pronzato, Identification of Parametric Models from Experimental Data, Springer Verlag, 1997.
[28] Z. Amribt, H. Niu, P. Bogaerts, Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures, Biochem. Eng. J. 70 (2013) 196–209.
[29] J. Ljunggren, L. Häggström, Specific growth rate as a parameter for tracing growth-limiting substances in animal cell cultures, J. Biotechnol. 42 (2) (1995) 163–175.
[30] Y.-H. Jeong, S.S. Wang, Role of glutamine in hybridoma cell culture: effects on cell growth, antibody production, and cell metabolism, Enzyme Microb. Technol. 17 (1) (1995) 47–55.
[31] F. Zhou, J.-X. Bi, A.-P. Zeng, J.-Q. Yuan, A macrokinetic and regulator model for myeloma cell culture based on metabolic balance of pathways, Process Biochem. 41 (10) (2006) 2207–2217.