

Dynamic metabolic models of CHO cell cultures through minimal sets of elementary flux modes

F. Zamorano^a A. Vande Wouwer^{a,1} R.M. Jungers^b G. Bastin^b

^a*Department of Automatic Control, University of Mons, Boulevard Dolez 31, 7000 Mons, Belgium - Email: francisca.zamorano/alain.vandewouwer@umons.ac.be*

^b*Centre for Systems Engineering and Applied Mechanics, Department of Mathematical Engineering, Catholic University of Louvain, 1348 Louvain-La-Neuve, Belgium - Email: georges.bastin@uclouvain.be*

Abstract

The concept of Elementary Flux Modes (EFMs) has been of central importance in a number of studies involving the analysis of metabolism. In [1] this concept is used to translate the metabolic networks of the different phases of CHO cell culture into macroscopic bioreactions linking extracellular substrates to products. However, a critical issue concerns the calculation of these elementary flux vectors, as their number combinatorially increases with the size of the metabolic network. In this study, a detailed metabolic network of CHO cells is considered, where the above-mentioned combinatorial explosion makes the computation of the elementary flux modes impossible. To alleviate this problem, a methodology proposed in [10] is used to compute a decomposition of admissible flux vectors in a minimal number of elementary flux modes without explicitly enumerating all of them. As a result, a set of macroscopic bioreactions linking the extracellular measured species is obtained at a very low computational expense. The procedure is repeated for the several cell life phases and a global model is built using a multi-model approach, which is able to successfully predict the evolution of experimental data.

Key words: elementary flux modes, dynamical modelling, metabolic networks, underdetermined systems, mammalian cells.

¹ Author to whom correspondence should be addressed

1 Introduction

Macroscopic models of bioprocesses have been used in many applications, ranging from simulation to estimation, optimization and control [4]. These models represent the conversion of substrates into products by a few macroscopic reactions, without taking the intracellular reaction network into consideration (black box representation).

These models can be derived using two main approaches. The first approach is essentially data-driven. Macroscopic models are derived solely from the experimental observation of the time evolution of a few extracellular components (substrates, products of interest, inhibiting compounds). Various techniques can be combined, including data analysis techniques such as principal component analysis to deduce the number of bioreactions and partial stoichiometry [5], and identification methods based on - whenever possible - decoupling techniques to estimate independently the stoichiometry and the kinetics (concept of C-identifiability) [6], [8]. In the second approach, the available prior knowledge about the metabolic network is exploited, and a macroscopic set of reactions is derived in agreement with the intracellular metabolism [7].

This is the second approach which is of interest in the present study, and particularly, the procedure devised in [14] where dynamic models are derived from the concept of Elementary Flux Modes (EFMs) for a metabolic network of CHO cells under balanced growth conditions. This latter assumption stipulates that the intracellular metabolites do not accumulate in the cell, or in other words, that the intracellular processes occur much faster than those happening outside the cell. In [1], this approach is further used to translate the metabolic networks of the different phases of the cell culture into macroscopic bioreactions linking extracellular substrates to products. However, a critical issue concerns the calculation of these elementary flux vectors, as their number combinatorially increases with the size of the metabolic network.

This latter point is one of the motivations behind this study, in which we consider a more detailed metabolic network of CHO cells developed by the authors in [17], where the above-mentioned combinatorial explosion makes the computation of the elementary flux modes impossible. To alleviate this problem, we apply a methodology to compute a decomposition of admissible flux vectors in a minimal number of elementary flux modes without explicitly enumerating all of them, as proposed by the authors in [10]. As a result, a set of macroscopic bioreactions linking the extracellular measured species is obtained at a very low computational cost. Further, the procedure is repeated for the different life cell phases (exponential growth, transition and death) to determine local dynamic models, which can then be assembled to form a global (piecewise) model for the entire culture. The multi-model approach has already

been applied successfully to describe the behavior of complex bioprocesses in other areas of applications such as wastewater treatment [15], or the culture of micro-algae in photo-bioreactors [11].

This paper is organized as follows. Section 2 introduces the general form of a dynamic model of batch cell cultures and the concept of Elementary Flux Modes (EFMs). The methodology for the computation of a minimal set of EFMs and the decomposition of an admissible flux distribution is presented in Section 3. A practical application of the methodology is presented in Section 4, where sets of macroscopic bioreactions are computed for each of the cell life phases of batch cultures of CHO-320 cells. Section 5 discusses the construction of a piecewise global dynamic model for the entire culture, based on the previous local models. Finally, Section 6 draws the main conclusions of this work.

2 Cell culture modelling

2.1 *Dynamics of a batch culture*

In general, cell cultivation in a batch process can be divided in at least three phases, according to the physiological states of the cells.

- the first phase corresponds to the exponential growth, where the concentration of the carbon source and all other substrates are in excess and there is sufficient dissolved oxygen allowing a rapid proliferation of the biomass. Lactate, alanine and ammonia are produced because of the high level of glucose and glutamine.
- the second phase is transition, where the sugar concentration decreases below a critical level and the produced lactate and alanine start to be consumed instead. There is sufficient dissolved oxygen in the medium in order to allow the oxidative pathways metabolize lactate and alanine and keep the cellular division, however in a less effective way.
- the third state corresponds to cellular death, where programmed cell death takes place upon exposure to stress encountered in the bioreactor. There could be various causes for apoptosis: nutrient depletion, waste byproduct accumulation, hypoxia, mechanical agitation, etc ([3]).

For a cell culture carried out in batch mode in a stirred tank reactor, the dynamics of substrates and products are described by:

$$\begin{aligned}\frac{dS}{dt} &= -v_s X(t) \\ \frac{dP}{dt} &= v_p X(t)\end{aligned}\tag{1}$$

where

- $X(t)$ is the biomass concentration,
- $S(t)$ is the vector of substrate concentrations,
- $P(t)$ is the vector of product concentrations,
- v_s is the vector of specific uptake rates,
- v_p is the vector of specific production rates.

Clearly, v_s and v_p are linear combinations of some of the (intracellular) metabolic fluxes v . Thus, by defining appropriate matrices N_s and N_p , the stoichiometric matrices for the extracellular substrates and final products, respectively, this relation can be expressed as:

$$\begin{aligned}v_s(t) &= N_s v(t) \\ v_p(t) &= N_p v(t).\end{aligned}\tag{2}$$

2.2 Metabolic network and elementary flux modes

The intracellular metabolism of living cells is usually represented by a metabolic network under the form of a hypergraph encoding a set of biochemical reactions. In this hypergraph, each node represents a particular intracellular metabolite and the edges represent the metabolic reactions or fluxes.

According to the pseudo steady-state assumption of metabolic flux analysis (MFA), it is assumed that the fluxes are balanced at each internal node, i.e. intracellular metabolites do not accumulate in the cell. This means that the net sum of production and consumption fluxes, weighted by their stoichiometric coefficients, is zero for each internal metabolite of the network. This steady-state balance around the internal metabolites is expressed by the algebraic relation:

$$\mathbf{N}\mathbf{v} = \mathbf{0} \quad \mathbf{v} \geq \mathbf{0}\tag{3}$$

where $\mathbf{v} = (v_1, v_2, \dots, v_n)^T$ is the n -dimensional column vector of fluxes and $\mathbf{N} = [n_{ij}]$ is the $m \times n$ stoichiometric matrix of the metabolic network (m is the number of internal metabolites and n is the number of fluxes). More precisely, a flux v_j denotes the rate of reaction j and a non-zero n_{ij} is the stoichiometric coefficient of the metabolite i in reaction j .

For a given metabolic network, the set S of possible flux distributions is the set of vectors \mathbf{v} that satisfy the linear system 3. This set S is the pointed polyhedral cone resulting from the intersection of the kernel of \mathbf{N} with the non-negative orthant. This implies that there exists a set of elementary flux vectors \mathbf{e}_i , the extreme rays (or edges) of this polyhedral cone, such that any flux distribution \mathbf{v} can be expressed as a non-negative linear combination of them:

$$\mathbf{v} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + \cdots + w_q\mathbf{e}_q \quad w_i \geq 0. \quad (4)$$

The $n \times q$ non-negative matrix \mathbf{E} with column vectors \mathbf{e}_i obviously satisfies $\mathbf{N}\mathbf{E} = 0$ and Equation (4) can be written in matrix form as

$$\mathbf{v} = \mathbf{E}\mathbf{w} \quad \text{with} \quad \mathbf{w} \triangleq (w_1, w_2, \dots, w_q)^T. \quad (5)$$

Thus, the elementary flux vectors are a way of representing the set of possible flux distributions. The dynamics of the concentration of each substrate and product in a batch reactor, where no exchange occurs with the outside environment, are written as follows:

$$\frac{d}{dt} \begin{pmatrix} S(t) \\ P(t) \end{pmatrix} = \begin{pmatrix} -N_s \\ N_p \end{pmatrix} vX \quad (6)$$

From 5 and 6, we obtain:

$$\frac{d}{dt} \begin{pmatrix} S(t) \\ P(t) \end{pmatrix} = \begin{pmatrix} -N_s \\ N_p \end{pmatrix} EwX \quad (7)$$

The product of stoichiometric matrices \mathbf{N}_s and \mathbf{N}_p times the elementary flux modes matrix \mathbf{E} yields the stoichiometric matrix for a set of macroscopic reactions, linking the extracellular substrates to the final products. Let us consider that the reaction scheme of the process involves N macroscopic reactions and M extracellular species, either substrates or products, with K being the $M \times N$ matrix for the stoichiometric coefficients.

$$K = \begin{pmatrix} -N_s \\ N_p \end{pmatrix} E \quad (8)$$

Then, if the vector ξ is defined as:

$$\xi = \begin{pmatrix} S(t) \\ P(t) \end{pmatrix}, \quad (9)$$

The dynamic model defined by the macroscopic bioreactions may be written

as:

$$\frac{d\xi}{dt} = K\mathbf{w}(t)X(t) = K\varphi(\xi, t) \quad (10)$$

where $\mathbf{w}(t)$ is the vector of the specific reaction rates w_i of the macroscopic bioreactions and φ is the vector of reaction rates.

3 Computation of the elementary flux modes and of minimal flux decomposition

3.1 Problem statement

A well known issue related to the EFMs representation is that the number of such vectors grows exponentially with the size of the network. This means that for detailed metabolic networks, such as the one considered in the following of this study, the computation of matrix \mathbf{E} becomes prohibitive.

In general, the decomposition of a flux distribution \mathbf{v} in the convex basis of elementary flux vectors \mathbf{e}_i does not necessitate the whole enumeration of the convex basis but requires only the knowledge of a few elementary vectors. Thus, the objective is to determine a minimal such decomposition. Nonetheless, when the vector \mathbf{v} is the solution of an underdetermined metabolic flux analysis problem, the situation is more complex, though it may be possible to find a decomposition with even less elementary flux modes. Indeed, it is not known a priori which vector, among all admissible flux distributions, is the one that can be decomposed in the minimal number of elementary flux modes. The information needed for computing these elementary vectors can be obtained directly from the stoichiometric matrix \mathbf{N} together with the extracellular measurements. Herein, this methodology is used to compute this decomposition without actually evaluating the whole convex basis, thanks to the convex programming techniques presented in [10].

3.2 Definition of some polytopes of interest

If we consider system 3 and take the constraints imposed by the extracellular measurements into account, it is possible to write

$$\begin{pmatrix} \mathbf{N} \\ \mathbf{N}_m \end{pmatrix} \mathbf{v} = \begin{pmatrix} \mathbf{0} \\ \mathbf{v}_m \end{pmatrix} \quad \mathbf{v} \geq \mathbf{0}. \quad (11)$$

for a given metabolic network and a given set of measurements. \mathbf{N}_m stands for the stoichiometric matrix of the extracellular species and \mathbf{v}_m is the vector of measurements.

As demonstrated in [10, 13], any admissible flux distribution \mathbf{v} can be expressed as a convex combination of $n-m$ elementary flux vectors \mathbf{e}_i . $n-m$ corresponds to the degrees of freedom of the system, if \mathbf{N} and \mathbf{N}_m are full rank matrices. Notice that the decomposition of \mathbf{v} in the convex basis $\{\mathbf{e}_i\}$ is not unique.

Moreover, if the number of measurements p is smaller than $n-m$, then there is at least one vector \mathbf{v}^* that can be expressed as a convex combination of only p elementary flux vectors. Hence, the objective is to determine such a decomposition in a minimal number of elementary flux vectors $\{\mathbf{e}_i\}$.

Using Equation (5), system 11 is equivalent to the system:

$$\begin{pmatrix} \mathbf{NE} \\ \mathbf{N}_m\mathbf{E} \end{pmatrix} \mathbf{w} = \begin{pmatrix} \mathbf{0} \\ \mathbf{v}_m \end{pmatrix} \quad \mathbf{w} \geq 0. \quad (12)$$

We observe that the first equation $\mathbf{NEw} = \mathbf{0}$ is trivially satisfied independently of \mathbf{w} since by definition $\mathbf{NE} = \mathbf{0}$. Hence, system 12 may be reduced to the second equation:

$$\mathbf{N}_m\mathbf{Ew} = \mathbf{v}_m \quad \mathbf{w} \geq 0. \quad (13)$$

or equivalently written

$$\begin{pmatrix} N_mE & -v_m \end{pmatrix} \begin{pmatrix} w \\ 1 \end{pmatrix} = 0. \quad (14)$$

In this form, it is clear that the set of admissible weighting vectors \mathbf{w} that satisfy Equation (13) constitutes a convex polytope that will be denoted \mathcal{H} . Therefore, there exists a set of appropriate edge vectors \mathbf{h}_i such that any arbitrary convex combination of the form:

$$\mathbf{w} = \sum_i \beta_i \mathbf{h}_i \quad \beta_i \geq 0 \quad \sum_i \beta_i = 1 \quad (15)$$

is necessarily an admissible \mathbf{w} satisfying Equation (13). The convex basis vectors \mathbf{h}_i have a critical property : the number of non-zero entries in these vectors is equal to the number of measurements p .

From a metabolic viewpoint, each vector \mathbf{h}_i is a solution \mathbf{w} of Equation (13). Vectors \mathbf{Eh}_i correspond to *minimal flux distributions* \mathbf{v} :

$$\hat{\mathbf{v}}_i = \mathbf{Eh}_i \quad \mathbf{v} \in \mathcal{F}. \quad (16)$$

Each minimal flux distribution \hat{v}_i represent the simplest pathways that satisfy the pseudo-steady state assumption and the constraints imposed by the extra-cellular measurements. Equation (16) implies that a minimal flux distribution (in terms of EFMs) can be obtained by different combinations of EFMs and in turn, of metabolic pathways. This will be illustrated further in the article, when we will assess the calculation procedure of the minimal set of EFMs.

3.3 Decomposing v in a convex basis

As already stated, the number of distinct extreme rays or cone vertices that are generated when computing the cone S may become very large because it combinatorially increases with the size of the underlying metabolic network. It is also the case for the number of vectors h_i that are vertices of the polytope \mathcal{H} .

We apply here the method presented in [10] to decompose a flux distribution \mathbf{v} in a minimal number ($p < n - m$) of elementary flux modes. To this end, we introduce yet another cone $\mathcal{K} \subset \mathbb{R}^p$. This cone is the projection of S by the matrix \mathbf{N}_m : $\mathcal{K} = \{\mathbf{y} = \mathbf{N}_m \mathbf{v} : \mathbf{v} \geq \mathbf{0}, \mathbf{N} \mathbf{v} = \mathbf{0}\}$.

We know that the vector \mathbf{v}_m is in \mathcal{K} because of Equation (11). So, \mathbf{v}_m can be expressed as a convex combination of p extreme rays y_i of cone \mathcal{K} (because \mathcal{K} has dimension p).

$$\mathbf{v}_m = \sum_i^p \alpha_i y_i \quad \alpha_i \geq 0 \quad \sum_i \alpha_i = 1 \quad (17)$$

Now, the extreme rays of \mathcal{K} are the projections of extreme rays \mathbf{e}_i of S under the matrix \mathbf{N}_m . This implies that the corresponding convex combination of the \mathbf{e}_i gives us the required \mathbf{v} . In other words, if y_i is an extreme ray of the projected cone \mathcal{K} , then \mathbf{e}_i is an extreme ray of cone S .

$$\mathbf{v}_m = \mathbf{N}_m \mathbf{v} \quad \Rightarrow \quad y_i = \mathbf{N}_m \mathbf{e}_i \quad (18)$$

As \mathbf{v}_m has been decomposed in p extreme rays in 17, a decomposition in the extreme rays of cone S is also achieved

$$\mathbf{v}_m = \sum_1^p \alpha_i \mathbf{N}_m \mathbf{e}_i = \mathbf{N}_m \sum_1^p \alpha_i \mathbf{e}_i \quad (19)$$

and thus, \mathbf{v} is decomposed in a minimal set of p elementary flux vectors.

$$\mathbf{v} = \sum_1^p \alpha_i \mathbf{e}_i \quad (20)$$

Table 1
Specific uptake/excretion rates for the three life phases

Specie	Exponential Growth Phase	Transition Phase	Death Phase
Glucose	$-1.6383 \pm 0.244e^{-1}$	-	-
Glutamine	$-4.7922 \pm 1.107e^{-2}$	$-1.4582 \pm 7.678e^{-3}$	$-8.9527 \pm 62.97e^{-4}$
Arginine	$-1.7381 \pm 1.659e^{-3}$	$-8.9108 \pm 0.271e^{-5}$	$5.1413 \pm 16.66e^{-5}$
Asparagine	$-1.2354 \pm 0.203e^{-3}$	$-1.7873 \pm 3.316e^{-5}$	$6.0603 \pm 10.76e^{-5}$
Aspartate	$-2.7112 \pm 4.304e^{-4}$	$-4.6172 \pm 4.601e^{-4}$	$-7.4483 \pm 20.77e^{-5}$
Isoleucine	$-1.7422 \pm 0.521e^{-3}$	$-4.1392 \pm 2.982e^{-4}$	$-1.7901 \pm 2.393e^{-4}$
Leucine	$-2.9556 \pm 0.610e^{-3}$	$-3.1471 \pm 2.109e^{-4}$	$-1.1150 \pm 8.286e^{-5}$
Lysine	$-3.0675 \pm 0.839e^{-3}$	$-2.7181 \pm 1.628e^{-4}$	$-1.9790e^{-5}$
Methionine	$-8.1777 \pm 1.777e^{-4}$	$-6.6621 \pm 6.668e^{-5}$	-
Phenylalanine	$-1.1747 \pm 0.309e^{-3}$	$-1.0902 \pm 0.832e^{-4}$	$-4.6531 \pm 18.38e^{-5}$
Serine	$-1.0054 \pm 0.499e^{-3}$	$-4.4716 \pm 3.295e^{-4}$	$1.5091 \pm 4.229e^{-4}$
Threonine	$-1.5358 \pm 0.928e^{-3}$	$-1.2195 \pm 2.679e^{-4}$	$-1.2157 \pm 7.073e^{-4}$
Tyrosine	$-8.7011 \pm 3.171e^{-4}$	$-8.5351 \pm 7.158e^{-5}$	$-1.2778 \pm 2.528e^{-4}$
Valine	$-2.0238 \pm 0.664e^{-3}$	$-2.7412 \pm 2.827e^{-4}$	$-1.5805 \pm 4.369e^{-4}$
Lactate	$2.9880 \pm 0.599e^{-1}$	$-2.0169 \pm 4.971e^{-2}$	$-3.8359 \pm 3.793e^{-2}$
NH_4^+	$3.8858 \pm 0.954e^{-2}$	$1.4428 \pm 8.118e^{-3}$	$1.5064 \pm 10.11e^{-3}$
Glycine	$2.6166 \pm 0.847e^{-3}$	$4.6293 \pm 14.47e^{-4}$	$-5.3266 \pm 22.34e^{-4}$
Alanine	$1.0273 \pm 0.144e^{-2}$	$-1.1855 \pm 56.37e^{-4}$	$-2.1682 \pm 1.527e^{-3}$
Glutamate	$3.0143 \pm 1.942e^{-3}$	$-9.7355 \pm 8.015e^{-4}$	$-9.0875 \pm 11.29e^{-4}$

For more details on the algorithm and the theory behind it, the reader is referred to references [9] and [10].

4 Macroscopic bioreactions for cultures of CHO cells

In this section we apply the methodology described above to three detailed (and underdetermined) metabolic networks describing the metabolism of CHO-320 cells. Each network represents the metabolism of one of the life phases of a cell in a batch culture: exponential growth, transition and death. For each of these networks, a minimal set of elementary flux modes is computed by applying the procedure described in Section 3. For reasons of space, the details of matrices \mathbf{N} for the growth, transition and death phases are not presented. To retrieve the list of reactions describing the different phases, the reader is referred to references [17] and [16].

To apply this procedure we need to define stoichiometric matrices \mathbf{N} and \mathbf{N}_m and the vector of extracellular measurements \mathbf{v}_m for each phase. The set of experimental data contains, respectively 19, 18 and 17 extracellular measurements for the exponential growth, transition and death phases. These vectors of experimental measurements \mathbf{v}_m are listed in Table 1.

The dimension of the vector \mathbf{v}_m will then determine the dimension of the matrix containing the minimal set of vectors \mathbf{e}_i (\mathbf{E}_{min}). Each elementary vector defines a metabolic path linking extracellular substrates to final products, which can be translated into a macroscopic reaction. Proceeding in this way, the set of 19 macroscopic reactions presented in Table 2 describes the main

Table 2
Macroscopic Reactions for the exponential growth phase

EFM	Macroscopic Reaction
e_1	$Tyr \rightarrow Glu + 4 CO_2$
e_2	$Glucose + 1.7 Gln \rightarrow 1.7 Lactate + 3.3 NH_4^+ + 6CO_2$
e_3	$Gln \rightarrow Glu + NH_4^+$
e_4	$Ser \rightarrow Gly$
e_5	$Asn \rightarrow Lactate + Urea$
e_6	$3.3 Glucose + 6.4 Gln + Asn + 1.9 Asp + 1.2 Arg + 1.4 Thr + 1.7 Lys + 1.6 Val$ $+ 1.3 Ile + 6.5 Leu + 1.7 Phe + Met + 1.2 Pro + 0.3 Trp + 0.5 His + 0.2 Eth + 0.4 Cho$ $\rightarrow 25.8 Biomass + 2.9 Ala + 5 NH_4^+ + 13.1 CO_2$
e_7	$15.2 Glucose + 7.7 Gln + 4.7 Asn + 14.2 Arg + 6.5 Thr + 13.5 Lys + 7.2 Val + 5.8 Ile$ $+ 10 Leu + 7.8 Phe + 4.6 Met + 5.7 Pro + 25.9 Trp + 2.5 His + 0.7 Eth + 2 Cho$ $\rightarrow 118.6 Biomass + 15.8 Ala + 8.6 Urea + NH_4^+ + 134.2 CO_2$
e_8	$Glucose + 1.7 Val \rightarrow 3.3 Lactate + 1.7 NH_4^+ + 4.3 CO_2$
e_9	$2.1 Glucose + 2.5 Gln + Asn + 1.2 Arg + 3.4 Thr + 1.5 Lys + 1.6 Val + 4.2 Ile$ $+ 5.1 Leu + 1.7 Phe + 1 Met + 1.2 Pro + 0.3 Trp + 0.5 His + 0.2 Eth + 0.4 Cho$ $\rightarrow 25.8 Biomass + NH_4^+ + 4.7 CO_2$
e_{10}	$15.2 Glucose + 7.7 Gln + 4.7 Asn + 9.4 Arg + 6.5 Thr + 6.5 Lys + 7.2 Val + 68.7 Ile$ $+ 10 Leu + 7.8 Phe + 4.6 Met + 5.7 Pro + 1.5 Trp + 2.5 His + 0.7 Eth + 2 Cho$ $\rightarrow 118.6 Biomass + 19 Lactate + 30.6 Ala + 3.8 Urea + NH_4^+ + 75.9 CO_2$
e_{11}	$Gln \rightarrow Ala + NH_4^+ + 2 CO_2$
e_{12}	$6.6 Glucose + 2.4 Gln + Asn + 1.2 Arg + 1.4 Thr + 8.1 Lys + 1.5 Val + 1.2 Ile$ $+ 2.1 Leu + 1.7 Phe + 2.4 Met + 1.2 Pro + 0.3 Trp + 0.5 His + 0.1 Eth + 0.4 Cho$ $\rightarrow 25.2 Biomass + 4.3 Gly + 3.9 NH_4^+ + 21 CO_2$
e_{13}	$Gln \rightarrow Lactate + Urea + CO_2$
e_{14}	$Glucose \rightarrow 1.7 Lactate + CO_2$
e_{15}	$6.8 Glucose + 1.7 Gln + Asn + 6.2 Arg + 1.4 Thr + 1.4 Lys + 1.6 Val + 1.3 Ile$ $+ 6.7 Leu + 1.7 Phe + Met + 1.2 Pro + 0.3 Trp + 0.5 His + 0.2 Eth + 0.4 Cho$ $\rightarrow 25.8 Biomass + 7 Lactate + 4.9 Urea + 6 NH_4^+ + 14.5 CO_2$
e_{16}	$49.6 Glucose + 7.7 Gln + 4.7 Asn + 22.7 Arg + 6.5 Thr + 37.9 Lys + 7.2 Val + 5.8 Ile$ $+ 10 Leu + 7.8 Phe + 4.6 Met + 5.7 Pro + 1.5 Trp + 2.5 His + 0.7 Eth + 2 Cho$ $\rightarrow 118.6 Biomass + 57.3 Gly + 17.1 Urea + NH_4^+ + 136.9 CO_2$
e_{17}	$1.2 Glucose + Arg \rightarrow 3 Lactate + Urea + 2 NH_4^+ + 3.2 CO_2$
e_{18}	$7.3 Glucose + 3.5 Gln + Asn + 1.2 Arg + 1.4 Thr + 8.1 Lys + 1.5 Val + 1.2 Ile$ $+ 2.1 Leu + 1.7 Phe + Met + 1.2 Pro + 0.3 Trp + 0.5 His + 0.1 Eth + 0.4 Cho$ $\rightarrow 25.2 Biomass + 4.9 Lactate + 8.7 NH_4^+ + 22.2 CO_2$
e_{19}	$Glucose + 1.7 Gln \rightarrow 3.3 Lactate + 3.3 NH_4^+ + 4.3 CO_2$

metabolic processes occurring during the growth phase.

Thus, the minimal set of EFMs obtained for the exponential growth phase has been translated into a set of macroscopic bioreactions, from which a general model can be deduced. At this point, it is worth noticing that each run of the model reduction algorithm will yield different minimal sets of EFMs, thus giving different sets of macroscopic reactions. The reader is reminded about vectors \mathbf{h}_i and Equation ((16)) which states that the pseudo-steady state assumption and the constraints imposed by the extracellular measurements can be satisfied by different minimal flux distributions \hat{v}_i . Hence, each time the calculation procedure is launched, a particular vector \mathbf{h}_i is found and in turn, a minimal flux distribution \hat{v}_i .

An estimation of the reaction rates for the macroreactions are obtained from Equation ((13)). As $\mathbf{N}_m \mathbf{E}$ is a $p \times p$ matrix, then \mathbf{w} is easily obtained from:

$$\mathbf{w} = (\mathbf{N}_m \mathbf{E})^{-1} \mathbf{v}_m \quad (21)$$

The resulting reaction rates w_i for each of the macroscopic reactions taking

Table 3
Reaction rates for the three sets of macroscopic reactions

Reaction Rate	Exponential Growth Phase	Transition Phase	Death Phase
w_1	$7.6104e^{-4}$	$1.4006e^{-4}$	$1.9790e^{-5}$
w_2	$3.3737e^{-3}$	$3.8043e^{-4}$	$1.9790e^{-5}$
w_3	$1.9371e^{-4}$	$4.1514e^{-5}$	$5.8637e^{-6}$
w_4	$9.2342e^{-4}$	$5.7292e^{-4}$	$2.0670e^{-4}$
w_5	$5.2333e^{-4}$	$8.3348e^{-7}$	$9.2354e^{-5}$
w_6	$1.7039e^{-4}$	$5.1261e^{-5}$	$2.4921e^{-5}$
w_7	$7.8766e^{-6}$	$2.9480e^{-4}$	$1.1288e^{-4}$
w_8	$6.1698e^{-4}$	$3.2716e^{-4}$	$3.6726e^{-2}$
w_9	$1.8203e^{-4}$	$3.3813e^{-2}$	$5.9150e^{-4}$
w_{10}	$4.0478e^{-6}$	$2.7660e^{-4}$	$6.2670e^{-5}$
w_{11}	$8.0719e^{-3}$	$5.4907e^{-5}$	$7.6960e^{-5}$
w_{12}	$9.4097e^{-5}$	$5.6892e^{-4}$	$7.3303e^{-7}$
w_{13}	$2.3533e^{-2}$	$2.6781e^{-4}$	$2.4188e^{-5}$
w_{14}	$1.7277e^{-1}$	$2.7978e^{-5}$	$2.1403e^{-4}$
w_{15}	$1.0689e^{-5}$	$2.4642e^{-5}$	$7.2050e^{-4}$
w_{16}	$1.5793e^{-5}$	$5.4348e^{-5}$	$7.4763e^{-5}$
w_{17}	$1.0180e^{-3}$	$7.0945e^{-5}$	$3.3575e^{-4}$
w_{18}	$8.1690e^{-6}$	$9.0659e^{-5}$	-
w_{19}	$6.4727e^{-3}$	-	-

Table 4
Macroscopic reactions for the transition phase

EFM	Macroscopic Reaction
e_1	$Tyr \rightarrow NH_4^+ + 9 CO_2$
e_2	$Gln \rightarrow 2 NH_4^+ + 3 CO_2$
e_3	$3 Leu + Met \rightarrow 2 Urea + 20 CO_2$
e_4	$Ser \rightarrow Gly$
e_5	$Asn \rightarrow Urea + 3 CO_2$
e_6	$13.7 Lactate + 2.2 Gln + Asn + 2.6 Asp + 1.2 Arg + 1.4 Thr + 1.4 Lys + 1.6 Val + 1.3 Ile + 2.2 Leu + 1.7 Phe + Met + 1.9 Ala + 4.5 Glu + 1.2 Pro + 0.3 Trp + 0.5 His + 0.2 Etn + 0.4 cho \rightarrow 25.8 Biomass + 0.7 Urea + 23.8 CO_2$
e_7	$Ala + Asp \rightarrow Urea + 4 CO_2$
e_8	$Val \rightarrow Gly + 2 CO_2$
e_9	$Lactate \rightarrow 3 CO_2$
e_{10}	$Ile + Leu \rightarrow Urea + 9 CO_2$
e_{11}	$Lys + 2 Phe \rightarrow 2 Urea + 18 CO_2$
e_{12}	$Gln \rightarrow Urea + 4 CO_2$
e_{13}	$Lys + 2 Glu \rightarrow 2 Urea + 10 CO_2$
e_{14}	$Lys + 2 Val \rightarrow 2 Urea + 10 CO_2$
e_{15}	$Thr + Ile \rightarrow Urea + 9 CO_2$
e_{16}	$Thr + 1.5 Lys \rightarrow 2 Urea + 9 CO_2$
e_{17}	$2 Asp + Lys \rightarrow 2 Urea + 8 CO_2$
e_{18}	$Lys + 2 Ile \rightarrow 2 Urea + 12 CO_2$

place during the exponential growth phase are listed in Table 3.

In the same way, a minimal set of elementary vectors for the transition phase is obtained. The number of extreme rays \mathbf{e}_i matches the number of entries in the vector \mathbf{v}_m . Thus, the 18 resulting elementary flux vectors are presented in Table 4, from which a set of macroscopic reactions is defined. Notice that the metabolic changes corresponding to this phase of the culture are reflected by the macroscopic reactions obtained. Lactate, Alanine and Glutamate are now consumed as substrates, and since glucose is depleted, it no longer appears as a substrate. The estimated reaction rates w_i are listed in Table 3.

The same procedure is now applied to the reaction network defining the

Table 5
Macroscopic reactions for the death
phase

EFM	Macroscopic Reaction
e_1	$Ala + Gly \rightarrow Asn$
e_2	$Lys + 2 Gly \rightarrow 2 Urea + 6 CO_2$
e_3	$Asp + 3 Leu + Pro \rightarrow Ser + Arg + 18 CO_2$
e_4	$Gly \rightarrow NH_4^+ + CO_2$
e_5	$Thr \rightarrow NH_4^+ + 3 CO_2$
e_6	$3 Ala + 1 Pro \rightarrow Arg + 8 CO_2$
e_7	$Glu \rightarrow Ser + 2 CO_2$
e_8	$Lactate \rightarrow 3 CO_2$
e_9	$Gln \rightarrow 2 NH_4^+ + 5 CO_2$
e_{10}	$2 Val \rightarrow Urea + 9 CO_2$
e_{11}	$2 Ile \rightarrow Urea + 11 CO_2$
e_{12}	$Asp + Glu \rightarrow Urea + 8 CO_2$
e_{13}	$2 Phe \rightarrow Urea + 17 CO_2$
e_{14}	$Glu \rightarrow NH_4^+ + 5 CO_2$
e_{15}	$2 Ala \rightarrow Urea + 5 CO_2$
e_{16}	$Tyr \rightarrow NH_4^+ + 9 CO_2$
e_{17}	$Ala \rightarrow NH_4^+ + 3 CO_2$

metabolism of the death phase of the culture. As vector \mathbf{v}_m includes 17 experimental measurements, the same number of elementary vectors are obtained. This set of extreme rays generate the corresponding macroscopic bioreactions presented in Table 5. Now that cells are dying, there is no production of biomass any longer and the metabolism is centered in the production of energy with CO_2 as main product. The resulting reaction rates w_i are presented in Table 3.

5 A piecewise dynamic model of CHO-320 cells

An estimation of the maximum reaction rates have been obtained for each of the cell life phase (see Table 3). To take account of possible substrate limitations, and guarantee concentration positivity during model simulation, it is suggested to modulate these maximum rates with Monod factors.

$$r_i = w_i \frac{S_1}{(k_{s_1} + S_1)} \frac{S_2}{(k_{s_2} + S_2)} \cdots \frac{S_z}{(k_{s_{n_i}} + S_{n_i})} \quad (22)$$

Subindex n_i indicates the number of substrates participating in reaction i .

Thus, the dynamical model can be rewritten as:

$$\frac{d\xi}{dt} = KrX. \quad (23)$$

In order to complete the model, it is necessary to select numerical values for the half-saturation constants of substrates. Our aim in this study is to propose a model structure and not to estimate these values from experimental data.

Clearly, our database is insufficient for this latter purpose. Here, we select somewhat arbitrary values equal to 0.1 mM, i.e., values small enough to not interfere during the growth phase but large enough to avoid stiffness difficulties in the simulation of the model differential equations. The same idea has been used in [14].

Consequently, a local dynamic model is obtained for each of the life phases. In Figures 1 and 2 the prediction of the three different models is presented. As expected, all three models fit well the available data in their respective time span.

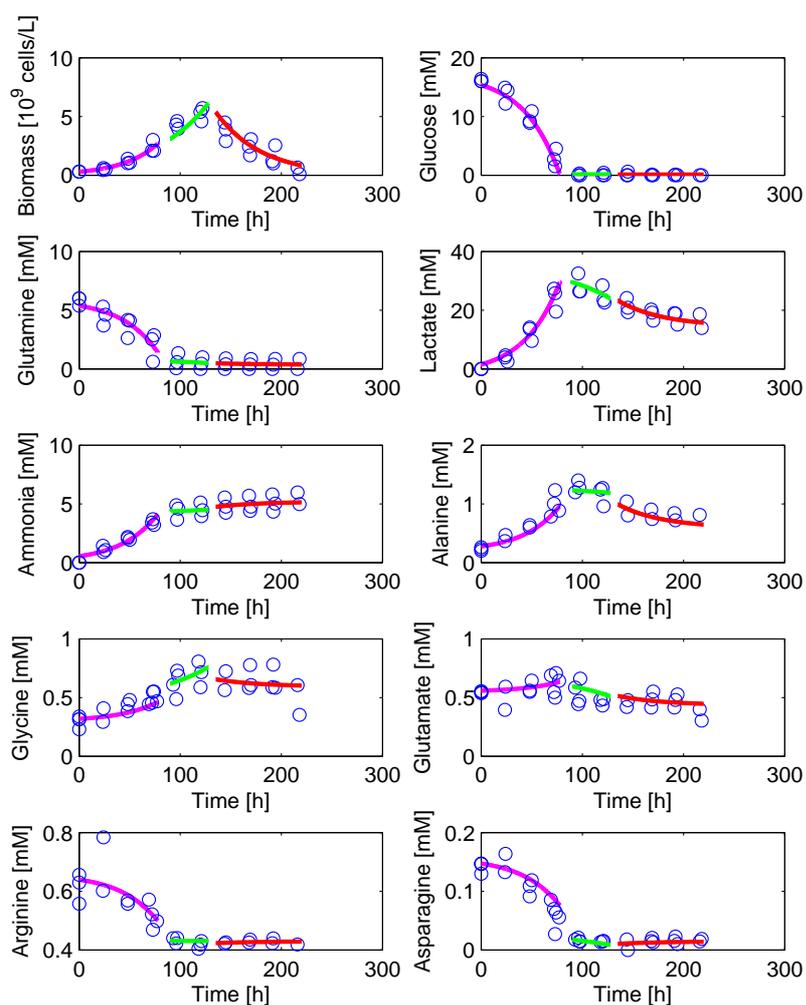


Fig. 1. Prediction of the three different models - Biomass and 9 first components - Magenta: growth phase model; green: transition phase model; red: death phase model.

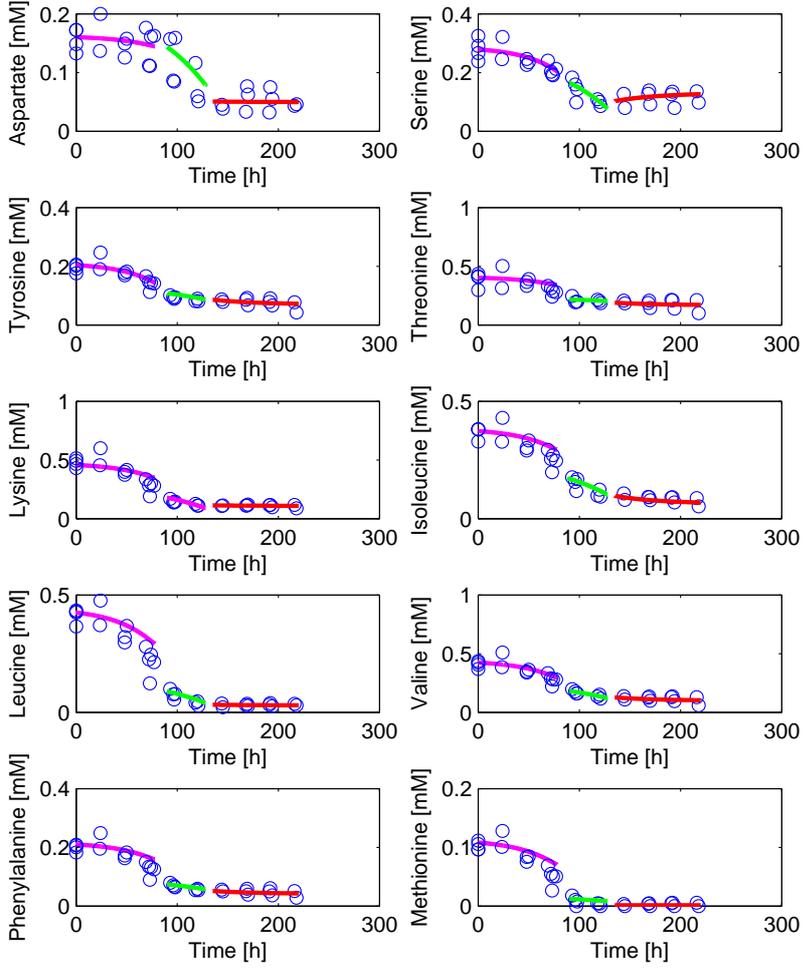


Fig. 2. Prediction of the three different models - 10 remaining components - Magenta: growth phase model; green: transition phase model; red: death phase model.

A global model describing the complete dynamics of a CHO-320 cell culture, can be defined as an interpolation between the three models obtained in the previous section for growth, transition and death phases. The influence of each model is controlled by means of weighting functions ϕ_g , ϕ_m and ϕ_d (see [12] for more on the multi-model approach), such that the global model is formulated as follows:

$$\frac{d\xi}{dt} = \phi_g \frac{d\xi_g}{dt} + \phi_m \frac{d\xi_m}{dt} + \phi_d \frac{d\xi_d}{dt}. \quad (24)$$

Many local basis functions could be used. One of the simplest option is provided by linear functions of time ϕ_g , ϕ_m and ϕ_d , as shown in Figure 3). In order to blend the three models, the first transition occurs in a time span starting from 75 hours until 95 hours, time of the culture at which glucose is depleted. The second transition starts at $t = 123$ hours and finishes at $t = 143$ hours, a

time range where some kind of stress in the culture medium triggers cellular apoptosis or programmed cell death. The time selection for the first model transition is derived from the fact that the last measurement points of the growth phase occurs at 72-74 hours and the first measurement points of the transition phase are at 96-98 hours. In the same way, the time selection for the second transition comes from the last measurement points of the transition phase and the first points of the death phase, at 120-122 hours and 144-145 hours, respectively. The simulation results are presented in Figures 4 and 5.

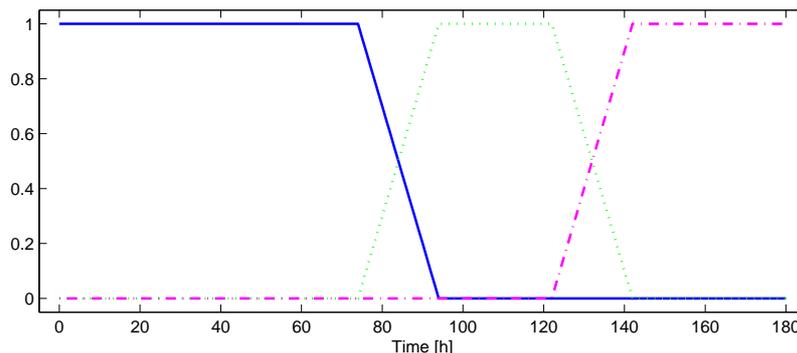


Fig. 3. Linear switching functions.

While the model reproduces quite well the evolution of cellular density and main substrates and products, it fails to provide good results for all metabolites. Indeed, at the end of the growth phase, the model stops predicting the consumption of certain amino acids such as Arginine, Asparagine, Threonine, Leucine, Isoleucine, Valine, Phenylalanine and Methionine. Hence, the transition phase model starts with wrong initial concentrations and is not able to catch up with the real data.

To alleviate the problem of the erroneous model prediction for certain amino acids, we search for those macroscopic reactions where these amino acids participate. It appears that all nine amino acids participate in almost exactly the same reactions. In addition, in all these reactions glucose appears as a substrate. The kinetic expressions of the reaction rates r are modeled by Monod kinetics and thus, they depend on glucose concentration as a multiplication factor. Consequently, the concentration of these amino-acids do not vary any longer, as the glucose concentration depletes.

Clearly, the early disappearance of glucose from the medium is the cause of this problem. The exponential growth phase model presented in Table 2 has been determined from the experimental measurements collected between 0 and 80 hours. Due to the reduced number of measurement points, the error in the determination of the specific uptake rate of glucose (and all other species) might be significant. Indeed, a smaller consumption rate of glucose would

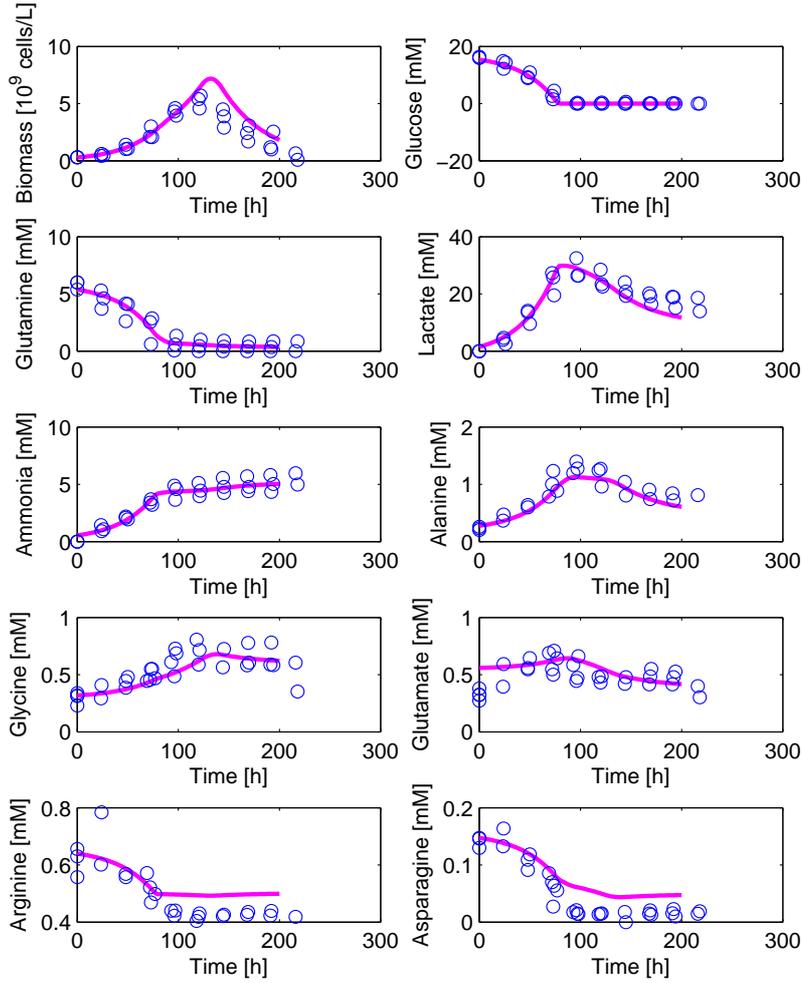


Fig. 4. Global model validation (using linear weighting functions).

maybe yield a macroscopic model capable of a better fit for the amino acids in question. Thus, we selected from Table 1 a smaller specific uptake rate of glucose within the confidence interval of the estimated value, so as to compute a new minimal set of EFMs, and in turn, a new model for the exponential growth phase. The set of macroscopic reactions obtained along with their corresponding reaction rates w_i are presented in Table 6.

The global model is constructed as before using linear functions of time. Now, the first transition starts at $t = 85$ hours until $t = 100$ hours. In this way, the overlapping of the exponential growth and transition phase models occurs later, allowing the first to have an influence on the global model for a longer time. The time span of the second transition remains identical, starting at $t = 123$ hours and finishing at $t = 143$ hours. The simulation results are

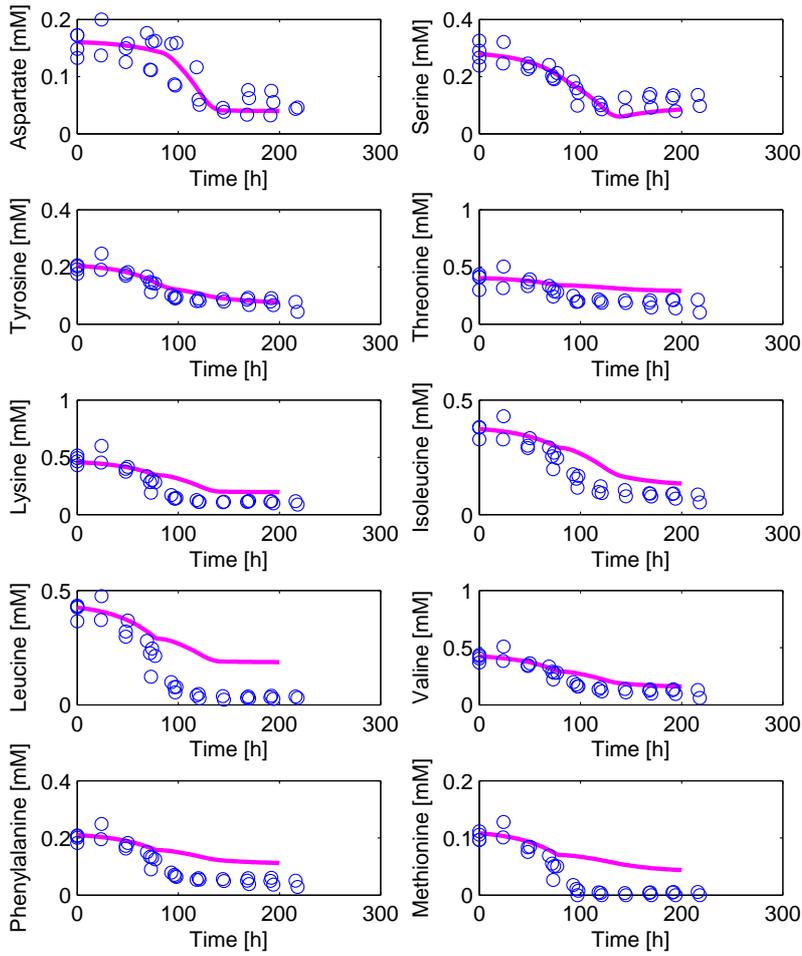


Fig. 5. Global model validation (using linear weighting functions).

presented in Figures 6 and 7.

6 Conclusions

Dynamic modeling of animal cell cultures is a delicate task that has attracted considerable attention in the last decades, with models ranging from low-dimensional macroscopic models to complex models mixing knowledge about the metabolic network and kinetic models [2].

In this study, a procedure for the derivation of macroscopic dynamic mod-

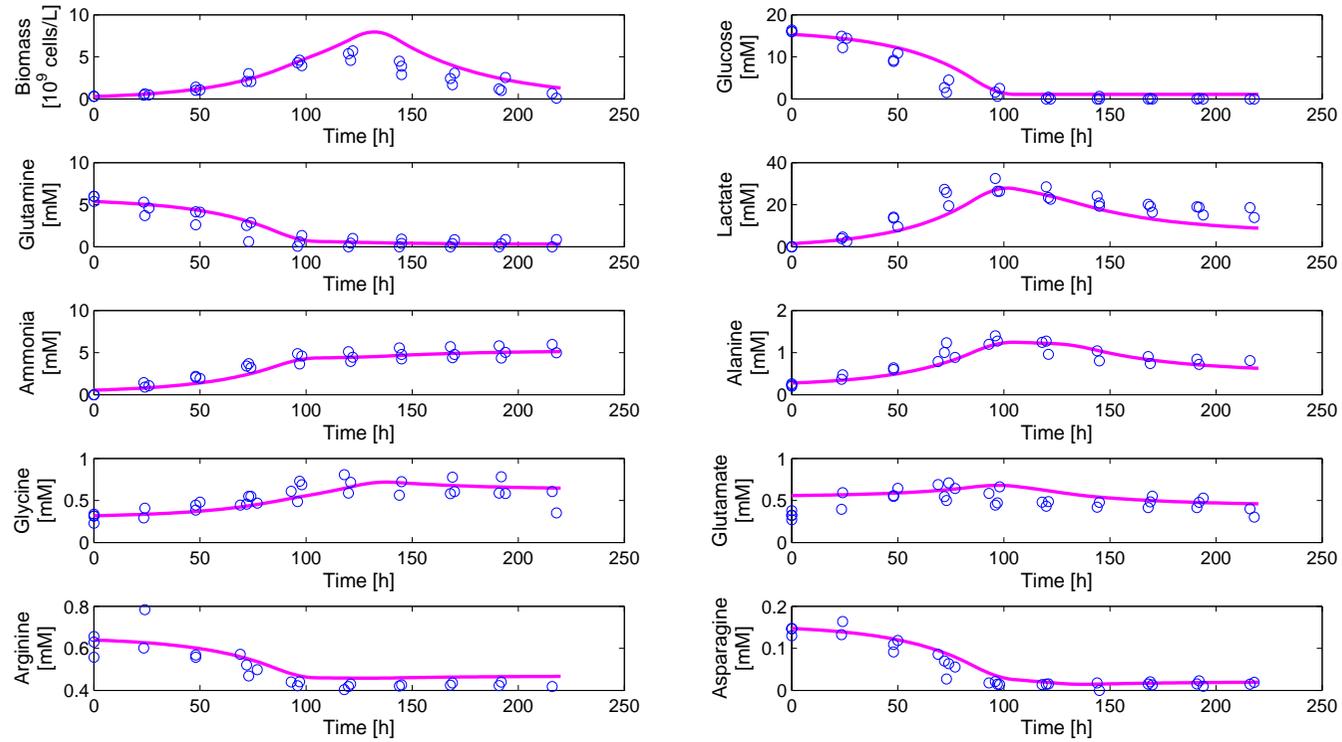


Fig. 6. Global model validation using a smaller glucose uptake rate.

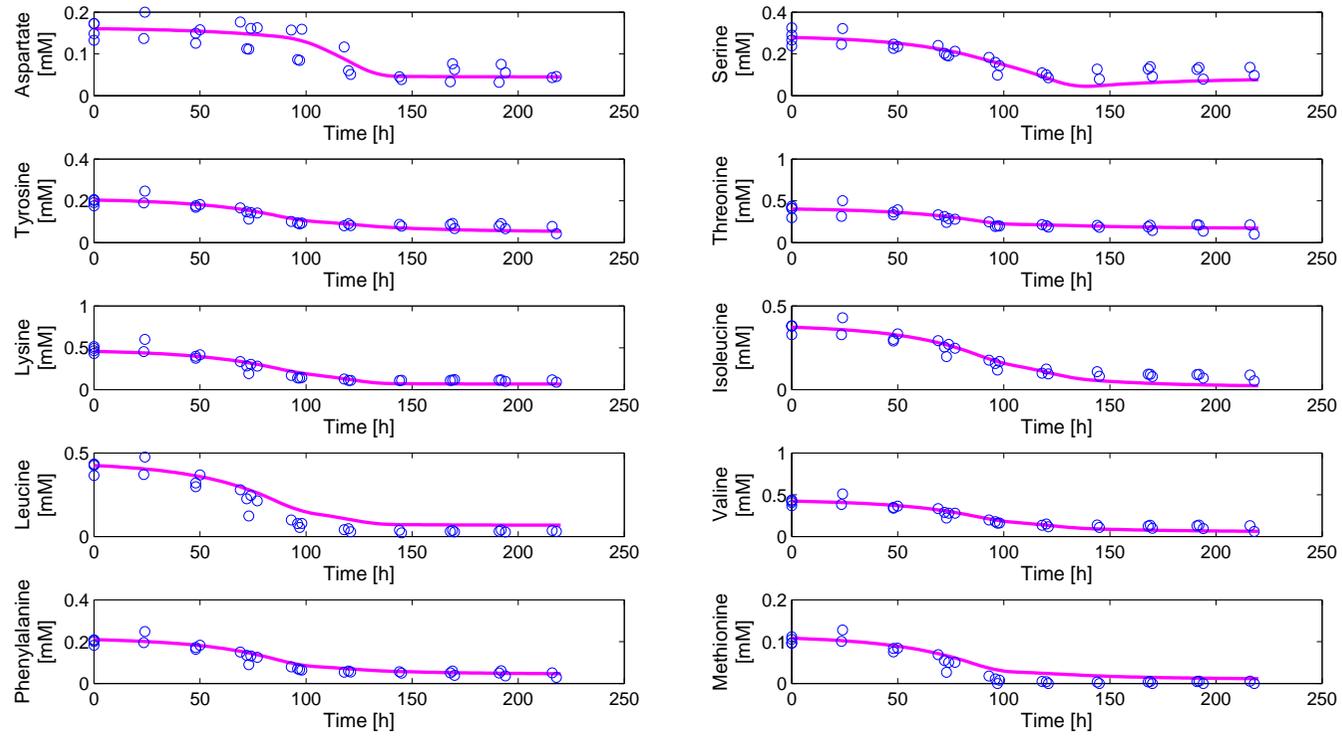


Fig. 7. Global model validation using a smaller glucose uptake rate.

Table 6
 Macroscopic Reactions for the exponential growth phase

EFM	Macroscopic Reaction	Reaction Rate w
e_1	$16.2 \text{ Glucose} + 3.1 \text{ Gln} + 1.9 \text{ Asn} + 3.4 \text{ Asp} + 2.2 \text{ Arg} + 1.4 \text{ Tyr} + 2.6 \text{ Thr}$ $+15.4 \text{ Lys} + 2.9 \text{ Val} + 2.3 \text{ Ile} + 4 \text{ Leu} + 1.7 \text{ Phe} + \text{Met} + 3.1 \text{ Pro} + 0.6 \text{ Trp}$ $+ \text{His} + 0.3 \text{ Eth} + 0.8 \text{ cho} \rightarrow 47.1 \text{ Biomass} + 14.9 \text{ Gly} + 43.6 \text{ CO}_2$	$3.3637e^{-3}$
e_2	$\text{Lysine} \rightarrow 2 \text{ NH}_4^+ + 6 \text{ CO}_2$	$1.2885e^{-3}$
e_3	$\text{Val} \rightarrow \text{Lactate} + \text{NH}_4^+ + 2 \text{ CO}_2$	$6.8944e^{-3}$
e_4	$\text{Gln} \rightarrow \text{Lactate} + 2 \text{ NH}_4^+ + 2 \text{ CO}_2$	$5.8354e^{-2}$
e_5	$\text{Glucose} \rightarrow 2 \text{ Lactate}$	$8.9550e^{-1}$
e_6	$\text{Ser} + \text{Arg} \rightarrow \text{Ala} + \text{Glu} + \text{NH}_4^+ + \text{Urea}$	$2.6505e^{-4}$
e_7	$\text{Gln} \rightarrow \text{Ala} + \text{NH}_4^+ + 2 \text{ CO}_2$	$3.2234e^{-3}$
e_8	$\text{Ile} \rightarrow \text{Glu} + \text{CO}_2$	$3.3120e^{-5}$
e_9	$\text{Glucose} \rightarrow 6 \text{ CO}_2$	$1.5911e^{-2}$
e_{10}	$\text{Thr} \rightarrow \text{Gly} + 2 \text{ CO}_2$	$6.4803e^{-5}$
e_{11}	$\text{Asn} + \text{Arg} \rightarrow 2 \text{ Ala} + 2 \text{ Urea} + 2 \text{ CO}_2$	$2.6428e^{-4}$
e_{12}	$\text{Ser} + 4 \text{ Arg} + \text{Met} \rightarrow 6 \text{ Ala} + 6 \text{ Urea} + 7 \text{ CO}_2$	$2.1242e^{-5}$
e_{13}	$\text{Tyr} + \text{Thr} \rightarrow \text{Urea} + 11 \text{ CO}_2$	$1.2274e^{-5}$
e_{14}	$\text{Ser} + 4 \text{ Leu} + \text{Met} \rightarrow 2 \text{ Ala} + 23 \text{ CO}_2$	$2.8599e^{-4}$
e_{15}	$\text{Tyr} \rightarrow \text{Ala} + 6 \text{ CO}_2$	$6.1860e^{-5}$
e_{16}	$\text{Ile} \rightarrow \text{Ala} + 3 \text{ CO}_2$	$1.3218e^{-4}$
e_{17}	$\text{Gln} \rightarrow \text{Urea} + 4 \text{ CO}_2$	$1.4200e^{-2}$
e_{18}	$\text{Phe} \rightarrow \text{Ala} + 6 \text{ CO}_2$	$1.2058e^{-4}$
e_{19}	$\text{Ser} + 2 \text{ Phe} + \text{Met} \rightarrow 2 \text{ Urea} + 23 \text{ CO}_2$	$4.9886e^{-6}$

els from detailed metabolic networks is presented, and discussed based on an application example related to batch cultures of CHO-320 cells. In particular, the relatively high complexity of the metabolic network makes impossible the computation of the complete set of elementary flux modes due to combinatorial explosion. An alternative procedure is therefore applied, where an admissible flux distribution is decomposed into a minimal set of elementary flux modes. The minimal set can be computed directly, without enumerating the full collection of EFMs. Model reduction based on this minimal decomposition provides sets of macroscopic bioreactions, as well as estimates of the maximum reaction rates. Dynamic models with suitable properties can therefore be obtained through the introduction of classical Monod factors. Piecewise models for the different cell life phases can also be easily constructed, using linear weighting functions (to switch from one phase to the other).

A global model can be obtained by a minor adjustment in the consumption rate of glucose. If glucose is considered to be consumed a little slower, the reaction rates of the macro reactions depending on glucose concentration become zero once the passage from the exponential growth phase model to the transition phase model is completed. The consideration of a smaller slope of glucose consumption allows the obtaining of a global model capable of reproducing the experimental data for the entire cell culture.

7 Acknowledgment

This article presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Program, initiated by the Belgian State, Science Policy Office. R. Jungers is a F.R.S.-FNRS fellow. The scientific responsibility rests with its authors.

References

- [1] A. Provost, G. Bastin, Y.-J. S., June 2007. From metabolic networks to minimal dynamic bioreaction models. In: 10th International IFAC Symposium on Computer Applications in Biotechnology. IFAC.
- [2] Ahn, W. S., Antoniewicz, M. R., 2011. Review - towards dynamic metabolic flux analysis in cho cell cultures. *Biotechnology Journal* 6, 1–14.
- [3] Arden, N., Betenbaugh, M. J., 2004. Life and death in mammalian cell culture: strategies for apoptosis inhibition. *Trends in Biotechnology* 22, 174–180.
- [4] Bastin, G., Dochain, D., 1990. On-line Estimation and Adaptive Control of Bioreactors.
- [5] Bernard, O., Bastin, G., 2005. On the estimation of the pseudo-stoichiometric matrix for macroscopic mass balance modelling of biotechnological processes. *Mathematical Biosciences* 193, 51–77.
- [6] Chen, L., Bastin, G., 1996. Structural identifiability of the yield coefficients in bioprocess models when the reaction rates are unknown. *Mathematical Biosciences* 132, 35–67.
- [7] Haag, J. E., Wouwer, A. V., Bogaerts, P., 2005. Dynamic modeling of complex biological systems: a link between metabolic and macroscopic description. *Mathematical Biosciences* 193, 25–49.
- [8] Hulhoven, X., Vande Wouwer, A., Bogaerts, P., 2005. On a systematic procedure for the predetermination of macroscopic reaction schemes. *Bioprocess Biosystem Engineering* 27 (5), 283–291.
- [9] Jungers, R., Zamorano, F., Blondel, V., Wouwer, A. V., Bastin, G., 2009. A fast algorithm for computing a minimal decomposition of a metabolic flux vector in terms of elementary flux vectors. *Vienna Conference on Mathematical Modelling - MATHMOD 2009*.
- [10] Jungers, R. M., Zamorano, F., Blondel, V., Wouwer, A. V., Bastin, G., 2011. Fast computation of minimal elementary decompositions of metabolic flux vectors. *Automatica* 47, 1255–1259.
- [11] Mocquet, C., Bernard, O., Sciandra, A., 2010. Cell cycle modelling of microalgae grown under a light-dark signal. *CAB 2010*.

- [12] Murray-Smith, R., Johansen, T. A., 1997. Multiple model approaches to modelling and control.
- [13] Provost, A., 2006. Metabolic design of dynamic bioreaction models. Ph.D. thesis, Université Catholique de Louvain.
- [14] Provost, A., Bastin, G., 2004. Dynamic metabolic modelling under the balanced growth condition. *Journal of Process Control* 14, 717–728.
- [15] Smets, I., 2002. Analysis and synthesis of mathematical algorithms for optimization and control of complex biochemical conversion processes. Ph.D. thesis, Department of Chemical Engineering.
- [16] Zamorano, F., 2010. Metabolic flux analysis of CHO cell cultures. Ph.D. thesis, Université de Mons.
- [17] Zamorano, F., Wouwer, A. V., Bastin, G., 2010. A detailed metabolic flux analysis of an underdetermined network of CHO cells. *Journal of Biotechnology* 150, 497–508.