

## Article

# Hybrid Dynamic Models of Bioprocesses Based on Elementary Flux Modes and Multilayer Perceptrons

Maxime Maton <sup>1</sup>, Philippe Bogaerts <sup>2</sup> and Alain Vande Wouwer <sup>1,\*</sup><sup>1</sup> Systems, Estimation, Control and Optimization (SECO), Université de Mons, 7000 Mons, Belgium<sup>2</sup> 3BIO-BioControl, Université Libre de Bruxelles, 1050 Brussels, Belgium

\* Correspondence: alain.vandewouwer@umons.ac.be

**Abstract:** The derivation of minimal bioreaction models is of primary importance to develop monitoring and control strategies of cell/microorganism culture production. These minimal bioreaction models can be obtained based on the selection of a basis of elementary flux modes (EFMs) using an algorithm starting from a relatively large set of EFMs and progressively reducing their numbers based on geometric and least-squares residual criteria. The reaction rates associated with the selected EFMs usually have complex features resulting from the combination of different activation, inhibition and saturation effects from several culture species. Multilayer perceptrons (MLPs) are used in order to undertake the representation of these rates, resulting in a hybrid dynamic model combining the mass-balance equations provided by the EFMs to the rate equations described by the MLPs. To further reduce the number of kinetic parameters of the model, pruning algorithms for the MLPs are also considered. The whole procedure ends up with reduced-order macroscopic models that show promising prediction results, as illustrated with data of perfusion cultures of hybridoma cell line HB-58.

**Keywords:** hybrid modeling; model reduction; dynamic models; metabolic network; elementary flux modes; identification; neural networks; multilayer perceptron; pruning; biotechnology; reaction systems



**Citation:** Maton, M.; Bogaerts, P.; Vande Wouwer, A. Hybrid Dynamic Models of Bioprocesses Based on Elementary Flux Modes and Multilayer Perceptrons. *Processes* **2022**, *10*, 2084.  
<https://doi.org/10.3390/pr10102084>

Academic Editor: Alessandro D' Adamo

Received: 2 September 2022

Accepted: 4 October 2022

Published: 14 October 2022

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



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Mammalian cell cultures are now widely established to produce therapeutic products of interest such as monoclonal antibodies, viral vaccines and proteins used in the treatment of genetic diseases. Following the current trends of Process Analytical Technologies (PAT) and Industry 4.0, the development of dynamic mathematical models and predictors or estimators is receiving ever-increasing attention in view of establishing digital twins of the bioprocesses.

The models can usually be classified into structured and unstructured models, where the first group considers the cellular metabolism, while the second neglects the intracellular activity and focuses on the evolution of extracellular species. Over the years, connections have been established between the two approaches, and intra- and extracellular information can be combined to develop reduced-order macroscopic models [1–4]. A central concept in the development of minimal bioreaction models is the notion of elementary flux modes (EFMs) [5], which can be seen as the simplest metabolic pathways linking extracellular substrates to products. The EFMs form a convex basis of the flux space, whose dimension unfortunately grows very quickly with the size of the metabolic network. Tackling this combinatorial explosion has been a major research concern in the last decades.

To alleviate this issue, a first approach is to consider metabolic networks of modest sizes, which can be obtained on the basis of detailed networks through metabolic flux analysis [6], for instance, by discarding insignificant fluxes [7]. In this way, elementary vectors can all be computed and enumerated using specific software tools such as METATOOL [8] and EFMTOOL [9].

Alternatively, when the number of EFMs is prohibitively large [10], methods to select EFM subsets are required. To this extent, Figueiredo et al. [11] identified the shortest elementary modes, Kaleta and his group [12] proceeded with subsystem analysis, Jungers [13] decomposed the flux distribution into a minimal number of elementary vectors, Machado and Tabe [14,15] developed random sampling, Soons [16,17] used ranking or controlled random search to select a subset of modes based on an optimization criterion and Oddsdotir [18,19] employed column generation techniques to identify a small number of modes. Most of the previous methods account for constraints related to cell-specific uptake or secretion rates to further constrict the flux space, as studied in [20,21].

Another strategy consists in reducing the number of elementary vectors from a larger initial pool so as to keep only a small set of representative modes. In connection with this, Hebing et al. [22] used an EFM reduction procedure based on a geometrical collinearity criterion. More recently, several procedures of EFM reduction have been developed by our research group, i.e., Abbate [23] selected the best EFM candidates based on the formulation of a linear optimization problem and Maton [24] developed a reduction methodology based on a combination of several criteria based on collinearity and a series of constrained least-square problems.

What makes the use of elementary flux modes attractive in bioprocess modeling is the possibility to derive minimal bioreaction schemes. Henceforth, it remains to model the kinetics using specific functions (such as Monod or Haldane laws or more complex nonlinear functions of the extracellular species) and parameter identification to establish a functional model from sets of experimental data. In the last decades, many studies addressed the issue of kinetic identification. However, the determination of kinetic structures is often based on arbitrary choices. Indeed, specific kinetic phenomena such as activation, saturation and inhibition can sometimes be described by different mathematical expressions with similar evolutions with respect to the culture species. Attempts have been made and achieved to propose general kinetic formalisms, as in [25,26], where power-law equations were used to represent overall reaction rates of complex biological systems [27–32]. Nevertheless, although these methods are effective and have the merit of being systematic, they do not capture a double component effect. More recently, generalized Monod equations were proposed in [20,33–35]; a systematic procedure to select the most likely kinetic structures using decision graphs, nested models and likelihood ratio tests was developed in [36]; and another systematic procedure to model the kinetics using generalized-kinetic functions and a three-step parameter identification procedure was detailed in [37,38].

All the previous approaches are based on kinetic functions incorporating factors representing various phenomena of activation, inhibition and saturation, and providing some biological interpretation. When the reaction rates are difficult to describe as they include influences from several culture species, an alternative approach is to resort to purely data-driven techniques such as neural networks [39–41]. Although these models present a high flexibility to capture nonlinear kinetic phenomena, they lack biological interpretation.

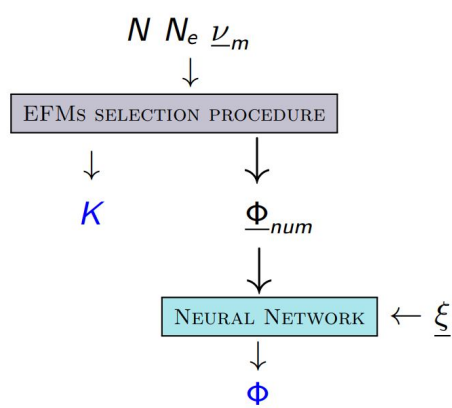
This latter situation corresponds to the starting point of this study. Indeed, the application of our EFM reduction algorithm [24,42] yields, on the one hand, a candidate stoichiometric matrix and, on the other hand, the time-evolution of the metabolic fluxes corresponding to the selected EFMs. These fluxes may have a complex time evolution resulting from the combination of several phenomena triggered by different culture species. Therefore, the objective of the present study is to explore the potential of recurrent neural networks, and particularly multilayer perceptrons (MLPs), to describe these fluxes. The main reason for this choice is their simplicity and the existence of efficient toolboxes such as the Deep Learning toolbox in Matlab. A disadvantage of MLPs, however, is that they have a fully connected structure with a quick increase in the number of parameters with the number of neurons in the hidden layer. In this study, it is desired to keep the number of parameters at the minimum and the use of a pruning algorithm is also explored. To support our findings, experimental data of perfusion cultures of hybridoma cell line HB-58 are used to derive a dynamic hybrid bioreaction model. This latter model is corroborated with the

mechanistic model of [43], in particular, in order to interpret how overflow metabolism is represented by the neural structure.

The paper is organized as follows. The next section presents the concept of hybrid modeling using elementary flux vectors for the derivation of minimal bioreaction models and neural networks for the description of the kinetics. In Section 3, information is given about cultures of hybridoma cell line HB-58, which are used as a representative case study. Numerical results are discussed and interpreted in Section 4. Finally, conclusions are drawn in Section 5.

## 2. Hybrid Modeling

The hybrid modeling methodology is sketched in Figure 1. Starting from the definition of a metabolic network with stoichiometric matrix  $N$ , the measurement configuration  $N_e$  and experimental uptake and excretion rates  $\underline{v}_m$ , this method first exploits mechanistic information based on the concept of elementary flux modes and an EFM reduction procedure to infer a minimal bioreaction model and its corresponding stoichiometric matrix  $K$ . Second, the procedure exploits the time evolution of the reaction rates  $\underline{\Phi}_{num}$ , which are provided by the EFM reduction algorithm, and applies a data-driven modeling of these rates using MLPs with the concentration of extracellular measurements  $\underline{\zeta}$  as inputs. The model components  $K$  and  $\underline{\Phi}$  can then be used in a predictive model. The full methodology is described in detail in the sequel.



**Figure 1.** Hybrid modeling using an elementary flux mode reduction procedure and neural networks.  $N$  is the stoichiometric matrix,  $N_e$  is the stoichiometric matrix of extracellular measurements,  $\underline{v}_m$  is the vector of experimental uptake and excretion rates,  $K$  describes the stoichiometry of the bioreactions,  $\underline{\phi}_{num}$  is the vector of reaction rates obtained numerically from the reduction procedure,  $\underline{\phi}$  is the vector of reaction rates modeled using NN and  $\underline{\zeta}$  is the vector of concentrations of culture species.

### 2.1. EFM Selection

#### 2.1.1. Metabolic Network Analysis

Cellular metabolism is defined as a set of chemical reactions, possibly catalyzed by enzymes, taking place within the cell and forming metabolic pathways. These intracellular reactions may be translated into a matrix representation defining a metabolic network as a  $m \times n$  stoichiometric matrix, denoted  $N$ . In this formalism,  $m$  is the number of internal metabolites and  $n$  represents the number of reactions. Considering the pseudo-steady state assumption, the following homogeneous system of linear equations is obtained:

$$N \underline{v} = 0 \quad (1)$$

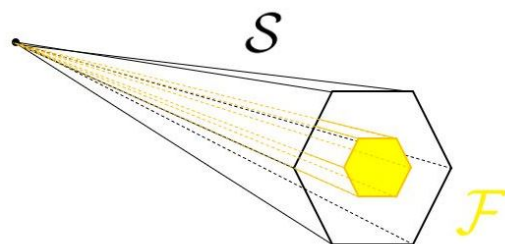
where  $\underline{v} \in \mathbb{R}^n$  gathers the fluxes of the network. Moreover, to express that the reactions have a net direction, the following constraint may be stated:

$$\underline{v} \geq 0 \quad (2)$$

Henceforth, the set of possible flux distributions  $\underline{v}$  defines a pointed polyhedral cone  $\mathcal{S}$  in the positive orthant. The edges of  $\mathcal{S}$  represent elementary flux vectors—the simplest metabolic pathways connecting extracellular substrates to final products without accumulation of metabolites.

As depicted in Figure 2, it is possible to further reduce the solution space by adding a set of linear constraints making use of experimental uptake and excretion rates  $\underline{v}_m$ :

$$\begin{pmatrix} N \\ N_e \end{pmatrix} \underline{v} = \begin{pmatrix} 0 \\ \underline{v}_m \end{pmatrix} \quad (3)$$



**Figure 2.** Cones of flux distributions.

Subject to the constraint in Equation (2), the set of solutions is included in the polytope  $\mathcal{F}$  ( $\mathcal{F} \subset \mathcal{S}$ ) and only specific combinations of modes provide a solution. In Equation (3),  $N_e$  is the stoichiometric matrix of extracellular measurements, a  $m_e \times n$  matrix where  $m_e$  stands for the number of measured extracellular species. The vector  $\underline{v}_m$  is inferred from the measurement of the time evolution of the extracellular concentrations, for instance, using regression based on the smoothed derivative of the concentration signals.

### 2.1.2. EFM Reduction Procedure

The EFM reduction method has been originally developed in [24,42], and only a brief overview is provided in this section. The 4-step method, illustrated in Figure 3, is divided into (i) the initial generation of the modes, (ii) the biological interpretation of the reaction scheme, (iii) a preliminary reduction up to  $\Omega$  modes and (iv) the selection of a minimal set of  $\Lambda$  EFMs.

- *Generation of the initial EFM set:* The first step concerns the generation of elementary flux vectors. If the size of the metabolic network is not too big, the whole set of modes can be computed and enumerated using software tools such as EFMTOOL. Otherwise, if the number of EFMs becomes prohibitive, alternative methods to identify only subsets are required. A fast generation algorithm [13] can be used that requires also the knowledge of experimental measurements of uptake and excretion rates. Nevertheless, regardless the EFMs generation method, a matrix of elementary flux modes  $E$  can be obtained.
- *Biological interpretation:* The second step consists in ensuring a biological interpretation of the matrix  $K$ , which is defined by

$$K = N_e E \quad (4)$$

and is a  $m_e \times n_{EFM}$  matrix whose columns provide the stoichiometry of each bioreaction. All the modes conducting to a macroreaction with no reactant or no product should be removed, yielding a reduced matrix of modes, denoted  $E^*$ . Note that reactions corresponding to maintenance are kept.

- *Main EFM reduction:* This step allows reducing the number of modes up to a target  $\Omega$ . This value is generally set close to, and sometimes slightly greater than,  $m_e$  to avoid computational issues during the final step of the procedure. First, collinearity between vectors is evaluated and the collinear modes are discarded. Second, an optimization-

based reduction is achieved where a randomly selected vector is removed if the following inequality is satisfied:

$$|\Xi^* - \Xi| < tol \quad (5)$$

where  $\Xi^*$  is the performance index for the candidate elimination,  $\Xi$  is the prior value of the indicator and  $tol$  is a tolerance value defined by the user. The selected performance index is a weighted constrained least-squares problem:

$$\Xi = \sum_{k=1}^M \left( K_e \underline{\phi}_k - \underline{v}_{m,k} \right) W^{-1} \left( K_e \underline{\phi}_k - \underline{v}_{m,k} \right)^T$$

$$\min_{\underline{\phi}_k} \Xi \quad s.t. \quad \underline{\phi}_k \geq 0 \quad (6)$$

where  $\underline{v}_{m,k}$  is the vector of uptake and excretion rates at every time step  $k$ ,  $\underline{\phi}_k$  corresponds to a time-varying decomposition of the flux  $\underline{v}_{m,k}$  into a reduced set of vectors stored in  $K_e$ , and  $W$  is a weighting diagonal matrix whose diagonal elements are  $\max_k v_{m,k,i}^2$  where  $v_{m,k,i}$  is the  $i^{th}$  element of the vector  $\underline{v}_{m,k}$  ( $i \in [1, n_{EFM}]$ ). This criterion represents how well the positivity constraints can be satisfied while reproducing the evolution of the extracellular fluxes. Note that  $K_e = N_e E_e$ , where  $E_e$  denotes a reduced matrix of EFMs.

- *Selection of a minimal bioreaction model:* For this final step, an even smaller number of EFMs is selected among the previous set of  $\Omega$  modes ( $\Lambda < \Omega$ ). This target  $\Lambda$  is chosen below the number of extracellular measured species  $m_e$  in order to derive macroscopic models with less reactions than components. This step is no longer based on random successive eliminations of modes but is a selection step of the best combination of  $\Lambda$  EFMs among the previous  $\Omega$  modes. Hence, the performance index  $\Xi$  of all possible EFM combinations is computed and the final set of modes is the one with the smallest value of the indicator, which represents the distance to the experimental data. Note that the number of possible combinations is given by

$$n_{comb} = \frac{\Omega!}{\Lambda! (\Omega - \Lambda)!} \quad (7)$$

As a consequence,  $\Omega$  has to be chosen small enough to avoid an unmanageable number of EFM combinations and, at the same time, large enough to have more flexibility in the selection, i.e., to have a sufficient pool of EFMs for the selection of the best EFM combination.

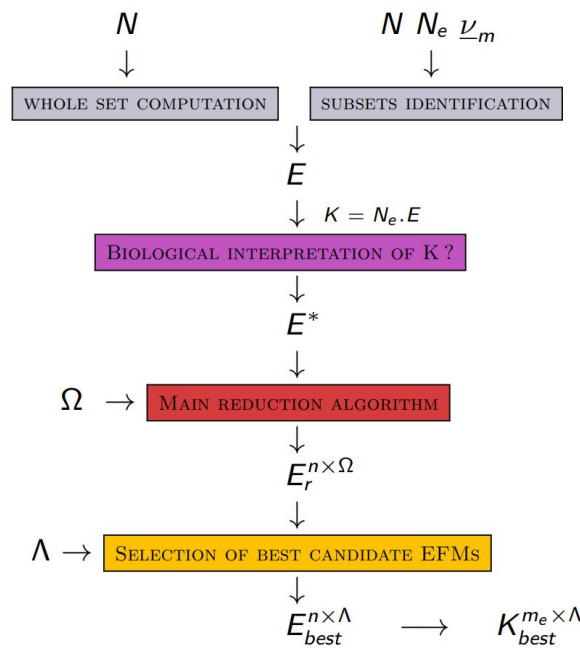
## 2.2. Dynamic Mass Balance Model

From the reduction procedure presented in the previous section, a stoichiometric matrix  $K$ , as well as the time evolution of the vector of reaction rates  $\Phi(t)$  have been obtained. The product of these two quantities defines the biological uptake and excretion rates of the culture species  $\underline{\zeta}$  in the following mass balance equation system:

$$\frac{d\underline{\zeta}}{dt} = K\Phi + D(\underline{\zeta}_{in} - \underline{\zeta}) \quad (8)$$

In this equation,  $D$  is a diagonal dilution matrix and  $\underline{\zeta}_{in}$  denotes the inflow concentrations.

To develop a predictive model, it is now necessary to represent the flux vector  $\Phi$  by a parametric model describing the influence of the extracellular culture species. In this study, MLPs will be used to describe the kinetic laws.

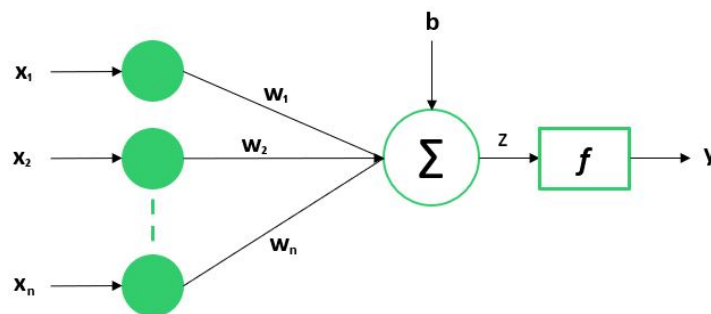


**Figure 3.** The modular EFM selection procedure.  $N$  is the stoichiometric matrix,  $N_e$  is the stoichiometric matrix of extracellular measurements,  $\underline{v}_m$  is the vector of experimental uptake and excretion rates,  $E$  is the matrix of EFMs and  $K$  describes the stoichiometry of the bioreactions.  $\Omega$  is a first target number for the reduced set of EFMs, close to the number of measured extracellular species  $m_e$ , and  $\Lambda$  is a second target number for the reduced set of EFMs, smaller than  $m_e$ .

NN Kinetic Modeling

A perceptron, as illustrated in Figure 4, involves a set of input signals  $x_i$  weighted with synaptic coefficients  $w_i$  plus a bias  $b$ :

$$z = b + \sum_{i=1}^n x_i \cdot w_i \tag{9}$$

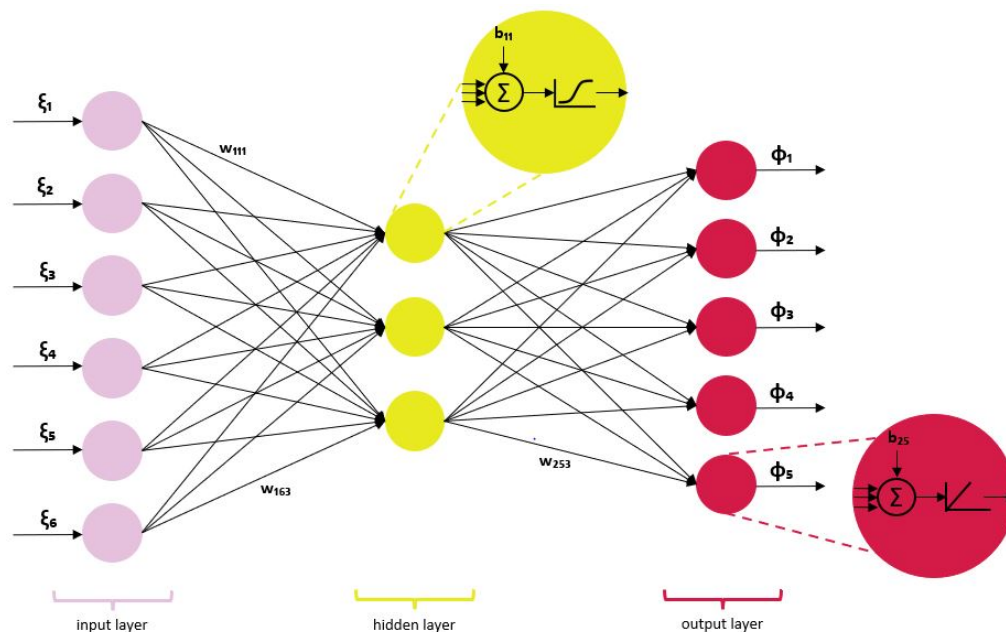


**Figure 4.** The structure of a perceptron:  $x_i$  represent the input signals,  $w_i$  denote synaptic coefficients,  $b$  is a bias,  $f$  is an activation function and  $y$  is the output signal.

This weighted sum enters an activation function  $f$  to predict the output signal  $y = f(z)$ . There exist many activation functions, with sigmoid functions being a popular choice. To represent complex nonlinear functions, the use of multilayer perceptrons (MLPs) is recommended, which are organized as follows:

- one input layer, which distributes the input values to the first hidden layer;
- one or several hidden layers of perceptrons;
- one output layer, which recovers the output of each perceptron of the last hidden layer.

Figure 5 shows a typical multilayer perceptron with one hidden-layer in the context of the present study, i.e., the input signals are the species concentrations  $\xi_i$  and the outputs are the elements of the flux vector  $\Phi$ . It is worth noting that the number of hidden layers constitutes the true computational engine of the MLP and the more hidden layers, the more powerful the artificial neural network but at the expense of a large number of parameters and the risk of overparametrization. Nevertheless, even one hidden-layer MLP can approximate the mapping of any continuous function [44].



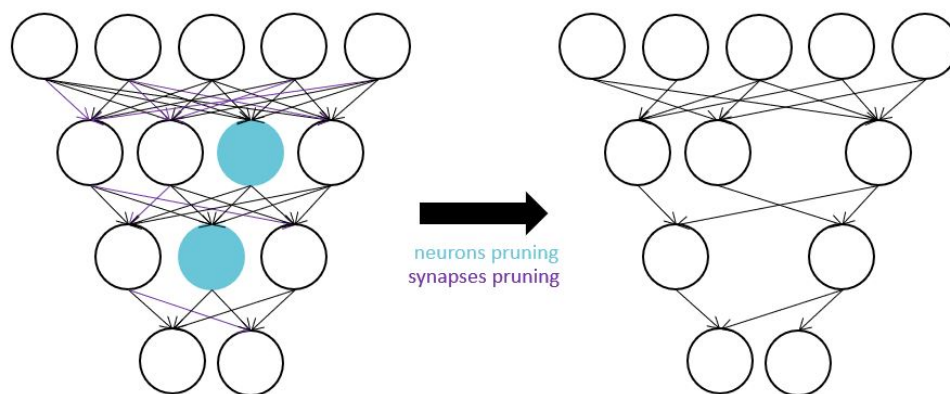
**Figure 5.** A multilayer perceptron with one hidden layer with 3 neurons; 6 input signals in the input layer, which are the concentrations of culture species; and 5 output signals in the output layer, which are the reaction rates of each bioreaction. The biases are denoted by  $b_i$  and activation functions in the hidden layer are sigmoids while activation functions in the output layer are identity functions.

The multilayer perceptron is known as a feedforward network as data flows in the forward direction from the input to the output layer. The neurons can be trained with a backpropagation learning algorithm computing the gradient of a loss function. Basically, the perceptron learning consists in adjusting the weights and the biases of the network in order to minimize the deviations between the outputs predicted by the network and the expected ones. For the training of MLPs, backpropagation is an efficient algorithm to estimate the gradients, which can be exploited in various methods, such as gradient descent or stochastic gradient descent, updating the parameters of the structure to minimize the loss function. More information on the learning algorithms can be found, for instance, in [45].

Before training the neural network, it is common practice to perform pre- and post-processing steps on the network inputs and outputs in order to make the training process more efficient. Depending on the magnitude of the input values, sigmoid activation functions may become saturated leading to small gradients and slow convergence of the training procedure. To tackle this issue, a normalization step is applied to input signals in the data set during a pre-processing step and the output signals can be reversely transformed back into the original range in a post-processing step.

MLPs are fully interconnected and the number of parameters rapidly increases with the number of hidden layers and neurons in these layers. Pruning algorithms discard redundant and unnecessary connections [46–50]. By removing synaptic connections with little influence, a sparsely connected network with multiple zero weights can be obtained. Methods to perform pruning of MLPs are proposed in [51]. Figure 6 illustrates the architecture of a network before and after different techniques of pruning. In this study, a simplified

version of the magnitude pruning algorithm proposed in [48] is developed. The idea is to assign a score to each parameter of the network (weights and biases) corresponding to its absolute value. Due to the pre- and post-processing, the score of each parameter is assumed to represent its relative importance to the accuracy of the trained network. Hence, a synaptic connection or a bias may be removed if its score is smaller than a threshold value defined by the user.



**Figure 6.** Architecture of an MLP before and after pruning—synapses and neurons pruning.

### 3. Case Study: Perfusion Cultures of Hybridoma Cell Line HB-58

Data used in this study come from the hybridoma cell line HB-58 enabling the production of IgG1 monoclonal antibodies, specific for mouse kappa light chain. Serum-free medium—prepared from DMEM/F12 (1:1) and completed with glutamine, 500  $\mu\text{mol}$  of ethanolamine, 10 mg of bovin insulin, cholesterol, Pluronic F-68, HEPES and other additives—was used. The cells were kept as suspension cultures in the medium in shake flasks at 37  $^{\circ}\text{C}$  in a 5%  $\text{CO}_2$  incubator. Experiments were conducted at the State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology (ECUST), in Shanghai [52].

Perfusion cultures were performed in a 2 L stirred bioreactor and were settled in a working volume of 1.8 L during the whole duration of the culture. Cultures were carried out in a controlled environment (36.8  $^{\circ}\text{C}$ , 40% DO, pH = 7.0  $\pm$  0.2). The perfusion phase started after 44–56 h of batch culture with a constant dilution rate of 0.0197  $\text{h}^{-1}$ . Cells were retained by a spin-filter (20  $\mu\text{m}$ ) and the stirring speed was fixed at 200 rpm. Data acquisition and process control were performed using the supervisory software MFCS/Win 3.0. Table 1 summarizes the operating conditions of the cultures.

**Table 1.** Summary of culture conditions for experiments in perfusion.

Experiment	$X_0$ ( $10^9$ cells/L)	$t_{fed}$ (h)	Feed Stream (mM)	
			$Glc_{in}$	$Gln_{in}$
1	0.19	54	11	5
2	0.23	56	15	11.5
3	0.36	48	28	4
4	0.36	44	28	9.5

#### 3.1. Metabolic Network

Depending on the level of description required, networks of different sizes might be considered. Henceforth, a metabolic network of hybridoma cells with 70 biochemical reactions and 44 internal metabolites is considered. The details of this metabolic network can be found in [53]; it includes the major reactions of central metabolism such as the pathways of glycolysis, the tricarboxylic acid cycle, amino acids metabolism, and the



synthesis of biomass and antibodies. The pentose phosphate pathway is not taken into account in the network definition because, in most tissues, 80 to 90% of glucose oxidation is achieved by glycolysis. The stoichiometric coefficients of the biomass and antibody synthesis are taken from literature [52]. Note that considering a metabolic reaction for the formation of biomass allows its prediction when deriving dynamical models.

### 3.2. Measurement Configuration

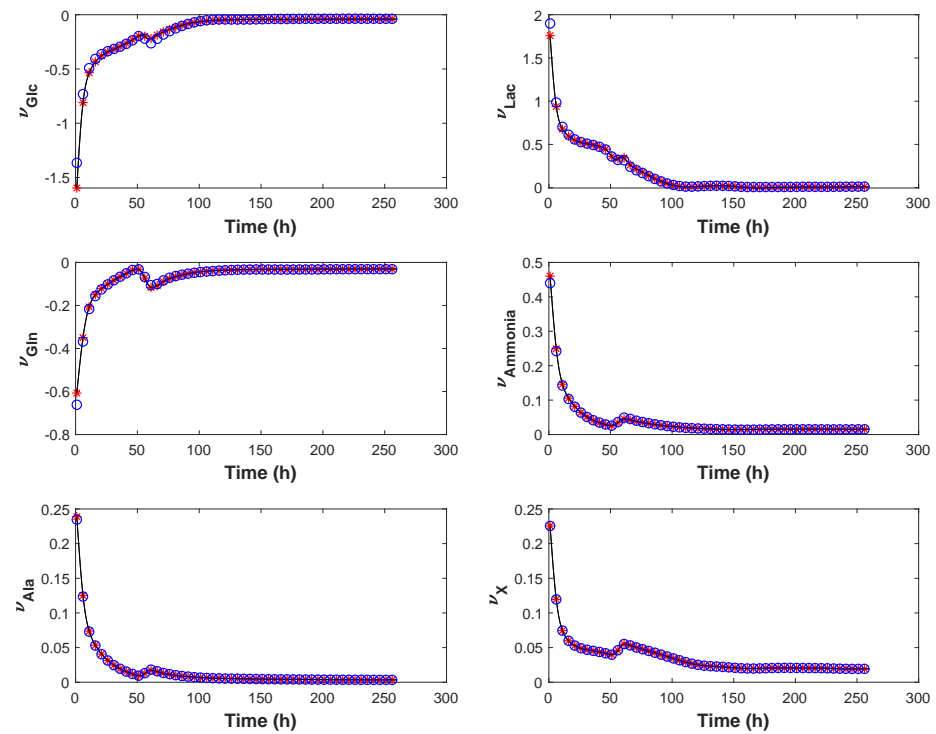
Only 6 extracellular measurements are accounted for—namely, glucose, lactate, glutamine, ammonia, alanine and biomass—gathered in vector  $\xi$ . This leads to the stoichiometric matrix of extracellular measured species  $N_e$ . Furthermore, the experimental uptake and excretion rates  $\underline{v}_m(t)$  may be computed using smoothing splines and differentiation methods.

## 4. Numerical Results

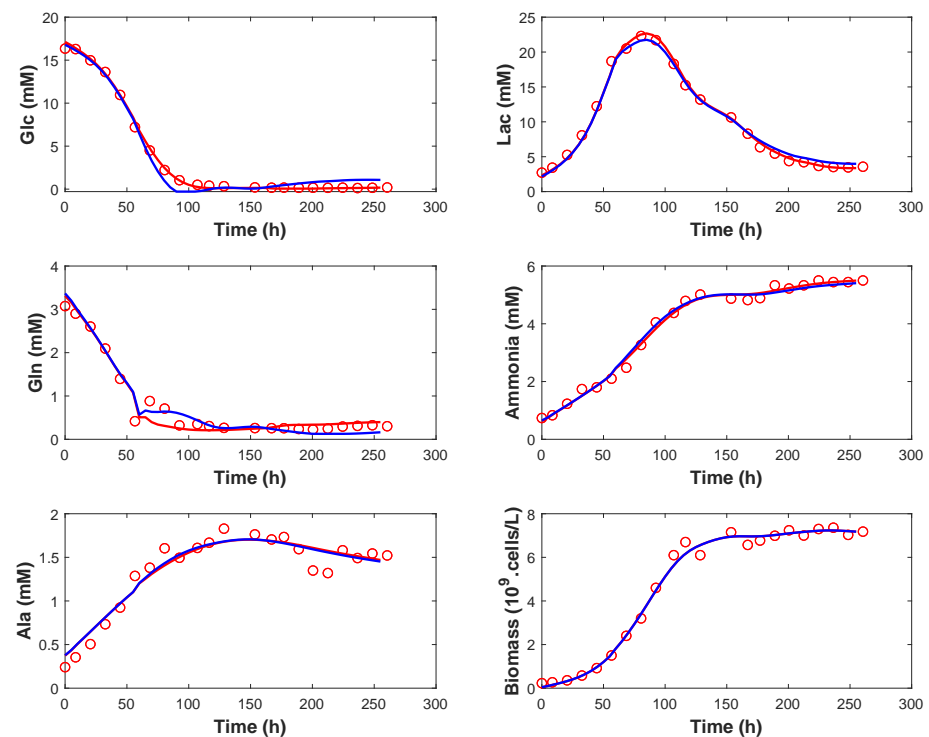
### 4.1. EFM Selection

The first step of the reduction procedure consists in computing an initial set of elementary flux vectors. As stated in Section 2.1.2, different approaches exist depending on the size of the metabolic network. In this case, using EFMTOOL as the EFM generator, an initial set of 22,563 modes is obtained. Alternative methods such as the one proposed in [13] identify only representative subsets of EFMs and allow getting a few hundred modes as a starting point. However, in order to prove the effectiveness of the reduction procedure, the whole set of elementary vectors is considered in this study since their number remains manageable. Then, for the purpose of ensuring a biological interpretation of the reaction scheme, any vector leading to a macroreaction with no biological meaning is discarded and the reduced matrix of modes  $E^*$  is made up of 21,874 vectors. The power of the methodology lies in the main reduction algorithm. With a collinearity indicator of 99% between the vectors, a drastic reduction to 214 modes is achieved. Next, an optimization-based reduction is performed with a target number  $\Omega$  equal to 10. The last step is the selection of a few modes from the previous set to derive models with fewer reactions than components. In this study, only five elementary flux modes are finally selected ( $\Lambda = 5$ ). Note that the target  $\Lambda$  could have been set to an even smaller value. However, because overflow metabolism is noticed in hybridoma cells, a further reduction may cause a loss of essential biological information.

A first direct validation of the reduced model can be achieved by comparing the experimental uptake and excretion rates  $\underline{v}_m(t)$  to the product  $K_e\Phi(t)$  computed in the optimization problem. This validation can be pursued by analyzing the prediction of the measured concentrations by integration of the computed fluxes and identification of the most likely initial conditions of the measured species. These results are shown in Figures 7 and 8, respectively. Almost no loss of information is noticed and the prediction of the concentrations is very satisfactory, highlighting the merits of the modular EFM selection procedure.



**Figure 7.** Time evolution of the experimental uptake and excretion rates (in  $\text{mM}\cdot\text{h}^{-1}$ ) in dataset # 2—numerical results (black curves);  $K_e\Phi$  for  $n_{EFM} = 10$  (red crosses);  $K_e\Phi$  for  $n_{EFM} = 5$  (blue bullets).

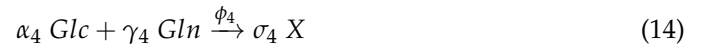


**Figure 8.** Time evolution of the measured concentrations in dataset # 2—experimental data points (red bullets); concentration profiles on the basis of  $n_{EFM} = \Omega = 10$  (red curves) and selecting  $n_{EFM} = \Lambda = 5$  (blue curves).

Besides the previous validations, it might be interesting to examine the final set of macroreactions obtained from the reduction algorithm:

$$K = \begin{bmatrix} 0 & -\alpha_2 & 0 & -\alpha_4 & -\alpha_5 \\ \beta_1 & 0 & 0 & 0 & \beta_5 \\ -\gamma_1 & -\gamma_2 & -\gamma_3 & -\gamma_4 & 0 \\ \delta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon_3 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 & 0 \end{bmatrix} \quad (10)$$

where  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $\delta_i$ ,  $\epsilon_i$  and  $\sigma_i$  are stoichiometric coefficients, listed in Table 2.  $K$  is a  $m_e \times \Lambda$  matrix with  $m_e = 6$  and  $\Lambda = 5$ . Equivalently, a macroreaction scheme can be drawn:



This reaction scheme makes sense and is validated by the study conducted in [43]. Indeed, glucose and glutamine are consumed with biomass growth and metabolites production defining the so-called respiratory metabolism. Moreover, the phenomenon of overflow metabolism is pointed out with lactate production from glucose excess and production of ammonia, lactate and alanine from glutamine excess. This point will be further discussed in Section 4.3.

**Table 2.** The stoichiometric coefficients of the reaction network.

Parameters	Value	Parameters	Value	Parameters	Value
$\alpha_2$	0.0150	$\beta_5$	0.1459	$\gamma_4$	0.0194
$\alpha_4$	0.0258	$\gamma_1$	0.2425	$\delta_1$	0.4851
$\alpha_5$	0.0729	$\gamma_2$	0.0092	$\epsilon_3$	0.2408
$\beta_1$	0.2425	$\gamma_3$	0.1204	$\sigma_4$	0.7925

#### 4.2. Kinetic Modeling

For the derivation of a dynamic model, it remains to describe the numerical signals  $\Phi$  coming from the reduction algorithm by neural networks. As illustrated in Figures 1 and 5, a multilayer perceptron will use as input signals the concentration of extracellular species  $\underline{\zeta}$  and as target values the reaction rates  $\Phi$  computed numerically in the reduction procedure. Therefore, there are six input signals, namely, the concentrations  $\zeta_i$  ( $i \in [1, 6]$ ) of the measured species and five output signals  $\Phi_i$  ( $i \in [1, 5]$ ), which are the specific rates of each reaction. Only one hidden layer is considered to limit the number of parameters (weights and biases), and the nonlinear activation functions in the hidden layer are hyperbolic tangent sigmoid functions while the activation functions in the output layer are identity functions. In summary, the following equations can be written:

$$z_i = b_{1i} + f\left(\sum_{j=1}^6 w_{1ij} \zeta_j\right), \quad i \in [1, Z], \quad (16)$$

$$\Phi_i = b_{2i} + \sum_{j=1}^Z w_{2ij} z_j, \quad i \in [1, 5], \quad (17)$$

with

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (18)$$

This function has the same shape as the classical sigmoid function but horizontal asymptotes are in  $-1$  and  $1$  so that the output values are squeezed into  $[-1, 1]$ . In the previous equations,  $Z$  is the number of neurons in the hidden layer. Hence, the number of parameters is  $5 + 12Z$ . Depending on the complexity of the problem, the number of neurons in the hidden layer can be increased at the expense of a significant increase in the number of parameters to identify. Table 3 summarizes the number of parameters of the network according to the number of neurons in the hidden layer. The rapid increase in the number of parameters with the number of neurons justifies the use of pruning algorithms.

**Table 3.** Number of network parameters for six input signals and five output signals.

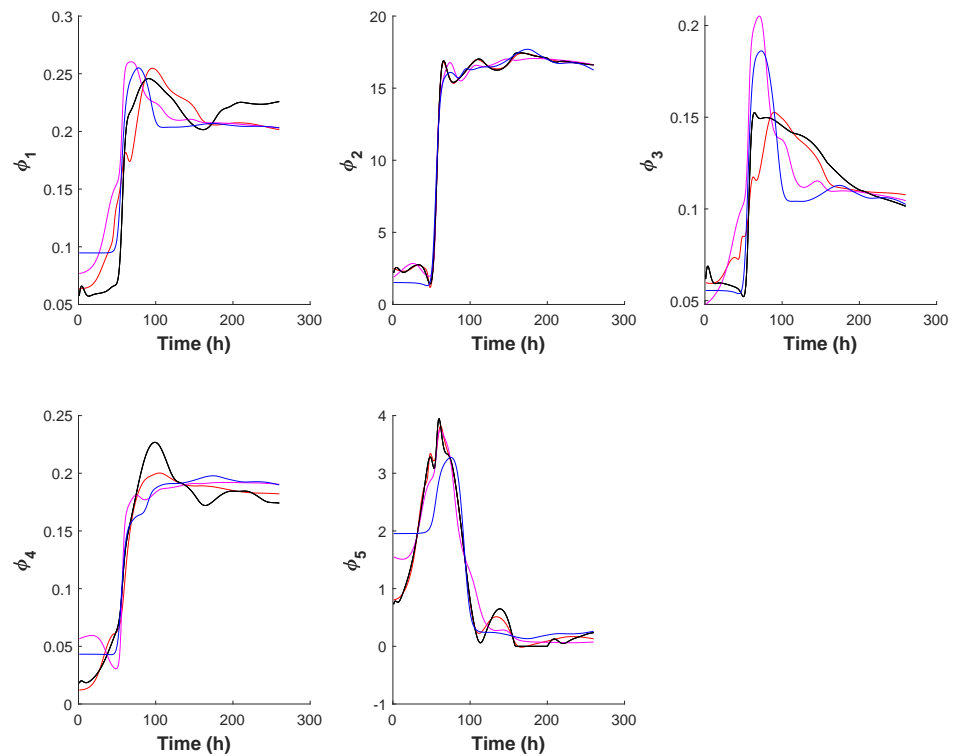
# Neurons	# Parameters
10	125
5	65
4	53
3	41
2	29
1	17

As a direct validation, Figure 9 shows the time evolution of the specific reaction rates  $\Phi$  produced by the reduction procedure (and used as target values) and the prediction with MLPs with different numbers of neurons in the hidden layer. As expected, the more neurons, the better the fitting to numerical results. Figure 10 depicts the corresponding concentration profiles. Although significant deviations appear in the reaction rates, they have a limited impact on the prediction of the time evolution of the concentrations. Hence, an MLP with only three neurons in the hidden layer is an acceptable compromise between the fitting to experimental data and the number of network parameters. Note that experimental data has been partitioned to avoid overfitting, i.e., 70% for training, 15% for validation and 15% for testing.

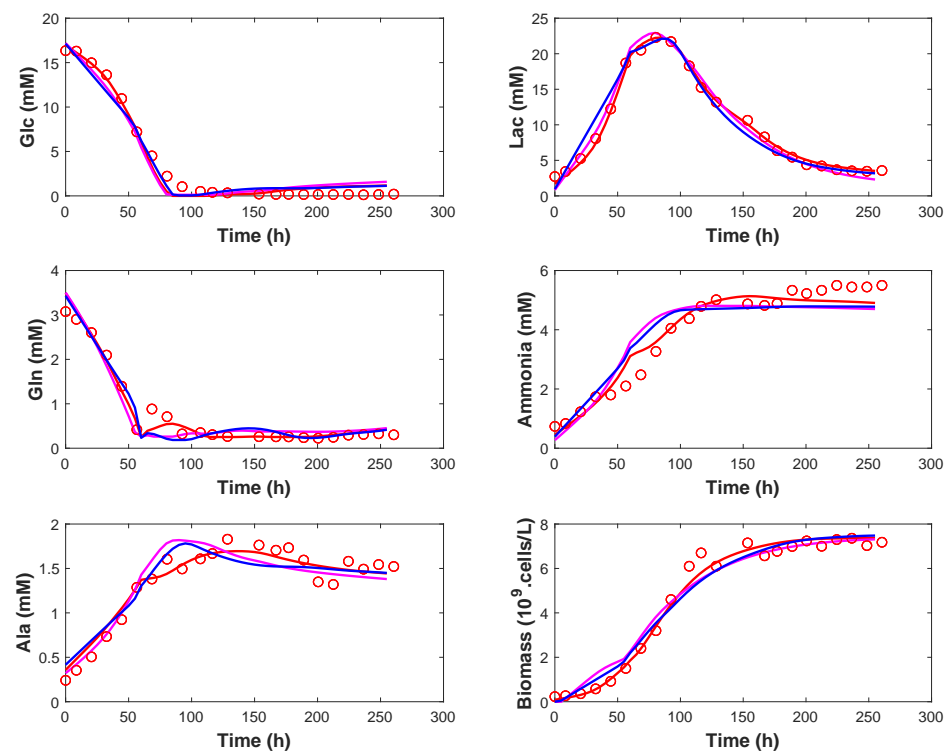
In order to improve the MLP prediction, instead of backpropagation, a global nonlinear identification of the weights and biases of the network can be performed using the Nelder–Mead simplex optimization algorithm in order to minimize a weighted nonlinear least-squares criterion:

$$\min_{\hat{\Theta}} J(\hat{\Theta}) = \sum_{j=1}^n \sum_{i=1}^{N_j} \left( \Phi_{ij}(\hat{\Theta}) - \Phi_{ij}^{num} \right) \Sigma^{-1} \left( \Phi_{ij}(\hat{\Theta}) - \Phi_{ij}^{num} \right)^T \quad (19)$$

where  $\hat{\Theta}$  is the vector of parameters to be re-identified containing the weights and biases of the neural network,  $\Phi_{ij}$  gives the network prediction at the  $i$ th time instant in the  $j$ th experiment,  $\Phi_{ij}^{num}$  is the vector of the corresponding numerical signals coming from the EFM reduction procedure and  $\Sigma$  is a normalization matrix in which diagonal elements are chosen as the squares of the maximum rate of each reaction.



**Figure 9.** Time evolution of the specific reaction rates (in  $\text{mM}\cdot\text{h}^{-1}$ ) in dataset # 2—target values (black curves); prediction with an MLP with ten neurons (red curves), five neurons (magenta curves) and three neurons (blue curves) in the hidden layer.



**Figure 10.** Prediction of the time evolution of the extracellular measured species in dataset # 2—numerical results from the optimization problem (black curves) using MLP with ten neurons (red curves), five neurons (magenta curves) and three neurons (blue curves) in the hidden layer.

To improve convergence, the optimization algorithm can be run several times using the parameters found in the previous round as initialization. The confidence intervals and variation coefficients of the estimated parameters can be obtained from the inverse Fisher information matrix [54]. The variation coefficients of the estimated parameters are given in Table 4.

Next, the magnitude pruning algorithm is exploited to reduce the number of parameters. The results of the pruning algorithm are shown in Table 5.  $W_1$  is a matrix collecting the synaptic connections between the input layer and the hidden layer,  $b_1$  represents the biases of the neurons in the hidden layer,  $W_2$  contains the weights between the hidden layer and the output layer and  $b_2$  are the biases of the neurons in the output layer. In this way, the number of parameters is reduced by 17% (41  $\rightarrow$  34 parameters).

**Table 4.** Coefficients of variation (in %) of the estimated weights and biases—colored boxes are removed synaptic connections or biases.

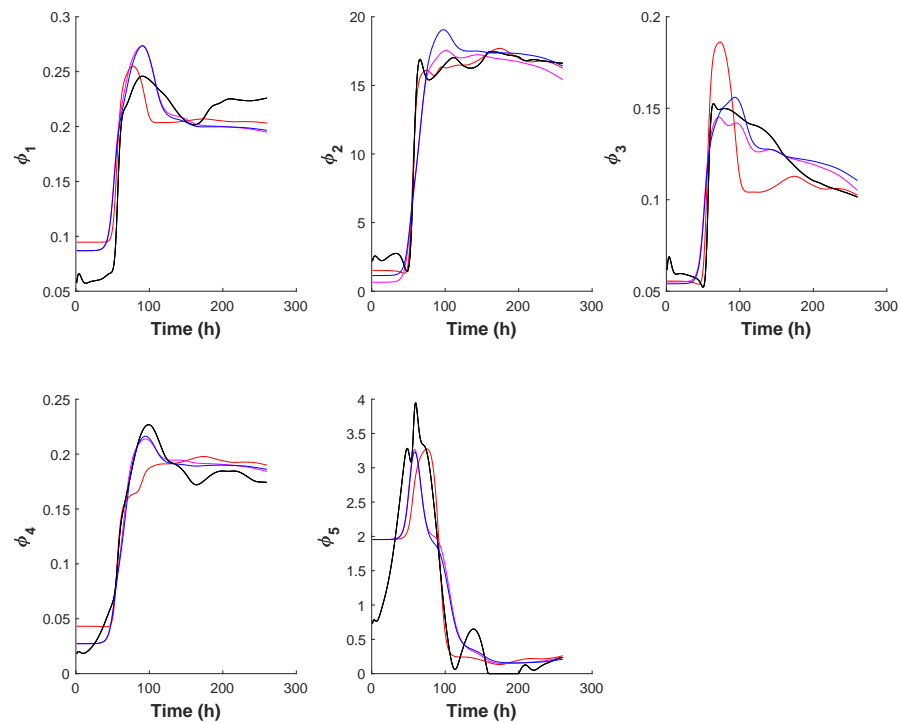
$W_1$						$b_1$	$W_2$			$b_2$
5.8	15.5	33.5	20	7.1	77	6.6	47.1	26.5	105.3	58.7
7.1	4.5	10.4	14.5	8.6	8.3	& 125.7	1	1	1.6	13.4
7.3	31.7	40.4	11.5	5.8	9.5	4.7	73.3	38.2	41.1	710.7
							32.8	50.9	82	169.6
							3.5	2.1	7.7	11.3

**Table 5.** Results of the magnitude pruning algorithm—colored boxes are removed synaptic connections or biases.

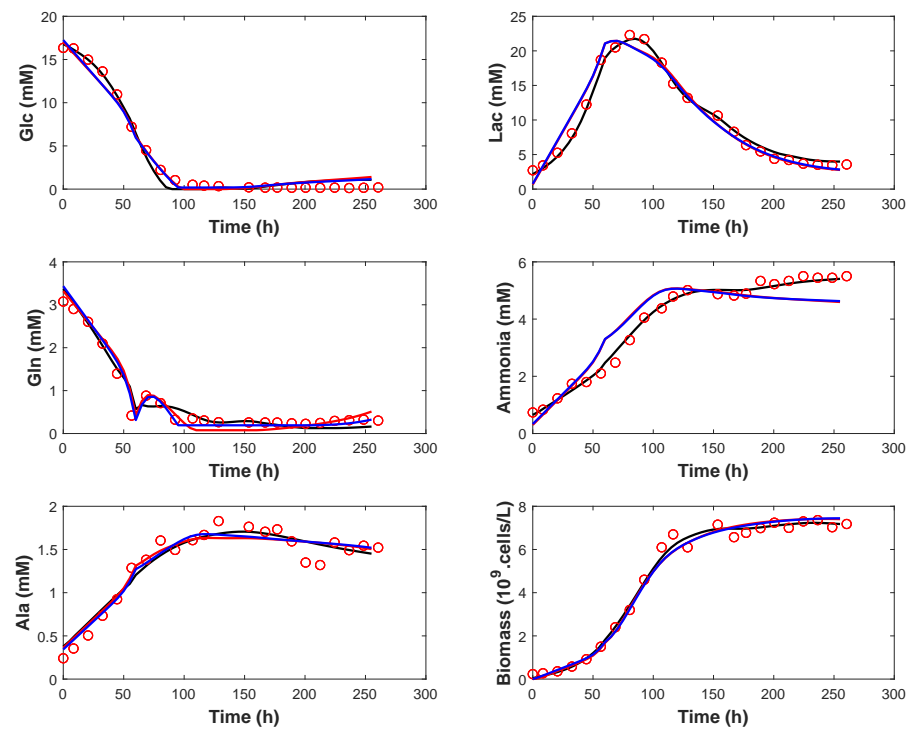
$W_1$						$b_1$	$W_2$			$b_2$

The time evolution of the reaction rates and the prediction of the measured concentrations using a pruned multilayer perceptron are illustrated in Figures 11 and 12, respectively. Figure 11 shows the added value of a global nonlinear re-identification of the parameters. Small deviations may be observed when using the pruned MLP for kinetic identification but the results remain very satisfactory, highlighting the merit of the simplified pruning method. Figure 12 validates the whole procedure providing good results for the prediction of the measured concentrations.

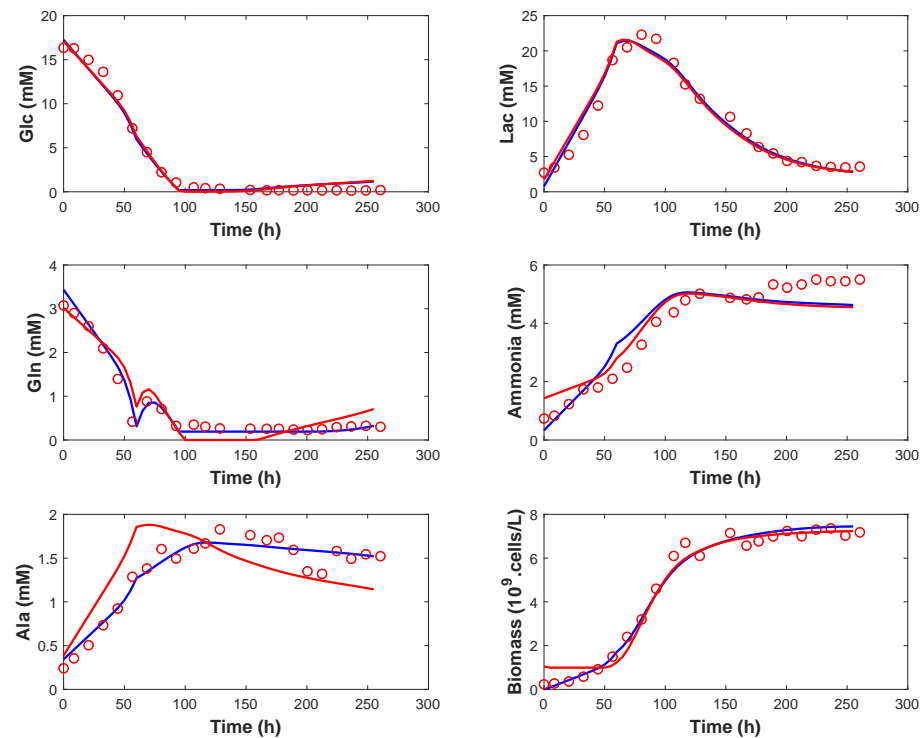
Furthermore, a pruning of the neural network can also be achieved on the basis of the coefficients of variation of the estimated parameters. Indeed, weights and biases identified with a coefficient of variation greater than 30% can be discarded without affecting the fitting of the experimental data. However, the elimination of the weights in  $W_2$  must be exercised with more care and only parameters identified with a coefficient of variation greater than  $\cong$  50% are neglected. Figure 13 shows the prediction of the time evolution of the extracellular measured species using this pruning strategy. It yields fewer parameters in the neural structure (41  $\rightarrow$  30 parameters) at the expense of larger deviations from experimental data. However, the results remain acceptable for deriving a macroscopic model for control and optimization purposes. Furthermore, the pruning strategy using the value of variation coefficients covers up most of the parameters defined as nonessential when using the magnitude pruning algorithm, validating the two pruning approaches.



**Figure 11.** Time evolution of the specific reaction rates (in  $\text{mM}\cdot\text{h}^{-1}$ ) in dataset # 2—target values (black curves); prediction with MLP with three neurons in the hidden layer (red curves), and with three neurons and nonlinear re-identification of weights and biases (magenta curves); pruned MLP with three neurons (blue curves).



**Figure 12.** Prediction of the time evolution of the extracellular measured species in dataset # 2—numerical results from the reduction procedure before kinetic identification (black curves); MLP with three neurons in the hidden layer and nonlinear re-identification of the parameters (red curves); pruned MLP with three neurons (blue curves).



**Figure 13.** Prediction of the time evolution of the extracellular measured species in dataset # 2—pruned MLP with three neurons using magnitude pruning algorithm (blue curves); pruned MLP with three neurons using a CV pruning strategy (red curves).

#### 4.3. Interpretation of Overflow Metabolism

Overflow metabolism is a phenomenon in which cells achieve incomplete oxidation of an abundantly supplied energy source, commonly glucose and glutamine, despite aerobic conditions. It results in the excretion of organic end-products that are often inhibitory. Overflow of glucose leads to the formation of lactate while overflow of glutamine results in the formation of ammonia. As mentioned in [55], glutamine is mainly excreted as alanine, proline and aspartate.

This phenomenon is essentially described by the reaction scheme deduced from the modular EFM selection procedure in Section 4.1. Indeed, in mammalian cell cultures, two main metabolic states can be pointed out: (i) the state of respiratory metabolism at low substrate uptake rates and (ii) the state of overflow metabolism at high substrate uptake rates. During respiratory metabolism, the consumption of glucose and glutamine leads to biomass growth and metabolite production such as lactate and ammonia. Respiratory growth is captured by reaction (14) for the substrate consumption and biomass growth, together with Equations (11), (13) and (15) for the metabolite production. Overflow of glucose leads to the production of lactate, as denoted by the sole reaction (15) (i.e., not coupled with reaction (14)), and excess of glutamine ends in the production of ammonia and lactate but also of other amino acids such as alanine among others as described by reactions (11) and (13). Lastly, the reaction (12) translates the consumption of glucose and glutamine for cell maintenance. The latter are two energy sources for the production of ATP and reduced pyridine nucleotides, essential for cell life. In that respect, all the chemical reactions deduced from the reduced matrix of EFMs have a biological meaning, as expected, consolidating the usefulness of the proposed reduction procedure.

The overflow metabolism phenomenon can be observed in the values of  $\Phi_i$  for  $i \in [1, 5]$  in Figure 11. Indeed,  $\Phi_5$  exhibits large values until about 100 h, before a significant decrease. This highlights a potential overflow metabolism on glucose during that first period of time, which vanishes at 100 h when glucose is almost completely depleted. This can be seen in



the glucose concentration–time profile in Figure 12. A similar decrease does not appear at the levels of  $\Phi_1$  and  $\Phi_3$ . This can be interpreted as an overflow metabolism on glutamine that takes place during the whole duration of the culture as the glutamine concentration remains above a value that must be greater than the critical level (see Figure 12).

## 5. Conclusions

This study establishes a complete procedure to derive a small macroscopic bioreaction model from metabolic networks using the concept of elementary flux vectors to infer a reaction scheme and neural networks to describe the reaction rates in terms of the culture species.

When the size of the metabolic network increases, depending on the level of description required, the number of elementary modes explodes and it is important to use systematic procedures to reduce and select the most informative modes in view of establishing the corresponding minimal bioreaction scheme. For this purpose, the modular reduction procedure includes several steps: (i) the initial generation of the modes (the whole set or only a representative subset), (ii) the biological interpretation of the reaction scheme, (iii) a reduction up to  $\Omega$  modes using collinearity between vectors, followed by a random elimination and (iv) the selection of a minimal set of  $\Lambda$  EFMs ( $\Lambda < m_e$ ) leading to models with fewer reactions than components.

Once such a minimal model has been derived, artificial neural networks can be exploited to model the specific rates of each chemical reaction. In this study, a multilayer perceptron with one hidden layer and a small number of neurons is successfully applied to this modeling task. To further reduce the number of parameters, two simple pruning algorithms are also applied with success. Pruning algorithms appear particularly useful because of the fully connected architecture of MLPs. Further investigation of a suitable compromise between the complexity of the neural network architecture and the use of more efficient pruning algorithms could be interesting research directions.

The hybrid physical–neural models show good predictive capability in a case study related to perfusion cultures of hybridoma cells. The hybrid model is biologically interpretable in terms of reaction schemes and the reproduction of important mechanisms such as overflow metabolism.

The main advantage of the proposed procedure is the rapid development of the hybrid dynamic models (once the experimental data have been collected) whose stoichiometry and kinetics can be obtained in a systematic way. The resulting dynamic models could be exploited for process monitoring (i.e., the development of software sensors) and model-based control, such as model predictive control. Further case studies are required to explore and consolidate these perspectives.

**Author Contributions:** Conceptualization, M.M., P.B. and A.V.W.; methodology, M.M., P.B. and A.V.W.; software, M.M.; formal analysis, M.M.; writing—original draft preparation, M.M. and A.V.W.; writing—review and editing, M.M., P.B. and A.V.W.; supervision, P.B. and A.V.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The experimental data are not available from the authors. Further information about the experimental studies can be found in [52].

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

## References

1. Hodgson, B.; Taylor, C.; Ushio, J.; Leigh, J. Intelligent modelling of bioprocesses: A comparison of structured and unstructured approaches. *Bioprocess Biosyst. Eng.* **2005**, *26*, 353–359.
2. Haag, J.; Vande Wouwer, A.; Bogaerts, P. Dynamic modeling of complex biological systems: A link between metabolic and macroscopic description. *Math. Biosci.* **2005**, *193*, 25–49.

3. Haag, J.; Vande Wouwer, A.; Bogaerts, P. Systematic procedure for the reduction of complex biological reaction pathways and the generation of macroscopic equivalents. *Chem. Eng. Sci.* **2005**, *60*, 459–465.
4. Baroukh, C.; Bernard, O. Metabolic modeling of *C. sorokiniana* diauxic heterotrophic growth. *IFAC-PapersOnLine* **2016**, *49*, 330–335.
5. Schuster, S.; Hilgetag, C. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* **1994**, *2*, 165–182.
6. Gao, J.; Gorenflo, V.; Scharer, J.; Budman, H. Dynamic metabolic modeling for a mAb bioprocess. *Biotechnol. Prog.* **2007**, *23*, 168–181.
7. Naderi, S.; Meshram, M.; Wei, C.; McConkey, B.; Ingalls, B.; Budman, H.; Scharer, J. Metabolic flux and nutrient uptake modeling of normal and apoptotic CHO cells. *IFAC Proc. Vol.* **2010**, *43*, 395–400.
8. von Kamp, A.; Schuster, S. Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics* **2006**, *22*, 1930–1931.
9. Terzer, M.; Stelling, J. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* **2008**, *24*, 2229–2235.
10. Klamt, S.; Stelling, J. Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep.* **2002**, *29*, 233–236.
11. Figueiredo, L.; Podhorski, A.; Rubio, A.; Kaleta, C.; Beasley, J.; Schuster, S.; Planes, F. Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics* **2009**, *25*, 3158–3165.
12. Kaleta, C.; de Figueiredo, L.; Schuster, S. Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Res.* **2009**, *19*, 1872–1883.
13. Jungers, R.; Zamorano, F.; Blondel, V.; Vande Wouwer, A.; Bastin, G. Fast computation of minimal elementary decompositions of metabolic vectors. *Automatica* **2011**, *47*, 1255–1259.
14. Machado, D.; Soons, Z.; Patil, K.R.; Ferreira, E.C.; Rocha, I. Random sampling of elementary flux modes in large-scale metabolic networks. *Bioinformatics* **2012**, *28*, i515–i521.
15. Tabe-Bordbar, S.; Marashi, S. Finding elementary flux modes in metabolic networks based on flux balance analysis and flux coupling analysis: Application to the analysis of *Escherichia coli* metabolism. *Biotechnol. Lett.* **2013**, *35*, 2039–2044.
16. Soons, Z.; Rocha, E.; Ferreira, I. Selection of elementary modes for bioprocess control. *Comput. Appl. Biotechnol.* **2010**, *11*, 156–161.
17. Soons, Z.; Ferreira, I.; Rocha, E. Identification of minimal metabolic pathway models consistent with phenotypic data. *J. Process Control* **2011**, *21*, 1483–1492.
18. Oddsdottir, H.; Hagrot, E.; Chotteau, V.; Forsgren, A. On dynamically generating relevant elementary flux modes in a metabolic network using optimization. *J. Math. Biol.* **2015**, *71*, 903–920.
19. Oddsdottir, H.; Hagrot, E.; Chotteau, V.; Forsgren, A. Robustness analysis of elementary flux modes generated by column generation. *Math. Biosci.* **2016**, *273*, 45–56.
20. Provost, A.; Bastin, G. Dynamic metabolic modelling under the balanced growth condition. *J. Process Control* **2004**, *14*, 717–728.
21. Zamorano, F.; Vande Wouwer, A.V.; Jungers, R.M.; Bastin, G. Dynamic metabolic models of CHO cell cultures through minimal sets of elementary flux modes. *J. Biotechnol.* **2013**, *164*, 409–422.
22. Hebing, L.; Neymann, T.; Thüte, T.; Jockwer, A.; Engell, S. Efficient generation of models of fed-batch fermentations for process design and control. *IFAC-PapersOnLine* **2016**, *49*, 621–626.
23. Abbate, T.; Fernandes de Sousa, S.; Dewasme, L.; Bastin, G.; Vande Wouwer, A. Inference of dynamical macroscopic models of cell metabolism based on elementary flux modes analysis. *Biochem. Eng. J.* **2019**, *151*, 107325.
24. Maton, M.; Vande Wouwer, A.; Bogaerts, P. Selection of a minimal suboptimal set of EFMs for dynamic metabolic modelling. *IFAC-PapersOnLine* **2021**, *54*, 667–672.
25. Savageau, M. Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.* **1969**, *25*, 365–369.
26. Savageau, M. Biochemical systems analysis. II. The steady-state solutions for n-pool system using a power-law approximation. *J. Theor. Biol.* **1969**, *25*, 370–379.
27. Voit, E.; Savageau, M. Equivalence between S-systems and Volterra-systems. *Math. Biosci.* **1982**, *78*, 47–55.
28. Savageau, M. Introduction to S-systems and the underlying power-law formalism. *Math. Comput. Model.* **1979**, *11*, 546–551.
29. Shiraishi, F.; Savageau, M. The tricarboxylic acid cycle in *Dictyostelium discoideum*. Formulation of the alternative kinetic representations. *J. Biol. Chem.* **1992**, *267*, 22912–22918.
30. Curto, R.; Sorribas, A.; Cascante, M. Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis: Model definition and nomenclature. *Math. Biosci.* **1995**, *130*, 25–50.
31. Torres, N.; Voit, E.; Gonzalez-Alcon, C. Optimization of nonlinear biotechnological processes with linear programming: Application to citric acid production by *Aspergillus niger*. *Biotechnol. Bioeng.* **1996**, *49*, 247–258.
32. Hernandez-Bermejo, B.; Fairen, V.; Sorribas, A. Power-law modeling based on least-squares minimization criteria. *Math. Biosci.* **1999**, *161*, 83–94.
33. Haag, J.; Vande Wouwer, A.; Remy, M. A general model of reaction kinetics in biological systems. *Bioprocess Biosyst. Eng.* **2005**, *27*, 303–309.
34. Naderi, S.; Meshram, M.; Wei, C.; McConkey, B.; I.B.; Budman, H.; Scharer, J. Development of a mathematical model for evaluating the dynamics of normal and apoptotic Chinese hamster ovary cells. *Biotechnol. Prog.* **2011**, *27*, 1197–1205.
35. Hagrot, E.; Oddsdottir, H.; Hosta, J.; Jacobsen, E.; Chotteau, V. Poly-pathway model, a novel approach to simulate multiple metabolic states by reaction network-based model—Application to amino acid depletion in CHO cell culture. *J. Biotechnol.* **2017**, *259*, 235–247.

36. Mailier, J.; Vande Wouwer, A. Identification of nested biological kinetic models using likelihood ratio tests. *Chem. Eng.* **2012**, *84*, 727–734.
37. Grosfils, A.; Vande Wouwer, A.; Bogaerts, P. On a general model structure for macroscopic biological reaction rates. *J. Biotechnol.* **2007**, *130*, 253–264.
38. Richelle, A.; Bogaerts, P. Systematic methodology for bioprocess model identification based on generalized kinetic functions. *Biochem. Eng. J.* **2015**, *100*, 41–49.
39. Montague, G.; Morris, J. Neural-network contributions in biotechnology. *Trends Biotechnol.* **1994**, *12*, 312–324.
40. Chen, L.; Bernard, O.; Bastin, G.; Angelov, P. Hybrid modeling of biotechnological processes using neural networks. *Control Eng. Pract.* **2000**, *8*, 821–827.
41. Vande Wouwer, A.; Renotte, C.; Bogaerts, P. Biological reaction modeling using radial basis function networks. *Comput. Chem. Eng.* **2004**, *28*, 2157–2164.
42. Maton, M.; Vande Wouwer, A.; Bogaerts, P. A systematic elementary flux mode selection procedure for deriving macroscopic bioreaction models from metabolic networks. *J. Process Control* **2022**, *118*, 170–184.
43. Amribt, Z.; Hongxing, N.; Bogaerts, P. Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures. *Biochem. Eng. J.* **2013**, *70*, 196–209.
44. Meyer-Baese, A.; Schmid, V. Foundations of neural networks. In *Pattern Recognition and Signal Analysis in Medical Imaging*, 2nd ed.; Academic Press: Cambridge, MA, USA, 2014; pp. 197–243.
45. Hecht-Nielsen, R. Theory of the backpropagation neural network. In *Neural Networks for Perception: Computation, Learning, and Architectures*; Academic Press: Cambridge, MA, USA, 1992; pp. 65–93.
46. Janowsky, S. Pruning versus clipping in neural networks. *Phys. Rev. A* **1989**, *39*, 6600.
47. Reed, R. Pruning algorithms: A survey. *IEEE Trans. Neural Netw.* **1993**, *4*, 740–747.
48. Hang, S.; Pool, J.; Tran, J.; Dally, W. Learning both weights and connections for efficient neural network. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 1135–1143.
49. Frankle, J.; Carbin, M. The lottery ticket hypothesis: Finding sparse, trainable neural networks. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
50. Tanaka, H.; Kunin, D.; Yamins, D.; Ganguli, S. Pruning neural networks without any data by iteratively conserving synaptic flow. *Adv. Neural Inf. Process. Syst.* **2015**, *33*, 6377–6389.
51. Silvestre, M.; Ling, L. Pruning methods to MLP neural networks considering proportional apparent error rate for classification problems with unbalanced data. *Measurement* **2014**, *56*, 88–94.
52. Niu, H.; Amribt, Z.; Fickers, P.; Tan, W.; Bogaerts, P. Metabolic pathway analysis and reduction for mammalian cell cultures—Towards macroscopic modeling. *Chem. Eng. Sci.* **2013**, *102*, 461–473.
53. Fernandes de Sousa, S.; Bastin, G.; Jolicoeur, M.; Vande Wouwer, A. Dynamic metabolic flux analysis using a convex analysis approach: Application to hybridoma cell cultures in perfusion. *Biotechnol. Bioeng.* **2015**, *113*, 1102–1121.
54. Dochain, D.; Vanrolleghem, P. Identification of bioprocess models. In *Bioprocess Control*; ISTE Ltd.: London, UK; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2008; pp. 47–76.
55. Quesney, S.; Marc, A.; Gerdil, C.; Gimenez, C.; Marvel, J.; Richard, Y.; Meignier, B. Kinetics and metabolic specificities of Vero cells in bioreactor cultures with serum-free medium. *Cytotechnology* **2003**, *42*, 1–11.