Dynamic estimations of metabolic fluxes with constraint-based models and possibility theory

Francisco Llaneras*, Antonio Sala, Jesús Picó.

AI2, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain (francisco.llaneras@gmail.com; jpico@ai2.upv.es; asala@isa.upv.es)

*Corresponding author: Dr. Francisco Llaneras
Universidad Politécnica de Valencia, Camino de Vera s/n, 46022
Phone: +34 654700098 — Email: francisco.llaneras@gmail.com

Living cells can be modelled by successively imposing known constraints that limit their behaviour, such as mass balances, thermodynamic laws or enzyme capacities. The resulting constraint-based models enclose all the functional states that the modelled *cells may exhibit. Then, predictions can be obtained from the models in two main ways: adding experimental data to determine the state of cells at given conditions (MFA) or invoking an assumption of evolved optimal behaviour (FBA). Both MFA and FBA predictions are typically performed at steady state. However, it is easy to take extracellular dynamics into account. This work explores the benefits of using possibility theory to get these dynamic predictions. It will be shown that the possibilistic methods (a) provide rich estimates for time-varying fluxes and metabolite concentrations, (b) account for uncertainty and data scarcity, and (c) give predictions relaxing the optimality assumption of FBA.* On the other hand, *these methods could serve as basis for monitoring and fault detection systems in industrial bioprocesses.*

Highlights

- Possibility theory is used to get dynamic estimates of fluxes and metabolites.
- The approach handles measurements errors, uncertainty, and data scarcity.
- Predictions can be obtained relaxing the assumption of cells evolved optimality.

Keywords

Constraint satisfaction problems, possibility theory, constraint-based model, metabolic flux analysis, flux balance analysis.

Abbreviations

MFA:	Metabolic Flux Analysis. Reviewed in (Llaneras, 2008).
FBA:	Flux Balance Analysis. Reviewed in (Palsson, 2006).
Poss-MFA:	Possibilistic approach to MFA. Introduced in (Llaneras, 2009).
Dynamic Poss-MFA:	Extension of Poss-MFA. Introduced in this paper.
Dynamic Poss-FBA:	Possibilistic approach to FBA. Introduced in this paper.
$\mathcal{MOC}$:	Model-based Constraints.
$\mathcal{MEC}$:	Measurement-based Constraints.
$\mathcal{CAC}$:	Capacity-based Constraints.

1. Introduction

To facilitate the study of biological systems under a system-level approach, constraint-based metabolic models have recently been assembled (Llaneras, 2008; Palsson, 2006). The cell metabolism can be viewed as a chemical "factory" that converts available raw materials into energy as well as building blocks needed to produce biological structures, maintain cells alive, and carry out various cellular functions. This process can be represented with a metabolic network that encodes the biochemical reactions taking place within the cells—the nodes represent the involved metabolites and the edges the reaction rates or metabolic fluxes; internal fluxes correspond to reactions within cells and exchange fluxes to exchanges of compounds between the cells and their environment (uptake of substrates and formation of products). The set of flux values at a given moment defines the metabolic state of the cells by representing which reactions are occurring and at which rate. An example of a metabolic network is given in figure 1.

These metabolic networks can be relatively complex and difficult to model. In principle, one could consider all the mechanisms operating in the metabolism, and this will lead to detailed quantitative predictions on cellular dynamics. However, this approach is difficult due to the lack of knowledge on the intracellular reactions and its parameters. As an alternative, the so-called constraint-based models impose mass balances, but disregard the dynamics of the (fast) intracellular reactions, and assume that most internal metabolites rapidly reach their steady-state. This way the state of the cells is represented without any information on the kinetics of the reactions. Cells are also subject to other constraints that limit their behaviour, such us thermodynamics or enzyme capacities. By assembling a set of constraints into a model, it is possible to determine which functional states can and cannot be achieved by a cell.

Constraint-based models are being used with different purposes (Llaneras, 2008), but this paper is focused on methods to estimate and predict cellular behaviour. We consider two main approaches to this problem: (a) metabolic flux analysis (MFA), which considers experimental data to perform estimations (Heijden, 1994), and (b) flux balance analysis (FBA), which assumes an evolved optimality of cells for the same purpose (Edwards, 2002). These predictions are typically static, aimed to study cells at a state of particular interest; however, several works taking extracellular dynamics into account can be found in the literature both in the context of MFA (Herwig, 2002; Takiguchi, 1997; Henry, 2007) and FBA (Mahadevan, 2002; Hjersted, 2009).

In this paper we extend the possibilistic framework introduced in (Llaneras, 2009) to address MFA and FBA considering extracellular dynamics. The paper is organised as follows: preliminaries on extracellular dynamics and possibility theory are given in section 2. In section 3.1 Possibilistic MFA is extended to get dynamic (time-varying) estimations of fluxes and metabolite concentrations; in section 3.2 it is shown how to formulate a FBA problem in the possibilistic framework. The methods described in section 3 are illustrated with two experimental case studies in section 4. Finally, some conclusions are withdrawn.

2. Preliminaries

2.1 Dynamics and constraint-based models

Let us start formulating a conventional constraint-based model. Consider a population of cells in an aqueous medium as environment. Given a metabolic network of the cells to be modelled (see figure 1 for an example), mass balances around the intracellular metabolites can be represented as follows:

$$\frac{d\mathbf{c}}{dt} = \mathbf{N} \cdot \mathbf{v} - \mu_x \cdot \mathbf{c} \tag{1}$$

where $\mathbf{c}$ is the m-dimensional vector of intracellular-metabolite concentrations, $\mathbf{v}$ the n-dimensional vector of fluxes through each reaction, μ_x is the specific growth rate of biomass cells, and $\mathbf{N}$ is the stoichiometric matrix linking fluxes and internal metabolites.

Typically, the internal metabolites are assumed to rapidly reach their steady-state, thus avoiding to consider the rarely known reaction kinetics. The dilution term $\mu_x \cdot \mathbf{c}$ is also often neglected. This way equation (1) turns into (2a), the so-called stoichiometric equation (Stephanopoulos, 1990). Other constraints can then be incorporated, such as the irreversibility of reactions. These can be represented by a set of inequalities (2b), where $\mathbf{D}$ is a diagonal matrix with $\mathbf{D_{ii}} = 1$ if the flux i is irreversible ($\mathbf{D_{ii}} = 0$ otherwise). This leads to the constraint-based models typically used to analyse the fluxes at given cellular steady-state:

$$\mathcal{MOC}_{ss} = \begin{cases} \mathbf{0} = \mathbf{N} \cdot \mathbf{v} & \text{(2a)} \\ \mathbf{0} \leq \mathbf{D} \cdot \mathbf{v} & \text{(2b)} \end{cases}$$

At this point one may want to account for the dynamics of the extracellular metabolites, i.e. those substrates and products that are exchanged between the cells and their environment. Mass balances around the extracellular metabolites can be established as follows:

$$\frac{d\mathbf{e}}{dt} = \mathbf{N_e} \cdot \mathbf{v} - D \cdot \mathbf{e} + \mathbf{F_e} \tag{3}$$

where $\mathbf{e}$ denotes the vector of extracellular-metabolites concentration, D is the dilution rate (inflow per volume), $\mathbf{F_e}$ is the net inflow/outflow of extracellular metabolites into the environment, and $\mathbf{N_e}$ is a selection matrix linking each metabolite e with the flux in $\mathbf{v}$ that represents its uptake or production rate. Without loss of generality, each extracellular metabolite in $\mathbf{e}$ is represented with two nodes, one intra- and one extracellular, so there is only one flux in $\mathbf{v}$ accounting for its total uptake or consumption. Biomass is also represented as an external metabolite and its synthesis with one flux in $\mathbf{v}$.

In summary, a constraint-based model considering the extracellular dynamics can be formulated with equations (2-3). This is the type of models we will use in section 3. Similar formulations can be found in the literature in the context of macroscopic models (Provost,

2004; Haag, 2005; Bastin, 2007) and also in the context of MFA (Herwig, 2002; Takiguchi, 1997; Henry, 2007) and FBA (Mahadevan, 2002; Hjersted, 2009).

2.2 Possibility theory and MFA

In a recent work we introduced Poss-MFA as a possibilistic framework to perform metabolic flux estimations using constraint-based models (Llaneras et al., 2009). The framework exploits the notion of "possibilistic constraint satisfaction problems", which was introduced in (Dubois, 1996). Similar optimization approaches to logic reasoning were previously explored in (Sala, 2001; Sala, 2008). The framework is based on two ideas: (1) represent knowledge with constraints satisfied to a certain degree, thus transforming the feasibility of a potential solution into a gradual notion of "possibility" that accounts for uncertainty, and (2) use computationally efficient optimisation-based methods to query for the "most possible" solutions.

This framework provides a simple and powerful way to deal with uncertainty both in the measurements and the model (e.g., imprecision and lack of knowledge), which is a typical difficulty in flux estimation problems. Herein we summarize the formulation of Poss-MFA, but further details can be found in (Llaneras et al., 2009; Llaneras, 2011).

Constraints: model and measurements. Let us start considering the constraints conforming the model ($\mathcal{MOC}$) that were given in equation (2). Then, we incorporate measurements of (some) extracellular fluxes as additional linear constraints, the measurement-based constraints ($\mathcal{MEC}$):

$$\mathcal{MEC} = \begin{cases} \mathbf{v_m} = \mathbf{w} + \varepsilon - \mu + \varepsilon_f - \mu_f \\ \quad 0 \leq \varepsilon, \mu \\ \quad 0 \leq \varepsilon_f \leq \varepsilon_f^M \\ \quad 0 \leq \mu_f \leq \mu_f^M \end{cases} \tag{4}$$

where $\mathbf{v_m}$ is the vector of actual fluxes for each measured flux, and $\mathbf{w}$ is the vector of measured values. Both differ due to errors and imprecision. This uncertainty will be represented by the slack variables ε, μ, ε_f and μ_f. The slack variables ε and ε_f represent negative errors in the measurement $\mathbf{w}$, whereas μ and μ_f represent positive ones. As will be explained below, ε and μ are "penalised" errors in a cost index (5), to assign a decreasing possibility to increasing errors, while ε_f and μ_f will be non-penalised, but bounded ones, thus defining a band of fully possible $\mathbf{v_m}$ values around the measured $\mathbf{w}$.

Possibility. Let us denote each candidate solution of (2)-(4) as $\delta = \{\mathbf{v}, \varepsilon, \mu, \varepsilon_f, \mu_f\}$ in some decision space Δ. The basic building block of possibility theory is a user-defined possibility distribution $\pi(\delta):\Delta \rightarrow [0,1]$. This function assigns a possibility for each solution ranging between impossible ($\pi=0$) and fully possible ($\pi=1$). Among different possible choices, a simple yet sensible way to build the function is introducing a linear cost index J to penalise large

deviations between actual fluxes and their measured values:

$$J(\delta) = \alpha \cdot \varepsilon + \beta \cdot \mu \tag{5}$$

and define the possibility of each solution δ as follows:

$$\pi(\delta) = e^{-J(\delta)} \quad \delta \in \Delta \tag{6}$$

This way equations (4-6) can be interpreted as representing the statement: «given a measured value w, the assertion $v_m = w$ is fully possible, and the more v_m and w differ, the less possible such situation is».

The actual possibility for each value of v_m depends on how the user defines the "sensor" accuracy coefficients in (4) and (5): the maximum bounds for ε_f^M and μ_f^M define an interval of fully possible values ($\pi=1$), and the possibility of v_m being out of this interval depends on the user-selected weights α and β in (5). Notice that measurements uncertainty can be non-symmetric, and that very complex descriptions can be achieved by adding slack variables.

Remark. Although there are alternatives, possibilistic conjunction in equation (6) is defined under a non-interactivity assumption, in analogy to statistical independence (Llaneras *et al.*, 2009). Under this assumption the joint possibility of two variables $\delta_1, \delta_2 \in \Delta$ can be expressed as the product of two univariate ones: $\pi(\delta_1, \delta_2) = \pi(\delta_1)\pi(\delta_2)$, with $J(\delta_1, \delta_2) = J(\delta_1) + J(\delta_2)$. This implies that the joint possibility of two non-fully possible event is lower than their independent possibilities. Consider this example, if $\pi(\mathbf{v}_4 = 3)$ is 0.9 and $\pi(\mathbf{v}_7 = 2)$ is 0.5, the joint possibility $\pi(\mathbf{v}_4 = 3 \cap \mathbf{v}_7 = 2)$ is 0.45.

Poss-MFA Steady-State flux estimation. The simplest flux estimate $\mathbf{v_{mp}}$ is given by the maximum possibility (minimum cost) solution $\delta_{mp} = \{\mathbf{v_{mp}}, \varepsilon_{mp}, \mu_{mp}, \varepsilon_{f,mp}, \mu_{f,mp}\}$ to the constraint satisfaction problem (2)-(4), which can be computed efficiently solving one linear programming problem (LP):

$$J^{\min} = \min_{\varepsilon, \mu, v} J \quad \text{s.t.} \begin{cases} \mathcal{MOC} \\ \mathcal{MEC} \end{cases} \tag{7}$$

Notice that point-wise estimates would be unreliable if multiple solutions were reasonably possible, so it is advisable to get interval estimates and distributions. Interval estimates for desired marginal (π) and conditional or *a posteriori* (γ) possibilities can be obtained, again, solving efficient LP optimisations (Llaneras, 2011).

For example, the interval of values with a conditional possibility higher than γ for a given flux is obtained solving three LP optimisations; first problem (7), and once $J^{\min}$ is available, the one below (lower bound) and the optimisation replacing minimum by maximum (upper bound):

$$v_{i,\gamma}^m = \min_{\varepsilon,\mu,v} v_i \quad \text{s.t.} \begin{cases} \mathcal{MOC} \cap \mathcal{MEC} \\ J - J^{\min} < \log \gamma \end{cases} \tag{8}$$

These possibilistic intervals have a similar interpretation to confidence intervals (*credible intervals*) in Bayesian statistics. In practice, getting estimates with conditional possibilities γ is equivalent to normalise the marginal possibility distributions to a maximum equal to one. In this paper we typically compute the interval of most possible values (γ=1) and other less possible intervals (γ=0.8, 0.5 and 0.1) to capture uncertainty.

Pos-MFA to evaluate consistency. Poss-MFA can also be used to evaluate the degree of consistency between a model ($\mathcal{MOC}$) and some experimental measurements ($\mathcal{MEC}$). Simply notice that the most "possible" solution (7) has an associated possibility value, $\pi^{mp} = \exp(-J^{\min})$ that grades consistency. Possibility equal to one must be interpreted as complete agreement between the model and the measurements, whereas lower values imply that there is certain degree of error in measurements, the model or both. The evaluation of consistency can be used to (a) conciliate a set of experimental measurements, (b) serve as basis for process monitoring and fault detection systems, or (c) validate a constraint-based model (Llaneras, 2011; Llaneras *et al.* 2009) and (Tortajada *et al.*, 2010).

3 Theory

3.1 Dynamic Possibilistic MFA

In this section we extend the Poss-MFA framework to take extracellular dynamics into account.

Imposing constraints. Let us consider a batch process (i.e., with no inflow nor outflow, and thus with $\mathbf{F_e} = D = 0$) during time [0, T], and assume measurements are sampled with some given sampling rate. The first step is to define the model-based constraints at each successive sampling instant *k*:

$$\mathcal{MOC}(k) = \begin{cases} \mathbf{0} = \mathbf{N} \cdot \mathbf{v}(k) & \text{(9a)} \\\\ \mathbf{D} \cdot \mathbf{v}(k) \geq \mathbf{0} & \text{(9b)} \\\\ \dfrac{\mathbf{e}(k) - \mathbf{e}(k-1)}{\Delta T} = \mathbf{N_e} \cdot \mathbf{v}(k) & \text{(9c)} \\\\ \mathbf{e}(k) \geq \mathbf{0} & \text{(9d)} \end{cases}$$

Equation (9a) and (9b) represent the mass balances around the internal metabolites and the reactions irreversibility, and are equivalent to the equation (2) but accounting for successive time instants. Equation (9c) represents the mass balance of extracellular metabolites (3) using a backward approximation of the derivatives for simplicity. Adding these mass balances

implies that the concentrations of extracellular metabolites $\mathbf{e}(k)$ are now variables of the problem. Equation (9d) constrains the concentrations to be positive. Hereinafter, we denote the set of system variables as $\mathbf{var}(k)$, with $\mathbf{var}(k) = \{\mathbf{v}(k), \mathbf{e}(k)\}$.

Once the model constraints have been defined, the second step will be to incorporate measurements at successive sampling instants as additional constraints, thus defining the measurement-based constraints, $\mathcal{MEC}(k)$:

$$\mathcal{MEC}(k) = \begin{cases} \mathbf{e_m}(k) = \mathbf{f_m}(k) + \varepsilon(k) - \mu(k) + \varepsilon_f(k) - \mu_f(k) & (10a) \\ \\ \varepsilon(k), \mu(k) \geq 0 & (10b) \\ \\ 0 \leq \varepsilon_f(k) \leq \varepsilon_f^M(k) & (10c) \\ \\ 0 \leq \mu_f(k) \leq \mu_f^M(k) & (10d) \end{cases}$$

where vector $\mathbf{e_m}(k)$ represents the actual (unknown) metabolite concentrations and $\mathbf{f_m}(k)$ their measured values. Notice that initial conditions $\mathbf{e}(0)$ should be given for non-measured metabolites.

Following the ideas in the preliminaries section, we define a cost index $J(k)$ for each time instant k,

$$J(\delta(k)) = \alpha(k) \cdot \varepsilon_1(k) + \beta(k) \cdot \mu_1(k) \tag{11}$$

and aggregate them by non-interactive conjunction defining a total cost index,

$$J_T = \sum_{k=0}^{t} J(k) \tag{12}$$

where $\alpha(k)$ and $\beta(k)$ are row vectors of user-defined coefficients that represent sensor accuracy associated with each measurement at each time instant (together with the defined bounds for the slack variables, $\varepsilon_f^M(k)$ and $\mu_f^M(k)$).

Estimation of fluxes and metabolites. Once that the constraint-based problem has been formulated, the fluxes and the metabolite concentrations can be estimated along the process duration. The straightforward approach to the problem is to consider all the operating constraints at each time instant simultaneously, thus taking into account all the available knowledge to get each estimate. Although this approach could be computationally intense if the sampling rate needs to be high, this is not the case in most situations (the issue is discussed below).

The simplest estimate is the solution with the highest possibility value of the defined constraint satisfaction problem (9-10), which is obtained solving an LP optimisation problem:

$$\mathrm{J}_{\mathrm{T}}^{\mathrm{min}} = \min_{\varepsilon(k),\mu(k),\mathbf{v}(k)} \mathrm{J}_{\mathrm{T}} \quad \text{s.t.} \left\{ \begin{array}{ll} \mathcal{MOC}(k) & \forall k \\ \mathcal{MEC}(k) & \forall k \end{array} \right. \tag{13}$$

The solution $\{\mathbf{var_{mp}}(k)\}_{k=1\ldots t}$ contains the "most possible" value for each flux and metabolite at each time instant. However, as explained above, interval estimates should be obtained to provide more reliable results in case multiple solutions are reasonably possible. For instance, to compute the interval of values with conditional possibility higher than γ for a flux (metabolite) at a given time, $[\,var_{i,\gamma}^{m}(k)\,,var_{i,\gamma}^{M}(k)\,]$ we solve three LP optimisations; first we solve (13), and subsequently the following problem (to get the lower bound) and the same one replacing minimum by maximum (to get the upper bound):

$$var_{i,\gamma}^{m}(k) = \min var_{\mathrm{i}}(k)$$

$$\text{s.t.} \left\{ \begin{array}{ll} \mathcal{MOC}(k) & \forall k \\ \mathcal{MEC}(k) & \forall k \\ \mathrm{J}_{\mathrm{T}}^{\mathrm{min}} - \log \pi(var_{mp}) < -\log \gamma \end{array} \right. \tag{14}$$

These possibilistic intervals provide a rich and concise estimation. Remember also that the whole possibility distribution of a particular variable can be reconstructed obtaining the intervals for a grid of possibilities, to say, γ=1, 0.9, 0.8, ... 0.1 (Llaneras *et al.*, 2009). Notice also that (14) can be used to estimate both metabolic fluxes $\mathbf{v}(k)$, and extracellular metabolite concentrations $\mathbf{e}(k)$. This means that it is possible to estimate the evolution of non-measured metabolites, such as the concentration of a product of interest or a by-product. An example will be given in a case study below.

Remark about computational efficiency. One feature of the Possibilistic framework is that the estimations are cast as efficient linear optimisation problems (LP). The most possible solution (13) is obtained solving just one LP, and each interval estimate (14) is obtained solving 2·*t* extra LP problems. There are widely-known and efficient software tools to solve these problems —with thousands of variables and constraints—, so the method is in principle suitable for large-scale models. However, as it has been described, the procedure could become computationally expensive if the extracellular metabolites needed to be measured with a high sampling rate, because the number of constraints grows with *t*. This situation will be rare because extracellular dynamics are typically slow. Anyway, if this is the case, the computational cost of the estimation procedure can be reduced adopting «finite-window» approaches (details are straightforward, omitted for brevity). The benefit of this «finite-window» formulation is that the possibilistic estimates are obtained solving LP problems that do not grow with the sampling rate. The price to pay is that, since less constraints are considered at each time, the solution space will eventually become larger and the intervals will be wider than those given by using a «whole-window» (14). This means that the "finite-window" approach may lead to less insight, although in no case it will lead to results deviated from those provided by the «whole-window» approach. Most of the works in the literature that account for extracellular dynamics follow approaches that can be seen as «finite-window»

with windows of size one (Herwig, 2002; Takiguchi, 1997; Henry, 2007; Mahadevan, 2002; Hjersted, 2009). However, using windows of convenient size, the users could chose the trade-off between computational efficiency and the more precise estimates of the «whole-window».

3.2. Dynamic Possibilistic FBA

Flux balance analysis (FBA) is a methodology to get predictions from a constraint-based model invoking an assumption of optimal cells behaviour (Edwards *et al.* 2002; Palsson, 2006). Basically, one particular cellular state among those that are feasible is promoted based on the assumption that cells have evolved to be optimal, and that their "objective" is known and can be expressed in convenient mathematical terms.

In this work we consider the particular case where extracellular dynamics are accounted for, as shown in (9), and FBA is used to predict fluxes and metabolite concentrations during a cultivation process (Mahadevan, 2002; Hjersted, 2009). The novelty of the approach is that possibility theory will be used. This provides a simple way to relax the optimality assumption by defining optimality in a gradual way: an optimal state is considered fully possible, and the more a state differs from it, the less possible such situation is considered. This allows, for example, to get dynamic predictions that capture alternate optima, and allow for a certain degree of sub-optimality.

Imposing constraints. Let us consider the same model constraints, $\mathcal{MOC}(k)$ at each time instant k, as done in (9). Measurements, however, are not incorporated. Instead, a few uptake fluxes are typically constrained based on (a) known capacities, (b) a kinetic expression or (c) their availability in the media. These constraints will be denoted as capacity-based constraints, $\mathcal{CAC}(k)$:

$$\mathcal{CAC}(k) = \mathbf{v}_{\mathbf{u}}^{\mathrm{m}}(k) \geq \mathbf{v}_{\mathbf{u}}(k) \geq \mathbf{v}_{\mathbf{u}}^{\mathrm{M}} \tag{16}$$

A hypothetical cells objective. FBA predictions are obtained by invoking an optimal use of resources, and therefore the cells objective must be defined (e.g., maximum growth). Typically the objective is expressed by means of a linear cost index $Z(k)$:

$$Z(k) = \mathbf{d} \cdot \mathbf{v}(k) \tag{17}$$

At this point FBA predictions at each time k could be obtained by maximising $Z(k)$ subject to the operating constraints. However, some refinements can be easily incorporated using the possibilistic framework.

Invoke optimality in a gradual way. To account for optimality in a gradual way, we will define the possibility of each candidate solution in terms of optimality. First, the slack variable $\phi(k)$ is introduced to represent sub-optimality as function of $Z(k)$ and its maximum $Z^{\mathrm{max}}(k)$,

$$\phi(k) = \frac{Z^{\mathrm{max}}(k) - Z(k)}{Z^{\mathrm{max}}(k)}, \quad 0 \leq \phi(k) \leq 1 \tag{19}$$

Possibility is then defined as a function of optimality using a cost index J_{opt},

$$J_{opt}(k) = \alpha_s \cdot \phi(k) \tag{20a}$$

$$\pi(k) = e^{-J_{opt}(k)} \tag{20b}$$

where α_s is the user-defined weight that links possibility and optimality.

Equation (20) implies that only cellular states **var**(k) maximising $Z(k)$ are fully possible (even if multiple optima may exist); and the more a state differs from the optimum, the less possible such situation is considered. The parameter α_s allows to define possibility as a "degree of belief", depending on the credibility that the user associates with different levels of sub-optimality. For instance, if the experimental evidence shows that the strain being modelled typically exhibits high growth optimality —for example—, the user should choose the parameter to strongly penalise sub-optimal states. Furthermore, the flexibility to define possibility as function of optimality —but not necessarily equivalent— makes it possible to consider other sources of uncertainty (e.g., parameter variability). Sometimes the degree of optimality to be shown by the modelled cells is not well-known. In these cases one could link possibility and optimality in an arbitrary but simply way: e.g. choosing $\alpha_s = -2 \cdot \log(0.5)$, possibility $\pi = 0.5$ is assigned to half optimal states, where $Z(k) = 0.5 \cdot Z^{max}(k)$. This way, possibility roughly indicates the degree of sub-optimality.

Getting dynamic predictions. Once optimality has been defined within the framework, predictions with desired sub-optimality (possibility) π can be obtained successively at each k following a two-step procedure.

At each time instant k, the first step provides the value of $Z^{max}(k)$, as follows:

$$
\textit{Step 1} \quad \left|
\begin{array}{l}
Z^{max}(k) = \max \; Z(k) \\[1em]
\text{s.t.} \begin{cases}
\quad \mathcal{MOC}(i) & \textit{1...k} \\
\quad \mathcal{CAC}(i) & \textit{1...k} \\
\phi(i) = \dfrac{Z^{max}(i) - Z(i)}{Z^{max}(i)} & \textit{1...(k-1)} \\
\alpha_s \cdot \phi(i) < \log \pi & \textit{1...(k-1)}
\end{cases}
\end{array}
\right. \tag{21}
$$

The last three constraints guarantee that the optimal solution at k does not violate the optimality π at previous time instants $\{1 \ldots k\text{-}1\}$.

In the second step, $Z^{max}(k)$ is used as reference to get the sub-optimal predictions as possibilistic intervals, $[\, var_{i,\gamma}^m(k) \,, var_{i,\gamma}^M(k)\,]$, solving a pair of LP optimisation problems:

Step 2

$$var_{i,\gamma}^{m}(k) = \min var_i(k)$$

$$\text{s.t.} \begin{cases} \mathcal{MOC}(i) & 1...k \\ \mathcal{CAC}(i) & 1...k \\ Z(k) = Z^{\max}(i)\cdot\left(1-\phi(i)\right) & 1...k \\ 0 \leq \phi(i) \leq 1 & 1...k \\ \alpha_s \cdot \phi(i) < \log \pi & 1...k \end{cases} \tag{22}$$

Bound $var_{i,\gamma}^{M}(k)$ is obtained by replacing minimum by maximum.

The two-step procedure can be repeated for different degrees of possibility thus getting a rich prediction that considers sub-optimality and captures alternate optima.
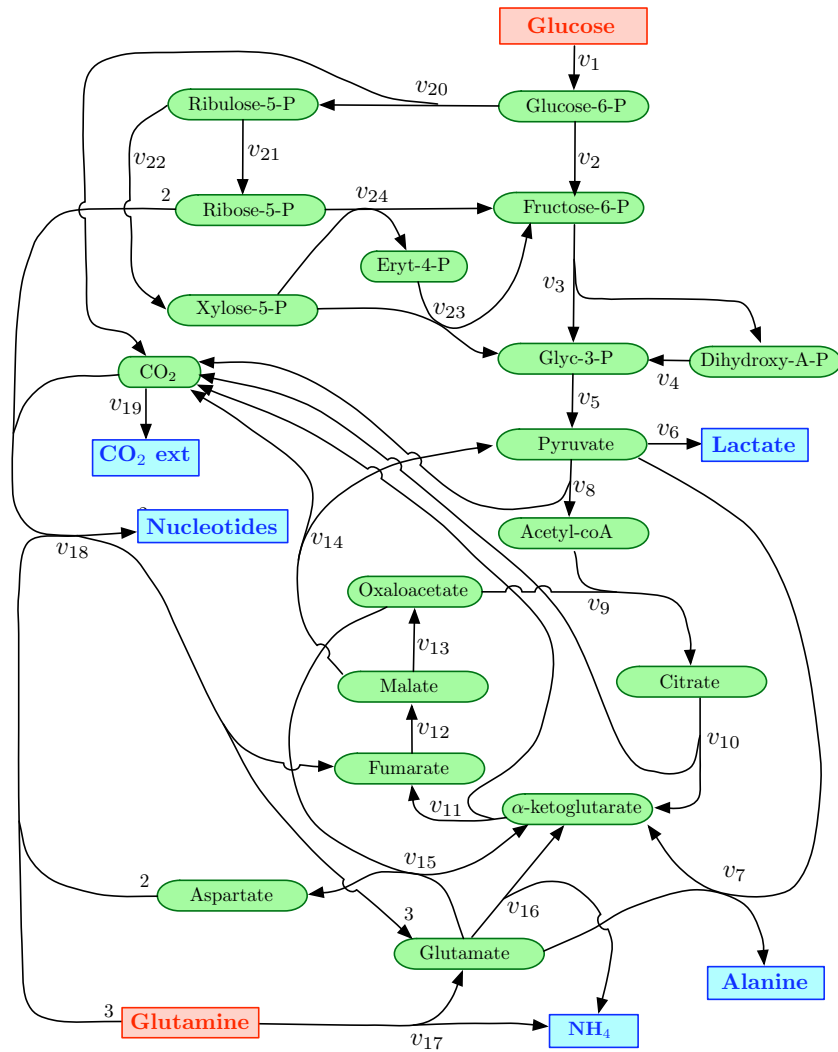

Figure 1. Intracellular metabolic network of CHO cells (Bastin, 2007).

4 Cases studies

Two experimental cases studies will be presented to illustrate the methods described so far. In section 4.1 dynamic Poss-MFA will be applied to monitor a CHO cells culture; in section 4.2 it will be applied to an *E. coli* batch process.

4.1 Dynamic Poss-MFA: monitoring of a CHO cells culture

Preparation. As first example we consider a cultivation of Chinese Hamster Ovary cells (CHO cells) in batch mode. A metabolic network that describes the metabolism concerned with the two main energetic nutrients, glucose and glutamine has been taken from (Bastin, 2007). The network, which is depicted in Figure 1, includes 31 reactions (24 internal ones, six exchanges and the biomass growth) and 25 metabolites (listed in Tables 1 and 2). There are no redundant mass balances, so the network has six degrees of freedom. The corresponding 25×31 stoichiometric matrix **N** is given in Table 3. The vector of reactions irreversibility, which defines the diagonal of the matrix **D**, is also given in Table 3. With matrices **N** and **D** the constraint-based model (5) is completely defined.
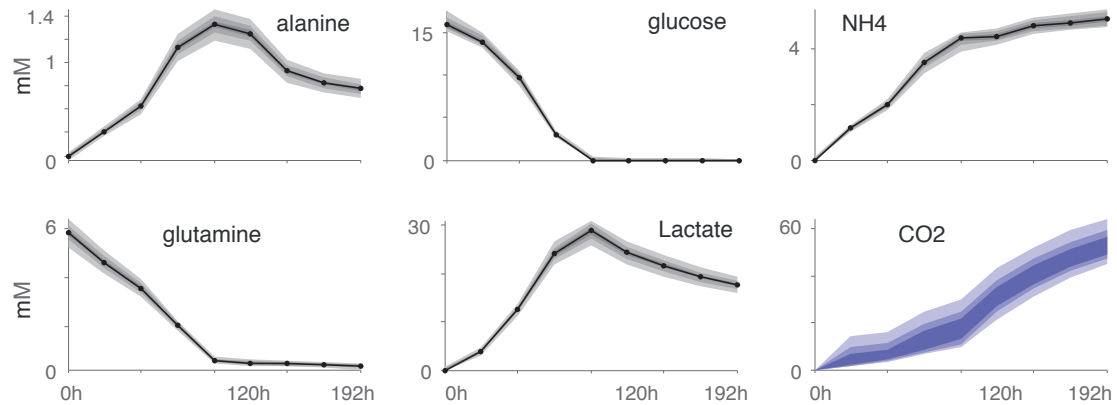
Figure 2. Measured and estimated metabolite concentrations during a cultivation of CHO cells. Measurements are denoted with black dots. The concentrations estimated with Poss-MFA for three degrees of possibility (π=1, 0.5, 0.15) are denoted with grey and blue areas.

Measurements of concentration for glucose (G), alanine (A), lactate (L), glutamine (Q) and ammonia (NH4), and biomass were taken from (Provost, 2006). Data were collected using a sample rate of 24 h. The uncertainty of the measurements is represented in possibilistic terms, as follows:

(i) Full possibility (π=1) is assigned to values with less than 5% of deviation.

(ii) Larger deviations are penalised, so values with a deviation equal to 20% have a possibility of π=0.1, and those with a deviation equal to 10% have a π≈0.5.

(iii) Symmetric uncertainty is considered (α=β).

To achieve this representation we must choose particular bounds $\varepsilon_f^M(k)$ and $\mu_f^M(k)$ and weights $a(k)$ and $\beta(k)$ for each measurement; condition (i) implies that $\varepsilon_f^M(k) = \mu_f^M(k) = 0.05\mathbf{w}(k)$, whereas (ii) defines $a(k)$, noticing that $0.2\mathbf{w}(k) = \mu_{20\%} + \mu_f^M(k)$, then $a(k) = -\log(0.1)/(0.2-0.05)/\mathbf{w}(k)$.

Dynamic Poss-MFA estimation of fluxes and metabolites. Dynamic Poss-MFA was applied to estimate the metabolic fluxes and the metabolite concentrations along the cultivation process (0-192h). We used (14) to compute three possibilistic interval ($\gamma=1$, 0.5, 0.15) for each variable at each time instant. This implies solving $2\cdot3\cdot9$ LP problems ($2\cdot p\cdot t$) for each variable.
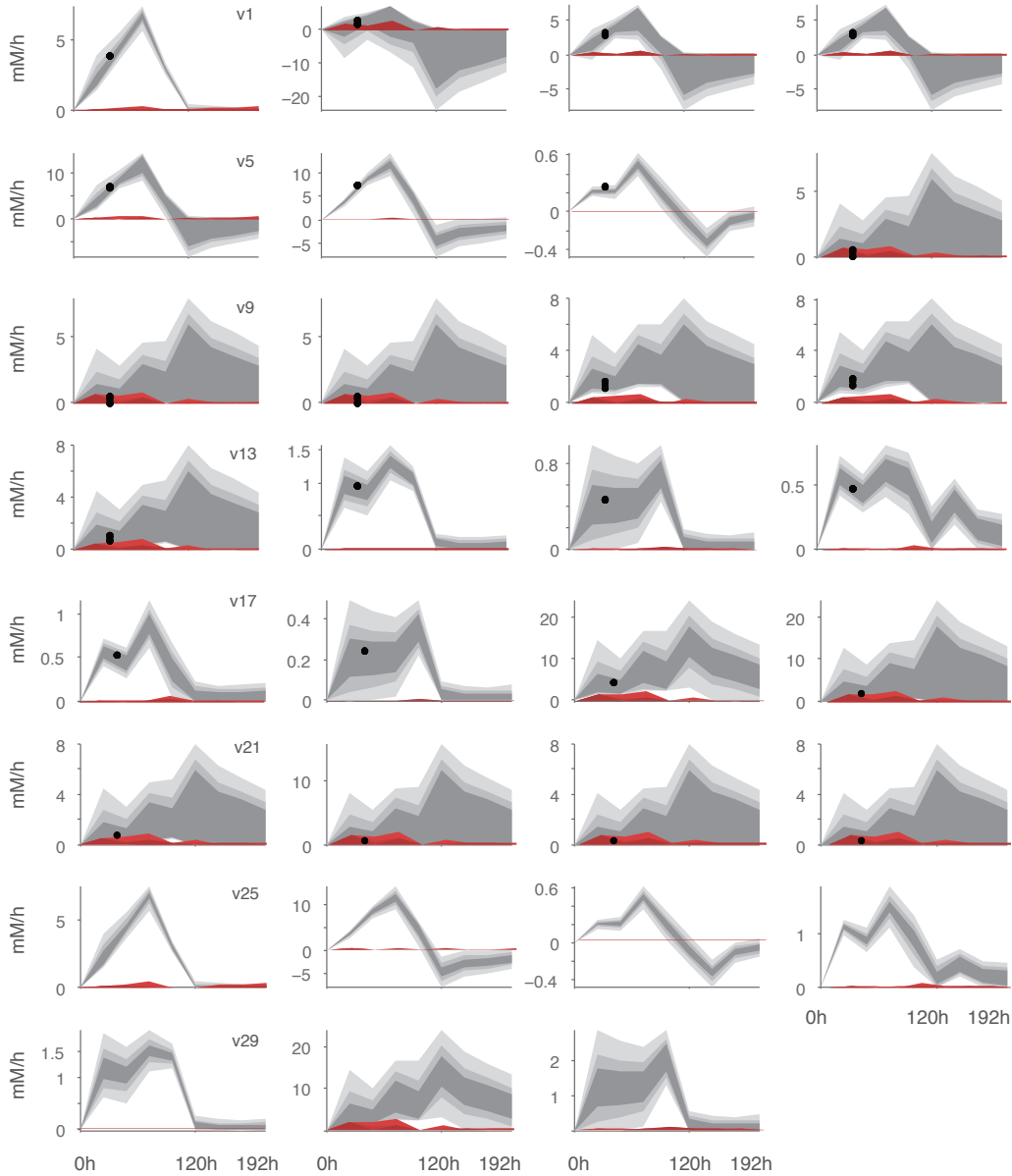
Figure 3. Estimated fluxes during a cultivation of CHO cells. The fluxes estimated with dynamic Poss-MFA for three degrees of possibility ($\pi=1$, 0.5, 0.15) are depicted in grey areas. The fluxes were also estimated with the «finite-window» approach for same three degrees of possibility; the oversize with respect to the «whole-window» approach is represented with red areas. The black dots represent the flux values estimated in Bastin (2007).

The evolution of the metabolite concentrations is depicted in Figure 2. The results show how it is possible to estimate non-measured metabolites over time, such as CO2. Notice that measured metabolites can also be estimated and improved, because considering them all together may reduce their uncertainty and correct errors if they are not completely consistent (although this was not the case in the example).

The estimated fluxes are depicted in Figure 3. This includes uptake or production rates for the extracellular metabolites (e.g., v_{25} or v_{26}) and internal fluxes (e.g., v_4 or v_{12}). It can be observed that some fluxes are estimated with precision (v_5 or v_7), whereas other estimates are wider (v_8 or v_{12}). Notice, however, that even the wider ones can be valuable: for instance, the estimates of v_{12} are wide, but they clearly indicate that reaction 12 is always active during exponential growth (0-120 h). For comparison purpose we consider the estimates in (Bastin, 2007) for all the internal fluxes (v_1 to v_{24}) at the growth phase (around time 72h). These are shown at Figure 3. The results are clearly similar to those obtained by us, although his intervals are narrower because (a) they assume that all reactions are irreversible during the growth phase, and (b) they do not consider uncertainty in the measurements (see, for example, the difference in v_1, the glucose uptake). Another possible flux estimation from the same data is given in (Provost, 2006b), which seems also compatible with our results and those of Bastin, even if the model used by Provost has some slight modifications.

Finally, we repeated the estimation using the «finite-window» approach described in section 3.1. It can be observed in figure 3 that the estimates are wider, as expected, although in this particular example the oversize is reasonably small. The oversize of the estimated areas (one per variable and degree of possibility) compared with those obtained with a «whole-window» approach is on average a 3.2% for fluxes (ranging from 0% to 12.4%) and a 4.3% for metabolites (ranging from 0% to 7.8%).

4.2 Dynamic Poss-FBA: predicting *E. coli* batch growth

Preparation. Finally we will illustrate the kind of results provided by dynamic Poss-FBA with an example of diauxic growth of *E. coli* on glucose and acetate. The example has been taken from Mahadevan *et al*. (2002). In that work the authors selected four pathways from a genome-scale reconstruction of *E. coli* and used them to define a simplified metabolic network with three extracellular metabolites, glucose (G), acetate (A) and oxygen (O), and biomass (*x*).

The four macro-reactions are the following:

$$
\begin{aligned}
v_1 : \quad & 39.43\ A + 35\ O2 \;\rightarrow\; x \\
v_2 : \quad & 9.46\ G + 12.92\ O2 \;\rightarrow\; x \\
v_3 : \quad & 9.84\ G + 12.73\ O2 \;\rightarrow\; 1.24\ A + x \\
v_4 : \quad & 19.23\ G \;\rightarrow\; 12.12\ A + x
\end{aligned}
\tag{23}
$$
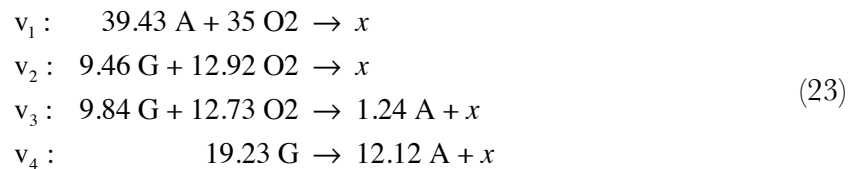
A constraint-based model accounting for these metabolites and biomass can be defined with the constraints $\mathcal{MOC}(k)$ and $\mathcal{CAC}(k)$. Let us consider a batch of 10 hours, divided in 21 intervals, so $k = [1, 2,..., 21]$. The reaction rates for each macro-reaction at each time instant

are represented with the vector $\mathbf{v}(k)=(v_1(k), v_2(k), v_3(k), v_4(k))^{\mathrm{T}}$. The first constraints in $\mathcal{MOC}(k)$ are the mass balances around extracellular species, which are the following:

$$\frac{G(k)-G(k-1)}{\Delta T}=\left(\begin{array}{cccc} 0 & -9.46 & -9.84 & -19.23 \end{array}\right)\cdot\mathbf{v}(k)$$

$$\frac{A(k)-A(k-1)}{\Delta T}=\left(\begin{array}{cccc} -39.43 & 0 & 1.24 & 12.12 \end{array}\right)\cdot\mathbf{v}(k)$$

$$\frac{O2(k)-O2(k-1)}{\Delta T}=\left(\begin{array}{cccc} -35 & -12.92 & -12.73 & 0 \end{array}\right)\cdot\mathbf{v}(k)+k_L a\big(0.21-O2(k-1)\big)$$

$$\frac{x(k)-x(k-1)}{\Delta T}=\left(\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array}\right)\cdot\mathbf{v}(k)$$

(24)

where G, A and O denote the metabolite concentrations (mM), and x the biomass concentration (g/L). The mass transfer coefficient for oxygen ($k_L a$) is 7.5 h⁻¹ accordingly to (Edwards *et al.*, 2001). The oxygen concentration in the gas phase is assumed to be a constant and equal to 0.21 mM.

Constraints are added to define the fluxes as irreversible and impose a positiveness condition to the metabolite concentrations:

$$\mathbf{D}\cdot\mathbf{v}(k)\geq 0$$

(25a)

$$\mathbf{e}(k)\geq 0,\ \mathrm{x}(k)\geq 0$$

(25b)

Then we define constraints $\mathcal{CAC}(k)$ to bind the glucose uptake with measured values of glucose concentration—notice that similar results were obtained when the uptake was modelled with Michaelis-Menten kinetics:

$$\frac{G(k)-G(k-1)}{\Delta T}=\frac{G_{\mathrm{m}}(k)-G_{\mathrm{m}}(k-1)}{\Delta T}$$

(26)

At this point the constraints operating at each time instant have been defined, $\mathcal{MOC}(k)$ constraints are defined with (24) and (25), and $\mathcal{CAC}(k)$ constraints with (26). The last step is to define the optimality criteria. According to Mahadevan et al. (2002), we assume that the cells' objective is maximization of growth and define the objective function as follows:

$$Z(k)=\left(\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array}\right)\cdot\mathbf{v}(k)$$

(27)

In order to analyze the degree of (growth) sub-optimality exhibited by the modelled cells, the parameter in (20) is defined as $\alpha_s = -2\cdot\log(0.5)$, so that possibility is $\pi=0.5$ when the biomass growth is 50% of the maximum feasible at time k. We will compute interval predictions for
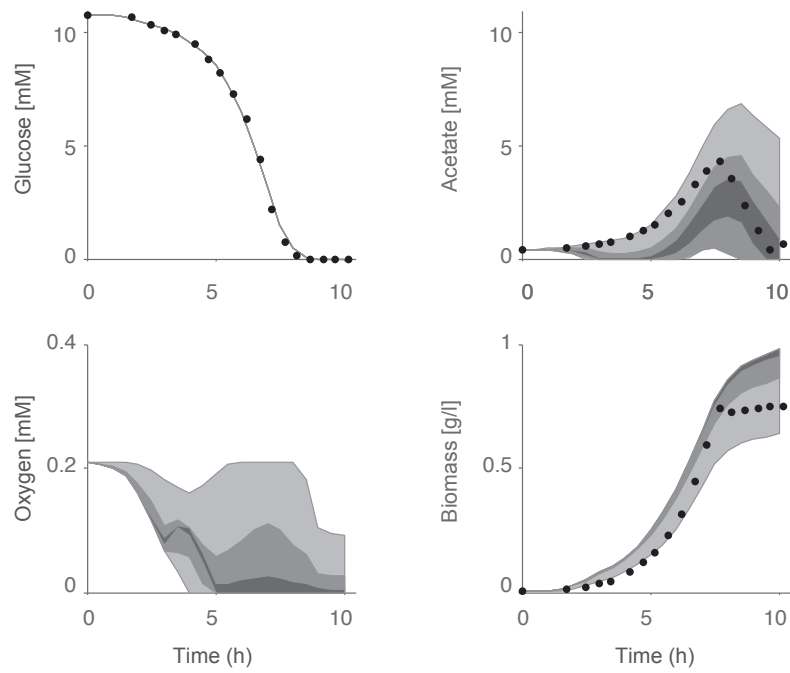
Figure 4. Measured and predicted metabolite concentrations during a cultivation of *E. coli*. Measurements are denoted with black dots. The concentrations estimated with Poss-FBA for three degrees of optimality (π=0.95, 0.8, 0.5) are denoted with grey areas. Recall that Poss-FBA only uses glucose measurements to perform the estimation.
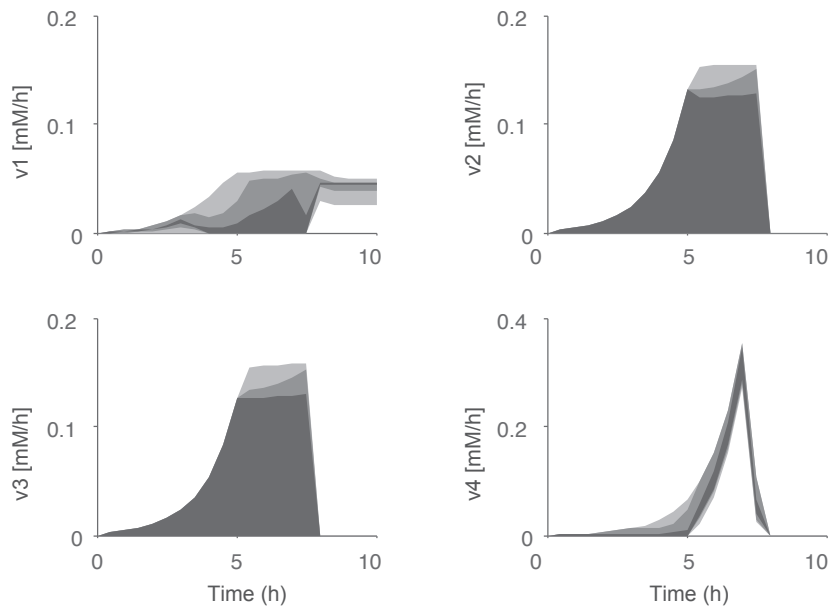
Figure 5. Estimated fluxes during a cultivation of *E. coli*. The fluxes estimated with Poss-FBA for three degrees of optimality (π=0.95, 0.8, 0.5), are denoted with grey areas.

possibilities {π=0.95, 0.8, 0.5}, thus capturing "practically" alternate-optimal trends, near optimal, and half optimal ones.

Dynamic Poss-FBA predictions. Finally we applied the two-step procedure described above (21-22) to get dynamic predictions for all the variables (fluxes and metabolites) and for three degrees of optimality (π=0.95, 0.8, 0.5). The results are depicted in Figures 4 and 5.

It can be observed that Poss-FBA detects alternative optima. It provides alternative predictions for v_2 and v_3 (Figure 5) even if only a slight sub-optimality is allowed (π=0.95). This is indeed sensible because these pathways have similar yields (i.e., they are nearly "exchangeable" in terms of biomass growth). This could indicate that both pathways can be efficiently used by the organism, or more likely, that the selection of one of them depends on phenomena not represented by the model (e.g., the choice depends on a secondary objective or its regulated by an environmental condition different from substrate availability). The results in Figure 4 also show that considering sub-optimality gives a richer prediction and better agreement with the actual concentrations. The cells behaviour can be reasonably captured with the simple model considered here, even if clearly the assumption of "maximisation of growth" is not perfect (during most of the batch the growth rate is only between 50% and 80% of the maximum).

The results show how considering sub-optimality and alternate optima provide an indication of the uncertainty of each prediction. For example, the assumption of "maximisation of growth" provides a more reliable prediction for biomass than for oxygen or acetate, for which wider ranges of values are reasonably possible. In summary, by grading optimality Poss-FBA indicates in a explicit way how dependent is each prediction on the underlying optimality assumption, which can be understood as a special type of model uncertainty.

5. Discussion

This paper has extended the possibilistic framework (Llaneras *et al.*, 2009) to get predictions from a constraint-based model accounting for extracellular dynamics. The two main approaches to get these predictions have been addressed: metabolic flux analysis (MFA), which incorporates experimental data to estimate the state of cells at given conditions, and flux balance analysis (FBA), which assumes evolved optimality of cells to predict their behaviour.

Features of Dynamic Possibilistic FBA. In the context of FBA we have shown that the possibilistic approach is a simple way to get rich predictions, both for fluxes and external metabolites. The novelty of dynamic Poss-FBA is that cells optimality assumption is incorporated in a relaxed or "softened" manner, so that the method gives not only a set of predicted values, but also an indication of how dependent each of these predictions is on the underlying assumption. Although there are other methods to do the same, Poss-FBA is a simple way to consider sub-optimality and capture alternate optima. Finally, by defining possibility as a "degree of belief", we provide a theoretical framework to investigate how to consider simultaneously different indicators of unfeasibility, such as sub-optimality, imprecision of measured variables, or uncertainty in models and parameters.

Features of Dynamic Possibilistic MFA. In the context of Poss-MFA it has been shown how to estimate time-varying fluxes and extracellular metabolite concentrations considering uncertainty and data scarcity. Although there are other methods accounting for extracellular dynamics (Provost et al. 2004; Herwig, 2002; Takiguchi, 1997; Henry, 2007), Dynamic Poss-MFA departs in three main ways.

First of all, most other methods are typically based on traditional MFA which point-wise estimates are unreliable if the estimation problem is not overdetermined, and thus require precise and sufficient measures that may be unavailable. Instead, Poss-MFA gives robust results even in scenarios of high uncertainty and lack of measurements, by computing distributions and interval estimates only as narrow as possible and considering reactions irreversibility. Poss-MFA achieves this following an approach similar to other interval MFA methods, such as the one used by Zamorano et al. (Zamorano et al., 2010), which exploit the notion of elementary modes to describe the possible fluxes and provide interval estimates, or the methods used by Ierapetritou and coworkers (Iver 2010a; 2010b) and by ourselves (Llaneras & Picó, 2007). In general, these interval estimates are more informative than point-wise ones in all instances when multiple flux values might be reasonably possible. In addition, by using possibility distributions, Poss-MFA considers measurements in a richer way and provide estimates for desired degrees of possibility. Dynamic Poss-MFA also inherits other benefits of the possibilistic framework, such as the ability to deal with non-gaussian errors or model imprecision, which have been illustrated before for static cases (Llaneras et al., 2009; 2010; 2011).

Another differential feature of our approach is that the measured extracellular-metabolite concentrations can be directly used as input, while other methods in the literature need a pre-step to calculate their average consumption/production rates (and their uncertainty, what can be particularly difficult) for each time instant by taking derivatives. Our intention was to keep the procedure simple and as purely "constraint-based" as possible, avoiding preliminary steps.

Another limitation of most dynamic MFA-wise methods is that the complete time-series of data are not analysed as a whole, but rather as unconnected sub-problems (one at each time instant). Our proposal takes into account, simultaneously, the constraints imposed by extracellular dynamics during the whole time-series, and therefore (a) it avoids the problem of how to propagate uncertainty along time without losing information, and (b) provides estimates that are more precise (narrower). This last advantage was illustrated in the first case study, where estimations with an unconnected «one-sized-finite-window» and a «whole-window» were compared.

Computational cost and large-scale. Regarding computation, one characteristic of the Possibilistic framework is that all the MFA and FBA estimations and predictions are cast as efficient (sparse) linear optimisation problems, for which widely-known tools exist. The method is therefore suitable for large-scale problems, with tens of thousands of decision variables (the number of them depends on sampling rate and time-window length). For convenience, dynamic Poss-MFA can be applied both with a «whole-window» or a «finite-window», thus fixing the size of the LP problems to be solved, so the users can select the trade-off between computational cost and narrower estimates.

Besides computational cost, it should be noticed that Poss-MFA applied to a large-scale and even genome-scale models will face the same challenge that every MFA-wise method: there may be many different solutions compatible with the few available measurements. In this situation, our method, thanks to the use of interval estimates, will be at least able to detect all

the solutions that are equally possible (or "similarly" possible) and depict them by means of intervals.

6. Conclusion

In summary, we have shown that a possibilistic framework can be useful to get rich dynamic predictions from a constraint-based model, both exploiting available experimental data to apply an MFA-wise method, or incorporating a rational hypothesis of evolved optimality of cells to follow an FBA-wise approach.

Table 1. List of exchange metabolites in CHO cells case study.

G	Glucose	Substrates	Q	Glutamine	initial substrates
L	Lactate	extracell. product	A	Alanine	extracell. product
NH4	Ammonia	extracell. product	CO2	Carbon dioxide	extracell. product
Nuc	Nucleotides	intracell. product			

Table 2. List of internal metabolites in CHO cells case study.

G6P	Glucose-6-phosphate	Mal	Malate
F6P	Fructosa-6-phosphate	Oxa	Oxaloacetate
G3P	Glyceraldehyde-3-phosphate	Glu	Glutamate
DAP	Dihydroxy-acetone Phosphate	Asp	Aspartate
Pyr	Pyruvate	Ri5P	Ribose-5-Phosphate
ACO	Acetyl-coenzyme A	Ru5P	Ribulose-5-Phosphate
Cit	Citrate	X5P	Xylose-5-Phosphate
aKG	α-ketoglutarate	E4P	Eryt-4-Phosphate
Fum	Fumarate	CO2i	Carbon dioxide (intracellular node)

Table 3. Stoichiometric matrix for CHO cells.

Irrevers.	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
Reaction	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	vG	vL	vA	vNH$_4$	vQ	vCO$_2$	vBio
1 G6P	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
2 F6P	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
3 G3P	0	0	1	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
4 DAP	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5 Pyr	0	0	0	0	1	-1	-1	-1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6 ACO	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7 Cit	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8 aKG	0	0	0	0	0	1	0	0	1	-1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9 Fum	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10 Mal	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11 Oxa	0	0	0	0	0	0	0	0	-1	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12 Glu	0	0	0	0	0	0	-1	0	0	0	0	0	0	-1	-1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13 Asp	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14 RU5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0
15 RI5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-2	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0
16 X5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	-1	0	0	0	0	0	0	0	0	0
17 E4P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0
18 CO2i	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	0	0	-1	-1	1	0	0	0	0	0	0	0	0	0	0	0
19 NUC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,17
20 G	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
21 L	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
22 A	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
23 NH4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	-1	0	0	0

| 24 | Q | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 25 | CO2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 |

7. Acknowledgements

8. References

1. Llaneras F. and Picó J. (2008). Stoichiometric modelling of cell metabolism. *J Biosci Bioeng*, 105 (1), 1-11.

2. Palsson BO (2006). *Systems biology: properties of reconstructed networks*. New York, USA: Cambridge University Press New York.

3. Heijden RT, Romein B, Heijnen JJ, Hellinga C, Luyben KC (1994). Linear Constraint Relations in Biochemical Reaction Systems: II. Diagnosis and Estimation of Gross. *Biotechnol Bioeng*, 43 (1), 11-20.

4. Edwards JS, Covert M, Palsson B (2002). Metabolic modelling of microbes: the flux-balance approach. *Environmental Microbiology*, 4:133-140.

5. Herwig C, von Stockar U (2002). A small metabolic flux model to identify transient metabolic regulations in Saccharomyces cerevisiae. *Bioprocess and Biosystems Engineering*, 24:395-403.

6. Takiguchi N, Shimizu H, Shioya S (1997). An on-line physiological state recognition system for the lysine fermentation process based on a metabolic reaction model. *Biotechnol Bioeng*, 55:170-181.

7. Henry O, Kamen A, Perrier M (2007). Monitoring the physiological state of mammalian cell perfusion processes by on-line estimation of intracellular fluxes. *Journal of Process Control*, 17:241-251.

8. Mahadevan R, Edwards JS, Doyle FJ (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophysics Journal*, 83:1331–1340.

9. Hjersted JL, Henson MA (2009). Steady-state and dynamic flux balance analysis of ethanol production by *Saccharomyces cerevisiae*. *IET Systems Biology*, 3(3):167–79.

10. Llaneras F, Sala A, Picó J (2009). A possibilistic framework for constraint-based metabolic flux analysis. *BMC Systems Biology*, 3:79.

11. Stephanopoulos GN, Aristidou AA, Nielsen J (1998). *Metabolic Engineering: Principles and Methodologies*. San Diego, USA: Academic Press.

12. Provost A and Bastin G (2004). Dynamic metabolic modelling under the balanced growth condition. *Journal of Process Control* 14(7):717–728.

13. Haag J, Vande Wouwer A, Bogaerts P (2005). Systematic procedure for the reduction of complex biological reaction pathways and the generation of macroscopic equivalents. *Chemical Engineering Science*, 60:459–465.

14. Bastin G (2007). Quantitative analysis of metabolic networks and design of minimal bioreaction models. *International Conference in Honor of Claude Lobry*.

15. Dubois D, Fargier H, Prade H (1996). Possibility theory in constraint satisfaction problems: handling priority, preference and uncertainty. *Applied Intelligence*, 6(4):287–309.

16. Sala A, Albertos P (2001). Inference error minimisation: fuzzy modelling of ambiguous functions. *Fuzzy Sets and Systems*, 121(1):95–111.

17. Sala A (2008). Encoding fuzzy possibilistic diagnostics as a constrained optimisation problem. *Information Sciences*, 178:4246–4263.

18. Llaneras F (2011). Interval and possibilistic methods for constraint-based metabolic models. PhD Thesis, Universidad Politécnica de Valencia, Valencia.

19. Tortajada M, Llaneras F, Picó J (2010). Validation of a constraint-based model of *Pichia pastoris* growth under data scarcity. *BMC Systems Biology*, 4:115.

20. Provost A, Bastin G, Agathos SN, Schneider YJ (2006). Metabolic design of macroscopic bioreaction models: application to Chinese hamster ovary cells. *Bioprocess and biosystems engineering* (2006) vol. 29 (5-6) pp. 349-366.

21. Provost A (2006b). Metabolic design of dynamic bioreaction models. PhD Thesis, Université catholique de Louvain, Louvain-la-neuve.

22. Zamorano F, Vande Wouwer A, and Bastin G (2010). A detailed metabolic flux analysis of an underdetermined network of CHO cells. *Journal of Biotechnology*, 150 (2010) 497-508.

23. Iyer V, Yang H, Ierapetritou M, Roth C (2010). Effects of glucose and insulin on HepG2-C3A cell metabolism. *Biotechnol Bioeng*,107: 347–356.

24. Iyer V, Ovacik MA, Androulakis IP, Roth CM, Ierapetritou MG (2010b). Transcriptional and metabolic flux profiling of triadimefon effects on cultured hepatocytes. *Toxicol appl pharmacol*. 248(3):165-77

25. Llaneras F and Picó J (2007). An interval approach for dealing with flux distributions and elementary modes activity patterns. *J Theor Biol*, 246:290–308.