Exploration of the Effect of Input Data on the Modeling of Cellular Objective in Flux Balance Analysis (FBA)

Carlos Eduardo García Sánchez[1] and Rodrigo Gonzalo Torres Sáez[2]

[1] Universidad Industrial de Santander, Escuela de Ingeniería Química, Bucaramanga, Colombia
[2] Universidad Industrial de Santander, Escuela de Química, Bucaramanga, Colombia
carlos.garcia6@correo.uis.edu.co, rtorres@uis.edu.co

Abstract. Flux Balance Analysis (FBA) is a technique that allows estimation of metabolic fluxes in established conditions, focusing the flux determination as an optimization problem. For this reason, it is important to use an appropiate objective function in order to adjust estimations of FBA with the real behaviour of the cell. The aim of this work was to examine the effect of a set of input data fluxes (among five diferent sets) on the accuracy of predictions obtained with FBA with application in seven different objective functions. In this study, ***Saccharomyces cerevisiae*** was selected as model microorganism, and its metabolism was represented at genomic scale using the model iMM904. Accuracy of obtained predictions was evaluated and compared with eight set of experimental data. Results showed that the objective function representing in a better way the cellular behaviour depends on the set of fluxes used as input data.

Keywords: Flux Balance Analysis, input data, objective functions, estimation accuracy.

1 Introduction

Flux Balance Analysis (FBA) is a technique of mathematical modeling, which allows determining a distribution of fluxes in a steady state through maximization of a supposed objective function [1]. Obtaining of a suited estimation by using of FBA requires a good metabolic model, a set of adjusted restrictions to real conditions and an objective function that generates realistic results. In other words, a mathematical function should model a cellular objective [1-5]. It also requires a set of input data that allows a feasible region for the model data [6]. Nevertheless, both quality of the obtained estimation and more appropriate objective function can depend on the number of fluxes used as input data in the FBA and chosen fluxes.

Different set of input data can result interesting for getting distinct objective in the cellular modeling. If it is wanted to analyze a cell as a biological system, the set of input data will consist probably of all uptake metabolic fluxes. On the other hand, if it is needed to predict production of secondary metabolites in a continuous culture, maybe it is necessary to add the specific growth rate to the input data, because this

L.F. Castillo et al. (eds.), *Advances in Computational Biology*,
Advances in Intelligent Systems and Computing 232,
DOI: 10.1007/978-3-319-01568-2_8, © Springer International Publishing Switzerland 2014

will be controlled by both mode of design and operation of the chemostat, and so on. From this way, it is necessary to broaden the searching for appropriate cellular objectives to conditions with different set of input data.

The aim of the present study was to explore the possible relationship between the set of selected data as input data in the application of FBA and the best objective function in this technique (that is, the one that allows lower errors in the estimation of exchange fluxes and specific growth rate). Therefore, based on sets of known experimental data, we compared the performance of different objective functions in the FBA varying input data and using the rest of measured fluxes for the evaluation of the quality of the estimations.

2 Data and Methods

We decided to use the yeast *Saccharomyces cerevisiae* as cell model, because is one of the most important microorganism for biotechnological applications. For the cell modeling, we selected the iMM904 model [7], which is a model at genomic scale and that includes 1228 metabolites in 1577 reactions with eight cell compartments.

As experimental data for carrying out evaluation of the performance of objective functions, we selected two dataset from continuous cultures. The first dataset was obtained from a number of aerobic experiments with $\mu = 0{,}1\ h^{-1}$ but with different oxygen uptake rates [8], and the second one was taken from anaerobic culture with four distinct specific growth rates [9]. These two dataset do not represent all different environmental conditions and growth that can be found. However, they attempt a range of conditions broad enough to obtain some conclusions from them, considering the objectives of the present study.

Data of aerobic growth contain measurements of both glucose and oxygen uptakes, production of ethanol, glycerol and carbon dioxide, and growth rate. The experimental configurations differ in the O_2 inlet percentages to the cell culture, being 0,5%, 1,0%, 2,8% y 20,9% of saturation of O_2 [8]. On the other hand, anaerobic experiments include measurements of substrate and oxygen uptake, and production of ethanol, glycerol, acetate, succinate, private and carbon dioxide, and specific growth rate (μ). The experiments were carried out with different μ (0,1, 0,2, 0,3 and 0,4 h^{-1}) [9].

Experimental data were classified in two categories: (i) the values used as equality restrictions in the FBA optimization problem were named 'input data', and (ii) the rest of fluxes were used in order to compare predictions obtained by FBA with experimental data, and named 'output data'. Different categories of input data were defined in order to determine if distinct initial conditions cause differences in the objective functions necessary to define cellular modeling. The categories were the following: uptake, uptake and μ, uptake and production, production and μ, and production. Each category makes reference to those fluxes that were taken as input data.

As objective functions were taken a set of representations of cellular objective that have been proposed in distinct previous studies on comparison of performance and generation of objective functions in FBA [3-4], [10-11]. The selected functions were

(I) maximization of biomass production, (II) minimization of glucose uptake, (III) minimization of NADH and NADPH in cytosol, (IV) maximization of biomass production plus minimization of carbon dioxide production, (V) maximization of biomass production plus minimization of NADH production in cytosol, (VI) maximization of biomass production plus minimization of NADH production in cytosol plus minimization of NADH and NADPH uptake in mitochondria, and (VII) minimization of ATP consumption in cytosol plus maximization of ATP transport from mitochondria to cytosol. By shortness, in Fig. 1 y Fig. 2 is kept roman numeration shown for objective functions.

The accuracy of estimations was evaluated with the error percentage on the estimation of the specific growth rate, and named Biomass error (see equation (1)), while the error percentage on the estimation of exchange fluxes that were experimentally measured, but not used as input data, were named Fluxes error (see equation (2)).

$$\text{Biomass error} = \left(\frac{\mu_{exp}-\mu_{est}}{\mu_{exp}}\right) * 100 \tag{1}$$

$$\text{Fluxes error} = \left(\frac{\|\vec{v}_{exp}-\vec{v}_{est}\|}{\|\vec{v}_{exp}\|}\right) * 100 \tag{2}$$

In the equations (1) and (2), the subscript "est" represents the estimations, and the subscript "exp" represents experimental data. In some categories of input data it was no possible to compute both errors, because of use of μ as input data, or of all exchange fluxes. Similarly, some objective functions do not appear in some Figures; this happens when some objective cannot be applied to some category of input data (for instance, maximization of biomass production cannot be used if μ is among the input data). It is important to note that sign Biomass error indicates if FBA overestimated (negative sign) or underestimate (positive sign) the specific growth rate of the cell.

3 Results and Discussion

In Fig. 1 are presented results obtained by simulation of aerobic experiments. For aerobic continuous cultures at low growth rates, it was found that among compared functions, maximization of biomass (function I) presents the lower errors using as input data the set 'uptake', while 'minimization of glucose uptake' (function II) showed better results using as input data 'production' and 'production and μ'. The function VII was clearly the second better in the category 'production'. Strangely, maximization of biomass (function I) gave worst estimations if only metabolite excretion fluxes are used as input data (that is, category 'production'). In the remaining category, 'uptake and μ', function VI showed the best performance.

The results for the anaerobic experiments are showed in Fig. 2. Again, maximization of biomass production (function I) is one of the best objective functions in the categories 'uptake' and 'uptake and production'; however, in the first category the function V gave slightly better estimations, and in the second with the functions IV

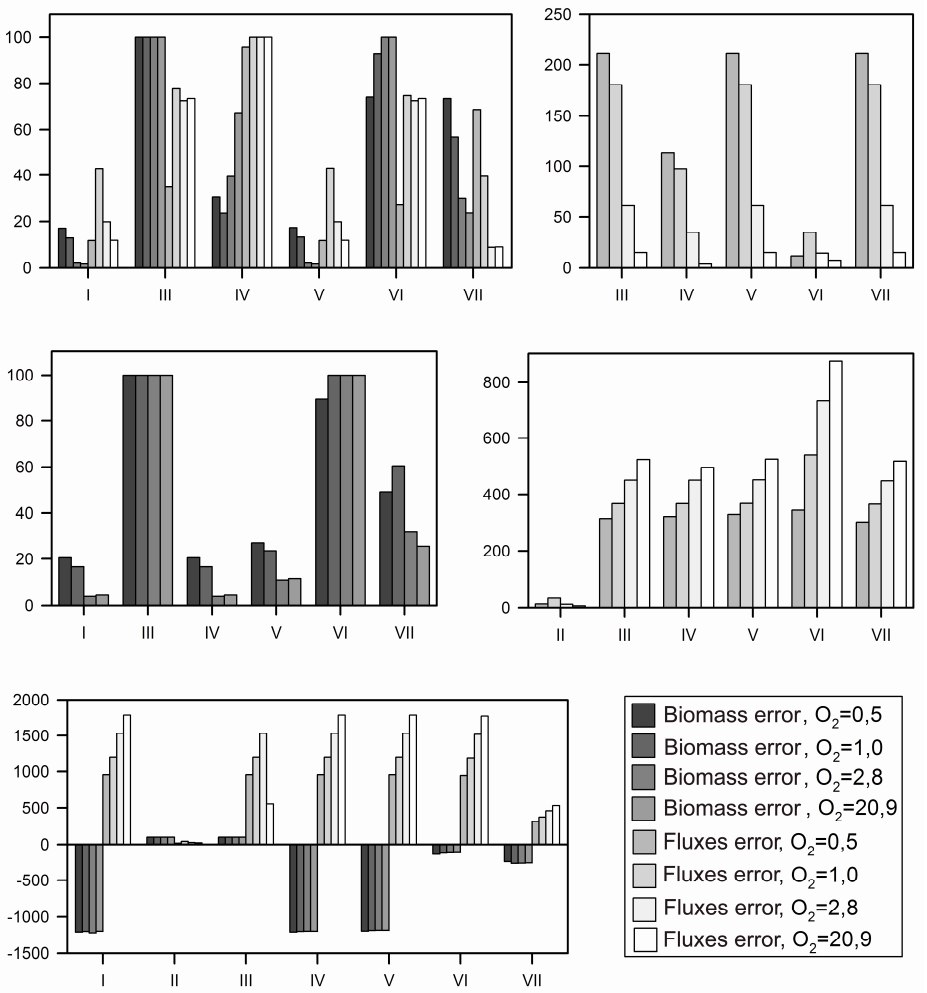

Fig. 1. Error percentages of the estimations obtained with FBA in aerobic growth conditions. Roman numbers represent seven different objective functions tested (See Data and Methods). The different panels show the obtained errors using the following input data: (**A**) uptake fluxes, (**B**) specific growth rate and uptake fluxes, (**C**) uptake and production fluxes, (**D**) specific growth rate and production, and (**E**) production fluxes.

and VII were obtained similar errors. In this case, with the function VII was obtained the best estimation for the category 'uptake and μ'. In the categories 'production' and 'production and μ', again the best function was the minimization of the glucose uptake (function II). In both situations, the function VII reached the second place on the quality of estimations.

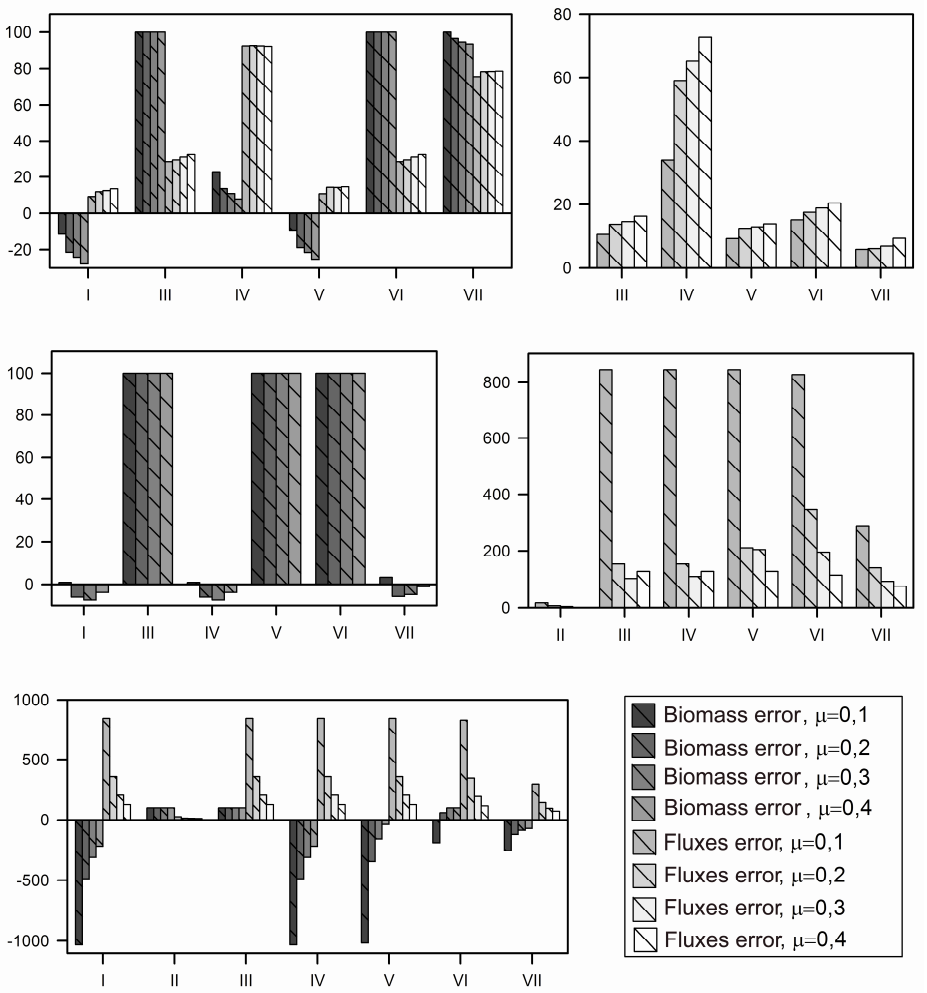

Fig. 2. Error percentages of the estimations obtained with FBA in anaerobic growth conditions. Roman numbers represent seven different objective functions tested (See Data and Methods). The different panels show the obtained errors using the following input data: (**A**) uptake fluxes, (**B**) specific growth rate and uptake fluxes, (**C**) uptake and production fluxes, (**D**) specific growth rate and production, and (**E**) production fluxes.

4 Conclusions

The results of this study indicate that the best objective function for representing the cellular behavior in FBA depends on the set of fluxes used as input data. Therefore, depending on the objective of the application of FBA (that is, according to the experimental condition to be represented, or the objective to be reached with the modeling) the objective function can be changed, being important to carry out a more detailed study of this problem, including more experimental data and different cell

conditions (among them, experiments of growth cell under batch conditions focusing on exponential phase), and using of a wide spectrum of different objective functions proposed for the analysis of cell modeling.

References

1. Raman, K., Chandra, N.: Flux Balance Analysis of Biological Systems: Applications and Challenges. Briefings in Bioinf. 10(4), 435–449 (2009)
2. Lee, J.M., Gianchandani, E.P., Papin, J.A.: Flux Balance Analysis in the Era of Metabolomics. Briefings in Bioinf. 7(2), 140–150 (2006)
3. Knorr, A.L., Jain, R., Srivastava, R.: Bayesian-based Selection of Metabolic Objective Functions. Bioinf. 23(3), 351–357 (2007)
4. Schuetz, R., Kuepfer, L., Sauer, U.: Systematic Evaluation of Objective Functions for Predicting Intracellular Fluxes in *Escherichia coli*. Mol. Syst. Biol. 3, 119 (2007)
5. Llaneras, F., Picó, J.: Stoichiometric Modelling of Cell Metabolism. J. of Bioscience and Bioeng. 105, 1–11 (2008)
6. Schilling, C.H., Schuster, S., Palsson, B.Ø., Heinrich, R.: Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-Genomic Era. Biotechnol. Prog. 15, 296–303 (1999)
7. Mo, M.L., Palsson, B.Ø., Herrgård, M.J.: Connecting Extracellular Measurements to Intracellular Flux States in Yeast. BMC Syst. Biol. 3, 37 (2009)
8. Wiebe, M.G., Rintala, E., Tamminen, A., et al.: Central Carbon Metabolism of *Saccharomyces cerevisiae* in Anaerobic, Oxygen-limited and Fully Aerobic Steady-state Conditions and Following a Shift to Anaerobic Conditions. FEMS Yeast Res. 8, 140–154 (2008)
9. Nissen, T., Schulze, U., Nielsen, J., Villadsen, J.: Flux Distributions in Anaerobic, Glucose-limited Continuous Cultures of *Saccharomyces cerevisiae*. Microbiology 143, 203–218 (1997)
10. Gianchandani, E.P., Oberhardt, M.A., Burgard, A.P., Maranas, C.D., Papin, J.A.: Predicting Biological System Objectives *de novo* from Internal State Measurements. BMC Bioinformatics 9, 43 (2008)
11. García Sánchez, C.E., Vargas García, C.A., Torres Sáez, R.G.: Predictive Potential of Flux Balance Analysis of *Saccharomyces cerevisiae* Using as Optimization Function Combinations of Cell Compartmental Objectives. PLoS ONE 7(8), e43006 (2012), doi:10.1371/journal.pone.0043006