Selection of Elementary Modes for Bioprocess Control

Zita I. T. A. Soons, Eugénio C. Ferreira, and Isabel Rocha

IBB – Institute for Biotechnology and Bioengineering - Centre of Biological Engineering, University of Minho,
Campus de Gualtar, 4710-057 Braga, Portugal, e-mail:{zita, irocha, ecferreira}@deb.uminho.pt

Abstract: The lack of model-based information in bioreactor monitoring and control can have a profound impact on biological systems. We therefore aim to develop a model using elementary modes (EMs) that represents the observed phenotype in given environmental conditions suited for bioprocess control. Challenges in the model development were the high number of possible phenotypes of stoichiometric models and the high computational intensity.

Two methods were compared to reduce the number of EMs to match the observed cellular phenotype. The first method is based on ranking modes and the second is a controlled random search (CRS) algorithm. Since we wish to obtain a biologically realistic subset of EMs, the objective function to be minimized is a trade-off between the error, efficiency of the modes, and model size.

The case study considered the central carbon metabolism of *Escherichia coli*. The original model containing 2706 modes for case 1 and 11718 for case 2 was reduced to a system of one for case 1 and three modes for case 2 giving a good correlation with the measured data. Furthermore, considering also intracellular besides extracellular metabolites, results in a better fit of the measured rates. Finally, the CRS outperformed the ranking algorithm.

Keywords: Elementary modes, model reduction, metabolism, *Escherichia coli*, biomass growth.

1. INTRODUCTION

Most mathematical models used for optimization and control of biotechnological processes are relatively simple and the complex interactions between the extracellular environment and the thousands of intracellular enzymes and metabolites are generally ignored. The lack of this information in bioreactor monitoring and control can have a profound impact on biological systems and lead to poor bioreactor control performance.

Nevertheless, the use of methods based on large models in process monitoring and control is nowadays limited due to their complexity and the lack of appropriate methodologies. The challenge of the development of a large-scale modelling strategy that predicts cellular phenotypes is not yet solved and will be addressed here in the view of bioprocess control.

Genome-scale stoichiometric models are currently the best approximation to a representation of the metabolic capabilities of the cell. However, stoichiometric models represent an infinite number of possible phenotypes and systems biology tools need to be applied such that the simulation matches the phenotypes in given conditions. Also, most tools in systems biology are designed for steady-state applications, whereas the aim of process control requires a dynamic approach. Although dynamics are not addressed explicitly in this work, the model is formulated such that it can be easily extended as such. Moreover, as a consequence of the complexity of the models, the computational intensity is high. The model simulations are too slow for some applications, such as online monitoring and control. Several model reduction approaches can be used to simplify models for use in process control, like the use of lumped reactions,

sensitivity analysis tools (Smets et al., 2002), singular perturbation theory (Bastin and Dochain 1990), and elimination of the dynamics of some processes based on their time scales (Haag et al., 2005).

A tool that has the potential to solve some of the above problems is Elementary Modes (EMs) (Schuster et al., 1999). EMs analysis identifies all minimal functional pathways inherent to a metabolic network. However, EMs analysis for large metabolic networks has the problem of combinatorial explosion of possible routes across the networks. In many situations, more EMs exist than necessary to construct all admissible flux distributions. Therefore, some of them can be taken as a generator set of the whole admissible region. A challenging task is how to select these EMs to describe a physiological state of interest. In literature, several approaches are described. Trinh et al. (2009) give an overview of this problem.

In the last years, several approaches that combine the use of EMs with experimental data have been used to predict cellular phenotypes and maximum production capabilities. Provost and Bastin (2004) achieved a model reduction by deriving a dynamic model based on EMs. The model is based on the elimination of intracellular rates to obtain a macroscopic model connecting substrates and products. The dynamic model, compatible with the underlying metabolic network, is built on these macro-reactions. The basic assumption is that the main dynamics are contained in the extracellular metabolites and that the intracellular metabolites are in steady-state. This approach of combining EMs with experimental data is the basis for the model in this work, subject to several modifications.

In modelling metabolic systems through EMs, usually some metabolites are considered "external" in the sense that they are available for uptake or can be secreted from the cell. Those metabolites are the sources of the network and their concentrations are assumed to be buffered. Internal metabolites have to be balanced with respect to production and consumption at steady state. In many cases, there are biochemical reasons to treat a metabolite as balanced or unbalanced (based, for example on known membrane transporters). Often, however, the classification is ambiguous, since the buffered condition can also be assumed for metabolites that are not secreted. Given that the higher the number of external metabolites the larger the number of EMs, Dandekar et al. (2003) propose a classification method of metabolites as external or internal that minimizes the number of EMs.

Our ultimate goal with a view on bioprocess control is, besides minimizing the number of EMs, capturing the essential process dynamics. In this work therefore, we propose to classify a metabolite as unbalanced or balanced based on time-scale separation, rather than on the physical presence of the metabolite, since the various reactions operate on different time scales, from milliseconds to hours or days. Hence, we classify the metabolites based on the time scale of interest: the slow metabolites are treated as unbalanced, the fast metabolites as balanced. In order to compare this methodology with the standard classification, we analyse two cases: in case 1 the distinction between extracellular (unbalanced) and intracellular (balanced) metabolites is the standard one while in case 2 it is based on time-scale separation. Time-scale separation in metabolic pathways has been studied by, amongst others, Delgado and Liao (1995).

An issue not addressed in most literature on the calculation of the elementary rates, is that of reversibility of EMs. It will be addressed in more detail in the methods section 2.1.

In the sequel, two methods will be compared to select a limited number of EMs matching the phenotype in given conditions. The first method is based on ranking modes and the second on a controlled random search (CRS) algorithm.

2. MODEL

2.1 Stoichiometric model and experimental data

The network model was reconstructed to represent *Escherichia coli* growing on glucose minimal media. The starting point this model was the reduced model for the central carbon metabolism of *E. coli* of Suthers et al. (2007). This model was modified in the following way: amino-acid metabolism is not included; biomass formation was modelled by acknowledging the metabolic drain from the central metabolic pathways according to Fischer et al. (2004); and oxidative phosphorylation was lumped and modelled according to Carlson and Srienc, (2004). The resulting model contains glycolysis, pentose phosphate pathway, TCA cycle, anaplerotic reactions, biomass formation, oxidative phosphorylation, maintenance energy, and membrane transport reactions. Energy requirements for biomass formation and energy production are also included. The model contains 45 metabolites and 48 reactions of which 23

are irreversible. The number of degrees of freedom is 10 (t there are linearly dependent balance equations).

The steady-state metabolite concentrations during continuous cultures (extracellular glucose (GlcEX), biomass, and acetate (AceEX); and intracellular glucose-6-phosphate (G6P), pyruvate (PYR), and 6-phosphogluconate (6PG)) were taken from Hoque et al. (2005); oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) were taken from Hua et al. (2003) and adjusted for the different biomass concentration and yield.

2.2 Model based on elementary modes

The classical dynamical model of a bioreactor can be established on the basis of a set of macro-reactions (Bastin and Dochain, 1990):

$$\frac{d\xi}{dt} = Kr(t) + u(t) \tag{1}$$

where ξ is the vector with concentrations of the chosen metabolites in the reactor liquid volume, $r(t)$ the vector of reaction rates, and $u(t)$ the net exchange of the metabolites with the outside. The stoichiometric matrix K of the macro-EM reactions reads (Provost and Bastin, 2004):

$$K = N \cdot E \tag{2}$$

where E is the $n_r \times n_{EMs}$ elementary mode matrix, computed using METATOOL 5.1.0 (Pfeiffer et al., 1999). Each column of E represents one EM. N denotes the corresponding $n_\xi \times n_r$ matrix of the chosen metabolites. At steady state, there is no accumulation of metabolites in the system and (1) can be simplified to:

$$K \cdot r = -u \tag{3}$$

Any steady-state flux pattern can be expressed as a nonnegative linear combination of these modes (section 3). In underdetermined or overdetermined systems the matrix K is not-invertible. Some authors applied the Moore-Penrose inverse to calculate the elementary rates (Poolman et al., 2004). An issue, not addressed in some literature on the calculation of the elementary rates, is that of reversibility of EMs. An EM is considered reversible if all its reactions are reversible. Conversely, EMs containing one or more irreversible reactions are irreversible. As a consequence the elementary rates of the irreversible modes should be greater than or equal to zero:

$$r_i \geq 0 \text{ if } K_i \text{ is irreversible} \tag{4}$$

Poolman et al. (2004) tackled this by simply removing the columns of K that lead to negative rates in irreversible modes and recalculating the assignment. Schwartz and Kanehisa (2005) took into account the reversibility constraints in a quadratic programming problem to calculate the elementary rates. Another method, based in linear programming, is the concept of the α-spectrum (Wiback et al., 2003). The α-spectrum encloses all possible solutions, but is not intended to find a reduced set of modes. In this work we compute the

rates r using a nonnegative least squares algorithm (Lawson and Hanson, 1974).

2.3 Two case studies

The proposed methods are illustrated on two cases. In case 1 the distinction between balanced and unbalanced metabolites in the EM computation is based on the standard definition. In case 2 it is based on time-scale separation. The vectors of the unbalanced metabolites (ξ_1) are given below.

Case 1:
$$\xi_1^T = \left[C_{GlcEX}, C_X, C_{AceEX}, C_{O_2}, C_{CO_2}\right] \quad (5)$$
$$u_1^T = \left[D\left(C_{GlcEX}^{in} - C_{GlcEX}\right), -\mu, -\mu \cdot C_{AceEX}, -CTR, OTR\right]$$

Case 2:
$$\xi_2^T = \left[C_{G6P}, C_{PYR}, C_{6PG}, C_{GlcEX}, C_X, C_{AceEX}, C_{O_2}, C_{CO_2}\right] \quad (6)$$
$$u_2^T = \left[-\mu \cdot C_{G6P}, -\mu \cdot C_{PYR}, -\mu \cdot C_{6PG}, D\left(C_{GlcEX}^{in} - C_{GlcEX}\right), -\mu, -\mu \cdot C_{AceEX}, -CTR, OTR\right]$$

Where C_{GlcEX} is the extracellular glucose concentration, C_X the biomass concentration, C_{AceEX} the extracellular acetate concentration, C_{O2} the oxygen concentration in the liquid, C_{CO2} the carbon dioxide concentration, D the dilution rate, which is equal to the specific growth rate μ, C_{G6P} the intracellular glucose-6-phosphate concentration, C_{PYR} the intracellular pyruvate concentration, and C_{6PG} the intracellular 6-phosphogluconate concentration. Note that in this particular case of continuous culture both the internal and external metabolites are in steady-state. For bioprocess control, e.g. fed-batch cultures, external metabolites are not considered in steady-state.

3. REDUCTION OF THE NUMBER OF EMS

For case 1 computation of the EMs using METATOOL gives 2706 EMs, of which 1622 are biomass producing modes. A way to select the EMs that describe a particular phenotype with biomass production could be based on the biomass yield on glucose and oxygen according to Carlson and Srienc (2004). Song and Ramkrishna (2009) chose a fixed number of EMs based on yield analysis using quadratic programming. Based on Fig. 1 (showing part of the data for illustration), a single EM or a linear combination of EMs could be selected to match experimental data.
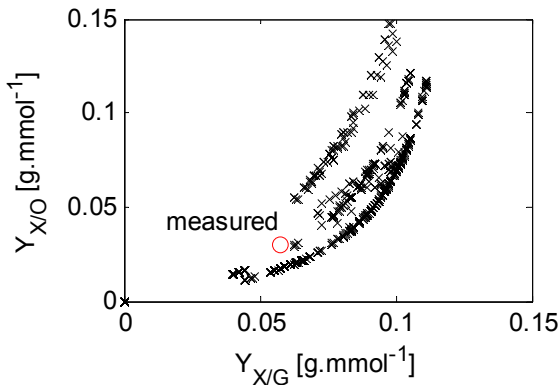

Fig. 1. Biomass yield on oxygen versus biomass yield on glucose for each EM (x). The circle (o) is the measured yield.

In this work, however, we wish to consider more metabolite measurement. Furthermore, we wish to consider all possible EMs (for instance also non-biomass producing modes), model size, and efficiency of the EMs. Hence, we reduce the number of EMs on the basis of an objective function that takes into account these aspects.

We aim to reduce the number of EMs such that we yield a realistic subset of K (corresponding to a subset of EMs) that matches the observed phenotype. In general, increasing the model size (or the number of selected EMs) is likely to improve the estimation errors. On the other hand, our assumption is that only a small number of EMs is active under defined process conditions. Besides, in our search for a biologically meaningful subset, we think that the more efficient EMs, in terms of investment in enzymes, are more likely to be active in practice. The proposed objective function therefore is a trade-off between the error $RMSE$, the average efficiency of the selected elementary modes P, and the model size MS:

$$J = RMSE + c_1 \cdot P + c_2 \cdot MS \quad (7)$$

where c_1 and c_2 are factors to weight the importance of efficiency and model size against the actual error. $RMSE$ denotes the weighted root mean squared error between the measured u_m and fitted rates u:

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum\left(\frac{u_m - u}{w}\right)^2} \quad (8)$$

where n is the number of measurements and w the weights, computed by the absolute value of u_m. Efficiency P is defined as the investment required to establish the EM, that is, to produce the enzymes and is calculated by the average sum of the absolute stoichiometry of the rates through the selected EMs (e) (modified from Stelling et al., 2002):

$$P = \frac{1}{MS} \cdot \sum|e_j| \quad (9)$$

where j denotes the reaction. The fluxes have to be normalized by the units of substrate consumed. Model size is defined by the number of nonzero components of the elementary rates r:

$$MS = \aleph(r) \quad (10)$$

We are now left with the problem of selecting the best subset of overall elementary modes K, which minimizes the objective function (7). Hereto we compare two algorithms: ranking (section 3.1) and CRS (section 3.2).

3.1 Ranking of elementary modes

The approach "ranking of EMs" is based on the idea of adding one column of K (representing the macro-reaction of one EM) to the current subset of K to give the largest improvement of the objective function until a minimal objective function is found. The algorithm uses this approach to expand the model, starting with a single term which minimizes the objective function.

3.2 Controlled random search algorithm

Price (1977) developed a CRS procedure, which searches for global minima in an iterative procedure. A drawback of the method is the computational time; nevertheless, it is more efficient than a pure random search. This procedure is applied here to select a limited number of EMs and their index from K on the basis of the objective function in an iterative search containing three steps:

1. First, a set of NT trial points is generated at random using n inputs (number of EMs and their column index) from the search domain V. The search domain V is defined by specifying limits on each of the n variables and a predetermined number of trial points NT. In this work, the constraints were set by a maximum of ten EMs and the maximum index defined by the number of calculated EMs. The objective function (7) is evaluated for each trial point and stored in array A together with the inputs.

2. Then the search starts by generating for each iteration a new trial point as follows: new points are generated by choosing $n+1$ random distinct points $RP_1, RP_2, \ldots RP_{n+1}$ from the set of NT stored points. The next trial point TP is computed from the centroid G of the n points $RP_1, \ldots RP_n$ minus the last point RP_{n+1}:

$$TP = 2 \cdot G - RP_{n+1} \qquad (11)$$

Provided that the new trial point TP satisfies the constraints the goal function is evaluated (J_{TP}).

3. The stored point M from matrix A with the greatest value is determined (J_M). J_{TP} is compared with the J_M in point M. If $J_{TP} < J_M$, M is replaced by TP in A.

Step two and three are repeated until the stop criterion is satisfied (all penalties J in the stored matrix A are identical, the maximum J is smaller than a certain value, or the maximum number of function evaluations is reached).

4. RESULTS AND DISCUSSION

The choice of the weighting factors c_1 and c_2 in (7) indicate the importance of efficiency and model size against the error and influences the selection of the number and index of the EMs. Since we intend to achieve a huge model reduction and to select biological relevant modes, we chose significant values for both ($c_1 = 2 \cdot 10^{-5}$, $c_2 = 1 \cdot 10^{-2}$). In section 4.1-4.3 we present the results for the specific cases. In section 4.4 we compare both methods.
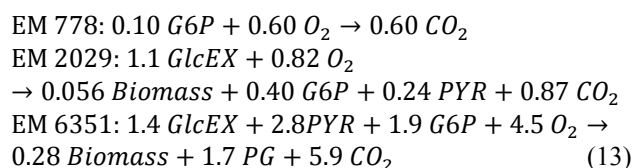
4.1 Ranking and CRS for case 1

Both algorithms reduced the original model containing 2706 EMs to a system based on one and the same mode for biomass growth. The corresponding macro-reaction that connects the extracellular substrates and the end-products is:

EM 267: $3.1\ GlcEX + 6.5\ O_2 \rightarrow 0.27\ Biomass + 7.3\ CO_2 + 0\ AceEX$ (12)
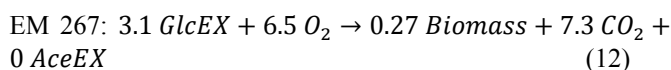
with an elementary rate of 0.50. Simulation of the reduced model gives an appropriate match with the data (Fig. 2A). Note that acetate formation is fitted to be zero in the reduced model. Application of the algorithms to a different data set, in which acetate is produced in significant amounts, would probably lead to the selection of different mode(s).

4.2 Ranking of EMs for case 2

The ranking algorithm reduced the original model containing 11718 EMs to a system based on three modes (Fig. 3). Since case 2 considers three additional metabolites as "unbalanced", the macro-reactions are now split into other macro-reactions containing the linkage with those metabolites as well (besides the linkage of the extracellular substrates with the extracellular products):

EM 778: $0.10\ G6P + 0.60\ O_2 \rightarrow 0.60\ CO_2$
EM 2029: $1.1\ GlcEX + 0.82\ O_2$
$\rightarrow 0.056\ Biomass + 0.40\ G6P + 0.24\ PYR + 0.87\ CO_2$
EM 6351: $1.4\ GlcEX + 2.8PYR + 1.9\ G6P + 4.5\ O_2 \rightarrow$
$0.28\ Biomass + 1.7\ PG + 5.9\ CO_2$ (13)

with elementary rates of 3.0; 1.5; and 0.057. Figure 3 shows that increasing the model size from one to three selected EMs improved the objective function. Adding more EMs did not improve the error sufficiently to compensate for the higher value of the objective function with the increased model size.
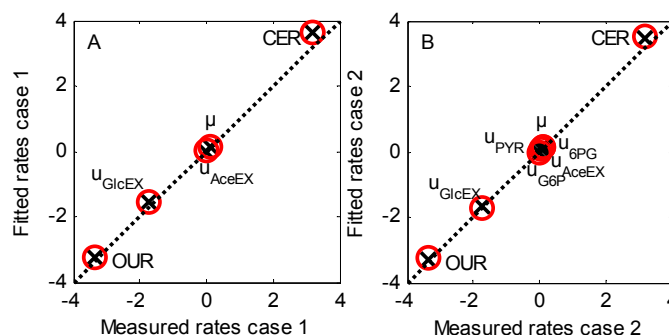

Fig. 2. Measured versus fitted rates for the reduced models using ranking (x) and CRS (O). A. case 1. B. case 2.
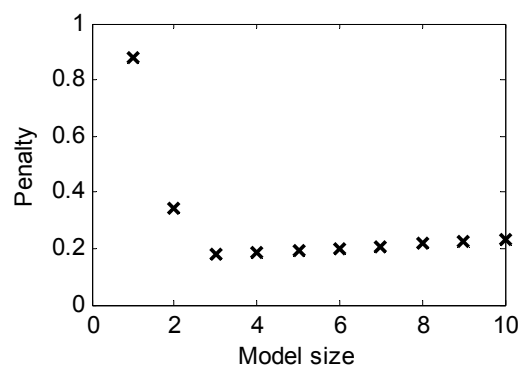

Fig. 3. Selection plot of the model size for case 2.

4.3 Controlled random search for case 2

The CRS procedure was applied several times and converged, during each run, to the same model size (3 EMs), but not to

the same set of EMs. All sets of EMs have close J (7), but not all the same. Convergence to global instead of local optima could be improved by increasing the number of trial points NT in the search (at the expense of increasing computational time) or by modifying the algorithm in different ways. As an example, Fig. 4 shows that the iterative search converges to a set of three EMs. The corresponding macro-reactions are the following, with elementary rates of 1.9; 0.059; and 0.41:
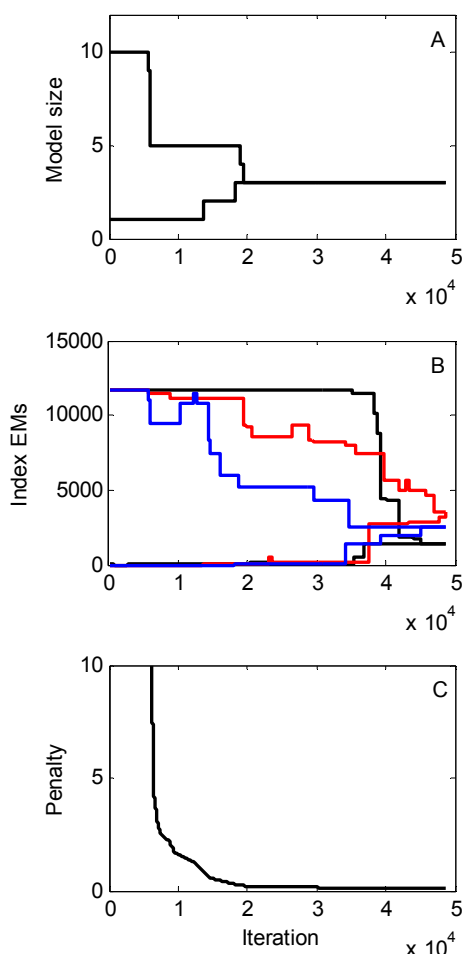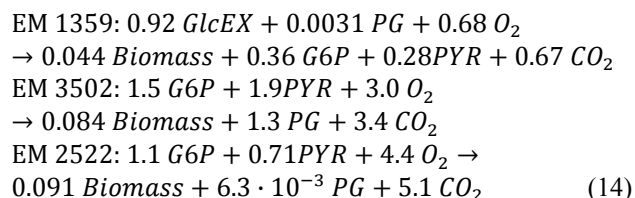
EM 1359: $0.92\ GlcEX + 0.0031\ PG + 0.68\ O_2$
$\rightarrow 0.044\ Biomass + 0.36\ G6P + 0.28 PYR + 0.67\ CO_2$
EM 3502: $1.5\ G6P + 1.9 PYR + 3.0\ O_2$
$\rightarrow 0.084\ Biomass + 1.3\ PG + 3.4\ CO_2$
EM 2522: $1.1\ G6P + 0.71 PYR + 4.4\ O_2 \rightarrow$
$0.091\ Biomass + 6.3 \cdot 10^{-3}\ PG + 5.1\ CO_2$ (14)

Fig. 4. Controlled random search to select a limited number of EMs for case 2. The lines represent the maximum and minimum values within the dataset A. A. Model size B. Index of the three selected EMs C. Penalty J (7).

4.4 Comparison of methods and discussion

An overview of the results from EMs reduction based on ranking EMs and CRS is shown in table 1 and Fig. 2. For the sake of a fair comparison of the results between cases 1 and 2, the penalty (7) was also computed using only the extracellular metabolites (J_{ext}). Note that J_{ext} was not used as a selection criterion and might therefore be slightly larger.

Extending the states vector of case 1 with a few intracellular metabolites resulted in an improved fit of the measured rates for the presented dataset. For some of the other tested datasets, the obtained fit was only slightly better using the extended states vector. We think that this result may be obtained due to the larger pool of EMs for case 2 and hence the higher possibility of finding EMs matching the data better: before reduction, case 2 contained 11718 EMs and case 1 2706.

The deteriorated results for case 1 in comparison with case 2 may suggest that the assumption behind case 1 was wrong, i.e., that the concentrations of G6P, PYR, and 6PG. are not balanced. However, since the differences for both cases were not that obvious on all datasets, more experiments are necessary to achieve more supported conclusions

Table 1. Overview of the results

	J	J_{ext}	*Model Size*
Ranking case 1	0.14	0.14	1
Ranking case 2	0.15	0.083	3
CRS case 1	0.14	0.14	1
CRS case 2	0.098	0.11	3

Table 1 furthermore shows that the CRS algorithm outperformed the ranking algorithm for a model size larger than one EM. If the optimal model size is one, both methods select the same EM, at the expense of a slightly larger computational time for the CRS. The ranking algorithm is based on expanding the model from the fixed "best" EM with the "next best" EM and so on. However, the set of EMs giving the best J does not necessarily include the best EM. So, the ranking algorithm may not find the best set, which the CRS is able to find. In other words, the ranking approach may easily fall in local minima. Conversely, the CRS may find the global minima. However, this is not guaranteed and the algorithm should be run several times to find the best solution.

This finding is in line with other authors. Crampin et al. (2004) already stated that the non-orthogonality of the matrix K means that the optimal subset of size $K+1$ is not necessarily the optimal subset K plus the "next best term" and that the selection process must therefore be iterative. Also Judd and Mees (1995) already stated: "It appears that finding the optimal model of size k is NP-hard - related to the feasible basis extension problem. If this is the case, we cannot expect to obtain the optimal solution easily."

We used data from one particular experiment (8 measured metabolites) to illustrate the results. Interestingly, the approach is flexible enough to be of use - at higher computational cost - if more datasets are simultaneously taken into account (e.g. different studies, at different conditions, etc.). In this way, the selection of EMs becomes less dependent on a single instance of measurements (more reliable) and the model will be able to represent a wider range of actual phenotypes. Furthermore, when inferring dynamics of the elementary modes ($r(t)$ in (1)) -to be able to model the time evolution of the metabolites ξ and ultimately control in real-time- informative datasets representing a variety of conditions are crucial for accurate estimation of the reaction

mechanisms and parameters of the elementary rates. A feature of the controlled random search algorithm (in contrast to the ranking algorithm) is that it can be easily extended with the simultaneous selection of reaction mechanisms.

As an alternative to elementary modes, the convex basis or a "minimal set of generating vectors" may be more suitable to represent a particular phenotype. The convex basis is a subset of the elementary modes and suffices to represent each feasible flux distribution in the network by a nonnegative linear combination of the convex basis vectors (Urbanczik, 2007; Song and Ramkrishna, 2009). Besides the computation of the EMs, METATOOL also facilitates the computation of the convex basis vectors. For instance, case 1 yields 2706 elementary modes as opposed to 433 generating vectors.

5. CONCLUSIONS

The development of methodologies to improve real-time process optimization, monitoring, and control based on large-scale metabolic models has the potential to raise process efficiency and productivity. This work is a first step towards the use of metabolic models in real-time by presenting a methodology to capture a large metabolic network by only a small number of elementary modes that are active under defined process conditions.

ACKNOWLEDGMENTS

REFERENCES

Bastin, G. and Dochain, D. (1990). *On-line estimation and adaptive control of bioreactors*, Elsevier, Amsterdam.

Carlson, R. and Srienc, F. (2004). Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Creation of overall flux states, *Biotechnol. Bioeng.*, 86, 149-162.

Crampin, E.J., McSharry, P.E., and Schnell, S. (2004). Extracting biochemical reaction kinetics from time series data, *Knowledge-Based Intelligent Information and Engineering Systems, Pt 2, Proceedings*, 3214, 329-336.

Dandekar, T., Moldenhauer, F., Bulik, S., Bertram, H., and Schuster, S. (2003). A method for classifying metabolites in topological pathway analyses based on minimization of pathway number, *Biosyst.*, 70, 255-270.

Delgado, J. and Liao, J. C. (1995). Control of metabolic pathways by time-scale separation, *Biosyst.*, 36, 55-70.

Fischer, E., Zamboni, N., and Sauer, U. (2004). High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived C-13 constraints, *Ana.l Biochem.*, 325, 308-316.

Haag, J.E., Wouwer, A.V., and Bogaerts, P. (2005). Systematic procedure for the reduction of complex biological reaction pathways and the generation of macroscopic equivalents, *Chem. Eng. Sc.*, 60, 459-465.

Hoque, M.A., Ushiyama, H., Tomita, M., and Shimizu, K. (2005). Dynamic responses of the intracellular metabolite concentrations of the wild type and pykA mutant *Escherichia coli* against pulse addition of glucose or NH3 under those limiting continuous cultures, *Biochem. Eng. J.*, 26, 38-49.

Hua, Q., Yang, C., Baba, T., Mori, H., and Shimizu, K. (2003). Responses of the central metabolism in *Escherichia coli* to phosphoglucose isomerase and glucose-6-phosphate dehydrogenase knockouts, *J. Bacteriology*, 185, 7053-7067.

Judd, K. and Mees, A. (1995). On Selecting Models for Nonlinear Time-Series, *Physica D*, 82, 426-444.

Lawson, C.L. and Hanson, R.J. (1974). *Solving least squares problems*, Prentice-Hall, Englewood Cliffs, New Jersey.

Pfeiffer, T., Sanchez-Valdenebro, I., Nuno, J. C., Montero, F., and Schuster, S. (1999). METATOOL: for studying metabolic networks, *Bioinformatics*, 15, 251-257.

Poolman, M.G., Venkatesh, K.V., Pidcock, M.K., and Fell, D.A. (2004). A method for the determination of flux in elementary modes, and its application to Lactobacillus rhamnosus, *Biotechnol. Bioeng.*, 88, 601-612.

Price, W.L. (1977). Controlled Random Search Procedure for Global Optimization, *Comp. J.*, 20, 367-370.

Provost, A. and Bastin, G. (2004). Dynamic metabolic modelling under the balanced growth condition, *J.Process Control*, 14, 717-728.

Schuster, S., Dandekar, T., and Fell, D.A. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering, *Trends in Biotechnology*, 17, 53-60.

Schwartz, J.M. and Kanehisa, M. (2005). A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes, *Bioinformatics*, 21, 204-205.

Smets, I., Bernaerts, K., Sun, J., Marchal, K., Vanderleyden, J., and Van Impe, J. F. (2002). Sensitivity function-based model reduction. A bacterial gene expression case study, *Biotechnol. Bioeng.*, 80, 195-200.

Song, H.S. and Ramkrishna, D. (2009). Reduction of a set of elementary modes using yield analysis, *Biotechnol. Bioeng.*, 102, 554-568.

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., and Gilles, E.D. (2002). Metabolic network structure determines key aspects of functionality and regulation, *Nature*, 420, 190-193.

Suthers, P.F., Burgard, A.P., Dasika, M.S., Nowroozi, F., Van Dien, S., Keasling, J.D., and Maranas, C.D. (2007). Metabolic flux elucidation for large-scale models using C-13 labeled isotopes, *Metabolic Eng.*, 9, 387-405.

Trinh, C.T., Wlaschin, A., and Srienc, F. (2009). Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism, *Applied Microbiol. Biotechnol.*, 81, 813-826.

Urbanczik, R. (2007). Enumerating constrained elementary flux vectors of metabolic networks, *IET Systems Biology*, 1, 274-279.

Wiback, S.J., Mahadevan, R., and Palsson, B.O. (2003). Reconstructing metabolic flux vectors from extreme pathways: defining the alpha-spectrum, *J. Theor. Biol*, 224, 313-324.