

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231390357>

When is the Quasi-Steady-State Approximation Admissible in Metabolic Modeling? When Admissible, What Models are Desirable?

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · AUGUST 2009

Impact Factor: 2.59 · DOI: 10.1021/ie900075f

CITATIONS

5

READS

21

2 AUTHORS:



[Hyun-Seob Song](#)

Pacific Northwest National Laboratory

55 PUBLICATIONS 481 CITATIONS

SEE PROFILE



[Doraiswami Ramkrishna](#)

Purdue University

287 PUBLICATIONS 6,668 CITATIONS

SEE PROFILE

When is the Quasi-Steady-State Approximation Admissible in Metabolic Modeling? When Admissible, What Models are Desirable?

Hyun-Seob Song and Doraiswami Ramkrishna*

School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47907

A strategy is presented for scrutiny of the quasi-steady-state (QSS) approximation in metabolic modeling that involves tracking dynamic data on extracellular products in so-called flux, as well as yield vector spaces when the substrate uptake flux is perturbed sinusoidally about some average value with specified frequency and amplitude. Criteria accrue for the diagnosis of any violation (and its extent) of the QSS approximation and are applied to data produced by simulations from models free from the QSS approximation. It is shown that, even when the assumption is admissible, the choice of the appropriate QSS model depends on the frequency and amplitude of the perturbation. Three different QSS models in the literature, viz., dynamic flux balance analysis, the macroscopic bioreaction model, and the hybrid cybernetic model, are evaluated, and circumstances are identified when one or more of the QSS models might be valid, as well as when none of them are.

1. Introduction

We are pleased to contribute to a festschrift for Mumbai University's Professor J. B. Joshi, whose forte has been to focus engineering fundamentals on understanding and improving engineering equipment. Although our article falls short of such a healthy goal, its claims to felicitating a colleague dedicated to engineering practice could hopefully be based on addressing an issue closely linked to model reduction for a quantitative study of bioreactors essential to the commercial exploitation of bioprocesses. Low-order metabolic models that might be facilitated by the study reported here could be attractive complements in the study of fermentation processes to the computational fluid dynamics of long interest to Professor Joshi.

Quantitative treatment of detailed metabolic processes has its roots in flux balance analysis (FBA), the edifice of which is the quasi-steady-state assumption (QSSA) of intracellular variables that quickly reach steady state with respect to varying external fluxes. The broad objective of this article is thus to examine the circumstances under which data obtained permit the application of the QSSA and whether suitable quasi-steady-state models (QSSMs) are available. The circumstances envisaged are those created by the effects of externally imposed dynamic changes in the substrate uptake flux on all extracellular variables, which can be examined in two alternative ways. One is to view the transient response in flux space (i.e., the space of all reaction rates), and the other is to consider the yield vector space pertaining to the yield of all extracellular products of metabolism per unit mass of substrate consumed. It can be readily shown that the QSSA will confine transient metabolism within a flux cone in flux space and within a finite convex hull in yield vector space. Thus, if the representation of transient metabolism in either of the foregoing spaces were to transgress their corresponding confines, a violation of the quasi-steady state (QSS) is implied. The extent to which such a transgression occurs will also provide a measure of how seriously the QSS is violated. Clearly, the risk of transgression is enhanced if the imposed temporal variation of the uptake flux is both substantial and rapid, as determined, for example, by the amplitude and frequency of a sinusoidal variation about some average. High amplitudes and frequencies would indeed endanger the QSSA,

and it is our objective to get a sense of the extent and speed of this change in the substrate uptake rate (characterized by a point in the amplitude–frequency plane).

It must now be apparent that, even if metabolic trajectories satisfy the QSSA by staying within the indicated confines, the existence of a satisfactory QSSM cannot be taken for granted. From the point of view of addressing this concern, we next survey the range of available QSSMs.

The measurement of external fluxes, enabled by measurements of substrate concentrations in the extracellular medium, together with stoichiometric coupling and optimization criteria such as maximization of biomass yield constitute the foundation of FBA. Also of interest are modeling approaches predicated on decomposing the metabolic network into so-called elementary modes (EMs), which are essentially a (teeming) collection of various pathway alternatives for converting external substrate into biomass and metabolic products.¹ Within such a framework, models address the distribution of substrate uptake through the different modes. In particular, the more recent cybernetic models of Ramkrishna and co-workers^{2–4} consider metabolic regulation, a prime aspect of biological processes, as the optimal distribution of uptake rates for the organism to accomplish a global goal such as maximizing the rate of growth or carbon uptake. The optimal goal is sought by investment of a fixed amount of resources required for (preferential) synthesis of enzymes that catalyze the various metabolic reactions. Young and Ramkrishna³ further model regulation of individual reactions in each EM through optimal distribution of available resources to maximize the flux through that mode. The above model framework requires elaborate measurements for the identification of model parameters. A somewhat simplified cybernetic framework is available from the work of Kim et al.,⁵ which imposes the QSSA for all intracellular variables, thus confining explicit consideration of regulation to only the uptake reactions that ensure the global goal of the organism. The resulting hybrid models, so named because they hybridize the cybernetic approach^{6,7} with the steady-state assumption for internal metabolites, contain significantly fewer parameters to be identified. However, the large number of EMs adds considerably to the kinetic parameters associated with substrate uptake rates, leading to an overparameterization problem, as measurements available for identification are restricted to the limited number of extracellular

* To whom correspondence should be addressed. E-mail: ramkrishn@ecn.purdue.edu. Tel.: (765) 494-4066. Fax: (765) 494-0805.

variables. In this connection, the work of Song and Ramkrishna⁸ shows how a substantial reduction can be effected on the number of EMs from an inspection of experimental data on the yield vector space for extracellular products. Such (or other) model-reduction strategies also enable a class of models proposed by Provost and co-workers, called macroscopic bioreaction models (MBMs),^{9,10} which consider an unregulated distribution of substrate uptake through the different EMs.

To summarize, a collection of models based on the QSSA are available, which we denote as (i) dFBA (where the prefix d signifies the dynamic adaptation of FBA), (ii) MBMs, and (iii) hybrid cybernetic models (HCMs). The models must be tested on experimental data produced by imposed sinusoidal changes in the organisms' environment. Indeed, not only are such data unavailable, but they can also be obtained only with extended experimental effort for which some guidance as to what might be expected from such experiments is most desirable. In this connection, it would appear that data produced by simulations of models of metabolic systems that are free from the QSSA would be sufficient. However, for such an exercise to have potential import for real metabolic systems, the dynamic models must have had a healthy established relationship with dynamic data. The cybernetic model of Young et al.⁴ satisfies this criterion by virtue of its success in accurately simulating two different strains of *Escherichia coli*. As another source of data, we simulate a kinetic model on yeast metabolism due to Teusink et al.¹¹ The primary concern is to examine whether the QSSA is acceptable for those full-order dynamic models and, if so, which specific QSSM is recommended.

This article begins with a brief description of full-order dynamic models to be considered later in the case studies (i.e., kinetic and cybernetic models) and shows how they are reduced to QSSMs under certain hypotheses including the QSSA. A discussion of the similarities and differences among the QSSMs follows to provide a clear understanding of individual models, which is essential for evaluating their applicability. Then, methodologies for testing the applicability of QSSMs in a hierarchical manner are proposed. The examples considered present various cases where QSSMs might or might not be applicable.

2. Metabolic Modeling Based on the QSSA

2.1. Full-Order Dynamic Metabolic Models. Cell growth (in a batch culture) can be described by the following set of differential equations

$$\frac{1}{c} \frac{dx}{dt} = S_x \mathbf{r} \quad (1)$$

$$\frac{d\mathbf{m}}{dt} = S_m \mathbf{r} - \mu \mathbf{m} \quad (2)$$

$$\frac{1}{c} \frac{dc}{dt} = \mu \quad (3)$$

where \mathbf{x} is the vector of n_x extracellular components, \mathbf{m} is the vector of n_m intracellular components, μ is the growth rate of biomass, c is the biomass weight defined per unit volume of the culture, and \mathbf{r} is the vector of n_r fluxes. The matrices S_x and S_m represent the $(n_x \times n_r)$ and $(n_m \times n_r)$ matrices stoichiometrically relating \mathbf{r} to the exchange and intracellular fluxes, respectively. The vector of extracellular species (\mathbf{x}) can be split into the vector of n_s substrates denoted by \mathbf{x}_s and the vector of n_p products denoted by \mathbf{x}_p . Similarly, the stoichiometric matrix S_x for extracellular species can be decomposed into the $(n_s \times n_r)$ matrix S_s for substrates and the $(n_p \times n_r)$ matrix S_p for

products. The growth rate μ can be obtained from the relation with the intracellular fluxes

$$\mu = \mathbf{h}^T \mathbf{r} \quad (4)$$

where the vector \mathbf{h} denotes the contribution of each flux to biomass growth, which is available in the biological literature.¹²

All fluxes within the metabolic network are tightly regulated to maintain their metabolic functions vital for survival under various environmental and genetic perturbations. Such a regulatory effect is, however, difficult to reflect fully in typical kinetic modeling where \mathbf{r} is related to extracellular and intracellular metabolites only using specific kinetic forms, that is, \mathbf{r} is represented as a function of \mathbf{x} and \mathbf{m} . On the other hand, the cybernetic approach models the fluxes as being regulated by controlling enzyme levels and their activities as follows

$$r_i = v_i e_i r_i^{\text{kin}} \quad (5)$$

where e_i is the level of enzyme catalyzing the i th reaction, v_i is the cybernetic variable representing the enzyme activity, and r_i^{kin} is the unregulated reaction rate, which is the part that is kinetically modeled. Consequently, \mathbf{r} is represented as a function of both enzyme levels and metabolite concentrations. The enzyme synthesis reaction is given by

$$\frac{de_i}{dt} = \alpha_i + u_i r_{E,i}^{\text{kin}} - (\beta_i + \mu) e_i \quad (6)$$

where the four terms on the right-hand side represent the constitutive and inducible synthesis rates, the degradation rate, and the dilution rate by growth, respectively; u_i is the cybernetic variable regulating the induction of the i th enzyme; and $r_{E,i}^{\text{kin}}$ is the unregulated part of the inducible rate of the i th enzyme synthesis. The cybernetic variables u_i and v_i are determined according to the cybernetic control laws established by Young,² Young and Ramkrishna,³ and Young et al.,⁴ where cellular regulation is described at the local and global levels using the concept of pathway analysis (i.e., EM analysis). In other words, the local cybernetic variables (\mathbf{u}_L and \mathbf{v}_L) are determined such that the flux through an individual EM is maximized, and the global cybernetic variables (\mathbf{u}_G and \mathbf{v}_G), which adjust the activation of EMs, are determined depending on their contribution to an overall objective function. The overall cybernetic variables are obtained by combining the local and global variables. This formulation is called Young's model (YM), which is now taken as a standard for modern cybernetic modeling. The concept of EMs is provided below in more detail. For a complete description of the YM, readers should refer to the original articles cited above.

2.2. QSSA. In metabolic modeling, it is often assumed that the intracellular metabolites are at a steady state while the extracellular metabolites are dynamic. This is due to the empirical observation that intracellular reaction fluxes are characterized by relatively fast dynamics in comparison to exchange fluxes (i.e., uptake and excretion fluxes). The QSSA is simply translated into the following stoichiometric balances

$$S_m \mathbf{r} = \mathbf{0} \quad (7)$$

This equation comes from eq 2 assuming that the dilution term by growth (i.e., $\mu \mathbf{m}$) is negligible.

2.3. Structure of QSSMs. QSSMs are constructed by combining eqs 1, 3, and 7. Note that eq 2 is replaced by eq 7 in QSS modeling, implying that intracellular fluxes immediately adapt themselves to varying exchange fluxes so that the QSS is maintained all the time. Because of this aspect, the intrinsic application of QSSMs is oriented toward predicting the con-

centration of extracellular metabolites. The estimation of intracellular flux distributions, which is also one of their important usages, is made relying on such external variables. Three QSSMs among other possible varieties are considered here, namely, dFBA, the MBM, and the HCM.

dFBA is an extension of FBA,¹³ one of the most widely used tools for analyzing metabolic systems, to dynamic modeling. FBA employs linear programming (LP) to obtain the solution of eq 7, which is often underdetermined (i.e., the number of unknowns is greater than the number of equations). The uptake fluxes are required to be given as an input to an LP problem where the rest of fluxes are calculated so as to maximize a prescribed objective function J (e.g., the biomass yield). dFBA has the same framework as FBA except that the uptake flux is described using a kinetic expression.^{14,15} The formulation of dFBA can be summarized as follows

$$\frac{1}{c} \frac{d\mathbf{x}_s}{dt} = -\mathbf{r}_s \quad (8)$$

$$\frac{1}{c} \frac{d\mathbf{x}_p}{dt} = \mathbf{r}_p \quad (9)$$

$$\frac{1}{c} \frac{dc}{dt} = \mu \quad (10)$$

where the vector of substrate uptake fluxes, \mathbf{r}_s , and the vector of product secretion fluxes, \mathbf{r}_p , correspond to $(-\mathbf{S}_s\mathbf{r})$ and $\mathbf{S}_p\mathbf{r}$, respectively. Given \mathbf{r}_s , \mathbf{r}_p and μ are determined by solving the following LP problem

$$\max_{\mathbf{r}} J = \sum_i \lambda_i r_i \quad (11)$$

subject to

$$\mathbf{S}_m\mathbf{r} = \mathbf{0}, \quad \mathbf{r}^{\min} \leq \mathbf{r} \leq \mathbf{r}^{\max} \quad (12)$$

where λ_i is the weight of the i th flux and \mathbf{r}^{\min} and \mathbf{r}^{\max} are the vectors for upper and lower bounds on the fluxes, respectively.

The formulation of the MBM and HCM, on the other hand, relies on convex analysis¹⁶ which shows that the solution space forms an unbounded polyhedral cone in flux space. The convex basis vectors spanning the solution space are EMs that correspond to the edge vectors of the cone under the situation that all exchange fluxes are irreversible, which can always be assumed by decomposing any reversible flux into two irreversible ones. Then, any feasible flux vectors satisfying the stoichiometric balances should lie on or inside the cone and thus can be represented by convex (or nonnegative) combinations of EMs

$$\mathbf{r} = \mathbf{Z}\mathbf{r}_M, \quad \mathbf{r}_M \geq \mathbf{0} \quad (13)$$

where \mathbf{Z} is the $(n_t \times n_z)$ matrix, the i th column (i.e., \mathbf{z}_i) of which represents the i th EM and \mathbf{r}_M is the vector of n_z weights of the EMs. Each element of \mathbf{r}_M also represents the vector of throughput fluxes associated with individual EMs. Substitution of eq 13 into eq 1 leads to

$$\frac{1}{c} \frac{d\mathbf{x}}{dt} = \mathbf{Z}_x \mathbf{r}_M \quad (14)$$

where the matrix \mathbf{Z}_x is introduced to denote $\mathbf{S}_x\mathbf{Z}$ for simplicity. The throughput fluxes \mathbf{r}_M can be interpreted as the uptake fluxes through EMs if \mathbf{Z}_x is normalized with respect to a reference substrate as assumed here. Then, the total uptake flux is given by the summation of the elements of \mathbf{r}_M , and the other fluxes are determined from stoichiometric ratios imposed by associated EMs.

The MBM and HCM differ from each other in evaluation of the individual uptake fluxes through EMs. For the MBM, the flux

vector \mathbf{r}_M is represented by a kinetic expression, whereas the HCM accounts for regulation of the uptake flux through that of the enzyme levels and their activities. Following the usual representation of reactions in cybernetic modeling, an individual uptake flux associated with the i th EM in the HCM is modeled as

$$r_{M,i} = v_{M,i} e_{M,i}^{\text{kin}} \quad (15)$$

and

$$\frac{de_{M,i}}{dt} = \alpha_{M,i} + u_{M,i} r_{ME,i}^{\text{kin}} - (\beta_{M,i} + \mu) e_{M,i} \quad (16)$$

where the subscript M denotes the mode composed of a set of local reactions. The interpretation of eqs 15 and 16 is the same as the interpretation of eqs 5 and 6.

Considering the role of \mathbf{r}_M as the weights of the EMs, the difference between the MBM and the HCM lies in the former neglecting regulation, whereas the latter takes the cybernetic view of regulation in splitting the total uptake rate into EMs. It is noteworthy that, because of the neglect of regulation, the distribution of substrate uptake among EMs varies much less in the MBM than in the HCM, producing substantial variations in some situations in the predictions of the two models.

Figure 1 illustrates how full-order and QSSMs are related to one another.

3. Methods for Evaluating the Applicability of QSSMs

A QSSM could be evaluated from direct comparison of the model predictions with the experimental data (i.e., extracellular measurements in this work). The applicability of a QSSM could then come into serious question if the model predictions depart substantially from the experimental data. This is a demanding criterion for complete approval of the model. One drawback of this test is that it is difficult to uncover the sources of error, if observed. The error of a QSSM might originate from various sources, such as (i) insufficiency of the metabolic network, (ii) inapplicability of the QSSA, (iii) inadequacy of other additional hypotheses introduced in individual QSSMs, (iv) poor parameter identification, or any combination thereof. This section provides a systematic methodology by which the individual error sources are examined in a stepwise manner.

3.1. Checking the Stoichiometric Model Using Steady-State Data. The evaluation of the applicability of QSSMs can begin with the stoichiometric model (or network model). Basically, the stoichiometric balances given by eq 7 should be satisfied by the experimental data collected at steady state. This condition can be restated as requiring that the extracellular measurements be represented by a convex combination of EMs, that is

$$\min_{\mathbf{w}_x} \|\mathbf{r}_{x,m}^{\text{ss}} - \mathbf{Z}_x \mathbf{w}_x\|_2 \leq \varepsilon_x, \quad \mathbf{0} \leq \mathbf{w}_x \quad (17)$$

where $\mathbf{r}_{x,m}^{\text{ss}}$ is the vector of n_x (exchange) fluxes taken from the metabolic system at steady state, \mathbf{w}_x is the vector of n_x weights of the EMs in flux space, ε_x is the tolerance for the error, and $\|\cdot\|_2$ represents the L-2 norm. [Note that, although it was already stated in eq 13 that \mathbf{r}_M has the meaning of the weights, a separate notation \mathbf{w}_x (or \mathbf{w}_y in eq 18) was used to indicate that the latter is not related to the kinetic description, unlike the former.]

Equivalently, the criterion of eq 17 can be rewritten with respect to yields, instead of fluxes, as follows

$$\min_{\mathbf{w}_y} \|\mathbf{y}_m^{\text{ss}} - \mathbf{Z}_x \mathbf{w}_y\|_2 \leq \varepsilon_y, \quad \|\mathbf{w}_y\|_1 = 1, \quad \mathbf{0} \leq \mathbf{w}_y \quad (18)$$

where \mathbf{y}_m^{ss} is the vector of n_x yields taken from the metabolic system, and \mathbf{w}_y is the vector of n_x weights to EMs in yield space,

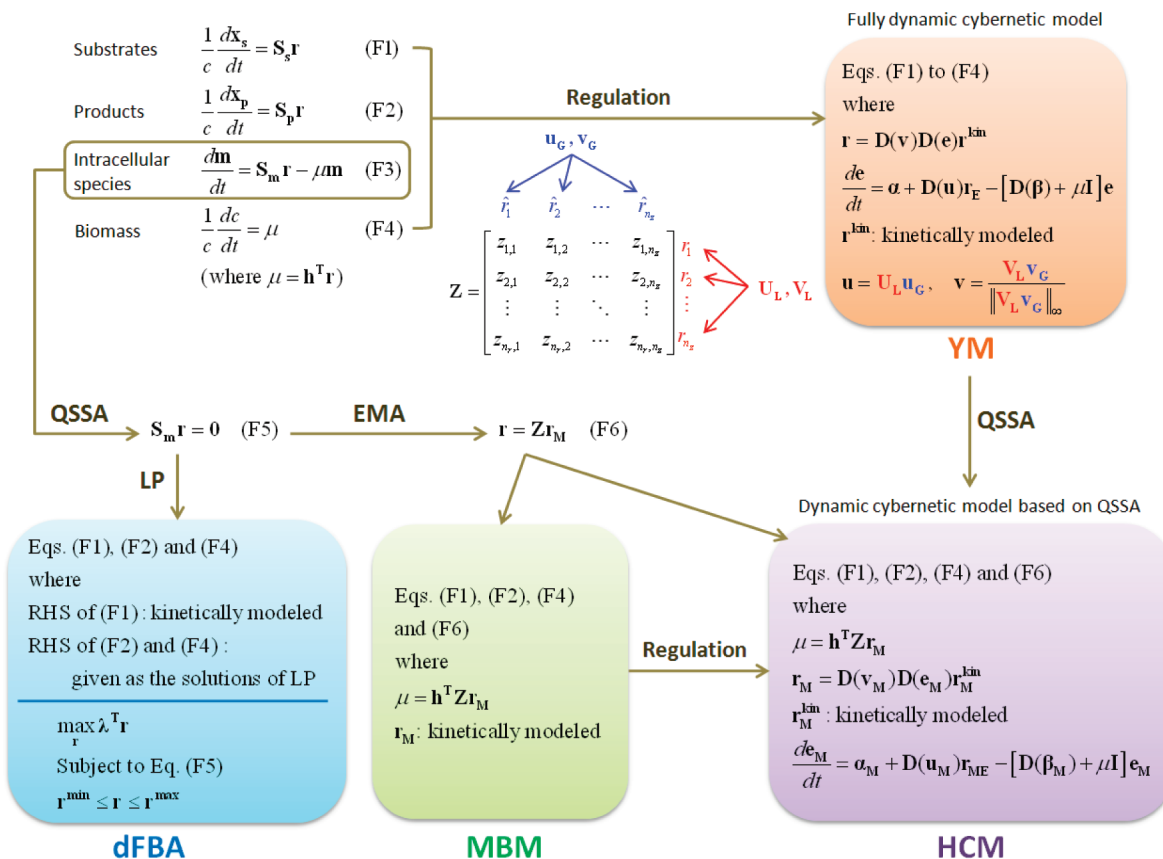


Figure 1. Connectivity of dynamic metabolic models: YM = Young's model, dFBA = dynamic flux balance analysis, MBM = macroscopic bioreaction model, HCM = hybrid cybernetic model, QSSA = quasi-steady-state approximation, EMA = elementary mode analysis, RHS = right-hand side, D = operator converting vector to diagonal matrix. See the Nomenclature section for the meanings of the notations and symbols.

ε_y is the tolerance, and $\|\cdot\|_1$ represents the L-1 norm. The steady-state yield vector y_m^{ss} is obtained by normalizing $r_{x,m}^{\text{ss}}$ with respect to the uptake flux of a reference substrate, r_s^{ss}

$$y_m^{\text{ss}} = r_{x,m}^{\text{ss}} / r_s^{\text{ss}} \quad (19)$$

It is noted that the vector y_m^{ss} always allocates a -1 value to the element corresponding to the reference substrate. Here, the positive/negative signs indicate the production/consumption of species. If the fluxes are defined on a carbon-mole (or mass) basis, the yield coefficients for the products are always confined to the range of 0–1.

Although eq 17 or 18 can serve equally well as a criterion, the latter is preferred for use in this work for convenience in the pictorial analysis. If neither criterion is met, the network should be examined for missing important pathways.

3.2. Evaluation of the QSSA Using Dynamic Data. The next stage is to evaluate the QSSA once the preceding criterion is satisfied. The violation of the QSSA (denoted by E_{QSSA}) can be quantified as

$$E_{\text{QSSA}}(t) = \min_{w_x(t)} \|r_{x,m}(t) - Z_x w_x(t)\|_2 \quad (20)$$

in flux space or

$$E_{\text{QSSA}}(t) = \min_{w_y(t)} \|y_m(t) - Z_y w_y(t)\|_2 \quad (21)$$

in yield space. The dynamic flux (or yield) data in eqs 20 and 21 can be obtained from the transient concentrations of extracellular species generated by perturbing the system.

Equations 20 and 21 measure the excursion of the trajectory from an unbounded polyhedral cone in flux space and a bounded

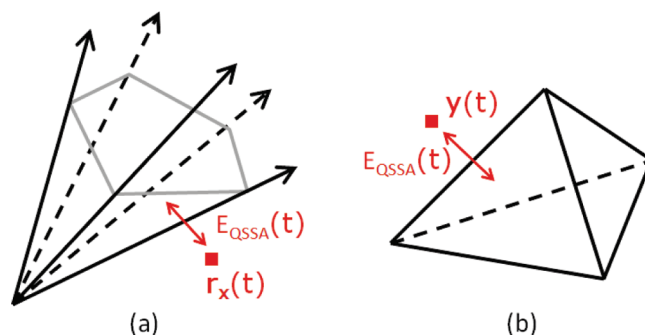


Figure 2. Error of the QSSA [i.e., $E_{\text{QSSA}}(t)$], which can be quantified as the distance (a) between the flux data and the cone in flux space or (b) between the yield data and the convex hull in yield space.

convex hull in yield space, respectively,⁸ as shown in Figure 2. Then, the QSSA can be approved if the following criterion is met

$$\max_t E_{\text{QSSA}}(t) \leq \varepsilon_x \quad \text{or} \quad \max_t E_{\text{QSSA}}(t) \leq \varepsilon_y \quad (22)$$

If the QSSA is shown to be violated, it might be necessary to forego it for metabolites showing slow dynamics. Then, QSSMs can be redesigned with the QSS assumption on a subset of intracellular metabolites, instead of on the full set.

3.3. Analysis of the Transient Data in Yield Space. Even when the QSSA is valid, this fact does not ensure the availability of QSSMs of high quality. This is because a QSSM necessarily involves other additional assumptions that might or might not be valid. Recall that dFBA postulates that metabolic fluxes are controlled so as to maximize the yield of a certain component (e.g., biomass). This hypothesis leads to determination of the

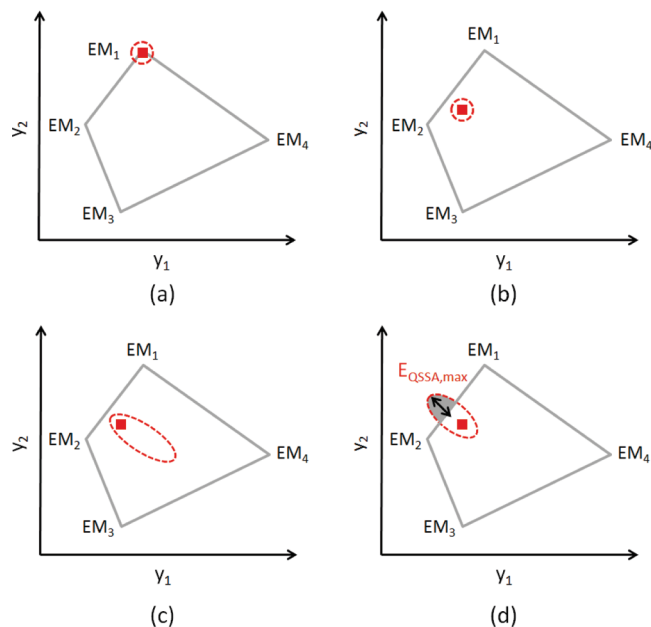


Figure 3. Presumed response of a metabolic system in two-dimensional yield space to external perturbations. The solid square and dashed circle in each plot denote the steady-state data and the area in which the transient trajectory moves under perturbation, respectively.

intracellular flux distribution and product concentrations from a single mode (possibly among many EMs with the same yield of the target product included in the objective function). On the other hand, the MBM and HCM view the uptake flux as being distributed among different EMs. The overall uptake flux is then estimated by convex combinations of EMs. As pointed out earlier, the difference between the HCM and the MBM lies in the former accounting for cellular regulation in distributing the uptake flux based on a global objective. A QSSM might or might not be applicable depending on the applicability of these hypotheses. Clearly, the selection of the best model among three candidates is a separate task from the evaluation of the QSSA.

The investigation begins with examining the characteristics of the transient data that might allow a preliminary screening of the candidate models. Several possible cases are delineated using a simple example illustrated in Figure 3.

Figure 3a features an application for which dFBA, the MBM, and the HCM could be considered. Two conditions should be satisfied for the application of dFBA. First, the steady-state yield data point must be found around a vertex of the convex hull, because the solution of LP is given by one of the extreme points.¹⁷ (Note that the case in which the number of LP solutions is infinite is not considered here.) Second, the trajectory of dynamic yield data must stay sufficiently close to the steady-state point even under perturbation. These requirements imply that one element of the weight vector \mathbf{w}_y in eq 21 is (close to) unity and the rest are (almost) zero. In Figure 3a, $w_{y,1} \approx 1$, and $w_{y,i} \approx 0$ ($i = 2-4$).

Figure 3b,c shows cases where dFBA cannot be applied even though the QSSA is admissible. Unlike the case in Figure 3a, three modes of EM₁–EM₃ are required to represent the steady-state data, that is, \mathbf{w}_y has three nonzero elements. The MBM and HCM appear to be applicable to the case of Figure 3b where the change of yield vector in transient is small (i.e., three nonzero elements of \mathbf{w}_y are almost constant). In Figure 3c, on the other hand, a dynamic shift among the weights occurs; that is, EM₁–EM₃ are activated for some period, and the whole set of EMs is activated for another period. In such a case where the

change of yield vector in transient is large, three possibilities exist: (i) both models work, (ii) only the HCM works, or (iii) neither one works. This should be revealed by rigorous model validation using the experimental data.

Finally, Figure 3d presents a case where no QSSM is applicable because of the violation of the QSSA, which can be detected at the preceding stage of checking criterion 22.

3.4. Summary. The decision on whether QSS modeling is appropriate involves several stages presented in Figure 4. Of main interest is whether the QSSA is admissible and, if it is, which among the class of QSSMs considered serves the best.

The QSSA is evaluated in terms of the fitting error between $\mathbf{y}_m(t)$ and $\mathbf{Z}_x \mathbf{w}_y(t)$. If this deviation is significant [i.e., $E_{\text{QSSA}}(t) > \varepsilon_y$], it implies that the assumption that (all) intracellular metabolites are at QSS is inadmissible. Possibly, some of the intracellular metabolites can show relatively slow dynamics so that they might not be governed by the QSSA. Thus, to eliminate this error source, the QSSA should be selectively applied to intracellular metabolites that show sufficiently fast dynamics.

If the error of the QSSA is negligible [i.e., $E_{\text{QSSA}}(t) \leq \varepsilon_y$], we move to the selection of the best model among the candidates, i.e., dFBA, the MBM, and the HCM. A qualitative analysis of the transient data in yield space might provide information on the appropriateness of the basic hypotheses of individual QSSMs, by which a preliminary screening might be possible. The validation of the selected model is to be made at the final stage by comparing model predictions with the system data.

It is important to note that our assessment of QSSMs has been restricted to discriminating between the behavior of the model and that of the real system in the presence of sinusoidal perturbations. However, in view of inherent nonlinearities, one might find differences between the real system and the QSSM with respect to bifurcation behavior for steady-state multiplicity and periodic oscillations. Although such nonlinear effects could be expected to manifest in a sinusoidally perturbed system, a thorough investigation of such differences in nonlinear behavior would require a concerted application of nonlinear methodology.

4. Case Studies

Using the proposed methods, two metabolic systems are examined as case studies in terms of whether they can be described by low-dimensional QSSMs. Because of the unavailability of the appropriate sets of experimental data, *in silico* systems that are free from the QSSA were used instead. One is the kinetic model for the metabolism of *Saccharomyces cerevisiae* considered by Teusink et al.,¹¹ and the other is the cybernetic model for the metabolism of *E. coli* developed by Young and co-workers.^{2,4} The metabolic networks of yeast and *E. coli* considered therein are given in Figure 5a,b, respectively. In both cases, glucose is fermented under anaerobic conditions.

The behaviors of two systems are profoundly different mainly because the latter accounts for cellular regulation whereas the former does not. Each system is put under sinusoidal fluctuation of uptake flux with different magnitudes and frequencies. The resulting flux data are converted to yield data following the idea suggested in the previous section.

4.1. Metabolic System Considered by Teusink et al.¹¹ The kinetic model considered by Teusink et al.¹¹ involves glucose uptake and secretion of various fermentation products such as glycogen, trehalose, ethanol, glycerol, and succinate (Figure 5a). The entire parameter set in this model for the synthesis of various metabolites was determined by *in vitro* experiments under conditions similar to those in the *in vivo* system. Thus,

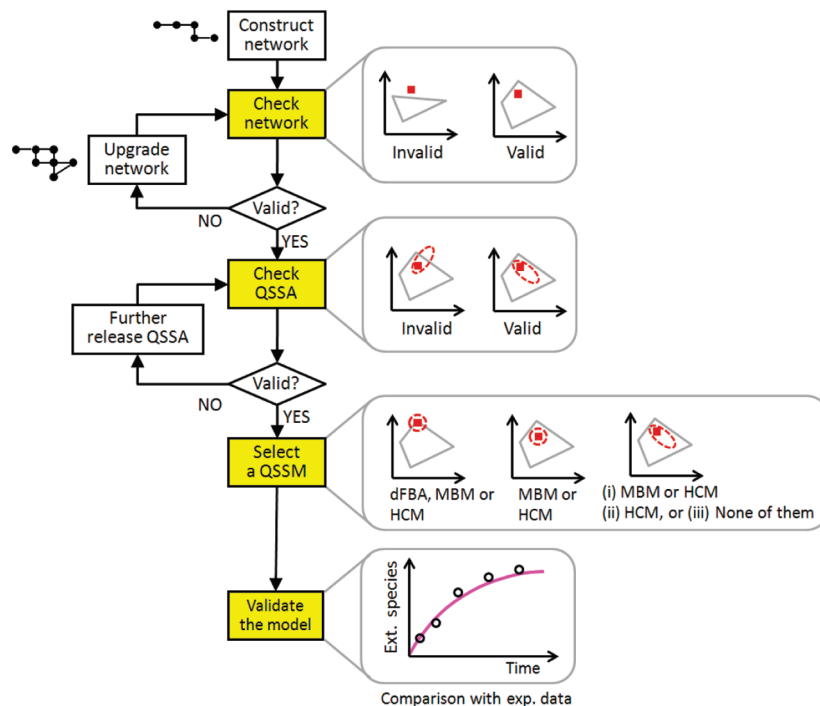


Figure 4. Hierarchical evaluation of the applicability of QSSMs.

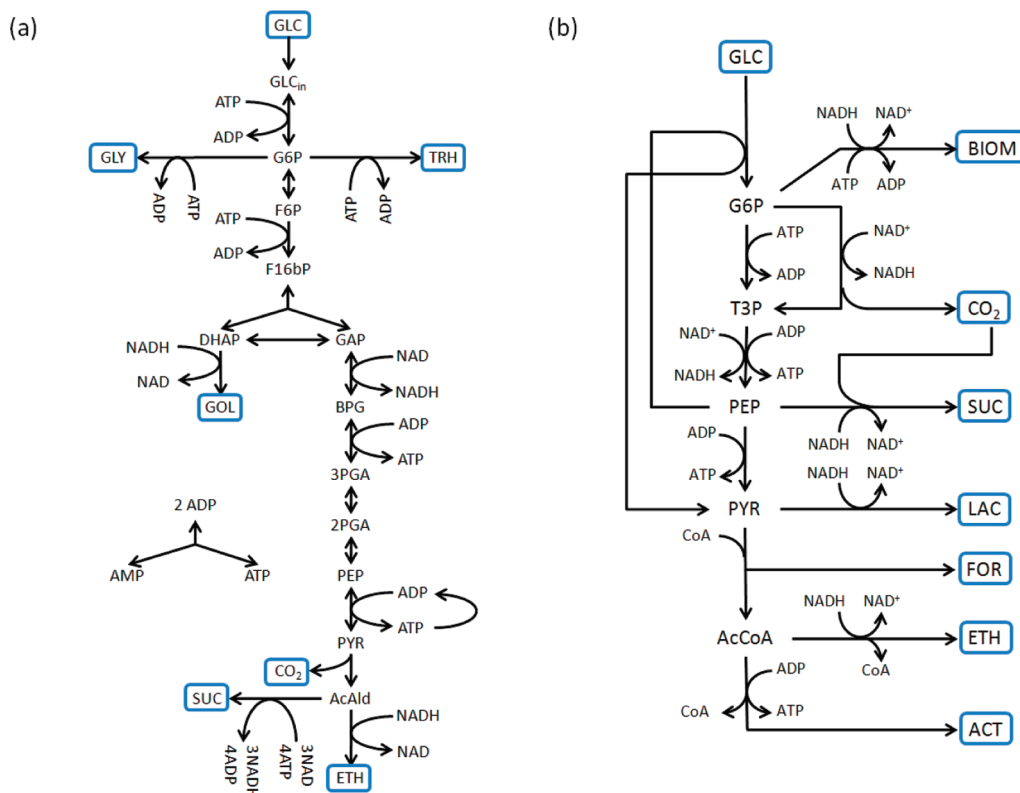


Figure 5. Metabolic networks considered in case studies that were redrawn from (a) Teusink et al.¹¹ and (b) Young et al.⁴ Slight modifications were made from the original networks: All exchange fluxes are set as irreversible in a, whereas the decomposition reaction of FOR into H₂ and CO₂ was neglected in b. The metabolites surrounded by boxes were treated as extracellular in decomposing the network into EMs.

the model presents an attractive source of transient data from a metabolic system, as the applicability of the QSSA is associated with the relative magnitudes of kinetic parameters for substrate uptake and intracellular reactions. Of course, insofar as regulatory effects are not included, the connection of the model to reality must remain incomplete.

The main concern is to examine whether the dynamics of this full-order kinetic model can be described by a reduced-order model based on the QSSA. The model was downloaded from JWS online, which serves as a repository for dynamic metabolic models,¹⁸ and was slightly modified such that exchange reactions are irreversible. Using METATOOL v.5.0,¹⁹

Table 1. Net Reactions of EMs for the Yeast Network Considered by Teusink et al.¹¹

EM	Net reaction
1	GLC = 2ETH + 2CO ₂
2	7GLC = 5GOL + SUC + 7ETH + 9CO ₂
3	3.5GLC = TRE + 3ETH + 3CO ₂
4	2GLC = GLY + 2ETH + 2CO ₂

the metabolic network was decomposed into four EMs as shown in Table 1. (Note that Schwartz and Kanehisa²⁰ obtained eight EMs from the same network system, which is ascribed to neglecting the involvement of cofactors in metabolic reactions.)

For the purpose of illustration, we confine ourselves to the prediction of only two products, ethanol (ETH) and glycerol (GOL). The steady-state data for the system at three different uptake fluxes were plotted in two-dimensional yield space of y_{GOL} and y_{ETH} . It is shown in Figure 6 that the positions of the steady-state yield data vary depending on the magnitude of the uptake flux. Because the steady-state data are away from a vertex of the convex hull, FBA (and subsequently dFBA) cannot be used to describe the system. Of course, insofar as this judgment against dFBA is based only on a comparison with data simulated from the model of Teusink et al.,¹¹ the conclusion is suspect. However, our objective was simply to show how such a conclusion can be drawn from metabolic data.

To evaluate the applicability of the QSSA, the dynamic data were generated using the following sinusoidal inputs (i.e., glucose uptake fluxes)

$$r_{\text{GLC}}(t) = r_{\text{ref}}[1 + \zeta \sin(\omega t)] \quad (23)$$

where r_{ref} is the reference uptake flux, ζ is the fluctuating magnitude, ω is the frequency, and the period of the input (T) is given by $2\pi/\omega$. In this study, r_{ref} was set to 4800 mmol (L of cytosol)⁻¹ h⁻¹ and ζ and ω were varied from 0.1 to 0.3 and from 15π to 600π h⁻¹, respectively. The resulting trajectories of the yield vector are shown in Figure 7. The small figures replicated in Figure 7 have the same scale (on a relative basis) as Figure 6, which is marked.

The trajectory of the yield vector under perturbation stays close to the steady state when the values of ζ and ω are small (i.e., $\zeta = 0.1$ and $\omega = 15\pi$), implying that the weight vector in eq 18 remains almost constant. The trajectories tend to move

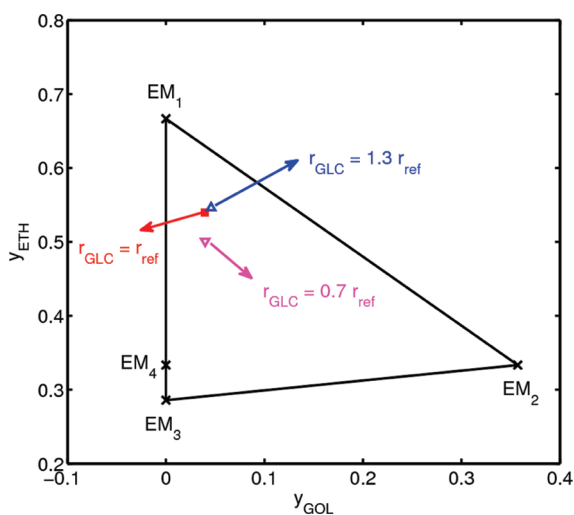


Figure 6. Locations of steady-state data generated from Teusink et al.'s model¹¹ in $y_{\text{GOL}}-y_{\text{ETH}}$ space. The steady-state yield value depends on the magnitude of the uptake flux, the reference value (r_{ref}) of which was set to 4800 mmol (L of cytosol)⁻¹ h⁻¹. The yield vector was defined on a carbon-mole basis.

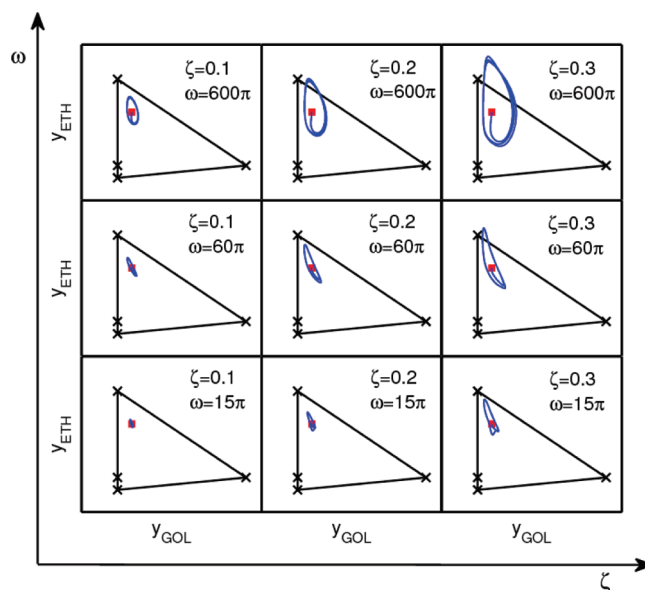


Figure 7. Transient response of Teusink et al.'s model¹¹ with respect to sinusoidal fluctuations of the uptake flux with different magnitudes (ζ) and frequencies (ω) in $y_{\text{GOL}}-y_{\text{ETH}}$ space. The yield vector was defined on a carbon-mole basis.

around more and more as the values of ζ and ω increase, and finally, a significant excursion is made to the exterior of the convex hull, for example, at $\zeta = 0.3$ and $\omega = 600\pi$. This shows that the QSSA is violated under such a severe perturbation to the system, implying that no QSSMs can be applied to describe the original system under such conditions. One possible remedy for this, as suggested earlier, is to forego the QSSA for metabolites showing slow dynamics and reexamine its validity for the rest of the system; this iterative process was not taken in this work.

Equation 23 reflects the imposed changes in the environment as oscillations in the substrate uptake rate. This choice was made simply to create circumstances in which the QSSA will be at risk and to show how the transient data will behave in yield space under these circumstances. In actual practice, sinusoidal variation in the substrate concentration (or dilution rate) might be more reasonably achieved, as conditions under which the total substrate uptake rate might not vary very much. However, at very low substrate concentrations, it is entirely possible for the uptake distribution among different elementary modes to vary substantially and produce the same transient effect as above. Hence, the transient response to a substrate concentration oscillation about a chosen low level might represent a closed curve of transient metabolic states sufficiently displaced from the steady state. It is this behavior that would have been ideal to display for simulated data, as actual experimental data under such transient changes are not available. Neither the model of Teusink et al.¹¹ in example 1 nor the YM in example 2 reflects such effects. Whereas Teusink et al.¹¹ were not concerned with EMs YM did not include numerous EMs for maintenance that would display the change in uptake rates among different modes.

4.2. Metabolic System Considered by Young and Co-workers. The metabolic modeling framework of Young and co-workers^{2,4} provides a fully dynamic description of a metabolic system accounting for cellular regulation. The model formulation includes the simultaneous consideration of local control that maximizes the flux through an individual EM and global control that modulates the competition of EMs to achieve the overall metabolic goal necessary to survive. The HCM has

Table 2. Net Reactions of EMs for the *E. coli* Network Considered by Young et al.⁴

EM	Net reaction
1	51.6GLC \rightarrow 89.0LAC + 2.0CO ₂ + BIOM
2	89.0GLC \rightarrow 30.0CO ₂ + 89.0LAC + 64.1SUC + BIOM
3	49.2GLC \rightarrow 4.1ACT + 4.1FOR + 80.9LAC + BIOM
4	49.2GLC \rightarrow 35.7CO ₂ + 39.8ACT + 39.8FOR + 9.5LAC + 35.7SUC + BIOM
5	35.8GLC \rightarrow 31.0ACT + 27.0ETH + 58.0FOR + BIOM
6	46.9GLC \rightarrow 33.3CO ₂ + 42.1ACT + 4.7ETH + 46.9FOR + 33.3SUC + BIOM
7	59.0GLC \rightarrow 89.0ETH + 89.0FOR + 46.5CO ₂ + BIOM
8	89.0GLC \rightarrow 89.0ETH + 89.0FOR + 51.4SUC + 20.8CO ₂ + BIOM

Table 3. Yield Matrix Z_x Derived from the *E. coli* Network Considered by Young et al.^{4 a}

	Z_{x1}	Z_{x2}	Z_{x3}	Z_{x4}	Z_{x5}	Z_{x6}	Z_{x7}	Z_{x8}
GLC	-1	-1	-1	-1	-1	-1	-1	-1
ACT	0	0	0.03	0.27	0.29	0.30	0	0
ETH	0	0	0	0	0.25	0.03	0.50	0.33
FOR	0	0	0.01	0.13	0.27	0.17	0.25	0.17
LAC	0.86	0.50	0.82	0.10	0	0	0	0
SUC	0	0.48	0	0.48	0	0.47	0	0.38
CO ₂	0.01	-0.06	0	-0.12	0	-0.12	0.13	0.04
BIOM	0.13	0.08	0.14	0.14	0.19	0.14	0.11	0.08

^a Yield vector defined on a carbon-mole basis.

its origin in the YM in that the global objective is retained while the internal fluxes are assumed to satisfy QSSA.

It is believed that the YM generates more physiologically meaningful data than a kinetic model because of its consideration of regulatory effects. The model was successfully validated using experimental data on anaerobic *E. coli* metabolism obtained from the wild-type and pta-ackA knockout strains where a pronounced metabolic shift was observed before and after the gene deletion. Moreover, it makes reasonable predictions for the behaviors of an additional knockout strain, as well as a strain with an exogenous gene overexpressed.

The metabolic network of *E. coli* is decomposed into eight modes (Table 2). In this example, the full yield coordinates including the whole product spectrum are considered for a more complete analysis. The response of the system was examined under perturbations with $r_{\text{ref}} = 50$ and 5 mmol (g of DW)⁻¹ h⁻¹, respectively, and ζ and ω were set 0.3 and 600 π h⁻¹. The QSSA can be considered acceptable for both cases within the tolerance of 1×10^{-3} .

In the case of $r_{\text{ref}} = 50$, the system does not show any appreciable fluctuations in yields or weights (results not shown). Such a robust behavior can be ascribed to the effect of cellular regulation. It is found that uptake through the fifth mode is dominant (i.e., $w_{y,5} = 0.88$) and remains virtually constant during the period of perturbation, thus providing support for dFBA under this circumstance as shown in Table 3. Furthermore, the MBM and HCM are also applicable for this case.

When the magnitude of the uptake flux is diminished (i.e., $r_{\text{ref}} = 5$), we observe periodic shift in weights (Figure 8) and slight fluctuation in yields (Figure 9). Because this behavior is not sufficiently described by dFBA, only the MBM or HCM can be considered for application. Some interesting features are observed in this case. First, only a subset of EMs is activated. In Figure 8, the first three modes are almost inactive throughout the perturbation, implying that the system can be modeled using a reduced set of EMs (i.e., EM₄–EM₈). Such a reduction of EMs is essential in designing the MBM and HCM to avoid the overparameterization problem that was mentioned earlier. This issue was handled in a more comprehensive way by Song and

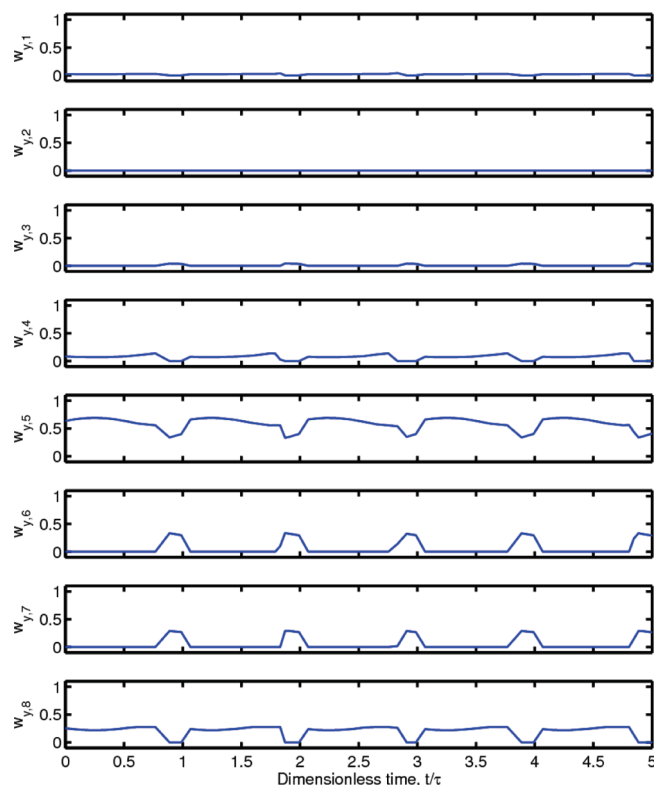


Figure 8. Dynamic change of the optimal weights that minimize the error between convex combination of EMs and yield data generated from the YM by imposing sinusoidal fluctuations of the uptake flux, i.e., $r_{\text{GLC}}(t) = 5[1 + 0.3 \sin(600\pi t)]$ mmol (g of DW)⁻¹ h⁻¹.

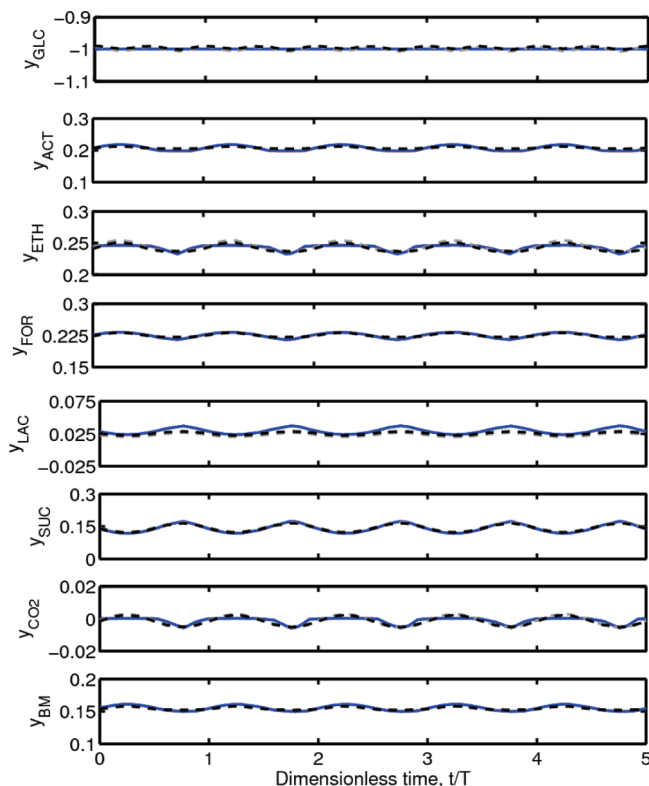


Figure 9. Fitting of MBM (dashed gray line) and HCM (dashed black line) to transient yield data generated from the YM (solid line) when $r_{\text{GLC}}(t) = 5[1 + 0.3 \sin(600\pi t)]$ mmol (g of DW)⁻¹ h⁻¹.

Ramkrishna.⁸ Second, Figure 8 also shows that different modes are periodically activated from one set to another. For the initial

three-quarters of the period, the active set is composed of EM₄, EM₅, and EM₈ and undergoes a switch over to EM₅–EM₇ during the last quarter. This example shows why the transient data are needed to model a metabolic system. Such dynamic data are critical in determining the parameter values associated with activated modes, as well as in identifying an active set. The dynamic switch of active modes in dynamic environments is a particularly interesting aspect of metabolic behavior observed here that models must address.

A comparison of the MBM and HCM is shown in Figure 9, which displays the fits of both models to transient yield data including that for the substrate with -1 as a soft constraint for its value. Both models were constructed using five modes from EM₄ to EM₈, instead of the whole set, based on the observation already made. No appreciable improvement was achieved by adding more modes to this subset. It turns out that the two models exhibit similar fitting capabilities in this example involving a single substrate where the fluctuation of the yield vector is mild.

We address two other examples where deviation of the two models might be expected. One is when the metabolic system undergoes the transition from feast to famine or vice versa. Then, the metabolic switch among EMs would be much more noticeable. This case was not considered here as the version of the YM used here did not include maintenance effects. The other is the sequential utilization of multiple substrates shown by Song et al.²¹ In a case study of recombinant yeast co-consuming glucose and xylose, the MBM and HCM (as well as dFBA) were compared. Four different sets of fermentation data, including the growth data on glucose, xylose, and mixtures at different compositions, were used for the evaluation of the models. As a result, it was found that the HCM showed superior performance and that the difference was marked in the mixture cases, where diauxic behaviors (i.e., sequential utilization of glucose and xylose) were observed.

5. Conclusions

It is established in this article that the methodology of examining trajectories of extracellular measurements in flux space and/or yield vector space can provide a successful diagnosis of whether or not QSSA is admissible. In this process, some resolution is possible between different QSSMs. A more detailed comparison of models is possible by comparison of metrics measuring model deviations from experimental measurements. Figure 4 notably presents a systematic map of the different steps that must be taken for the evaluation of the QSSA and, when the QSSA is valid, for the appropriate choice of QSSM among various alternatives. Although various error sources were discussed for QSSMs, case studies for a demonstration of the same are left for future work.

Acknowledgment

The authors are pleased to acknowledge a special grant from the Dean's Research Office at Purdue University for support of the current work, and NSF GOALI Program (BES-0000961) for support of earlier work which provided the basis of the present work. We also thank Prof. Jamey Young at Vanderbilt University for the help in building up a code to reproduce the simulation results of his Ph.D. thesis.

Nomenclature

Acronyms

dFBA = dynamic flux balance analysis

EM = elementary mode

EMA = elementary mode analysis

FBA = flux balance analysis

HCM = hybrid cybernetic model

MBM = macroscopic bioreaction model

QSS = quasi-steady state

QSSA = quasi-steady-state approximation

QSSM = quasi-steady-state model

Symbols

c = cell dry weight per unit volume of the culture

D = operator converting vector to diagonal matrix

E = error

e = vector of enzyme levels

h = vector correlating fluxes to the biomass growth rate

J = metabolic objective function

m = vector of intracellular species concentrations

n = number

r = vector of fluxes

t = time

T = period of sinusoidal input

u = vector of cybernetic variables

v = vector of cybernetic variables

w = vector of weights

x = vector of extracellular species concentrations

y = vector of yields

z = column vector of the matrix Z

Z = elementary mode matrix

Greek Letters

α = vector of constitutive rates of enzyme synthesis

β = vector of enzyme degradation rates

ε = tolerance

λ = flux weight

μ = biomass growth rate

ω = frequency of sinusoidal fluctuation

ζ = magnitude of sinusoidal fluctuation

Subscripts

E = enzyme

G = global

L = local

M = elementary mode

m = intracellular metabolites or measurements

p = products

ref = reference

s = substrates

x = extracellular metabolites or exchange fluxes

y = yields

Superscripts

kin = kinetic model

max = upper bound

min = lower bound

ss = steady state

T = transpose

Metabolites

ACT = acetate

BIOM = biomass

CO₂ = carbon dioxide

ETH = ethanol

FOR = formate

GLC = glucose

GLY = glycogen
 GOL = glycerol
 LAC = lactate
 SUC = succinate
 TRH = trehalose

Literature Cited

- (1) Schuster, S.; Fell, D. A.; Dandekar, T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* **2000**, *18* (3), 326–332.
- (2) Young, J. D. A system-level mathematical description of metabolic regulation combining aspects of elementary mode analysis with cybernetic control laws. Ph.D. Thesis, Purdue University, West Lafayette, IN, 2005.
- (3) Young, J. D.; Ramkrishna, D. On the matching and proportional laws of cybernetic models. *Biotechnol. Prog.* **2007**, *23* (1), 83–99.
- (4) Young, J. D.; Henne, K. L.; Morgan, J. A.; Konopka, A. E.; Ramkrishna, D. Integrating cybernetic modeling with pathway analysis provides a dynamic, systems-level description of metabolic control. *Biotechnol. Bioeng.* **2008**, *100* (3), 542–559.
- (5) Kim, J. I.; Varner, J. D.; Ramkrishna, D. A Hybrid Model of Anaerobic *E. coli* GJT001: Combination of Elementary Flux Modes and Cybernetic Variables. *Biotechnol. Prog.* **2008**, *24* (5), 993–1006.
- (6) Kompala, D. S.; Ramkrishna, D.; Tsao, G. T. Cybernetic Modeling of Microbial Growth on Multiple Substrates. *Biotechnol. Bioeng.* **1984**, *26* (11), 1272–1281.
- (7) Kompala, D. S.; Ramkrishna, D.; Jansen, N. B.; Tsao, G. T. Investigation of Bacterial Growth on Mixed Substrates—Experimental Evaluation of Cybernetic Models. *Biotechnol. Bioeng.* **1986**, *28* (7), 1044–1055.
- (8) Song, H.-S.; Ramkrishna, D. Reduction of a set of elementary modes using yield analysis. *Biotechnol. Bioeng.* **2009**, *102* (2), 554–68.
- (9) Provost, A.; Bastin, G. Dynamic metabolic modelling under the balanced growth condition. *J. Process Control* **2004**, *14* (7), 717–728.
- (10) Provost, A.; Bastin, G.; Agathos, S. N.; Schneider, Y. J. Metabolic design of macroscopic bioreaction models: Application to Chinese hamster ovary cells. *Bioprocess Biosyst. Eng.* **2006**, *29* (5–6), 349–366.

(11) Teusink, B.; Passarge, J.; Reijenga, C. A.; Esgalhado, E.; van der Weijden, C. C.; Schepper, M.; Walsh, M. C.; Bakker, B. M.; van Dam, K.; Westerhoff, H. V.; Snoep, J. L. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* **2000**, *267* (17), 5313–5329.

(12) Stephanopoulos, G. N.; Aristidou, A. A.; Nielsen, J., *Metabolic Engineering: Principles and Methodologies*; Academic Press: San Diego, 1998.

(13) Reed, J. L.; Palsson, B. O. Thirteen years of building constraint-based in silico models of *Escherichia coli*. *J. Bacteriol.* **2003**, *185* (9), 2692–2699.

(14) Hjersted, J. L.; Henson, M. A.; Mahadevan, R. Genome-scale analysis of *Saccharomyces cerevisiae* metabolism and ethanol production in fed-batch culture. *Biotechnol. Bioeng.* **2007**, *97* (5), 1190–1204.

(15) Mahadevan, R.; Edwards, J. S.; Doyle, F. J. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* **2002**, *83* (3), 1331–1340.

(16) Rockafellar, R. T. *Convex Analysis*; Princeton University Press: Princeton, NJ, 1970.

(17) Dantzig, G. B. *Linear Programming and Extensions*; Princeton University Press: Princeton, NJ, 1963.

(18) Olivier, B. G.; Snoep, J. L. Web-based kinetic modelling using JWS Online. *Bioinformatics* **2004**, *20* (13), 2143–2144.

(19) von Kamp, A.; Schuster, S. Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics* **2006**, *22* (15), 1930–1931.

(20) Schwartz, J. M.; Kanehisa, M. Quantitative elementary mode analysis of metabolic pathways: the example of yeast glycolysis. *BMC Bioinformatics* **2006**, *7*, 186.

(21) Song, H.-S.; Morgan, J. A.; Ramkrishna, D. Systematic development of hybrid cybernetic models: Application to recombinant yeast co-consuming glucose and xylose. *Biotechnol. Bioeng.*, published online Mar 19, 2009, <http://dx.doi.org/10.1002/bit.22332>.

Received for review January 16, 2009

Revised manuscript received April 24, 2009

Accepted April 27, 2009

IE900075F