Standard Review

Toward a comprehensive description of microbial processes through mechanistic and intelligent approaches

Pratap R. Patnaik

Institute of Microbial Technology, Sector 39-A, Chandigarh-160036, India.
IMTECH communication No. 048/2007. E-mail: pratap@imtech.res.in. Fax: +91-172-2690132/585.

Accepted 21 August, 2008

Microbial processes functioning in bioreactors under realistic conditions are subject to incomplete dispersion and the presence of noise from the environment and within the cells. These factors complicate the development of good quantitative descriptions of microbial reactors. Most analyses have therefore focused on either the intra-cellular or the extra-cellular processes, ignoring or simplifying the other facet. The resulting models are thus useful only for the intended purposes and in limited domains, but they do not include a comprehensive description of all features. These models have employed one or more of three main approaches to develop quantitative descriptions – mechanistic, cellular intelligence (or cybernetic), and artificial intelligence. Models using judicious combinations of two or more methods have wider and more versatile applicability. However, no model has accommodated both intra-cellular and extra-cellular noise in a macroscopic description of a nonideal bioreactor. Based on a review of recent studies, such a conceptual model is presented here. It combines all three approaches in a flexible design.

Key words: Microbial processes, dispersion, noise, modeling approaches, comprehensive description.

INTRODUCTION

Cellular processes are often more complex than chemical processes, even when both generate the same outputs. Many factors contribute to this complexity. At a fundamental level, living cells sustain more complex and intricate networks of reactions than many chemical reactions. Cellular reaction networks are not always completely deciphered, thus underlining their complexity and necessitating simplifying methods to formulate workable kinetic models (Gombert et al., 2000; Varner and Ramkrishna, 1999). Unlike chemical reactions, metabolic reactions are regulated by structurally and functionally complex molecules such as DNA, RNA and enzymes, many of whose concentrations are small but important. The low concentrations also make these molecules sensitive to noise at the genetic level (Kaern et al., 2005), which in turn has an impact on the observed behavior of the process (Haag et al., 2005).

Cellular metabolic processes also respond to environmental changes in ways that are difficult to capture through models constructed on chemical kinetic principles alone. The lag phase behavior of cultures transferred from one medium to another, the responses to abrupt changes in input streams, and cellular dynamics in the presence of external noise are some examples. Such observations may however be described quantitatively by "intelligent" models. By contrast with the static nature of chemical kinetic models, intelligent models employ either inherent or artificial intelligence. Inherent intelligence is the basis of the so-called cybernetic models (Dhurjati et al., 1985; Patnaik, 2000) which attribute to living cells the ability to retain information, understand it and adjust their responses on the basis of past experience. Artificial intelligence has been invoked largely to model macroscopic microbial behavior under the influences of noise and spatial variations in a bioreactor. Methods such as artificial neural networks, fuzzy logic and genetic algorithms then provide more faithful representations of varied cellular dynamics than mechanistic models do (Hodgson et al., 2004; Patnaik, 2006a).

While intelligent descriptions of microbial processes in bioreactors have decisive advantages over mechanistic models under realistic conditions, they have limitations too. Cybernetic models tend to be quite complex, often resulting in large sets of differential equations, and some-
times more than one cybernetic goal seems to explain the observed behavior equally well (Patnaik, 2000; Straight and Ramkrishna, 1994). A pivotal feature of cybernetic models is the presence of regulatory key enzymes, but it has not always been possible to establish a correspondence between these and the enzymes actually detected. A common criticism of artificial intelligence methods is that they are too empirical and do not incorporate the physiological features of cellular processes. Consequently, it becomes difficult to provide physical interpretations for the parameters of these models and relate them to kinetic and thermodynamic features (James et al., 2002; Lubbert and Jorgensen, 2001).

Despite their inadequacies, mechanistic models are simple and are derived through biological and physical principles. Therefore they have greater physiological closeness to internal cellular processes; as a result, their parameters can be attributed physical meaning and can be manipulated by introducing mutational, genetic or operational changes (Haag et al., 2005; Hodgson et al., 2004). Given the different strengths of mechanistic models and intelligent models, it might be useful to combine the two approaches to develop composite (or hybrid) models for microbial reactors. These composite models should, in principle, turn out to be adaptive, flexible and self-regulating like intelligent models, but also possess the fundamental biological basis of mechanistic models.

The idea of composite or hybrid models is not new. Many previous studies (Coleman et al., 2003; Galvanauskas et al., 2004; James et al., 2002; Patnaik, 2003a) have combined two or more modeling approaches to optimize and control bioreactors for different microbial cultures. However, no investigation has yet been reported of an approach that includes both intra-cellular and extra-cellular noise as well as cybernetic kinetics and bioreactor nonidealities. Since all these are important features of real microbial processes, this communication provides a perspective of how the development of existing methods of modeling can lead to such a composite descriptive framework.

**NONIDEAL FEATURES OF CELLS AND REACTORS**

**Cellular noise and complexity**

Cells synthesize products as a result of the expression of specific genes. Molecules such as DNA, mRNA and proteins are involved in gene expression. These molecules are usually present in concentrations sufficiently low for stochastic effects to become significant (Kaern et al., 2005; Raser and O'Shea, 2005). As a result, the amount of a particular protein that a gene synthesizes fluctuates from cell to cell in a population and with time for each cell. These fluctuations (termed genetic noise) arise from a number of sources but they may be categorized broadly as either (a) intrinsic or (b) extrinsic.

Intrinsic noise refers to fluctuations associated with promoter activation or deactivation and the synthesis and decay of mRNA and proteins. Extrinsic noise pertains to fluctuations in gene products such as RNA polymerase, ribosomes and certain proteins. Although external to the relevant genes, these fluctuations act on the genes, thus complicating both genetic expression and the identification of the two contributing effects. Nevertheless, Elowitz et al.'s (2002) two-reporter assay provides an elegant method to differentiate and quantify these two types of genetic noise.

While it may be possible to measure intrinsic noise and extrinsic noise and relate them to biochemical parameters (Swain et al., 2002), it still remains difficult to quantify individual repositories of intrinsic noise. Intrinsic noise may be present at any of three levels: (a) individual genes, (b) reaction pathways in a network, and (c) the cells as a whole. Each of these sources affects particular intra-cellular processes but all three interact as shown in Figure 1.

Since cellular processes involve complex networks of reactions regulated at the genetic level, they should be able to withstand stochastic effects so that the cellular machinery can function without being destabilized. In other words, cells should be sufficiently robust to both intrinsic and extrinsic noise. Complexity and robustness are inter-related, and examples abound in biological and ecological systems (Carson et al., 2006). Robustness is the maintenance of specific characteristics of system behavior in the face of perturbations (Carson and Doyle, 2002; Kitano, 2004a; Stelling et al., 2004). Kitano (2004a) has argued that complex evolvable systems are necessarily robust; since evolution is a fundamental trait of living cells, they too are robust.

Many factors contribute to robustness, of which feedback is a prominent example. While positive feedback amplifies fluctuations and negative feedback attenuates them, the former also increases phenotypic diversity in a population of cells (Becskei and Serrano, 2000; Kaern et al., 2005; Rao et al., 2002). In a heterogeneous population, different phenotypes have different survival probabilities under given environmental conditions, and this has important implications for disease control (Balaban et al., 2004). Since both negative and positive feed-back have beneficial as well as detrimental effects, it may be worthwhile to design gene networks that incorporate the helpful features of both; research on HIV (Richman, 2001) and cancer (Kitano, 2004b) indicates that this is possible.

The evolvability of living systems is also a core concept underlying cybernetic models of microbial processes. The cybernetic approach (Dhurjati et al., 1985; Varner and Ramkrishna, 1999) attributes to living cells the ability to learn from their experiences and accordingly respond optimally to changing circumstances. Cybernetic models have not only provided more faithful representations of the dynamic behavior of microbial reactors (Dhurjati et al., 1985; Hodgson et al., 2004; Namjoshi and Ramkrishna, 2001; Patnaik, 2000) but also explained uncommon patterns of behavior that were considered aberrant by mechanistic modeling approaches (Narang et al., 1997;
Heterogeneity and noise in bioreactors

Populations of cells are usually cultivated in bioreactors under conditions that favor specific objectives. Typical objectives are the growth of the cells themselves, the formation of particular products, and the removal of harmful components from the environment of the cells. To achieve these objectives, it is possible to have elaborate controls of small bioreactors used in the laboratory but practical difficulties and high costs limit both measurements and control in large bioreactors. Therefore, larger reactors are less ‘ideal’ than small vessels and generate less of the product(s). Two significant nonideal features are: (a) spatial variations within the reactor, as a result of incomplete mixing or dispersion and (b) the influx of noise from the environment. The latter feature is obviously more likely in continuous flow and fed-batch operations; nevertheless, kinetic and thermodynamic considerations often favor the choice of such operating modes (Liden, 2001).

While spatial heterogeneity increases with reactor size, even small bioreactors can have significant gradients on a microscopic scale (Larsson et al. 1996). This makes the optimal positioning of sensors difficult and expensive. Moreover, the presence of spatial variations and the influx of noise from the environment create differences among the cells and in the distribution of fluxes along the pathways of metabolic networks, in the yields of products and sometimes in the stability of the fermentation. These effects are illustrated by numerous studies of the production of ethanol by Saccharomyces cerevisiae in continuous cultures (Garhyan and Elnashaie, 2004; Garhyan et al., 2003; Patnaik, 2005). Recall here that extrinsic noise also creates differences between cells in a population (Kaern et al., 2005; Raser and O’Shea, 2005). Since both extrinsic noise and external (environmental) noise have an impact on the cells, this raises an intriguing question: will the two sources of noise amplify or nullify each other? Depending on the operating conditions, either effect is possible: noise may drive a culture from a monotonic stationary state to an oscillating or a chaotic state and, alternately, proper filtering of noise can restore stable states from chaotic conditions (Garhyan and Elnashaie, 2004; Garhyan et al., 2003; Patnaik, 2006b).

The complexities described above make it difficult to formulate mechanistic models and control policies based on them that are sufficiently accurate, simple, flexible, adaptive, fast and robust. These difficulties have motivated the use of artificial intelligence (AI) methods for on-line estimations of important variables and optimal control of bioreactors. Artificial neural networks (Gadkar et al., 2005), genetic algorithms (Hodgson et al., 2004) and fuzzy logic (Arnold et al., 2002) have been used in different applications. Although they are superior to classical methods of modeling and control, AI methods are, in effect, ‘black box’ entities with poor physiological foundations. Not surprisingly, therefore, they are often difficult to train, do not always yield unique models and have limited capability outside their training domains. These weaknesses have generated hybrid models that combine AI or cybernetic models with mechanistic equations. Many applications of hybrid models have been reviewed recently (Galvanauskas et al., 2004). While establishing the usefulness of hybrid approaches, they also reveal areas that are still uncertain and require further inquiry. These are discussed in later sections.

A CONCEPTUAL BASIS FOR MODEL DEVELOPMENT

The formulation of a comprehensive model for a bioreactor begins at the cellular level, integrates this with...
the extra-cellular environment, establishes a quantitative
description of this environment through mass balances,
and incorporates non ideal features. The processes
inside the cells themselves are sufficiently complex to
require a hierarchical portrayal; Figure 2 presents such a
portrait. Many studies (Dun and Ellis, 2005; Shimizu, 2002;
Wang et al., 2006) have discussed cellular functions from
the perspective of Figure 2, so that discussion will not be
repeated here. An overriding feature of this conceptual
depiction is the emergence of ‘omic’ structures as build-
blocks for intra-cellular processes. Genomics des-
cribes the primary functions at the level of genes. Ensem-les of genes may differ from one genome to another,
thus controlling the expression of different proteins throu-
g through DNA transcription processes (transcript-tomics).
The expressed proteins differ in their structures, functions,
stability and interactions within individual mole-cules and
between molecules; these features are the domain of
proteomics.

Macroscopic manifestation of these intra-cellular facets
begins at the metabolomic stage, where metabolic regu-
laratory networks and fluxes along pathways constituting
these networks are analyzed. The flow distributions are
controllable and are critical to the product distributions
obtained in microbial cultures. Thus metabolomics pro-
vides a vital interface between genome-level processes
and those at the bioreactor level. The hierarchical struc-
ture of the ‘omic’ domain continues through the meta-
bolome to the bioreactor, as illustrated in Figure 2 (Ortoleva
et al., 2003; Wang et al., 2006). Note that from the
microscopic to the macroscopic level the number of
variables reduces from up to \(10^5\) (reflecting the complex-
ity of cellular processes) to less than 10. This reduction
is important because practical monitoring and control of a
bioreactor can be done efficiently only with a limited num-
ber of variables.

Multi-cellular processes that differ so widely in comple-
xity, time scales and the number of variables may under-
standably be described by more than one approach.
Different workers have adopted different approaches and
used different assumptions, according to the conditions of
the system being studied and the objective. There appear
to be five main approaches, with differences and simi-
larities:

i.) Equation-oriented approach.
ii.) Signal-oriented approach.
iii.) Cellular intelligence approach.
iv.) Artificial intelligence approach.
v.) Composite (or hybrid) approach.

These are discussed below, followed by a proposal to
combine some of them to formulate a composite compre-
hensive model.

**EQUATION-ORIENTED MODELING**

The underlying premise for equation-oriented models is
that biological processes follow the same laws as chemi-
ical processes. This implies that, as for chemical reac-
tions, mass balances and kinetic equations can be formu-
lated on the basis of measurements of inputs, outputs
and intermediates. We recapitulate here that cellular
systems involve reactions inside the cells and transport
between (a) the cells and the extra-cellular broth, (b)
through the broth itself and (c) sometimes between the
broth and the external environment. On this basis two
streams of modeling have evolved independently. One
relates to the intra-cellular metabolic processes (Gombert
and Nielsen, 2000; Varner and Ramkrishna, 1999) and
the other pertains to bioprocesses (Bailey, 1998; Lubbert
and Jorgensen, 2001). However, a comprehensive descrip-
tion of a microbial system should encompass both
streams, and Haag et al. (2005) work illustrates how this
may be done.

Consider a set of reactions that follow the stoichiometry

\[
\sum_{i=1}^{m} \kappa_{S_i} S_i \rightarrow \sum_{i=1}^{n} \kappa_{P_i} P_i
\]

Where \(S_i\) is the i-th substrate and \(P_i\) the i-th product. In a
perfectly mixed continuous flow stirred tank bioreactor,
the mass (or molar) balances for each component in the
extra-cellular fluid may be written as

\[
\frac{d(c_{ex})}{dt} = \overline{K}_{ex} \overline{r} V + c_{in} Q_{in} - c_{ex} Q_{out}
\]

Where \(c_{ex}\) is the vector of exit concentrations, \(c_{in}\)
the vector of inlet concentrations, \(Q_{in}\) the inflow rate, \(Q_{out}\)
the outflow rate, \(V\) the volume of material in the bioreactor and t
the elapsed time, \(\overline{r}\) contains the rate terms for the
concentrations \(c_{ex}\) , and \(\overline{K}_{ex}\) is a vector of stoichiometric
rate constants.

Normally \(Q_{out} = Q_{in}\) and hence \(V\) is constant. Then, with
the dilution rate defined as \(D = Q_{in}/V\), Eqn. (2) becomes

\[
\frac{dc_{ex}}{dt} = \overline{K}_{ex} \overline{r} + (c_{in} - c_{ex}) / D
\]

Haag et al. (2005) also accounted for the common obser-
vation that, at any time, some cells are active (or viable)
and others are inactive (or dead). Ignoring detailed me-
chanisms, they considered simply that, overall, dead cells
\((X_d)\) arise irreversibly from viable cells \((X_v)\). Then the
mass balances for the biomass may be expressed as:

\[
\frac{dc_{x_v}}{dt} = \overline{k}_{x_v}^T \overline{r} - Dc_{x_v} = (\mu - k_d) c_{x_v} - Dc_{x_v}
\]

where \(\overline{k}_{x_v}\) is the reaction rate matrix, \(\mu\) is the
specific growth rate, and \(k_d\) is the death rate constant.
Here $\bar{K}_{x_d}$ and $\bar{K}_{x_d}$ contain the respective reaction rate constants, and $k_d$ is the rate constant for the decay of viable cells. Combining Eqs. (4) and (5), the total specific growth rate may be obtained from

$$\frac{dc_{x_d}}{dt} = \bar{K}_{x_d} R - Dc_{x_d} = k_d c_{x_d} - Dc_{x_d}$$

(5)

The equality $\mu_{tot} = \mu$ is understandable since only viable cells contribute to the growth. Metabolic reactions inside the cells contribute to changes in the viable cell mass. Therefore, similar to Equation (2), global balances for the intra-cellular metabolites, $\bar{c}_i$, may be written as

$$\frac{d(\bar{c}_i c_{x_i} V)}{dt} = \bar{K}_i \bar{q} c_{x_i} V - \bar{c}_i k_d c_{x_i} V - \bar{c}_i c_{x_i} Q_{out}$$

(7)

The vector $\bar{q}$ contains the metabolic fluxes along the reaction pathways. Equation (7) may be recast in the form of Equation (3) to obtain

$$\frac{d\bar{c}_i}{dt} = \bar{K}_i \bar{q} - \bar{c}_i \mu$$

(8)

Since the control volume of the bioreactor alone is a closed system, the mass flows between the cells and their environment have to be balanced. This leads to

$$\bar{c}_{ex} (t) = \bar{K}_{ex} \bar{c}_{ex} (\bar{c}_{ex}, \bar{c}_i) + \bar{g}_{ex} (\bar{c}_{ex}, t)$$

(9)

$$\bar{c}_i (t) = \bar{K}_i \bar{q} (\bar{c}_{ex}, \bar{c}_i) - \bar{c}_i \mu (\bar{c}_{ex}, \bar{c}_i)$$

(10)

with $\bar{g}_{ex}$ containing the terms representing exchanges with the environment outside the bioreactor. In the simplest case without gaseous inflow or outflow,

$$\bar{g}_{ex} (\bar{c}_{ex}, t) = [\bar{c}_{in} (t) - \bar{c}_{ex} (t)]D$$

(11)

This model is rigorous but complex. Moreover, it depends on intra-cellular concentrations, which are usually difficult to measure, and it ignores intra-cellular regulatory processes (Dhurjati et al., 1985; Patnaik, 2000). These limitations have been exposed in studies of hybridoma cultures (Namjoshi et al., 2003; Zupke and Stephano-poulos, 1995).

**SIGNAL-ORIENTED APPROACH**

As an alternative to the “bag full of enzymes” (Lengeler et al., 1999) approach of equation-oriented mechanism-based modeling, the signal-oriented approach views a cell as a mosaic of function units with signals flowing between them. The content and the nature of flows determine the ultimate functions of a cell. Three biological criteria are used to demarcate these units: (i) the presence of an enzymatic network with a common physiological task, (ii) control of this network at the genetic level by a common regulatory unit, such as an operon or a regulon, usually organized in a hierarchical way, and (iii) the coupling of this regulatory network through a signal transduction system.

Based on these concepts, Kremling and coworkers (Kremling et al., 2000) proposed a signal-oriented description of cellular dynamics. They illustrated their method with *Escherichia coli*. In their application each functional unit is characterized by two “coordinates”. The structural coordinate is described by the number and type of inputs and outputs. For example, a functional unit may describe transcription processes connecting a pool of nucleotides with the RNAs. The second coordinate is behavioral and it is expressed by mathematical equations describing the structural object. As might be expected, functional units differ in their complexity and response times.

Kremling et al. (2000) also assigned to each unit an indicator molecule (called an alarmone) whose level of activity controls the activities of superimposed regulatory networks. It is worth remembering this concept in order to draw a parallel with the cybernetic approach to be described later.

Each function may be composed of one or more elemental submodels or model objects. These are basically of three types: (i) substance storages, (ii) substance transformers and (iii) signal transformers. Substance storage devices may either contain genetic information (as in the cases of DNA, RNA and proteins) and may not (e.g. intermediate metabolites). Signal transformers form the central nervous system of the signal-oriented approach. As Kremling et al. (2000) point out: they provide the crucial links between the reception of stimuli, either from inside the cells or from the external environment, and the cellular responses.

Signal transformers also help to differentiate between metabolic flows and signal flows. Metabolic networks comprise metabolic and regulatory subnetworks. The metabolic subnetwork contains the metabolic fluxes whereas the regulatory component describes the signal transduction processes. The signal-oriented approach may be illustrated by the simple example of the synthesis of a protein (Kremling et al., 2000). This requires the processes of transcription, translation and replication, which are modeled as shown in Figure 3. Each process has three components—substance storage, substance transformer and signal transformer. As seen, transcription and translation provide unidirectional signal transfer. The signal transformer of the transcription cascade processes information about DNA sequences, regulatory proteins,
etc. and determines the transcription efficiency. The product of transcription - mRNA - is an input to the signal transformer of the translation stage, where the translation efficiency is determined for the synthesis of the final protein.

The chemosensory system of *E. coli* (Figure 4) demonstrates the application of the signal-oriented approach to a more complex problem. Contrary to Fickian diffusion, *E. coli* move up a chemical gradient. A chemotactic path comprises alternate straight-line movements ("runs") and corrective changes of direction ("tumbles"). The movements are effected by long helical flagellae attached to rotary motors embedded in the cell surface. When the motors spin counter-clockwise, they propel the cells along straight paths. Clockwise rotations generate tum-

**CELLULAR INTELLIGENCE APPROACH**

Ramkrishna and associates proposed a different perspective of microbial metabolism and growth. Like Kremling and coworkers, they were motivated by the inability of mechanistic models to portray and predict certain features of microbial cultures. For example, mechanistic kinetics accounts for steady state variations with the dilution rate in continuous flow bioreactors by invoking an unproven maintenance term but still fails to handle the transient approach to a steady state. Another case is diauxic growth on mixture of two substrates, where the
The mechanistic approach cannot explain why one substrate is ignored until the other is exhausted (Ramkrishna, 2003).

To account for such apparent oddities, Ramkrishna and coworkers (Dhurjati et al., 1985; Straight and Ramkrishna, 1994; Varner and Ramkrishna, 1999) proposed that living cells have internal regulatory controls that enable the cells to exercise judgment in a given set of conditions. An alternate way to describe this is to say that cells possess some rudimentary intelligence that enables them to learn from their experiences and respond to environmental changes in a manner that is most favorable to themselves.

Figure 5 represents a flow sheet of the key stages in a cybernetic modeling framework (Dhurjati et al., 1985). A typical cell contains an “adaptive machinery” that controls metabolic transformations in response to extra-cellular variations. The extra-cellular soup is viewed as a resource pool (of substrates and other nutrients), whose constituents are allocated optimally to different metabolic pathways such that a desired objective (such as cell growth) is maximized. Once the essential proteins are synthesized, a “permanent machinery” carries out the metabolic reactions for replication of cellular material. The third component is a “regulator”, and it embodies a crucial feature of cybernetic modeling that distinguishes it from mechanistic modeling. The regulator controls the distribution of resources to achieve the maximization objective referred to above. The cells choose their objectives such that their survival is favored at all times. In this sense, the cybernetic approach formalizes the evolutionary approach that Demain (1971) had recognized nearly four decades ago.

Cybernetic modeling begins with the concept of a key enzyme whose synthesis and activity are regulated cybernetically. The utilization of each substrate (in a mixture) is regulated by a corresponding key enzyme. To explain the cybernetic process in simple terms, let \( n \) substrates \( S_1, S_2, \ldots, S_n \) contribute to the synthesis of an equal number of proteins \( P_1, P_2, \ldots, P_n \). Let \( R_i \) be the allocation rate of \( S_i \) to \( P_i \), and \( R \) the total allocation rate. Then \( u_i = R_i/R \) is the fractional allocation \( S_i \) to \( P_i \). Based on Mandelstam and McQuillen’s (1968) work, Dhurjati et al. (1985) considered \( R \) to be constant, thereby resulting in the constraint

\[
\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} R_i / R = 1 \tag{12}
\]

The total rate of production of cell mass, \( X \), is the sum of the concentrations from the individual substrates:

\[
\frac{dx}{dt} = -\sum_{i=1}^{n} Y_{si} \frac{ds_i}{dt} \tag{13}
\]

where \( Y_{si} \) is the yield of cell mass per unit mass of \( S_i \) consumed. The rate of consumption of \( S_i \) depends, among other factors, on its key enzyme \( E_i \), whose activity, \( e_i \), varies with time. For Monod kinetics,

\[
\frac{ds_i}{dt} = \frac{\mu_{\text{mij}} e_i s_i X}{Y_{si} (K_{si} + s_i)} \tag{14}
\]

Here \( \mu_{\text{mij}} \) is the maximum specific growth rate on \( S_i \) and \( K_{si} \) is a saturation constant. Equation (14) differs from a classic Monod equation by including \( e_i \), whose rate of change is simply the difference between its synthesis and degradation rates:

\[
\frac{de_i}{dt} = r_{E_i} - (\beta_i + r_x) e_i \tag{15}
\]

The first term represents synthesis and the second is for degradation.

At this point the cybernetic approach invokes its central concept. The activities of the key enzymes are regulated by a set of cybernetic variables \( \lambda^{s}_{ij} \) such that they are proportional to the returns from the respective enzymes. If \( r_{ij} \) is the rate of formation of the \( i \)-th product from the \( j \)-th resource (or substrate) then:

\[
\lambda^{s}_{ij} = \frac{r_{ij}}{\max_{k}(r_{ik})} \tag{16}
\]
Equation (16) applies to “substitutable” substrates, that is, where any one of a set of substrates may contribute to a synthetic pathway for a product. For “complementary” substrates, where each substrate has a choice of pathways,

$$\lambda_{ij} = \frac{r_{ij} / P_{ij}}{\max_k (r_{kj} / P_{kj})}$$  \hspace{1cm} (17)

Expounding on these fundamental ideas, many workers have expanded the cybernetic approach and applied it to different systems. Their studies, like Kremling et al. (2000, 2001) examples of signal-oriented modeling, expose the value as well as some weaknesses of cybernetic descriptions.

By incorporating regulatory controls that enable the cells to utilize information gained from experience and thereby respond intelligently to external conditions, cybernetic modeling is able to overcome the rigidity and the limitations of mechanistic modeling. It has successfully portrayed lag phase behavior, cellular responses to changes in dilution rate, and both diauxic and triauxic growth, features which have been difficult to describe by mechanistic methods (Bapat et al., 2006; Namjooshi and Ramkrishna, 2001; Ramkrishna, 2003). However, as mentioned above, the cybernetic approach too has weaknesses. One weakness is that quite complex models may be required to describe adequately metabolic dynamics of multi-cellular systems, especially under non-ideal conditions. A second difficulty is the inability sometimes to identify a unique cybernetic goal, thus creating uncertainty about the cellular response itself. Recent reviews (Patnaik, 2000, 2001a, 2008) have discussed these aspects in detail and suggested combining cellular intelligence with other methods.

**ARTIFICIAL INTELLIGENCE APPROACH**

The successes of artificial intelligence (AI) methods in different disciplines and particularly in remote sensing and control, together with the difficulties of obtaining rapid on-line acquisition of intra-cellular data perhaps motivated the use of AI for microbial processes. Many recent applications provide a testimony to the value of AI in cellular systems.

AI was initially employed in microbial cultures for broadly two purposes (Schugerl, 2001). One is for estimations of time-dependent variables that are difficult or/and expensive to monitor by instrumental methods. The second class of applications was for bioreactor control. The latter use of course required on-line data, provided either by AI methods or by sensory hardware, but it depended also on good models of biological processes. However, it is often difficult to formulate mathematical models that are sufficiently simple, accurate and flexible to be useful under realistic conditions. This difficulty generated a third class of applications of AI, for bioprocess modeling and optimization, and it has also been a driving force for cybernetic models.

The early applications of AI have been reviewed by Patnaik (1998) and by Lubbert and Simutis (1998), whereas more recent work has been discussed by Komives and Parker (2003). These applications have employed different AI methods, notably artificial neural networks, fuzzy logic and genetic algorithms. Often two or more techniques have been used in conjunction, sometimes combined also with classical mathematical models for certain features.

Neural networks seem to be the most favored method to represent microbial processes. They have been used to portray both cell growth and related dynamics (Acuna et al., 1998; Valdez-Castro et al., 2003) of microbial behavior in bioreactors affected by incomplete mixing of the broth and the inflow noise from the environment (Patnaik, 2002), and for early detection of different types of process faults (Vora et al., 1997). However, neural networks, being essentially “black box” devices, suffer from weak physiological links with the biological process, thereby creating difficulties under nonideal conditions (Chen and Rollins, 2000). Other AI methods such as fuzzy logic and genetic algorithms avoid some of these problems but have others of their own.

For instance, the choice of the membership function in fuzzy logic or the fitness function in genetic algorithms is not always known uniquely, and competing candidates may perform equally well within the experimental errors. Nevertheless, both methods have been successful in specific situations. Arnold et al. (2002) study of the adaptation of microorganisms to an inhibiting factor in an industrial brewing process provides a good practical example of the usefulness of fuzzy logic in formalizing the intuitive knowledge of skilled vinegar brewers. Hodgson and coworkers (2005) used genetic programming to reach the interesting conclusion that, for *Saccharopolyspora erythrea* cultures, constrained mathematical forms were superior to flexible unconstrained models, even though no prior knowledge of the fermentation was used.

The strengths and the weaknesses of individual AI methods suggest the possibility of combining them such that the composite model minimizes the overall weakness and/or capitalizes on the strengths. Diverse applications illustrate the validity of this approach for different microbial systems. In one of the early studies, Ye et al. (1994) coupled fuzzy logic with a feed-forward neural network to control β-galactosidase production by recombinant *E. coli*. More complex devices were employed by Coleman et al. (2003) and Fellner et al. (2003). The former maximized the production of green fluorescent protein by *E. coli* through a combination of decision-free analysis, a neural network and a hybrid genetic algorithm. In a novel hybrid network, the latter authors introduced a fuzzy node, a differential equation node and chemometric node into a back propagation network to obtain on-line estimates of diacetyl alcohol in a brewery fermenter.
All the investigations cited so far have used one or more AI methods but no mechanistic model. While avoiding the weaknesses of mechanistic modeling, they have also lost its physiological relevance. Recent work by Galvanauskas et al. (2004) and by Patnaik (2003a, 2006b) has highlighted the benefits of integrating some mechanistic information with AI methods. Patnaik (2003a, 2006b) has also shown how neural networks may be designed to filter the inflow noise into a bioreactor, an important function for industrial applications.

The basic structure of a hybrid model is portrayed in Figure 6. A set of mechanistic models (e.g. Monod kinetics or substrate inhibition equations) and one or more AI models contribute to the complete kinetic description. Their outputs are fed to a bioreactor model, which may be a set of macroscopic equations or purely AI descriptors or a combination. Figure 6 incorporates Galvanauskas et al. (2004) recommendation to use a weighted combination of the kinetic components but in a more generalized manner. \( \alpha_j \) here is the weight assigned to the j-th kinetic model; obviously \( \sum \alpha_j = 1 \). Their values may be determined iteratively.

These composite (or hybrid) neural models still do not account for intra-cellular noise. However, as discussed earlier, noise inside the cells has a vital impact on the metabolic processes (Kaern et al., 2005; Kitano, 2004a; Raser and O’Shea, 2005; Stelling et al., 2004). Moreover, since substrates supplied from outside participate in metabolic reactions, external noise enters the cells and interacts with internal noise. Although the nature of these interactions are not yet clear, we do have the interesting information that neither source of noise per se is entirely detrimental and optimal filtering of either, and preferably both, can indeed enhance cellular functions (Andrews et al., 2006; Patnaik, 2006b; Rao et al., 2002). These observations and the benefits of combining AI methods with equation-oriented models presents the possibility of designing composite architectures that optimally blend cellular intelligence, artificial intelligence and mechanistic models in a framework that links intra-cellular and extra-cellular processes. This concept is discussed below.

**CONCEPTUAL DEVELOPMENT OF A COMPREHENSIVE MODEL**

Kaern et al. (2005) have provided an elegant exposition of the different sources of noise in a microbial culture and their inter-relationship. Their concept is captured diagrammatically in Figure 1, where the influx of external noise (from the environment) has been added. Environmental noise enters through feed streams, permeates the culture broth and penetrates the cells through the diffusional transfer of substrates. Within the cells, external noise encounters intra-cellular noise. Of the two types of intra-cellular noise, extrinsic noise contributes more substan-
tially to stochasticity in gene expression than intrinsic noise (Elowitz et al., 2002; Kaern et al., 2005; Raser and O’Shea, 2005; Stelling et al., 2004)]. Whereas intrinsic noise causes differences between reporter genes in a single cell, extrinsic noise creates differences between cells. The latter thus has a direct effect on the behavior of a population of cells in a bioreactor and is therefore significant for bioreactor dynamics.

Kaern et al. (2005) and other analysts of cellular noise (Elowitz et al., 2002; Raser and O’Shea, 2005; Rao et al., 2002; Thattai and van Oudenaarden, 2004) concentrated on fluctuations in cellular components and did not consider the impact of environmental noise on these processes. On the other hand, Patnaik (2002, 2003a, 2006b) and others (Coleman et al., 2003; Galvanauskas et al., 2004; Gadkar et al., 2005; James et al., 2002) studying bioreactor problems focused on observable macroscopic variations, either without intra-cellular detail or by lumping molecular-level fluctuations into simple mathematical descriptions. Both approaches have value in their respective spheres but neither is complete, so here we will try to evolve a way to concatenate them to develop a comprehensive description of microbial processes.

The development of even a conceptual strategy needs to recognize the central role of yet another source of fluctuations that interfaces those at the cellular level with macroscopic fluctuations induced by external variations. This is the binding noise discussed by Andrews et al. (2006). Although their work pertains to chemotaxis, the idea and their model should be applicable to any cellular system. The mechanism of chemotaxis involves the binding of chemical ligands to receptor clusters protruding from the surfaces of cells. Since small molecules in low concentrations are involved, stochastic fluctuations are associated with this process (Kaern et al., 2005; Raser and O’Shea, 2005). Andrews et al. (2006) argued that bacterial cells possess an inherent mechanism to filter this noise optimally. To model this mechanism, they relied on Yi et al.’s (2000) observation that the presence of an integral feedback system (Figure 7a) imparts robust adaptation to the cells. Such a mechanism may be modeled by a differentiator in series with a low-pass filter (Figure 7b). In a more general context of any cellular reaction system, the ligand-receptor binding may be replaced by the binding of substrate molecules to active sites on the cell surface, thus retaining the validity of Andrews et al.’s (2006) depiction.

Feedback circuits are the most common device to regulate noise at the genetic level. Negative feedback attenuates fluctuations whereas positive feedback amplifies them. While many examples of feedback regulation in biological systems are known and many of their mechanisms have been identified (Kaern et al., 2005; Raser and O’Shea, 2005; Rao et al., 2002; Thattai and van Oudenaarden, 2004), the construction of gene circuits that impart the desired feedback features is still at an early stage. However, Becskei et al. (2001) designs of negative (Becskei and Serrano, 2000) and positive (Becskei et al., 2001) feed-back modules in E. coli to control the fluctuations of a green fluorescence protein used as a marker of gene expression indicate the feasibility of synthetic gene circuits. Interestingly, simple negative feedback corresponds to a low-pass filter, which also forms part of an integrated feedback model (Andrews et al., 2006). This similarity between the two main modes of cellular feedback stabilization makes it convenient to use similar noise filtering methods.

The macroscopic broth in which the cells are immersed is subject to noise from the environment as well as dispersion in the broth itself. Although environmental noise, carried mainly by inlet streams, may be reduced by classical algorithmic filters such as the extended Kalman filter, the low-pass Butterworth filter and the cum-sum filter, the static nature of these devices restricts their adaptability to varying disturbances, and thus undermines their effectiveness. Neural networks as filtering devices perform much better (Gadkar et al., 2005; Patnaik, 2001b, 2002). However, these networks may be difficult to train, have limited extrapolation capability and have little organic connection with the metabolic processes. These limitations have led to the development of hybrid filters that combine neural networks for some variables with algorithmic filters for others (Galvanauskas et al., 2004; Hodgson et al., 2004; James et al., 2002; Patnaik, 2006b).

The development of hybrid filters seems to have been motivated by the success of similar modules for bioreactors per se. While good mixing can be achieved in small laboratory-scale bioreactors, this is difficult in large reactors. The presence of spatial gradients (Larsson et al., 1996; Liden, 2001) and the limitations as well as the

![Figure 7](image-url)
Figure 8. Flow sheet of a conceptual composite model for a microbial system. The model includes both intra-cellular and extra-cellular processes. Previous studies suggest the following representations for the components: Box 1 = auto-associative neural network, Box 2 = Differentiator + Low-pass filter, Box 3 = Optimization algorithm or AI system, Box 4 = Feed-forward neural network, Box 5 = Flexible combination of mechanistic, cybernetic and AI models, and Box 6 = Macroscopic balances + Elman neural network.

Costs of instrumental methods make it impractical to have on-line measurement of all important variables (Mandenius, 2004; Sonnleiter, 2000). Off-line measurements can be too slow for rapid data feedback and control, especially in response to disturbances. Moreover, the time-varying nature of external noise and of dispersion in the broth often invalidate conventional kinetic equations developed on the basis of laboratory-scale observations. As a result control strategies that require a good model of a bioprocess have limited validity in nonideal situations. Hybrid models overcome these weaknesses by enabling rapid on-line estimations of variables that are difficult to measure and by implementing "intelligent" control policies.

Most hybrid models of microbial processes (Coleman et al., 2003; Galvanauskas et al., 2004; James et al., 2002; Lubbert and Jorgensen, 2001) have used combinations of one or more AI methods with mathematical equations. However, both AI and equation-oriented descriptions ignore internal regulatory controls that cybernetic models recognize. Therefore Patnaik (2006c) recently proposed a conceptual approach that models bioreactors through arrays of neural networks, cybernetic models and AI models. He did not, however, include either binding noise or genetic noise. Since these are not isolated from each other or from extra-cellular noise and dispersion, a more holistic model should embrace all of these. Figure 8 presents a flow sheet for such a model.

Inflow streams are first filtered to prune environmental noise present in them (box 1). This filter generates the same output variables as it receives but with reduced noise. An auto-associative neural network is germane to this requirement and many applications (Patnaik, 2001b, 2003a, 2006b) have shown that it is superior to algorithmic filters and other neural configurations. On entering the cytoplasm, the substrate molecules bind to active sites on the cell surfaces; box 2 in Figure 8 contains the binding noise filter, which may be represented by a differentiator coupled to a low-pass filter (Andrews et al., 2006; Rao et al., 2002). The cellular reactions, leading to biomass growth and the formation of products, begin after the binding process. This is a critical phase, both biologically and for model building. Kinetic equations derived from observations of ideal laboratory-scale fermentations often cannot be translated directly to large non-ideal bioreactors (Liden, 2001; Lubbert and Jorgensen, 2001). AI and cellular intelligence (cybernetic) models provide more accurate, flexible and faithfull descriptions. Since each method has strengths and weaknesses, a judicious combination of more than one approach is often recommended (Arnold et al., 2002; Coleman et al., 2003; Galvanauskas et al., 2004; Hodgson et al., 2004; James et al., 2002; Patnaik, 2003a; Ramkrishna, 2003). The sequence of blocks labeled (5) illustrates this idea. For generality, the microbial process is considered to have an arbitrary number of ‘n’ variables, each of whose rates of change is described by an appropriate cellular model.

The bioreactor itself may be represented by either macroscopic mass balances or by neural networks or by a combination of the two. If neural networks are used, a recurrent network of the Elman (1990) type turns out to be the best. The reason is that an Elman network has internal feedback loops that mimic the internal recirculation streamlines in a reactor with finite dispersion (Patnaik, 2003a, 2006c). The entire process is under neural control (box 4) because this is more efficient than adaptive PID control (Gadkar et al., 2005; Hisbullah et al., 2002; Patnaik, 2003a). Box (3) is a critical component in that it compares input and output data and thereby adjusts the settings of the controller dynamically. For a PID controller this involves manipulating the gain and the integral and differential time constants, whereas a neural controller requires adjustments of the weights associated with the inter-neuron signal flows. Although an AI routine may be employed for the optimizer, algorithmic multivariable optimization methods may also be acceptable if the variables of interest do not have widely different dynamics.
CONCLUDING REMARKS

Owing to the complexity of the processes involved, quantitative descriptions of cellular systems functioning in realistic situations have focused on either microscopic intra-cellular phenomena or a macroscopic observable behavior. Either approach generates workable models that are useful for their intended limited purposes but not beyond. Macroscopic models account for dispersion in the culture broth, transport processes and the inflow of noise from the environment, but they tend to ignore or simplify metabolic detail, intra-cellular controls and noise within the cells. Models focused on the cells per se describe the latter features in detail while lumping or ignoring their macroscopic manifestations.

Given that cellular processes differ widely, and even a given process behaves differently under different conditions, it is worthwhile to develop comprehensive descriptions that incorporate both intra- and extra-cellular processes under realistic conditions. The resulting model(s) may then be tailored for specific applications. The versatility and usefulness of composite models are suggested by examples of the use of two or more descriptive methodologies to simulate and optimize microbial processes. These approaches are broadly of three types – mechanistic (or equation-oriented), cellular intelligence (or cybernetic) and artificial intelligence.

Based on the topologies of the hybrid models that have been used for microbial reactors, a conceptual framework for a comprehensive composite model is devised. The flow sheet for such a model accommodates (a) fluid dispersion in the broth, (b) environmental noise, (c) genetic noise and (d) ligand or substrate binding noise in a flexible manner. Each phenomenon may described by a suitable modeling approach and, according to each application, one or more of them may be simplified or dispensed with or assigned a suitable weightage. This flexibility also allows the relative contributions of different processes to the overall behavior of a microbial reactor to be adjustable on-line, a feature that is useful when flux distributions or product patterns or morphology or other relevant characteristics of the cells change as a fermentation progresses.

REFERENCES

Kremling A, Bettenbrock K, Laube B, Jahreis K, Lengeler JW, Gilles ED


