SRC-TR-87-60

CHEMICAL PROCESS SYSTEMS LABORATORY

Modeling of Biological Systems with Time Delays

by

Nam Sun Wang
Gregory Stephanopoulos

RESEARCH REPORT

CHEMICAL PROCESS SYSTEMS ENGINEERING LABORATORY

MODELING OF BIOLOGICAL SYSTEMS

WITH TIME DELAYS

Nam Sun Wang and Gregory Stephanopoulos

A CONSTITUENT LABORATORY OF
THE SYSTEMS RESEARCH CENTER

THE UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742

MODELING OF BIOLOGICAL SYSTEMS

WITH TIME DELAYS

by

Nam Sun Wang
Department of Chemical and Nuclear Engineering
University of Maryland
College Park, MD 20742
and
Gregory Stephanopoulos
Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

September 11, 1986

KEYWORDS

Time delay, Kernel, Modeling, Integro-differential equations.

SUMMARY

Unlike chemical reactor dynamics, microbial behavior depends not only on the present state of the environment surrounding a microorganism but, more importantly, on its past history, as well. Herein lies a major obstacle in the modeling of a biological process with a simple set of equations. By incorporating a culture's past history in the form of a time-delay kernel, a novel approach to bioprocess identification and modeling is formulated. A time delay kernel is included in the state equations, and a generalized method the mathematical simplification via the transformation of an integro-differential equation to a set of first order ODE's is developed. The resulting model possesses the combined advantages of the simplicity of an unstructured, lumped-parameter model and the power of a complex structured model. The experimental determination of the kernel will be discussed, with an emphasis on the on-line parameter estimation, control, and optimization of a biochemical reactor.

I. INTRODUCTION

It has been pointed out[1,2] that two important problems in the optimal design and operation of a biological reactor are the lack of reliable biological sensors and the lack of simple mathematical models with satisfactory predictive capability. The sensor inadequacy is especially accute in the areas of continuous measurement of cell mass and substrate/product concentrations, which are among the most fundamental state variables in nearly all fermentation systems. The relatively poor state of instrumentation means that the current measurements are discrete in time and frequently contain a high level of noise which must be filtered out before they are to be used to control a bioreactor.[3,4,5,6]

As far as the existing models are concerned, they are either inadequate during transient operation (lumped parameter models), very complicated and time consuming for control and optimization purposes (single-cell models), or contain a large number of directly indeterminable parameters (structured models) for any practical application. Despite significant modeling efforts, simple, descriptive , and easy to construct models are not yet available.

One of the main purposes of advance measurement and modeling capabilities should be the satisfactory control of biological processes. This control action should be viewed more as a mechanism to safeguard the process against various types of disturbances rather than as a mean of improving performance. To be sure, there are situations in which the performance of a process can be dramatically improved with the proper control but these cases are less likely to be found in bioprocesses aiming at the volume production of fuels and chemicals.

This paper attemps to address mainly the second problem regarding the modeling of a biochemical reactor. A new approach to bioprocess identification and modeling will be outlined. The proposed approach considers the effect on rates

and yields of not only the present state of the system but also the previous history through the concept of a kernel integral. The set of the resulting integro-differential equations is then shown to be equivalent to a set of first order ordinary differential equations representing a generalized structured model. These simple ordinary differential equations then can be relatively easily manipulated with the well developed mathematical techniques to yield insightful information on the dynamics of the system, including the analysis of the stability of steady states, etc. Furthermore, size reduction techniques are outlined which can lead to a low dimensional, directly observable model while preserving at the same time the biological significance of various parameters.

II. NEW MODELING APPROACH

Mathematical models are needed for control purposes and bioreactor design. They are the condensed version of our knowledge about a system, and their sophistication can vary widely. A useful model should be properly balanced with respect to its mathematical complexity and its ability to capture the essential features for the intended purpose. It should also be simple enough to permit direct determination of its key parameters by performing feasible experimental procedures. The validity of a complex model is especially questionable when it contains a large number of parameters whose values cannot be experimentally evaluated. The success of models in engineering has always depended on the valid use of approximations and assumptions in reducing the complexity of the real world to simple and manageable mathematical abstraction, and biochemical engineering is no exception in this respect.

In the following section, a modeling approach is introduced which combines the simplicity of an unstructured model with the power of a complex structured model. The essence of the approach is the inclusion of a time delay kernel in the equation

describing the dynamics of a bioreactor. As an introductory example, consider the familiar case of a continuous bioreactor operation modeled by a lumped-parameter two-state-variable model, namely:

$$\frac{dx}{dt} = -Dx + \mu(s)x \tag{1}$$

$$\frac{ds}{dt} = D(s_f - s) - \frac{1}{Y_s}\mu(s)\times \tag{2}$$

where we assume that the specific growth rate, μ, of biomass, x, is a function of the limiting substrate, s. The above model, and for that matter almost all other models presently in use, states that the behavior of the biomass-substrate system depends only on the present state, and there is no provision for the past history of the microorganism. It has been recognized for a long time, however, that the observed response of a cell population at a certain time instant is the composite result of various biological processes that were initiated at different time instants in the past as a response to the instantaneous environmental conditions prevailing at each particular time. These various processes result in a present overall specific growth rate that can be described with the introduction of a time delay kernel, $k(t,h)$, in the specific growth rate:

$$\frac{dx}{dt} = -Dx + \left[\int_{-\infty}^{t} \mu[s(h)]k(t,h)dh\right] x \tag{3}$$

$$\frac{ds}{dt} = D(s_f - s) - \frac{1}{Y_s}\left[\int_{-\infty}^{t} \mu[s(h)]k(t,h)dh\right] x \tag{4}$$

The idea of a variable's dependence on its past history has been in existence for quite some time.[7,8,9] In ecological studies, the interaction of prey-predator has been described by Volterra models, which include a kernel associated with one of the states of the system. As shown later, our handling of the kernel is more general in the sense that the shape of the kernel is not restricted. For a linearized time-invariant system, k no longer depends on t and the integration variables h separately

but on the difference $t - h$.

$$\frac{dx}{dt} = -Dx + \left[\int_{-\infty}^{t} \mu[s(h)]k(t-h)dh\right] x \tag{5}$$

$$\frac{ds}{dt} = D(s_f - s) - \frac{1}{Y_s}\left[\int_{-\infty}^{t} \mu[s(h)]k(t-h)dh\right] x \tag{6}$$

Non-dimensionalization can be carried out to simplify the above equations without any loss of generality. If time is scaled with reference to D^{-1} and concentrations are scaled with reference to s_f, then we have:

$$\frac{dx}{dt} = \left[-1 + \int_{-\infty}^{t} \mu[s(h)]k(t-h)dh\right] x \tag{7}$$

$$\frac{ds}{dt} = 1 - s - \frac{1}{Y_s}\left[\int_{-\infty}^{t} \mu[s(h)]k(t-h)dh\right] x \tag{8}$$

The kernel $k(t)$ is usually referred to as the impulse response function and can be interpreted as a weighing factor as shown schematically in Figure 1. Since it can be generally assumed that future states have no effect on the present, $k(t)$ can be implicitly set to zero for $t < 0$. Note that, strictly speaking, $k(t)$ is not a time delay probability distribution function and $k(t) < 0$, i.e. negative weighing, is possible. The μ in the integrand of Equations (7) and (8) is the specific growth rate would have realized if the system operated at a steady state characterized by the corresponding value of s for a prolonged period of time; it is the true specific growth rate in the absence of time delay effects. The presently observed apparent value of the specific growth rate is given by the integral of Equations (7) or (8) and can be conceptualized as a string of impulses each of which is felt by the system over a period of time according to the impulse transfer function. The main questions, of course, are how can one determine the appropriate kernel form and what is the biological significance of the latter. These two points will be discussed later.

Various possibilities exist for the functional form of $k(t)$. See Figure 2. One can set $k(t)$ to be a delta function, $\delta(t)$, meaning that both the future and the past

have absolutely no weight on the specific growth rate and that the present instant carries all the weight. The integral $\int_{-\infty}^{t} \mu[s(h)]k(t-h)dh$ reduces to $\mu[s(t)]$ in this case, and Equations (7) and (8) reduce to the conventional unstructured model of Equations (1) and (2).

Another possibility is to assume that there is a fixed time lag in the response of the system, i.e. $k(t) = \delta(t-\tau)$, meaning that the specific growth rate depends on the substrate concentration at a discrete time instant τ units before the present. The state equations in this case are reduced to:

$$\frac{dx}{dt} = [-1 + \mu[s(t-\tau)]]\, x \tag{9}$$

$$\frac{ds}{dt} = 1 - s - \frac{1}{Y_s}\mu[s(t-\tau)]x \tag{10}$$

Analysis of this case can be performed by using the theories of differential-difference equations.[10] The relatively simple system of Equations (9) and (10) can be successfully analyzed, but, because the mathematical theories of differential-difference equations are not as well developed as ordinary differential equations, some problems may arise in the integration and general analysis of this type of differential-difference equations especially in slightly more complicated systems.

A more general approach is to express an arbitrary function $k(t)$ in terms of a series of base functions which permit the transformation of the integro-differential equations into a set of simple first order equations. This is accomplished by approximating an analytical function $k(t)$ as a summation of exponential distribution functions of order m or less:

$$k(t) = a_0 k_0(t) + a_1 k_1(t) + a_2 k_2(t) + \cdots + a_m k_m(t) \tag{11}$$

where the general expression for the nth exponential distribution function is:

$$k_n(t) = \begin{cases} \frac{T_{-1}}{n!}\left(\frac{t}{T}\right)^n e^{\frac{-t}{T}} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{12}$$

The first two exponential distribution functions are sometimes used in ecological studies and they have special names.

$$n = 0 \quad k_0 = T^{-1} e^{\frac{-t}{T}} \quad \cdots\cdot \text{weak generic delay} \tag{13}$$

$$n = 1 \quad k_1 = T^{-2} t e^{\frac{-t}{T}} \quad \cdots\cdot \text{strong generic delay} \tag{14}$$

Some of the properties of the exponential distribution functions are shown in Figure 3, and the first few of these exponential distribution functions are shown in Figure 4. Note that if these functions are normalized with respect to the average delay, $(n+1)T$, then one can see that the peak at the average delay becomes higher and narrower as n increases. (See Figure 5.) It can be shown that as $n \longrightarrow \infty$, $k_n(t) \longrightarrow \delta(t-\tau)$, where τ is the average delay. In this limiting case, the state equations are again reduced to Equations (9) and (10).

At close inspection, the nth order exponential distribution function is identical to the residence time distribution function of a system of n-CSTR's in series in the modeling of a chemical reactor. Accordingly, if $k(t)$ is expressed as the sum of m exponential functions, the observed specific growth rate at time t, expressed as $y(t) \equiv \int_{-\infty}^{t} \mu[s(h)]k(t-h)dh$, will be the weighed sum of m integrals, $\int_{-\infty}^{t} \mu[s(h)]k_j(t-h)dh \quad j = 1,2,...,m$, i.e.,

$$\begin{aligned} y(t) &= \sum_{j=0}^{m} a_j \left[\int_{-\infty}^{t} \mu[s(h)]k_j(t-h)dh \right] \\ &= \int_{-\infty}^{t} \mu[s(h)][\sum_{j=0}^{m} a_j k_j(t-h)]dh \\ &= \int_{-\infty}^{t} \mu[s(h)]k(t-h)dh \end{aligned} \tag{15}$$

The weighing factors a_j and the delay time constant T are chosen in such a way as to fit the observed transient of the specific system in a shift-up or shift-down experiment. A small value of m usually gives a very satisfactory fit.

Quite significantly, we are not bounded by the limited functional shapes of each individual exponential distribution function. By expressing the kernel as a linear combination of these base functions, any sufficiently smooth continuously differentiable function can be represented if a sufficiently large number of base functions are used. This is because the approach is essentially the same as expanding the function $e^{\frac{t}{T}}k(t)$ by a power series. Theoretically, m could be extended to ∞, but two or three terms should be sufficient under most circumstances in practice. For example, a linear combination of $k_0(t)$ and $k_1(t)$ results in

$$k(t) = (a_0 T^{-1} + a_1 T^{-2} t) e^{\frac{-t}{T}} \tag{16}$$

where $a_0 + a_1 = 1$ so that the kernel is normalized to unity. Some of the shapes of $k(t)$ generated by a combination of these two base functions are shown in Figure 6.

The reason for choosing exponential distribution functions is that they permit easy and elegant transformation of a set of integro-differential equations into a set of simple ordinary differential equations. These exponential distribution functions possess the property that each and every one of them is the solution to the following differential equation:

$$\sum_{i=0}^{n+1} T^i \binom{n+1}{i} \frac{d^i k_n(t)}{dt^i} = 0 \tag{17}$$

with initial conditions:

$$\begin{cases} \frac{d^i k_n(0)}{dt^i} = 0 & i = 0, 1, 2, \cdots, n-1 \\ \frac{d^n k_n(0)}{dt^n} = T^{-(n+1)} \end{cases}$$

For example, $k_0(t) = T^{-1} e^{\frac{-t}{T}}$ satisfies

$$T\frac{dk_0}{dt} + k_0 = 0 \tag{18}$$

with initial condition:

$$k_0(0) = T^{-1}$$

Similarly, $k_1(t) = T^{-2}te^{\frac{-t}{T}}$ satisfies

$$T^2\frac{d^2k_1}{dt^2} + 2T\frac{dk_1}{dt} + k_1 = 0 \tag{19}$$

with initial conditions:

$$k_1(0) = 0$$

$$\frac{dk_1(0)}{dt} = T^{-2}$$

The above properties of the exponential distribution functions can be used to eliminate the kernel from the integro-differential equations (7)-(8) and convert them into a larger, but mathematically identical, set of first order ordinary differential equations.

If we treat the integral containing the kernel as a new function $y_n(t)$,

$$y_n(t) \equiv \int_{-\infty}^{t} \mu(h)k_n(t-h)dh \tag{20}$$

then differentiating $y_n(t)$ with respect to t $n+1$ times with the help of Liebnitz's rule yields:

$$\sum_{i=0}^{n+1} T^i \binom{n+1}{i} \frac{d^i y_n(t)}{dt^i} = \mu(t) \tag{21}$$

The set of the resulting differential equations is one order higher than the kernel originally contained inside the integral. This higher order differential equation can be easily transformed into a set of first order differential equations through some well known canonical transformations. Thus, for a simple nth order kernel $k_n(t)$, the integral is transformed to the following set of equations:

$$\begin{aligned} \frac{dy}{dt} &= z_1 \\ \frac{dz_1}{dt} &= z_2 \\ &\cdot \\ &\cdot \end{aligned} \tag{22}$$

$$\frac{dz_{n-1}}{dt} = z_n$$

$$\frac{dz_n}{dt} = T^{-(n+1)}\left[-\sum_{i=0}^{n} T^i \binom{n+1}{i} z_i + \mu(t)\right]$$

Because of the linear properties of the differential and integration operators, a linear combination of more than one base functions of $k_n(t)$ will leave the approach unchanged. Thus if a first order kernel has the form

$$k(t) = (a_0 T^{-1} + a_1 T^{-2} t) e^{\frac{-t}{T}} \qquad (a_0 + a_1 = 1) \tag{23}$$

Then, $y(t) = \int_{-\infty}^{t} \mu[s(h)] k(t-h) dh$ can be converted to a second order ordinary differential equation:

$$T^2 \frac{d^2 y(t)}{dt^2} + 2T \frac{dy(t)}{dt} + y(t) = \mu(t) + a_0 T \frac{d\mu(t)}{dt} \tag{24}$$

which can be further transformed to a mathematically equivalent set of two first order ordinary differential equations.

$$\frac{dy(t)}{dt} = z \tag{25}$$

$$\frac{dz(t)}{dt} = -2T^{-1} z - T^{-2} y + T^{-2} \mu(t) + a_0 T^{-1} \frac{d\mu(t)}{dt} \tag{26}$$

For example, with the kernel of Equation (23), the system dynamic equations (7) and (8) are now:

$$\frac{dx}{dt} = (y-1)x \tag{27}$$

$$\frac{ds}{dt} = 1 - s - \frac{1}{Y_s} yx \tag{28}$$

$$\frac{dy}{dt} = z \tag{29}$$

$$\frac{dz}{dt} = -2T^{-1} z - T^{-2} y + T^{-2} \mu[s(t)] + a_0 T^{-1} \frac{d\mu}{ds}\frac{ds}{dt} \tag{30}$$

$$= -2T^{-1} z - T^{-2} y + T^{-2} \mu(s) + a_0 T^{-1} \frac{d\mu(s)}{ds} - a_0 T^{-1} \frac{d\mu(s)}{ds} s - a_0 T^{-1} \frac{1}{Y_s} \frac{d\mu(s)}{ds} yx$$

Since, as it was mentioned, a first order kernel is usually sufficient in describing bioreactor dynamics, the dependence of the specific growth rate (and other similar culture parameters) on the past history of the culture can thus be described with only two additional differential equations. This increase in the dimensionality of the system is a small price to pay considering the significantly enhanced predictive capabilities of the model.

The dynamics of a chemostat culture in the presence of time delay kernels has been analyzed with the use of linearized stability analysis and bifurcation theories. The full spectrum of dynamic behavior, including damped oscillations (when a 0th order kernel is included) and sustained oscillations (when a 1st order kernel is included) can be predicted. The inclusion of kernels in other variables such as Y_s and x can also be analyzed in a similar manner. Furthermore, product formation, although not considered in this paper, can also be similarly studied. More detailed and complete results on the effect of time delay on the stability, classical process control consideration, and optimal control formulations will be the subject of forthcoming publications.

The experimental determination of the kernel has also been investigated for various transient situations. Shown in Figures 7a and 7b are the computer simulated responses of a biochemical reactor described by Equations (5) and (6). From the noisy transient data of $\mu(t)$ and $y(t)$ when the dilution rate is shifted up from 0.3 hr^{-1} to 0.7 hr^{-1}, the kernel was reconstructed by minimizing the mean square deviation of the $y(t)$ predicted by the kernel away from the observed $y(t)$. The resulting kernel is shown in Figure 7c. Here, $\mu(t)$ is assumed to be the true specific growth rate in the absence of the effects of time delay. Given $s(t)$, this true specific growth rate $\mu(t)$ is obtained, in actuality, from a μ versus s curve constructed from a series of steady state experiments, in which the time delay effects are eliminated.

For the purpose of this simulation, the μ versus s curve is assumed to follow the Monod model; however, it need not be so. Since the frequency response function or the pulse response function can be considered merely as another representation of the impulse response function (i.e. the kernel), sinusoidal or pulse methods can also be utilized to determine experimentally the shape of the kernel. The above example represents the worst case of estimating the impulse response function from a pulse experiment. Much better agreement between the true kernel and the estimated one can be achieved if an impulse can be applied to the system; the agreement is also considerably better if the noise level is decreased. (Of course, the reconstructed kernel coincides with the true one in the absence of noise.)

The above simulation study suggests that the first step in the experimental determination of the kernel is to construct a μ versus s curve through a series of steady state runs. During a transient experiment in which the dilution rate or the feed substrate concentration is shifted up or down, the substrate concentration can be continuously estimated as a function of t, as shown in the previous sections. Furthermore, by referring to the μ versus s curve, $\mu[s(t)]$ can be generated continuously as well. The estimation scheme presented earlier can also be used to provide a continuous estimate of the instantaneous specific growth rate $y(t)$, and, finally, the kernel is generated.

Currently, work is under way in order to determine experimentally the shape of the kernel for a continuous culture of *S. cerevisiae*. Shift-up, shift-down, and sinusoidal perturbation experiments of the dilution rate, the substrate feed concentration, the pH, or the temperature are being performed. The on-line measurements with the aid of the parameter and state estimation algorithms described in the previous section will be used to determined the shape of the kernel function, and this new approach to bioprocess identification and modeling will be tested in terms of

the model's capability in predicting the microbial behavior, including the more general and revealing behavior, such as the occurrence of sustained oscillation, under different conditions.

The difference between a complex structured model and a simple unstructured model is analogous to that between statistical and classical thermodynamics. Whereas a structured model tries to explain the observed phenomena through a large set of differential equations in terms of the more fundamental variables such as the concentrations of various intermediates; unstructured models are usually composed of those variables that can be physically "seen" or "felt" more readily and are, thus, more comprehensible to human minds. The proposed modeling approach herein attempts to retain the general form of an unstructured model so as to facilitate simple physical interpretation of the variables by such familiar terms as the specific growth rate. At the same time, this modeling approach attempts to incorporate only those metabolic intermediates that are important to the dynamics of the system and to reduce the order of a complicated structured model through the analysis of eigenvalue-eigenvector of a linearlized system. How this can be accomplished is briefly outlined below.

In general, a dynamic system (including a structured model) can be described by a set of first order differential equations:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{31}$$

where $\mathbf{x}$ is the state vector and $\mathbf{u}$ is the input to the system. For a system linear in the state variables, the above equation can be written as:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{g}(t) \tag{32}$$

The fundamental-matrix solution to the above differential equation is expressed by

the following Lagrange formula:

$$\mathbf{x}(t) = \int_{-\infty}^{t} \mathbf{K}(t,h)\mathbf{g}(h)dh \tag{33}$$

If the linearization matrix $\mathbf{A}(t)$ is constant, then this solution further reduces to:

$$\mathbf{x}(t) = \int_{-\infty}^{t} \mathbf{K}(t-h)\mathbf{g}(h)dh \tag{34}$$

where $\mathbf{K}$ is the fundamental matrix of Equation (31). Thus, the appearance of a kernel in Equations (3)-(6)is spontaneous; it arises mathematically during the process of solving a set of differential equations. The eigenvalue and eigenvector of the matrix $\mathbf{A}(t)$ can be analyzed to simplify and to reduce the dimension of the system by retaining only the first few most important modes and eliminating the remaining nonsignificant modes.

If the unstructured part of the system (i.e. biomass, substrate, product, etc.) are included in the state variable $\mathbf{x}(t)$, then the state variable can be grouped according to those that appear in the unstructured model ($\mathbf{x}_1(t)$) and those that are contained only in the structured model ($\mathbf{x}_2(t)$).

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{35}$$

The linearization matrix $\mathbf{A}(t)$ and the non-homogeneous forcing function $\mathbf{g}(t)$ can be partitioned similarly:

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{A}_{11}(t) & \mathbf{A}_{12}(t) \\ \mathbf{A}_{21}(t) & \mathbf{A}_{22}(t) \end{bmatrix} \tag{36}$$

$$\mathbf{g}(t) = \begin{bmatrix} \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \end{bmatrix} \tag{37}$$

With this partition, Equation (32) becomes:

$$\frac{d\mathbf{x}_1(t)}{dt} = \mathbf{A}_{11}(t)\mathbf{x}_1(t) + \mathbf{A}_{12}(t)\mathbf{x}_2(t) + \mathbf{g}_1(t) \tag{38}$$

$$\begin{aligned} \frac{d\mathbf{x}_2(t)}{dt} &= \mathbf{A}_{21}(t)\mathbf{x}_1(t) + \mathbf{A}_{22}(t)\mathbf{x}_2(t) + \mathbf{g}_2(t) \\ &= \mathbf{A}_{2}2(t)\mathbf{x}_2(t) + \tilde{\mathbf{g}}(t) \end{aligned} \tag{39}$$

Thus, the unstructured model's equivalent of the structured model described by Equation (31) is:

$$\frac{d\mathbf{x}_1(t)}{dt} = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, t) \tag{40}$$

where $\mathbf{x}_2$ is the delay kernel defined by:

$$\mathbf{x}_2 = \int_{-\infty}^{t} \mathbf{K}_{22}(t,h)\tilde{\mathbf{g}}(h)dh \tag{41}$$

where $\mathbf{K}_{22}(t,h)$ is the fundamental matrix to $\mathbf{A}_{22}(t)$ of Equation (39). As can be seen from the preceding equations, the time delay kernel arises quite natually as a conseqence of reducing a larger set of dynamic equations in a structured model by a smaller set of dynamic equations in a unstructured model.

III. DISCUSSION

Microbial behavior depends not only on the present state of the environment but on past histories as well. This is the main reason for the inadequacy of the simple set of Equations (16) and (17). The dependence of a culture on its past history is manifested in the presence of a lag phase in the beginning of a batch cultivation. It is also present during the transients of continuous fermentors resulting from, among others, a shift-up of nutrient concentration. For example, the diauxic batch fermentation of glucose producing ethanol as the intermediate product exhibits a lag phase before the glucose begins to be consumed; another lag phase is also present before the ethanol is taken up, as shown in Figure 8. In fact, it is this second lag that is most often used to detect the presence of more than one substrates. Figure 9 shows that although glucose concentration was suddenly increasing at 0 hour, the observed apparent specific growth rate did not start to increase until 1.1 hour later. Such lag has often been explained in terms of the need to synthesize the necessary pools of enzymes and intermediates before the rate of substrate utilization is adjusted to the changed conditions. The importance and the presence of time lag have been recognized for many years, and in this paper we have attempted to offer

a simple mathematical means by which the idea of time delay can be incorporated into the existing models without drastically increasing the complexity of the models. Furthermore, by expanding the delay kernel in a series of exponential distribution functions, the integro-differential equations can be easily reduced to a set of first order ordinary differential equations for which the mathematical theories are well developed and various established techniques are available to analyze them.

Much benefit can be derived from the recognition of time delay. It is a well known fact that time delay can cause, among other undesirable problems, serious instability difficulties if it is neglected in a control strategy. Furthermore, an optimal control scheme may not be truly optimal if time delay is not properly considered. Figure 10 shows a hypothetical run in a bioreactor. The data collected during the short transient period after the start up can be used to update the shape of the kernel and other model parameters. Based on the updated model and model parameters and objective functions, an optimal path can be calculated by an on-line computer. Occasionally, deliberate excursions can be introduced to update the kernel and model parameters if they are suspected of gradual changes during a long steady state run. A way in which a simple but powerful model such as the one proposed herein and the state of parameter estimation scheme discussed in the previous sections can be used is in the combined forward and feedback control of a bioreactor. Shown schematically in Figure 11 is an interactive estimation-control optimization scheme in which the on-line measurement on a bioreactor is passed through an estimation-filter block to rid of the noise and to yield a set of on-line estimates for the state variables and growth parameters. These estimates are used as the basis for feedback control as well as for on-line process modeling. The biochemical process is continuously modeled, new values of the model parameters are estimated, and the biological model itself, including the shape of the kernel,. is constantly updated. This can be accomplished by tracking the control history

and comparing the deviation of the actual state away from the predicted values. Although such an ideal scheme does not exist presently, the state and parameter estimation and the new approach to modeling proposed herein is a step toward the realization of such a scheme.

ACKNOWLEDGEMENT

A partial financial support for this work provided by ARCO's Young Faculty Investigator Award is gratefully acknowledged.

REFERENCES

1. Cooney, C. L., Computer application in fermentaion technology – a perspective, *Biotechnol. Bioeng. Symp.*, **9**, 1, 1979.

2. Wang, N. S. and Stephanopoulos, G., Computer applications to fermentation processes, *CRC Critical Reviews in Biotechnology*, **2**, No. 1, 1, 1984.

3. Stephanopoulos, G. and San, K.-Y., Studies on on-line bioreactor identification. I. Theory, *Biotechnol. Bioeng.*, **26**, 1176, 1984.

4. San, K.-Y. and Stephanopoulos, G., Studies on on-line bioreactor identification. II. Numerical and experimental studies, *Biotechnol. Bioeng.*, **26**, 1189, 1984.

5. Grosz, R., Stephanopoulos, G., and San, K.-Y., Studies on on-line bioreactor identification. III. Sensitivity problems with respiratory and heat evolution measurements, *Biotechnol. Bioeng.*, **26**, 1198, 1984.

6. San, K.-Y. and Stephanopoulos, G., Studies on on-line bioreactor identification. IV. Utilization of pH measurements for product estimation, *Biotechnol. Bioeng.*, **26**, 1209, 1984.

7. May, R. M., *Stability and Complexity in Model Ecosystems*, Princeton Univ. Press, Princeton, 1973.

8. Cushing, J. M., *Integrodifferential Equations and Delay Models in Population Dynamics*, Springer-Verlag, Berlin, 1977.

9. MacDonald, N., Time delays in chemostat models, in *Population Dynamics*, CRC Press, Boca Raton, FL, 1982, chapt 2.

10. Bellman, R. and Cooke, K. L., *Differential-Difference Equations*, Academic Press, New York, 1963.

FIGURE CAPTIONS

Figure 1. The interpretation of a time-invariant kernel integral which relates the input to the output of a linear system.

Figure 2. Three frequently used functional forms of $k(t)$: (a) delta function without time delay; (b) delta function with a discrete time delay τ; (c) general distributed time delay. Note the direction of past and future are the reverse of the conventional time plots.

Figure 3. Some properties of the exponential distribution functions.

Figure 4. Exponential distribution function of order n.

Figure 5. Exponential distribution function of order n normalized with respect to the average delay.

Figure 6. Linear combination of the 0th order exponential distribution function, $k_0(t)$, and the 1st order exponential distribution function, $k_1(t)$; $k(t) = a_0 k_0(t) + a_1 k_1(t)$, $a_0 + a_1 = 1$.

Figure 7. (a) Simulated input (i.e. the specific growth rate in the absence of time delay effects) as a function of time in a continuously operated bioreactor described by the state equations (20) and (21) after a shift-up in the dilution rate from 0.3 hr^{-1} to 0.7 hr^{-1}. (Parameters used: $\mu = \frac{0.5s}{0.1+s}$; $s_f = 5.0$; $Y_s = 0.5$; noise level in measurement = 5(b) Simulated output (i.e. the observed specific growth rate containing time delay effects) as a function of time. (Upper smooth curve: the true value of $y(t)$; lower smooth curve: the calculated value of $y(t)$ based on the estimated kernel function of (c). (c) The true and the estimated shapes of the kernel.

Figure 8. Batch fermentation of glucose by *S. cerevisiae*. Note that there are two distinct region of high growth activities.

Figure 9. On-line estimates and off-line measurements in a dilution rate step-up experiment of continuous glucose-limited cultivation of *S. cerevisiae* with ethanol formation. Shift is at 0 hr.[6] (a) Cell biomass concentration. (b) Glucose concentration. (c) Ethanol concentration. (d) Specific growth rate.

Figure 10. Use of transient data for the determination of kernel and model parameters during the start-up of a bioreactor and the subsequent utilization of model in control and optimization.

Figure 11. Block diagram of the measurement- estimation- modeling- optimization- control configuration.

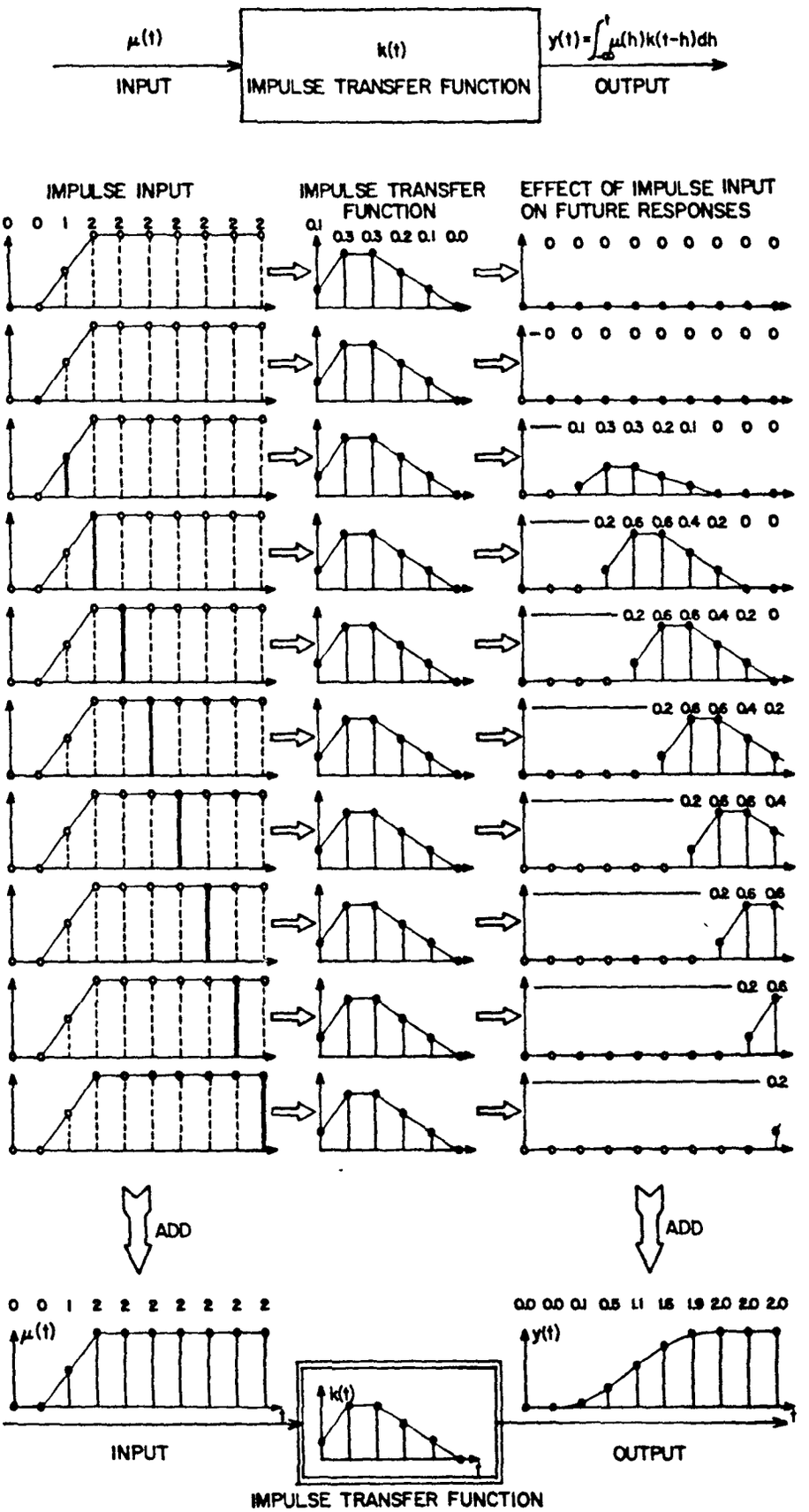

μ(t)
INPUT
k(t)
IMPULSE TRANSFER FUNCTION
y(t)=∫μ(h)k(t-h)dh
OUTPUT
IMPULSE INPUT
IMPULSE TRANSFER FUNCTION
EFFECT OF IMPULSE INPUT ON FUTURE RESPONSES
0 0 1 2 2 2 2 2 2 2
0.1 0.3 0.3 0.2 0.1 0.0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0.1 0.3 0.3 0.2 0.1 0 0 0
0.2 0.6 0.6 0.4 0.2 0 0
0.2 0.6 0.6 0.4 0.2 0
0.2 0.6 0.6 0.4 0.2
0.2 0.6 0.6 0.4
0.2 0.6 0.6
0.2 0.6
0.2
ADD
ADD
0 0 1 2 2 2 2 2 2 2
μ(t)
0.0 0.0 0.1 0.5 1.1 1.6 1.9 2.0 2.0 2.0
y(t)
k(t)
INPUT
OUTPUT
IMPULSE TRANSFER FUNCTION

Figure 1

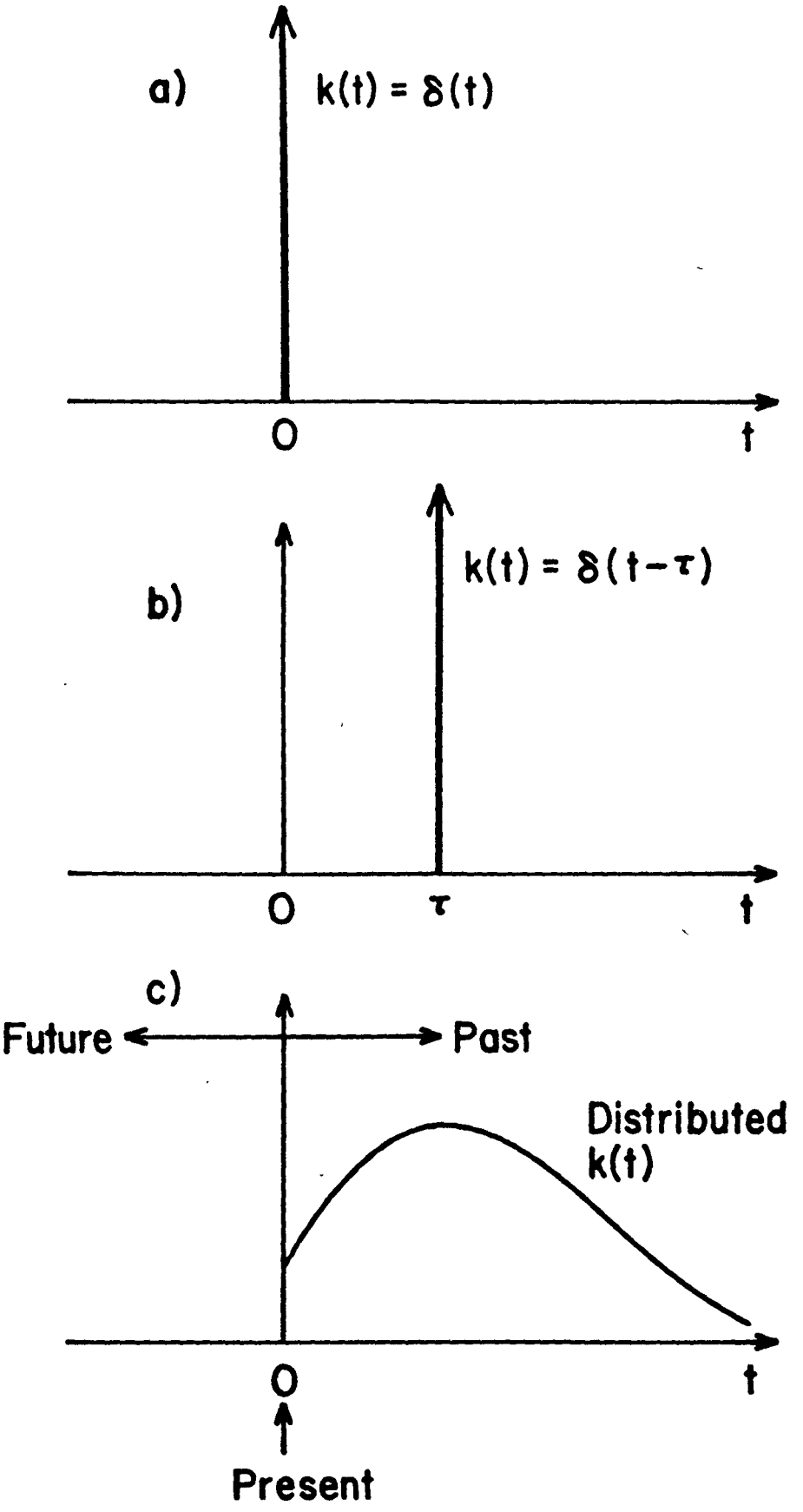
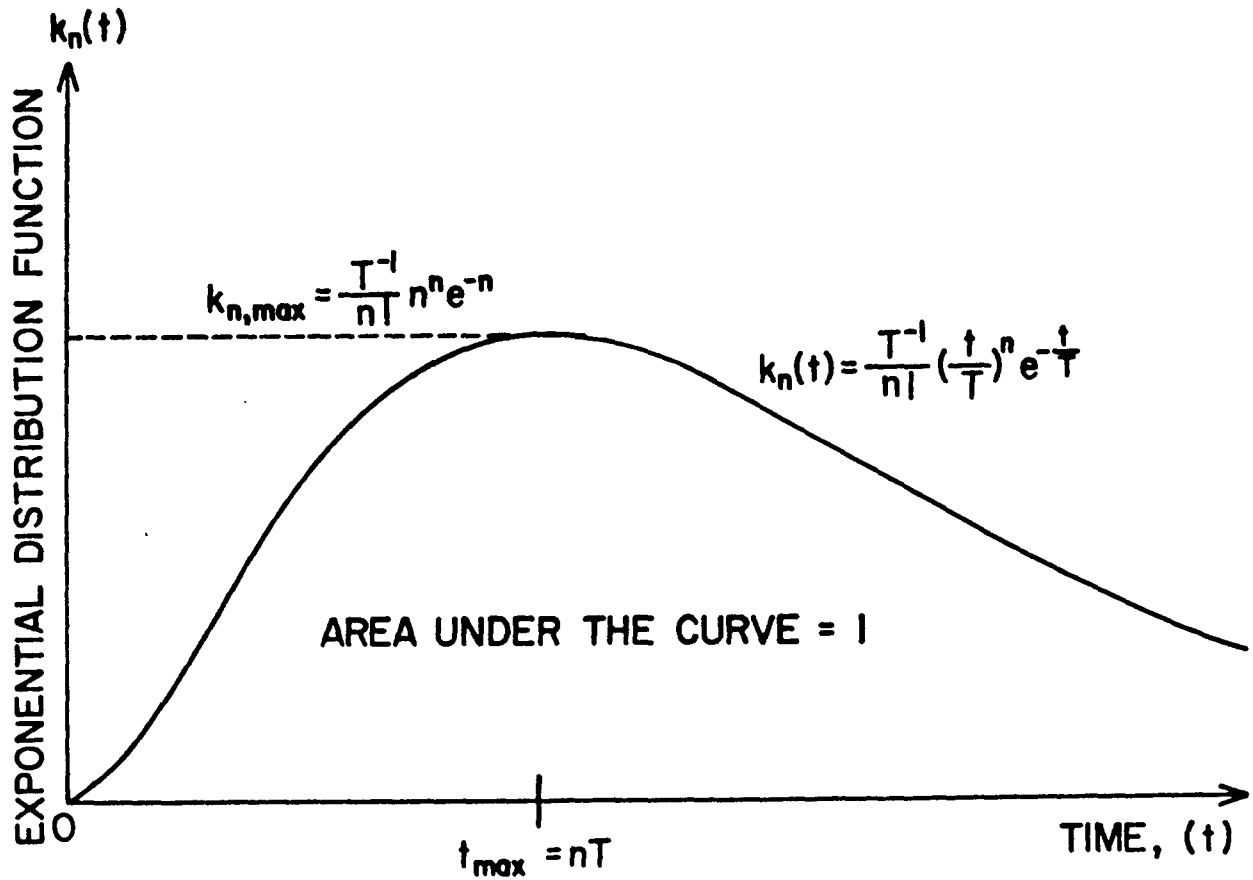


Figure 2



$$\text{ith Moment of } k_n(t) = \int_0^\infty t^i k_n(t)dt = \frac{(n+i)!}{n!} T^i$$

0th Moment = Area under the Curve = 1

1st Moment = Average Delay = τ = $(n+1)T$

2nd Moment = = $(n+1)(n+2)T^2$

Variance of k_n = Average Spread = σ^2 = $(n+1)T^2$ = $\frac{\tau^2}{n+1}$

Figure 3

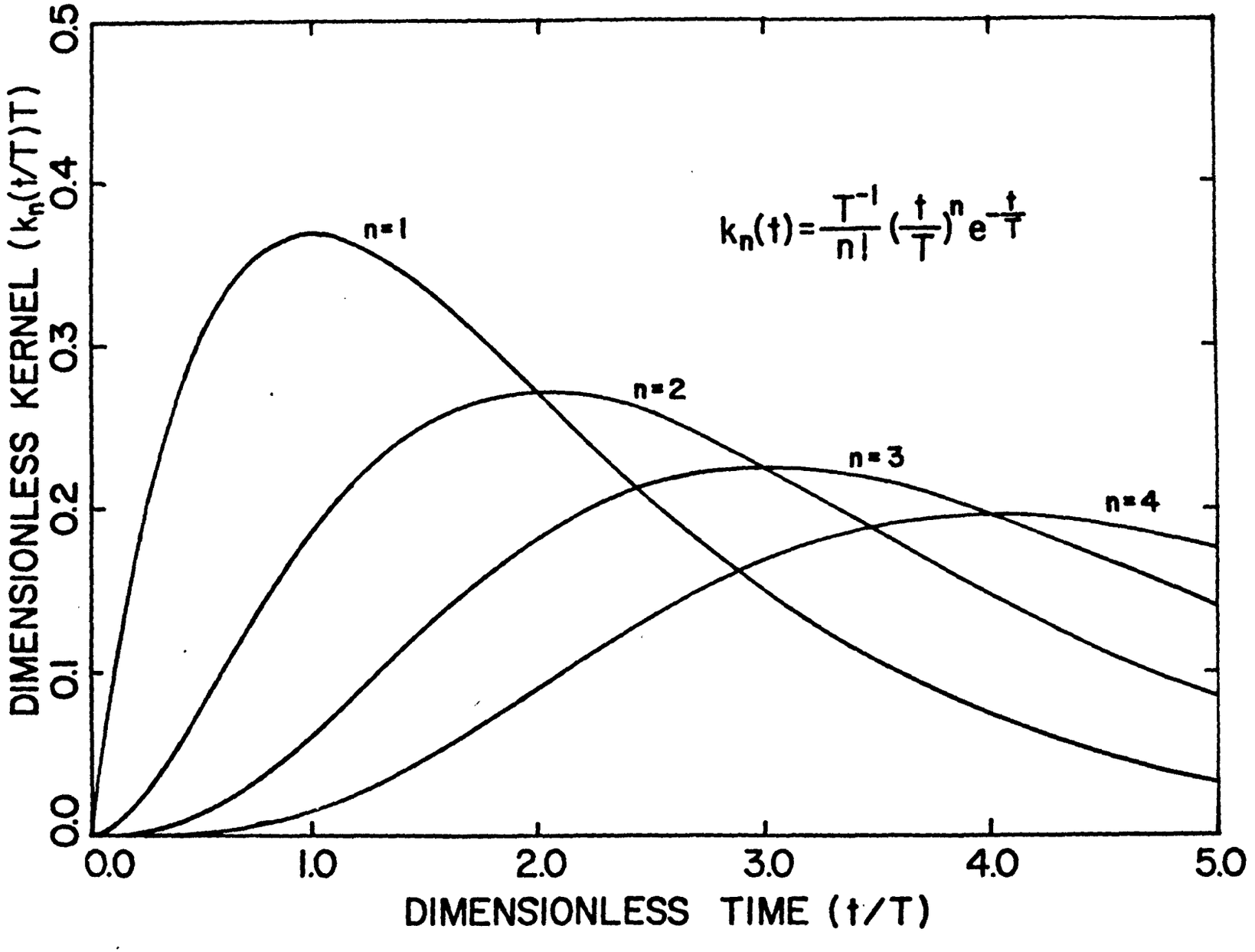

DIMENSIONLESS KERNEL ($k_n(t/T)T$)
0.0
0.1
0.2
0.3
0.4
0.5
0.0
1.0
2.0
3.0
4.0
5.0
DIMENSIONLESS TIME (t/T)
n=1
n=2
n=3
n=4
$k_n(t)=\frac{T^{-1}}{n!}(\frac{t}{T})^n e^{-\frac{t}{T}}$

Figure 4

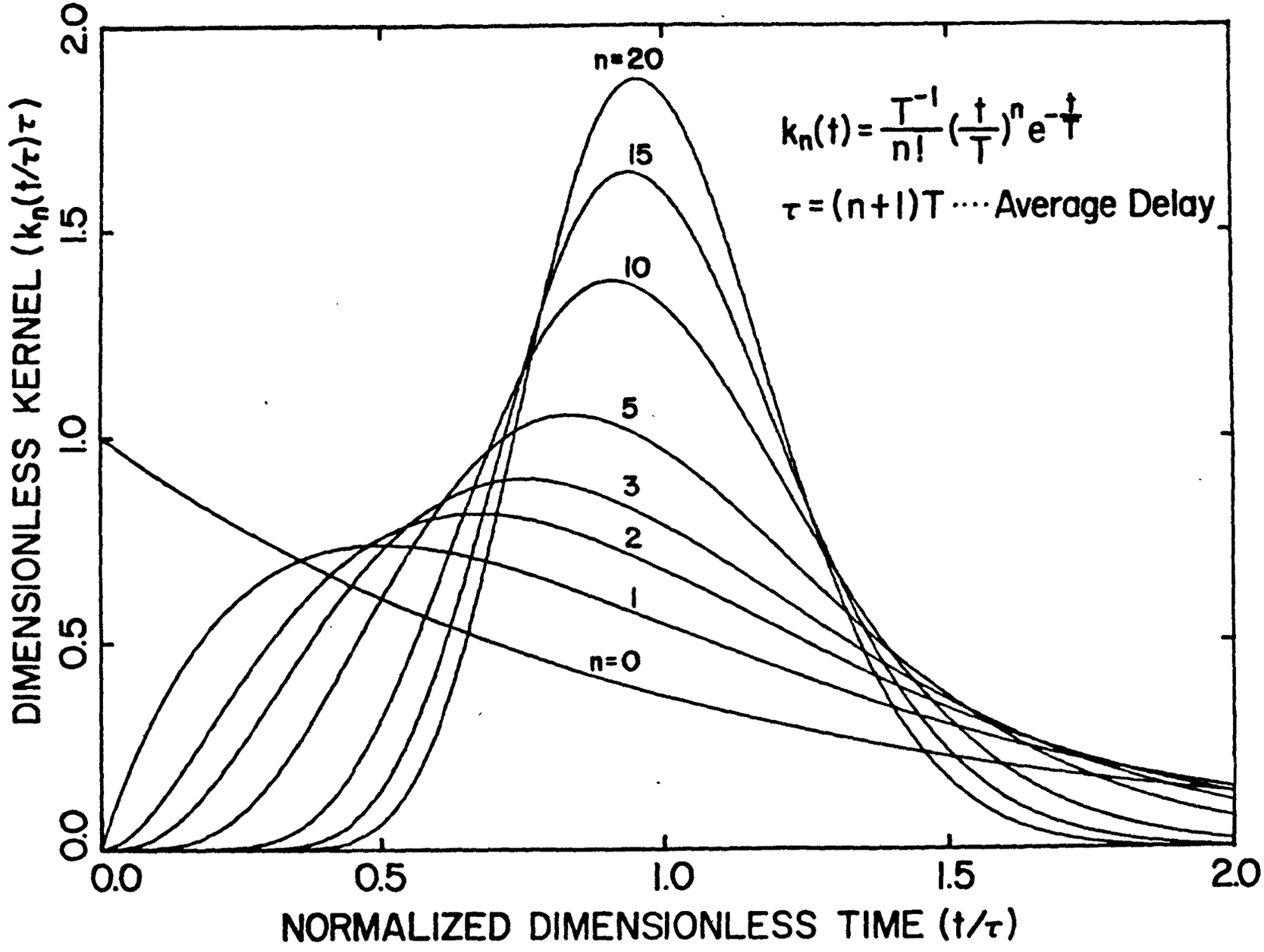

DIMENSIONLESS KERNEL ($k_n(t/\tau)\tau$)
NORMALIZED DIMENSIONLESS TIME (t/τ)
$k_n(t) = \frac{T^{-1}}{n!}(\frac{t}{T})^n e^{-\frac{t}{T}}$
$\tau = (n+1)T$ ···· Average Delay
n=20
15
10
5
3
2
1
n=0
0.0
0.5
1.0
1.5
2.0

Figure 5

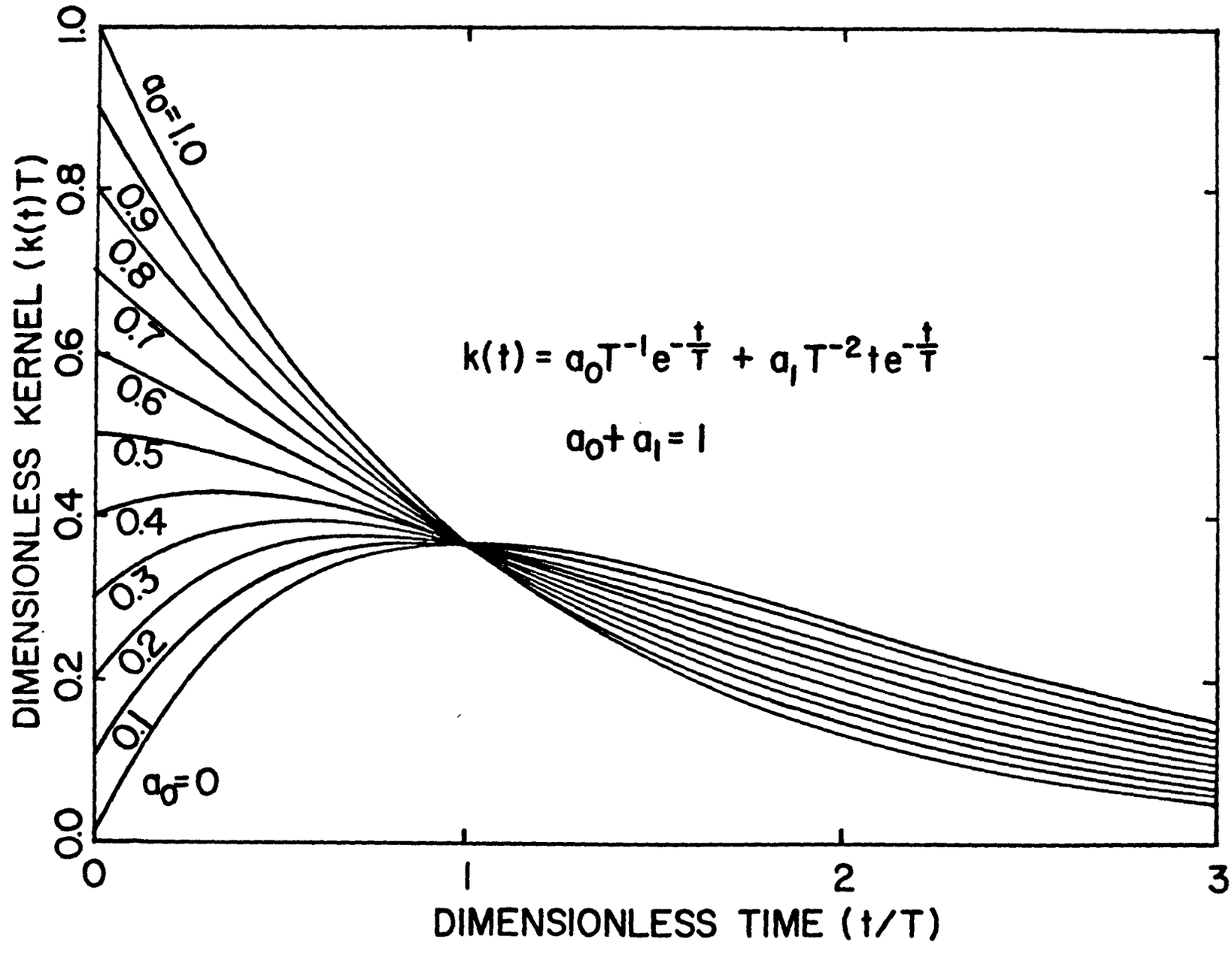

DIMENSIONLESS KERNEL (k(t)T)
DIMENSIONLESS TIME (t/T)
$k(t) = a_0 T^{-1} e^{-\frac{t}{T}} + a_1 T^{-2} t e^{-\frac{t}{T}}$
$a_0 + a_1 = 1$
a_0=1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
a_0=0
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3

Figure 6

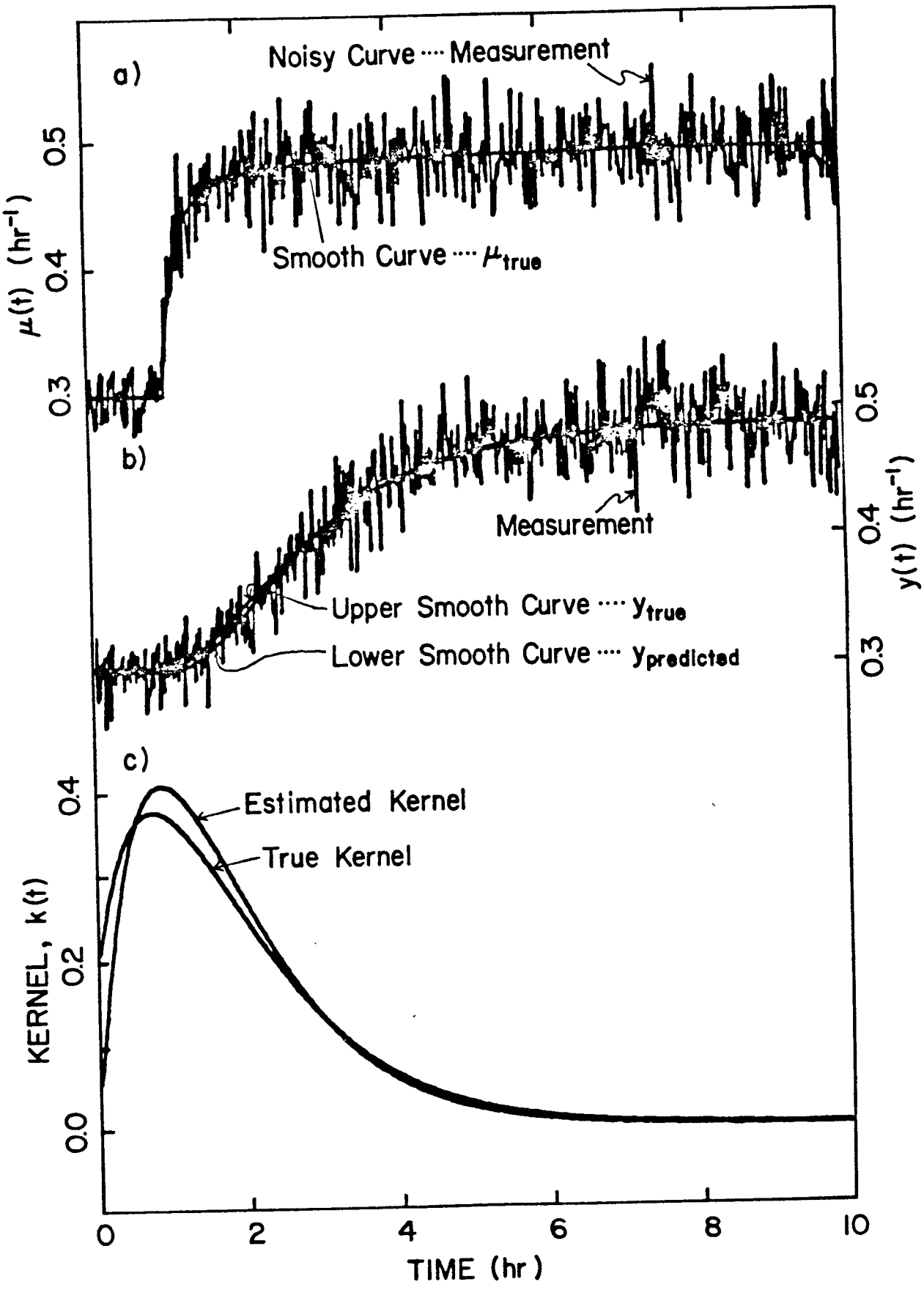

a)
Noisy Curve ···· Measurement
Smooth Curve ···· μ_{true}
$\mu(t)$ (hr^{-1})
0.3
0.4
0.5
b)
Measurement
Upper Smooth Curve ···· y_{true}
Lower Smooth Curve ···· $y_{predicted}$
$y(t)$ (hr^{-1})
0.3
0.4
0.5
c)
Estimated Kernel
True Kernel
KERNEL, k(t)
0.0
0.2
0.4
0
2
4
6
8
10
TIME (hr)

Figure 7

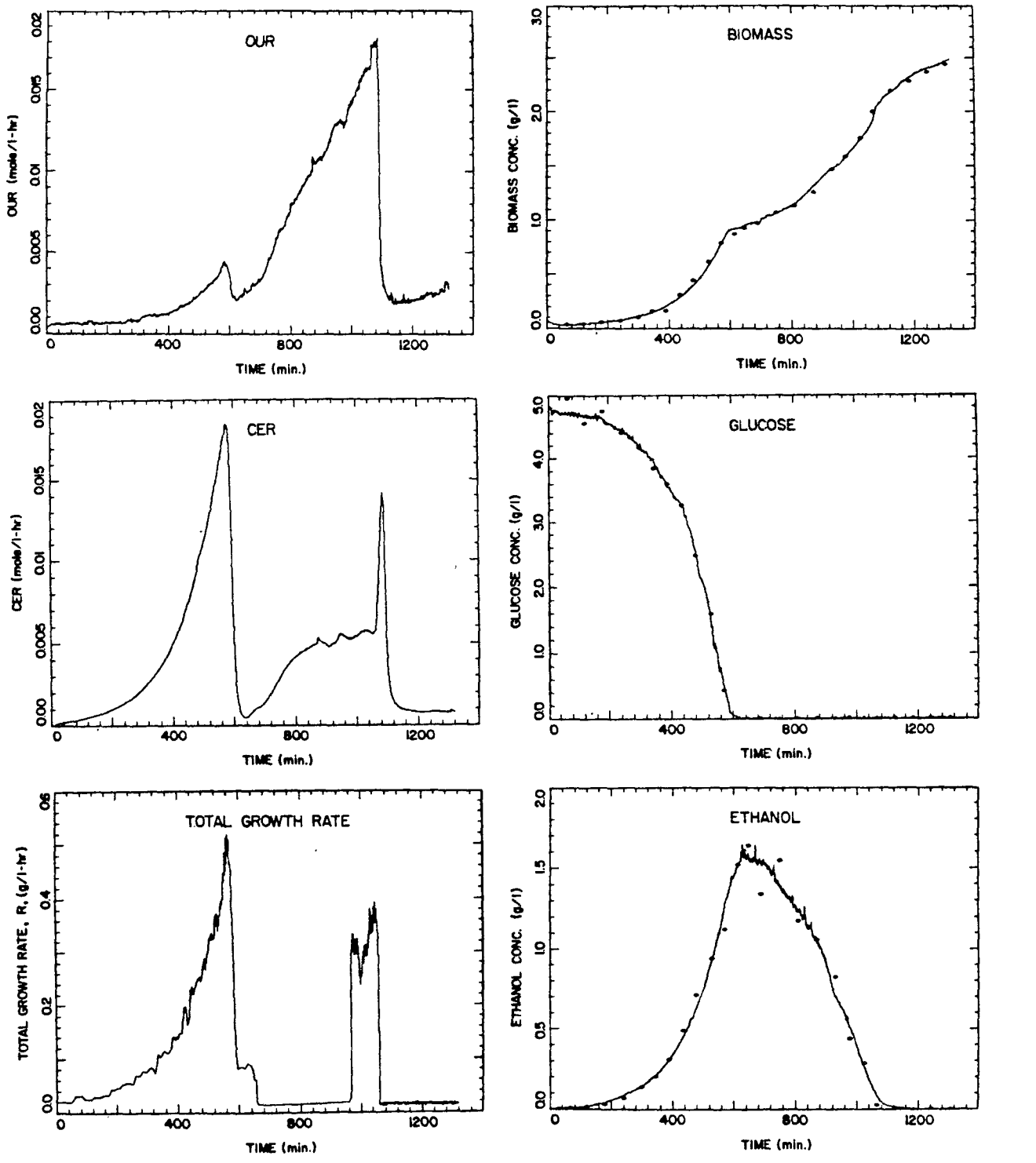

OUR
OUR (mole/l-hr)
0.00
0.005
0.01
0.015
0.02
0
400
800
1200
TIME (min.)
BIOMASS
BIOMASS CONC. (g/l)
0.0
1.0
2.0
3.0
0
400
800
1200
TIME (min.)
CER
CER (mole/l-hr)
0.00
0.005
0.01
0.015
0.02
0
400
800
1200
TIME (min.)
GLUCOSE
GLUCOSE CONC. (g/l)
0.0
1.0
2.0
3.0
4.0
5.0
0
400
800
1200
TIME (min.)
TOTAL GROWTH RATE
TOTAL GROWTH RATE, R, (g/l-hr)
0.0
0.2
0.4
0.6
0
400
800
1200
TIME (min.)
ETHANOL
ETHANOL CONC. (g/l)
0.0
0.5
1.0
1.5
2.0
0
400
800
1200
TIME (min.)

Figure 8

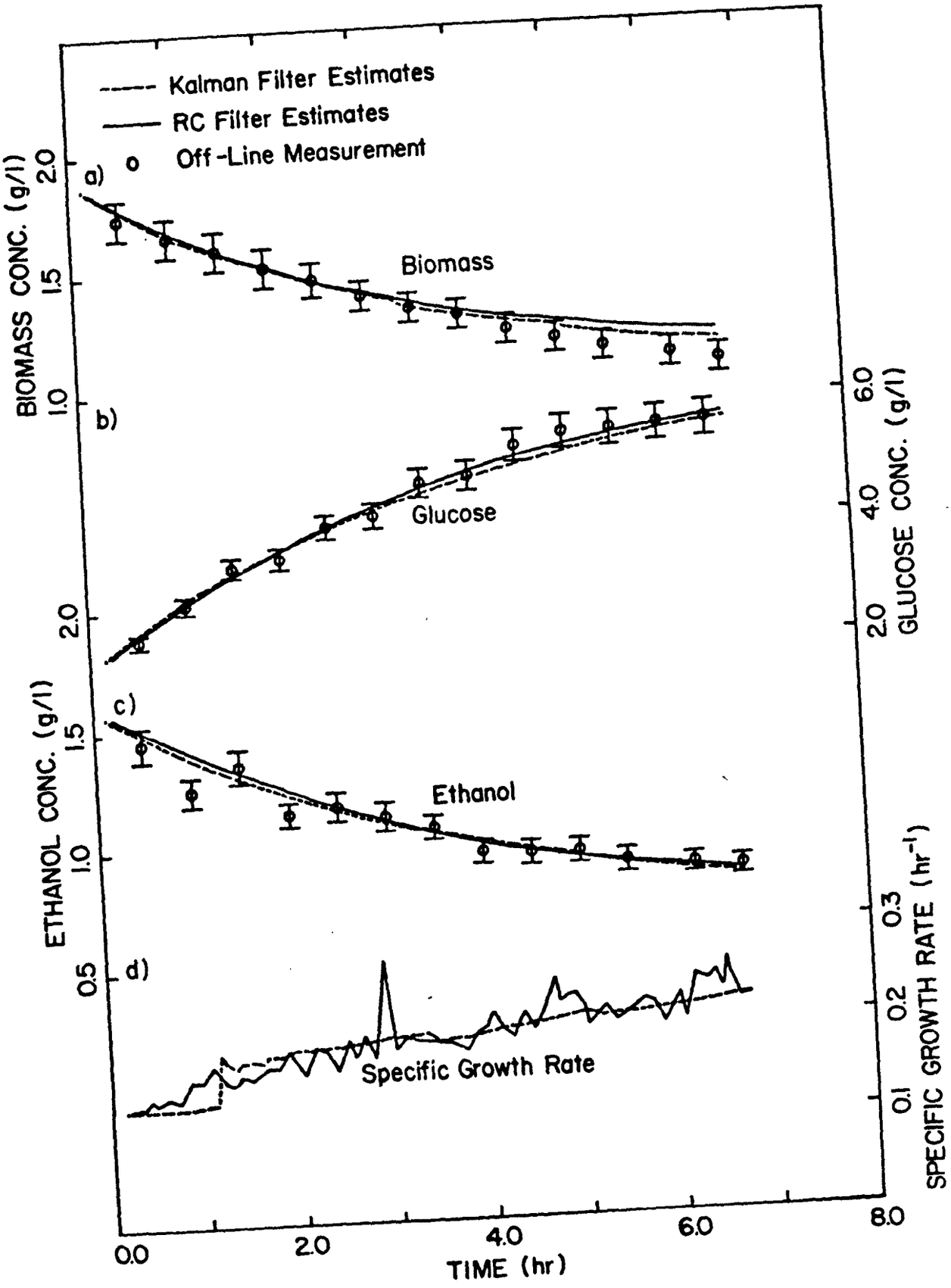

Kalman Filter Estimates
RC Filter Estimates
Off-Line Measurement
a)
Biomass
b)
Glucose
c)
Ethanol
d)
Specific Growth Rate
BIOMASS CONC. (g/l)
GLUCOSE CONC. (g/l)
ETHANOL CONC. (g/l)
SPECIFIC GROWTH RATE (hr⁻¹)
TIME (hr)
0.0
2.0
4.0
6.0
8.0
0.5
1.0
1.5
2.0
0.1
0.2
0.3
2.0
4.0
6.0

Figure 9

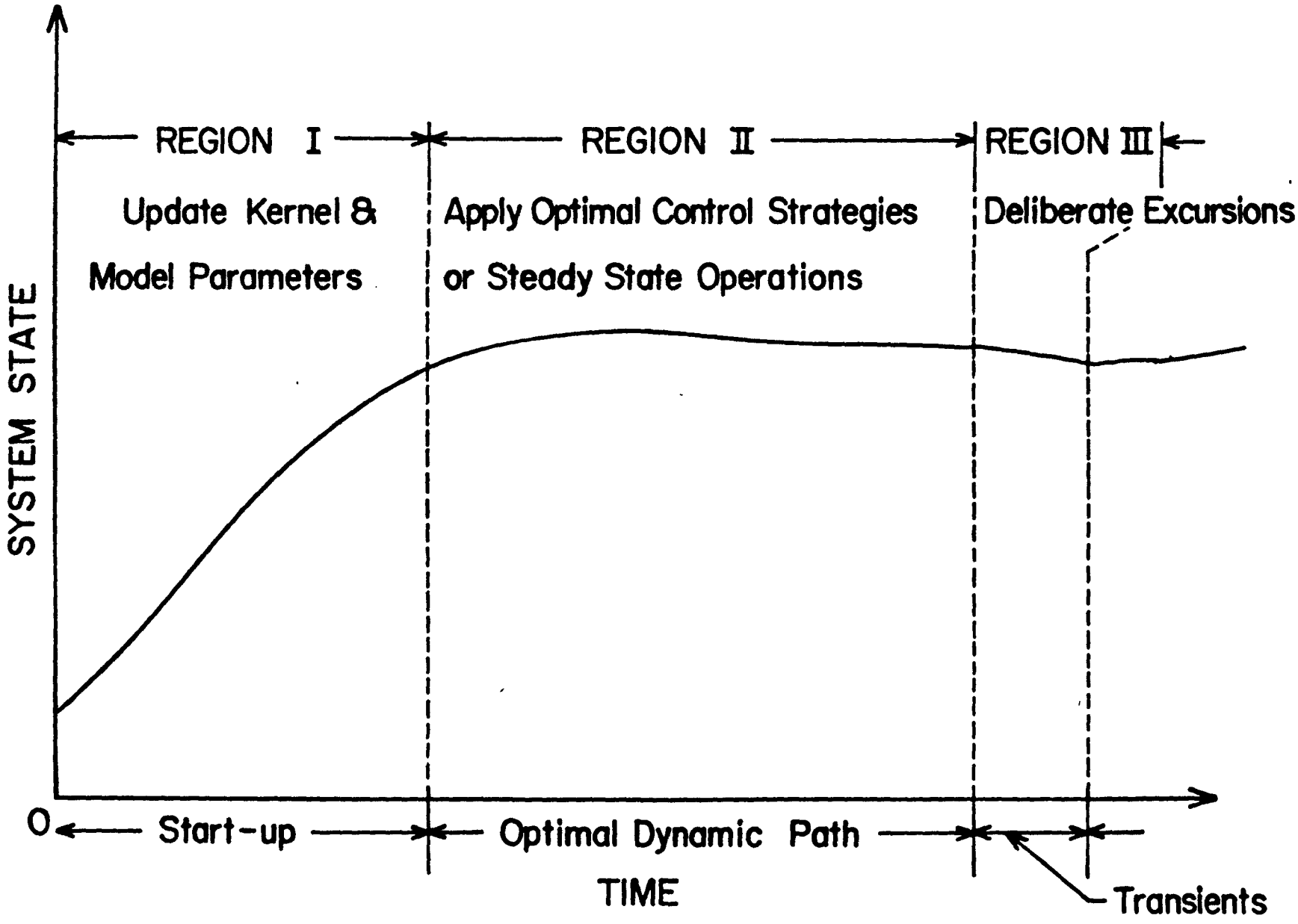

REGION I
REGION II
REGION III
Update Kernel &
Model Parameters
Apply Optimal Control Strategies
or Steady State Operations
Deliberate Excursions
SYSTEM STATE
0
Start-up
Optimal Dynamic Path
TIME
Transients

Figure 10

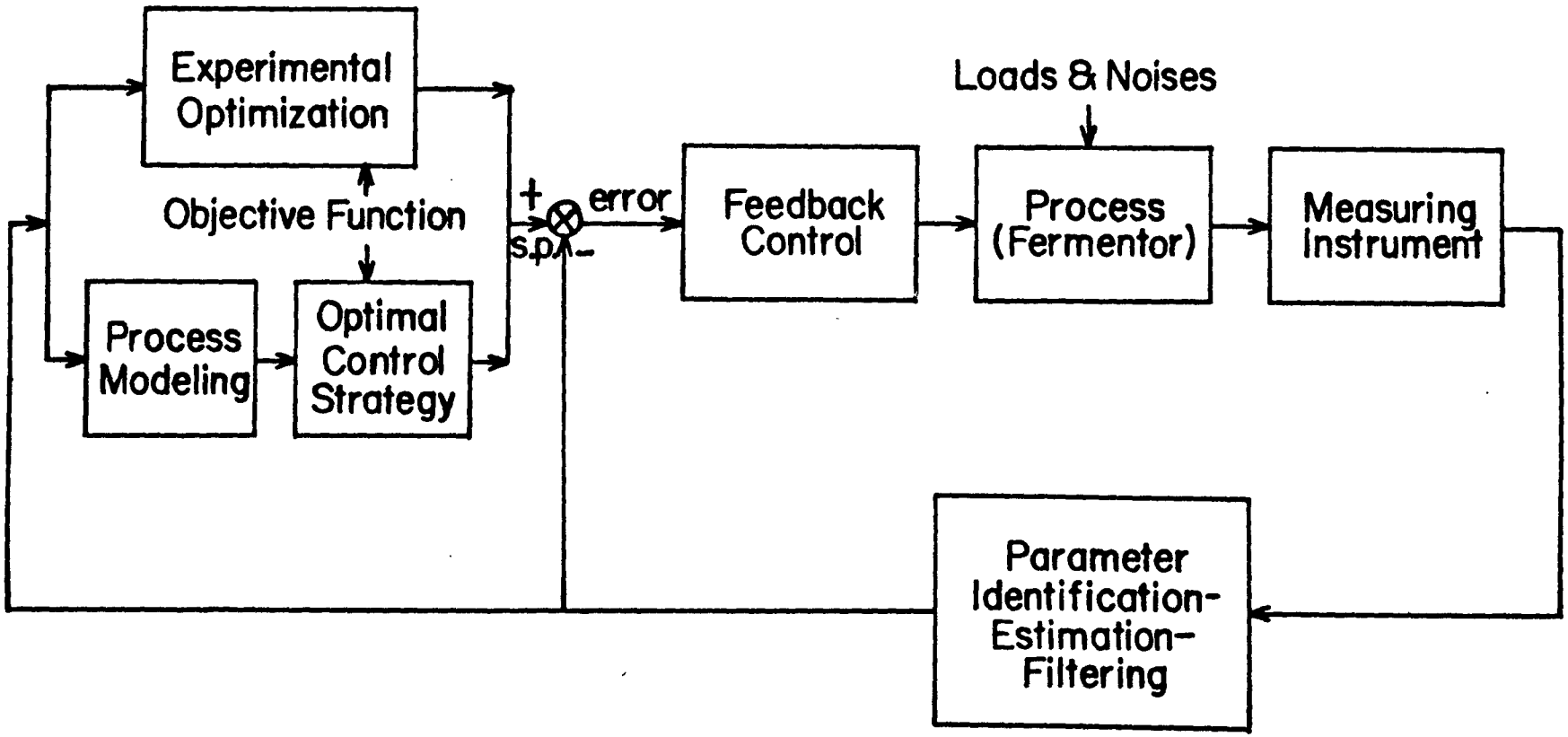

Experimental Optimization
Objective Function
Process Modeling
Optimal Control Strategy
+
s.p.
-
error
Feedback Control
Loads & Noises
Process (Fermentor)
Measuring Instrument
Parameter Identification-Estimation-Filtering

Figure 11