Abstract

DOUGHERTY, DANIEL PATRICK **Deterministic and Semi-Mechanistic Approaches in Predictive Fermentation Microbiology** (Under the direction of Sharon R. Lubkin)

Predictive fermentation microbiology utilizes deterministic and stochastic mathematical models to study the growth dynamics of microorganisms. If the components of such models represent known or hypothesized biological growth processes then these models can be used to refine existing hypotheses or generate new hypotheses about the factors controlling growth.

Special techniques must be used when fitting such models to experimental data. Methods are suggested for model re-parameterization and model fitting which improve the estimation of model parameters. Once estimates of model parameters have been made, temporal and multivariate sensitivity analyses can assess important relationships among the model parameters.

A deterministic dynamic model of batch growth by a homofermentative lactic acid bacterium growing in a variable temperature environment was derived. This model predicts cell growth as well as changes in the chemical composition of the medium. This model was fit to experimental data. Analysis of the model revealed a quantitative reversal in parameter sensitivities across temperatures. Although mechanistic, this model neglected the effects of pH , organic acid dissociation and ionic strength of the medium. It is shown that these chemical dynamics are important and can be modeled through a convenient semi-mechanistic approach. The ability to model these chemical dynamics appropriately allows for a modeling framework in which the acid tolerance strategies commonly exhibited by bacteria can be studied.

Deterministic and Semi-Mechanistic Approaches in Predictive Fermentation Microbiology

by

DANIEL P. DOUGHERTY

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

BIOMATHEMATICS PROGRAM

Raleigh

2002

APPROVED BY:

Sharon R. Lubkin Chair of Advisory Committee	Frederick Breidt Jr. Co-chair of Advisory Committee
Marcia L. Gumpertz	Zhilin Li

Biography

I was born in Fairfax County, Virginia to George J. and Joyce B. Dougherty. My father was a Lieutenant Colonel in the Army Corps of Engineers and my mother was, and still is, a teacher of English to international students. My father's military career resulted in my family and I being stationed in Frankfurt, Germany from where we traveled most of Europe during my early childhood. I spent most of my teenage years in Hampton, Virginia from where I graduated Kecoughtan High School in 1992.

Originally wanting to become a physical therapist, I studied Biology at Hampton University (Hampton, Virginia) and at James Madison University (Harrisonburg, Virginia). During my education at James Madison University, I developed an interest in Microbiology. I was fortunate to be able to conduct research as an undergraduate under the direction of Bruce Wiggins on the identification of sources of fecal pollution in agricultural watersheds. After taking Calculus courses from John Klippert and Peter Kohn at JMU, I decided that I also liked Mathematics. I graduated in 1997 with a Bachelors of Science in Biology and a Minor in Mathematics. After graduation, I spent a Summer at the Savannah River Ecology Laboratory in Aiken South Carolina.

In 1997, I began my study in the Biomathematics Program of the Department of Statistics at North Carolina State University. I received a Masters degree in 1999 for work on the numerical solution of partial differential equations relating to microbial chemotaxis. For the past 2 years, I have been working with the U.S. Department of Agriculture's Food Science Research Unit who have a research station at North Carolina State University.

Acknowledgements

There are many people who have helped in the creation of this dissertation. First, I would like to thank my advisor and dissertation chair Sharon R. Lubkin who has consistently found time to give advice and steer me clear of fruitless paths in my research. Her constructive criticisms and editorial candor have no doubt greatly improved the quality of this dissertation.

Many people at the USDA ARS Food Science Research Unit have been instrumental in helping me conduct the research in this dissertation. These include my advisor and dissertation co-chair Frederick Breidt Jr. who has always provided insight into the microbiology of the experiments we were conducting and was always optimistic about the progress we were making. Roger McFeeters has consistently provided many important insights into the biochemical aspects. Roger Thompson has been instrumental in carrying out the analytical chemistry and providing advice on technical procedures. I also wish to thank Janet Hayes and Laura Reina for their technical assistance.

I am also deeply indebted to my family who have supported me both financially and emotionally throughout my entire education. Without their support and encouragement I would have never considered attempting to obtain a Doctoral degree.

Lastly, I would like to thank some fellow Biomathematics students who have helped make my stay at NCSU particularly memorable. These include Cindy Chemilewski, Elizabeth Brooks, Sarah Hardy, John Fieberg, Nikkala Pack, Virginia Passour, Jason Pirone, Doug Robinson, Kyle Shertzer, Denis Thompson, Andrea Vogl, Tao Wang, and Jihao Zhou.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> *The study of the growth of bacterial cultures does not constitute a specialized subject or branch of research: it is the basic method of Microbiology. It would be a foolish enterprise, and doomed to failure, to attempt reviewing briefly a "subject" which covers actually our whole discipline. Unless, of course, we considered the formal laws of growth for their own sake, an approach which has repeatedly proved sterile.*
>
> Monod, J. (1949). The growth of bacterial cultures. *Annu. Rev. Microbio.* **3**, 371–394

1.1 Early Models of Population Growth

Since Jacques Monod made the above statement in 1949, the field of microbiology has greatly enlarged and diversified. The advent of molecular techniques, genomics, and proteomics as well as our enhanced understanding of cellular and sub-cellular fluid mechanics have created new avenues of research which in fact do not require

a centralized focus on population growth. Microbiology has perhaps become more mathematical than ever in order to tackle these new areas of research. At the same time, mathematical modeling of microbial population growth has begun to emerge as a distinct subdisclipline. The purpose of the review which follows is to give an account of how mathematical techniques for predicting microbial growth evolved in conjunction with experimental advances.

The development of mathematics greatly preceded the discovery of microorganisms. It is not surprising, therefore, that the earliest mathematical models of population growth were aimed at predicting human and animal population growth. It is possible that the first instance of a mathematical equation being used to describe the growth a biological population was in 1202 when Fibonacci derived his famous recursion relation. Eqn. 1.1 was found by Fibonacci to predict the number of rabbits in a population that reproduced according to the following simple assumptions (Edelstein-Keshet, 1988, p.5)

- Each rabbit reaches sexual maturity after 1 month.
- Each pair of sexually mature rabbits produces 1 new pair of rabbits every month
- The mortality of the rabbits is 0.

$$N_{t+2} = N_t + N_{t+1}, t = 1, 2, ... \tag{1.1}$$

where $N(1) = 0$ and $N(2) = 1$ by convention. The ideas engendered in this recursive relation are simple. The model suggests that the current reproductive capacity of a population depends on its current size and there is a simple geometric relationship relating the current population size to the new population size.

There are of course several problems with the Fibonacci series as a model of population growth. First, members of the population reproduce in unison at equal discrete intervals of time. Members of microbial and human populations, however, exhibit reproductive events asynchronously and need not always give rise to 2 offspring. The second concern, and perhaps the more important, is that the model assumes that there are no factors limiting the population such as death, disease, access to food etc.

The first objection to Fibonacci's model could not be adequately addressed for some 500 years. Continuous models of population growth were not possible until the advent of Calculus. During the 2 years between 1665 and 1667 Isaac Newton derived the basic foundations of Calculus while the first published work on Calculus did not surface until Leibniz independently published the basic rules of Calculus in 1684 (Boyce and DiPrima, 1988, pp. 110-111). A century later, the first continuous model of population growth would come into use.

Malthus in 1798 described the growth of human populations by stating only that "Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio." The Calculus allowed mathematicians of the time to write Malthus' assumptions as the differential equation Eqn 1.2. In this form it is clear that "Malthusian growth" implies that the rate of population growth is related to the current population size by a constant proportion.

$$\frac{dN}{dt} = \mu N \tag{1.2}$$

The "Mathusian" growth model was the first continuous model of population growth. The validity and implications of the model are still debated among economists and social scientists. The population of humans has, in fact, showed no signs of leveling

off since the time of Malthus. Indeed, some have argued the population of humans is growing even faster than exponentially (Cohen, 1995; Johansen and Sornette, 2001).

While Malthus clearly recognized that unchecked population growth could not continue indefinitely, he did not provide any model to describe limited population growth. Arguably, the first model to account for restricted population growth was introduced in 1838 by the Belgian mathematician Pierre François Verhulst (Pearl, 1925, p. 4). Known as the discrete Logistic equation (or simply the Verhulst equation) Eqn. 1.3 states that encounters between individuals in the population will result in fatality at a rate of α while the population grows in a geometric fashion similar to Eqn.1.2.

$$N_{t+1} = \mu N_t (1 - N_t) \tag{1.3}$$

$$= \mu N_t - \mu N_t^2 \tag{1.4}$$

where N_t is the population scaled by its maximum potential density. In this model, a dynamic steady state is possible ($N \equiv 1$) in which reproduction is exactly balanced by mortality.

One of the initial benefits from the development of Calculus was a thorough understanding of the properties of logarithms[1]. For example, Gompertz (1825) wanted to gain better predictions of human mortality for the purposes of adjusting annuities. He supposed that the specific mortality rate, defined by $-\frac{d\ln(N)}{dt}$ (or equivalently $-\frac{1}{N}\frac{dN}{dt}$), could be expressed in the form

$$-\frac{d\ln(N)}{dt} = aq^t \tag{1.5}$$

[1]Logarithms were originally discovered by Napier in 1614

where a and q are constants and t represents age.

By utilizing the properties of logarithms we have that

$$\begin{aligned} \ln(N) &= -\int aq^t dt \\ &= -\int ae^{t\ln q} dt \\ &= -\frac{a}{\ln q} e^{t\ln q} + C \\ &= -\frac{a}{\ln q} q^t + C \end{aligned} \tag{1.6}$$

So that

$$\begin{aligned} N &= \exp(-\frac{a}{\ln q} q^t + C) \\ &= KB^{q^t} \end{aligned} \tag{1.7}$$

where K is a constant and $B = -a/\ln q$. When used to model population growth, Eqn. 1.5 is often re-written in the following way.

$$\frac{dN}{dt} = aq^t N \tag{1.8}$$

$$= ae^{-\alpha t} N \tag{1.9}$$

where a represents the specific growth rate at $t = 0$ and α is a constant controlling the rate at which the population reaches steady-state. This makes clear the distinction between Gompertz growth and Malthusian growth. *In Gompertz growth, the growth rate decreases exponentially over time and is independent of the population size.* This

form is often integrated and re-parameterized into an explicit form

$$N = C \exp\left(\exp\left(-B\left(t - M\right)\right)\right) \tag{1.10}$$

where B is the relative growth rate at $t = M$, C is the upper asymptote of the population as $t \to \infty$ and M is the time at which the growth rate is maximal (McMeekin et al., 1993, p. 45). Although the Gompertz model, in its many re-parameterizations, is an improvement over unlimited population growth, it does not take into account environmental factors which may also limit growth.

1.2 The Advent of Microbiological Models

Calculus was not the only major scientific break-though made during the 1600's. In 1676, a 44 year old haberdasher living in Delft, Holland peered through a home-made glass lens and observed bacteria, fungi and other microorganisms. Antony van Leeuwenhoek's discovery of what he originally called "animalcules" marked the beginning of modern Microbiology. While some advances in taxonomy were made, bacteria were not effectively cultured until nearly 200 years later.

Microorganisms provided scientists with unique opportunities for studying population dynamics. Unlike many other populations, microbial populations could generally be grown quickly, inexpensively and reproducibly. There was an impetus to quantify, characterize, and compare microbial growth among various populations and across different experimental conditions. Mathematical models were able to provide means for accomplishing these goals. Unlike human or animal populations, however, microbial populations generally exceed the resources of their experimental environments

within a day's time. There was clearly a need for models more advanced than those of Malthus or Verhulst.

The advantage in using Calculus to model the growth of bacteria over time did not escape microbiologists. For example, the microbiologists M'Kendrick and Pai (1910) suggested a continuous model based on the assumption that the rate of microbial growth was proportional to the amount of available food left in the environment. They also made the simplifying assumption that the amount of resources required for maintenance was negligible when compared to the amount spent on growth. The resulting model was

$$\frac{dN}{dt} = \beta \left(S_o - N \right) N \tag{1.11}$$

where S_o is the original concentration of nutrition available and β is a rate constant. Note that the nutrient concentration has been scaled by the amount of nutrient required to produce a single cell. Thus, the units of S and N are taken to be equal.

Eqn. 1.11 reconciled a problem which the Verhulst equation shares with the Fibonacci sequence, namely that the population grows at discrete times. This equation, generally referred to as the Logistic equation, is however normally attributed to Pearl. Pearl and Reed (1920) drew on new theory concerning the Calculus of autocatalyzed chemical reactions and re-cast the discrete Verhulst equation as a continuous model for the growth of human populations. Pearl also allowed for a second parameter (K) to describe the limiting population size.

$$\frac{dN}{dt} = \mu N \left(1 - \frac{N}{K} \right) \tag{1.12}$$

Table 1.1: The 6 phases of microbial growth as originally defined by Monod (1949). Note that the definitions are given in terms of instantaneous rates of change. One or more of these phases may be absent and each phase could occur multiple times.

1	lag phase	growth rate null
2	acceleration phase	growth rate increases
3	exponential phase	growth rate constant
4	retardation phase	growth rate decreases
5	stationary phase	growth rate null
6	phase of decline	growth rate negative

Eqn. 1.11 is clearly Eqn 1.12 re-parameterized with $\mu = \beta S_o$ and $K = S_o$.

Clearly, it was common knowledge among microbiologists of the early 1900's that if microorganisms were grown in rich media then, at least initially, the population growth could be described by Eqn. 1.2 (hence the term exponential growth). However, a clear understanding of the subsequent phases of microbial growth was missing due to the lack of consistent and rigorously defined set of criteria. Monod (1949) recognized 6 distinct phases of microbial growth (see Table 1.1). Monod, furthermore, gave each definition in terms of the instantaneous rate of change of the microbial density. Defining the phases of growth using the concept of a derivative allowed direct exploration of microbial growth by more advanced mathematical models and techniques.

While providing mathematical definitions for the phases of growth, Monod is also recognized for his work in predictive microbiology and especially in continuous culture of bacteria. In particular, Monod (1942) introduced a mathematical model of bacterial growth limited by a single substrate. But unlike M'Kendrick and Pai, Monod drew on the saturating kinetics of enzymes put forth by Michaelis and Menten (1913) and

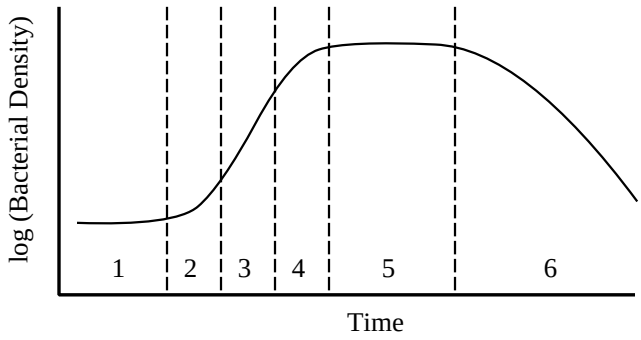

Figure 1.1: The 6 phases of microbial growth. 1) *lag phase* 2) *acceleration phase* 3) *exponential phase* 4) *retardation phase* 5) *stationary phase* 6) *phase of decline.*

suggested that the specific growth rate should be considered a hyperbolic function of substrate concentration.

$$\frac{dN}{dt} = \frac{\alpha}{Y}\left(\frac{S}{K_s + S}\right) N \tag{1.13}$$

Eqn. 1.13 blends the assets of Eqn. 1.11 and Eqn. 1.12. Here S is the concentration of the substrate in the environment, Y is a parameter relating the amount substrate required to produce a new cell, α is the maximum rate of substrate utilization and K_s is the half-saturation constant for substrate utilization. When used in conjunction with an equation describing the substrate utilization and exponential death of the cells, say

$$\frac{dN}{dt} = \frac{\alpha}{Y}\left(\frac{S}{K_s + S}\right) N - \delta N \tag{1.14}$$

$$\frac{dS}{dt} = -\alpha\left(\frac{S}{K_s + S}\right) N \tag{1.15}$$

then the Monod equation can describe the exponential, retardation, stationary, and death phase as well as the concentration of the limiting resource.

A limitation of Monod's model is that growth often ceases in stationary phase with a large amount of residual resource. Clearly, in these cases, substrate limitation is not the only limiting factor. Ramkrishna et al. (1967) expanded Monod's model to include the production of growth inhibiting substances such as organic acids which are the waste products of many fermentations.

$$\frac{dN}{dt} = \frac{\alpha}{Y}\left(\frac{S}{K_S + S}\right)N - \delta IN \tag{1.16}$$

$$\frac{dS}{dt} = -\alpha\left(\frac{S}{K_S + S}\right)N \tag{1.17}$$

$$\frac{dI}{dt} = \gamma\left(\alpha\left(\frac{S}{K_S + S}\right)\right)N + \tau\delta IN \tag{1.18}$$

Here I is the concentration of the inhibitor substance and S, N have the same meanings used above. The inhibiting substance is produced in proportion to the number of cells dying at time t while cells are killed by their interaction with the inhibitor at a rate δ. Thus, Ramkrishna's formulation assumes that inhibition occurs by increasing the death rate of the microorganism.

A different approach to modeling the inhibition of growth was taken by Levenspiel (1980) who simply multiplied Monod's single-substrate limitation term by a saturating inhibition term.

$$\frac{dN}{dt} = \alpha\left(1 - \frac{I}{\mathrm{MIC}_I}\right)^n \frac{S}{K_s + S}N \tag{1.19}$$

Here I is the concentration of the inhibiting substance, MIC_I is the largest concentra-

tion of inhibitor at which growth can still occur, and n is a constant. In the Levenspiel modification, death of the microorganism only occurs for $I > \text{MIC}_I$.

In addition to ignoring growth inhibition, the Monod model also ignores the process of nutrient uptake. In order for microorganisms to utilize a substrate they must either actively channel or passively acquire the substrate within the cytoplasm. Droop (1973) modeled microorganisms (originally algae) by considering an internal quota of substrate (Q). Specifically, Q represents the amount of substrate "captured" by the organism. The substrate uptake is considered to be mediated by an enzyme and therefore follows saturation kinetics. The essential Droop model is as follows

$$\frac{dN}{dt} = \alpha\left(1 - \frac{Q_{\min}}{Q}\right)N - \delta N \tag{1.20}$$

$$\frac{dQ}{dt} = -\mu_{\max}\left(\frac{S}{K_s + S}\right) - \alpha\left(1 - \frac{Q_{\min}}{Q}\right)Q \tag{1.21}$$

$$\frac{dS}{dt} = -\mu_{\max}\left(\frac{S}{K_s + S}\right)N \tag{1.22}$$

where $\mu_{\max}$ is the maximum specific uptake rate,α is the maximum specific growth rate, $Q_{\min}$ is the minimum Q required for growth and K_s is a half saturation constant for substrate uptake. Although the Droop and Monod models are conceptually different formulations of microbial growth, (Burmaster, 1979) has shown that in fact under mild restrictions their solutions are equivalent at steady-state.

Nevertheless, the Droop model high-lighted the need for a new approach to modeling the ecology of populations in dynamic environments. Building on the basic idea of non-constant yield function Kooijman (2001, 2000) and others have developed the theory of dynamic energy budget models. The basic idea of a dynamic energy budget model is that Q is partitioned so that some fraction is reserved for growth

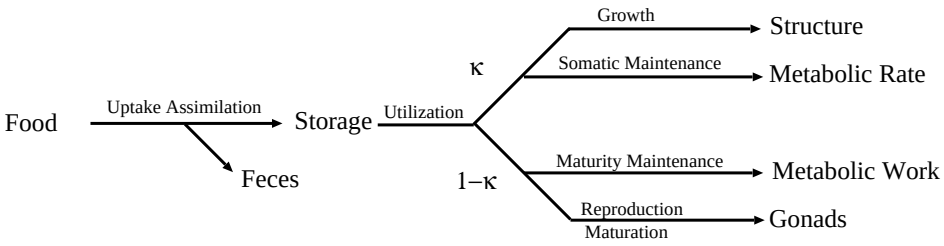

Figure 1.2: Diagrammatic depiction of Kooijman's κ–rule model.

processes but another fraction is reserved for maintenance. Thus, the yield is not constant but that the way in which the yield is perturbed occurs in a deterministic manner. The prototypical dynamic energy budget model is Kooijman's κ–rule model (See Figure 1.2). Kooijman (2000) has been able to show that the Monod and Droop models are simplified cases of the κ–rule model.

1.3 Modeling Lag Phase

The models discussed in the previous section gave methods for modeling exponential, stationary and death phase (see Table 1.1). The purpose of this section is to describe methods for modeling lag phase. Lag phase corresponds to a period of metabolic adaptation to new or stressful environmental conditions. Temperature, for example, has a major effect on the duration of lag phase (Dufrenne et al., 1997). In order to maintain optimal membrane fluidity bacteria must modify compounds such as lipids and cholesterols. While energy is being diverted to synthesizing these compounds cell division slows or ceases altogether. Similar responses occur during exposure to acid or heat stress where special heat or acid "shock proteins" are synthesized which increase tolerance of these stresses.

Lag phase can be identified graphically on a semi-log plot of cell density over time.

The end of lag phase can be identified as the time at which the second derivative of log cell density with respect to time is maximized (Buchanan and Cygnarowicz, 1990).

$$\text{lag duration} \quad = \quad \arg\max_{t} \left(\frac{\partial^2 \log(CFU mL^{-1})}{\partial t^2} \right) \tag{1.23}$$

This method for calculating lag phase is illustrated in Figure 1.3. Estimations of the second derivative can be very sensitive to the spacing of observations in time and measurement error. A robust method for estimating the second derivative is to use local cubic regression in conjunction with iteratively re-weighted least squares (Fan and Gijbels, 1996).

Lag phase can be modeled using delay differential equations (for an introduction see MacDonald (1982)). However, to be mechanistic, a lag phase model must relate the duration of lag phase to current environmental conditions (McKellar and Knight,

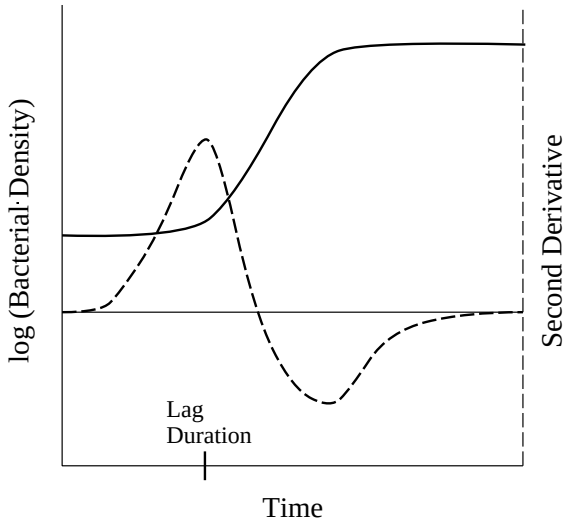

Figure 1.3: Determination of lag phase duration using the second derivative method. The end of lag phase is the time at which the second derivative of the log cell density has its maximum.

2000). The use of delay differential equations is also limited in that non-constant and non-linear delays can not be solved easily with standard software (although see Shampine and Thompson (2001)). Such equations, for example, can be re-cast as partial differential equations treating "delay time" as an extra dimension. This greatly increases the numerical complexity however. The mathematical analysis of such delay differential equations also requires special attention. In particular, one must take care in specifying the lag "history". In general, it seems that delay differential equations are more popular in models of continuous culture in which population oscillations are observed.

For constant lag, a delay differential equation for bacterial growth with a lag phase can be written as a two compartment model (Baranyi, 1998). Specifically one equation describes the population of cells still in lag phase and another equation describes the population in exponential phase.

$$\frac{dN_{\text{lag}}}{dt} = -\nu N_{\text{lag}} \tag{1.24}$$

$$\frac{dN_{\text{exp}}}{dt} = \mu N_{\text{exp}} + \nu N_{\text{lag}} \tag{1.25}$$

$$N_{\text{lag}}(0) = N_o \tag{1.26}$$

$$N_{\text{exp}}(0) = 0 \tag{1.27}$$

Other approaches to modeling lag involve modifications of existing models such as the Monod, Droop and others presented earlier. For example, Gibson et al. (1987) suggested a modification of the Gompertz model (Eqn. 1.7) where $\log(N)$ was modeled

and a parameter A was used to set the value of the lower asymptote

$$\log(N) = A + Dexp\left(-exp\left(-B(t-M)\right)\right) \tag{1.28}$$

One limitation of this model is that for data in which there is no lag phase a negative estimate for A can be obtained. This, however, usually does not present a practical problem.

Baranyi et al. (1995) suggests a model which combines aspects of the Droop model (Eqn. 1.20) and logistic equation (Eqn. 1.12). They posit the existence of a substance q which represents the physiological state of a cell. They assume that during exponential phase, this substance should increase exponentially.

$$\frac{dq}{dt} = \nu q \tag{1.29}$$

$$\frac{dN}{dt} = \mu_{\max}\frac{q}{1+q}\left(1-\frac{N}{N_{\max}}\right)N \tag{1.30}$$

where $\mu_{\max}$ is given by $a(T-T_{min})^2$ known as Ratkowsky's square root model (McMeekin et al., 1993, p.94). Here $\mu_{\max}$ is the maximum specific growth rate of the microorganisms under the current environmental conditions, ν is specific growth rate of the substance q (Baranyi et al. (1995) take it to equal $\mu_{\max}$), $N_{\max}$ is the carrying capacity of the culture, b is the change in growth rate per unit T and T_{min} is the extrapolated temperature at which no growth occurs. Note that T_{min} is not intended to coincide with the true minimum temperature for growth. They found that accurate predictions of lag phase and growth during dynamic shifts in temperature depended most strongly on obtaining accurate estimates for the initial value of q.

1.4 Comment on Predictive Food Microbiology

As noted by McMeekin et al. (1993, p.6) the importance of predictive microbiology in the food industry arises out of the fact that the goals of predictive microbiology coincide with those of HACCP (Hazard Analysis and Critical Control Points) systems for identification of risks in food production processes. These include identifying important organism(s), identifying the variables which have the greatest effect on growth, predicting the number of organism at every point during growth, and re-evaluating current culturing practices and ecological theories. It is not surprising that one of the most active area of research for predictive microbiology is in the field of Food Science (Baranyi and Roberts, 1995; Pruitt and Kamau, 1993; Skinner et al., 1994).

An area of interest in predictive food microbiology is the effect that temperature has on the growth rate of bacteria. As mentioned above, temperature affects the duration of lag phase and can therefore be a factor in determining shelf life and spoilage. Also of concern is the affect the food micro-environment has on the survival of food-borne pathogens. Models which restrict growth based on the accumulation of toxic end-products such as Levenspiel's modification of the Monod equation or Ramkrishna's model can be used to predict the outcome of competitive interactions between microorganisms in food if expanded to include equations for the different microorganisms present. This may include, for example competition between a pathogen and a naturally present microorganism.

The purpose of this thesis is to develop new techniques for predicting the growth of microbial populations in variable environments. In Chapter 2 we discuss mathematical techniques for the construction of mathematical models and methods for

estimating the parameters of such models from experimental data. In Chapter 3 a mathematical model for the batch growth of lactic acid bacteria in variable temperature fermentations is developed. We find that one of the limitations of this model is that is does not account for the inhibitory effects of pH and organic acid accumulation in a mechanistically accurate manner. The goal of Chapter 4 is therefore to give a thorough development of a flexible technique for modeling the dynamics of pH and organic acid accumulation in fermentations. Finally in Chapter 5 we apply this technique to the dynamic model.

Chapter 2

Model Parameterization, Model Fitting and Model Analysis

The models presented in Chapter 1 are prototypical and based on fairly restrictive assumptions. As such they are often adapted or included within larger systems of equations. It is the goal of this chapter to discuss issues involved in applying such models as tools in the analysis and characterization of experimental data. First, methods of model re-parameterization are discussed which may improve estimation properties of the model parameters. Second, practical issues encountered in estimating the parameters of dynamical models from data are discussed. Lastly, some techniques for analyzing a model once the parameters have been estimated are given.

2.1 Techniques for Model Re-parameterization

The manner in which parameters enter in a model is not unique. Depending on the discretion of the investigator, algebraic manipulations can lead to parameterizations

which for one reason or another are considered advantageous. In this section, 2 methods of re-parameterization are discussed and suggestions for their application when formulating models are given.

2.1.1 The Method of Nondimensionalization

According to the Π theorem of Buckingham (1914), through a process known as nondimensionalization, a model can be reduced to a form which possesses all of the qualitative behavior of the dimensional model but has the fewest number of parameters possible, each of which is dimensionless. This method therefore has some potentially nice properties. These are:

1. Improvement in the scaling (numerical properties) of the problem
2. Reduction in the number of free parameters
3. Elucidating or incorporating the important relationships among parameters

Unfortunately, nondimensionalization does not provide a unique mapping from the dimensional parameter space to the nondimensional space. For ease of exposition and in order to highlight some of the pitfalls, consider the equation of a line.

$$Y = \beta_o + \beta_1 X \tag{2.1}$$

Let's consider this line to relate the height of a particular object (Y) as measured in meters to its weight (X) measured in kilograms. For now, it will be assumed that $\beta_o = 3\text{m}$ and $\beta_1 = 10\text{mkg}^{-1}$. Looking at Fig. 2.1 it is clear that the scales of the variables (X and Y) are rather disparate. Supposing that this is undesirable for the

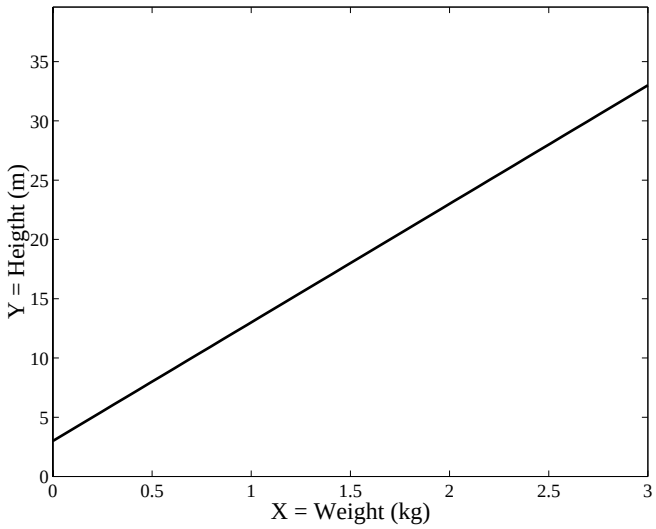

Figure 2.1: Line with slope of 10 meters kilogram $^{-1}$ and intercept of 3 meters. Note the disparity in scales between the variables X and Y.

application it would be nice to have an equation that exhibits all of the qualitative features of Eqn. 2.1 but permits the variables to be represented on nearly equal scale. This can be accomplished through nondimensionalization.

Let $y = y'y^*$ and $x = x'x^*$ where x' and y' are characteristic measurement values (i.e. units) and x^* and y^* are dimensionless variables. Then

$$y = \beta_o + \beta_1 x \tag{2.2}$$

$$y^* y' = \beta_o + \beta_1 x^* x' \tag{2.3}$$

$$y^* = \left[\frac{\beta_o}{y'}\right] + \left[\frac{\beta_1 x'}{y'}\right] x' \tag{2.4}$$

Where dimensionless parameter clusters have been enclosed in [ ]. The units of the data Y are scaled by setting $Y' = \beta_o$. This has the effect of dividing the heights by the intercept 3 m. The units of X are scaled by setting $X' = \beta_o/\beta_1$. This has

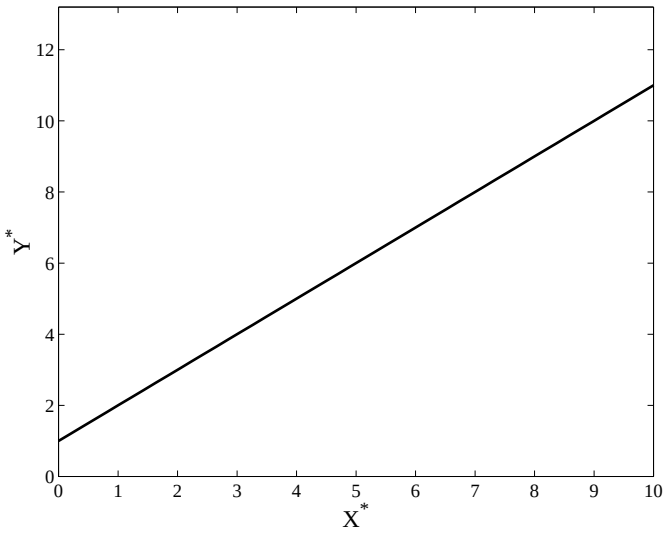

Figure 2.2: Nondimensionalized line with slope of 1 and intercept of 1. Note the parity of scales of the nondimensional variables X^* and Y^*.

the effect of dividing the weights by $3m/10\ kg$. These scaling relationships provide a nondimensional equation of a line having both slope and intercept equal to 1.

The result is a functional relationship which can certainly be thought of as the "canonical" form of a line. As promised by Buckingham's Π theorem, it was possible to eliminate 2 parameters by choosing appropriate scaling relationships with the 2 state variables.

Unfortunately, when the objective is inference, this nondimensionalization is not useful. Indeed, there are no parameters to be estimated. Furthermore, this nondimensionalization can be singular. In particular, if it were true that an object of height zero also had weight zero ($\beta_o = 0$), certainly not an unreasonable possibility, then the inverse transformation would have an isolated singularity (i.e. $y' = \frac{\beta_o}{0} = \infty$). Certainly, choosing a model with more terms, for example a quadratic, would leave some parameters in the model after nondimensionalization. However, in order to perform

inference on a space of nondimensional parameters the data must be scaled as set out in the nondimensionalization. This can not be done without knowing the value of the dimensional parameters involved in the scaling.

2.1.2 The "Parameters in the Denominator" Method

While it turns out that nondimensionalization is not useful in the context of parameter estimation there are other possibilities for model re-parameterization that may be beneficial prior to model fitting. In many modeling applications, functional forms which asymptote (or saturate) are used as surrogates for complicated biological processes which are enzyme mediated or which for one reason or another are thought to follow saturation kinetics. By far the most well-known functional form for saturation kinetics is the Michaelis-Menten equation. The standard Michaelis-Menten equation (see graph in Fig. 2.3) is

$$V_o = \frac{V_{\max}\ [S]}{K_{\mathrm{M}}\ + [S]} \tag{2.5}$$

where V_o is the initial rate of reaction, $V_{\max}$ is the maximum rate of reaction and K_{M} is the value of $[S]$ at which $V_o = V_{\max}\ /2$. There are also various modifications of the Michaelis-Menten equation useful for modeling situations where an inhibitor substance reduces the rate (V_o). For example, competitive inhibition leads to an increase in the apparent K_{M} , non-competitive inhibition leads to a decrease in the apparent $V_{\max}$ and uncompetitive inhibition leads to an increase in the apparent K_{M} and a decrease in the apparent $V_{\max}$.

A practical problem which has received a great deal of attention is the poor estimation properties of the parameters ($V_{\max}$, K_{M}) of the Michaelis-Menten form (Bates et al.,

1987; Ratkowsky, 1986; Goovaerts et al., 2001; Bentzen and Taylor, 1991). The problem arises from the correlation between the model parameters. For example, in the limit as $[S]$ approaches zero the Michaelis-Menten equation reduces to that of a line. Therefore, at low substrate concentrations it is only the *ratio* of $V_{\max}$ to K_{M} that matters. Furthermore, it is often not possible for an experimentalist to ensure that all regions of the curve (see Fig. 2.3) are represented in the data (Bentzen and Taylor, 1991). This is especially true when saturation kinetics are embedded in a dynamical systems model where the dynamics control the levels of "substrate". Recall for example the Droop model (Eqn. 1.20) where the growth rate was determined by the concentration of acquired substrate (Q). Since an increase in substrate acquisition results in an increase in cell division, saturating levels of Q may never be attained. As is clear from the examples given in Chapter 1, many of the equations used in mathematical models of bacterial growth involve Michaelis-Menten or related saturation kinetics forms. As such, even if the number of data points is large, the information about each of the parameters in the Michaelis-Menten form may not be sufficient to provide accurate estimation.

The reason for this is as follows. Numerical methods used for estimating the parameters in nonlinear equations require estimates of the gradient of the objective function with respect to the parameters in the model. If the surface generated by the objective function is smooth, has only a single minimum and is locally quadratic about the minimum then the estimation procedure is generally efficient (Press et al., 1992). Unfortunately, rather than being locally quadratic about the true values of the parameters, the objective surface of the Michaelis-Menten equation is highly collinear. This leads to inefficient estimation and bias and large uncertainty in the parameter estimates.

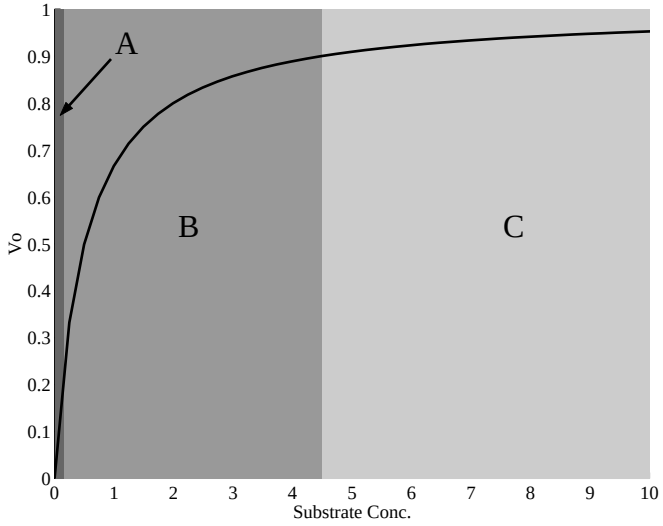

Figure 2.3: Standard Michaelis-Menten relationship between substrate concentration and the initial reaction velocity. For this plot $V_{\max} = 1$ and $K_{\mathrm{M}} = 0.5$. Regions A, B and C correspond to the first-order, non-linear and zero-order regions of the curve.

In an effort to improve the estimation procedure, Ratkowsky (1986) suggested a standard reparameterization technique for the functional forms common in reaction kinetics models (Hougen-Watson equations). This re-parameterization is called the "parameters in the denominator" method. For example, Eqn. 2.5 can be re-parameterized by dividing numerator and denominator by V_{max}. Doing this gives

$$V_o = \frac{[S]}{\alpha + \beta\,[S]} \tag{2.6}$$

where $\alpha = \frac{K_{\mathrm{M}}}{V_{\max}}$ and $\beta = \frac{1}{V_{\max}}$. Note that the denominator is linear in the parameters. Other common Michaelis-Menten forms and their "parameters in the denominator" re-parameterizations are given in Table 2.1. Note that if the denominator is to be linear in the parameters the re-parameterization of the non-competitive form of Michaelis-Menten results in 4 parameters. The introduction of an additional parame-

ter which is by definition correlated with the other parameters may not outweigh the cost of optimizing an additional parameter. For this reason, an alternate "parameters in denominator" version is given.

Ratkowsky (1985) demonstrated the superiority of this re-parameterization method in the estimation of kinetics parameters from 17 data sets taken from the literature. This analysis of the "parameters in the denominator" method however, did not consider the properties of the estimates when data only from a particular region of the saturation curve are available. Thus, to compare the properties of the standard parameterization to the "parameters in the denominator" re-parameterization a simulation study was performed.

Regions associated with the first order ($[S] << K_{\mathrm{M}}$), nonlinear($[S] \approx K_{\mathrm{M}}$) and zero order ($[S] >> K_{\mathrm{M}}$) portions of the Michaelis-Menten curve were selected. The boundaries of regions A, B, and C were set at substrate concentrations [0,0.75), (0.75,9.5) and (9.5,49.5] respectively. Within each region data was simulated and parameters estimated. Specifically, within a particular region of Fig. 2.3, 25 substrate concentration levels were generated equally-spaced over the region. The velocity (V_o) of each point was given by Eqn.2.5 with $V_{\mathrm{max}} = 1$ and $K_{\mathrm{M}} = 0.5$ plus a normally distributed deviate with mean 0 and standard deviation 0.02. To compare the standard parameterization to the parameters in denominator re-parameterization the following statistics were computed for each of the parameters in each of the substrate regions: empirical variance (EV), average asymptotic variance (AAV), average bias (AB) and the mean squared error (MSE). Averaged were calculated using 100 simulated data sets.

As can be seen in Table 2.2 the "parameters in the denominator" method gives less correlated parameter estimates and slightly more precise (see Table 2.3) estimates in

Table 2.1: Covariance reducing re-parameterizations of common Michaelis-Menten type kinetics equations. $F = 1 + I/\mathrm{K}_I$, $\alpha = K_\mathrm{M}\ /V_\mathrm{max}$,$\beta_1 = 1/V_\mathrm{max}$ $\gamma = K_\mathrm{M}\ /(V_\mathrm{max}\ \mathrm{K}_I\ )$, and $\delta = 1/(\mathrm{K}_I\, V_\mathrm{max}\ )$

Kinetics	Standard	Re-parameterized
Uninhibited	$V_o = \frac{V_\mathrm{max}\ [S]}{K_\mathrm{M}\ +[S]}$	$V_o = \frac{[S]}{\alpha+\beta[S]}$
Competitive	$V_o = \frac{V_\mathrm{max}\ S}{FK_\mathrm{M}\ +S}$	$V_o = \frac{[S]}{\alpha+\beta 2[S]} + \gamma\,[I]$
Uncompetitive	$V_o = \frac{V_\mathrm{max}\ S}{K_\mathrm{M}\ +FS}$	$V_o = \frac{[S]}{\beta_0+\beta[S]+\delta[I][S]}$
Non-competitive	$V_o = \frac{V_\mathrm{max}\ S}{F(K_\mathrm{M}\ +S)}$	$V_o = \frac{[S]}{\alpha+\beta[S]+\gamma[I]+\delta[I][S]}$
	or	$V_o = \frac{[S]}{F(\alpha+\beta[S])}$

Table 2.2: Correlations among parameter estimates of the uninhibited Michaelis-Menten form obtained from 25 data points generated uniformly in the region indicated with a simulated error variance of 0.02. The true V_max and K_M were 1 and 0.5 respectively. The boundaries of regions A, B and C were set at substrate concentrations [0,0.75), (0.75,9.5) and (9.5,49.5] respectively. Reported values are the average of 100 trials.

	Region		
Model	A	B	C
Traditional	0.981	0.835	0.889
Re-parameterized	-0.925	-0.788	-0.884

Table 2.3: Comparison of emperical variance (EV), average asymptotic variance (AAV), average bias (AB) and mean squared error (MSE) for uninhibited Michaelis-Menten. Reported values are the average of 100 trials.

Region	Statistic	V_{max}	K_M	α	β
	EV	2.33×10^{-3}	2.17×10^{-3}	5.76×10^{-4}	2.42×10^{-3}
A	AAV	2.16×10^{-3}	2.02×10^{-3}	5.37×10^{-4}	2.25×10^{-3}
	AB	3.55×10^{-3}	4.59×10^{-3}	3.88×10^{-3}	-5.90×10^{-3}
	MSE	2.34×10^{-3}	2.19×10^{-3}	5.91×10^{-4}	2.45×10^{-3}
	EV	7.47×10^{-5}	1.06×10^{-3}	8.41×10^{-4}	7.49×10^{-5}
B	AAV	6.48×10^{-5}	9.69×10^{-4}	7.76×10^{-4}	6.47×10^{-5}
	AB	-5.07×10^{-4}	-2.23×10^{-3}	-1.78×10^{-3}	4.34×10^{-4}
	MSE	7.50×10^{-5}	1.06×10^{-3}	8.44×10^{-4}	7.51×10^{-5}
	EV	6.19×10^{-5}	3.35×10^{-2}	3.25×10^{-2}	6.22×10^{-5}
C	AAV	6.98×10^{-5}	3.33×10^{-2}	3.22×10^{-2}	7.01×10^{-5}
	AB	8.11×10^{-4}	1.52×10^{-2}	1.61×10^{-2}	-8.73×10
	MSE	6.26×10^{-5}	3.37×10^{-2}	3.28×10^{-2}	6.29×10^{-5}

all 3 regions of interest. Analogous simulation studies were performed for the 3 common Michaelis-Menten inhibition forms by allowing the concentration of the inhibitor substance to range from 0 to 9*K_I . In all cases the "parameters in the denominator" version was estimated more efficiently than the traditional parameterization. For the non-competitive and un-competitive Michaelis-Menten forms (defined in Table 2.1), the Levenberg-Marquardt algorithm failed to converge for the standard parameterization in many cases but converged in all cases for the re-parameterized version. The results for the un-competitive and non-competitive results will not be discussed further since the superiority of the parameters in denominator method is clear in these cases. For competitive inhibition a simulation study analgous to that above was conducted. The true V_{max} and K_M were 1 and 0.5 respectively. The boundaries of regions A, B and C were set at substrate concentrations [0,0.75), (0.75,9.5) and (9.5,49.5] respectively. The value of K_I was set at 2 and the concentration of inhibitor substance was allowed to vary from 0 to 9*K_I . Table 2.4 shows that the "parameters in the denominator" produces less correlated parameter estimates no matter which region (A,B or C) the data are sampled from. The summary statistics in Table 2.5 suggest, however, only a slight improvement in the reparameterized version of the model.

These simulation studies indicate that the "parameters in the denominator" versions of the models are more identifiable and have better convergence properties but perform only slightly better in terms of variance and bias. Certainly in dynamic systems models convergence is major consideration. A limitation of this simulation study is the assumption of independent and identically distributed Gaussian error. Typically the error in initial rate data is proportional to $V_o$or may have some other error structure depending on the method of measuring V_o. In such cases Ratkowsky (1986)

Table 2.4: Correlations among competitive Michaelis-Menten parameter estimates obtained from 25 data points generated uniformly in the region indicated with a simulated error variance of 0.02. The true $V_{\max}$, K_{M} and K_{I} were 1 0.5, and 2 respectively. The region boundaries were set at substrate concentrations 0.75, 9.5 and 49.5. Reported values are the average of 100 trials.

		Region		
Model	Parameters	A	B	C
Traditional	$V_{\max}$ and K_{M}	0.977	0.736	0.730
	$V_{\max}$ and K_I	0.805	0.552	0.658
	K_{M} and K_I	0.900	0.960	0.992
Re-parameterized	α and β	-0.912	-0.683	-0.721
	α and γ	-0.091	-0.256	-0.491
	β and γ	-0.133	-0.321	-0.098

recommends using weighted least squares rather than either linearization on a double-reciprocal plot or use of a transformation. Since this method of re-parameterization is easy to accomplish and seems superior in several respects to the standard parameterization, its use is recommended in dynamical systems modeling even if the full range of dynamics are not directly observable in the data.

Table 2.5: Comparison of emperical variance (EV), average asymptotic variance (AAV), average bias (AB) and mean squared error (MSE) (for uninhibited Michaelis-Menten. Reported values are the average of 100 trials.

Region	Statistic	V_{max}	K_M	K_I	α	β	γ
	EV	1.03×10^{-3}	1.01×10^{-3}	5.23×10^{-3}	2.66×10^{-4}	9.96×10^{-4}	8.18×10^{-6}
A	AAV	1.20×10^{-3}	1.17×10^{-3}	5.79×10^{-3}	3.10×10^{-4}	1.19×10^{-3}	8.02×10^{-6}
	AB	-5.77×10^{-3}	-6.43×10^{-3}	-1.60×10^{-2}	-2.49×10^{-3}	3.39×10^{-3}	4.58×10^{-4}
	MSE	1.07×10^{-3}	1.05×10^{-3}	5.48×10^{-3}	2.73×10^{-4}	1.01×10^{-3}	8.39×10^{-6}
	EV	6.35×10^{-6}	1.76×10^{-4}	3.34×10^{-3}	1.64×10^{-4}	6.94×10^{-6}	2.82×10^{-6}
B	AAV	1.02×10^{-5}	2.29×10^{-4}	3.72×10^{-3}	1.96×10^{-4}	1.02×10^{-5}	3.10×10^{-6}
	AB	4.10×10^{-4}	3.09×10^{-3}	1.33×10^{-2}	1.20×10^{-3}	-1.39×10^{-4}	-1.82×10^{-4}
	MSE	6.52×10^{-6}	1.85×10^{-4}	3.52×10^{-3}	1.66×10^{-4}	6.96×10^{-6}	2.85×10^{-6}
	EV	4.50×10^{-6}	3.37×10^{-3}	6.09×10^{-2}	2.46×10^{-3}	3.69×10^{-6}	1.76×10^{-5}
C	AAV	4.47×10^{-6}	3.19×10^{-3}	5.72×10^{-2}	3.12×10^{-3}	4.49×10^{-6}	1.78×10^{-5}
	AB	9.97×10^{-4}	3.88×10^{-2}	1.56×10^{-1}	1.65×10^{-3}	-1.83×10^{-5}	7.54×10^{-4}
	MSE	5.50×10^{-6}	4.87×10^{-3}	8.54×10^{-2}	2.46×10^{-3}	3.69×10^{-6}	1.81×10^{-5}

In conclusion, the following guidelines are suggested for applying model re-parameterizations when the goals of modeling include inference about 1 or more model parameters.

1. Parameters to be estimated can be arranged in clusters so as to provide new aggregate parameters. Parameter aggregation should be carried out so as to produce parameters with good estimation properties (e.g. reduced covariance with other parameter estimates).

2. A nondimensional parameter cluster should be eliminated (i.e. set to 1) only if all parameters within the cluster are known. This will allow the data on the variables involved to be scaled appropriately prior to model fitting.

3. If a nondimensional parameter cluster is to be eliminated prior to parameter estimation, then the data values of the variables involved in the scaling must be likewise scaled.

4. Subsequent to parameter estimation, the inverse scaling relationships must be appropriately applied to parameter estimates as well as their standard errors. This may require series approximations to the expectations and standard errors of the original parameters (for example see Casella and Berger (1990)).

2.2 Model Fitting

2.2.1 Parameter Optimization

The simulation study presented in the previous section was carried out using a non-linear least squares algorithm. It will be shown in this section that, unfortunately, standard non-linear least squares algorithms are usually not suitable for estimating

the parameters of many dynamical systems models. Furthermore, it will be shown that the objective surfaces corresponding to such problems may contain large relatively flat regions.

To illustrate some of the difficulties, consider the classic Lotka-Volterra system for the interaction between a predator (N_1) and its prey (N_2).

$$\frac{dN_1}{dt} = \kappa N_1 N_2 - \delta N_1 \tag{2.7}$$

$$\frac{dN_2}{dt} = \alpha N_2 - \gamma N_1 N_2 \tag{2.8}$$

with initial conditions $N_1(0) = N_{10}$ and $N_2(0) = N_{20}$. Here α is the growth rate of the prey, γ is the rate of capture, κ is the conversion efficiency and δ is the predator death rate. The solution of such a problem can be obtained by numerical integration. Two approaches can be taken to maintain numerical stability and to maintain the error of numerical integration below some acceptable tolerance. First, a time-step can be chosen small enough so that these conditions are always satisfied. This is computationally wasteful however since a larger time-step could be taken without sacrificing accuracy in regions where the solution is changing slowly. The second approach is to adaptively modify the time-step in the integration procedure – allowing larger time-steps where the solution is changing slowly and smaller time-steps where the solution is changing rapidly (Press et al., 1992). Since it is difficult to know ahead of time what the minimum time-step requirement will be, the adaptive procedure is the only reasonable method for obtaining accurate solutions for most applications.

While the adaptive procedure is beneficial in terms of maintaining accuracy during the integration, its use has dire consequences when used within iterative parameter estimation routines such as non-linear least-squares algorithms (Hairer et al., 1987,

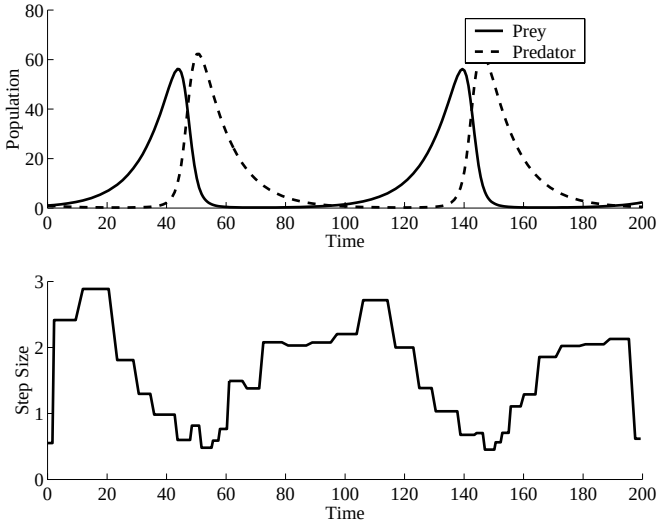

Figure 2.4: The Lotka-Volterra model with $\alpha = 0.01$, $\gamma = 0.009$, $\kappa = 0.1$, and $\delta = 0.1$. In the bottom figure, the time-steps (in units of time) used by the solver are plotted against time.

p. 183). Figure 2.4 shows a simulation of the Lotka-Volterra system and the corresponding time-steps used by the adaptive integration procedure. These figures clearly show that in regimes where the rate of change of a variable is larger, the adaptive ODE solver takes smaller time steps. Conversely where the solution is changing more slowly, the solver can maintain accuracy with a larger time step. Note the discontinuous nature of the jumps in the step-size. Suppose the Lotka-Volterra model was to be fit to data on predator and prey populations which had been taken over time. Suppose also that a nonlinear least squares method was to be used to do the parameter estimation. In many non-linear least squares algorithms it is necessary to calculate the gradient of the objective function with respect to the parameters. This is used in determining a favorable direction in parameter space in which to proceed with the minimization of the least squares objective function. Fig. 2.5 shows the partial derivative of the prey-predator population vector with respect to the parameter γ

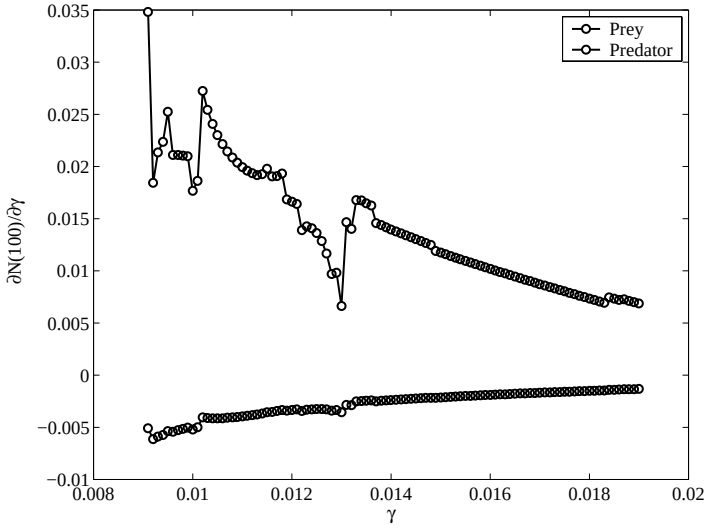

Figure 2.5: Partial derivative of the solution with respect to γ at t=100. The non-smoothness of the gradient with respect to perturbations of the parameter spells disaster for most non-linear least squares algorithms.

at $t = 100$. This choice of $t = 100$ is arbitrary and similar results are observed at any other time point and with any of the other parameters. A centered finite difference was used with step size of 2.e-6 and the derivative estimates did not change if the step size was decreased any further. The non-smoothness of the gradient with respect to perturbations of the parameter arises from the internal switching of the integration procedure between time-step sizes. This spells disaster for non-linear least squares algorithms that depend on evaluation of gradients in the objective function. Depending on the "jump" the algorithm will respond as if it had just found a value much better or much worse than a nearby point. In either case, the optimization will likely get trapped. The other problem which is encountered in estimating parameters in dynamical systems is the sum of squared errors surface can be flat or can be convoluted possessing many local minima. The sum of squares error surface for the Lotka-Volterra system possesses both.

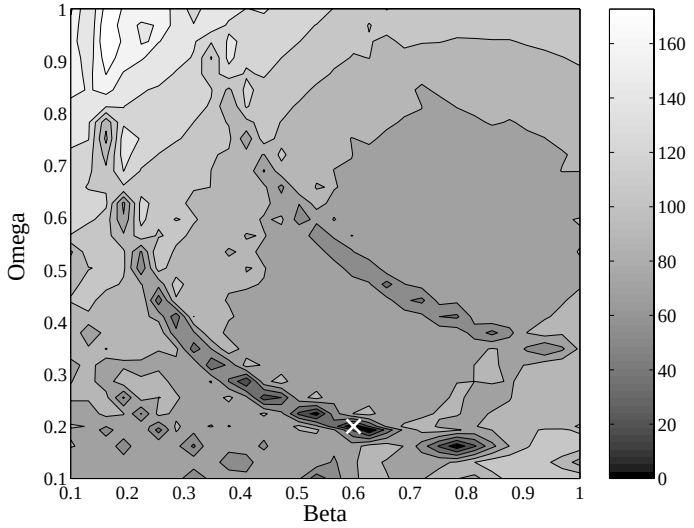

Figure 2.6: Contour plot of the SSE surface of the 2-parameter Lotka-Volterra model. The true parameter set ($\beta = 0.6, \Omega = 0.2$) is indicated by the white "X".

To illustrate these points the above Lotka-Volterra model will be nondimensionalized so that it contains only 4 nondimensional model parameters Ω, β and 2 initial conditions N_{10}^* and N_{20}^*

$$\frac{dN_1^*}{dt^*} = N_1^* N_2^* - \Omega N_1^* \tag{2.9}$$

$$\frac{dN_2^*}{dt^*} = \beta N_2^* - N_1^* N_2^* \tag{2.10}$$

where $\Omega = \delta t'$ and $\beta = \alpha t'$. This will simplify viewing the sum of squared errors surface. Clearly, Ω or β could be eliminated by scaling time and have a model in just 1 parameter, but this would mask the complexity of the fitness landscape for the 4 parameter model too much. The residual sum-of-squares are used as the fitness criterion. The calculated sum of squared errors surface is plotted in Fig. 2.6. The true parameter set, indicated by the white "X", is nestled in a locally quadratic minimum

as desired. The problem is that there are many such depressions in this surface with nearly the same minimum value. The appearance of troughs in the objective surface are characteristic of periodic data in which the modes of the candidate solution "slide" past the modes of the data as a particular parameter is varied. There are also large flat areas to contend with. These aspects of the sum of squared errors surface make finding the optimal parameter set nearly impossible for any direct-search method.

When presented with such a surface, it is advisable to use a search method which does not rely on local gradients and is not easily "fooled" by local minima. Search methods which embody these characteristics include Simulated Annealing, Nelder-Mead Simplex and genetic algorithms (Press et al., 1992). While these methods are less likely to be trapped in local minima they are not guaranteed to find the global minimum. Also, unlike most non-linear least squares algorithms, these methods generally do not produce estimates of the parameter covariance matrix at convergence. Estimates of parameter variance and covariance must therefore be calculated by some other method subsequent to parameter estimation.

The Lotka-Volterra model is a pedagogical example of how difficult fitting a dynamical model to data can be. Without the use of an adaptive ODE solver the accuracy of the solution is unknown. Use of a fixed ODE solver is not recommended since its use would almost surely result in biased parameter estimates. The number of local minima and the "flatness" of the surface about the minimum will depend in large part on the number of model parameters and the degree of the non-linearity in the model's functional forms. The use of an iterative method which does not use gradients (Nelder-Mead Simplex, genetic algorithms etc.) is recommended. In the next section, some objective functions are introduced that may be more appropriate to fitting dynamic model systems to data.

2.2.2 Objective Functions for Fitting Dynamical Systems

Assuming that a suitable minimization algorithm can be found, it is still necessary to specify the objective function to be minimized. When the dynamics of a single population or quantity (state variable) is to be modeled and the underlying model to be fit can be assumed to be true and has a known error structure, there are many suitable measures of goodness-of-fit common in the statistical literature. The most well-known are sum-of-squares error (SSE) and the method of maximum-likelihood.

When the underlying model can not be assumed to be true, as is always the case in mechanistic models, the situation is quite different. Because the underlying model can not be assumed to be true there is almost always systematic lack of fit in the model predictions. The goodness-of-fit measures presented in Table 2.6 were compiled by (Ross, 1996) and focus on particular qualities of the fit that may be desirable in a mechanistic model context. These include measures that can be used to balance over-prediction and under-prediction, minimize bias or minimize relative errors.

These measures may be limited when used in model fitting contexts where there are multiple state variables. In particular, these measures do not address the fact that the state variables may be on very different scales (recall the earlier discussion of nondimensionalization). If these measures are used as presented, they can bias an optimization toward state variables with greater variability and observations with larger relative magnitude.

To deal with the issue of multiple state variables having different characteristic scales (Omlin et al., 2001) suggested the following weighted least-squares measure as

Table 2.6: Goodness-of-fit measures. In the explanation of each measure, P represents the predicted values and O represents the observed values and N represents the number of observations.

AF	Accuracy factor	$10^{\sum(\lvert\log_{10}(P/O))/N)\rvert}$	Gives a multiplicative factor to generate an "expected range" of residuals. For example if AF = 1.25 the the expected range of deviations is (P/1.25,P*1.25).
BF	Bias factor	$10^{\sum(\log_{10}(P/O)/N)}$	BF = 1 indicates no bias, BF > 1 indicates over-prediction. BF < 1 indicated under-prediction.
MARR	Mean abs. relative residual	$\frac{(1/N)\sum(\lvert(O-P)/O\rvert)}{0.01}$	MARR is the average percent difference of predictions from observed.
MRPR	Mean relative percent residual	$\frac{(1/N)\sum((O-P)/O)}{0.01}$	MRPR > 0 indicate under-prediction. MRPR < 0 indicate over-prediction.
RMSR	Root mean square residual	$(\sum((O-P)^2)/N)^{1/2}$	Will be zero when both the bias and variability of the prediction are zero. Penalizes equally for over- prediction and under-prediction so is usually considered inappropriate when predictions are constrained to an interval.

a goodness-of-fit objective function in model fitting.

$$WSS(\theta) \quad = \quad \sum_{k=1}^{n_y} \sum_{j=1}^{n_{t,fit}} \sum_{i=1}^{n_{y_k,z}} \left(\frac{y_{\text{meas,k,j,i}} - y_k(z_{y_k,i}, t_j, \theta)}{sc_{y_k}} \right)^2 \tag{2.11}$$

Here $y_{\text{meas,k,j,i}}$ is the measured value of the k^{th} state variable y_k under the i^{th} treatment scenario $z_{y_k,i}$ at time t_j and $y_k(z_{y_k,i}, t_j, \theta)$ is the corresponding model prediction given model parameters θ. The quantity which allows this measure to be used across all state variables is sc_{y_k}. Omlin et al. describes sc_{y_k} as a scaling factor "...to give all measured variables similar influence on the estimates of parameters." Unfortunately Omlin et al. give no guidance on how to select such scales.

Experimentation has been done with taking sc_{y_k} equal to the range taken across all observations on y_k and with sc_{y_k} equal to the standard deviation over all observations on y_k. *While there was no clear advantage to either approach, the range may be considered inferior since it is less robust to outliers.* In either case, division by sc_{y_k} leads to a non-dimensional measure of fit by normalizing the residuals on each state variable so that they are all of the same order of magnitude. Since use of the range may be inferior, the use of the standard deviation as the normalizing quantity is recommended. Furthermore to make measures of fit less biased due to unequal sampling frequencies across the data sets the scaling can be taken to be $n_{yk}(sc_{y_k})$. That is, each residual is normalized by the number of observations on the variable y under conditions k.

In conclusion, the use of an iterative optimization procedure is suggested which does not rely on gradients to optimize WSS with a scaled standard deviation as a scaling factor as a suitable objective function for model fitting of systems with multiple state variables.

2.3 Model Analyses

Methods to evaluate model behavior and robustness after parameters have been obtained are now considered. The most common approach is to measure the sensitivity of model output to perturbations in the parameters. In this approach, a small, usually relative, perturbation about the estimated value of a parameter is used to generate new model predictions. If the new model predictions are close to those for the unperturbed parameter then the model is said to be robust with respect to that parameter.

In the case of bacterial growth, for example, one might be interested in determining the sensitivity of cell density predictions. The sensitivity of the $\log_{10}$ of cell density can be defined as

$$\frac{\partial \log_{10} N}{\partial P_i} \approx \frac{\log_{10} N \mid_{P_i+0.05P_i} - \log_{10} N \mid_{P_i-0.05P_i}}{0.1P_i} \tag{2.12}$$

It is often the case, however, that the sensitivity of model output to perturbations in 1 parameter depends to a large extent on the value of other parameters. To assess dependencies in a sensitivity analysis one can use a technique similar to a multi-factor ANOVA (Swartzman and Kaluzny, 1987).

A conceptually straight-forward way of designing such an experiment is to use a 3^n factorial arrangement of parameter perturbations where n is the number of parameters in the analysis. The problem here, typical of factorial designs, is that a very large number of simulations are needed for even moderate n.

There are several designs common in the response surface literature that can be used to estimate the interactions with fewer number of design points. The design used is a central composite design (Keuhl, 2000). A central composite design is constructed

by first generating a 2^n factorial and then "face points" in each of the cardinal directions. There are also c "center points" placed at the location of the unperturbed parameter set. A central composite design requires $2^n + 2n + c$ design points compared with the 3^n required by the factorial design. Thus, when there are more than 2 parameters the central composite design can require fewer points (depending on c). The utility of $c > 1$ is that it allows a better estimate of measurement error in an experimental setting. In a sensitivity analysis, however, there is no measurement error as such, so it is assumed that $c = 1$ (Swartzman and Kaluzny, 1987). A central composite design for a 2 parameter model is displayed in Fig. 2.7. This design has 9 points which is the same as the 3^2 factorial. For all dimensions higher than 2, however, the central composite offers a saving. A central composite design for a 3 parameter model is displayed in Fig. 2.8. This design uses 15 points while the factorial requires $3^3 = 27$. The central composite design permits calculation of "main effects", "first-order interactions" and "quadratic effects." This allows assessment of the magnitude, direction and concavity of the sensitivity surface with respect to the model parameters.

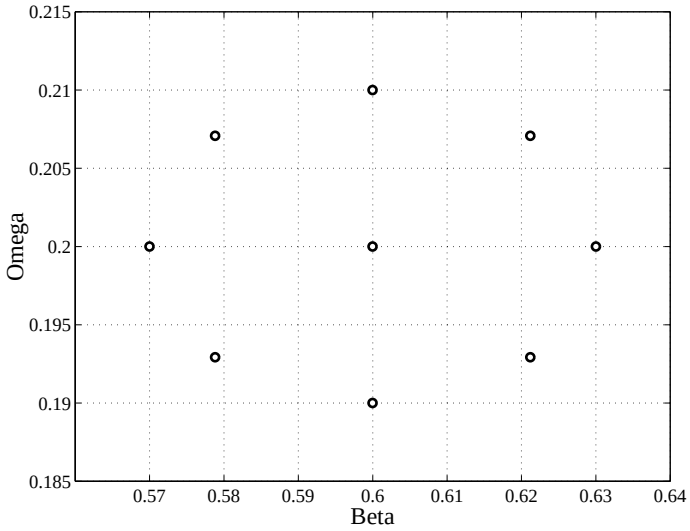

Figure 2.7: Central composite design for 2 parameter multivariate sensitivity analysis.

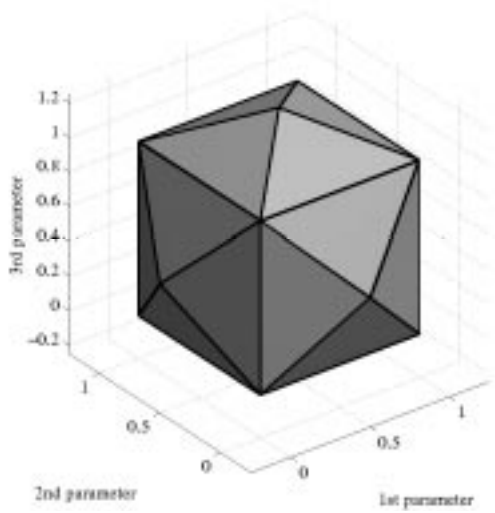

Figure 2.8: Central composite design for 3 parameter multivariate sensitivity analysis. Although it is not clear from this diagram, there is also a design point in the center of the design.

Chapter 3

An Energy-Based Dynamic Model for Variable Temperature Vegetable Fermentation by Lactococcus lactis

In this chapter, some of the approaches to predictive microbiology described in Chapters 1 and 2 are applied and extended. In particular, a mechanistic mathematical model is developed for predicting the progression of batch fermentation of cucumber juice by the homolactic bacterium *Lactococcus lactis* under variable environmental conditions. In order to overcome the deficiencies of currently available models, a dynamic energy budget approach is used to model the dependence of growth on current as well as past environmental conditions. Using parameter estimates from independent experimental data, the model is able to predict the outcomes of three different temperature shift scenarios. Sensitivity analyses elucidate how temperature

affects the metabolism and growth of cells through the lag, exponential, stationary and death phases of fermentation and reveal that there is a qualitative reversal in the factors limiting growth between low and high temperatures. The model has an applied use as a predictive tool in batch culture growth. It has the added advantage of being able to suggest plausible and testable mechanistic assumptions about the interplay between cellular energetics and the modes of inhibition by temperature and end product accumulation.

3.1 Derivation of Model Equations

There is some evidence to suggest that the adenine nucleotide content (ATP + ADP + AMP) of bacteria partially reflects the capacity of the cell for growth and is directly affected by environmental stresses such as pH, temperature and starvation (Mercade et al., 2000; Metge et al., 1993; Chibib and Tholozan, 1999; Jetton et al., 1991). While adenylate phosphates certainly contribute to the cell's total "energy", this measure may not represent absolutely the cell's ability to grow and cope with environmental stresses (Kooijman, 2001). Therefore, a formalism similar to that of Eqn. 1.20 is adopted and the internal quota of "energy" is denoted Q and the per-cell fraction is denoted q.

Bacterial growth should cease when q reaches 0. But the growth rate of the bacteria should be limited to some maximal rate α as the internal pool of "energy" gets large. Conversely, the death rate of cells will approach a maximum rate δ_1 when q is zero but should vanish when q is large. These assumptions are satisfied by the

following equation governing the growth of the bacteria whose density is $N(t)$.

$$\frac{dN}{dt} = \alpha(\frac{q}{k_{q1}+q})N - \delta_1 exp(-\delta_2 q)N \tag{3.1}$$

The rate at which the death rate goes to zero for increasing q is controlled by δ_2. Note that in this formulation, the *observed* growth rate is the actual growth rate minus the death rate.

Homolactic bacteria such as *Lactococcus lactis* derive the vast majority of their "energy" by consuming the glucose present in the growth medium. Since an initial pool of "energy" is needed to initiate glycolysis, the following saturating kinetics equation is suggested for the consumption of glucose (S) by bacteria.

$$\frac{dS}{dt} = -\mu_1(\frac{q}{k_{q2}+q})NS \tag{3.2}$$

Thus, as q grows large, the rate of glucose consumption approaches μ_1. k_{q2} is the value of q at which the glucose consumption rate is half its maximum μ_1. In this application, malic acid is present and malolactic conversion to lactic acid accounts for about 1/3 of the total lactic acid produced during the fermentation. Therefore, an equation similar to the above for the malolactic reaction is included.

$$\frac{dM}{dt} = -\mu_2(\frac{q}{k_{q3}+q})NM \tag{3.3}$$

μ_2 is the maximum malolactic rate and k_{q3} is the half-saturation value. The stoichiometry of homolactic conversion of glucose to lactic acid (L) and malolactic conversion

of malic acid to lactic acid is well known (Jay, 1992). It may state that

$$\frac{dL}{dt} = 2\,[\text{glucose consumed}] + 1\,[\text{malic reduced}] \tag{3.4}$$

where the glucose consumed and malic reduced are $-dS/dt$ and $-dM/dt$ respectively.

The internal pool of "energy" will increase in proportion to the glucose consumed. However, "energy" must be used to cope with acidic conditions and temperature changes. The following equation is a tentative sketch of the dynamics of the "energy" state of the population (Q).

$$\begin{aligned} \text{Change in Q per hour} &= \text{Glucose consumed} - \text{Acid Toxicity} \\ &\quad -\text{Temperature adaptivity} \end{aligned} \tag{3.5}$$

Specifically, the contribution to Q is taken to be proportional to the glucose consumed. The acid toxicity is taken to be proportional to the total concentration of lactic acid (ionic and molecular forms).

When glucose in the medium is brought inside the cell (Eqn. 3.2) to be converted to "energy", an adjustment is required in order to account for the change in volume from the fermentation flask to cell. Since the volume of cells is not known precisely, this adjustment factor is confounded with the glucose-to-energy conversion parameter. It should be noted that this is actually a nice feature of the model because having to know the precise volume of a cell would severely limit the extensibility of the model.

Modeling the "acid toxicity" term in Equation 3.5 is a difficult task and there is currently a great deal of debate about what is the ultimate cause of acid toxicity. There are clearly at least 2 main components to acidity for lactic acid bacteria.

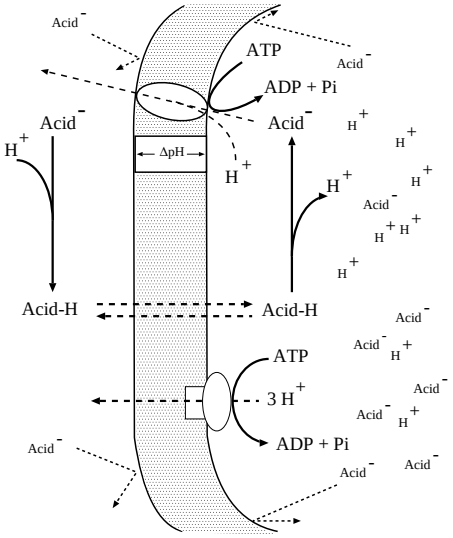

Figure 3.1: Accumulation of organic acid anions and proton imposes a drain on the intracellular energy of a cell. A single ATP is hydrolyzed to transport out 3 protons. Organic acid is actively transported out of the cell. The mechanism of acid transport shown is electo-neutral symport.

The first is pH and the second is the intra-cellular concentration of lactic acid (see Figure 3.1). These 2 quantities are themselves related by the Henderson-Hasselbach equation (Butler and Cogley, 1998). However, in a complex growth medium many other acids may be present and issues regarding shifts in buffer capacity and ionic strength become important. Dealing with these issues will be delayed until Chapter 4. For now it is assumed that dependence on total lactic acid is sufficient to explain end product toxicity and suggest the functional form

$$\text{Acid Toxicity} = \kappa q N L \tag{3.6}$$

$$= \kappa Q L \tag{3.7}$$

where κ is the per-cell energy cost of lactic acid.

Focus is now on modeling the "temperature adaptivity" term in Eqn. 3.5. Although the goal of Van Impe et al. (1995) to model temperature dynamically is shared a different approach than the modified Ratkowsky model must be taken. First, temperatures (T) are limited below the optimal temperature of the bacterium (T_{opt}) since these are the ranges relevant to food production. As discussed in Chapter 1, it is reasonable to assume that large deviations from the optimal temperature require large physiological adjustments.

$$\text{Temperature adaptivity} \quad = \quad -\tau(T - T_{opt})^2 |\frac{dT}{dt}|Q \tag{3.8}$$

Based on purely enzyme kinetic theory, lower temperatures should also result in reductions in the growth, glucose consumption, malic acid conversion and the acid toxicity rates (recall the square root model of Ratkowsky). These trends were noted by Chibib and Tholozan (1999) as well as in these data. A rate's reduction due to temperature (T) can be modeled by dividing the rate by an inhibition term F_T of the form

$$F_T \quad = \quad 1 + \left(\frac{T - T_{opt}}{K_{T^*}}\right)^2 \tag{3.9}$$

$$\tag{3.10}$$

Where K_{T^*} is a temperature sensitivity term (smaller values translate into greater sensitivity). Thus a deviation from optimal temperature is considered to manifest itself as a classical *non-competitive* inhibitor of these reactions (Cornish-Bowden,

1995).

The resulting model consists of a system of five differential equations with dependent variables for cell density (N), the extracellular concentrations of glucose (S), lactic acid (L) and malic acid (M) and intra-cellular "energy" (Q). Independent variables are time t and temperature T. The model parameters, their biological interpretations, units and estimates are given in Table 3.2. The differential equations are as follows:

$$
\begin{aligned}
\textbf{CFU:} \quad & \frac{dN}{dt} = \alpha\left(\frac{q}{k_{q_1}+q}\right)N - \delta_1 exp(-\delta_2 q)N \\
\textbf{Glucose:} \quad & \frac{dS}{dt} = -\frac{\mu_1}{F_{T_1}}\left(\frac{q}{k_{q_2}+q}\right)NS \\
\textbf{Malic Acid:} \quad & \frac{dM}{dt} = -\frac{\mu_2}{F_{T_1}}\left(\frac{q}{k_{q_3}+q}\right)NM \\
\textbf{Energy:} \quad & \frac{dQ}{dt} = \beta\left[\frac{\mu_1}{F_{T_1}}\left(\frac{q}{k_{q_2}+q}\right)NS\right] - \frac{\kappa}{F_{T_2}}LQ \\
& \qquad -\gamma\left[\alpha\left(\frac{q}{k_{q_1}+q}\right)N\right] - \tau\frac{dT}{dt}(T-T_{opt})^2Q \\
\textbf{Lactic Acid:} \quad & \frac{dL}{dt} = 2\left[\frac{\mu_1}{F_{T_1}}\left(\frac{q}{k_{q_2}+q}\right)NS\right] \\
& \qquad +\left[\frac{\mu_2}{F_{T_1}}\left(\frac{q}{k_{q_3}+q}\right)NM\right]
\end{aligned}
\tag{3.11}
$$

Table 3.1: State variables of the model[a].

Symbol	*Meaning*	*Units*	*Observed Range*
t	Time	h	0–256
T	Temperature	°C	10–30
N	Cell density	CFU mL^{-1}	$1.0x10^6$–$2.0x10^9$
S	Glucose in the medium	mM	0–35
M	Malic acid in the medium	mM	0–12
Q	Intracellular energy	fraction of initial "energy"	0–5
L	Lactic acid in the medium	mM	0–35

[a] Q/N was normalized to 1 at the time of inoculation (t=0). Therefore, Q/N measures energy as a fraction of the energy quota that is typically present in an overnight culture.

where

$$q = Q/N\text{and} \tag{3.12}$$

$$F_T = 1 + \left(\frac{T - T_{opt}}{K_{T^*}}\right)^2 \tag{3.13}$$

Table 3.2: Model parameter descriptions and estimates.

Parameter	*Description (unit)*	*Estimate*
α	Maximal bacterial growth rate (h^{-1})	3.42
μ_1	Maximal glucose consumption rate (CFU mL^{-1})(mM^{-1})	8.95×10^{-10}
μ_2	Maximal malate conversion rate (CFU mL^{-1})(mM^{-1})	2.41×10^{-9}
kq_1	Value of q at which cell growth rate equals $\alpha/2$ (unitless)	8.03
kq_2	Value of q at which glucose consumption rate equals $\mu_1/2$ (unitless)	1.07
kq_3	Value of q at which glucose consumption rate equals $\mu_2/2$ (unitless)	1.35×10^{-2}
δ_1	Death rate when $q = \infty(h^{-1})$	7.02×10^{-2}
δ_2	Death rate sensitivity to changes in q (unitless)	3.73×10^{3}
β	Conversion rate of glucose into energy (mM^{-1})	6.80×10^{8}
γ	Energy required for cell division (CFU mL^{-1})	2.07
κ	Energy cost per 1 mM lactic acid ($\text{mM}^{-1}h^{-1}$)	0.220
KT_1	Sensitivity of metabolic processes to deviations from optimal temp (°C)	8.11
KT_2	Sensitivity of lactic acid inhibition to deviations from optimal temp (°C)	6.28
T_{opt}	Optimal temp (°C)	37 (assumed)
τ	Energy cost for transient temp adjustment (°C $^{-1}$)	1.60×10^{-3}

3.2 Materials and Methods

3.2.1 Bacterial strains and media.

Lactococcus lactis subsp. *lactis* strain LA221 transformed with the chloramphenicol resistance gene pGK12 (Breidt and Fleming, 1998) was obtained from the USDA culture collection. LA221 was grown on M17 broth (Difco Laboratories, Detroit, Mich.) containing 1.5% agar (Difco), 1% glucose (Sigma Chemical Co., St. Louis, Mo.) and 5 ug mL^{-1} chloramphenicol. Growth experiments were conducted in cucumber juice medium. Cucumbers were pureed and pressed to render raw juice, which was frozen until needed. The raw juice was thawed and then clarified by heating at 85°C for 5 min, followed by centrifugation at 12,000 rpm for 20 min and filter sterilized. The CJ medium was prepared by adding 600 mL juice to 400 mL dH2O. The diluted juice was supplemented with 2% NaCl and then filter-sterilized as described by Daeshel et al. (1984).

3.2.2 Fixed temperature experiments

Overnight cultures were prepared by growing LA221 in CJ at 30°C . Water-jacketed jars (Wheaton, Millville, N.J.) were filled with 200 mL of fresh CJ and inoculated at 1×10^6 CFU mL^{-1} of bacterial culture. Each flask was sealed with a silicone stopper that contained a sterile syringe sample port, through which an 18-gauge, 10 cm needle was passed. The growth medium was kept well-mixed by a magnetic stirrer. Compressed nitrogen was humidified by sparging through deionized water, filtered (0.2 um Millex-FG50 filter, Millipore Corp., Bedford, Mass.), and released into the headspace of the fermenter jars at a rate of 1.3 L h^{-1}. Temperature during the

fermentation was controlled by a circulating water bath (NESlab RTE-211; NESlab, Portsmouth, N.H.). The temperature of the growth medium was monitored directly by sterile thermocouples inserted through the silicone stoppers and recorded by a microcomputer (OM-3000; Omega, Stamford, Conn.). Growth observations at 10, 20 and 30°C included quantification of CFU mL^{-1}, glucose, malic acid, and lactic acid concentrations. Growth at a particular temperature was monitored until all phases of growth had been observed.

The data consisted of 2 replicates (jars) at each temperature. Each pair of replicates was carried out side-by-side. Because only a single water bath was available to maintain the desired temperatures, the experiments for different temperatures were carried out on different days. The sampling of the jars at the different temperatures resulted in a total of 22 sample points over time at 10°C (a total of 44 samples), 14 samples points over time at 20°C (a total of 28 observations) and 13 sample points over time at 30°C (a total of 26 observations). For each sample, cell density, glucose concentration, malic acid, lactic acid and pH were measured and recorded. Thus the data consisted of 98 5×1 vector-valued observations.

3.2.3 Biological assays

When a sample was desired, a sterile disposable syringe (1 cc, BectonDickinson, N.J.) was used to withdraw a 1-mL sample from the fermentation flask sample port. Cells were removed from 1-mL samples by centrifugation at 13,000 rpm for 1.5 min. High performance liquid chromatography (HPLC) analyses of the supernatant quantified total lactic acid, glucose and malic acid. HPLC was carried by the single injection procedure of McFeeters (1993). The pH of the medium was determined using an

electronic pH meter (IQ 200; IQ Scientific Instruments, Inc., San Diego, Calif.). Cell density (CFU mL^{-1}) was determined by the spiral plate count method using an Autoplate 4000 Automated Spiral Plater (Spiral Biotech, Bethesda, Md.) and a Protos Plus Colony Counter (Bioscience International, Rockville, Md.).

3.2.4 Statistics and programming

All computing was performed on a 300 Mhz Ultra Sparc 10 processor (Sun Microsystems, Palo Alto, Calif.). MATLAB^{TM}(ver. 5.3, MATLAB functions are indicated in all-caps) software was used to solve the system of equations 3.11 and for all other programming. The equations were solved using the adaptive stiff ODE solver (ODE15S). The ODE solver often required temperatures and rates of temperature change at times different from when they were actually sampled. This problem was overcome simply by using linear interpolation (INTERP1) to estimate the temperature values at the times requested. The derivative of temperature with respect to time was calculated prior to parameter estimation using a centered finite difference approximation (GRADIENT) of the temperature data. During the solution of the ODEs, the gradient was obtained by linear interpolation to the desired time point.

A single set of parameters was estimated from the fixed temperature data with the intent that the fitted model would predict the data from each of the fixed-temperature experiments equally well. Specifically, parameters were estimated using all of the fixed temperature data at once via a weighted least-squares method (see Chapter 2 p. 39). In the following, it is assumed that i indexes the tempertures, j indexes the runs within a particular temperature and k indexes the observations within a run. For each of the a temperatures there are b_i runs and within the j^{th} run there

are c_j observations. The squared residual of a system variable was normalized by the standard deviation of the observed data on that variable across all temperatures (s_*). The number of samples taken during the fermentations was not constant across temperatures or runs. Therefore, to make the fits have equal weight, the sum of squared errors for a particular temperature (i) was divided by the number of sample points taken for that temperature ($\sum_{j=1}^{b_i}(c_j)$). The weighted squared residuals were then summed to give a measure of total lack of fit. The following equation explicitly describes the formulation of the weighted sum of squared errors.

$$\begin{aligned} WSS \;=\; & \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{i,j}} \frac{(\log_{10}(\text{Obs. CFU/mL}_{i,j,k}) - \log_{10}(\text{Pred. CFU/mL}_{i,k}))^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{CFU/mL}})} \\ & + \frac{(\text{Obs. Malic}_{i,j,k} - \text{Pred. Malic}_{i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{Malic}})} \\ & + \frac{(\text{Obs. Lactic}_{i,j,k} - \text{Pred. Lactic}_{i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{Lactic}})} \\ & + \frac{(\text{Obs. Glucose}_{i,j,k} - \text{Pred. Glucose}_{i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{Glucose}})} \end{aligned} \tag{3.14}$$

Note in the above that the log of cell density was employed. The data for the cell density was heteroscedastic, with sample variance approximately proportional to the magnitude of the density. A flexible distribution that has this property (constant coefficient of variation) is the Gamma distribution. Schaffner (1998) gives further evidence why cell count data should be considered to have a Gamma distribution. To approximately normalize the variance a log transform of the cell density was employed. The HPLC data (lactic acid, malic acid and glucose), however, were assumed to have normally distributed errors. Initial conditions were also estimated from the data for each of the state variables except for Q which was initialized to the cell density so

that $q(0) = 1$.

Since an adaptive ODE solver was used, neither the Jacobian nor the Hessian of the likelihood function depends smoothly on perturbations in the parameters (Hairer et al., 1987). As a result, traditional direct-search gradient-based algorithms failed. Differential Evolution (Storn, 1996), a genetic algorithm, was used to achieve a parameter set which approximately minimized WSS, followed by as many Nelder-Mead Simplex (FMINSEARCH) iterations as required to obtain convergence. A function tolerance of 1×10^{-4} and a parameter tolerance of 1×10^{-4} were used as the convergence criteria of the Nelder-Mead Simplex procedure.

3.2.5 Validation studies

Validation of the calibrated model was accomplished by comparing predictions and data from variable environment experiments. In all cases, two or more independent replicates of the fermentations were carried out. In the first scenario, the temperature of the medium was maintained at 30°C for 3.75 h, and then it was dropped to 10°C . This temperature change was accomplished over a period of about 15 min. In order to compare the model's predictions about the effect of the previous energy state on growth, a second scenario was conducted. In this scenario, cells were grown at 30°C for 54.42 h. Then 100 μL of fermented broth were inoculated into 200 mL of fresh CJ also at 30°C for 68.5 h. In a third scenario cells were grown at 30°C for 35.75 h and then 100 μL of fermented broth was inoculated into fresh medium coinciding with a temperature drop from 30 to 10°C .

The data consisted of 2 replicates (jars) for each scenario. Each pair of replicates was carried out side-by-side. Because only a single water bath was available to main-

tain the desired temperatures, the experiments for different scenarios were carried out on different days. For the first scenario, 17 sample points over time (a total of 34 observations) were taken. For the second scenario, 19 samples were taken during the first period of growth and 21 samples taken during the second period of growth (a total of 80 observations). In the third validation scenario, 19 samples were taken during the first period of growth and 25 samples were taken during the second period (a total of 88 observations).

3.2.6 Sensitivity analyses

Sensitivity of a particular parameter was calculated as the relative change in the model prediction for a 10% perturbation of that parameter with all other parameters fixed at their estimated values (Haefner, 1996). The interest was in determining the sensitivity of cell density ($\log_{10}(N)$) to perturbations in the i^{th} parameter (P_i) was calculated as the centered finite approximation to $\frac{\partial \log_{10} N}{\partial P_i}$ given in Eqn.3.15.

$$\frac{\partial \log_{10} N}{\partial P_i} \approx \frac{\log_{10} N \mid_{P_i + 0.05P_i} - \log_{10} N \mid_{P_i - 0.05P_i}}{0.1 P_i} \tag{3.15}$$

Positive values of this measure indicated that when a parameter is perturbed upward the model prediction for N is higher than when the unperturbed parameter is used. Conversely, negative sensitivities indicated that a positive perturbation resulted in a reduction in the predicted value for N. The advantage of this approach was that it could be performed at each point in time over a simulated fermentation. When the resulting sensitivity curve was plotted over time, it was easy to determine the relative

importance of the model parameters at each of the different phases of bacterial growth.

Also interesting were the interactions between parameters that affected model predictions. To examine interactions in sensitivities, a "multiple parameter sensitivity analysis" (Swartzman and Kaluzny, 1987) was used. Obtained from the analysis were the main effects, interaction terms, and higher order terms (analogous to a multi-way ANOVA). A multiple parameter sensitivity analysis was performed on the observed growth rate. The model 3.11 tracks both cell growth as well as cell death. Therefore, to determine the observed growth rate a smoothing spline (CSAPS) was fit to the cell density data from which the first derivative was calculated (FNDER). The largest positive value of the first derivative during the exponential growth phase was taken as the observed growth rate. For the ANOVA, a central composite design (Keuhl, 2000) was used which allowed estimation of all main effects and first-order interactions. The analysis of the observed growth rate was limited to the eight parameters determined to be important in the temporal analyses. The resulting central composite design required $2^8 + 2(8) + 1 = 273$ function evaluations. P-values, it should be noted, were meaningless in this context since there was technically no error in simulated data (Swartzman and Kaluzny, 1987).

3.3 Results

3.3.1 Model development and calibration

The model links the mechanisms of nutrient acquisition, end-product accumulation, temperature adaptation, and cell growth. In particular, the rates of glucose consumption, malic acid reduction, and cell growth and death all depend on the current

intra-cellular "energy" level (q). The convention of assigning the initial value of q to 1 was used because the fermentations were inoculated from overnight cultures of approximately the same age. This meant that q should be interpreted as a measure of the fraction of the initial intra-cellular "energy" typical in an overnight bacterial culture. The state variables of the model are summarized in 3.2, and the model parameters along with their biological interpretations and estimated values are summarized in 3.1.

To calibrate the model, 30°C was used as a reference point, but growth was also observed at 20 and 10°C . It is important to note, however, that all overnight cultures were grown at 30°C prior to inoculation for an experiment. Thus, inoculation at lower temperatures constituted a temperature shock to which the bacteria had to adapt. Dependence on the rate of temperature change, as well as the magnitude of the temperature change, was manifested in the model in the equation for cellular energy, as well as in the rate-specific reductions in the glucose consumption rate, malic acid reduction rate, and lactic acid inhibition rate. As can be observed in Fig. 3.2, the shift to low temperatures produces a lag-phase effect. The lag-phase was less pronounced at 20°C (Fig. 3.3). and cells inoculated into fresh medium at 30°C showed no lag (see Fig. 3.4).

Model validation and analysis.

After estimating the model parameters from fixed temperature data, the model was validated against three temperature scenarios. The first scenario consisted of variable temperature regime. Bacteria were grown for 3.75 h at 30°C , and then the temperature was brought down to 10°C over a period of approximately 15 min. Growth at 10°C was continued for another 60 h. Model predictions closely matched the experi-

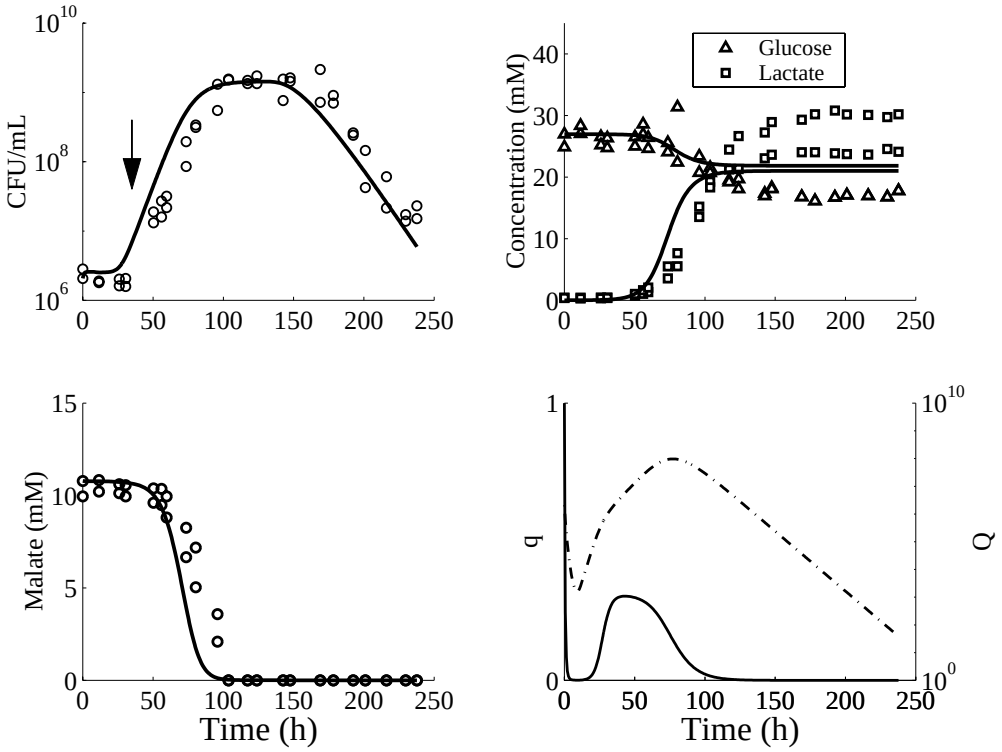

Figure 3.2: Predictions of the calibrated model at 10°C . Note that the culture was incubated at 30°C prior to inoculation. The duration of lag phase is indicated by the arrow. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q.

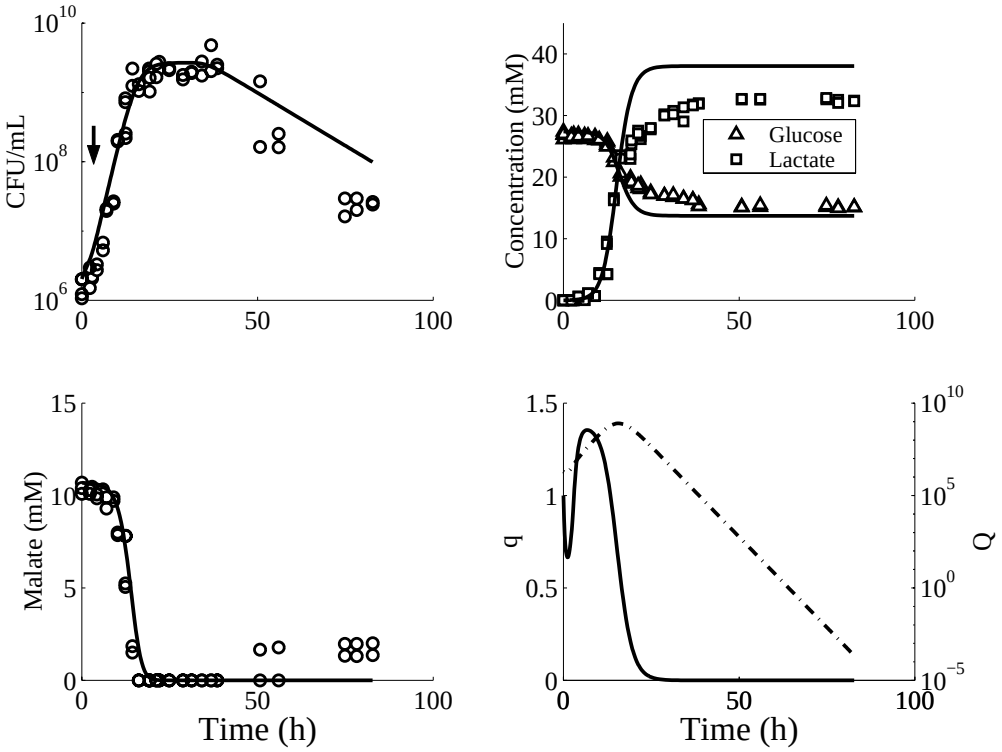

Figure 3.3: Predictions of the calibrated model at 20°C . Note that the culture was incubated at 30°C prior to inoculation. The duration of lag phase is indicated by the arrow. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q.

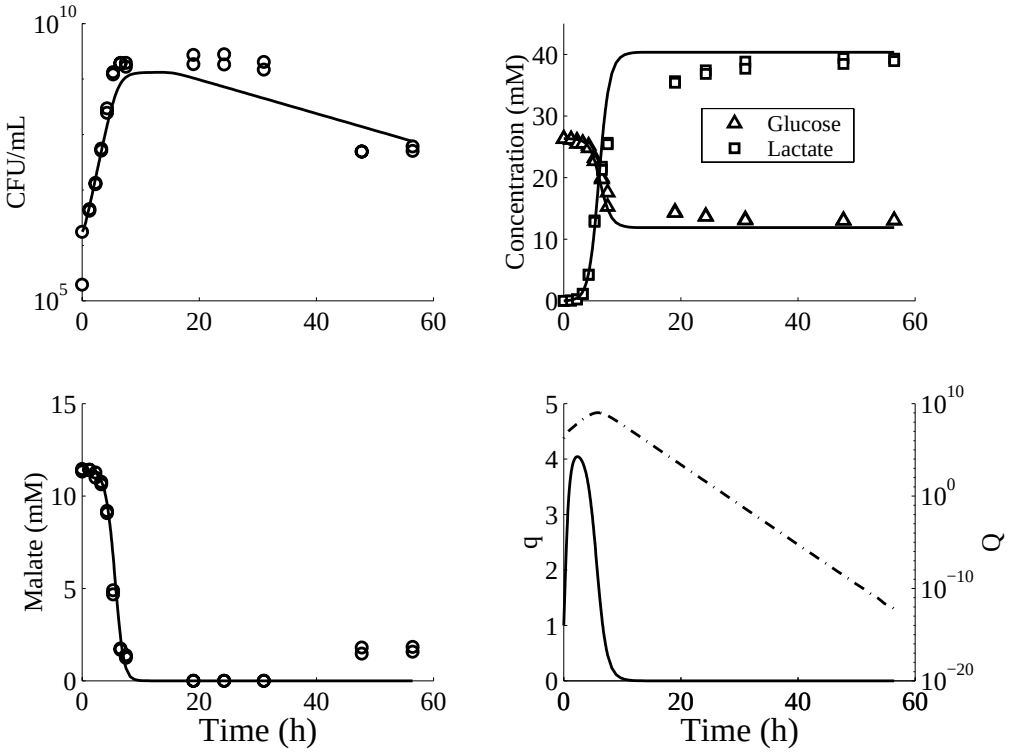

Figure 3.4: Predictions of the calibrated model at 30°C . The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q.

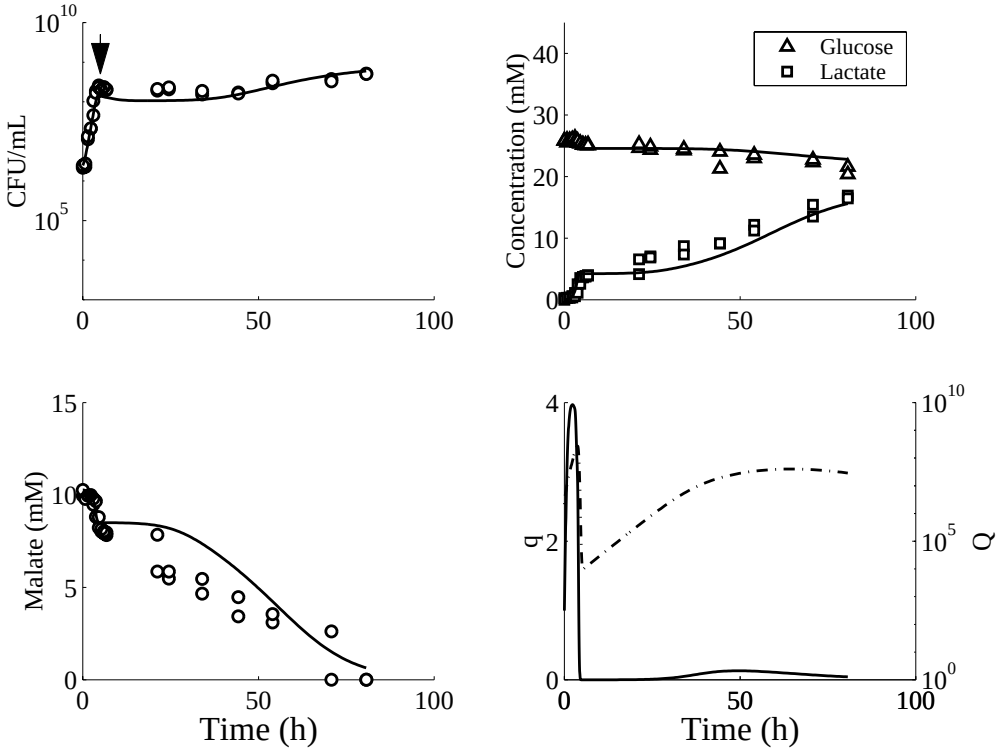

Figure 3.5: Variable temperature validation (scenario #1). After 3.75 h, the temperature was reduced from 30°C to 10°C over a period of about 15 min. Arrow indicates temperature shift. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q.

mental data. It can be seen in Fig. 3.5 that the model accurately predicted the period of rapid growth at 30°C , as well as the pronounced lag after the temperature shift.

A second scenario (Fig. 3.6) involved a batch growth at 30°C , followed by re-inoculation into fresh medium also at 30°C . This scenario was meant to gauge the ability of the bacterial population to rapidly recover from a toxic previous growth environment once placed into fresh medium. At the point of re-inoculation, the model predictions for the cell counts and internal pool of energy were adjusted for

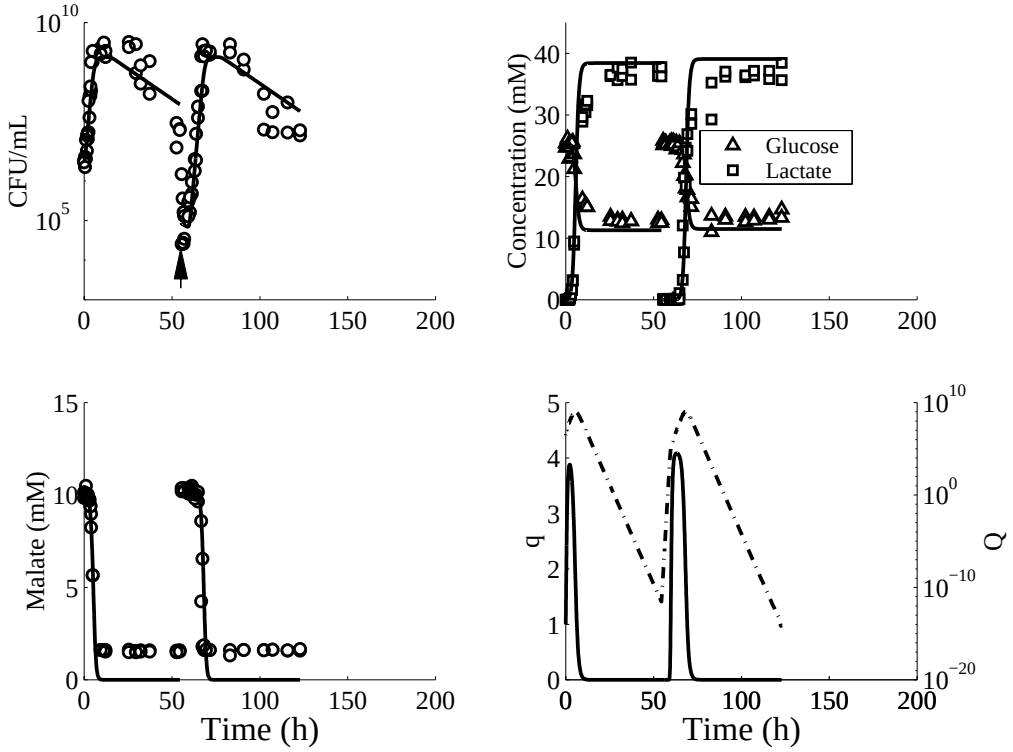

Figure 3.6: Re-inoculation validation experiment (scenario #2). After 55 h, the cells were re-inoculated into fresh medium. The temperature was held at 30°C throughout the entire experiment. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q.

dilution and used as initial conditions for a second simulation run. The energy stores of the bacteria were rapidly replenished once placed in fresh medium. There was no noticeable lag phase, and the second profile nearly duplicated the behavior of the first profile. In the third production scenario, the shift in temperature from 30 to 10°C coincided with a re-inoculation into fresh medium. The results (see Fig. 3.7) suggest an overall robustness of the model to perturbations in temperature and initial conditions.

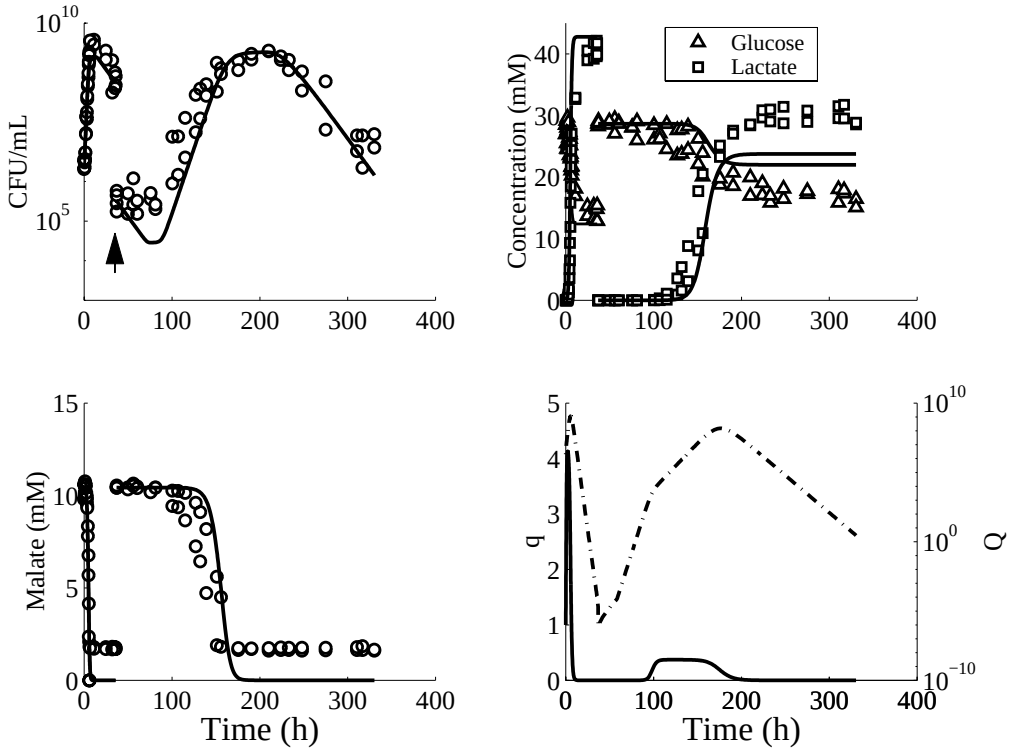

Figure 3.7: Model predictions for scenario #3. Re-inoculation into fresh medium at 3.75 h coincided with a shift in temperature from 30°C to 10°C . The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line $= Q$.

3.3.2 Sensitivity analyses

A sensitivity analysis was carried out to gain insight into which components of the model are most important with respect to growth regulation in response to changes in temperature. For example, in Figs. 3.8–3.10 a "temporal sensitivity analysis" was used to measure sensitivity with respect to predicted cell density for the fixed temperature fermentations. In general, the magnitudes of the sensitivities increase as the temperature decreases. At 10°C , cell counts during exponential and stationary phase are positively affected by positive perturbations in β, μ_1, K_{T1} and kq_1, but are negatively affected by positive perturbations in α, τ, γ, and kq_2. δ_1 is important only in death phase, where it had an increasingly negative effect. There was a reversal in the signs of all parameter sensitivities near the transition from stationary to death phase (Fig. 3.8). This reversal occurs to a lesser extent at 20°C (Fig. 3.9) and is absent in the analysis at 30°C (Fig. 3.10). Particularly striking are the curves for parameters α and kq_1, which represent growth rate and energy (respectively), and experienced changes in sign from the analysis at 10°C to the analysis at 30°C . Positive perturbations in κ (energy cost of acid stress) have a relatively larger negative effect at 30°C (Fig. 3.10) than in the lower temperatures (Fig. 3.8), and the positive perturbations of kq_1 were larger in magnitude, compared to other parameters, during the exponential and stationary phases at 30°C (Fig. 3.10).

While this kind of analysis is useful in exploring the effect of a single parameter, understanding the interplay between parameters was also important. It was found that the exponential and stationary phases were most sensitive to perturbations in the parameters. Therefore, a multiple sensitivity analysis was conducted of the maximal observed growth rate to gain a better understanding of how the model behaves

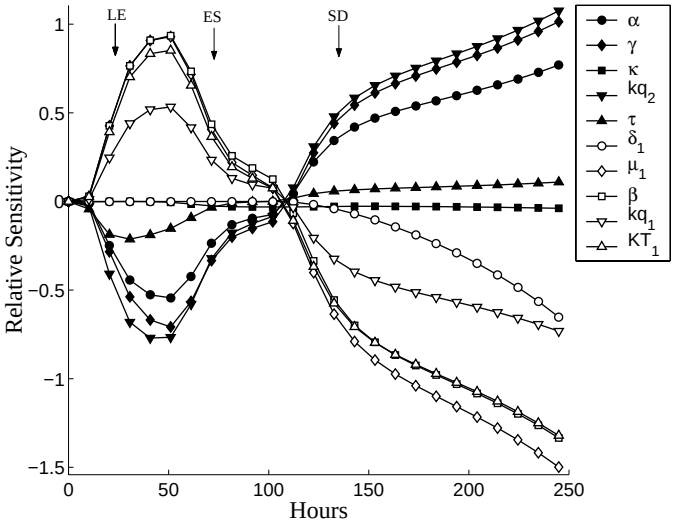

Figure 3.8: Temporal sensitivity profiles of cell density [log10(CFU mL^{-1})] with respect to the model parameters at 10°C (only the 10 most sensitive model parameters shown). LE, ES and SD are reference points indicating approximate times of transition between lag, exponential, stationary and death phase, respectively.

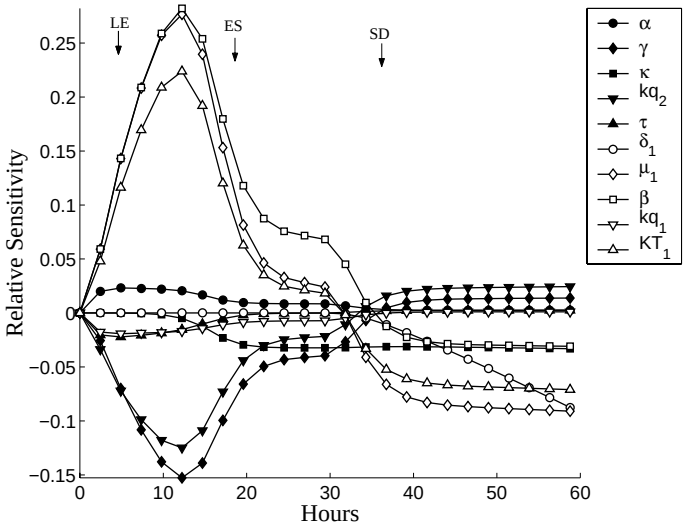

Figure 3.9: Temporal sensitivity profiles of cell density [log10(CFU mL^{-1})] with respect to the model parameters at 20°C (only the 10 most sensitive model parameters shown). LE, ES and SD are reference points indicating approximate times of transition between lag, exponential, stationary and death phase, respectively.

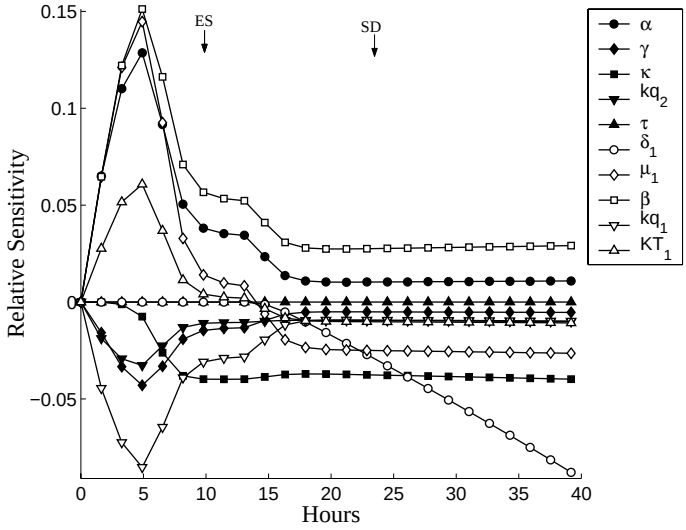

Figure 3.10: Temporal sensitivity profiles of cell density [log10 (CFU mL^{-1})] with respect to each of the model parameters at 30°C (only the 10 most sensitive model parameters shown).

during these phases of growth. The main effects, interaction terms, and second-order effects for simulations at 10 and 30°C are reported in Tables 3.4 and 3.3, respectively. Comparison of the data in Tables 3.4 and 3.3 reveals that μ_1 and β are the most important parameters in determining the growth rate. The fact that the signs of α and kq_1 in Table 3.3 are the opposite of what they were in Table 3.4 suggests that growth is limited in a fundamentally different manner at 10°C than at 30°C . This is also supported by the fact that τ does not interact with any of the (other) parameters at 30°C while at 10°C the interactions are relatively strong. The observed growth rate, irrespective of temperature, was most strongly influenced by the glucose consumption rate (μ_1), glucose-to-energy conversion rate (β), and maximal growth rate (α). The growth rate is least affected by the energetic cost of reproduction (γ) and the energy half-saturation constant for glucose consumption (k_{q2}). The energetic cost of

temperature adaptation (τ) is only important in the 10°C analysis (Table 3.4), where it tends to manifest a negative effect on growth rate through its negative interactions with μ_1, β, K_{T1}, and kq_1.

Table 3.3: Multiple sensitivity (index) analysis for growth rate at 30°C . The parameters considered here were also found to be important in determining cell density during exponential and stationary phase (see Figure 3.10). Listed are the coefficients describing the sensitivity surface about the estimated parameters. Constant terms indicate the mean trend or "main effect" due to that parameter alone, whereas the interaction terms are a measure of how a perturbation in one of the parameters affects model sensitivity to the other parameter involved in the interaction. The terms on the main diagonal indicate concavity in the sensitivity surface with respect to that particular parameter. The larger values are underlined.

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	0.3300	-0.1491	0.0626	0.0299	-0.1202	0.2789	0	0.0379	0.0257
β	0.3310		-0.1445	0.0310	-0.1192	0.2786	0	0.0377	0.0259
KT_1	0.1405			-0.1019	-0.0594	0.1260	0	0.0165	0.0106
kq_1	-0.1666				0.0758	-0.0386	0	0.0790	0.0506
α	0.2721					-0.1565	0	-0.1458	-0.0938
τ	-0.0000		*Symmetric*				0.0100	0	0
γ	-0.0931							0.0089	-0.0189
kq_2	-0.0738								-0.0073

Table 3.4: Multiple sensitivity (index) analysis for growth rate at 10°C . Comparison with Table 3.3 reveals that μ_1 and β are the most important parameters in determining the growth rate. The fact that the signs of α and kq_1 are the opposite of what they were in Table 3.3 suggests that growth is limited in a fundamentally different manner at 10°C than at 30°C . The larger values are underlined.

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	0.1754	-0.0435	0.1134	0.1044	-0.0863	0.0804	-0.0120	-0.0254	-0.0037
β	0.1756		-0.0447	0.1049	-0.0861	0.0802	-0.0123	-0.0256	-0.0038
KT_1	0.1607			-0.0549	-0.0792	0.0738	-0.0106	-0.0227	-0.0044
kq_1	0.0972				-0.1364	0.1964	-0.0071	0.1399	0.1408
α	-0.1019					-0.0383	0.0070	-0.1375	-0.1410
τ	0.0005		*Symmetric*				0.0266	0.0065	0.0127
γ	-0.1454							0.0434	-0.0490
kq_2	-0.1276								0.0073

Table 3.5: Multiple sensitivity (index) analysis for observed death rate at 10°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	-0.0069	0.0394	0.158	0.150	0.120	-0.124	-0.007	-0.142	-0.141
β	-0.0066		0.0383	0.145	0.116	-0.119	-0.0069	-0.137	-0.136
KT_1	-0.0063			0.0375	0.110	-0.113	-0.0065	-0.130	-0.128
kq_1	-0.0051				0.0345	-0.092	-0.0053	-0.104	-0.103
α	0.0053					0.0346	0.0053	0.107	0.106
τ	0.0003		*Symmetric*				0.0333	0.006	0.0061
γ	0.006							0.0357	0.122
kq_2	0.0059								0.0357

Table 3.6: Multiple sensitivity (index) analysis for observed death rate at 30°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	-7.88x10^{-7}	-3.40x10^{-6}	9.44x10^{-6}	-2.86x10^{-5}	-1.47x10^{-5}	-1.38x10^{-5}	-3.25x10^{-17}	3.30x10^{-6}	2.61x10^{-5}
β	-3.02x10^{-8}		-5.76x10^{-6}	7.91x10^{-6}	5.73x10^{-6}	-1.84x10^{-6}	-5.10x10^{-16}	-2.31x10^{-5}	-5.77x10^{-6}
KT_1	-7.85x10^{-7}			-3.44x10^{-6}	-1.27x10^{-5}	-1.40x10^{-5}	9.78x10^{-17}	2.94x10^{-6}	2.28x10^{-5}
kq_1	-7.34x10^{-7}				-5.56x10^{-6}	-2.02x10^{-5}	-1.84x10^{-16}	7.56x10^{-7}	1.18x10^{-5}
α	-8.65x10^{-7}					-8.46x10^{-6}	1.63x10^{-16}	2.12x10^{-6}	1.91x10^{-5}
τ	4.94x10^{-18}		*Symmetric*				-3.34x10^{-6}	-6.29x10^{-16}	-2.06x10^{-16}
γ	-1.92x10^{-7}							-9.38x10^{-6}	-7.44x10^{-6}
kq_2	5.31x10^{-7}								-3.62x10^{-6}

Table 3.7: Multiple sensitivity (index) analysis for max cell density at 10°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	1.13	-2.46	-4.04	-3.72	-3.63	3.66	-2.93x10^{-3}	3.70	3.98
β	1.54		-2.62	-3.63	-3.58	3.61	2.50x10^{-3}	3.64	3.89
KT_1	1.01			-2.10	-3.30	3.31	-0.0203	3.35	3.58
kq_1	0.801				-1.97	3.45	-3.53x10^{-3}	3.48	3.51
α	-0.810					-1.23	-4.72x10^{-4}	-3.50	-3.55
τ	-1.90e-05		*Symmetric*				-0.141	-5.53x10^{-3}	-2.35x10^{-3}
γ	-1.30							-1.09	-3.55
kq_2	-1.04								-1.33

Table 3.8: Multiple sensitivity (index) analysis for max cell density at 30°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	0.0795	-0.172	-0.209	-0.0762	-0.0350	0.0943	$1.39\text{x}10^{-13}$	0.0778	0.125
β	0.481		-0.362	-0.0821	-0.0563	0.1048	$1.39\text{x}10^{-13}$	0.0891	0.136
KT_1	0.0194			-0.0294	-0.0221	0.0434	$2.78\text{x}10^{-14}$	0.0419	0.0498
kq_1	-0.260				0.0840	0.0572	$2.00\text{x}10^{-13}$	0.0933	0.0315
α	0.317					-0.220	$-9.99\text{x}10^{-14}$	-0.1242	-0.0317
τ	$3.90\text{x}10^{-15}$		*Symmetric*				0.0182	$3.28\text{x}10^{-13}$	$4.44\text{x}10^{-14}$
γ	-0.123							$3.00\text{x}10^{-3}$	-0.0618
kq_2	-0.0972								0.0131

Table 3.9: Multiple sensitivity (index) analysis for terminal lactic acid concentration at 10°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	0.724	-1.1098	-1.51	-1.39	-1.41	1.41	$4.68\text{x}10^{-6}$	1.38	1.54
β	0.714		-1.08	-1.36	-1.39	1.39	$-6.19\text{x}10^{-5}$	1.37	1.50
KT_1	0.655			-0.963	-1.28	1.28	$-5.89\text{x}10^{-4}$	1.25	1.38
kq_1	0.557				-0.945	1.39	$2.40\text{x}10^{-4}$	1.37	1.40
α	-0.563					-0.403	$1.53\text{x}10^{-4}$	-1.37	-1.41
τ	$-1.53\text{x}10^{-5}$		*Symmetric*				-0.0214	$-4.39\text{x}10^{-4}$	$-1.43\text{x}10^{-5}$
γ	-0.60278							-0.365	-1.39
kq_2	-0.66791								-0.418

Table 3.10: Multiple sensitivity (index) analysis for terminal lactic acid concentration at 30°C .

Parameter	*Constant Term (Main Effect)*	*Interaction Terms*							
		μ_1	β	KT_1	kq_1	α	τ	γ	kq_2
μ_1	<u>0.191</u>	-0.171	<u>-0.144</u>	-0.0610	-0.0349	0.0388	$7.21\text{x}10^{-16}$	0.0455	0.132
β	<u>0.184</u>		<u>-0.162</u>	-0.0582	-0.0340	0.0375	$3.12\text{x}10^{-14}$	0.0459	<u>0.122</u>
KT_1	0.0787			-0.0727	-0.0147	0.0163	$1.16\text{x}10^{-18}$	0.0202	0.0515
kq_1	0.0137				-0.0246	0.0331	$2.22\text{x}10^{-14}$	0.0442	0.0228
α	-0.0102					-0.0128	$-1.32\text{x}10^{-14}$	-0.0484	-0.0222
τ	$-5.60\text{x}10^{-16}$		*Symmetric*				$4.45\text{x}10^{-4}$	$4.02\text{x}10^{-14}$	$4.85\text{x}10^{-15}$
γ	-0.0430							-0.00124	-0.0441
kq_2	<u>-0.155</u>								0.0276

3.4 Discussion

There is generally an inverse relationship between model complexity and model robustness (Baranyi and Roberts, 1995; Haefner, 1996). Kooijman (Kooijman, 2000), however, argues that "a fair comparison of models should be based on the number of parameters per variable described, not on absolute number." A model was developed which predicts not only cell density (N) but also reliably predicts glucose (S), lactic acid (L), malic acid (M), and gives substantive and experimentally verifiable predictions of intracellular energy using only 14 model parameters. In this regard, the model's complexity is on par with that of currently available models. The parameter γ is small and its elimination from the model was considered. Numerical experiments, however, showed that it was not possible to obtain sharp resolution in lag phase or stationary phase without this parameter.

Included in Figs. 3.2–3.7 are the predicted internal energy profiles (Q) and the per-cell internal energy ($q = Q/N$) profiles. The model clearly predicted reductions in "energy" available for growth immediately after a temperature shift or when end-product inhibition ensues in stationary phase. Currently, no experiments have been conducted with *L. lactis* (LA221) to support this, and this is the subject of future work. However, Mercade et al. (2000) has shown that the yield (Y_{ATP}) of ATP of *L. lactis* decreased from 11.5 g mol^{-1} at a pH of 6.6 to 5 g mol^{-1} at a pH of 4.4, thus demonstrating an energy drain due to acidic conditions. Jetton et al. (1991) has shown that starved cells of *Methanothrix soehngenii* contained relatively high levels of AMP (2.2 nmol mg^{-1} protein), but essentially no ADP or ATP during acetate degradation. Addition of new substrate, however, quickly brought the ATP levels back up to concentrations of about 1.4 nmol mg^{-1} protein. The Gram-negative

bacterium *Pectinatus frisingensis* has been shown by Chibib and Tholozan (1999) to experience decreases in ATP, ADP and AMP concentrations during cold shocks (30 to 20°C). The bacteria returned to a pre-shock metabolic state when returned to 30°C in the presence of glucose. Metge et al. (1993) showed that the total adenine nucleotide content for a species of *Pseudomonas* decreased from $153\text{x}10^{-20}$ moles cell^{-1} during exponential phase to $56\text{x}10^{-20}$ moles cell^{-1} during stationary phase.

Interpretation of the temporal sensitivity analyses in Figs. 3.8–3.10 is straightforward. By conducting the temporal sensitivity analyses at different temperatures, it was possible to see how temperature affects the relative importance of each parameter in relation to model predictions. The sensitivity of α (the maximum growth rate of the cells) became negative at low temperatures, κ (the parameter which controls the energetic cost of cell division) became relatively unimportant at low temperatures, and τ became important only at low temperatures. These results suggest that growth at colder temperatures was limited primarily by the requirements for temperature adaptation, while growth at 30°C was limited primarily by acid stress. These results also suggest that at low temperatures it was more advantageous to divert energy to temperature adaptation. This idea is also supported by the large negative sensitivity of parameter γ (at 10°C) seen in Fig. 3.8, since it is this parameter which controls how much energy is spent on reproduction.

The sign reversal of nearly every parameter sensitivity at the end of stationary phase in Fig. 3.8 is striking but entirely reasonable. This characteristic of the sensitivity analysis comes from the fact that the factors that promote strong and rapid growth also promote rapid end product accumulation and precipitate cell death. At low temperatures, this phenomenon was enhanced by the increased energy demand required for temperature adaptation. In general, 8 of the 14 model parameters were

important in determining growth during exponential and stationary phase. These were β, μ_1, K_{T1}, and kq_1, which have positive sensitivities, and α, τ, γ and kq_2, which have negative sensitivities. Not surprisingly, these are the parameters that control growth (α,kq_1, and γ), sugar utilization (β, μ_1, and kq_2), and temperature adaptivity (KT_1 and τ) in the model.

The multiple sensitivity analysis of these parameters revealed that μ_1 and β were the most important parameters in determining the observed growth rate. The qualitative shift in parameter sensitivity suggests that the observed growth rate was limited in a fundamentally different manner at 10°C than at 30°C . In particular, these results support the previous suggestion that growth was limited at low temperatures by the demand for temperature adaptation, while at warmer temperatures end-product accumulation was the primary limiting force. For example, the parameter controlling energy cost for temperature adaptation, τ, had virtually no bearing on the model predictions at 30°C .

Previous researchers (Brockelhurst et al., 1995; Bovill et al., 2000; Fu et al., 1991; Van Impe et al., 1992, 1995), have developed models to predict growth during continuous changes in temperature. These models use an empirical function, such as Gompertz or Ratkowsky relationships to describe temperature induced lag phase. In the model, temperature is variable that controls the predicted metabolic activity of the cell. Using this mechanistic approach, it is possible to predict how changes in cell physiology produce a temperature induced lag phase. While some systematic lack of fit was observed, Figs. 3.2–3.7 demonstrate the qualitative agreement of predicted and experimental results. Understanding how physiological changes affect growth with varying temperatures may lead to a rational method for selecting biocontrol or starter cultures.

In the formulation of the WSS it was assumed that the errors of one variable were independent of the errors of another variable. This assumption allowed us to combine the sum-of-squares of all the data in an additive manner. That the variables themselves are not independent does not necessitate the dependence of the residuals. However, it is likely that the underlying model is not absolutely correct – simplifying assumptions were employed in the formulation of the model. The assumed independence of model errors is therefore likely false. For example, under-estimation of glucose consumption tended to coincide with under-estimation of lactic acid production. Additionally, the single-injection method of (McFeeters, 1993) was used to quantify the lactic acid, malic acid and glucose. In this method, a single sample is processed to quantify all 3 components. Although different detectors were used to quantify the acids and the sugars, measurement errors in such a case are more likely to be correlated.

A more realistic framework for the dependence of errors in a coupled system of equations is suggested by Phillips (2000). In this approach, a covariance matrix of the errors is estimated. This covariance matrix can contain covariances among the errors of diffent variables as well as covariances of observations taken over time on the same experimental unit. Unfortunately, Phillips development is for linear systems of simultaneous equations. The equations of the fermentation model are non-linear differential equations and not amenable to the matrix formulation given by Phillips. One must also weigh the cost of the many additional parameters which must be estimated from the data in order to fit complicated error structures. Alternatively, one may take the approach of (Wood, 2001) and (Jost and Ellner, 2000) in which the desired distributional assumptions of normality and independence are met more closely by adding a smoothing spline component to the model to account for systematic lack

of fit. Such approaches are termed "semi-mechanisic" or "partially specified" and currently being explored with regard to the fermentation model.

In this chapter, it has been shown that the energetic costs of temperature adaptation can explain lag phase. The model predicted lag phase, death phase and maximal growth rates. A quadratic temperature inhibition function (I_T) was used in the model, however, the functional form may be improved to allow for temperature effects above and below the optimal temperature for growth (T_{opt}). Model components did not vary independently of one another, and all affected and depended on the internal pool of cellular energy (q). It was this dynamic energy budget aspect of the model that allowed growth predictions across a range of continuously varying environmental conditions for lag, exponential, and stationary phases of batch culture. The model was validated using broth fermentations. This work may serve as the basis for modeling more complex fermentation systems. Future research will be aimed at experimentally determining intra-cellular ATP concentrations during batch growth of LA221.

Chapter 4

Semi-Mechanistic Models for Buffer Systems

4.1 Introduction

In the preceding chapter, a mathematical model was developed for batch fermentation in a variable temperature environment. The main features of this model are that it can predict the dynamics of the cell population as well as changes in the chemical composition of the medium during temperature shocks. This model is not without mechanistic errors however. The model is unable to predict pH . Thus, the difference in toxicity between the neutral and dissociated forms of an organic acid in the fermentation medium is not addressed. The acid toxicity was taken to be proportional to the total lactic acid concentration. But this is a convenient simplifying assumption which will be shown is not correct. In this chapter, a more refined approach is provided for modeling the underlying chemical equilibria that determine among other things the pH , the buffer capacity and ion concentrations in fermentation media. Using

this method, a model is produced which more accurately reflects the biology of acid toxicity.

To illustrate some of the difficulties involved in modeling the chemical dynamics, a somewhat naive experiment will be presented which investigates the relationship between pH , organic acid ion concentrations and the growth rate of the lactic acid bacterium *Lactococcus lactis*. After discussing the results of this experiment and its shortcomings, a flexible modeling framework is presented for dealing with the chemical dynamics and suggest how to incorporate such processes into the model presented in the preceding chapter.

4.2 Effects of pH and Protonated Acid Concentration on Growth Rate

Many fermentations involve the reduction of a sugar to a weak acid. The dissociation equilibrium of an weak acid (HA) having a single dissociable proton (H^+) is described by

$$\mathrm{K}_a \quad = \quad \frac{\left[\mathrm{H}^+\right] \gamma_+ \left[A^-\right] \gamma_-}{\left[HA\right] \gamma_o} \tag{4.1}$$

where K_a is the dissociation constant of the acid and the γ_i are activity coefficients which depend on factors such as the temperature and salt content of the solution (Pitzer, 1991). The dimensionless quantity pH is similarly defined in terms of activity coefficients

$$\mathrm{pH} \quad = \quad -\log_{10}(\left[\mathrm{H}^+\right] \gamma_+) \tag{4.2}$$

Thus, the dissociation of an acid depends in a direct way on the pH . For the time being, it is assumed that the γ_* are unity but a more general case will be dealt with shortly.

As indicated above, the dissociation of a neutral molecule results in the production of charged molecules called ions. The presence of ions can influence the chemical and biological processes occurring in the solution. Actually, this activity affects the dissociation equilibria of all species in solution. It turns out that the values of the γ_i in Eqn.4.1 will depend in large part on the concentration of ions (Pitzer, 1991).

The standard measure of the chemical activity of ions in a solution is ionic strength which is defined by

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2 \tag{4.3}$$

In this expression, n is the number of ionic species in solution, c_i is the concentration of of the i^{th} ionic species and z_i^2 is its charge squared. Because solutions are electrically neutral the condition that $\sum_{i=1}^{n} c_i z = 0$ (charge balance) should also be met. In practical terms, this means that one must remember to include all ions present in the solution in the calculation of ionic strength not just the ones being measured. It should be clear from this brief introduction that pH , acid ion concentration and ionic strength are inter-related. It is not clear, however, how these factors are related with respect to the growth rate of a bacterium in the solution.

To investigate the effects of pH and protonated acid on the growth of bacteria *Lactococcus lactis* was grown in a rich medium at constant at various pH and protonated lactic acid concentrations ($[HL]$). The initial ionic strength was set at 0.342 to ensure that changes in ionic strength due to fermentation are negligible and I can be

assumed constant. To determine the relationship between pH and protonated acid a factorial experimental design was used. The primary difficulty in setting up this kind of experiment is the inter-dependencies among pH , lactic acid dissociation and ionic strength mentioned earlier. Thus, to achieve desired levels of $[HL]$ the total amount of lactic acid ($[L^-] + [HL]$) added to the medium was adjusted and the medium was supplemented with $NaCl$ to maintain I at 0.342. The concentrations of lactic acid and NaCl required were calculated for each different pH and desired $[HL]$.

4.2.1 Materials and Methods

Cultures and Media

Lactococcus lactis subsp. *lactis* strain LA221 transformed with the chloramphenicol resistance gene pGK12 (Breidt and Fleming, 1998) was obtained from the USDA culture collection. LA221 was grown on M17 broth (Difco Laboratories, Detroit, Mich.) containing 1.5% agar (Difco), 1% glucose (Sigma Chemical Co., St. Louis, Mo.) and 5 ug mL^{-1} chloramphenicol. Growth experiments were conducted in M17 broth supplemented with 1% glucose (GM17) and in M17 broth supplemented with 1% glucose and 50mM succinic acid (GSM17). NaCl was added when necessary to GSM17 to maintain constant ionic strength at 0.342 M across all treatments.

Succinic Acid Ionic Strength Contribution

Succinic acid has 2 dissociable protons. The 2 pK_a of succinic acid are ($pK_{a1} = 4.22$ and $pK_{a2} = 5.64$). The following equilibria and mass balance equations were used to determine the equilibrium concentrations of dissociated and undissociated forms of

succinic acid.

$$\begin{aligned}
\left[C_4H_5O_4^-\right] &= K_{a1}\frac{[C_4H_6O_4]}{H^+} && (4.4)\\
\left[C_4H_4O_4^{2-}\right] &= K_{a2}\frac{\left[C_4H_5O_4^-\right]}{H^+} && (4.5)\\
50mM &= [C_4H_6O_4] + \left[C_4H_5O_4^-\right] + \left[C_4H_4O_4^{2-}\right] && (4.6)
\end{aligned}$$

Equation 4.4 and 4.5 were substituted into 4.6 to give the following solutions to the above equations

$$\begin{aligned}
[C_4H_6O_4] &= \frac{50\text{mM}\left[\text{H}^+\right]^2}{\left[\text{H}^+\right]^2 + K_{a1}\left[\text{H}^+\right] + K_{a1}K_{a2}} && (4.7)\\
\left[C_4H_5O_4^-\right] &= \frac{0.05\text{ M}K_{a1}\left[\text{H}^+\right]}{\left[\text{H}^+\right]^2 + K_{a1}\left[\text{H}^+\right] + K_{a1}K_{a2}} && (4.8)\\
\left[C_4H_4O_4^{2-}\right] &= \frac{0.05\text{ M}K_{a1}K_{a2}}{\left[\text{H}^+\right]^2 + K_{a1}\left[\text{H}^+\right] + K_{a1}K_{a2}} && (4.9)\\
& && (4.10)
\end{aligned}$$

Lactic Acid Ionic Strength Contribution

Lactic acid has a single dissociable proton with a pK_a of 3.86. The concentration of the lactate anion was calculated by

$$\left[L^-\right] = \frac{CK_{al}}{\left[\text{H}^+\right] + K_{al}} \quad (4.11)$$

where C is the total lactic acid concentration ($[L^-] + [HL]$).

Required NaCl determination

After calculating the succinate and lactate ion concentrations using Eqns 4.7– 4.10 and Eqn. 4.11 for a particular set of conditions, the amount of NaCl required to maintain the ionic strength at 0.342M was determined.

$$\begin{aligned} \text{NaCl M} &= 0.342 - \text{Ionic strength from succinic} \\ &\quad - \text{Ionic strength from lactic acid} \end{aligned}$$

Experimental Design

A factorial design was utilized with pH levels at 4.0, 4.3, 4.6, 4.9, 5.2, and 5.5 and protonated lactic acid levels at 1.0, 2.0 ,3.0, 4.0 and 5.0 mM. For each combination of levels, a sterile 15 mL screw-top tube was filled with 5 mL of 2X GSM17 broth and the calculated volume of lactic acid and NaCl stock. The pH of each tube was then adjusted to the desired level using 0.3N HCl and 0.3N NaOH. The final volume in each tube was brought to 10mL using sterile de-ionized water. Corrective adjustments of the pH were made using 0.3N HCl and 0.3N NaOH as needed. Table 4.2.1 provides the concentrations of NaCl, succinic acid and lactic acid required to generate the treatment levels.

Growth Experiments

L. lactis (LA221) was cultured overnight in GM17 broth to a cell density of $\approx 1 \times 10^9$. To each 15 mL screw-top tube twice washed cells from the overnight stock was resuspended at a concentration of $\approx 1 \times 10^6$CFU mL^{-1}. From each tube, 3 200 μL aliquots were taken to fill 3 wells (assigned at random) of a 96 well microtitre

Table 4.1: Calculated levels for the factorial experimental design.

Desired pH	*Desired* $[HL]$ *(mM)*	*Ionic Contrib. Succ.*	*Ionic Contrib. Lact.*	*Salt to add (M)*	*Lactic acid to add (mM)*
4	1	0.020	0.001	0.321	2.479
4.3	1	0.030	0.003	0.309	3.951
4.6	1	0.042	0.006	0.294	6.888
4.9	1	0.056	0.012	0.275	12.749
5.2	1	0.071	0.023	0.247	24.442
5.5	1	0.089	0.047	0.206	47.774
4	2	0.020	0.003	0.319	4.958
4.3	2	0.030	0.006	0.306	7.902
4.6	2	0.042	0.012	0.288	13.777
4.9	2	0.056	0.023	0.263	25.498
5.2	2	0.071	0.047	0.224	48.885
5.5	2	0.089	0.094	0.159	95.547
4	3	0.020	0.004	0.318	7.437
4.3	3	0.030	0.009	0.303	11.854
4.6	3	0.042	0.018	0.282	20.665
4.9	3	0.056	0.035	0.251	38.247
5.2	3	0.071	0.070	0.201	73.327
5.5	3	0.089	0.140	0.112	143.321
4	4	0.020	0.006	0.316	9.916
4.3	4	0.030	0.012	0.300	15.805
4.6	4	0.042	0.024	0.276	27.554
4.9	4	0.056	0.047	0.239	50.996
5.2	4	0.071	0.094	0.177	97.769
5.5	4	0.089	0.187	0.066	191.094
4	5	0.020	0.007	0.315	12.396
4.3	5	0.030	0.015	0.297	19.756
4.6	5	0.042	0.029	0.270	34.442
4.9	5	0.056	0.059	0.228	63.745
5.2	5	0.071	0.117	0.154	122.211
5.5	5	0.089	0.234	0.019	238.868

plate. Each well was overlaid with 80 μL of mineral oil. Optical density of the wells was measured every 30 minutes for 4 days. The temperature of the micro-well plate was maintained at $30°C$.

Statistical Analysis

Treatments were randomly allocated to wells of a 96 well plate according to the factorial design discussed above. The ln of the transmittance data for each well of the 96 well plate was plotted against time and smoothed by local polynomial regression (Fan and Gijbels, 1996). The specific growth rate was then measured as the maximum positive observed slope along the growth curve. Stepwise polynomial regression analysis on the rates was performed using pH and protonated lactic acid concentration as candidate factors.

4.2.2 Results

The stepwise regression analysis suggested that polynomials higher in order than 2 did not explain significantly more variability than the following quadratic model

$$\text{Rate}_i = \beta_0 + \beta_1 \text{pH} + \beta_2 [\text{HL}] + \beta_3 \text{pH} [\text{HL}] + \beta_4 \text{pH}^2 + \beta_5 [HL]^2 + \epsilon_i \quad (4.12)$$

$$\epsilon \sim N(0, \sigma^2) \quad (4.13)$$

The results of the regression analysis using this model are summarized in Table 4.2.2. Parameter estimates from the analysis were used to plot the fitted surface in Figure 4.2.2.

Table 4.2: Quadratic regression analysis of growth rate data. Significant (*) and highly significant (**) factors are indicated.

Parameter	Estimate	T for H0: Parameter = 0	$Pr > \mid T \mid$	Std Error of Estimate
β_0	-3.15	-8.83	0.0001**	0.357
β_1	1.04	7.37	0.0001**	0.141
β_2	0.209	14.83	0.0001**	0.014
β_3	-0.0486	-18.37	0.0001**	0.003
β_4	-0.0760	-5.45	0.0001**	0.014
β_5	1.92×10^{-3}	2.56	0.0132*	7.5×10^{-4}

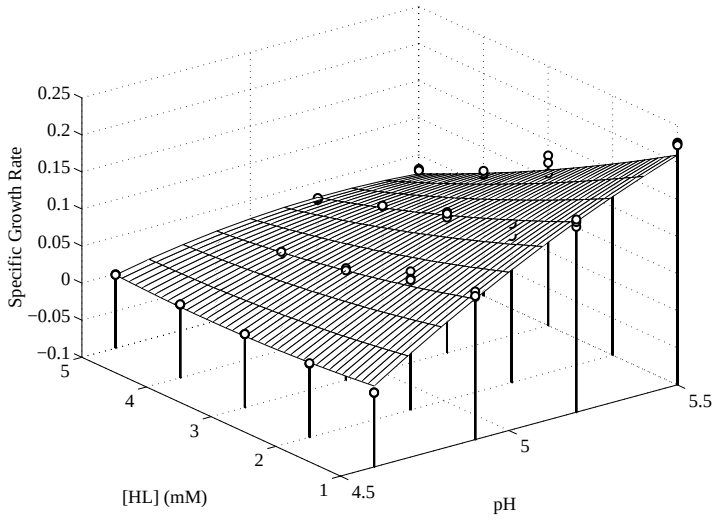

Figure 4.1: Experimental growth rate data along with fitted quadratic polynomial.

4.2.3 Discussion

The quadratic model fit the data well and higher order models did not provide any significant improvement in fit. The strongest effects were pH , protonated lactic acid and the interaction between these two components. The high significance attributed to the pH $[HL]$ interaction strongly suggest that the physiological effects of protonated lactic acid and pH are not independent (P-value = 0.003).

This result suggests that care must be taken in designing experiments meant to measure the effects of acids or pH on growth. For example, the standard practice of calculating minimal-inhibitory values (MIC) of an acid can not be done by simply adjusting the concentration of the acid alone. For example, the MIC for pH is only accurate at a particular level of protonated lactic acid and vice versa. Care should be taken when reporting MIC values in the literature.

There were, however, several limitations of the above analysis. First no account of the ionic strength contribution from the growth medium was made. It was also assumed that the ionic strength was constant throughout the experiment. Actually, the ionic strength will change as fermentation acids are produced by the bacteria. Secondly the pK_a used in determining the required concentrations of lactic acid and the ionic strength contributions from the lactate and succinate anions were based on tabulated pK_a. Such tabulated pK_a are obtained by extrapolating experimental estimates obtained at various ionic strengths to $I = 0$.

The Davies equation can be used to adjust zero-extrapolated pK_a to reflect the effects of ionic strength (Butler and Cogley, 1998) by estimating the γ_i. For the charged ion pairs in the numerator of Eqn 4.1 the geometric mean of the single ion activities is calculated by

$$-\log(\gamma_{\pm}) \quad = \quad A|z_+ z_-| \left(\frac{\sqrt{I}}{1+\sqrt{I}} - bI \right) \tag{4.14}$$

where the z_i are the charges on the molecules and $A = 1.825 \times 10^6 (\epsilon T)^{-3/2}$. ϵ is the dielectric constant of the solvent ($\epsilon = 78.3808$ for water), T is the temperature in Kelvin and b is the "salting-out" parameter. However, if the ion is neutral (as in the

denominator of Eqn 4.1) the activity is predicted by

$$-\log(\gamma_o) \quad = \quad -bI \tag{4.15}$$

Thus, the zero-extrapolated pKa commonly tabulated (pK_a^o) are updated using the formula

$$\mathrm{pK}_a \quad = \quad \mathrm{pK}_a^o - 2\log(\gamma_\pm) + \log(\gamma_o) \tag{4.16}$$

In the case of molecules with multiple pK_a , the $\log(\gamma_o)$ term is replaced by $Az^2\left(\frac{\sqrt{I}}{1+\sqrt{I}} - bI\right)$ when the denominator term corresponds to a charged molecule. Here z corresponds to the charge on the ion in the denominator.

If it is assumed that the ionic strength in the medium was 0.342 M and the temperature is 25°C the Davies equation adjusts the pK_a of lactic acid to 3.58. The percent error in predicting the protonated acid using the zero-extrapolated K_a (K_{a0}) instead of the adjust K_a ($K_{a\mathrm{adj}}$) is given by

$$\%error \quad = \quad 100\frac{\left[\mathrm{H}^+\right] + K_{a\mathrm{adj}}}{K_{a\mathrm{adj}}}\left(\frac{K_{a0}}{\left[\mathrm{H}^+\right] + Ka_0} + \frac{K_{a\mathrm{adj}}}{\left[\mathrm{H}^+\right] + K_{a\mathrm{adj}}}\right) \tag{4.17}$$

Thus at pH 5 the error in $[HL]$ would have been about 1.0% and at pH 4 the error would have been about 20.0%. This error would have propagated into the NaCl calculations as well.

Clearly the levels of protonated lactic acid used in the experiment were subject to rather large relative error. The conclusion about the independence of pH and $[HL]$ on growth rate are likely still valid since the ordering of the levels would not have

been the same. *The parameter estimates for the effects due to these components however can not be trusted.* In the next section, methods for overcoming some of the limitations encountered in this experiment are discussed.

4.3 pH , Ionization, Chirality, and Bacterial Growth

As was demonstrated in the previous section, the pH (Eqn. 4.2) of a growth medium directly affects the metabolic activities of the bacteria growing in it. There are several biological reasons why this would be. The generation of energy (in the form of ATP) requires a positive pH gradient across the bacterial membrane and a negative membrane electrical potential ($\triangle\psi$). Bacteria usually maintain these gradients using electron transport processes following glycolysis. However, if the external pH drops this gradient may dissipate and make it impossible for the cells to generate ATP. Glycolysis is not, however, the only mechanism by which bacteria can generate a pH gradient. For example, bacteria which can decarboxylate malic acid, an acid common in wines and fruits, to lactic acid may exhibit increased growth rates over similar bacteria grown in media without malic acid (Poolman et al., 1991; Versari et al., 1999; Loubière et al., 1992). Malolactic fermentation proceeds in the cytoplasm with the utilization of a proton during decarboxylation of malate to lactate. The resulting increase in cytoplasmic pH is thought to be capable of contributing to a positive $\triangle$pH across the cell membrane and thereby enhancing ATP production. For example, when *Leuconostoc mesenteroides* cells were grown in the presence of malate they showed a modest increase in growth rate over the control (Versari et al., 1999). Because malolactic fermentation proceeds by decarboxylation, it can also contribute to the buffering of the media which can reduce the effect of organic acid accumula-

tion (McFeeters et al., 1982).

Although a negative pH gradient constitutes a direct drain on cellular energy, bacteria can be adversely affected by low pH in other ways. For example, pH has a direct effect on the oxidation state of amino acids which can disrupt the function of enzymes and other proteins. Similarly, pH affects the relative concentrations of charged and uncharged forms of organic acids (recall Eqn. 4.1). This is important since the uncharged forms of organic acids are generally nonpolar and able to diffuse across the cell membrane. *Russell (1992) has shown that modest $\triangle pH$ can drive the concentration of organic acid anions within a bacterium to molar levels.* High concentrations of intracellular organic acid anion such as this can have a deleterious effect on bacteria.

Bacteria have developed various mechanisms for dealing with acid stress. For example, Diez-Gonzalez and Russel (1997) have shown that cells of *Escherichia coli* O157:H7 allow their internal pH to decrease during a fermentation. This prevents the establishment of a large $\triangle$pH across the cell membrane thereby reducing the net influx of weak acid anions. For example *Lactobacillus crispatus* maintains its pH approximately 1 pH unit above ambient (Benthin and Villadsen, 1995). In contrast, cells of *Lactococcus spp.* actively maintain their pH in the range of 5.2-5.4 (Mercade et al., 2000). McDonald et al. (1990) found that *Lactobacillus plantarum* maintains a moderate $\triangle$pH in spite of high organic acid concentrations. They suggest that this contributes to *L. plantarum*'s propensity to dominate the late stages of fermentation. Many bacteria are known to have inducible acid stress responses which confer increased acid tolerance. For example, if the pathogen *E. coli* O157:H7 is exposed to acetate, butyrate or propionate at neutral pH some 26 genes are up-regulated conferring increased acid tolerance at pH 3 (Arnold et al., 2001). Conversely, formate, a

major end-product of *E. coli*'s mixed acid fermentation, was found to suppress acid tolerance.

Benthin and Villadsen (1995) determined that the chirality of a fermentation acid (e.g. D- versus L-lactic acid) can greatly affect the relative toxicity of the acid. In experiments with *Lactobacillus delbruekii* subsp. *bulgaricus* D-lactic acid was found to be much more inhibitory than L-lactic acid. The main inhibitory effect was a prolonged lag phase and a decrease in the maximum cell density, however, they found no reduction the specific growth rate or biomass yield. It must be stressed, however, that these studies were performed at a constant pH of 6.0. At this pH , virtually all of the lactic acid ($pK_a = 3.86$) in the media would have been in the anionic form. The most likely mode of action of D-lactic acid was therefore as a competitive inhibitor of the malate-lactate transporter which transports lactic acid out of the cytoplasm. Thus, their results support the hypothesis that it is ultimately organic acid anion accumulation within the cytoplasm which is the ultimate cause of acid toxicity. The chirality of the acid, however, can greatly increase the rate at which acid accumulates within the cytoplasm by inhibiting transport. Benthin and Villadsen caution that the type of stereoisomer (D- or L-) and the extent to which acid export out of the cell is inhibited varies from species to species.

The ability of microorganisms to cope with acid stress will translate into competitive fitness when multiple species/strains are grown together. For example in continuous alcoholic fermentation by yeast Bayrock and Ingledew (2001) found that contaminating *Lactobacillus paracasei* were able to produce some 20 g/L of lactic acid before reaching steady state. This led to a rapid yeast decline and caused a 44% reduction in ethanol yield. Similarly van Beek and Preist (2002) found that whiskey fermentations are characterized by 3 distinct phases. In the first the yeast dominate

but are eventually overcome by *Lactobacilli*. In the final stages, the acid tolerant *L. acidophilus* and *L. crispatus* appeared and dominated. Martens et al. (1999) found that pH was the most important factor in determining the dominance of *L. curvatus* over *Enterobacter cloacae* when these 2 bacteria were grown together in MRS broth.

4.4 Functional Forms for Common Acid Tolerance Strategies

As mentioned above, there are 2 common strategies that bacteria use for dealing with the toxic effects of organic acids. The first strategy is based on maximizing the potential to make energy in the form of ATP. This is done by maintaining the pH of the cytoplasm above that of the medium. A problem with this strategy is that it promotes organic anion accumulation within the cytoplasm. The second common strategy is based on varying the cytoplasmic pH with the pH of the media. This reduces the risk of organic anion accumulation but still allows a modest $\triangle$pH sufficient for ATP synthesis.

One of the problems encountered with modeling such strategies is that quantifying intracellular levels of organic acid, and intracellular pH is problematic. *Fortunately, with only a few mild assumptions, it is possible to easily model these 2 acid tolerance strategies with only knowledge about the extracellular concentrations or organic acids and pH.*

A mathematical model for the intracellular partitioning of a toxic ionizing molecule at constant intracellular pH is given by Kooijman (2000, p.193-194). This model can be used to estimate the toxic effects of organic acids and bases. Letting cytoplasmic

lactic acid and its anion be given by $[HL]_0$ and $[L^-]_0$ respectively and the quantities in the medium be given by $[HL]_1$ and $[L^-]_1$ the partition coefficient is defined as $P_{01} = \frac{[L^-]_0+[HL]_0}{[L^-]_1+[HL]_1}$. This may be written as

$$P_{01} = \frac{a + b\left(1 + \frac{[L^-]_1}{[HL]_1}\right)}{1 + \frac{[L^-]_1}{[HL]_1}} \tag{4.18}$$

where a and b are the lumped constants $\frac{1+\frac{[L^-]_0}{[HL]_0}}{k_1+k_2\frac{[L^-]_0}{[HL]_0}}k_1\frac{\rho_0^2}{\rho_1^2}$ and $\frac{1+\frac{[L^-]_0}{[HL]_0}}{k_1+k_2\frac{[L^-]_0}{[HL]_0}}k_2\frac{\rho_0^2}{\rho_1^2}$ respectively (Kooijman, 2000). Here k_1 is the rate of export out of the cell, k_2 is the rate of uptake and ρ_0 is the affinity coefficient for the substance in the cytoplasm and ρ_1 is the affinity coefficient for the substance in the medium. Thus, a simple 2 parameter model can explain the accumulation of organic acid by a cell regulating its pH at a constant value.

If, on the other hand, the cytoplasmic pH is allowed to vary with the pH of the medium, the expressions a and b can no longer be considered lumped constants. However, the assumption that the cytoplasmic pH is maintained in constant proportion to the pH of the medium still allows the intracellular quantity $\frac{[L^-]_0}{[HL]_0}$ to be eliminated from a and b. Mathematically the derivative control of intracellular pH can be expressed as

$$\frac{[\mathrm{H}^+]_1}{[\mathrm{H}^+]_0} = \chi \tag{4.19}$$

where χ is a constant. This may be re-written using the definition of K_a as

$$\frac{[L^-]_0}{[HL]_0} = \chi\frac{[L^-]_1}{[HL]_1} \tag{4.20}$$

Substituting this into Eqn. 4.18

$$P_{01} = \frac{1+\chi\frac{[L^-]_1}{[HL]_1}}{k1+k2\chi\frac{[L^-]_1}{[HL]_1}}\frac{k1+k2\frac{[L^-]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}\frac{\rho_0^2}{\rho_1^2} \tag{4.21}$$

It should be noted that the dimensionless quantity $\chi\frac{[L^-]_1}{[HL]_1}$ may be relatively small or large so it can not be immediately eliminated. To make this clear note that

$$\chi\frac{[L^-]_1}{[HL]_1} = \frac{\mathrm{K}_a}{\left[\mathrm{H}^+\right]_0} \tag{4.22}$$

If $\chi\frac{[L^-]_1}{[HL]_1}$ is large, then $\left[\mathrm{H}^+\right]_0 << \mathrm{K}_a$ and

$$P_{01} \approx \left(\frac{1}{k_2}\frac{\rho_0^2}{\rho_1^2}\right)\frac{k_1+k_2\frac{[L^-]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}} \tag{4.23}$$

If on the other hand $\chi\frac{[L^-]_1}{[HL]_1}$ is small, then $\left[\mathrm{H}^+\right]_0 >> \mathrm{K}_a$ and

$$P_{01} \approx \left(\frac{1}{k_1}\frac{\rho_0^2}{\rho_1^2}\right)\frac{k_1+k_2\frac{[L^-]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}} \tag{4.24}$$

Lastly if $\chi\frac{[L^-]_1}{[HL]_1} = 1$ then $\left[\mathrm{H}^+\right]_0 = \mathrm{K}_a$ and

$$P_{01} \approx \left(\frac{2}{k_1+k_2}k_2\frac{\rho_0^2}{\rho_1^2}\right)\frac{k_1+k_2\frac{[L^-]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}} \tag{4.25}$$

These findings are summarized in Table 4.3. Thus, it has been found that a hyperbolic functional form results in each of the limiting cases of derivative pH control. This is the

Table 4.3: Partition coefficients for derivative regulation of pH . The parameters c–h are lumped parameters.

Condition	*Approx.* P_{01}	Lumped Model
$\left[\mathrm{H}^+\right]_0 << \mathrm{K}_a$	$\frac{1}{k_2}\frac{\rho_0^2}{\rho_1^2}\frac{k_1+k_2\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$	$\frac{c+d\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$
$\left[\mathrm{H}^+\right]_0 >> \mathrm{K}_a$	$\frac{1}{k_1}\frac{\rho_0^2}{\rho_1^2}\frac{k_1+k_2\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$	$\frac{e+f\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$
$\left[\mathrm{H}^+\right]_0 = \mathrm{K}_a$	$\frac{2}{k_1+k_2}k_2\frac{\rho_0^2}{\rho_1^2}\frac{k_1+k_2\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$	$\frac{g+h\frac{\left[L^-\right]_1}{[HL]_1}}{1+\frac{[L^-]_1}{[HL]_1}}$

same functional form that was derived for when the internal pH is constant although the parameterizations have completely different interpretations.

4.4.1 A Model for Competitive Inhibition of Acid Transporters

As discussed by Benthin and Villadsen (1995) the presence of a stereoisomer can act as an inhibitor of acid export. Competitive inhibition of the membrane transporters can be modeled by dividing the appropriate partition coefficient (a P_{01} taken from Table 4.3 for example) by a Michaelis-Menten type competitive inhibition term. This approach to modeling competitive inhibition of the lactate export mechanism is illustrated by Eqn. 4.26 where L^* represents the concentration of the inhibiting stereoisomer in the media and K_I is a sensitivity coefficient.

$$P_{01}^* = \frac{P_{01}}{1+\frac{K_{L^*}}{L^*}} \tag{4.26}$$

4.5 Semi-Mechanistic Approaches to Modeling pH and Buffer Capacity

The simplifications permitted under derivative pH regulation shed light on why modeling acid toxicity as simply proportional to total extracellular acid is incorrect. Recall that in Chapter 3 the following functional form was assumed to describe the energetic cost associated with lactic acid.

$$\text{acid toxicity} \quad = \quad -\kappa([L^-] + [HL])Q \tag{4.27}$$

In particular, it is clear from Eqn. 4.21 that modeling acid toxicity as proportional to the total acid concentration implicitly assumes that the intracellular and extracellular pH are equivalent and that the partition coefficient (P_{01}) is constant. Since *Lactococcus lactis* maintains its internal pH at ≈ 5.1 (Russell, 1992) the toxicity in the late stages of fermentation (when typically pH $\rightarrow 3.9$) will be systematically underestimated rather than overestimated. Clearly a different approach is needed.

The preceding sections demonstrated that a more accurate method of predicting pH and acid dissociation is required. Also, if such a method were available, it could be used to model the common acid tolerance strategies of bacteria in a mechanistically accurate way.

A unified approach to modeling changes in pH during batch fermentation will now be developed. A flexible semi-mechanistic approach is presented which uses local polynomial regression to model the buffering influence of complex or undefined components in the growth media while defined or measurable components are modeled using expressions based on acid-base theory. When applied to homolactic batch fermenta-

tion, this method can be implemented using only a single titration of the unfermented medium to give accurate predictions throughout a fermentation. Because models of the buffering capacity of a growth medium can be parameterized independently from the growth data, this approach may reduce the number of parameters which need to be estimated from growth data. The semi-mechanistic approach is validated by seeing how well it predicts pH given the outputs from an existing fermentation model. This approach may be applicable to modeling other bacterial fermentations, including competitive growth among strains with different acid tolerance strategies.

4.5.1 Existing Methods for Modeling pH

Mathematical modeling has become a popular method for exploring the effect of pH on the dynamics of growth and the outcome of inter-specific competition (Breidt and Fleming, 1998; McDonald and Sun, 1999; Versari et al., 1999; Martens et al., 1999; Malakar et al., 1999). Clearly, the extent to which the accumulation of an end product will affect the progression of a fermentation will depend in a large part on how much it influences the pH of the growth medium. A common approach is to model pH as a simple function of the end product with adjustable parameters. The parameters of such models are estimated from experimental growth data for a particular species of bacteria growing in a specific medium. For example, Martens et al. (1999) studied the competitive interactions between *Lactobacillus curvatus* and *Enterobacter cloacae*. They used the following simple nonlinear functional form to predict pH

$$\mathrm{pH} \quad = \quad \frac{\mathrm{pH}_i + a_1 P}{1 + a_2 P} \tag{4.28}$$

where P is the concentration of lactate, pH_i is the initial pH and a_1 and a_2 are adjustable parameters which must be estimated from experimental data.

They then used the pH predictions to modify the growth rate of *L. curvatus* using Eqn. 4.29

$$\mu_l = 4\mu_{\mathrm{opt}l}\frac{S}{K_{Sl}+S} \tag{4.29}$$

$$+\frac{(\mathrm{pH} - \mathrm{pH}_{\mathrm{min}l})(\mathrm{pH}_{\mathrm{max}l} - \mathrm{pH})}{(\mathrm{pH}_{\mathrm{max}l} - \mathrm{pH}_{\mathrm{min}l})^2} - b_2 P \tag{4.30}$$

Here S is the concentration of glucose and P is the concentration of lactate, $\mathrm{pH}_{\mathrm{min}l}$ is the minimum pH for growth, $\mathrm{pH}_{\mathrm{max}l}$ is the maximum pH for growth, K_{Sl} is the half saturation constant of glucose with respect to the maximum growth rate $\mu_{\mathrm{opt}l}$ and b_2 is an adjustable parameter. Martens et al. found that the inhibition term for lactate was not required for *Enterobacter cloacae* and they used an equation analogous to Eqn. 4.29 to model the specific growth rate.

$$\mu = 4\mu_{\mathrm{opt}e}\frac{S}{K_{Se}+S} \tag{4.31}$$

$$+\frac{(\mathrm{pH} - \mathrm{pH}_{\mathrm{min}e})(\mathrm{pH}_{\mathrm{max}e} - \mathrm{pH})}{(\mathrm{pH}_{\mathrm{max}e} - \mathrm{pH}_{\mathrm{min}e})^2} \tag{4.32}$$

They found that the minimum pH for growth for *E. cloacae* was 5.6 while that for *L. curvatus* was 4.3 indicating that *E. cloacae* is more sensitive to pH. Eqn 4.28 clearly depends on the concentration of lactate alone. Unfortunately, it is not clear how to extend this formula for a situation in which several acid (or base) species are present.

Change in pH has also been modeled in the context of differential equations. Breidt and Fleming (1998) were interested in describing the competitive interactions between *Lactococcus lactis* and *Listeria monocytogenes*. Their model consisted of a

system of 5 differential equations. In their approach, the concentration of hydrogen ions was modeled using a simple logistic form.

$$\frac{dP}{dt} = \rho C(1 - \frac{P}{kp}) - \kappa \frac{dM}{dt} \tag{4.33}$$

where P represents the hydrogen ion (i.e. proton) concentration, C is the concentration of total lactic acid, M is the malic acid concentration which removes κ protons as it is fermented and kp is a saturation parameter. The $\kappa\frac{dM}{dt}$ term describes the alkalinization of the medium by malolactic fermentation and $\rho C(1 - \frac{P}{kp})$ describes acidification of the medium by homolactic fermentation. They were then able to calculate the concentrations of protonated and un-protonated acids using the Henderson-Hasselbach equation (Butler and Cogley, 1998). A weakness of this approach is that $\frac{dP}{dt}$ is not zero when $\frac{dC}{dt}$ is zero. Thus, this approach to modeling pH may not be appropriate if, for example, a lag phase is expected.

There have been only a few attempts to predict pH in complex dynamic environments. Horiuchi et al. (2001) used a 3-layer neural network with a back-propagation algorithm to model changes in product distribution as a function of pH in chemostat cultures. This method involved training the neural network during controlled step changes in pH. Once trained, the neural network could predict transient changes in product distribution for continuously varying pH. Van Vooren et al. (2001) has recently introduced a method called "automatic model building" to accomplish similar goals. This method attempts to infer the constituents of a solution by estimating individual pK_a and concentrations of potential components from titration data. Limits can be placed on the number of components to be estimated.

Here a method is presented that is straightforward, broadly applicable, and allows

the modeler to link pH dynamics directly to the metabolic activities of bacteria. The method uses a generalization of standard acid-base calculations so that calculations involving simple buffers with known dissociation constants can be executed conveniently. When complex buffers such as those common in foods and in microbiological media must be considered. Local polynomial regression is used in order to extend the standard approaches. To make these techniques relevant for dynamic systems modelling, a simple algorithm is also presented for updating pH during the course of a simulated fermentation.

4.5.2 Computational Approach

The basic formula useful in modeling pH and buffer capacity will now be presented. The approach is to use algebraic expressions of the dynamics rather than differential equations. Without any loss of flexibility it is assumed that the pH is fixed and known in all calculations. Further justification for this assumption is given in the discussion of generalized titrations.

Brønsted-Lowry Acids and Bases

The general concept of what kind of molecule constitutes and "acid" or "base" is given by the Lewis theory of acids and bases. This approach considers and acid to be an electron pair acceptor (for example boron triflouride BF_3) and a base to be an electron pair doner (for example trimethylamine $N(CH_3)_3$). A more limited concept is the Brønsted-Lowry theory which regards an acid as a proton (H^+) doner (for example (HCl) and a base as a proton acceptor (for example NaOH). The Lewis theory of acids and bases is considerably more complex than that due to Brønsted-Lowry.

Fortunately for most microbiological experiements, the simpler Brønsted-Lowry is usually sufficient. In the development which folows it will always be assumed the acid or base reactions can be explained with the Brønsted-Lowry theory and that the results of this development are not sufficient to explain reactions involving strict Lewis acids or bases.

Van Slyke (1922) introduced buffer capacity as a measure of the ability of a buffer to resist change in pH with the addition of base during a titration. He defined the buffer capacity (β) of a solution as

$$\beta \equiv \frac{\partial C_b}{\partial \mathrm{pH}} \tag{4.34}$$

where C_b is the number of moles of base added per liter of solution. Van Slyke (1922) was also the first to demonstrate the additivity of buffer capacities. For example if a solution consists of n compounds each of which contribute to β, then a decomposition of the form

$$\beta = \beta_1 + \beta_2 + \ldots + \beta_n \tag{4.35}$$

is admissible where β_i is called the ith partial buffer capacity component.

When the dissociation constants for various compounds are known, the partial buffer contribution of that compound can be expressed explicitly. For example, the buffer capacity of an acid with a single dissociable proton has the form

$$\beta = \ln(10)\left(C\frac{K_a\left[\mathrm{H}^+\right]}{(\left[\mathrm{H}^+\right]+K_a)^2}\right) + \ln(10)\left(\frac{K_w}{\left[\mathrm{H}^+\right]} + \left[\mathrm{H}^+\right]\right) \tag{4.36}$$

where C is the concentration of the acid, K_a is the dissociation constant of the single

proton and K_w is the ion-product constant of water ($K_w \approx 1.8 \times 10^{-16}$ at 25°C and zero ionic strength). Here $\ln(10)(K_w/\left[H^+\right] + \left[H^+\right])$ is the partial buffer contribution of water and $\ln(10)C(K_a\left[H^+\right])/(\left[H^+\right] + K_a)^2)$ is the partial contribution due to the acid. For an acid with 2 dissociable protons the expression for β becomes

$$\beta = \ln(10)C\left(\frac{K_{a1}\left[H^+\right]^3 + 4K_{a1}K_{a2}\left[H^+\right]^2 + K_{a1}^2K_{a2}\left[H^+\right]}{\left(\left[H^+\right]^2 + K_{a1}\left[H^+\right] + K_{a1}K_{a2}\right)^2}\right) \tag{4.37}$$

$$+\ln(10)\left(\frac{K_w}{\left[H^+\right]} + \left[H^+\right]\right) \tag{4.38}$$

where K_{a1} and K_{a2} are the acid dissociation constants for the primary and secondary dissociable protons respectively. *Notice that the partial contribution of a diprotic acid is not equivalent to the sum of 2 monoprotic acids (Butler and Cogley, 1998, p.134).*

Similar formulae may be derived for buffers with any number of dissociable groups. Eqn. 4.39 is a computational formula useful in the computation of the partial buffer capacity of a compound with an arbitrary number (n) of dissociable groups.

$$C_b = \ln(10)\left[H^+\right]\frac{C\sum_{i=1}^{n}\left[(n+1-i)\left[H^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]}{\sum_{i=1}^{n+1}\left[\left[H^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]} \tag{4.39}$$

Theorem 1. *The partial contribution of a single n-protic acid having molar concentration C in a solution containing C_b moles of base is*

$$\frac{C\sum_{i=1}^{n}\left[(n+1-i)\left[H^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]}{\sum_{i=1}^{n+1}\left[\left[H^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]} \tag{4.40}$$

Proof of 4.39. Proof is by induction. For clarity, X_n^{j-} is defined to be a molecule having j dissociated groups out of a total of n and C_k to equal $\sum_{p=1}^{k}[X_k^p]$. Following the development by (Butler and Cogley, 1998, pp.130-134) it is noted that for a monoprotic weak acid ($n = 1$), in a solution containing strong acid at a concentration of C_a the equilibrium and conservation equations

$$K_{a1} = \frac{\left[\mathrm{H}^+\right] X_1^{1-}}{X} \tag{4.41}$$

$$C = [X_1] + \left[X_1^{1-}\right] \tag{4.42}$$

$$K_w = \left[\mathrm{H}^+\right] \left[OH^-\right] \tag{4.43}$$

along with the the charge balance

$$C_b = \left[X_1^{1-}\right] + \left[OH^-\right] + C_a - \left[\mathrm{H}^+\right] \tag{4.44}$$

$$\tag{4.45}$$

gives

$$C_b = \frac{CK_{a1}}{\left[\mathrm{H}^+\right] + K_{a1}} + \frac{K_w}{\left[\mathrm{H}^+\right]} + C_a - \left[\mathrm{H}^+\right] \tag{4.46}$$

Here $K_w/\left[\mathrm{H}^+\right] - \left[\mathrm{H}^+\right]$ is the partial contribution due to the dissociation of water and C_a is the partial contribution due to strong acid. Therefore, the partial contribution of the weak acid is

$$C_{b1} = \frac{K_{a1}}{\left[\mathrm{H}^+\right] + K_{a1}} \tag{4.47}$$

which is 4.39 evaluated at $n = 1$. Now, assuming that (4.39) is true for $n = m$, which

gives

$$C_{bm} = \frac{C\sum_{i=1}^{m}\left[(m+1-i)\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1-i}(Kaj)\right]}{\sum_{i=1}^{m+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1-i}(K_{aj})\right]} \tag{4.48}$$

it must be proven that 4.39 is true when $n = m+1$.

Note that

$$C_{b(m+1)} = C_{bm} + (m+1)\left[X_{m+1}^{(m+1)-}\right] \tag{4.49}$$

From C_{bm} it is clear that

$$\left[X_{m+1}^{(m+1)-}\right] = C_{(m+1)}\frac{\prod_{j=1}^{m+1}(Kaj)}{\sum_{i=1}^{m+1+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1+1-i}(K_{aj})\right]} \tag{4.50}$$

so that

$$C_{b(m+1)} = \frac{C_{m+1}\sum_{i=1}^{m}\left[(m+1-i)\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1-i}(Kaj)\right]}{\sum_{i=1}^{m+1+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1+1-i}(K_{aj})\right]} \tag{4.51}$$

$$+(m+1)C_{(m+1)}\frac{\prod_{j=1}^{m+1}(Kaj)}{\sum_{i=1}^{m+1+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1+1-i}(K_{aj})\right]} \tag{4.52}$$

$$= \frac{C\sum_{i=1}^{m+1}\left[(m+1+1-i)\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1+1-i}(K_{aj})\right]}{\sum_{i=1}^{m+1+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{m+1+1-i}(K_{aj})\right]} \tag{4.53}$$

which is 4.39 evaluated at $n = m+1$ completing the proof. □

Along with the fact that $d\left[\mathrm{H}^+\right]/d\mathrm{pH} = -\ln(10)\left[\mathrm{H}^+\right]$, Eqn. 4.39 can be differentiated with respect to $\left[\mathrm{H}^+\right]$ to give a formula for the buffer capacity of an n-protic

acid.

$$\beta = \ln(10)\left[\mathrm{H}^+\right]\left[\frac{BA' - AB'}{B^2}\right] \tag{4.54}$$

where

$$\begin{aligned}
A &= C\sum_{i=1}^{n}\left[(n+1-i)\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right] \\
B &= \sum_{i=1}^{n+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right] \\
A' &= C\sum_{i=2}^{n}\left[(n+1-i)(i-1)\left[\mathrm{H}^+\right]^{i-2}\prod_{j=1}^{n+1-i}(K_{aj})\right] \\
B' &= \sum_{i=2}^{n+1}\left[(i-1)\left[\mathrm{H}^+\right]^{i-2}\prod_{j=1}^{n+1-i}(K_{aj})\right]
\end{aligned}$$

Note that Eqn. 4.54 reduces to Eqn. 4.36 for $n = 1$ and reduces to Eqn. 4.37 for $n = 2$.

Van Slyke (1922) points out that the formulae corresponding to a base of concentration C is obtained by simply replacing the K_a in the preceding expressions with Kw/K_b. Also, for a compound having both acid and base components (amphoteric) the expression for β is simply the summation of its acid and base components. If the titration is by acid then Eqn. 4.39 can still be used assuming the convention that C_b take on negative values with magnitude equal to the number of moles of *acid* added per liter of solution.

Ionic Strength

The partial ionic strength contribution can be obtained by a slight modification of the formula for Cb. As defined in Eqn. 4.3

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2 \tag{4.55}$$

therefore Eqn. 4.39 can be modified in the following way to give the partial ionic strength contribution (I).

$$I = \frac{0.5C\sum_{i=1}^{n}\left[(n+1-i)^2\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]}{\sum_{i=1}^{n+1}\left[\left[\mathrm{H}^+\right]^{i-1}\prod_{j=1}^{n+1-i}(K_{aj})\right]} \tag{4.56}$$

The ionic strength contributions from water however must be calculated as

$$I = 0.5\left(\frac{\mathrm{K}_w}{\left[\mathrm{H}^+\right]} + \left[\mathrm{H}^+\right]\right) \tag{4.57}$$

Concentrations of All Ionic Species

By a similar alterations to Eqn. 4.39 formulae for the concentrations of all ionic species can be found. This is done by simply realizing that each term in the summation relates to the concentration a particular ion. Due to its derivation from a charge balance expression, each term in the summation in Eqn. 4.39 refers to an ion concentration multiplied by its charge $(n+1-i)$. To obtain the concentrations of ions the $n+1-i$ term is simply set to 1 and each term is taken as an element in a list rather than summed. The concentration of the uncharged form may be gotten by subtraction from the total concentration (C) the sum of the concentrations of all the ionic species.

Generalized Titrations

Normally, the titrants used in performing titrations are strong acids or strong bases. Because these titrants dissociate essentially completely, their contribution to C_b is easily calculated. In many instances (e.g. food grade applications) strong acids or bases are undesirable and weak acids or bases are used instead. In order to still use the data from such titrations in calculating buffer capacities, a more general approach to calculating C_b must be taken. When a weak acid or weak base is used as the titrant the process of inferring a relationship between pH and C_b will be referred to as a *generalized titration*.

The contribution to C_b from a weak acid or base will vary depending on the extent to which it dissociates and thereby depend on the pH and the ionic strength. For clarity, let C_b (pH) indicate C_b at a particular pH . If the partial contribution *of the weak titrant* to C_b (pH) is denoted by C_τ(pH) then

$$C_b \text{ (pH)} \quad = \quad -C_\tau(\text{pH }) \tag{4.58}$$

That is, the moles of base added is the negative of the moles of base that would be required to bring the buffer τ to the pH in question. Similarly, for complex buffers containing more than 1 buffering compound τ_1,τ_2,..., the additivity of the partial C_b can be exploited and the following definition made

$$C_b \text{ (pH)} \quad = \quad -(C_{\tau_1}(\text{pH }) + C_{\tau_2}(\text{pH }) + \ldots + C_{\tau n}(\text{pH })) \tag{4.59}$$

That is, the moles of base added is equal to the negative of the sum of the con-

tributions from the weak acids or bases. To illustrate these points consider Fig. 4.2 which is a plot of the partial C_b of a 0.01M NaOH ($pK_b \approx$ -3) solution over the pH range 2–12. Over this interval the contribution to C_b remains a constant 0.01 M because NaOH is a strong base and its dissociation is essentially complete over this pH range. In contrast consider Fig. 4.3 which gives the partial contributions to C_b from a 0.01M solution of gluconic acid ($pK_a \approx 3.6$). At low pH, this weak acid is completely protonated and therefore makes no contribution to C_b . Conversely at high pH all of the gluconic acid is dissociated and its contribution to C_b is equivalent to the negative of its concentration (i.e. the conjugate base of a weak acid is a strong base). *Clearly to determine the moles of base available for titration the dissociation of the weak titrant at the particular pH in question needs to be considered.* Thus, for generalized titrations, determination of a mathematical relationship between C_b and pH requires a function which maps pH to C_b (an approach which at first may seem counter-intuitive to those acquainted only with strong acid or strong base titrants). It is therefore the convention to treat pH as the predictor variable and C_b as the response variable.

Dynamic Systems Modeling

The buffer capacity measure is also useful for dynamic system modeling. It can provide a simple mechanism for updating pH predictions over time. To develop an updating formula, simply re-write β as

$$\frac{\partial \text{pH}}{\partial t} = \frac{1}{\beta}\frac{\partial C_b}{\partial t} \tag{4.60}$$

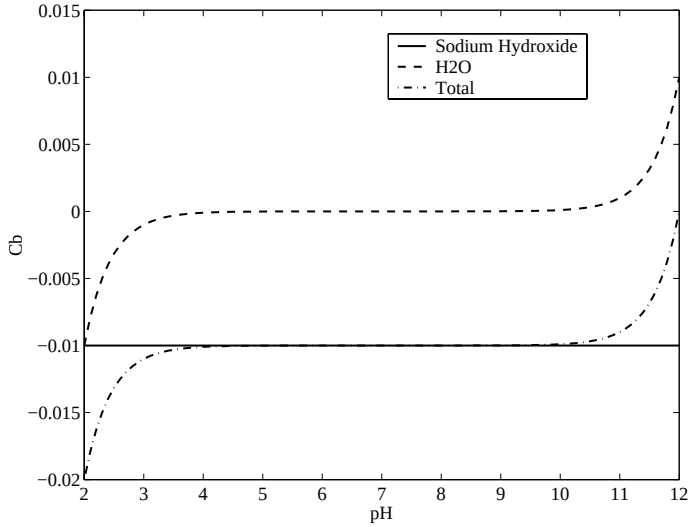

Figure 4.2: Partial contributions to C_b from the components of a 0.01M NaOH solution. The partial C_b (solid line) is nearly constant and represents the amount of the species available for titration.

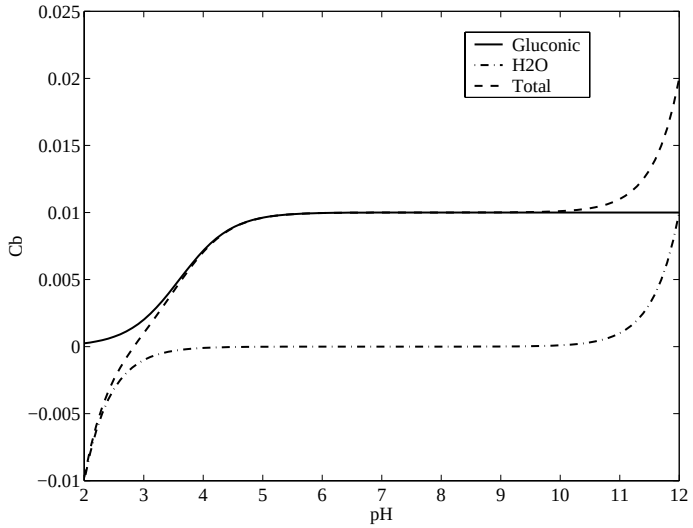

Figure 4.3: Partial contributions to C_b derived from the components of a 0.01M gluconic acid solution. The partial C_b (solid line) is the relevant quantity in a generalized titration since it represents the amount of the species available for titration.

It may be assumed that $\left[\mathrm{H}^{+}\right]$ is constant during the small time interval ∂t so that the partial contribution to the change in pH due to a small change in the concentration of an n-protic acid is

$$\frac{\partial \mathrm{pH}}{\partial t} = \frac{1}{\beta_t}\left(\frac{\frac{dC}{dt}\sum_{i=2}^{n+1}\left(\prod_{j=1}^{i-1}K_{a_j}\left[\mathrm{H}^{+}\right]^{n+1-i}\right)}{\sum_{i=1}^{n+1}\left(\prod_{j=1}^{i-1}K_{a_j}\left[\mathrm{H}^{+}\right]^{n+1-i}\right)}\right) \tag{4.61}$$

β_t is the buffer capacity given the concentration of the acid at time t and dC/dt is the instantaneous rate of change in the concentration of the acid. β_t indicates that the buffer capacity is calculated given the concentrations of weak acids and bases at time t. Consider this simplistic Monod type model for single substrate limited batch fermentation. The bacteria (N) produce a weak acid (L) as a metabolic waste product of glucose (G) fermentation. The production of lactic acid should change (decrease) the pH .

$$\begin{aligned}
\text{Cell Density:}\quad \frac{dN}{dt} &= \frac{\mu_{\max}}{Y}\left(\frac{G}{K_G+G}\right)N-\delta N\\
\text{Glucose:}\quad \frac{dG}{dt} &= -\mu_{\max}\left(\frac{G}{K_G+G}\right)N\\
\text{Lactic Acid:}\quad \frac{dL}{dt} &= \mu_{\max}\left(\frac{G}{K_G+G}\right)N\\
\mathrm{pH}:\quad \frac{d\mathrm{pH}}{dt} &= \frac{1}{\beta_t}\frac{dC_b\,(\frac{dL}{dt})}{dt}
\end{aligned} \tag{4.62}$$

where $\mu_{\max}$ is the maximum glucose consumption rate, Y is the cell yield, δ is the specific death rate and K_G is the half-saturation constant for glucose consumption.

While this model gives the basic mathematical formalism there are obstacles which need to be overcome to make this approach viable. The paramount difficulty in applying this approach directly is the fact that microbial growth media is often quite

complex consisting of undefined components such as yeast extract, peptic digests etc. In order to predict pH continuously in time a way for calculating β_t at any required time is needed.

Local Polynomial Modeling of Buffers

Eqns. 4.39 and 4.54 are applicable so long as the concentrations of all acids and bases and their dissociation equilibria are known. Unfortunately, these requirements are usually not met in biological applications due to the complexity and undefined nature of most growth media and biological buffers. Although the buffer capacities of such solutions can be quite complex chemically, it is reasonable to assume that the relationship between C_b and pH should be smooth and continuous. While a titration can be used to determine the relationship between C_b and pH there is error in measurement. It is also not possible to observe the effect of adding a small amount of base at every possible pH value. A flexible regression procedure is required which will provide a continuous differentiable prediction function for C_b given pH .

In order to obtain such a predictor, local polynomial regression is used (for an introduction to the subject see (Fan and Gijbels, 1996, pp.57-107)) to predict pH as a function of C_b . In this type of regression analysis, a model of the form

$$Y_i = X_i\beta_i + \epsilon_i \tag{4.63}$$

is assumed as in standard linear least squares but the procedure is carried out locally by regressing only on a neighborhood of points about Y_i. This is practically

implemented by weighted least squares where the weighted equations

$$W_i Y_i = W_i X_i \beta_i + \epsilon_i \tag{4.64}$$

result in the least-squares estimator

$$\hat{Y}_i = \left(X_i (X_i' W_i X_i)^{-1} X_i' W_i \right) Y_i + \epsilon_i \tag{4.65}$$

The weights in W are generally functions of a distance metric between the prediction point and all other data points. These functions, called smoothing kernels, typically give more weight to points closer to the prediction point X_i. Many common kernels also give zero weight to points farther than a distance h from the prediction point. This distance h is called the bandwidth. The magnitude of the bandwidth controls the number of linearly independent $\hat{\beta}_i$ and thus controls the overall model complexity.

There are several reasons for choosing local polynomial regression over other smoothing methods.

- Automatically controls for model complexity
- Automatically adjusts for unevenly spaced design points
- Behaves well at the boundaries of the data
- Estimates of the derivative are easily obtained

The ease with which derivative estimates are obtained is especially nice since it allows for the straightforward calculation of β from experimental titration data. Predictors of C_b and β for mixtures of undefined and defined buffers consist of summations

of local polynomial regression estimators for the complex components and algebraic predictors (i.e. Eqn. 4.54) of defined buffer components as described above. Thus, the local polynomial regression procedure provides a transparent means of extending the computational formula given for buffers with known dissociation constants to complex buffers containing undefined components.

One important consideration relating to the estimation of derivatives from experimental data is the influence of spurious data points. For example, most pH probes manifest drift in measurement during use and must be re-calibrated. Re-calibrations should also be performed whenever the pH exceeds the calibration pH . The occurrences of drift, re-calibration or incorrectly recorded data can lead to small jumps in the titration data that can create spurious jumps in the estimation of β. *For this reason, iteratively re-weighted least squares is used to down-weight unlikely data points.*

Generalized Additive Modeling of Buffer Mixtures

The smoothing afforded by local polynomial regression allows titration relations of complex buffers to be included in pH and buffer capacity calculations. However, in many instances buffers solutions are mixtures of defined as well as undefined components. In this section, some formal statistical approaches for modeling such mixtures are developed.

Consider the titration of weak acid at concentration C in a complex buffer. In the previous section, the additivity of C_b was noted. This will be used to advantage here where the solution contains simple as well as complex buffer components. Let $C_{b\ i}$ be the sum of the contributions from the simple and complex buffer components

at pH_i.

$$C_{b\ i} = \frac{CKa_1}{\left[\mathrm{H}^+\right]_i + Ka_1} + \mathcal{F}(\mathrm{pH}_i) + \epsilon_i \tag{4.66}$$

where $\epsilon_i \sim N(0, \sigma^2)$ and $\mathcal{F}(\mathrm{pH}_\rangle)$ is the contribution from the complex media in which the weak acid is dissolved.

To fit this model to titration data the following algorithm is employed. The simple buffer contributions, which are assumed to be known without error, are subtracted from the observed C_b . The result of this subtraction is the partial contribution to C_b from the unknown components. $\mathcal{F}(\mathrm{pH}_i)$ is then regressed on the residuals using local polynomial regression. Currently a global smoothing parameter is determined by generalized cross-validation. As the smoothing parameter is decreased, $\mathcal{F}(\mathrm{pH}_i)$ tends to interpolate the n data points and approaches a $(n-1)^{th}$ order polynomial. For larger values of the smoothing parameter $\mathcal{F}(\mathrm{pH}_i)$ smooths the data and approaches a p^{th} order polynomial where p is the underlying order of polynomial model. Since estimates of the first derivative of C_b are needed for calculating the buffer capacity, p is taken to be 3. Although a linear or quadratic polynomial could also give derivative estimates, this choice has been shown to reduce variance in the prediction of the derivative (Fan and Gijbels, 1996).

This procedure has several benefits. First even if the concentrations C of the weak acid are in error the prediction of the over-all model will still fit the data well (albeit with potentially an increase in variance). Secondly, the extent of smoothing by $\mathcal{F}(\mathrm{pH}_i)$ can be used as a measure of the importance of the contributions of the simple buffer components. In other words, $\mathcal{F}(\mathrm{pH}_i)$ can be used in statistical inference.

In particular, the significance of each buffer's contribution to the overall buffering

of the solution can be determined. This is carried out by re-estimation of $\mathcal{F}$ after the particular defined component in question is removed from the model. The new fit is then compared to the old and significance of the indicates significance of the buffering component. In addition one can test for the contribution to buffering of the undefined components by testing the standard null hypothesis that $\mathcal{F}(\mathrm{pH}_i)$ is everywhere 0.

Basic methods of inference for local polynomial regression estimators was given by Cleveland et al. (1988). Let $\mathcal{H}$ represent the projection matrix which maps the observations to predictions (the Hat matrix)

$$\hat{y} = \mathcal{H}y \tag{4.67}$$

In the case of local polynomial regression the Hat matrix is of the form

$$\mathcal{H} = X(X'WX)^{-1}X'W \tag{4.68}$$

Letting $R_{\mathcal{H}} = (I - \mathcal{H})$

$$\hat{\epsilon} = R_{\mathcal{H}}y \tag{4.69}$$

Thus the variance covariance matrix for the vector $\hat{y}$ is $\sigma^2\mathcal{H}\mathcal{H}'$ and the variance covariance matrix for ϵ is $\sigma^2 R_{\mathcal{H}}R'_{\mathcal{H}}$. Using the fact that the trace of a projection matrix is equivalent to its rank, the following degrees of freedom can be calculated and an F test for the difference between the fit of a null model and alternative model can be formulated. Consider, for example, $\mathcal{H}1$ to represent the smooth which results by regressing on observed C_b after <u>all</u> measured buffer contributions are adjusted for and let $\mathcal{H}2$ represent the smooth which results when 1 or more of the measured

components is not adjusted for (ignored). In order to test for the significance of these components the following procedure is carried out.

$$\nu_1 = tr(R_{\mathcal{H}2} - R_{\mathcal{H}1}) \tag{4.70}$$

$$\nu_2 = tr((R_{\mathcal{H}2} - R_{\mathcal{H}1})(R_{\mathcal{H}2} - R_{\mathcal{H}1})' \tag{4.71}$$

$$\delta_1 = tr(R_{\mathcal{H}1}) \tag{4.72}$$

$$\delta_2 = tr(R_{\mathcal{H}1}R'_{\mathcal{H}2}) \tag{4.73}$$

Then to test model 2 against model 1 formulate the variance ratio

$$\hat{F} = \frac{\frac{RSS_2 - RSS_1}{df2 - df1}}{\frac{RSS_1}{df1}} \tag{4.74}$$

$$= \frac{(y'R_{\mathcal{H}2}y - y'R_{\mathcal{H}1}y)/\nu_1}{(y'R_{\mathcal{H}1}y)/\delta_1} \tag{4.75}$$

which is distributed approximately as F with numerator degrees of freedom ν_1^2/ν_2 and δ_1^2/δ_2 (Cleveland et al., 1988).

Ionic Strength and Temperature Adjustments

Recall that ionic strength is a measure of the electrical interactions occurring within a solution. For relatively dilute solutions (see Pitzer (1991) for more general cases) ionic strength is adequately defined by

$$I = \sum_{i=1}^{n}(c_i z_i^2) \tag{4.76}$$

where c_i is the concentration of the i^{th} ion in the solution and z_i is its charge.

Debye and Hückle derived thermodynamic equations governing the properties of

non-ideal solutions of electrolytes. However, the Debye-Huckle equations are well known to behave poorly for ionic strengths above 0.1 (Butler and Cogley, 1998).

The standard Davies equation (see Eqn. 4.14) can predict with acceptable accuracy ion activities for ionic strengths up to 0.5. At higher ionic strengths there is an increased tendency for ion-ion pairing. This reduces the number of ions contributing to the ionic strength. In such regimes, the Davies equation over-estimates the decrease in the pK_a (see Fig. 4.4). Higher ionic strengths can be modeled by reducing the "salting out" parameter b. However this reduces accuracy at low ionic strength. Samson et al. (1999), fortunately, give a simple modification of the standard Davies equation that reduces b linearly in proportion to the ionic strength.

$$-\log(\gamma_i) = Az_i^2\left(\frac{\sqrt{I}}{1+\sqrt{I}} - ((-1/30)I + b)I\right) \tag{4.77}$$

Samson et al. suggest that this modified version yields reliable results up to an ionic strength of 1.2 M.

Although the Davies and Samson equations can be used to calculate the activity coefficients, they require ionic strength as an input. The activity coefficients, however, affect the extent of dissociation in the buffer components and therefore affect ionic strength. In order to achieve optimal activity coefficient predictions an iterative refinement procedure can be used. It was found that 3-4 iterations are usually sufficient for convergence of activity coefficient estimation.

To demonstrate the flexibility of this approach a theoretical buffer was simulated containing 6 dissociable protons. The pK_a of the acid at $I = 0$ are 3,5,7,9,11, and 12. Eqn. 4.77 was used to predict the activity coefficients over the pH range [1,13]. First the buffer capacity was simulated to verify that all peaks were resolved (see Fig. 4.5).

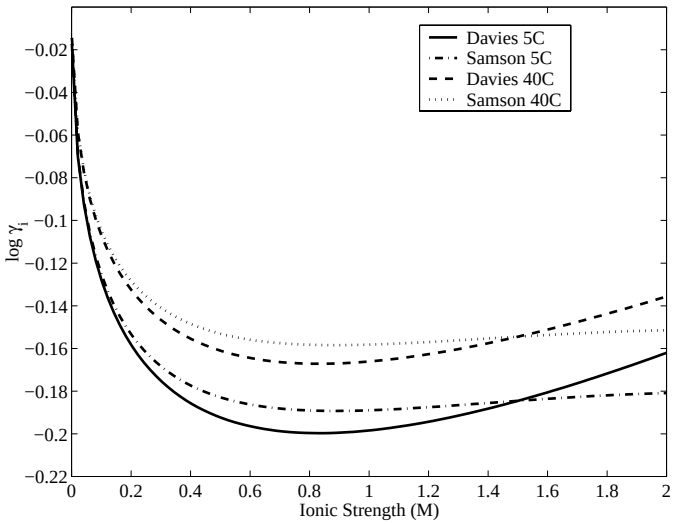

Figure 4.4: Comparison of the Davies equation with the modification due to Samson et al. (1999).

Next the ionic strength changes that would occur during a titration of such a solution (see Fig. 4.6) were determined.

Ionic Strength of Complex Buffers

In this section a useful lower bound for the ionic strength contribution from a complex buffer is introduced. Consider, the charge balance of a solution of NaOH and C moles L^- monoprotic weak acid HA.

$$\left[Na^+\right] + \left[\mathrm{H}^+\right] \quad = \quad \left[OH^-\right] + \left[A^-\right] \tag{4.78}$$

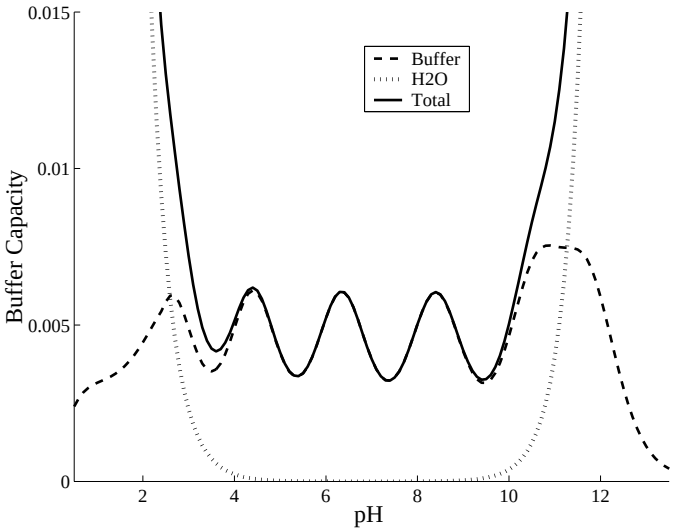

Figure 4.5: Partial buffer capacity diagram of a 15 mM solution of a theoretical acid with 6 pK_a. To account for the changes in ionic strength, the pK_a's were modified by their activities as calculated by the Davies equation. The pK_a of the acid at I=0 are 3,5,7,9,11, and 12.

which is re-arranged to give C_b .

$$\left[Na^+\right] = \left[OH^-\right] + \left[A^-\right] - \left[\mathrm{H}^+\right] \tag{4.79}$$

$$C_b = \frac{CK_a}{K_a + \left[\mathrm{H}^+\right]} + \frac{\mathrm{K}_w}{[OH^-]} - \left[\mathrm{H}^+\right] \tag{4.80}$$

Comparing this to the ionic strength of the solution

$$I = 0.5(\left[Na^+\right] + \left[OH^-\right] + \left[A^-\right] + \left[H^+\right]) \tag{4.81}$$

it is noticed that $I = C_b + \left[\mathrm{H}^+\right]$. Unfortunately this relationship does not hold when compounds with multiple dissociable groups or salts containing ions with charges greater than 1 or -1 are present. However, in such cases it will always be the case

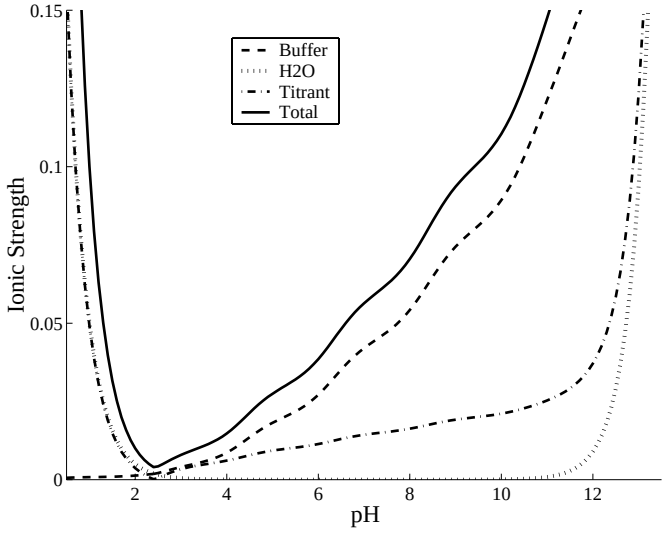

Figure 4.6: Partial ionic strength diagram of a 15 mM solution of a theoretical acid with 6 pKa. To account for the changes in ionic strength, the pK_a's were modified by their activities as calculated by the Davies equation. The pK_a of the acid at I=0 are 3,5,7,9,11, and 12. Clearly, as the pH increases, the predominant species are more negatively charged and have greater affect on the ionic strength.

that $I \geq C_b + \left[\mathrm{H}^+\right]$. **This fact leads us to suggest $C_b + \left[\mathbf{H}^+\right]$ as a lower bound for the ionic strength in complex buffers.** This is viewed as a convenient improvement over the zero ionic strength assumption for complex buffers containing undefined components.

4.5.3 Initial pH prediction

A natural question to ask after constructing such a buffer is what should the initial pH of such a solution be if it were actually constructed. Initial pH determination may be accomplished by determining the pH at which 4.39 is zero. Practically speaking this is the pH at which no acid or base has been added. This is efficiently done for simple or complex buffers using scalar zero–finding algorithm.

4.5.4 Discussion

The approach that has been taken in modeling pH and buffer capacity in complex buffers is an example of semi-mechanistic modeling. Semi-mechanistic models contain both explicit as well as partially specified functional relationships (Wood, 2001) and are analogous to statistical models such as partial splines and generalized additive models. Typically a spline or some other "automatic" smoothing approach is used to model parts of a system about which little is known or which can not be easily quantified. Semi-mechanistic models usually also contain side-conditions which limit the complexity and smoothness of this component. The approach is to use generalized cross-validation to select the degree of smoothing.

Semi-mechanistic modeling is advantageous in the context of predictive fermentation biology because of the batch-to-batch variation that is typical in the growth media

or food product. For example, sunlight and high temperatures can result in variability in sugar content, acidity, oil levels and mineral content in fruits. The stage of vegetable and fruit maturity also has a direct impact on the chemical state of the food. Future research will focus on including within this framework uncertainties about the measured concentrations of defined components, the values of their dissociation constants and the relationship between temperature, ionic strength solvent properties and their dissociation constants. In the next Chapter, this semi-mechanistic approach is applied to modeling the pH changes during the fermentation of vegetable broth.

Chapter 5

Application of Semi-Mechanistic Buffer Modeling to Batch Fermentation

In the previous chapter, a semi-mechanistic method was developed for predicting pH in complex solutions containing both defined and undefined components. This method allows prediction of pH during a fermentation in a mechanistically realistic manner. The ability to predict pH reliably at any given time during a fermentation allows the common acid toxicity tolerance strategies exhibited by bacteria (recall Table 4.3) to be modeled mechanistically. It is the goal of this chapter to determine if such semi-mechanistic models of buffers can be embedded within a system of ODEs describing batch growth.

Specifically, the semi-mechanistic approach will be applied as part of a revision of the energy-based model for batch fermentation by *Lactococcus lactis* that was presented in Chapter 3. Except where noted, the data and fitting procedure used are

the same as that described for the "fixed temperature experiments" in Chapter 3. We also present a variation of the model useful for modeling the batch growth of the pathogen *Listeria moncytogenes*. Comparison of the model parameter estimates indicate some key differences in the growth kinetics between the two species.

5.1 Re-evaluation of the Energy-Based Dynamic Model of *Lactococcus lactis*

The system of equations 3.11 introduced in Chapter 3 was modified to incorporate the "parameters in the denominator" reparameterization of the Michaelis-Menten terms mentioned in Chapter 1. In addition, the inhibition term due to total lactic acid (L) was replaced by inhibition due to the intra-cellular concentration of lactic acid (IP). Depending on the medium, *Lactococcus lactis* maintains its intracellular pH nearly constant at a pH above 5.2 (Russell, 1992; Foucaud et al., 1995; Siegumfeldt et al., 2000). As a consequence, the energetic cost associated with the intracellular concentration of lactic acid can be modeled as proportional to the partition of lactic acid $[[HL] + [L^-]]\, P_{01}$ (recall Eqn. 4.18). A computationally convenient representation for inclusion of this functional form in a system of differential equations is obtained by rewriting $[[HL] + [L^-]]\, P_{01}$ as

$$\frac{k_1 + k_2 \frac{[L^-]}{L-[L^-]}}{1 + \frac{[L^-]}{L-[L^-]}} L \tag{5.1}$$

where $L - [L^-]$ has replaced $[HL]$. If the constant of proportionality relating the partition to the energetic cost is lumped together with k_1 and k_2, 2 new lumped pa-

rameters result. These 2 new parameters are called κ_1 and κ_2 respectively. The result is an expression relating the intracellular lactic acid concentration to the extracellular concentration.

$$IP = \frac{\kappa_1 + \kappa_2 \frac{[L^-]}{L-[L^-]}}{1 + \frac{[L^-]}{L-[L^-]}} L \tag{5.2}$$

Taking the derivative of the resulting expression with respect to time gives

$$\frac{dIP}{dt} = \kappa_1 \frac{dL}{dt} + (\kappa_2 - \kappa_1) \frac{d\,[L^-]}{dt} \tag{5.3}$$

where IP is the intracellular concentration of lactic acid (both anion and undissociated form), L is the extracellular concentration of lactic acid (both anion and undissociated form) and $[L^-]$ is the extracellular concentration of lactic acid anion. It is important to note, however, that $\frac{d[L^-]}{dt}$ is really just a function of $\frac{dL}{dt}$ and $\left[\mathrm{H}^+\right]$. This is so because $[L^-] = \frac{L\mathrm{K}_a}{\left[\mathrm{H}^+\right]+\mathrm{K}_a}$. Therefore, within the context of a numerical solution of the ODE system, $\frac{d[L^-]}{dt}$ is obtained through the evaluation of $\frac{\mathrm{K}_a}{\left[\mathrm{H}^+\right]+\mathrm{K}_a} \frac{d[L]}{dt}$.

$$\frac{dIP}{dt} = \kappa_1 \frac{dL}{dt} + (\kappa_2 - \kappa_1) \frac{\mathrm{K}_a}{\left[\mathrm{H}^+\right] + \mathrm{K}_a} \frac{d\,[L]}{dt} \tag{5.4}$$

The differential equations representing the changes in pH and the intracellular lactic acid concentration over time were added to the model. Remembering that the "parameters in the denominator" reparameterization is being used, the revised model

Table 5.1: State variables of the revised model[a].

Symbol	*Meaning*	*Units*	*Observed Range*
t	Time	h	0–256
T	Temperature	°C	10–30
N	Cell density	CFU mL^{-1}	$1.0x10^6$–$2.0x10^9$
S	Glucose in the medium	mM	0–35
M	Malic acid in the medium	mM	0–12
Q	Intracellular energy	fraction of initial "energy"	0–20
L	Lactic acid in the medium	mM	0–35
IP	Intracellular concentration of lactic acid	mM	0–600

[a] Q/N was normalized to 1 at the time of inoculation (t=0). Therefore, Q/N measures energy as a fraction of the energy quota that is typically present in an overnight culture.

is expressed by the following set of equations

$$
\begin{aligned}
\textbf{CFU:} \quad & \frac{dN}{dt} = \frac{q}{\Omega_1 + \nu_1 q} N - \delta_1 exp(-\delta_2 q) N \\
\textbf{Glucose:} \quad & \frac{dS}{dt} = -\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \\
\textbf{Malic Acid:} \quad & \frac{dM}{dt} = -\frac{1}{F_{T_1}} (\frac{q}{\Omega_3 + \nu_3 q}) NM \\
\textbf{Energy:} \quad & \frac{dQ}{dt} = \beta \left[\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \right] - \frac{IP}{F_{T2}} Q \\
& \quad -\gamma \left[\alpha \left(\frac{q}{k_{q_1} + q} \right) N \right] - \tau \frac{dT}{dt} I_T Q \\
\textbf{Lactic Acid:} \quad & \frac{dL}{dt} = 2 \left[\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \right] \\
& \quad + \left[\frac{1}{F_{T_1}} (\frac{q}{\Omega_3 + \nu_3 q}) NM \right] \\
\textbf{pH} \quad & \frac{d\mathrm{pH}}{dt} = \frac{1}{B_t} \frac{\partial C_b \ (dL/dt, dM/dt, \mathrm{pH})}{\partial t} \\
\textbf{IP:} \quad & \frac{dIP}{dt} = \kappa_1 \frac{dL}{dt} + (\kappa_2 - \kappa_1) \frac{\mathrm{K}_a}{[\mathrm{H}^+] + \mathrm{K}_a} \frac{d[L]}{dt}
\end{aligned}
\tag{5.5}
$$

where

$$
\begin{aligned}
& q = Q/N, \\
& \Omega_1 = \tfrac{k_{q1}}{\alpha}, \Omega_2 = \tfrac{k_{q2}}{\mu_1}, \\
& \Omega_3 = \tfrac{k_{q3}}{\mu_2}, \nu_1 = \tfrac{1}{\alpha} \\
& \nu_2 = \tfrac{1}{\mu_1}, \nu_3 = \tfrac{1}{\mu_2}, and \\
& F_{T*} = 1 + \left(\tfrac{T - T_{opt}}{K_{T*}} \right)^2
\end{aligned}
\tag{5.6}
$$

A description of the variables and their units is given in Table 5.1. B_t represents the

buffer capacity of the solution evaluated at time t and C_b $(dL/dt, dM/dt, \text{pH})$ is the moles of base added. As discussed in the previous chapter, B_t and C_b $(dL/dt, dM/dt, \text{pH})$ are calculated using local polynomial regression estimates, obtained from titration data, as well as the corresponding theoretical relations for lactic acid and malic acid.

5.1.1 Model Parameter Optimization

As it was pointed out above, the application of Eqn. 5.5 requires estimation of the ODE model parameters as well as a semi-mechanistic model of the growth medium's buffering. The parameterization of the buffering component of the model must be carried out prior to estimation of the model parameters. To make completely clear the order of operations, the following procedural outline summarizes the overall computational approach used in estimation of the parameters of the ODE system given by Eqn. 5.5.

Note that in the following, the titration and development of a semi-mechanistic model of the growth medium may be done in advance of or subsequent to the growth experiment(s). In particular, the estimation of the buffer model is carried out on titration data and is independent of the estimation of the parameters in the ODE system which is carried out on growth data. An important consequence of this is that the number of parameters needing to be estimated from the growth data does not increase when pH prediction is incorporated into the modeling framework.

1. Develop semi-mechanistic model of unfermented cucumber juice medium
   (a) Perform titration of medium to obtain observed C_b ($C_b{}^{\text{obs}}$) at various pH
   (b) Quantify malic acid and lactic acid in the medium by HPLC to obtain contributions to C_b from measurable buffer components ($C_b{}^{\sim}$)

(c) Subtract measurable contributions to C_b from malic and lactic acid ($C_b{}^* = C_b{}^{obs} - C_b{}^{\sim}$)

(d) Carry out local polynomial regression on $C_b{}^*$ using pH as the predictor variable

(e) Let the predicted titration ($\hat{C_b}$) be given by $\hat{C_b} = C_b{}^* + C_b{}^{\sim}$

2. Fit ODE system to the growth data by minimizing a WSS (see Eqn. 3.14)

(a) Solve ODE system obtaining predictions of CFU/mL, glucose, pH , energy, intracellular lactic acid and extracellular malic and lactic acid.

(b) Calculate WSS (recall Eqn. refeqn:wssenergy) using data for CFU/mL, glucose, pH , and extracellular malic and lactic acid

Note that in the above procedure, the pH is predicted using the ODE's current prediction of $\frac{dL}{dt}$ and $\frac{dM}{dt}$ in Eqn. 4.61. *The observed lactic and malic acid are not used in updating the pH prediction.* It is not advantageous to use the observed lactic and malic acid directly in solution of the ODE system since values at times requested by the numerical integration routine are not available. Interpolation, as was used in the temperature data could be used to obtain intermediate values but unlike the temperature data the HPLC measurements are subject to greater variability and may give less reliable estimates. *In short, using observed data to modify the pH would only serve to decouple the pH equation from the system and offers no computational advantage.*

After several preliminary numerical experiments it was determined that the lumped parameter κ_1 was always estimated near zero. This is biologically reasonable since κ_1 represents the flux of the molecular (uncharged) form of lactic acid across the mem-

brane at low extracellular pH . Since the pK_a of lactic acid is ≈ 3.86 the majority of the lactic acid is present in the anionic form in the medium. Therefore, κ_1 was eliminated from the model and the following ODE system was used.

$$
\begin{aligned}
\textbf{CFU:} \quad & \frac{dN}{dt} = \frac{q}{\Omega_1 + \nu_1 q} N - \delta_1 exp(-\delta_2 q) N \\
\textbf{Glucose:} \quad & \frac{dS}{dt} = -\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \\
\textbf{Malic Acid:} \quad & \frac{dM}{dt} = -\frac{1}{F_{T_1}} (\frac{q}{\Omega_3 + \nu_3 q}) NM \\
\textbf{Energy:} \quad & \frac{dQ}{dt} = \beta \left[\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \right] - \frac{IP}{F_{T2}} Q \\
& \qquad -\gamma \left[\alpha \left(\frac{q}{k_{q_1} + q} \right) N \right] - \tau \frac{dT}{dt} I_T Q \\
\textbf{Lactic Acid:} \quad & \frac{dL}{dt} = 2 \left[\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \right] \\
& \qquad + \left[\frac{1}{F_{T_1}} (\frac{q}{\Omega_3 + \nu_3 q}) NM \right] \\
\textbf{pH} \quad & \frac{d\mathrm{pH}}{dt} = \frac{1}{\beta_t} \frac{\partial C_b \ (dL/dt, dM/dt, \mathrm{pH})}{\partial t} \\
\textbf{IP:} \quad & \frac{dIP}{dt} = \kappa_2 \frac{\mathrm{K}_a}{[\mathrm{H}^+] + \mathrm{K}_a} \frac{d[L]}{dt}
\end{aligned}
\tag{5.7}
$$

where

$$
\begin{aligned}
& q = Q/N, \\
& \Omega_1 = \tfrac{k_{q1}}{\alpha}, \Omega_2 = \tfrac{k_{q2}}{\mu_1}, \\
& \Omega_3 = \tfrac{k_{q3}}{\mu_2}, \nu_1 = \tfrac{1}{\alpha} \\
& \nu_2 = \tfrac{1}{\mu_1}, \nu_3 = \tfrac{1}{\mu_2}, and \\
& F_{T*} = 1 + \left(\tfrac{T - T_{opt}}{K_{T*}}\right)^2
\end{aligned}
\tag{5.8}
$$

5.1.2 Results

The parameter estimates for the revised model Eqn. 5.7 are given in Table 5.2. Parameters were converted fro the "parameters in the denominator" parameterization back to their original parameterization. The estimate for α was obtained as $1/\nu_1$, the estimate for μ_1 was obtained as $1/\nu_1$, the estimate for kq_1 was obtained as $\alpha\Omega_1$ and the estimate for kq_2 was obtained as $\alpha\mu_1$. As can be observed in Fig. 5.1, the shift to low temperatures produces a lag-phase effect which is well-matched by model predictions. There is noticeable systematic lack of fit with regards to the cell density predictions. Conversely, the predictions for pH , lactic and malic acid are extremely good. Of particular note is the predictions for pH at 20°C (see Fig. 5.2) and 10°C (see Fig. 5.1). At these temperatures, the malolactic reaction proceeds relatively faster than the lactic acid fermentation because the net glucose consumption decreases during lag phase. There is a corresponding momentary increase in pH . Fig. 5.3 shows the calibrated model fit for the 30°C scenario.

Table 5.2: Model parameter descriptions and estimates for the revised model of *L. lactis*.

Parameter	*Description (unit)*	*Estimate*
α	Maximal bacterial growth rate (h^{-1})	1.52
μ_1	Maximal glucose consumption rate (CFU $\mathrm{mL}^{-1}\mathrm{h}^{-1}$)	3.64×10^{-8}
μ_2	Maximal malate conversion rate (CFU $\mathrm{mL}^{-1}\mathrm{h}^{-1}$)	2.18×10^{-7}
kq_1	Value of q at which cell growth rate equals $\alpha/2$ (unitless)	9.26
kq_2	Value of q at which glucose consumption rate equals $\mu_1/2$ (unitless)	2.23
kq_3	Value of q at which malate consumption rate equals $\mu_2/2$ (unitless)	2.59×10^{-7}
δ_1	Death rate when $q = \infty(h^{-1})$	2.22×10^{-2}
δ_2	Death rate sensitivity to changes in q (unitless)	1.87
β	Conversion rate of glucose into energy (mM^{-1})	7.54×10^{8}
γ	Energy required for cell division (CFU mL^{-1})	2.45
κ_2	Energy cost per 1 mM IP ($\mathrm{mM}^{-1}h^{-1}$)	4.24
KT_1	Sensitivity of metabolic processes to deviations from optimal temp (°C)	1.26
KT_2	Sensitivity of lactic acid inhibition to deviations from optimal temp (°C)	1.30
T_{opt}	Optimal temp (°C)	37 (assumed)
τ	Energy cost for transient temp adjustment (°C $^{-1}$)	1.37×10^{-6}

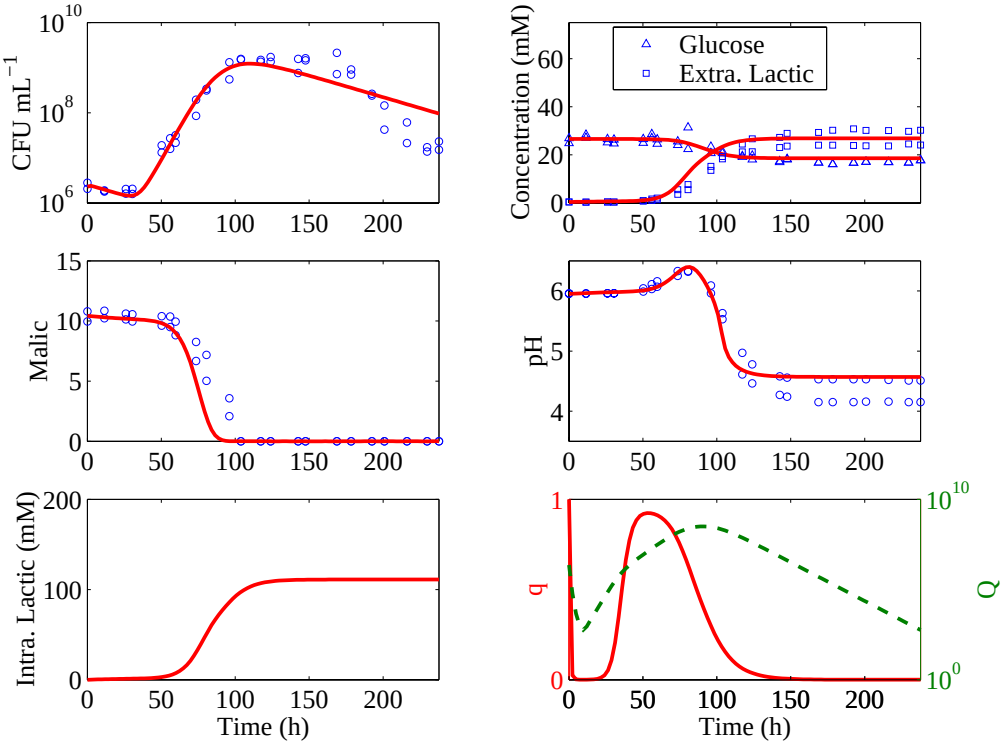

Figure 5.1: Predictions of the calibrated *L. lactis* model at 10°C . Note that the culture was incubated at 30°C prior to inoculation. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q

Temporal sensitivity analyses of the revised model revealed that a few parameters dominated the model predictions. Therefore the 5 most sensitive parameters were ranked in order of decreasing value. The 5 most influential parameters at 10°C in terms of predicting cell density were τ, δ_1 K_{T1}, γ and δ_2. At 30°C the 5 most sensitive model parameters were δ_1, K_{T2}, ν_1, K_{T1} and κ_2.

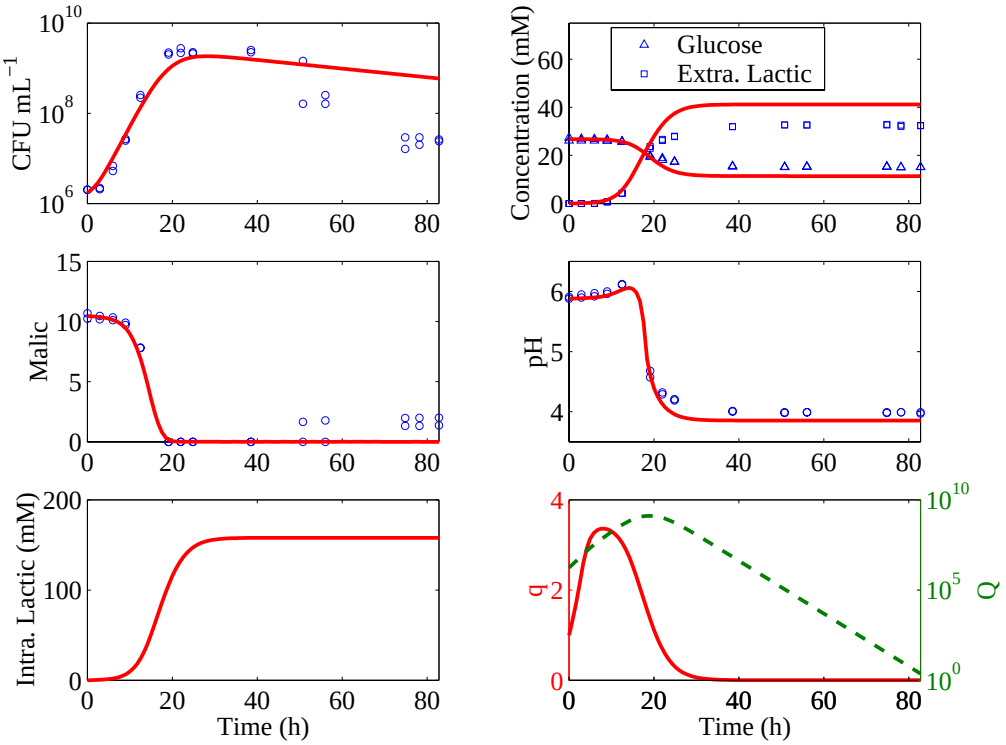

Figure 5.2: Predictions of the calibrated *L. lactis* model at 20°C . Note that the culture was incubated at 30°C prior to inoculation. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q

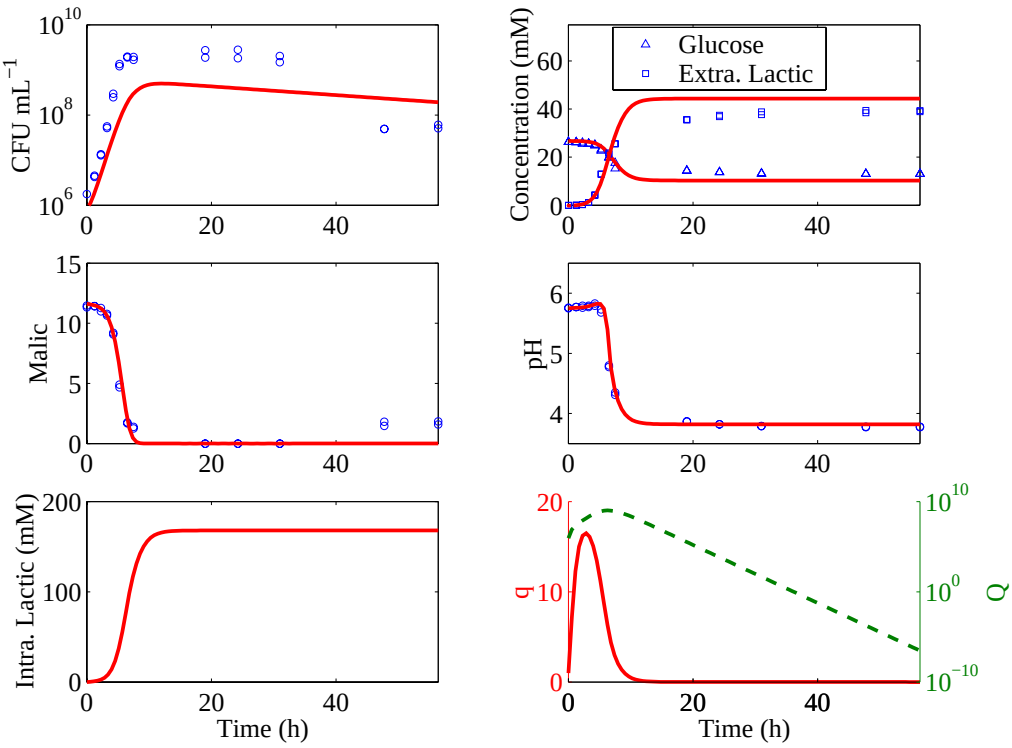

Figure 5.3: Predictions of the calibrated *L. lactis* model at 30°C . Note that the culture was incubated at 30°C prior to inoculation. There was no discernable lag phase. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q

5.1.3 Discussion

The goal of this research was to determine if by embedding a semi-mechanistic model of pH dynamics into a mechanistic model of batch fermentation reasonable estimates of pH and fermentation acids could be obtained. It is clear from the simulation results in Figs. 5.1–5.3 that semi-mechanistic method of modeling pH worked very well. In particular, fine details such as the increase in pH during lag phase due to the malolactic fermentation at 10°C were well–matched by the model predictions (see Fig. 5.1). Perhaps even more satisfying is that highly accurate predictions of not only pH but concentrations of lactic acid and glucose were obtained at 10, 20 and 30°C (see Figs. 5.1–5.3).

The optimal fit of the model did however leave something to be desired. The model did not predict well in all cases stationary phase cell densities or rates of cell death during death phase. The model does, however, predict the qualitative trends nicely. It is hoped that when more growth data are available the goodness of fit can be assessed more rigorously. Future efforts will be focused on refining the equations for cell growth so that the details of stationary phase and death phase are modeled in a more robust manner. While it is likely that the model does contain some mechanistic errors it is also possible that batch-to-batch variability in the medium is contributing to the variations in bacterial growth kinetics.

One possible source of model error is that the titration data used to parameterize the semi-mechanistic components of the model did not result from titrations of the same batch of media used in the growth studies. These growth experiments were conducted prior to the development of the semi-mechanistic approach so a titration of newly prepared CJ was used. While the model predictions for pH and fermentation

acids are promising, more accurate accounting of the buffer of the medium could be obtained by titration of the medium actually used in each run of the growth experiment.

In the next section we attempt to modify this model of *L. lactis* for the food pathogen *Listeria monocytogenes*.

5.2 Simplified Energy-Based Dynamic Model for *Listeria monocytogenes* Growth in Vegetable Broth

The goal of this section is to determine how well the semi-mechanistic method for predicting pH will work on another bacterial species. *Listeria monocytogenes* is a pathogen which has been implicated in spontaneous abortion of fetuses and can lead to fatality in the elderly, very young and immuno-compromised individuals (Wijtzes et al., 1993). There is interest, therefore, in determining environmental factors which inhibit the growth of this organism. *Listeria* has the potential to be a food-borne pathogen and therefore its growth at abuse temperatures should be studied. Growth experiments were conducted with *Listeria monocytogenes* (B0164) at 30°C and 10°C .

5.2.1 Materials and Methods

Cultures and Media

Listeria monocytogenes strain B0164 transformed with the erythromycin resistance gene pGKe (Breidt and Fleming, 1998) was obtained from the USDA culture collec-

tion. B0164 was grown on TSA broth (Difco Laboratories, Detroit, Mich.) containing 1.5% agar (Difco), 1% glucose (Sigma Chemical Co., St. Louis, Mo.) and 5 ug mL^{-1} erythromycin. Growth experiments were conducted in cucumber juice medium. Cucumbers were pureed and pressed to render raw juice, which was frozen until needed. The raw juice was thawed and then clarified by heating at 85°C for 5 min, followed by centrifugation at 12,000 rpm for 20 min and filter sterilized. The CJ medium was prepared by adding 600 mL juice to 400 mL dH2O. The diluted juice was supplemented with 2% NaCl and then filter-sterilized as described by Daeshel et al. (1984).

Experimental Set-up

Overnight cultures were prepared by growing B0164 in CJ at 30°C . The data consisted of 2 replicates (jars) from which 25 sample points over time (a total of 50 observations) were taken. Water-jacketed jars (Wheaton, Millville, N.J.) were filled with 200 mL of fresh CJ and inoculated at 1×10^6 CFU mL^{-1} of bacterial culture. Each flask was sealed with a silicone stopper that contained a sterile syringe sample port, through which an 18-gauge, 10 cm needle was passed. The growth medium was kept well-mixed by a magnetic stirrer. Compressed nitrogen was humidified by sparging through deionized water, filtered (0.2 um Millex-FG50 filter, Millipore Corp., Bedford, Mass.), and released into the headspace of the fermenter jars at a rate of 1.3 L h^{-1}. Temperature during the fermentation was controlled by a circulating water bath (NESlab RTE-211; NESlab, Portsmouth, N.H.). The temperature of the growth medium was monitored directly by sterile thermocouples inserted through the silicone stoppers and recorded by a microcomputer (OM-3000; Omega, Stamford, Conn.). Growth observations at 10 and 30°C included quantification of CFU mL^{-1}, glucose, malic acid, pH and lactic acid concentrations. Growth at a particular

temperature was monitored until all phases of growth had been observed.

Computational Details

The methods for solving the ODE's and updating the pH were the same as those mentioned in the previous section. However, as it will be discussed below, *L. monocytogenes* (B0164) did not ferment malic acid. The malic acid equation was therefore eliminated from the model. For the purposes of updating pH estimates the average measured malic acid level was used.

As discussed below, *L. monocytogenes* (B0164) was found to form biofilms at 10°C . Therefore only data from the 30°C run could be used. Since only growth at 30°C was used, the temperature adaptation terms from the original model were also eliminated (i.e. τ was set to zero and, F_{T*} was set to 1). The model resulting from

these simplifications is given by the following system of equations.

$$
\begin{aligned}
\textbf{CFU:} \quad & \frac{dN}{dt} = \frac{q}{\Omega_1 + \nu_1 q} N - \delta_1 exp(-\delta_2 q) N \\
\textbf{Glucose:} \quad & \frac{dS}{dt} = -\frac{q}{\Omega_2 + \nu_2 q} NS \\
\textbf{Energy:} \quad & \frac{dQ}{dt} = \beta \left[\left(\frac{q}{\Omega_2 + \nu_2 q} \right) NS \right] - (IP)Q \\
& \qquad -\gamma \left[\alpha \left(\frac{q}{k_{q_1} + q} \right) N \right] \\
\textbf{Lactic Acid:} \quad & \frac{dL}{dt} = 2 \left[\frac{1}{F_{T_1}} \left(\Omega_2 + \nu_2 q \right) NS \right] \\
& \qquad + \left[\frac{1}{F_{T_1}} \left(\frac{q}{\Omega_3 + \nu_3 q} \right) NM \right] \\
\textbf{pH:} \quad & \frac{d\mathrm{pH}}{dt} = \frac{1}{B_t} \frac{\partial C_b \ (dL/dt, dM/dt, \mathrm{pH})}{\partial t} \\
\textbf{IP:} \quad & \frac{dIP}{dt} = \kappa_1 \frac{dL}{dt} + (\kappa_2 - \kappa_1) \frac{d[L^-]}{dt}
\end{aligned}
\tag{5.9}
$$

where

$$
\begin{aligned}
& q = Q/N \\
& \Omega_1 = \tfrac{k_{q1}}{\alpha}, \Omega_2 = \tfrac{k_{q2}}{\mu_1} \\
& \nu_1 = \tfrac{1}{\alpha}, \nu_2 = \tfrac{1}{\mu_1} \\
& F_T = 1 + \left(\tfrac{T - T_{opt}}{K_{T^*}} \right)^2
\end{aligned}
\tag{5.10}
$$

Because of the simplifications of the model with regards to malolactic fermentation a slightly different WSS function was used to optimized the model parameter estimates. The following equation explicitly describes the formulation of the weighted sum of squared errors used in fitting the simplified model to the *L. monocytogenes* data. It is similar to that given by Eqn. 3.14. Recall that i indexes the temperatures,

j indexes the runs within a particular temperature and k indexes the observations within a run.

$$
\begin{aligned}
WSS \;=\; & \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{i,j}} \frac{(\log_{10}(\text{Obs. CFU/mL}_{i,j,k}) - \log_{10}(\text{Pred. CFU/mL}_{i,k}))^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{CFU/mL}})} \\
& + \frac{(\text{Obs. Lactic}_{t,j,p} - \text{Pred. Lactic}_{i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{Lactic}})} \\
& + \frac{(\text{Obs. Glucose}_{t,j,p} - \text{Pred. Glucose}_{i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{Glucose}})} \\
& + \frac{(\text{Obs. pH}_{\,i,j,k} - \text{Pred. pH}_{\,i,k})^2}{\sum_{j=1}^{b_i}(c_{i,j})(s_{\text{pH}})}
\end{aligned}
\tag{5.11}
$$

Note that malic acid is no longer included in the calculation of WSS.

5.2.2 Results

During the 10°C fermentations, the formation of thick mucoid strands in the fermentation flask was noted. Phase contrast microscopy revealed that *L. monocytogenes* (B0184) formed biofilms. The films displayed a clear delineation between an inner layer of sessile cells and an outer layer of motile cells (see Fig. 5.4). Since the formation of biofilms prevents the accurate determination of cell density, only fermentations at 30°C were considered.

The results of the 30°C fermentation are displayed in Figure 5.5. These data indicate that *L. monocytogenes* experiences a long stationary phase at 30°C and that its death rate subsequent to stationary phase is relatively small compared to its growth rate. Parameter estimates for the model 5.9 are given in Table 5.3. For ease of comparison, the estimates for the *L. monocytogenes* model are given along side those that were obtained for *L. lactis*.

Table 5.3: Comparison of model parameter estimates *L. lactis* and *L. monocytogenes*.

Parameter	*Description (unit)*	*L. lactis*	*L. mono.*
α	Maximal bacterial growth rate (h^{-1})	1.52	1.48
μ_1	Maximal glucose consumption rate (CFU $\text{mL}^{-1}h^{-1}$)	3.64×10^{-8}	3.77×10^{-11}
μ_2	Maximal malate conversion rate (CFU $\text{mL}^{-1}h^{-1}$)	2.18×10^{-7}	–
kq_1	Value of q at which cell growth rate equals $\alpha/2$ (unitless)	9.26	4.86
kq_2	Value of q at which glucose consumption rate equals $\mu_1/2$ (unitless)	2.23	3.41×10^{-2}
kq_3	Value of q at which malate consumption rate equals $\mu_2/2$ (unitless)	2.59×10^{-7}	–
δ_1	Death rate when $q = \infty(h^{-1})$	2.22×10^{-2}	6.60×10^{-2}
δ_2	Death rate sensitivity to changes in q (unitless)	1.87	354.2
β	Conversion rate of glucose into energy (mM^{-1})	7.54×10^8	3.29×10^9
γ	Energy required for cell division (CFU mL^{-1})	2.45	2.05
κ_2	Energy cost per 1 mM IP ($\text{mM}^{-1}h^{-1}$)	4.24	4.10
KT_1	Sensitivity of metabolic processes to deviations from optimal temp (°C)	1.26	–
KT_2	Sensitivity of lactic acid inhibition to deviations from optimal temp (°C)	1.30	–
T_{opt}	Optimal temp (°C)	37 (assumed)	–
τ	Energy cost for transient temp adjustment (°C $^{-1}$)	1.37×10^{-6}	–

Figure 5.4: Phase contrast image of *Listeria monocytogenes* biofilm. Cells which appear dark are motile while lighter cells are sessile.

5.2.3 Discussion

By modifying the model introduced in Chapter 3 to include the dynamic prediction of pH and acid stress only 2 additional model parameters were required. These were κ_2 and the initial condition for IP (assumed to be zero). This savings is largely due to the fact that the parameters relating pH to the buffering of the growth medium were derived independently using data from HPLC quantification and titration of the medium. This suggests that a great advantage of this approach over current methods for predicting pH is that the number of new parameters which must be estimated from the growth data is minimized.

Although more experimental data is needed, these preliminary results suggest that the modified model of batch growth is able to provide useful diagnostic information

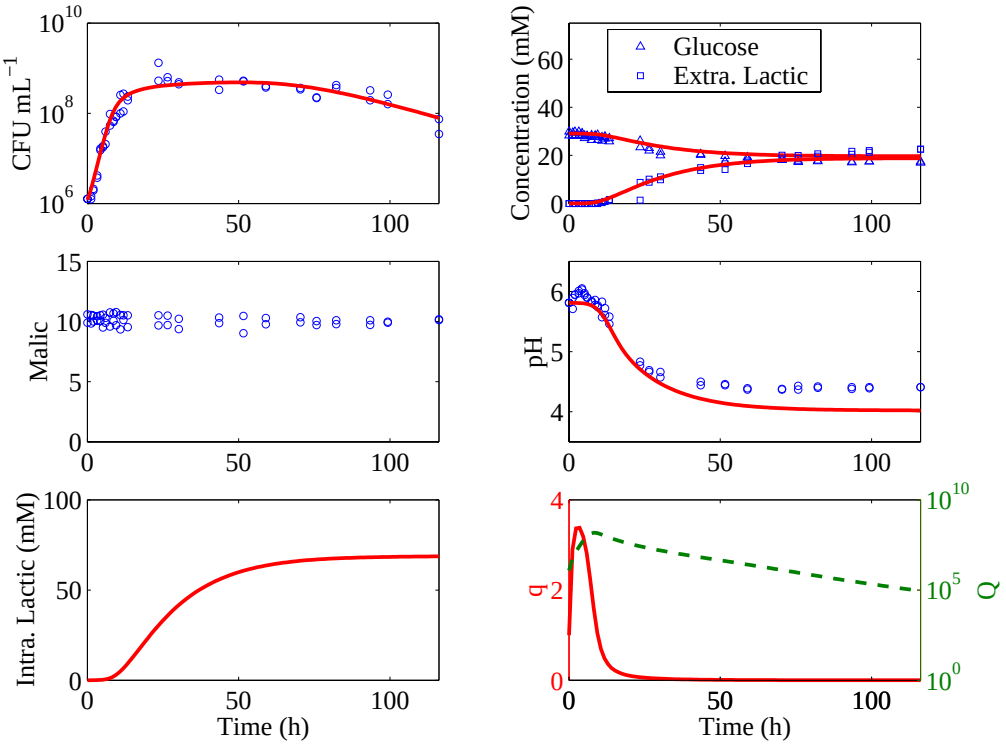

Figure 5.5: Predictions of the calibrated *L. monocytogenes* model at 30°C . Note that the culture was incubated at 30°C prior to inoculation. The symbols ∘ (number of CFU/milliliter or concentration of malic acid in millimolars), △ (concentration of glucose in millimolars), and □ (concentration of lactic acid in millimolars) represent experimental values, and the curves represent predicted values. Dashed line = Q

about the differences between the two bacteria. The maximum cell density observed for *Listeria* was approximately 3×10^8 CFU mL^{-1} and was lower than that for *Lactococcus*. The tendency of *L. monocytogenes* to form biofilms may explain why the cell counts were approximately one order of magnitude less than the maximum counts observed for *Lactococcus lactis* (LA221). If this is the case, and the cell densities were equal then a majority of *L. monocytogenes* cells (approx. 90%) would have been engaged in biofilms.

Comparison of Figures 5.3–5.5 reveals that the growth kinetics of *L. monocytogenes* were very different from those of *L. lactis*. In particular, the specific growth rate (α) of *L. monocytogenes* was about 3% less than that of *L. lactis*. The estimated specific rate of lactic acid influx (κ_2) was about 3% less in *L. monocytogenes*. Also, perhaps due to this increased acid resistance, the stationary phase of *L. monocytogenes* was much longer. The stationary phase of *L. lactis* was about 20 h while that for *L. monocytogenes* was nearly 100 h. Greater acid resistance in *L. monocytogenes* also corresponded with a higher estimate of δ_2 and a lower estimate γ than what were found for *L. lactis*.

An interesting feature of the parameter estimates given in Table 5.3 is that the maximum specific rate of glucose consumption (μ_1) is three orders of magnitude less than that for *L. lactis*. Visual inspection of the data, however, would indicate that glucose utilization by *L. monocytogenes* is approximately one order of magnitude less than that of *L. lactis*. This apparent discrepancy derives from the fact that the model for *L. lactis* under estimates the stationary phase cell density at 30 °C (see Fig. 5.3) by approximately 1 order of magnitude and that k_{q1} for *L. lactis* is nearly twice that estimated for *L. monocytogenes*. The model for *L. lactis* does not under-estimate the maximum cell densities at 20 and 10°C . This therefore is an example where lack-

of-fit of a model (or bias) can skew interpretations of model parameter comparisons. Although model parameter comparisons correctly indicate a smaller specific rate of glucose consumption by *L. monocytogenes* the difference is exaggerated by the underestimates of cell density by the *L. lactis* model at 30°C .

Comparisons between the parameter estimates for the *L. lactis* and *L. monocytogenes* models indicate that *L. monocytogenes* survives longer during stationary phase by reducing the net accumulation of lactic acid and also reducing the energy required to grow. It would be interesting to compare experimentally measured intracellular lactic acid concentrations to the model predictions of IP. Future work will be focused in this area as well as assessing model performance for bacteria with different acid tolerance strategies and different fermentative pathways.

Chapter 6

Conclusions

In this thesis, new techniques for predicting the growth of microbial populations in variable environments have been developed. In Chapter 2, currently available mathematical techniques for the construction of mathematical models and methods for estimating the parameters of such models from experimental data were reviewed. The "parameters in denominator" method appears to produce models with more identifiable parameters (less correlated) and to thereby improve convergence. Because the re-parameterization is easy to implement its use should be encouraged. This is especially true in mechanistic models where data over the entire functional range may not be available.

The models of batch fermentation of *Lactococcus lactis* and *Listeria monocytogenes* provide predictions on several variables each on perhaps different scales and with different error structures. The weighted least squares approach used to estimate the parameters in these models assumes independence and the cell densities were log transformed to allow all variables to follow a normal distribution more closely. While more complicated (realistic) approaches to modeling the error structure of this system

of differential equations should be investigated this is beyond the scope of the present research. Furthermore, such investigations will require substantially more data in order to estimate the many parameters (variance components) that would be required. In order to combine the model fits to data on each of the variables and across different treatments, a weighted sum of squares which uses the standard deviation multiplied by the number of observations for a given treatment as a weighting factor was recommended. To minimize this weighted sum of squares, a genetic algorithm was found to be superior to direct search gradient based methods. It is the discrete adaptiveness of ODE solvers which necessitates the use the of a genetic algorithm. Fixed step ODE solvers are not recommended since the error of integration is not controlled.

In Chapter 3, a mathematical model for the batch growth of lactic acid bacteria in variable temperature fermentations was developed and fit to experimental data. This model was limited because the way in which acid toxicity was modeled did not coincide with current biological theory. More useful models, however, require the ability to predict pH continuously during a fermentation and in response to the accumulation of fermentation acids. Chapter 4 dealt with this issue and a semi-mechanistic method was developed. In this method, local polynomial regression is used to model the components of the media which can not be measured while measured components whose dissociation constants are known are modeled using traditional acid-base theory. An important advantage of this method is the ease with which the dissociation equilibria of the acids and bases can be adjusted for ionic strength and temperature.

The semi-mechanistic procedure also lends itself nicely to testing for the importance of buffer components with respect to the overall buffering of the medium. One of the limitations of this inference procedure (outlined in Section 4.5.2) is that the precision of estimates of the concentrations and dissociation constants of defined or

measurable components is not addressed. Future research will be directed at incorporating *a priori* estimates of measurement variability of such components into the inference procedures.

The ability to predict pH continuously in complex growth media allows greater flexibility in modeling acid toxicity. In particular, functional forms for some of the common acid toxicity responses observed in bacteria were developed. Finally in Chapter 5 the semi-mechanistic approach was applied in conjunction with a model of intracellular lactic acid accumulation to a more advanced model of batch fermentation. This model gave good predictions of the extracellular pH and fermentation acids. Furthermore, the model was able to predict the intracellular concentration of lactic acid. Future research will be directed at assessing the accuracy of these theoretical predictions.

Lastly, a model for the batch growth of *Listeria monocytogenes* was developed based on the more advanced model of *L. lactis* just mentioned. Similar quality of predictions were obtained for this system even though the growth kinetics of *L. monocytogenes* were qualitatively different from *L. lactis*. Future research will be directed at modeling the competitive growth of these 2 species as well as other species which possess different acid tolerances and fermentative pathways.

Bibliography

Arnold, C. N., J. McElhanon, A. Lee, R. Leonhart, and D. A. Siegele (2001). Global analysis of *Escherichia coli* gene expression during the acetate-induced acid tolerance response. *J. Bacteriol.* **183**(7), 2178–2186.

Baranyi, J. (1998). Comparison of stochastic and deterministic concepts of bacterial lag. *J. Theor. Biol.* **192**, 403–408.

Baranyi, J. and T. A. Roberts (1995). Mathematics of predictive food microbiology. *Int. J. Food. Microbiol.* **26**, 199–218.

Baranyi, J., T. Robinson, A. Kaloti, and B. Mackey (1995). Predicting growth of *Brochothrix thermosphacta* at changing temperature. *Int. J. Food. Microbiol.* **27**, 61–75.

Bates, J. H., D. A. Bates, and W. MacKillop (1987). On the difficulties of fitting the double Michaelis-Menton equation to kinetic data. *J. Theor. Biol.* **125**, 237–241.

Bayrock, D. and W. M. Ingledew (2001). Changes in steady state on introduction of a *Lactobacillus* contaminant to a continuous culture ethanol fermentation. *J. Ind. Microbiol. Biot.* **27**, 39–45.

Benthin, S. and J. Villadsen (1995). Different inhibition of *Lactobacillus delbrukii* subsp. *vulgaricus* by d- and l-lactic acid: effects on lag phase, growth rate and cell yield. *J. Appl. Bacteriol.* **78**, 647–654.

Bentzen, E. and W. D. Taylor (1991). Estimating michaelis-menten parameters and lake water phosphate by the rigler bioassay: Importance of fitting technique, plankton size, and substrate range. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 73–83.

Bovill, R., J. Bew, N. Cook, M. D'Agostino, N. Wilkinson, and J. Baranyi (2000). Predictions of growth for *Listeria monocytogenes* and *Salmonella* during fluctuating temperature. *Int. J. Food. Microbiol.* **59**, 157–165.

Boyce, W. E. and R. C. DiPrima (1988). Calculus. John Wiley and Sons.

Breidt, F. and H. P. Fleming (1998). Modeling of the competitive growth of *Listeria monocytogenes* and *Lactococcus lactis* in vegetable broth. *Appl. Environ. Microb.* **64**(9), 3159–3165.

Brockelhurst, T., G. Mitchell, Y. Ridge, R. Seale, and A. Smith (1995). The effect of transient temperatures on the growth of *Salmonella typhimurium* lt2 in gelatin gel. *Int. J. Food. Microbiol.* **27**, 45–60.

Buchanan, R. L. and M. L. Cygnarowicz (1990). A mathematical approach toward defining and calculating the duration of the lag phase. *Food. Microbiol.* **7**, 237–240.

Buckingham, E. (1914). On physically similar systems. *Phys. Rev.* **4**, 345–376.

Burmaster, D. E. (1979). The continuous culture of phytoplankton: Mathematical equivalence among three steady-state models. *Am. Nat.* **113**(1), 123–134.

Butler, J. N. and D. R. Cogley (1998). Ionic Equilibrium: Solubility and pH Calculations. John Wiley and Sons.

Casella, G. and R. L. Berger (1990). Statistical Inference. Wadsworth and Brooks/Cole.

Chibib, N.-E. and J.-L. Tholozan (1999). Effect of rapid cooling and acidic ph on cellular homeostasis of *Pectinatus frisinggensis*, a strictly anaerobic beer-spoilage bacterium. *Int. J. Food. Microbiol.* **48**, 191–202.

Cleveland, W. S., S. J. Devlin, and E. Grosse (1988). Regression by local fitting: Methods, properties, and computational algorithms. *J. Econometrics.* **37**, 87–114.

Cohen, J. E. (1995). Population growth and earth's human carrying capacity. *Science.* **269**(5222), 341–346.

Cornish-Bowden, A. (1995). Fundamentals of Enzyme Kinetics. Portland Press Ltd.

Daeshel, M., R. McFeeters, H. Fleming, T. Klaenhammer, and R. Sanozky (1984). Mutation and selection of *Lactobacillus plantarum* strains that do not produce carbon dioxide from malate. *Appl. Environ. Microb.* **47**, 419–420.

Diez-Gonzalez, F. and J. B. Russel (1997). The ability of *Escherichia coli* o157:h7 to decrease its intracellular ph and resist the toxicity of acetic acid. *Microbiology* **143**, 1175–1180.

Droop, M. R. (1973). Some thoughts on nutrient limitation in algae. *J. Phycol.* **9**, 264–272.

Dufrenne, J., E. Delfgou, W. Ritmeester, and S. Notermans (1997). The effect of previous growth conditions on the lag phase time of some foodborne pathogenic micro-organisms. *Int. J. Food. Microbiol.* **34**, 89–94.

Edelstein-Keshet, L. (1988). Mathematical Models in Biology. McGraw-Hill Inc.

Fan, J. and I. Gijbels (1996). Local polynomial modelling and its applications. Chapman and Hall.

Foucaud, C., M. Herve, J. Neumann, and D. Hemme (1995). Glucose metabolism and internal ph of *Lactococcus lactis* subsp. *lactis* cells utilizing nmr spectroscopy. *Lett. Appl. Microbiol.* **21**, 10–13.

Fu, B., P. Taoukis, and T. Labuza (1991). Predictive microbiology for monitoring spoilage of dairy products with time-dependent integrators. *J. Food. Sci.* **56**(5), 1209–1215.

Gibson, A. M., N. Bratchell, and T. A. Roberts (1987). The effect of sodium chloride and temperature on the rate and extent of *Clostridium botulinum* type a in pasteurized pork slurry. *J. Appl. Bacteriol.* **62**, 479–490.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determing the value of the life contingencies. *Philos. T. Roy. Soc. B* **115**, 513–583.

Goovaerts, P., J. Semrau, and S. Lontoh (2001). Monte carlo analysis of uncertainty attached to microbial pollutant degradation rates. *Environ. Sci. Technol.* **35**, 3924–3930.

Haefner, J. W. (1996). Modeling Biological Systems: Principles and Applications. Chapman and Hall.

Hairer, E., S. Nórsett, and G. Wanner (1987). Solving Ordinary Differential Equations. Springer-Verlag.

Horiuchi, J., S. Kikuchi, M. Kobayashi, T. Kanno, and T. Shimizu (2001). Modeling of ph response in continuous anaerobic acidogenesis by an artificial neural network. *Biochemical Engineering Journal* **9**, 199–204.

Jay, J. M. (1992). Modern Food Microbiology. Chapman and Hall, 4 edn.

Jetton, M. S., A. J. Stams, and A. J. Zehnder (1991). Adenine nucleotide content and energy charge of *Methanothrix soehngenii* during acetate degradation. *Fems. Microbiol. Lett.* **84**, 313–318.

Johansen, A. and D. Sornette (2001). Finite-time singularity in the dynamics of the world population, economic and financial indices. *Physica A* **294**, 465–502.

Jost, C. and S. P. Ellner (2000). Testing for predator dependence in predator-prey dynamics: a non-parametric approach. *P. Roy. Soc. Lond. B Bio.* **267**.

Keuhl, R. (2000). Design of experiments: statistical principles of research design and analysis. Duxbury, 2 edn.

Kooijman, S. (2000). Dynamic energy and mass budgets in biological systems. Cambridge University Press, 2 edn.

Kooijman, S. (2001). Quantitative aspects of metabolic organization: a discussion of concepts. *P. Roy. Soc. Lond. B Bio.* **356**, 331–349.

Levenspiel, O. (1980). The monod equation: A revisit and a generalization to product inhibition situations. *Biotechnol. Bioeng.* **22**, 1671–87.

Loubière, P., P. Salou, M.-J. Leroy, N. D. Lindley, and A. Pareilleux (1992). Electrogenic malate uptake and improved growth energetics of the malolactic bacterium *Leuconostoc oenos* grown on glucose-malate mixtures. *J. Bacteriol.* **174**(16), 5302–5308.

MacDonald, N. (1982). Time delays in chemostat models. In M. J. Bazin (Ed.), Microbial Population Dynamics, CRC Press, Inc., vol. CRC Series in Mathematical Models in Microbiology, chap. 2.

Malakar, P., D. Martens, M. Zwietering, C. Béal, and K. van't Reit (1999). Modelling the inter-actions between *Lactobacillus curvatus* and *Enterobacter cloacae*: Ii. mixed cultures and shelf life predictions. *Int. J. Food. Microbiol.* **51**(1), 67–79.

Malthus, T. (1976). An essay on the principle of population. In P. Appleman (Ed.), An essay on the principle of population: text, source and background criticism.

Martens, D., C. Béal, P. Malakar, M. Zwietering, and K. van't Reit (1999). Modelling the inter-actions between lactobacillus curvatus and enterobacter cloacae: I. individual growth kinetics. *Int. J. Food. Microbiol.* **51**(1), 53–65.

McDonald, K. and D.-W. Sun (1999). Predictive food microbiology for the meat industry: A review. *Int. J. Food. Microbiol.* **52**, 1–27.

McDonald, L., H. Fleming, and H. Hassan (1990). Acid tolerance of *Leuconostoc mesenteroides* and *Lactobacillus plantarum*. *Appl. Environ. Microb.* **56**(7), 2120–2124.

McFeeters, R. (1993). Single-injection hplc analysis of acids, sugars, and alcohols in cucumber fermentations. *J. Agr. Food. Chem.* **41**, 1439–1443.

McFeeters, R., H. Fleming, and R. Thompson (1982). Malic acid as a sourcve of carbon dioxide in cucumber juice fermentations. *J. Food. Sci.* **47**, 1862–1865.

McKellar, R. and K. Knight (2000). A combined discrete-continuous model describing the lag phase of *Listeria monocytogenes*. *Int. J. Food. Microbiol.* **54**, 171–180.

McMeekin, T., J. Olley, T. Ross, and D. Ratkowsky (1993). Predictive Microbiology, vol. 5: Innovation in Microbiology Series. Research Studies Press Ltd.

Mercade, M., N. Lindley, and P. Loubière (2000). Metabolism of *Lactococcus lactic* subsp. *cremoris* mg 1363 in acid stress conditions. *Int. J. Food. Microbiol.* **55**, 161–165.

Metge, D. W., M. H. Brooks, R. L. Smith, and R. W. Harvey (1993). Effect of treated-sewage contamination upon bacterial energy charge, adenine nucleotides, and DNA content in a sandy aquifer on cape cod. *Appl. Environ. Microb.* **59**(7), 2304–2310.

Michaelis, L. and M. L. Menten (1913). Die kinetik der invertinwirkung. *Biochem. Z* **49**, 333–369.

M'Kendrick, A. G. and M. K. Pai (1910). The rate of multiplication of micro-organisms. *P. Roy. Soc. Edinb.* **31**, 649–655.

Monod, J. (1942). Recherches sur la croissance des cultures bactériennes. Herman et Cie.

Monod, J. (1949). The growth of bacterial cultures. *Annu. Rev. Microbio.* **3**, 371–394.

Omlin, M., R. Brun, and P. Reichert (2001). Biogeochemical model of lake z urich: sensitivity, identifiability and uncertainty analysis. *Ecological Modeling* **141**, 105–123.

Pearl, R. (1925). The Biology of Population Growth. Alfred A. Knopf.

Pearl, R. and L. J. Reed (1920). On the rate of growth of the population of the United States since 1970 and ites mathematical representation. *P. Natl. Acad. Sci. USA* **6**(6), 275–288.

Phillips, G. D. A. (2000). An alternative approach to obtaining nagar-type moment approximations in simultaneous equation models. *J. Econometrics.* **97**, 345–364.

Pitzer, K. S. (1991). Activity Coefficients in Electrolyte Solutions. CRC Press.

Poolman, B., D. Molenaar, E. J. Smid, T. Ubbink, T. Abee, P. P. Renault, and W. N. Konings (1991). Malolactic fermentation: electrogenic malate uptake and malate/lactate antiport generate metabolic energy. *J. Bacteriol.* **173**(19), 6030–6037.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press.

Pruitt, K. and D. Kamau (1993). Mathematical models of bacterial growth, inhibition and death under combined stress conditions. *J. Ind. Microbiol.* **12**, 221–231.

Ramkrishna, D., A. Fredrickson, and H. Tsuchiya (1967). Dynamics of microbial propagation: Models considering inhibitors and variable cell composition. *Biotechnol. Bioeng.* **9**, 129–170.

Ratkowsky, D. (1985). A statistically suitable general formulation for modelling catalytic chemical reactions. *Chem. Eng. Sci.* **40**, 1623–1628.

Ratkowsky, D. A. (1986). A suitable parameterization of the Michaelis-Menton enzyme reaction. *Biochem. J.* **240**, 357–360.

Ross, T. (1996). Indices for performance evaluation of predictive models in food microbiology. *J. Appl. Bacteriol.* **81**, 501–508.

Russell, J. (1992). Another explanation for the toxicity of fermentation acids at low ph: anion accumulation versus uncoupling. *J. Appl. Bacteriol.* **73**, 363–370.

Samson, E., G. Lemaire, J. Marchand, and J. Beaudoin (1999). Modeling chemical activity effects in strong ionic solutions. *Comp. Mater. Sci.* **15**, 285–294.

Schaffner, D. (1998). Predictive food microbiology gedanken experiment:why microbial growth data require a transformation? *Food. Microbiol.* **15**, 185–189.

Shampine, L. F. and S. Thompson (2001). Solving ddes in matlab. *Appl. Numer. Math.* **37**(4), 441–458.

Siegumfeldt, H., K. B. Rechinger, and M. Jakobsen (2000). Dynamic changes of intracellular ph in individual lactic acid bacterium cells in response to a rapid drop in extracellular ph. *Appl. Environ. Microb.* **66**(6), 2330–2335.

Skinner, G., J. Larkin, and E. Rodehamel (1994). Mathematical modeling of microbial growth: a review. *J. Food. Safety.* **14**, 175–217.

Storn, R. (1996). On the usage of differential evolution for function optimization. *North American Fuzzy Information Processing Society (NAFIPS)* pp. 519–523.

Swartzman, G. and S. Kaluzny (1987). Ecological Simulation Primer. MacMillan Publishing.

van Beek, S. and F. G. Preist (2002). Evolution of the lactic acid bacterial community during malt whisky fermentation: a polyphasic study. *Appl. Environ. Microb.* **68**(1), 297–305.

Van Impe, J. F., B. M. Nicolaï, T. Martens, J. D. Baerdemaeker, and J. Vandewalle (1992). Dynamic mathematical model to predict microbial growth and inactivation during food processing. *Appl. Environ. Microb.* **58**(9), 2901–2909.

Van Impe, J. F., B. M. Nicolaï, M. Schellekens, T. Martens, and J. D. Baerdemaeker (1995). Predictive microbiology in a dynamic environment: a system theory approach. *Int. J. Food. Microbiol.* **25**, 227–249.

Van Slyke, D. (1922). On the measurement of buffer values and on the relationship of buffer value to the dissociation constant of the buffer and the concentration and reaction of the buffer solution. *J. Am. Chem. Soc.* **52**, 525–570.

Van Vooren, L., M. Van De Steene, J. Ottoy, and P. Vanrolleghem (2001). Automatic buffer capacity model building for the purpose of water quality monitoring. *Water. Sci. Technol.* **43**(7), 105–113.

Versari, A., G. P. Parpinello, and M. Cattaneo (1999). *Leconostoc oenos* and malolactic fermentation in wine: a review. *J. Ind. Microbiol. Biot.* **23**, 447–455.

Wijtzes, T., P. McClure, M. Zwietering, and T. Roberts (1993). Modelling bacterial growth of *Listeria monocytogenes* as a function of water activity, ph and temperature. *Int. J. Food. Microbiol.* **18**, 139–149.

Wood, S. N. (2001). Partially specified ecological models. *Ecol. Monogr.* **71**(1), 1–25.