CONTINUOUS BIOCHEMICAL PROCESSING:

INVESTIGATING NOVEL STRATEGIES TO PRODUCE SUSTAINABLE FUELS

AND PHARMACEUTICALS

A Dissertation

by

JONATHAN PATRICK RAFTERY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	M. Nazmul Karim
Committee Members,	Arul Jayaraman
	Mahmoud El-Halwagi
	Wolfgang Bangerth
Head of Department,	M. Nazmul Karim

August 2017

Major Subject: Chemical Engineering

ABSTRACT

Biochemical processing methods have been targeted as one of the potential renewable strategies for producing commodities currently dominated by the petrochemical industry. To design biochemical systems with the ability to compete with petrochemical facilities, inroads are needed to transition from traditional batch methods to continuous methods. Recent advancements in the areas of process systems and biochemical engineering have provided the tools necessary to study and design these continuous biochemical systems to maximize productivity and substrate utilization while reducing capital and operating costs.

The first goal of this thesis is to propose a novel strategy for the continuous biochemical production of pharmaceuticals. The structural complexity of most pharmaceutical compounds makes chemical synthesis a difficult option, facilitating the need for their biological production. To this end, a continuous, multi-feed bioreactor system composed of multiple independently controlled feeds for substrate(s) and media is proposed to freely manipulate the bioreactor dilution rate and substrate concentrations. The optimal feed flow rates are determined through the solution to an optimal control problem where the kinetic models describing the time-variant system states are used as constraints. This new bioreactor paradigm is exemplified through the batch and continuous cultivation of β-carotene, a representative product of the mevalonate pathway, using *Saccharomyces cerevisiae* strain mutant SM14.

The second goal of this thesis is to design continuous, biochemical processes capable of economically producing alternative liquid fuels. The large-scale, continuous

production of ethanol via consolidated bioprocessing (CBP) is examined. Optimal process topologies for the CBP technology selected from a superstructure considering multiple biomass feeds, chosen from those available across the United States, and multiple prospective pretreatment technologies. Similarly, the production of butanol via acetone-butanol-ethanol (ABE) fermentation is explored using process intensification to improve process productivity and profitability. To overcome the inhibitory nature of the butanol product, the multi-feed bioreactor paradigm developed for pharmaceutical production is utilized with *in situ* gas stripping to simultaneously provide dilution effects and selectively remove the volatile ABE components. Optimal control and process synthesis techniques are utilized to determine the benefits of gas stripping and design a butanol production process guaranteed to be profitable.

ACKNOWLEDGEMENTS

While this work is a result of many years of challenging work, much of this thesis would not be possible without the support of so many people. I would like to personally thank many of the people responsible for holding me up, keeping me sane, and pushing me to complete this momentous milestone in my life.

First, I would like to thank those with direct contributions to my development as a researcher and as an academic. Chief among those is my advisor, Dr. Karim, without whom I would not have had the ability to achieve all I have. When I first came to Texas A&M in 2012 I was unsure of where my life would be heading, but through some stroke of luck I was able to land with an advisor who pushed me to my limits and made me consider quitting more than once when worried I would not be able to meet expectations, but also provided the support and guidance needed to overcome those brief moments of weakness and succeed. I would also like to thank the members of my committee, Dr. Jayaraman, Dr. El-Halwagi, and Dr. Bangerth, who so graciously provided their advice and guidance throughout and put me in position to be ready for the trials and tribulations of an academic. I would also like to thank my funding sources, including the National Science Foundation, the TIAS Foundation, and the Graduate Teaching Fellowship program for providing me with the means to survive and thrive.

I would also like to thank the people that have been there in the trenches with me over the course of this process. Members of the Karim group, including Melanie, Chiru, Tejasvi, Laura, Carolina, and Xinghua, never failed to offer support during the low times and made for excellent companions during the good times. Other members of the fifth

floor of the Jack E. Brown building, including Michelle, Quint, Katie, Nandita, Dorsa, Martin, Ravi, and Daniel, built a community and network of close friendships that will continue well beyond the time when we are no longer at TAMU. I also need to give a special thanks to Greg and Amanda, two very special people from my years at UMass, that have not only willingly accepted but embraced the annoyance of listening to my constantly complain about simulations that didn't work, long weekends spent in the lab, and most of all the terrible, ceaseless summers in Texas. The network of friends that I have been miraculously been able to foster has been a buoy for me throughout my time at Texas A&M, and I know each one of your will go on to do remarkable things.

Last, but most certainly not least, I would like to thank my family who have supported me not only through this adventure but since the very beginning. Mom, the love and passion you have put into teaching has inspired me to embark on my own journey in the profession, that more intelligently at the collegiate where I do not have to deal with parent. Dad, the hardworking and selfless mentality you continuously displayed growing up has been one of the cornerstones of my approach to life, knowing that whatever needs to get done gets done both with respect to work and interpersonal relationships. Stephen, thank you for providing that spark of competitiveness, the drive to be the best I can be even if it may have come at your expense more than once. I would also like to thank my grandmother, Diane, for not only being willing to put up with me calling as I drove to work day after day, but for also showing unconditional support both emotionally and monetarily. Without your support throughout the years I would not be the man I am today, never mind the researcher.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor M. Nazmul Karim, Professor Arul Jayaraman, and Professor Mahmoud El-Halwagi of the Department of Chemical Engineering and Professor Wolfgang Bangerth of the Department of Mathematics.

The experiments used to gather data necessary for the analyses depicted in Chapter IV were conducted in part by Melanie DeSessa, M. Carolina Ordoñez, and Tejasvi Jaladi of the Department of Chemical Engineering.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a HEEP fellowship from the Texas A&M Institute for Advances Studies (TIAS) under the direction of Dr. Christodoulos Floudas and the Graduate Teaching Fellowship from the College of Engineering at Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

1.1. Motivation

Petroleum based products have become a key component of daily life across the globe since the Industrial Revolution. Electricity for commercial and residential use has relied extensively on coal and natural gas resources, many commodity chemicals and pharmaceuticals are produced from petroleum-derived feedstocks, and the transportation industry depends on gasoline, diesel, and jet fuel to successfully passage both people and products globally. However, the use of fossil fuel based energy sources has been linked to an increase in global warming, leading to a push for cleaner, more sustainable sources for energy and chemical feedstocks. For these renewable methods to supplant the ingrained dependence on fossil-based resources, it is important for engineers to develop processes that can reliably and inexpensively produce these necessary products at the levels that meet current demands.

While wind, solar, nuclear, and other renewable sources have been examined as renewable methods for sustainable electricity generation, these methods do not currently meet the requirements to supplant the fuels and chemical precursors provided by petroleum. Biochemical conversion strategies, however, have been identified as a potential approach to fill this niche. Inroads in biochemical engineering over the last few decades, specifically in the areas of genetic engineering, have provided a powerful toolkit to produce many chemicals through the genetic modification of cellular organisms. Equipped with the ability to harness and adapt the powers of cellular

metabolism, processes are being developed for the commercialization of many biologically-derived products.

To this point, processes for biochemically derived fuels and chemicals have struggled to achieve the reliability and economically attractive heights of the petrochemical industry. The traditional batch methods used in biochemical production are challenged by batch-to-batch variability and high operating and capital costs which have led to soaring, uncompetitive prices for renewable fuels such as ethanol and butanol. Similarly, the same batch processing methods have limited supplies in the biologic and nutraceutical industry, resulting in unaffordable prices attributed to the inability to meet increasing demand.

One way to cut costs for consumers is to cut costs for suppliers, which can be achieved by transitioning from an unreliable batch-based processing method to the continuous processing method found in traditional chemical engineering applications. However, these new continuous processes need to be able to achieve productivities that rival or surpass those of batch systems in a way that is reliable and not subject to the same variability found in those processes. Process systems engineering, composed of strategies for process modeling, process design, and process control, offer the techniques necessary to investigate the effects of continuous production in biochemical processing. These continuous biochemical strategies must be able to deal with multiple types of biological products, i.e. intercellular and extracellular, and provide economic incentive to stimulate transition in the industry. By coupling the state-of-the-art biochemical methods with process systems engineering, inroads can be made on developing

dependable, sustainable processes that can meet demand and continue to deliver the quality of life to society that was once achieved through the petrochemical industry.

1.2. Goals and outline of thesis

The development of novel continuous methods for the biochemical production of alternative fuels and pharmaceuticals is the theme of this thesis. The specific achievements of this work can be categorized as follows: *(1) develop models describing the non-trivial kinetics of products in the mevalonate pathway (beta-carotene, taxadiene, bisabolene, etc.) in genetically modified organisms, (2) design a novel continuous processing strategy that improves productivity and profitability for biological processes using a multi-feed inlet configuration and knowledge of reaction kinetics, (3) demonstrate the benefits of continuous production of pharmaceuticals using the kinetic models developed in Goal 1, (4) develop a process synthesis framework for determining the optimal strategy for cellulosic ethanol production, and (5) examine the efficacy of the continuous processing strategy that combines the continuous bioprocessing paradigm in Goal 2 and the process synthesis framework in Goal 4 for biobutanol production*.

In Chapter II, fundamentals of process systems engineering and mathematical programming are introduced. Algorithms for solving unconstrained optimization problems are introduced to lay the groundwork for algorithms that consider optimization problems with equality and inequality constraints. Methods for linear and nonlinear optimization in the presence of integer variables are also considered. The review of optimization strategies concludes with a discussion of global optimization methods.

Applications of mathematical programming in the context of chemical engineering are discussed, specifically in the areas of process modeling, process synthesis, and process control.

Chapter III of this thesis provides a discussion of the growth and prominence of biotechnology. Advancements in genetic engineering that have led to development of high-productivity and high-tolerance strains will be examined. Special attention is placed on the growth of the pharmaceutical and alternative energy industries, where batch production strategies are traditionally utilized. The benefits and challenges of using continuous processing strategies in bio-based industries is also proposed.

In Chapter IV, the kinetic modeling of genetically modified *Saccharomyces cerevisiae*, or baker's yeast, is explored. The addition of several non-native genes to the mevalonate pathway, resulting in the production of beta-carotene, is discussed. Experimental procedures are explained for the aerobic, batch cultivation of the baker's yeast, including procedures for measuring the key components of glucose, biomass, ethanol, acetic acid, and beta-carotene. A two-step parameter estimation, to estimate growth-rate parameters and then estimate kinetic parameters, is performed using nonlinear regression using an objective function based on the coefficient of determination (R^2) value. Local and global stability analyses are also conducted to determine the accuracy and importance of the models and their parameters.

Chapter V proposes a continuous production strategy for pharmaceutical production based on a multi-feed bioreactor configuration. This configuration is an improvement over traditional single-feed processes as it allows for the independent

manipulation of the bioreactor dilution rate and inlet substrate concentration by modulating the flowrates. An optimal control formulation is discussed that looks to maximize process productivity while maintaining steady state subject to the kinetics of the desired biochemical reaction through the manipulation of feed flowrates. A case study for the continuous production of beta-carotene is presented using the kinetics developed in Chapter IV and using a three-feed configuration for the addition of glucose and ethanol as substrates and media as a diluent. Optimal control policies and the resulting system state profiles are found for one-, two-, and three-feed configurations and the economics for each are assessed as the difference between the cost of materials and the profit from the salable product beta-carotene. Results will demonstrate the improved economic potential of continuous bioprocessing using multiple feed configurations over using only a single feed.

Chapter IV presents the development of a process synthesis framework for the economical production of continuous, biologically-produced ethanol using multiple cellulosic biomass feeds. Process alternatives considered in the synthesis framework include biomass selection and pretreatment method, effectively determining the cost-effective strategy for providing the raw material glucose from lignocellulosic biomass. A case study using this framework is presented for the United States. Biomass choices are selected from those available around the United States, with prominently studied methods for the pretreatment of each type considered. Results of this case study show that ethanol can be produced in an economically favorable manner from raw materials that do not influence food production. Additionally, the effect of using multiple biomass

feeds on the ethanol generation, electricity generation, and required harvesting area is characterized.

Chapter VII introduces the continuous paradigm discussed in previous chapters to the production of alternative liquid fuels, specifically biobutanol. Specifically, the economic viability of using the multi-feed bioreactor discussed in Chapter V coupled with *in situ* gas stripping will be examined. First, the optimal control problem will be formulated using the flowrate of the liquid feeds and stripping gas as manipulated variables to maximize the profitability of the intensified bioreactor system. The steady state results of the optimal control analysis will be used to develop a large-scale process synthesis problem like that considered for ethanol in Chapter VI. This synthesis problem will consider the cellulosic biomass types, pretreatment options, and number of butanol bioreactors as process alternatives. Results will show that the economic production of butanol is feasible and attractive, providing sustainable alternatives to currently used fossil fuels.

In Chapter VIII, novel methods for the pedagogy of process systems engineering methodologies in process integration are developed. This chapter begins with a brief overview of the currently employed process integration methods, namely direct recycle, mass integration, and heat integration, and proposes the inclusion of the newly developed financial integration to move from a linear teaching method to a more holistic approach. Basics of the graphical approach to financial integration are explained, and a new method for algebraically analyzing financial integration problems is postulated. The holistic teaching method is exemplified in two case studies, first in relation to a

generalized traditional chemical engineering process and then to a biochemical process for the conversion of lignocellulosic biomass to biofuels.

Chapter IX summarizes the results of the previous chapters. This chapter also re-emphasizes the importance of developing continuous strategies for sustainable design, specifically using biological conversion methods to produce alternative liquid fuels and pharmaceuticals. Finally, future research directions inspired by the work demonstrated in this thesis will be discussed.

CHAPTER II

PROCESS SYSTEMS ENGINEERING: FUNDAMENTALS AND APPLICATIONS

2.1. Process Systems Engineering (PSE): An Introduction

Process systems engineering (PSE) is the branch of chemical engineering that deals with the design, operation, optimization, and control of chemical process systems. The goals of PSE are (1) to ensure the economic preservation of commodity chemical systems via the improvement of product quality or process efficiency in addition to a reduction of costs associated with a process and (2) to propagate the growth of the specialty chemicals industry via the discovery of new products or new methods of delivering those and preexisting products to market in a fast and economically efficient manner [1]. Through the development of general methodologies, the PSE community looks to improve many chemical industries.

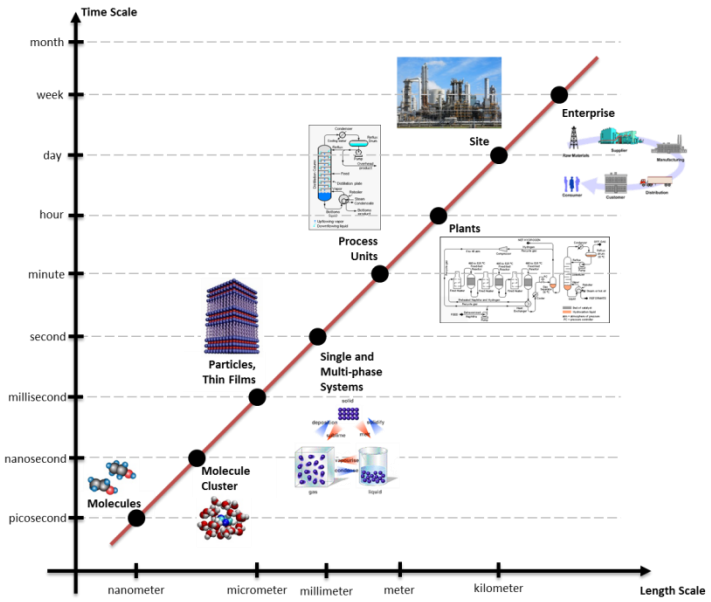

Figure 2.1. The chemical supply chain considered in process systems engineering [1].

Figure 2.1 depicts the various multi-scale applications of PSE, ranging from nanometer-scale molecular processes that operate on the order of picoseconds to kilometer-scale chemical production processes that operate over a period of days or weeks [1]. Examples of the applications of process systems engineering strategies include the modeling of complex chemical and physical phenomena, the design and synthesis of novel chemical processes, the control of processes to operate within set safety limits, and the integration of processes to conserve resources such as raw materials and energy. These applications will be discussed in more detail in Section 2.3.

The main challenge facing the PSE community is developing general methodologies capable of describing, designing, and controlling systems with such vast time and length scales. To this end, process systems engineering strategies are rooted in the solution of mathematical programs that minimize or maximize an objective subject to a set of constraint functions governing the physical and chemical phenomena of the system of interest. Typical objectives in chemical engineering applications include maximizing profitability, maximizing sustainability and minimizing environmental impacts, and minimizing safety hazards, among others. A generic mathematical program is shown below in Equation 2.1, with *Z* defining the objective value of interest and *x* and *y* defining the continuous and integer variables of the system, respectively [2].

$$\begin{aligned} &\text{minimize} && Z = f(x,y) \\ &\text{subject to} && g(x,y) \leq 0 \\ &&& h(x,y) = 0 \\ &&& x \in X \\ &&& y \in Y \quad \text{integer} \end{aligned} \tag{2.1}$$

Here, the equality constraints $h(x, y)$ often describe material and energy balances that must be satisfied for the system to be considered optimal. The inequality constraints $g(x, y)$ are typically used for limiting the feasible space of the potential solution due to physical limitations of the chemical process; for example, restricting the operating temperature of a reactor or the density of a mixture to a desired, feasible range. While continuous variables x are common in chemical systems, often denoting temperatures, pressures, and compositions, the integer variables y are used to make decisions, i.e. whether a processing unit would be included in an optimal process design or how many reactors are needed to achieve a target production rate.

The type of optimization problem given by Equation 2.1 is defined by the characteristics of the functions f, g, and h and on the existence of the variable set y. If the integer variables are not present, then Equation 2.1 takes the form of either a linear program (LP) if the functions f, g, and h are linear with respect to the continuous variables x or a nonlinear program (NLP) otherwise. If the integer variables y are present then Equation 2.1 is deemed a mixed-integer (MIP) program, which can either be linear (MILP) or nonlinear (MINLP) depending on the previously explained criteria for the functions f, g, and h. As an additional consideration, if any of the functions f, g, and h take the form of differential equations with respect to time then the program is referred to as a dynamic program or optimal control problem. While the adaptability of the functionality for f, g, and h is powerful for describing many different chemical and physical processes, it also requires solution methods capable of solving a mathematical program of any type to optimality.

2.2. Methodologies for Solving Mathematical Programing Problems

2.2.1. Local Optimality Criteria, the Impact of Constraints, and Global Optimality

Before discussing methods for finding the optimal solutions to mathematical programs, it is important to understand the criteria that makes a feasible point a candidate for optimality. The definition of an optimizer x^* for a function f in a set of feasible points S is shown in Equation 2.2 below:

$$f(x^*) \leq f(x) \text{ for all } x \in S \text{ such that } \left||x - x^*|\right| < \varepsilon \tag{2.2}$$

For unconstrained problems, any feasible value x may be the optima of a single-variable function if the derivative of $f(x)$ is necessarily zero. This can be seen in the example in Figure 2.2(a) below. Using the example of a quadratic function, we can see that the value of x is less than all nearby values at the origin, where the derivative of this function is also zero. However, to sufficiently determine that the feasible optima is indeed the optimum and not a saddle point, the second derivative of $f(x)$ must be necessarily positive definite. Referring again to Figure 2.2(a), we can see that the second derivative is 2, so the point (0,0) meets the necessary and sufficient condition to be the optimum solution. These necessary and sufficient conditions extend to unconstrained multi-variable systems, where the Jacobian matrix of partial first derivatives must be equivocally zero and the Hessian matrix of partial second derivatives must be positive definite, i.e. all eigenvalues are positive, at an optimal point x^*.

The addition of constraints to an optimization problem results in complications when determining the optimal solution. We exemplify this with Figure 2.2(b) above, where a linear equality constraint and an inequality constraint on the feasible values of x

are added to the unconstrained problem previously discussed. In this example, the optimal solution must lie within the feasible region of x, eliminating the optimal solution in the unconstrained problem, and must also lie at the intersection between the objective function x^2 and the linear equality constraint $2x$. This leads to a single feasible point, which is also the optimal value of x^*, at a value of 2; however, unlike the unconstrained problem, the optimal solution is not at the point where the derivative of the objective function is equal to zero. In the existence of constraints, new necessary and sufficient conditions and new methods must be established to determine optimality.

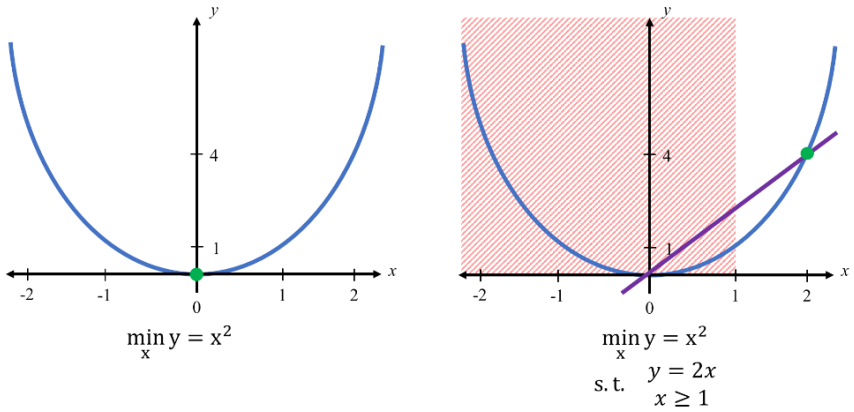

Figure 2.2. Example of an (a) unconstrained optimization problem and (b) constrained optimization problem. The red shaded region in (b) denotes the infeasible solution space of the optimization problem due to the inequality constraint. The green dots signify the optimal solution in each case.

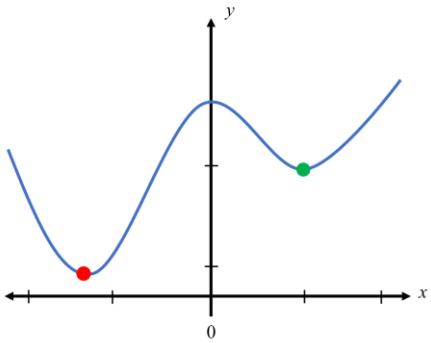

Figure 2.3. Example of an objective function with multiple optimal values. The local and global optimal solutions are designated in green and red, respectively.

In addition to the complication of constraints, many objective functions have a higher degree of nonlinearity when compared to the quadratic function shown in Figure 2.2. As such, these functions often have multiple local minima, like the function shown in Figure 2.3. This function has two local minima, one marked in red and another in green, with the red minimizer being one true global minimum of the function *f(x)*. It is often important to be able to determine the global optimum of an objective function; this can be done through stochastic methods such as a Monte Carlo search or genetic algorithms, but with no guarantee of success [3-5]. Often more deterministic methods that can guarantee convergence to the global solution are desired. These methods will be discussed in more detail in later sections.

Most chemical engineering applications require the consideration of both equality, i.e. mass and energy balance, and inequality, i.e. operating limits, constraints during analysis. These constraints may include integer variables for decision-making purposes or differential equations for temporally- or spatially-variant systems. Often in chemical engineering applications it is important to find the global minimum of a solution, which results in the highest level of profitability or sustainability or conversely the lowest level of safety risk associated with a process. Methods for solving optimization problems, from simple unconstrained, univariate problems to those with the complexity relevant to chemical engineering applications, will be discussed in the following sections.

2.2.2. Continuous Systems

2.2.2.1. Optimization Methods for Determining Local Solutions

With an understanding of the optimality criteria and the additional challenges faced with constraints and objectives with multiple optima, we can begin to develop methods for solving mathematical programs. Regardless of the complexity of the optimization problem, all solution strategies follow the general algorithm shown in Figure 2.4 below [6]. The algorithm begins by supplying an initial guess of the solution x_0. Then, for every iteration $k = 0,1, \dots, n$ the following four steps are taken. First, the point x_k is checked for optimality and the algorithm is stopped if optimality is determined true. If the point x_k is not optimal then a search direction p_k is determined. Additionally, a step length α_k ensures that a nontrivial reduction is made in the function value. Finally, the new point x_{k+1} is calculated from Equation 2.3 below.

$$x_{k+1} = x_k + \alpha_k p_k \tag{2.3}$$

Two methods can be used to calculate α_k, the backtracking method and the Wolfe condition. The backtracking method, shown as Equation 2.4, sets the value α_k to a set value and requires the check of a simple inequality to ensure that the solution of the iteration is at least some fraction μ of the decrease predicted by a linear approximation of the function. If a reduction is achieved the value of α_k is used, otherwise α_k is reduced and the check is performed again. In each iteration, the step size α_k is reset to its initial value.

$$f(x_k + \alpha p_k) \leq f(x_k) + \mu \alpha_k p_k^T \nabla f(x_k) \tag{2.4}$$

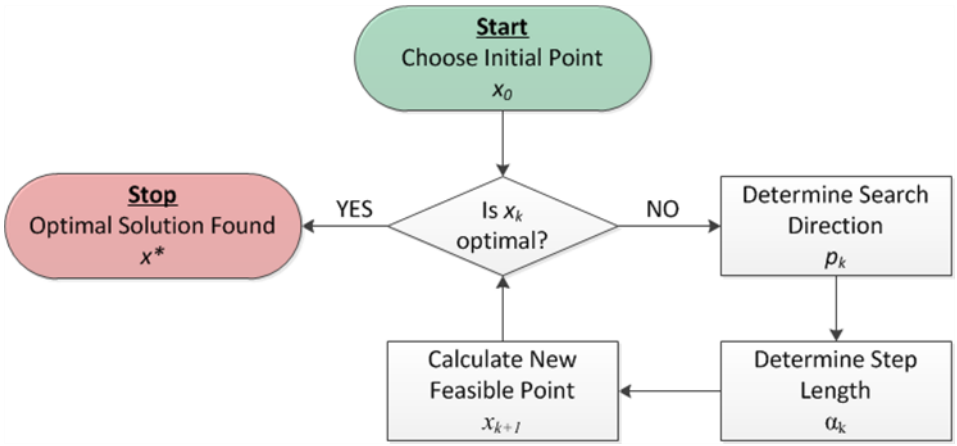

Figure 2.4. General framework for solving general optimization problems.

Alternatively, a more efficient but more complicated way of calculating the step size α_k is the Wolfe condition shown in Equation 2.5. This condition is the result of the solution of a one-dimensional optimization problem that minimizes the function $f(x_k + \alpha_k p_k)$. The need to solve the Wolfe condition leads to more complex algorithms than the solution to the backtracking method, but often leads to better algorithms by further restricting the feasible space of α_k. To guarantee a decrease in the function value, the Wolfe condition is paired with the sufficient decrease condition in Equation 2.4 with the criteria that $\mu < \eta$. Illustrations for both the backtracking method and the Wolfe condition are shown in Figure 2.5.

$$|p_k^T \nabla f(x_k + \alpha_k p_k)| \leq \eta |p_k^T \nabla f(x_k)| \tag{2.5}$$

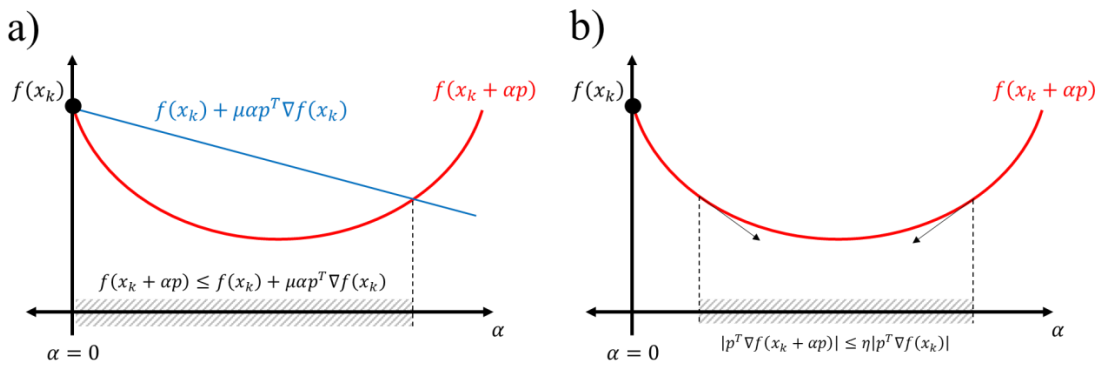

Figure 2.5. Line search methods for determining the sufficient decrease in function value: (a) the backtracking method and (b) the Wolfe condition (adapted from Griva et al.) [6].

The construction of algorithms capable of handling problems found in chemical engineering begin with ways of estimating the α_k and p_k values in Equation 2.3. Three methods to determine these values will be discussed for unconstrained problems consisting of only an objective function and will be used as a foundation upon which all solution strategies for more complicated problems are developed. The addition of constraints results in the need to reformulate the problem into a new, single objective function to allow for the use of the unconstrained methodologies. Further complexity is added for problems with integer variables, which is handled by systematically relaxing the integer constraints to be continuous, allowing the use of other methods previously discussed. Specific methods for unconstrained, constrained, and integer optimization problems will now be discussed in more detail.

Unconstrained Optimization Algorithms

When solving an optimization problem of the simplest form, i.e. a minimization problem with no constraints, the logical solution is to construct a solver that heads

strictly "downhill" in the feasible space. The simplest method for solving unconstrained optimization problems, the gradient descent or steepest-descent method, follows this logic exactly by equating the search direction p_k to the negative gradient of the function $f(x)$, or $-\nabla f(x)$. Developed by Cauchy in 1847, This method requires that the objective function is differentiable and converges to a solution linearly [7]. Additionally, because this method does not account for the curvature of the objective function, it is often necessary to include a line search to calculate α_k to guarantee convergence to a solution.

Modifications have been considered to compensate for and speed up the slow convergence rate of the gradient descent method. The most commonly used modification is to use the curvature of the objective function $f(x)$ at the point x_k to determine a modification for the step length p_k. To do this, the second-order Taylor series of the objective function $f(x_k)$ at iteration *k+1* is used, as shown in Equation 2.6. The definition of a stationary point is then used by setting the derivative of Equation 2.6 with respect to the step length p_k to zero. The solution to Equation 6 gives the step length p_k as $\nabla f(x_k)[\nabla^2 f(x_k)]^{-1}$.

$$f(x_{k+1}) = f(x_k + p_k) \approx f(x_k) + f'(x_k)p_k + \frac{1}{2}p_k^T[\nabla^2 f(x_k)]p_k \tag{2.6}$$

$$\frac{df(x)}{dp_k} = 0 = \nabla f(x_k) + \nabla^2 f(x_k)p_k \quad \rightarrow \quad p_k = \frac{-\nabla f(x_k)}{\nabla^2 f(x_k)} \tag{2.7}$$

The method that considers the curvature of the function $f(x)$ in its solution algorithm by using the step length p_k shown in Equation 2.7 is known as Newton's method. By considering the curvature in the search direction, unconstrained optimization problems solved with Newton's method achieve quadratic convergence to the optimal

solution x^*. However, this improved convergence speed comes at the cost of needing the objective function to be twice differentiable, as the calculation of the Hessian matrix of second derivatives $\nabla^2 f(x)$ is required. Additionally, as the number of variables is increased the order of arithmetic operations increases rapidly, making Newton's method impractical to use. Newton's method is often seen as the ideal method for solving unconstrained problems but is often not used in practice in its native form [6].

Methods have been developed to increase the practicality of Newton's method by using the matrix B_k as an estimate of the Hessian matrix in the step length p_k, as defined in Equation 2.8 below. The matrix B_k must be positive definite to achieve the goal of minimization and many different methods have been considered to estimate and update the matrix B_k for every iteration [8-12]. One of the most popular methods to estimate B_k is called the BFGS method, named after its developers Broyden, Fletcher, Goldfarb, and Shanno, and is shown in Equation 2.9.

$$p_k = \frac{-\nabla f(x_k)}{B_k} \tag{2.8}$$

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \tag{2.9}$$

Two new parameters, s_k and y_k, are needed to complete the update for B_k and are defined as $s_k = x_{k+1} - x_k$ and $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$. Quasi-Newton methods are the most common methods for unconstrained optimization found in commercially available solvers as the estimation of the Hessian matrix avoids unnecessary calculations. However, Quasi-Newton methods are implemented at the expense of

reducing the convergence speed from quadratic to super-linear, still an improvement over the gradient-descent method.

All the methods previously discussed require the objective function to be at least continuous in $\mathbb{R}^N$ to calculate the necessary first order derivatives with respect to all N variables. However, for objective functions it may not be possible or practical to compute first derivatives; for example, the function may be discontinuous at certain points, making the derivative undefined at points, or first derivatives may be difficult to calculate even if they do exist. Alternative, derivative-free methods can be used to optimize objective functions of this type [13]. The most common derivative free method is the Nelder-Mead method, which searches a simplex of sample points x_{k+1}^i around the current point x_k and accepts a "downhill" direction based on the evaluation of the objective function and other criteria at each x_{k+1}^i [14].While these methods often converge more slowly than their derivative based counterparts, and are not suited well for high-dimensional problems, they do offer strategies for handling cases with problems where derivatives are not available [6].

Constrained Optimization Algorithms

To this point methods that can solve optimization problems of a single objective functions have been established. In practice, however, these types of problems are rare and often are accompanied by constraints that restrict the feasible space where the optimal solution resides. Methods for solving constrained problems have been developed that handle constraints by reformulation into a single objective function form that can be solved using the established methods previously discussed. Three of these methods,

penalty methods, barrier methods, and the method of Lagrange multipliers, will be discussed in more detail.

Perhaps the simplest way of handling equality constraints is to include increases in the value of the objective function when a constraint is violated. This thinking gives rise to a subset of penalty methods that repurpose the constraints into a second term $\rho\psi(x)$ of the objective function as shown in Equation 2.10. This function $\psi(x)$ is a function of the equality constraints $h_i(x)$ and has the property described in Equation 2.11; the function must be equivalently zero at points x in a region that does not violate the constraint and greater than zero otherwise. Additionally, the value of the weighting factor ρ, also termed the penalty parameter, is strictly positive and larger values of ρ force the solution of the reformulated problem toward the feasible region. Typical penalty method algorithms solve the constrained optimization problem by using a lenient, low-valued penalty parameter and solve successive minimization problems using the previous solution x^* and increasingly larger penalty parameters ρ_k until the solution x^* is no longer sufficiently decreasing.

$$\pi(x) = f(x) + \rho\psi(x) \tag{2.10}$$

$$\begin{aligned} \psi(x) &= 0, \quad \text{if } x \text{ is feasible} \\ \psi(x) &> 0, \quad \text{otherwise} \end{aligned} \tag{2.11}$$

Penalty function methods vary with the selection of the function $\psi(x)$. Two of the most common penalty functions are shown as Equation 2.12 and Equation 2.13. In the first case, the quadratic-loss penalty is used as proposed by Courant [15]. The second is a more general penalty function, for which the value of γ is greater than 1; when γ is

2, the quadratic-loss penalty function is recovered. An illustrative example of a constrained optimization problem solved with the penalty function method is shown in Figure 2.6.

$$\psi(x) = \frac{1}{2}\sum_{i=1}^{m}[h_i(x)]^2 \tag{2.12}$$

$$\psi(x) = \frac{1}{\gamma}\sum_{i=1}^{m}|h_i(x)|^{\gamma} \qquad \gamma \geq 1 \tag{2.13}$$

In the same vein as penalty methods, barrier methods are used to solve problems with inequality constraints by modifying the objective function in a way that limits the feasible space to that allowed by the constraint functions $g_i(x)$. However, instead of enforcing a penalty when in an infeasible region, this method works by adding a secondary term $\mu\phi(x)$ to the objective function $f(x)$ that creates a barrier that results in an undefined function evaluation if the infeasible region is entered. The barrier function $\phi(x)$ has the property that it must approach an infinite value as the function $g_i(x)$ approaches zero; additionally, the value of the optimizer x^* will approach the true optimal solution as the value of the weight parameter μ approaches zero. General algorithms using the barrier method begin with a large value of μ, solve the reformulated optimization problem as an unconstrained problem, and continue to resolve the unconstrained problem with a monotonically decreasing value of μ until a sufficiently small decrease in the minimizer x^* is achieved.

$$\beta(x) = f(x) + \mu\phi(x)$$

$$\phi(x) \to \infty \quad \text{as} \quad g_i(x) \to 0_+$$

As with penalty methods, various barrier method algorithms vary with the selection of the barrier function $\phi(x)$. One choice, the logarithmic barrier function, is shown in Equation 2.14 [16, 17]. This method restricts the value of $g_i(x)$ to be greater than zero as the logarithmic function is undefined for negative inputs. Another option is the inverse function shown in Equation 2.15. This function is small for large values of $g_i(x)$ but gets asymptotically larger as the value of the constraint approaches zero. An example of the barrier function solution method is shown in Figure 2.6 below.

$$\phi(x) = -\sum_{i=1}^{m} \log\big(g_i(x)\big) \tag{2.14}$$

$$\phi(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)} \tag{2.15}$$

While penalty and barrier function methods are useful, they are often subject to complications if implemented poorly. First, the reformulated unconstrained optimization problems for both methods can become increasingly difficult to solve due to the ill-conditioned Hessian matrix of $\pi(x)$ and $\beta(x)$ as ρ and μ are increased and decreased, respectively [6]. Additionally, penalty and barrier methods often result in solutions that are close but not exactly the solution to the original constrained problem; in fact, the penalty method often gives a solution that is in the infeasible space as seen in Figure 2.6 below. However, an additional method can be utilized that avoids these problems and determines the exact solution to the constrained optimization problem in the presence of both equality and inequality constraints simultaneously.

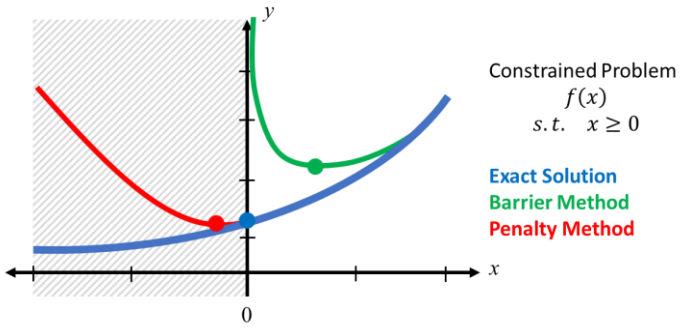

Figure 2.6. Illustrative example of the penalty (red) and barrier (green) methods for solving constrained optimization problems. The exact solution is shown in blue.

A more complex method for solving constrained optimization problems was developed by Joseph Louis Lagrange in 1761 and, like the penalty and barrier methods, uses a reformulation of the problem into a single objective function [18]. This method begins by defining a new function called the Lagrangian function, or $L(x, \lambda, \mu)$, that is the linear combination of the objective function and the constraints of the system. This is shown as Equation 2.16 below, with the coefficients λ and μ being termed the Lagrange multipliers for the *p* equality constraints and *q* inequality constraints, respectively.

$$L(x, \lambda) = f(x) + \sum_{j=1}^{n} \lambda_j h_j(x) + \sum_{i=1}^{m} \mu_i g_i(x) \tag{2.16}$$

Due to the reformulation of the original objective function and constraints into a single Lagrangian function with new parameters, new criteria for optimality are necessary. These criteria are the Karush-Kuhn-Tucker (KKT) conditions shown in Equation 2.17 [19, 20]. The first criterion is that the value of the Jacobian of the Lagrangian function with respect to x must be zero at the stationary point (x^*, λ^*, μ^*). Additionally, all Lagrange multipliers for the inequality constraints μ_j must be positive

or zero; the Lagrange multipliers λ_i for the equality constraints can take on any real value. The third condition is called the complementary slackness condition and requires that either the inequality constraint $g_i(x)$ is active, i.e. has a value of zero, or its associated eigenvalue is zero. Finally, the Hessian of the Lagrangian function must be positive definite, as was the case for the unconstrained objective function $f(x)$.

$$\begin{gathered} \nabla_x L(x^*, \lambda^*, \mu^*) = 0 \\ \mu^* \geq 0 \\ \mu^* g(x^*) = 0 \\ Z^T [\nabla^2 L(x^*, \lambda^*, \mu^*)] Z \text{ is positive definite} \end{gathered} \tag{2.17}$$

With a new formulation and a new set of optimality criteria for systems with equality and inequality constraints, an algorithm for solving the constrained optimization problem can be considered. Termed sequential quadratic programming (SQP), the algorithm first solves for the search direction for the variables x, or p_k, and the Lagrange multipliers, or ν_k, by solving a quadratic program. This quadratic program is composed of a second-order Taylor series of the Lagrangian function at (x_k, λ_k) subject to constraints that equate a linear approximation of $g(x_k + p_k)$ to zero as shown in Equation 2.18. Here, the Lagrange multipliers for the equality and inequality constraints are considered in λ_k and the search direction ν_k. This search direction is then used to update the values of the x_k and λ_k. The steps are repeated until the KKT optimality criteria in Equation 2.17 are satisfied. This methodology for constrained optimization corresponds to the application of Newton's method for unconstrained optimization shown in Figure 2.4, and the same simplifications can be used to improve the algorithm with respect to the number of calculations needed.

$$\begin{aligned} \min q(p) &= \frac{1}{2} p^T [\nabla^2_{xx} L(x_k, \lambda_k)] p + p^T [\nabla_x L(x_k, \lambda_k)] p \\ s.t. \quad & [\nabla g(x_k)]^T p + g(x_k) = 0 \end{aligned} \tag{2.18}$$

2.2.3. Mixed-Integer Systems with Integer Variables

To this point, algorithms for solving nonlinear programs (NLPs) with equality and inequality constraints have been developed. However, many chemical engineering applications require decision-making, planning, and scheduling, a feat not always feasible with only continuous variables. This leads to the inclusion of integer variables more suited to these tasks, resulting in complex optimization formulations like that shown in Equation 2.19.

$$\begin{aligned} \min_{x,y} Z &= f(x,y) \\ s.t. \quad & g(x,y) \leq 0 \\ & h(x,y) = 0 \\ & x \in \mathbb{R}^N \\ & y \in [0,1] \end{aligned} \tag{2.19}$$

To this point, the algorithms that have been discussed can handle nonlinearities, inequality constraints, and equality constraints, but are not capable of dealing with problems with added integer variables, termed mixed-integer programs (MIP). This section will cover the ideas and algorithms available for these mixed-integer optimization problems. The methods for mixed-integer linear programs (MILP) with linear objective function $f(x,y)$ and linear constraints $g(x,y)$ and $h(x,y)$ will first be discussed, with emphasis placed on the branch-and-bound technique as the most common method used for large mixed-integer problems. The solution strategies for

MILP problems will then be used as a basis for the more complex mixed-integer nonlinear programming (MINLP) algorithms.

2.2.3.1. Linear Systems with Integer Variables – Mixed-Integer Linear Programs (MILP)

The solution strategies for mixed-integer linear programs can be separated into four broad categories: (1) cutting plane methods, (2) decomposition methods, (3) logic-based methods, and (4) branch-and-bound methods [21]. In cutting plane methods, new constraints are added to the optimization problem that reduce the feasible region until a 0-1 optimal solution is obtained [22-25]. Decomposition methods, beginning with those developed by Benders in 1962, utilize variable partitioning, duality, and relaxation methods to exploit and decompose the mathematical structure of the models [24, 26-28]. Logic-based methods utilize disjunctive constraints or symbolic inference techniques that define rules for the behavior of the binary variables and reduce the problem size [29-32]. Finally, branch-and-bound methods utilize a binary tree to represent the 0-1 combinations of binary variables, partitioning the feasible region into subdomains systematically [33-35]. An example of this method for a system of three binary variables is shown in Figure 2.7. This strategy is the most widely used for MILP problems, and will now be explained in more detail.

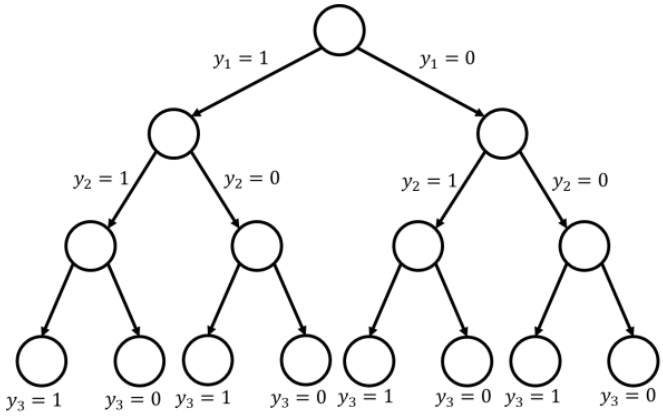

Figure 2.7. Binary tree representation of the branch-and-bound algorithm for solving mixed-integer linear programs (MILP).

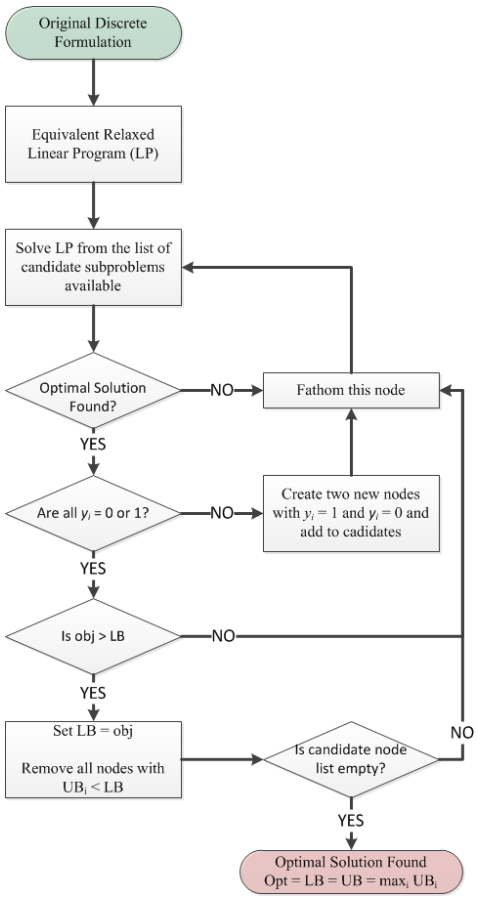

Figure 2.8. Branch and bound algorithm for the solution to mixed-integer linear programs.

The general branch and bound algorithm for solving MILP problems is shown in Figure 2.8. The algorithm begins by solving a relaxation of the original MILP by removing the restrictions on all integer variables and formulating a linear program (LP) that can be solved using traditional methods such as the simplex method. The most common relaxation method is to allow the binary variables *y* to take any value between zero and one; this will result in an optimal solution of the relaxed LP with fractional values for most if not all binary variables. If the relaxed problem does not have an optimal solution, the problem is deemed infeasible.

$$\begin{aligned} \min_{x,y} Z &= f(x,y) \\ s.t. \quad & g(x,y) \leq 0 \\ & h(x,y) = 0 \\ & 0 \leq y \leq 1 \end{aligned} \tag{2.20}$$

For a feasible problem, a binary variable y_i that has taken a fractional value in the relaxed problem of the initial node is chosen as the branch variable. This results in two new daughter nodes with a set value of y_i, one with a new formulation where $y_i = 0$ and a second where $y_i = 1$. One of these formulations is chosen as the new candidate subproblem (CS) and is solved using a similar relaxation to that in Equation 2.20. This node must then pass through a series of checks to determine if a new optimal solution of the master MILP has been found. First, the relaxed linear program of the CS must have a feasible solution. If it does not, then the node is considered infeasible, the value to which y_i was set cannot be true in the solution to the master MILP, and the node is fathomed; the daughter nodes of any fathomed node do not need to be considered further.

If an optimal solution to the candidate problem exists, the values of the remaining binary variables are checked. If any of the binary variables take a fractional value in the optimal solution, then this variable is selected as another branch point and two new daughter nodes are created for each instance of the binary value. However, if all values are binary, then this solution satisfies the MILP master problem. If the objective function value of the CS is greater than the current optimal solution, the objective value of the CS becomes the new optimal solution, otherwise the node is fathomed and a new CS is chosen from the unfathomed nodes. In this way, the lower bound on the solution is continuously increased until the optimal solution is found. The checks on each node continue until no candidate sub-problems remain and a solution to the master MILP problem has been determined.

While this algorithm guarantees that the entire feasible region of the master MILP is searched for the globally optimum solution, it is subject to three techniques that have been considered for choosing subsequent candidate subproblems in the binary tree: (1) the Last-In-First-Out (LIFO) approach, also known as the depth-first with backtracking strategy, (2) the breadth-first strategy, and (3) the best-bound strategy. In the LIFO strategy, the next candidate subproblem is taken to be the daughter node of the current subproblem, leading the solver down a single branch of the binary tree before investigating other branches [21]. When the current branch is fathomed, the solver backtracks to another branch and continues down that branch of LP problems. Alternatively, the breadth-first strategy choses the next CS as another node at the same level of the binary tree before heading down any specific branch [21]. Finally, the best-

bound strategy takes a hybrid approach, choosing the next node on the list that has the lowest lower bound [21]. However, it should be noted that there is a trade-off between each node-selection strategy and there is therefore no rule that predict fast solvers toward optimal solutions.

2.2.3.2. Nonlinear Systems with Integer Variables – Mixed-Integer Nonlinear Programs (MINLP)

Designing many chemical engineering applications, including those involving reaction kinetics, fluid mechanics, and phase equilibrium, require the coupling of integer variables for design decisions with inherently complex, nonconvex, nonlinear models. These systems cannot be described by an MILP formulation without loss of accuracy by linearization or including many additional constraints and integer variables using disjunctions that characterize the nonlinearity of these models. Finding solutions to MINLP formulations like those in Equation 2.21 pairs the difficulty experienced with solving usually nonconvex nonlinear programs (NLP) and the complexity of solving mixed-integer linear problems (MILP). This section will describe methods for determining the optimal solution for mixed-integer nonlinear programs (MINLPs) of the following form:

$$\begin{aligned} \min_{x,y} Z &= c^T y + f(x) \\ & g(x) \leq 0 \\ & h(x) = 0 \\ & Ax = a \\ s.t. \quad & Cx + By \leq b \\ & Ey \leq e \\ & x^L \leq x \leq x^U \\ & y \in \{0,1\}^n \end{aligned} \tag{2.21}$$

While the branch-and-bound strategies discussed for MILPs can be applied, finding the solution to the relaxed NLPs in each node would require much more computational effort than the relaxed linear programs found at each node for MILPs. Instead, a more elegant solution strategy has been developed that couples the solving of an NLP subproblem with fixed binary variables and an MILP subproblem to determine an upper bound and lower bound, respectively, to the MINLP master problem. This general algorithm is shown in Figure 2.9. If the lower bound is found to be greater than the upper bound, then the optimal solution is found; alternatively, if a solution does not exist for either the NLP or the MILP then the master MINLP is deemed infeasible. Specific MINLP solvers are differentiated by their formulation strategy for the NLP and MILP subproblems. Two specific and widely-used methods, the Outer Approximation (OA) method and General Benders Decomposition (GBD)s, will be discussed in more detail.

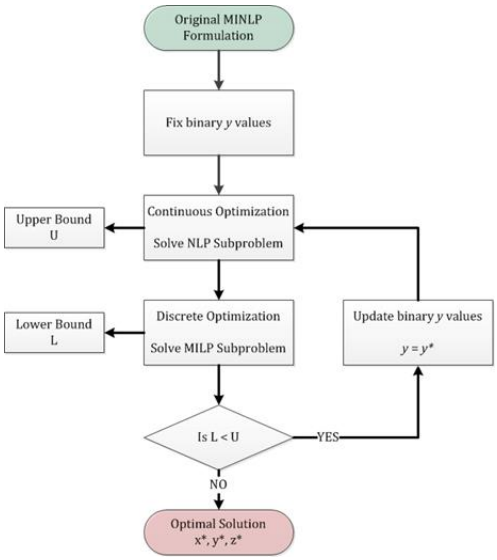

Figure 2.9. General algorithm for the solution of mixed-integer nonlinear programs (MINLP).

Both OA and GBD utilize an NLP subproblem by setting the set of binary variables and solving the resulting NLP using methods discussed in Section 2.2.2. Instead, the differences in these algorithms lie in how the MINLP problem is reformulated into an MILP subproblem. The OA algorithm, developed by Duran and Grossmann in 1986, formulates the MILP subproblem by linearizing the objective function and inequalities [36, 37]. This is done by calculating a diagonal relaxation matrix T^k from the value of the Lagrange multipliers λ^k calculated as the solution to the NLP subproblem, as shown in Equation 2.22. Using this and the first order Taylor series approximation of the objective function, defined as α_{OA} and the inequality constraints, the MINLP is reformulated to the MILP shown in Equation 2.23. This reformulation is used to solve for the next iteration of the binary variables *y*, and the cycle between NLP and MILP continues until the upper and lower bounds converge and an optimal solution is found.

$$T^k = \{t_{ii}\} = \begin{cases} +1, & if\ \lambda_i^k > 0 \\ -1, & if\ \lambda_i^k < 0 \\ 0, & if\ \lambda_i^k = 0 \end{cases} \tag{2.22}$$

$$\begin{aligned} &\min_{x,y} z_{OA}^k = \alpha_{OA} \\ &\alpha_{OA} \geq c^T y + f(x^k) + \nabla f(x^k)^T (x - x^k) \\ &T^k\Big(\nabla h(x^k)^T (x - x^k)\Big) \leq 0 \\ s.t. \quad &g(x^k) + \nabla g(x^k)^T (x - x^k) + By \leq 0 \\ &Ay \leq a \\ &\sum_{i \in B^k} y_i - \sum_{i \in N^k} y_i \leq |B^k| - 1 \end{aligned} \tag{2.23}$$

Alternatively, GBD uses a similar but more sophisticated approach to develop the MILP reformulation. Instead of using separate linearization for the objective function

and inequality constraints, this method uses a linear approximation given by the Lagrangian at the NLP solution (x^k, λ^k) [31, 38, 39]. The specific formulation of this Lagrangian is shown in Equation 2.24 below. Unlike the OA algorithm, no adjustments are needed to include the equality constraints $h(x)$ as the extra term $\lambda^k h(x^k)$ in the Lagrangian reformulation is necessarily zero. Using this Lagrangian definition, the original MINLP is reformulated into the MILP subproblem by minimizing the value of α_{GB} which is necessarily greater than or equal to the Lagrangian in Equation 2.24. This reformulation is shown in Equation 2.25. As with the OA algorithm, this reformulation is used to determine the new values of the binary variables *y* for the next iteration of the NLP subproblem.

$$L(x^k, y, \lambda^k) = c^T y + f(x^k) + (\lambda^k)^T (Cx^k + By - b) \tag{2.24}$$

$$\begin{aligned} &\min_y z_{GB}^k = \alpha_{GB} \\ &\alpha_{GB} \geq c^T y + f(x^k) + \sum_{i=1}^{t} \lambda_i^k (g_i(x^k) + b_i y_i) \\ s.t. \quad &Ay \leq a \\ &\sum_{i \in B^k} y_i - \sum_{i \in N^k} y_i \leq |B^k| - 1 \end{aligned} \tag{2.25}$$

When considering the two algorithms available to solve MINLP problems, it is important to discuss the trade-offs when using either the GBD and OA method. The first point of note is the requirements of the MILP formulations from each decomposition strategy. The MILP formulation for GBD in Equation 2.23 is only a function of the integer variables *y*. However, it can be seen in Equation 2.25 that the OA algorithm requires that the MILP reformulation be a function of both the continuous variables *x*

and the binary variables *y*. This results in a larger MILP formulation with an increased computational requirement to solve in the case of the OA algorithm, but also allows for tighter lower bounds and therefore fewer NLP solutions to be calculated. Inversely, the MILP formulation for GBD is smaller than that of OA, being an almost pure integer problem, but the complete delineation between the continuous and discrete spaces require more iterations of the algorithm and therefore more NLPs to be solved. While both methods are effective in solving MINLP problems, the computational trade-offs should be considered when choosing a solution method.

2.2.4. Differential Systems

To this point, all solution strategies for optimization problems have dealt with algebraic systems of equations. However, many applications in chemical engineering are interested in how systems perform over time, and therefore require the use of differential equations to define these system dynamics. The optimization of systems of time-dependent equations is a subset of optimization theory known as optimal control, defined by the problem formulation shown in Equation 2.26. Here, the variables $x(t)$ are the system states that are functions of time and q is the set of control parameters or manipulated variables. Typically, the state and control variables are connected by an ordinary differential equation (ODE) defined as $L(x, y)$. Finally, constraints can be put on the states in the form of $g\big(x(t)\big)$ and on the control parameters in the form of $h(q)$.

$$\begin{aligned} \min_{x=x(t),q} \; & z(x,t) = f(x(t),q) \\ & L(x(t),q) = 0 \\ s.t. \quad & g\big(x(t)\big) \geq 0 \\ & h(q) \geq 0 \end{aligned} \tag{2.26}$$

While the inclusion of differential equations into the optimization problem adds some complexity, the general strategy is to reformulate the problem into an algebraic system and allow for the use of the methods previous discussed in Section 2.2.3. The following sections will be devoted to discussing direct solution strategies for optimal control problems. Indirect methods based on deriving optimality criteria have also been developed, and the reader is direct to literature for more information on those methods [40-42].

2.2.4.1. Decomposition Strategies

The simplest direct method for reformulating an optimal control problem into a system of algebraic equations is by discretizing the differential terms into their finite difference representation. In its simplest form, the time horizon T is divided into N segments of length Δt such that $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$. Finite difference methods, for example the forward Euler method, are then applied using this time step to all differential and integral terms in the system of equations. An example of a generic ODE and integral reformulation is shown as Equations 2.27 and 2.28.

$$\frac{dx}{dt} \rightarrow \frac{x^n - x^{n-1}}{\Delta t} \tag{2.27}$$

$$\int_0^T F\big(x(t),t,q(t)\big) \rightarrow \sum_{n=1}^{N} \Delta t \cdot F\left(\frac{x^n + x^{n+1}}{2}, t_n, \frac{q^n + q^{n-1}}{2}\right) \tag{2.28}$$

When introduced correctly, the introduction of finite difference methods results in a system of algebraic equations where one time step is inherently dependent on the previous time step through the discretization. Three techniques can be developed from the methods of discretization that will be discussed in more detail: (1) single shooting, (2) multiple shooting, and (3) collocation on finite elements.

The single shooting method is the most basic of discretization methods for optimal control problems. For this method, the control actions $q(t)$ are discretized to values q_i over the entire time length *T*. The dynamics of the problem are converted to a system of nonlinear equations by discretizing the differential equation constraints using a time-marching numerical integration technique such as forward or backward Euler or more advanced Runge-Kutta methods [43-45]. Likewise, the objective function is discretized using a numerical integration strategy. The optimal values of the control actions q_i can then be determined by solving the NLP formed by these reformulated objective function and constraints. An illustrative example of the single shooting method is shown in Figure 2.10. While useful, this method is often susceptible to calculation errors in the discretization of the differential equation, and as a result does not work well with differential systems of high nonlinearity [45].

An alternative to the single-shooting method is the multiple-shooting method. In this method, the time domain *T* is divided into *N+1* intervals t_i in which the single-shooting method is used. In this problem, a new set of variables s_i are used to denote the initial condition for the state variables $x(t, q)$ in each interval. To guarantee the continuity of the solution $x(t, q)$ in the final solution, a new set of equality constraints

are added to the optimization formulation as shown in Equation 2.29. More accurate solutions are obtained using this method as the numerical integration is done over a shorter time course than the single-shooting method, resulting in less susceptibility to numerical error. Additionally, parallelization can be used when solving the solutions to the individual time domains [46]. However, these improvements to the algorithm comes at the cost of increasing the size of the NLP relaxation.

$$s_{i+1} = x_i(t_{i+1}, s_i, q_i) \tag{2.29}$$

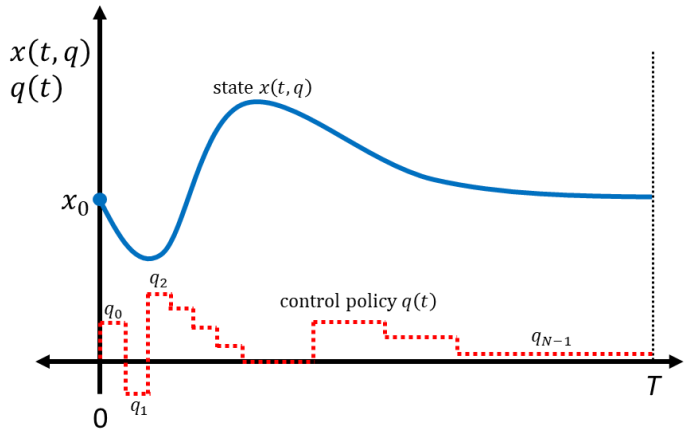

Figure 2.10. Illustrative example of the single shooting method for solving optimal control problems.

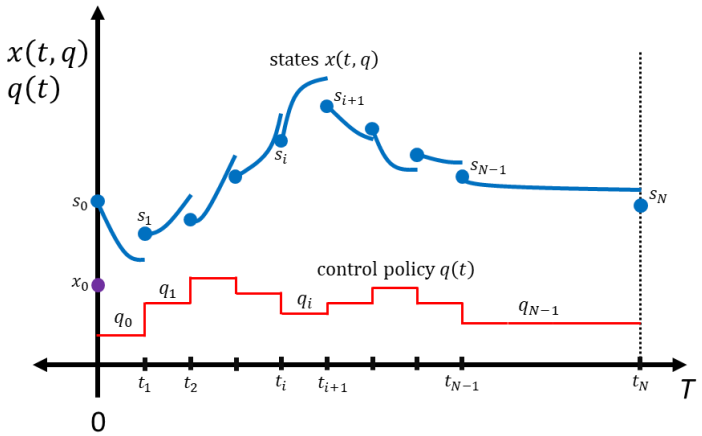

Figure 2.11. Illustrative example of the multiple-shooting method for solving optimal control problems.

An alternative to the shooting methods, collocation on finite elements discretizes both the manipulated variables $q(t)$ and the state variables $x(t)$. However, unlike the shooting methods, the state variables are not approximated using a discretization of the differential terms but are instead approximated using a specified functional form. Discretization using collocation on finite elements begins with the division of the time domain *T* into *S* subintervals $[t_{s-1}, t_{s+1}]$ finite elements where $t_S = t_f$.

Early discretization within each finite element was done using implicit Runge-Kutta methods due to their stability properties when compared to explicit methods [43, 44, 47, 48]. In 1988, Reddien first proposed using local orthogonal collocation methods that fixed the degree of the interpolating polynomial and varies the number of segments [49]. Reddien's work used Gauss-Legendre points and a cubic spline as a basis function, but other researchers have also considered using Chebychev polynomial, Lagrange polynomials, and higher-order Gauss-Lobatto methods as an alternative [50-53].

Alternatively, pseudospectral methods have been developed that allow for the variation of the degree of the interpolating polynomial while fixing the number of segments in the collocation. To do this, a global polynomial of variable degree is used to estimate the states and collocation is performed at chosen points. The use of global collocation methods has been shown to increase the convergence rate compared to local collocation methods [45]. While not covered here in detail, global collocation strategies have been developed for Gauss-Lobatto quadrature [54-56], Legendre-Gauss quadrature [57], and Legendre-Gauss-Radau quadrature [58, 59].

2.2.5. Methods for Finding Global Solutions

To this point, all the algorithms that have been discussed for solving optimization problems have resulted in locally optimal solutions. For linear convex nonlinear systems, this local solution is in fact the global solution for the optimization problem. However, many applications in chemical engineering are modeled using highly nonlinear models, often resulting in nonconvexity when used in optimization applications. As a result, methods are needed to deterministically guarantee the global solution to nonconvex systems will be found.

The most widely used strategy for guaranteeing convergence to a global solution was developed by McCormick in 1975 [60]. In his method, McCormick proposed using an iterative strategy of convex relaxations, branching, and bounding to limit the feasible region in which the global solution can lie. An example of this method is shown in Figure 2.12. First, the feasible region of the global solution is underestimated with a convex function as shown in Figure 2.12a. Then, the upper bound is calculated using a local solver and the lower bound is calculated by solving the relaxed problem. Assuming the upper bound is larger than the relaxation, the feasible region is split into two regions. Convex relaxations are again calculated in the upper bound and lower bounds in these two sub-regions are calculated as shown in Figure 2.12c. If the lower bound of any region is greater than the upper bound of another region, the regions with the greater upper bounds are fathomed and removed from the feasible region as shown in Figure 2.12d. This proves is continued until the global solutions is found to a reasonable level.

Many researches have expanded on McCormick's method to develop algorithms that determine global solutions using various relaxation methods [61-65]. Two

algorithms capable of solving NLP and MINLP models to global optimality using variations of this technique are commercially available: the **A**lgorithms for Co**NT**inuous/**I**nteger **G**lobal **O**ptimization of **N**onlinear **E**quations (ANTIGONE) software package developed by Misener and Floudas [66-68] and the **B**ranch **A**nd **R**educe **O**ptimization **N**avigator (BARON) software package developed by Tawarmalani and Sahinidis [65].

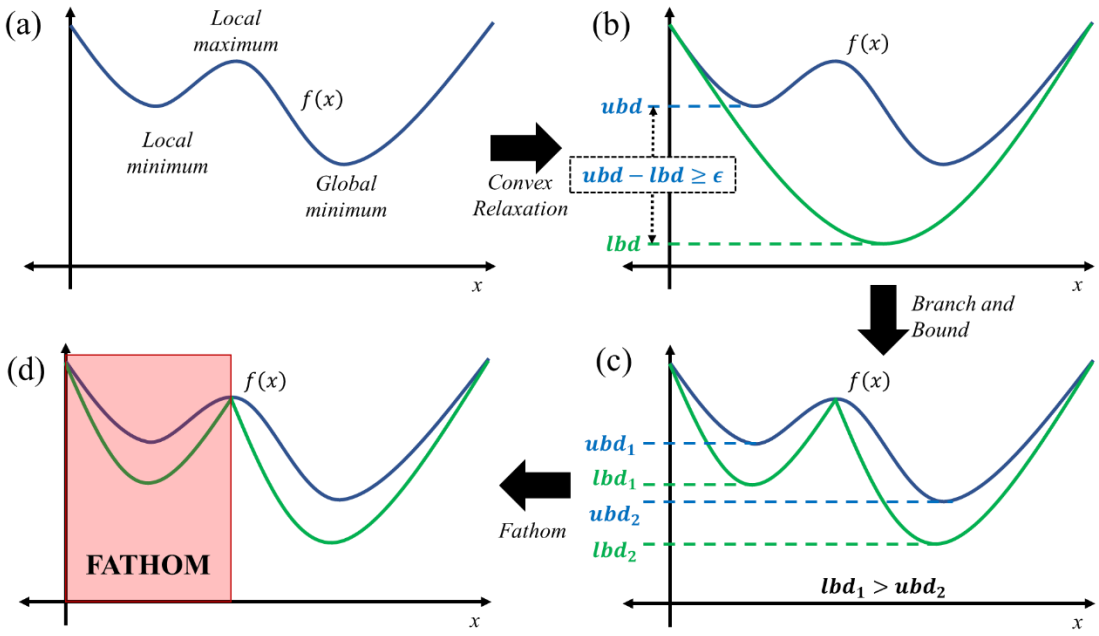

Figure 2.12. General algorithm for using convex relaxations for global optimization.

When discussing global optimization strategies, it would be remiss not to touch upon the stochastic methodologies that have been developed to find global solutions to optimization problems. This is not intended to be a review of all stochastic methods, but will touch upon the prevalent method found in literature: the Monte Carlo Method and its extension of simulated annealing. The Monte Carlo search method uses a large

sample of "randomly selected" points to search the feasible region, comparing the functional value at each successive point, and keeping the lesser value of the function [5]. This algorithm uses an exponential barrier to select search directions and allow the search to continue uphill with some desired probability *P*, as seen in Equation 2.30. The parameter of the exponential barrier can be used as a tuning parameter for the method. Alternatively, this method can be coupled with a local Newton's method search at various points to avoid using a very large sampling number and deterministically find the locations of the various local minima to compare.

$$\exp\left[-\frac{f(x_{k+1})-f(x_k)}{T_k}\right] \geq P \tag{2.30}$$

One variation of the Monte Carlo method is simulated annealing, a method first used to study the thermodynamics of phase transitions in crystals [4]. This method utilized the same structure as the Monte Carlo method but reduces the "annealing temperature" T_k is slowly reduced as the search is completed. This allows the algorithm to explore the feasible region in the early search but then slowly heads downhill toward a minimum as the search progresses. The key to the simulated annealing strategy is how the parameter T_k is deceased. If the temperature is decreased too quickly, then there is a chance of being stuck in a local minimum that is akin to be trapped in a glass state in a crystal. However, if the temperature is not reduced quickly enough then this method is no better than the general Monte Carlo search method.

2.3. Areas of Interest in Process Systems Engineering

In chemical engineering, optimization algorithms like those discussed in detail in

Section 2.2 are used in the study of process systems. More specifically, nonlinear optimization strategies are used to develop comprehensive models of chemical reaction kinetics, fluid dynamics, and mass and energy balance models, among other phenomena. These models are then used to design large-scale, commercial processes to produce commodity chemicals, specialty chemicals, and energy. Finally, once processes have been designed and integrated, process systems engineers are tasked with its safe operation that guarantees product quality by developing successful control strategies.

Each of these process systems areas require the development and solution of a usually nonlinear, nonconvex mathematical program that utilize the methods previously discussed to determine optimal operating conditions. Often it is desired to find the optimal solution to the program as this results in the highest profit margin, lowest environmental impact, or lowest safety risk. More specific formulations and a discussion on specific applications of each topic is discussed in more detail in the next sections.

2.3.1. Process Modeling: Developing Insight to Complex Phenomena

At the heart of chemical engineering practice, and more specifically process systems engineering application, lay mathematical models that capture the principles of these phenomena. These models form the foundation upon which all chemical processes are analyzed and designed by examination and fitting of experimental data to derive a mathematical description of a process. While accurate models are needed, they do not necessarily rely on a first-principles or knowledge-based approach that requires the models to follow an understood and accepted formulation. In fact, data-driven and

empirical modeling is often capable of achieving the required accuracy through deriving correlations in process data. This section will be focused on the ideas of first-principle models and data-driven models, the benefits and challenges faced with each, and recent applications of both modeling strategies.

2.3.1.1. Knowledge-based Modeling

First-principles or knowledge-based modeling is the most commonly experienced form of chemical process modeling. This method begins with the formulation of a model of a known functional form $f(x,p)$ as a function of the input data x and the adjustable model parameters p. An unconstrained nonlinear program like that shown as Equation 2.31 is developed to minimize the differences between experimentally determined data points and the model predictions. Additional constraints can be added to restrict the feasible values of the parameters p based on prior knowledge. Solutions to this program can be found using Newton's method previously discussed.

$$\min_{p} z = \sum_{i=1}^{N} \left(y_{exp,i} - f(x_i, p) \right)^2 \tag{2.31}$$

While knowledge-based modeling is the most prevalent form of mathematical modeling in literature, it suffers from a few drawbacks based on the selection of $f(x,p)$. Often the chemical or physical processes being investigated is not completely understood, making determining the function form of $f(x,p)$ difficult. Mathematical models of a preconceived form may also suffer from complexity issues; models can either be too simple, therefore not capturing all necessary phenomena, or too complex, resulting in large, complicated models that over-fit the data and misrepresent the studied

phenomena [69, 70]. Methods can be used to assess the level of overfitting or underfitting, including local and global sensitivity analysis techniques that assess the "correctness" of both the fit parameters p and the function $f(x, y)$, respectively [71-73]. Despite these drawbacks, first-principles models provide structured models that can be easily utilized in process design and control applications.

2.3.1.2. Data-Driven Modeling

As an alternative to knowledge-based models, researchers have proposed using data-driven and empirical models to describe chemical processes. Data-driven models forgo using models of set functionality, instead using data analysis to derive general models that capture trends and correlations. One important method for data-driven modeling is the artificial neural network (ANN), a regression-based method that represents the input-output behavior of the data after estimation of the model parameters, a process known as training.

Neural network models try to simulate the ability of the brain to process information by constructing a set of nodes, called neurons, that systematically process input data to produce output data. Each neuron takes the incoming information, performs some operation on it, and passes the output result to the next neuron. The typical operation performed on the data in each node is the nonlinear weighted sum, shown in Equation 2.32, where the activation function *K* represents a predefined function such as a sigmoidal function or a hyperbolic tangent function and w_{ij} and $\widehat{w}_{jk}$ represent the weights between nodes [74].

$$y_j = K\left(\sum_i w_{ij} x_i\right) \tag{2.32}$$

The connection between neurons in an ANN is shown in Figure 2.13 below and begins with a set of input neurons, displayed in green, to which the input data x is introduced. In the center of the network is one or more layers of hidden neurons, shown in blue, and the network is completed with a layer of output neurons displayed in red. The ability to change the number of hidden layers, the activation function at each node, and the interconnectivity between each node make ANNs very powerful modeling tools. In chemical engineering, these types of models have been used in process identification and process control applications, among others [75-85].

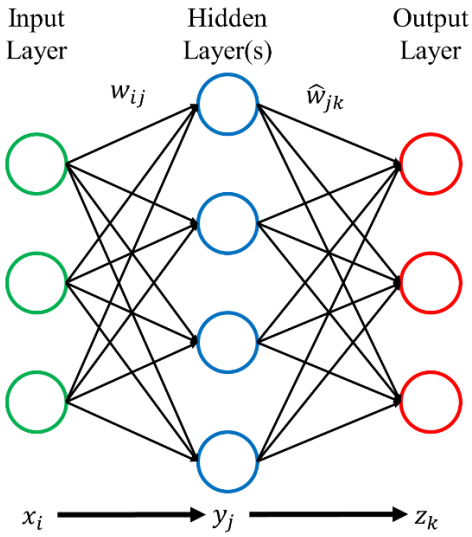

Figure 2.13. Artificial neural network representation for the development of process models.

2.3.2. Process Synthesis: Deterministic Methods for Designing New Processes

Once individual processes have been understood through the development of mathematical models, processes can be developed that take advantage of these reaction

and separation phenomena to convert raw materials to produce desired products. However, the sheer number of combinations between possible reaction and separation methods available results in many possible arrangements. Often, processes should be designed that meet some optimal criteria, typically maximizing process profitability, maximizing process productivity or minimizing environmental impact. This idea lends itself nicely to the formulation of a process synthesis problem as a mixed-integer programming problem like those discussed in Section 2.2.3. This formulation includes the objective function describing the desired optimization criteria, energy and mass balance equality constraints, and inequality constraints on the allowable operating conditions on the considered unit operations. Integer variables are included in the formulation to allow for the selection of necessary processing equipment while ignoring the unnecessary equipment in the final design. Using this framework, typical MILP or MINLP solvers can be utilized to find the optimal process topology, optimal operating conditions, and the resulting value of the objective function from many processing alternatives.

Process synthesis strategies have been used to design a variety of systems. Early use of these methods was for the development of optimal subsystems within a process, such as heat exchanger networks or distillation networks [32, 86-93]. As computer power increased, so too did the size of the synthesis problem. In recent years, process synthesis methods have been applied to many processes relevant to the fight against global warming, including carbon capture [94, 95] and processes capable of producing energy from renewable sources [96-100]. Intensified processes that combine reaction

and separation processes are also of interest to the process synthesis community, as these processes reduce the material and energy requirements and result in a cheaper production of commodity goods [101-104]. Ultimately, process synthesis strategies offer a powerful tool toward designing novel, sustainable processes that can reserve natural resources and produce goods at a large but affordable scale.

2.3.3. Process Control: Maintaining Product Quality and Process Safety Standards

Once an optimal process has been designed at steady state operation using process synthesis methods, it is important to maintain this operation in the instance of disturbances. Typical process output looks like Figure 2.14, where the state variables depicted in blue oscillate around a steady state, shown in green. This field of process control looks to use planned perturbations in manipulated variables to offset the effect of process disturbances on the process states and maintain the process states between desired control limits, shown in red. The use of process control strategies also allows controlled shifts in processing set points to increase or decrease production in the wake of market demands.

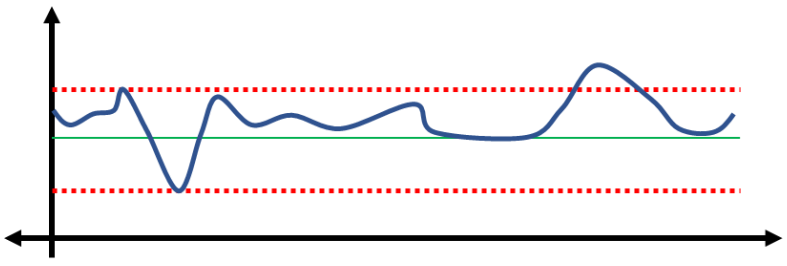

Figure 2.14. Typical variation in process operation when process control is applied to a process output. The red lines denote the allowable control limits, the green line the desired steady state, and the blue line the value of the process state.

Traditional process control strategies, and those widely used in industry, rely on the feedback of the process states $y(t)$ to the controller to determine changes in manipulated variables $u(t)$ to adjust for disturbances $d(t)$ or changes in the set point y_{sp} [105]. The general feedback strategy is shown in Figure 2.15. These process states are compared to the set point and an adjustment in the manipulated variable is calculated from the controller as a function of the error $e(t)$. Standard controllers use a function that considers a combination of proportional, integral, and derivative terms of the error to calculate the change in the manipulated variable, with the specific functionality shown in Equation 2.33.

$$u(t) = Ke(t) + \frac{K}{\tau_I}\int_0^t e(\tau)\, d\tau + k\tau_D \frac{de(t)}{dt} \tag{2.33}$$

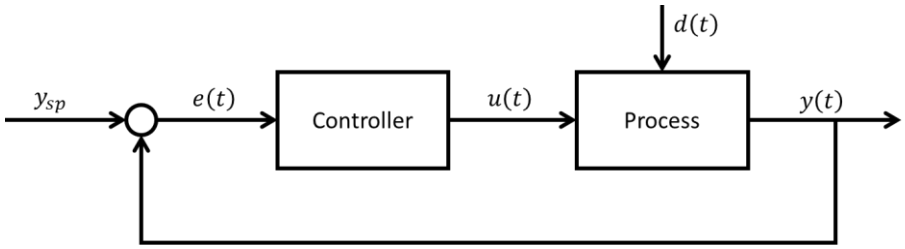

Figure 2.15. Standard feedback control strategy. State outputs y(t) are compared to a set point y_{sp} to calculate an error e(t). The error is used to calculate a manipulated input u(t) that adjusts the process for disturbances d(t) in process operation.

The main idea of classical control theory, as described above, is to improve the closed-loop stability of the chemical process to maintain a desired steady state; the choice of the steady state at which to operate is left as a decision to the user. However, there is incentive for determining the optimal steady state based on profitability, productivity, or safety. In recent years, determining these optimal steady states has been

achieved by formulating the problem as a dynamic optimization problem and solving for the desired manipulated variables *a priori*. The dynamic optimization problems formulated from optimal control problems are solved using the discretization methods discussed in Section 2.2.4. Though operating in an open-loop fashion, and therefore not able to consider process disturbances like that of the closed-loop classical control formulation, these optimal control problems provide the target set points for which closed-loop control can be utilized. After the set points have been determined, classical control methods or more comprehensive model-based control (MPC) strategies like those shown in Figure 2.16 can be used to maintain operation at this optimal steady state [106, 107].

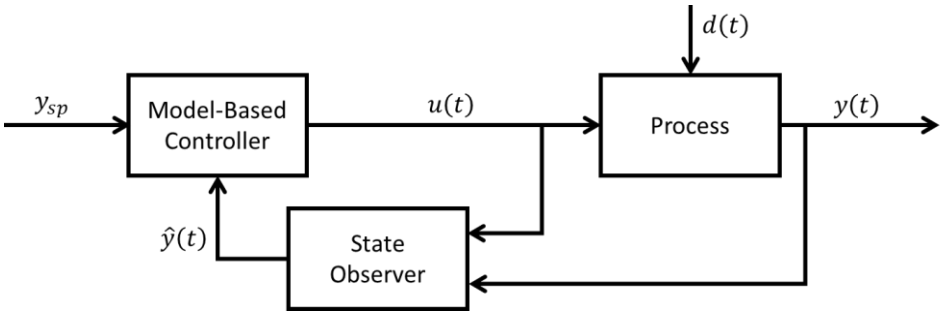

Figure 2.16. Model-based control (MPC) strategy for closed-loop process control. Measurements y(t) and inputs u(t) are used by a state observer to estimate the system states $\hat{y}(t)$. A state-feedback controller is used to generate changes in the manipulated variables u(t) based on set point y_{sp} and state estimation information.

Optimal control methodologies have been used in chemical engineering for a variety of challenges. Batch processes were among some of the first applications of optimal control strategies as it was determined that dynamically adjusting the temperature of these reaction and separation processes could result in higher productivities [52, 108-111]. Many other separation processes, particularly distillation

and crystallization, have been shown to benefit from manipulation of available process parameters to improve product recovery and reduce energy requirements [112-114]. More recently, bioprocesses have been the focus of research in optimal control, as small perturbations in temperature can lead to an enhancement in fermentation yields for batch and fed-batch processing methods [115-117]. Ultimately, dynamic optimization optimal control provides an important tool for studying chemical engineering problems by achieving the full potential of process systems.

2.4. Conclusion

This chapter has illustrated the many algorithms available for solving optimization problems of varying complexity. Detailing the ability to solve mathematical programs with nonconvexity, nonlinearity, equality and inequality constraints, and integer variables allows the area of process engineering to model and optimize intricate chemical processes. To this point, the studies of process systems engineering have been mainly focused on traditional, continuous processes that deliver commodity chemicals or energy in the form of fossil fuels. However, the development of alternative processes based on the emergence of biological production strategies has gained interest among the engineering community. The next chapter will discuss the potential benefits and underlying challenges found in bio-based processing strategies, with emphasis placed on the renewable energy and pharmaceutical industries.

CHAPTER III

BIOCHEMICAL PROCESSES AS PLATFORMS FOR SUSTAINABLE ENGINEERING

3.1. An Introduction to Biochemical Processing

Throughout history, humans have been reliant on biological products. The understanding of standard agricultural practices resulted in an increase in available food resources and the settling of a traditionally nomadic people. Many plants and microbes produce chemicals that are medicinal in nature, providing treatments to many illnesses and increasing the average life expectancy. Biological materials have also been used as a source of fuel, beginning with wood used for fires and steam engines and eventually shifting to coal and oil used to provide electricity and power internal combustion engines. Faced with present day challenges of global warming, declining energy availability, and rising health care costs, humans may need to again rely on biological methods to develop sustainable systems that increase availability of energy and medicinal resources.

In this chapter, the opportunities and challenges of harnessing the power of biologically produced molecules are discussed. Attention will be paid to the scientific advancements in genetic engineering and the resulting impact on the understanding of cellular metabolism and enzyme kinetics. Additional focus will be placed on the use of these advancements to develop biochemical processing methods, defined as any chemical process that uses complete living cells or their components in the conversion of a raw material to a final purified product, to increase availability of sustainable, clean

sources of energy and pharmaceuticals. Challenges innate to biochemical systems that rely on batch operation will be addressed, and the opportunities for continuous bioprocess development using process systems engineering strategies will be discussed. This interconnection between bioprocessing and optimization will set the stage for the work presented in the following chapters.

3.2. Fundamentals of Biochemical Processes

To effectively utilize biochemical methods in chemical processes, it is important to understand the fundamentals of cellular functions. Since the discovery of penicillin by Alexander Flemming in 1928, engineers have been looking to understand and utilize biochemical processes to produce necessary proteins and chemicals [118]. In the decades following, microbial systems have been used to produce enzymes such as catalase, cellulase and hemicellulose, and lactase, or primary and secondary metabolites such as ethanol, citric acid, vitamins, and various antibiotics. Genetic engineering has also been used to develop recombinant organisms capable of producing metabolites not native to the species, including insulin, streptokinase, and various vaccines [119-125].

This section describes the fundamental aspects of biochemical processes used to produce high-value pharmaceuticals and commodity chemical products. First, cellular metabolism, including the critical reaction pathways for aerobic and anaerobic growth, is discussed. Methods for systematically reprogramming these metabolic pathways by modifying the genetics of host microbial systems to produce non-native products are summarized. These methods are the pivotal discovery in microbiology for making biological processes viable. A mathematical understanding of these biochemical

interactions is needed, whether genetically altered or natural. To this end, modeling paradigms for both cellular metabolism and enzyme catalysis are presented. Finally, recovery and purification methods for biologically derived products will be briefly explored as the impetus for driving changes in the paradigm of biochemical processing practices.

3.2.1. Cellular Metabolism and Genetic Engineering: Reprogramming Cellular Processes to Produce Desired Products

3.2.1.1. Cellular Metabolism

Cellular metabolism is the set of all biochemical reactions that are necessary for the production and utilization of energy within the cell. These reactions convert a substrate, such as glucose, to a variety of organic molecules necessary for cellular function and growth, such as nucleotides, amino acids, fatty acids, sugars, and energy in the form of adenosine triphosphate (ATP) needed to power the synthesis of these vital molecules. The reactions of cellular metabolism can be broken down into two categories based on the presence of oxygen: aerobic when oxygen is present and anaerobic otherwise. The mechanisms for each category will be briefly discussed to achieve an understanding of the complexity of cellular metabolism. As many desired products, including the cells themselves, are the result of these reaction pathways, it is important for process systems purposes to develop models to quantify the vast array of reaction networks found in the cell.

Regardless of the presence of oxygen, the beginning of the metabolic process using a glucose substrate is the Embden-Meyerhof-Parnas (EMP) pathway, also known

as glycolysis [126]. This pathway produces a small amount of energy for future cellular reactions; however, the more important result of glycolysis is the conversion of a single molecule of glucose into two molecules of pyruvate through multiple enzymatic steps, a key metabolite that will be used as a raw material to produce many desired chemicals. Under anaerobic conditions, the pyruvate produced during glycolysis will be converted to organic acid or alcohol products. Under aerobic conditions, however, these pyruvate molecules are used for the acylation of coenzyme-A to form the molecule acetyl-CoA.

The reduction of acetyl-CoA to the molecule citrate serves as the entry point into the tricarboxylic acid (TCA) cycle, also known as the Krebs cycle. The primary goal of the TCA cycle is to produce the reduced form of the molecules nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide (FADH), providing electrons necessary for powering future biosynthetic pathways and ATP generation during aerobic respiration. The cyclic conversion of citric acid in the TCA cycle also provides the carbon skeletons for amino acid synthesis, which play a critical role in the production of proteins; a specific type of protein, enzymes, will be discussed later due to their relevance to biochemical processing applications.

Under anaerobic conditions, the pyruvate produced during glycolysis takes an alternative path to produce the energy required for biochemical synthesis. Due to the lack of oxygen that acts as an electron acceptor during aerobic growth, pyruvate molecules are converted to their final product through a series of balanced reduction-oxidation reactions for energy production. This series of reactions, termed fermentation, often result in many commercially relevant products, such as ethanol, butanol, acetic

acid, lactic acid, isopropanol, and glycerol, among others. A general flowchart of cellular metabolism, including the aerobic and anaerobic pathways, is shown in Figure 3.1.

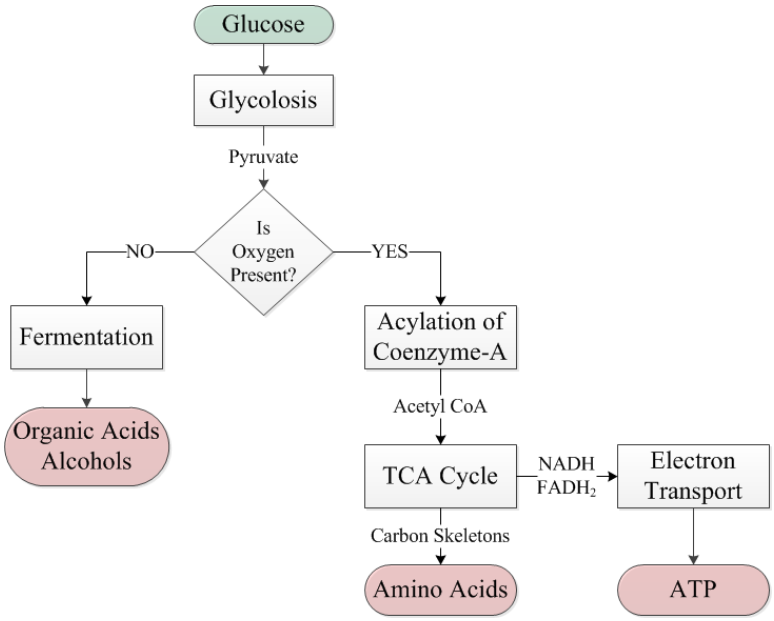

Figure 3.1. Generic flowchart of the metabolic reactions occurring in the cell.

3.2.1.2. Genetic Engineering

Even with an understanding of cellular metabolism, its use in chemical processes would require the selection of a specific organism that can produce a specific product at the necessary quantity to be economically viable. This seemingly impossible task would be much more realistic if the genes for the desired product from one species could be expressed in another, allowing for the design of bioprocesses around cell strains with known qualities but which produce non-native proteins. It would be even more helpful if modifications could be made within an organism that allows for the increased expression of genes that promote production and the decreased expression of genes for alternative pathways. With recent discoveries in microbiology, these what-if scenarios have become

a reality, allowing for the expression of foreign genes in specified host cell lines. Two main methods are utilized to achieve this gene expression: inserting a circular piece of DNA known as a plasmid into a cell, and directly modifying the genomic DNA of the cell. The use of a plasmid allows for the over-expression of a gene by adding multiple copies of the desired DNA sequence in a single construct, potentially resulting in higher production rates of desired proteins or metabolites, but is susceptible to stability issues from either being "kicked-out" of the cell or losing function after a certain number of replications. Alternatively, the direct modification of the genome provides a more stable change by altering the chromosomal DNA directly at the cost of an increased difficulty of adding multiple copies of the desired gene, resulting in a lower level of expression. While not specifically a focus of the work within this thesis, the understanding of cellular genetics and the tools developed for genetic engineering are vital to the success of any biochemical engineering process. As such, a brief history of the field, including an explanation of the two methods for genetic modification, will be given.

The process of genetic engineering using plasmid expression consists of four general steps: (1) identification, isolation, and amplification of the gene of interest from a cell line, (2) combination of the gene with a promoter region, termination region, and selectable marker gene into a plasmid construct, (3) transformation of the desirable cells with the constructed plasmid, and (4) selection of cells with successful transformation based on the choice of the selectable marker. Genetic engineering using plasmids does not affect the genomic DNA of the cell but instead inserts a new construct of stable

DNA that can be utilized to produce compounds not native to the chosen cellular species.

Identification and isolation of the desired gene begins with the use of restriction enzymes, or enzymes that cut DNA at specific sights called restriction sights [127]. After this digestion process fragments of DNA of varying sizes are produced, and the identification of the desired fragment and further generate of more copies of the fragment is needed, a process called amplification. Assuming the nucleotide sequence of the DNA is known for the desired gene, electrophoresis can be used to separate the fragments and the fragment of the desired length can be isolated [128]. After isolation, polymerase chain reaction (PCR) is used to amplify the gene of interest by using thermal cycling to promote DNA deconstruction followed by DNA reconstruction processes [129, 130].

With the gene of interest isolated, it now needs to be placed in a stable structure that can be introduced into the desired cell line. This structure, termed a construct or plasmid, often contains a gene promoter region, a gene terminator region, and an additional gene called a selective marker [126, 131, 132]. The promoter region acts as a binding site for the enzyme responsible for DNA transcription into RNA, RNA polymerase, and its associated transcription factors. These regions also allow modulation of the rate at which a certain gene is transcribed, which in turn modulates the level of production of a specific product. The terminator region, conversely, acts as a signal for the transcription process that the complete sequence of the gene has been transcribed. The last piece of the construct, the selective marker, is included to indicate the success of

the upcoming transformation into the species of choice. Often a gene for an antibiotic resistance, the existence of this marker changes the phenotype of the engineered cell line; a successfully modified cell with an antibiotic resistance will survive in the presence of typically lethal antibiotics while unmodified cells will die [133-136]. By combining the desired gene, promoter, terminator, and selective marker into a complete construct using enzymes called DNA ligase enzymes, the final plasmid can be introduced to the chosen cell line [137].

The process of introducing the plasmid construct into the desired cell line is called transformation and occurs naturally in some bacterial cell lines [138, 139]. However, most nonbacterial cell lines require some external mechanism to accept the plasmid. One method is to treat the cells with a salt wash and heat treatment, opening pores within the cell wall to allow for diffusion of the plasmid [140, 141]. An additional method to open pores to allow cellular transformation is to introduce an electric field to the cells, a process known as electroporation [140, 142]. After using either one of these transformation techniques, the success of the transformation can be verified using the selective marker, successive generations of the transformed cells can be grown to stabilize the transformed system.

The alternative to introducing a plasmid is to directly modify the genome of the cell. There have been many methods developed to accomplish this, but the most recently developed CRISPR/Cas method is the simplest and most efficient method yet to target the modification of the chromosomal DNA directly [143]. The CRISPR/Cas method is designed to mimic the immune system of single-cell organisms that uses a two-

component system of guide RNA (gRNA) and a Cas protein to provide resistance to foreign genetic elements [144, 145]. Genetic alterations to the target DNA can be made simply by altering the genomic sequence of the gRNA, which in turn guides the Cas protein to the location on the DNA that is to be modified. This targeted, stable approach to genetic modifications has made it the preferred method for gene knockouts, gene repression, and gene activation. Additionally, scientists have developed ways to edit multiple sites on the target DNA by using multiple gRNA sequences in a single CRISPR/Cas complex [146, 147]. The CRISPR/Cas method has already been used for many applications, and the mechanisms behind this revolutionary tool are still being studied and developed for uses in more systems [148-152]. Ultimately, the development of the genetic engineering tools discussed have been imperative to the bioprocessing community, allowing modifications at the cellular level to modulate the productivity of large scale systems at the cellular level, adding to the modifications and methods already available at the process level.

3.2.2. Developing Models for Bioreactions: Converting Raw Materials to Desired Products

3.2.2.1. Methods for Quantifying Metabolic Processes

To use the knowledge of both aerobic or anaerobic cellular metabolism and genetic engineering to develop large-scale bioprocesses, models are required that capture the complexity of cellular metabolism shown in Figure 3.1. One modelling method is to use chemical structured models that preserve the stoichiometry of all reaction pathways in the description of the kinetics. These models are often large sets of dynamic equations

that conserve the mole balances of the entire system of reaction pathways by predicting the flux of metabolites, denoted v, through the reaction pathways [153, 154]. An example of a single flux balance equation is shown in Equation 3.1, with the S denoting the stoichiometry for the specific reaction. A steady-state assumption is often applied to these large sets of flux balance equations, allowing for the determination of the flux values by solving the resulting linear set of equations. While these methods are very detailed and have found applications in predictive genetic engineering, they are often too complicated to be useful in biochemical process designs [155, 156].

$$\frac{\mathrm{dx}}{\mathrm{dt}} = \mathrm{S} \cdot \mathrm{v} = 0 \tag{3.1}$$

When considering models for biochemical process design, it is often easier to capture the dynamics of macroscopic, measurable components without considering the vast networks of intracellular reactions. These models, termed unstructured models, only consider the biomass growth, substrate consumption, and production of major desired products and byproducts. The equation for biomass growth is shown in Equation 3.2, where the rate of change of biomass formation is autocatalytic at a rate proportional to the current biomass concentration X. The proportionality constant μ is termed the growth rate and is a function of the substrate concentration S as shown in Equation 3.3. The first term in the Equation 3.3, called the Monod equation based on his work on bacterial culture growth in 1949, is necessary to describe the stages of cellular growth: (1) the lag phase where the cells are adjusting to their environment, (2) the exponential growth phase, and (3) the stationary phase encountered on the depletion of the substrate [157]. The second term in the equation is used to model the death phase that follows the

stationary phase. A pictorial description of the cellular growth phases is shown in Figure 3.2.

$$\frac{\mathrm{dX}}{\mathrm{dt}} = \mu X - k_d X \tag{3.2}$$

$$\mu = \frac{1}{X}\frac{dX}{dt} = \frac{\mu_{max} S}{K_S + S} \tag{3.3}$$

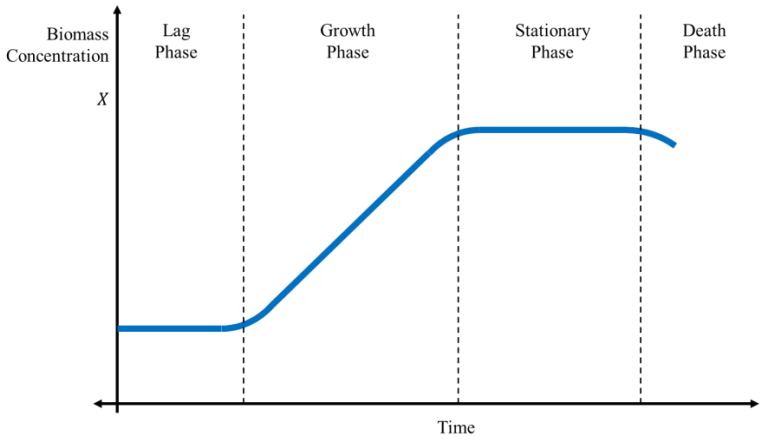

Figure 3.2. Typical stages of biomass growth in batch culture.

Cellular growth is accompanied by substrate consumption and formation of other metabolic products. Equations 3.4 and 3.5 are unstructured modeling equations used to describe these phenomena. The rate of substrate depletion occurs proportionally to the rate of biomass formation μX; the proportionality constant is equal to the reciprocal of a yield coefficient Y_{XS}. Likewise, the rate of product formation is also proportional to the amount of biomass in the culture broth. However, unlike the substrate utilization, the amount of product formed may be growth-associated, or a function of the growth rate μ, or non-growth. This is accounted for in Equation 3.5 with the first and second term, respectively, where the coefficients α and β serve as proportionality constants. When coupled, Equations 3.2, 3.3, 3.4, and 3.5 serve as a complete macroscopic description of

cellular metabolism and is much more conducive to analyzing, designing, and controlling biochemical processes.

$$\frac{dS}{dt} = -\frac{\mu X}{Y_{XS}} \tag{3.4}$$

$$\frac{dP}{dt} = \alpha\mu X + \beta X \tag{3.5}$$

Equations 3.2 through 3.5 describe the growth of cellular species in an ideal environment. It is often the case, however, that cellular growth is hindered due to a high substrate or product concentration, a process known as substrate or product inhibition. Many substrates, including glucose, and desirable products formed in the anaerobic pathway of cellular metabolism, including ethanol, butanol, and acetic acid, are often inhibitory to cellular growth at high concentrations [126]. Cellular processes may also be downregulated due to a toxic component in the growth media. Cellular inhibition can be classified as noncompetitive or competitive regardless of its cause, and the Monod equation used for describing the growth rate must be modified to account for this inhibition. Equation 3.6 below is the common modification for noncompetitive substrate or product inhibition, where I is the concentration of the inhibition compound, which may be the substrate, product, or toxic compound, and K_I is the inhibition parameter. A similar equation is available for competitive inhibition, which is shown as Equation 3.7.

$$\mu = \frac{\mu_{max} S}{\left(1 + \frac{K_S}{S}\right)\left(1 + \frac{I}{K_I}\right)} \tag{3.6}$$

$$\mu = \frac{\mu_{max} S}{K_m\left(1 + \frac{I}{K_I}\right) + S} \tag{3.7}$$

3.2.2.2. Methods for Quantifying Enzymatic Reactions

Certain special products of cellular metabolism that garners interest in bioprocess design are enzymes. While many enzymes are used within the cellular metabolism pathways, these proteins are of interest in bioprocesses for their ability to catalyze chemical reactions. Like transition metals used for traditional chemical catalysts, enzymes work by lowering the activation energy required to convert a raw material to a desired product. Enzymatic reactions typically proceed at a much higher rate than chemically catalyzed reactions under ambient conditions due to their specificity, versatility, and effectiveness, a characteristic that can be utilized to design more efficient processes [126]. Simple enzyme catalysis is defined by the chemical pathway shown in Equation 3.8 below, where *ES* is the complex formed between the enzyme and its specific substrate in a first step followed by the formation of the product *P* and release from the enzyme *E*.

$$E + S \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} ES \overset{k_2}{\rightarrow} E + P \tag{3.8}$$

To efficiently implement the benefits of enzymes in chemical processes, a mathematical description of these kinetics is required.

The first mathematical description for enzymatic reactions was developed by Henri in 1902 and again by Michaelis and Menten in 1913 [158, 159]. These mathematical models describe the rate of formation of the product *P* and are qualitatively like the Langmuir-Hinshelwood kinetics that describe traditional metallic catalysis reactions and the Monod equation previously discussed that describes the cellular growth rate. Two methods can be used to develop the kinetic equations, the

initial rapid equilibrium assumption developed by Henri and Michaelis and Menten, and the quasi-steady-state assumption later developed by Briggs and Haldane in 1925 [159, 160]. Both methods begin with time-dependent models of the product and enzyme-substrate complex concentrations shown in Equations 3.9 and 3.10 below. An additional mass balance is used to account for the total amount of enzyme in the system, which is conserved throughout the reaction. In the equilibrium assumption, the definition of the equilibrium constant is used along with Equation 3.11 to determine an algebraic solution forES; this is then substituted into Equation 3.9 to successfully derive Equation 3.12. Alternatively, the quasi-steady-state hypothesis can be used by setting Equation 3.10 equal to zero and solving for ES; this method also leads to Equation 3.12 when substituted into Equation 3.11.

$$\frac{dP}{dt} = k_2(ES) \tag{3.9}$$

$$\frac{d(ES)}{dt} = k_1(E)(S) - k_{-1}(ES) - k_2(ES) \tag{3.10}$$

$$E = E_0 - ES \tag{3.11}$$

$$\frac{dP}{dt} = \frac{V_m S}{K_m + S} \qquad V_m = k_2 E_0 \tag{3.12}$$

The model in Equation 3.12 describing the production of product P is a function of the substrate concentration S and the total enzyme concentrationE_0. It also contains two parameters that need to be estimated, V_m and K_m. While the parameter estimation procedures discussed in Chapter II can be used alongside experimental data, three other methods are available based on restructuring Equation 3.12 into different linear forms

and comparing to linearized data. The resulting models after each linearization method is shown as Equation 3.13, 3.14, and 3.15 and their corresponding linear regression plots are referred to as the Lineweaver-Burk plot [161], the Eadie-Hofstee plot [162, 163], and the Hanes-Woolf plot [164, 165], respectively. In each plot the values of K_m and V_m are found as either an x-intercept, a y-intercept, or the slope of the regression. Examples of the three data fitting methods are shown in Figure 3.3. below.

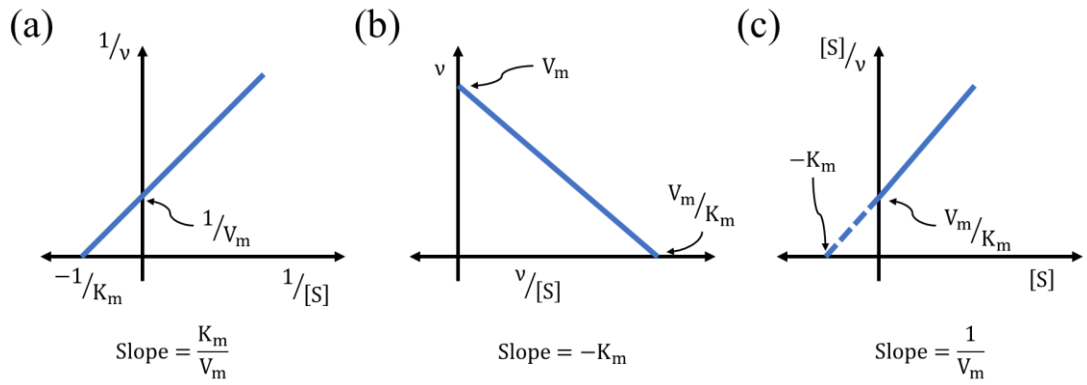

Figure 3.3. (a) Lineweaver-Burk, (b) Eadie-Hofstee, and (c) Hanes-Woolf linear regression methods for the determination of Michaelis-Menten parameters K_m and V_m.

$$\frac{1}{\nu} = \frac{1}{V_m} + \frac{K_m}{V_m}\frac{1}{S} \qquad \nu = \frac{dP}{dt} \tag{3.13}$$

$$\nu = V_m - K_m\frac{\nu}{S} \qquad \nu = \frac{dP}{dt} \tag{3.14}$$

$$\frac{S}{\nu} = \frac{K_m}{V_m} + \frac{S}{V_m} \qquad \nu = \frac{dP}{dt} \tag{3.15}$$

To this point methods for modeling simple enzyme kinetics have been discussed, but enzymatic processes are often more complex. Many enzymes have multiple substrate binding sites that must be bound in specific order for the reaction to proceed, a process known as allosteric or cooperative binding [126]. Like cellular systems, enzymes are

often inhibited by binding with certain compounds other than the desired substrate; however, inhibition in enzymatic processes is accounted for in the product formation equation directly instead of through the growth rate. Some compounds such as heavy metals will result in permanent inhibition when binding with the enzyme active site, while others will bind reversibly [126]. While the methodology used holds for these more complex systems, inhibition will result in slight variations in Equation 3.12. Allosteric enzymes will result in a cooperativity coefficient n being introduced as an exponent to the substrate concentration $[S]$.

Reversible inhibition will also result in changes to Equation 3.12 based on how the inhibition occurs. The first method of reversible inhibition is competitive inhibition, where the inhibitory molecules compete with the substrate for the active site of the enzyme and results in a modification of Equation 3.12 shown in Equation 3.16.

$$\begin{matrix} E + S \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} ES \overset{k_2}{\rightarrow} E + P \\ E + I \overset{K_I}{\Longleftrightarrow} EI \end{matrix} \qquad \frac{dP}{dt} = \frac{V_m S}{K_m\left(1 + \frac{I}{K_I}\right) + S} \tag{3.16}$$

Alternatively, noncompetitive inhibitors do not bind to the active site but rather another site on the enzyme, hindering the binding of the substrate molecule by causing a deformation of the substrate active site; this results a modified equation shown in Equation 3.17.

$$\begin{matrix} E + S \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} ES \overset{k_2}{\rightarrow} E + P \\ E + I \overset{K_I}{\Longleftrightarrow} EI \overset{K_m'}{\Longleftrightarrow} ESI \Longleftrightarrow ES \end{matrix} \qquad \frac{dP}{dt} = \frac{V_m S}{K_m\left(1 + \frac{I}{K_I}\right) + S} \tag{3.17}$$

In certain cases, uncompetitive inhibition occurs, which is the selective binding of an inhibitor to the enzyme-substrate complex, preventing the reaction from proceeding. Equation 2.18 is the result of these additional uncompetitive inhibition reaction pathways.

$$E + S \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} ES \overset{k_2}{\rightarrow} E + P \qquad ES + I \overset{K_I}{\Longleftrightarrow} ESI \qquad \frac{dP}{dt} = \frac{\left(\frac{V_m}{1 + \frac{I}{K_I}}\right) S}{K'_m \left(\frac{1}{1 + \frac{I}{K_I}}\right) + S} \tag{3.18}$$

Finally, a high substrate concentration can result in inhibition. This results in a modification to the Michaelis-Menten equation as shown in Equation 3.19.

$$E + S \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} ES \overset{k_2}{\rightarrow} E + P \qquad ES + S \overset{K_{S_I}}{\Longleftrightarrow} ES_2 \qquad \frac{dP}{dt} = \frac{V_m S}{K_m + S + \frac{S^2}{K_{S_I}}} \tag{3.19}$$

3.2.2.3. Industrial Bioprocessing Applications for Metabolism and Enzyme Kinetics

Researchers have taken advantage of recombinant systems and enzyme kinetics to design and analyze a variety of biochemical processes, including the production of biofuels, detergents, foods, and pharmaceuticals. Much of the recent interest in enzyme reactions is to produce simple sugars via hydrolysis of starch or cellulose. Enzymes used in this case can hydrolyze and break apart the respective α-linkages or β-linkages between the glucose monomers. Resulting simple sugars are often used in the fermentative production of alternative liquid fuels [166-171]. Economical production processes for several types of enzymes, including cellulases and lipases, are also of great

interest for developing biological detergents [172-175]. In the food industry, enzymes are often used to make products such as high fructose corn syrup, fruit juice, and lactose free milk [176-179]. Work has also been done on more sustainable production methods for enzymes used in pharmaceutical applications, such as protease or penicillin amylase, while also reducing the costs of these medicinal products [180-186]. Ultimately, the understanding of enzymatic pathways and the resulting kinetics is vital to the production of processes that span the fuels, food, and health-care industries.

Many systems have been developed to produce cells as the desirable product for medicinal use. These cells are administered as whole-cell therapies for the treatment of many different health problems. Many researchers have investigated the development of bioprocesses to produce stem cells capable of differentiating into several specialized cell types, providing treatments for many genetic, immunological, and degenerative diseases, as well as a variety of cancers [187-192]. Another system of interest is the *ex vivo* production of red blood cells that can increase the supply of rare blood types available for needed transfusions [193]. Work has also been done on developing cellular therapies for the tissue engineering of skin and for use in therapies for lung diseases [194, 195].

The models and understanding available for enzyme kinetics and cellular metabolism allow the development of biochemical conversion processes to produce many desirable products. However, strategies to separate the desired products from the unconcentrated effluent of the reactor are needed to develop complete chemical processes. Though not a large focus for the work presented in this thesis, for

completeness the next section will briefly introduce the separation processes commonly used for the recovery of biological products.

3.2.3. Bio-separation: Recovering Desired Products from Biological Residues

Once a cell line has been successfully transformed and a bioreactor process has been implemented to produce a desired compound, separation processes are necessary to recover the product in a purified form. One of the challenges of many biological products, specifically those in the pharmaceutical area, is the requirement of highly purified product to allow for the successful administration to patients. Also, while some products are extracellular and found in the culture broth, many intracellular products are not secreted from the cell and require additional processing steps to break open, or lyse, the cells to retrieve the product. The recovery process consists of up to five basic steps depending on the nature of the desired product: (1) cell concentration, (2) cell disruption, (3) isolation of the product and removal most of the water, (4) purification or removal of contaminating chemicals, (5) final preparation [126].

The processing limitations and desired quality of the final product often put a large burden on the separation train, usually consisting of multiple processes in series to guarantee purity. Like traditional chemical engineering systems, separation processes in biochemical process can account for 50-80% of the total manufacturing costs [196-200]. While work has been done to improve the separation processes, the high separation cost is often attributed to the low productivity of bioreactions [201-203]. To improve further upon these separation costs, advancements need to be made in methods for producing the desired product in the bioreactor at a level which increases process productivity.

3.3. Meeting Consumer Demand: The Need for Continuous Manufacturing

Traditional bioprocessing has been performed in a batch mode, where the reaction proceeds for a designated amount of time after initially charging the reactor with a specific amount of raw materials. Once the time has passed, the contents of the bioreactor are emptied to proceed to downstream separation and the bioreactor is cleaned, sterilized, and filled with a new batch of reagents to begin the reaction cycle again.

While currently the industry standard in bioprocess, batch modes of operation face inherent production challenges. Inhibiting and toxic compounds often build up during batch fermentation as the reaction progresses, causing loss of cellular or enzymatic activity; fed-batch production methods have been investigated to mitigate this problem by offering the ability to systematically introduce substrate while simultaneously provide dilution effects; these fed-batch systems often suffers from other, similar disadvantages as batch production, most notably batch-to-batch variability [204-208]. Processes operating in batch and fed-batch mode typically employ multiple reactors operating cyclically to ensure that some product is always available for downstream processing, reducing process down-time and allowing for operation in a pseudo-continuous manner at the cost of increased capital expenditure [209]. This configuration also requires a large operating cost as enough raw materials are needed on-hand to operate multiple reactors simultaneously [210]. For small-scale production, these methods of production are acceptable; however, as product demand increases, conventional wisdom says that batch processing methods are inadequate [211].

For large-scale production, bioprocessing should emulate the continuous processing strategy used by the commodity chemical industries. Continuous processes differ from batch and fed-batch processing methods by introducing raw materials to the reactor via an inlet flow and removing product from the bioreactor via an outlet flow. When equal, the inlet or outlet flowrate is normalized by the volume of the bioreactor and is referred to as the dilution rate D and is synonymous with the space velocity of traditional chemical systems [212]. The inclusion of an outlet stream facilitates the gathering of measurements necessary for implementing online process control and fault detection strategies, mitigating the batch-to-batch reliability issues of batch or fed-batch strategies [209]. Continuous production typically results in decreased capital and operating cost requirements due to the operation of smaller reactors in fewer number to achieve the desired process throughput [213]. Continuous bioprocessing methods have also been shown to be environmentally friendlier by eliminating the need for the cleaning and sterilization step between batches. The United States Food and Drug Administration (USFDA) has been initially hesitant to adopt standards for continuous processing due to challenges with contamination and the ability to successfully allocate "batch numbers" that can help identify the product in case of a processing issue. However, these initial hesitations have since been alleviated as continuous processes have been identified as a method to improve product supply, product quality, process flexibility, and process robustness for biomanufacturing [204, 213]. Figure 3.4 compares the batch and continuous processing paradigms from a process systems perspective.

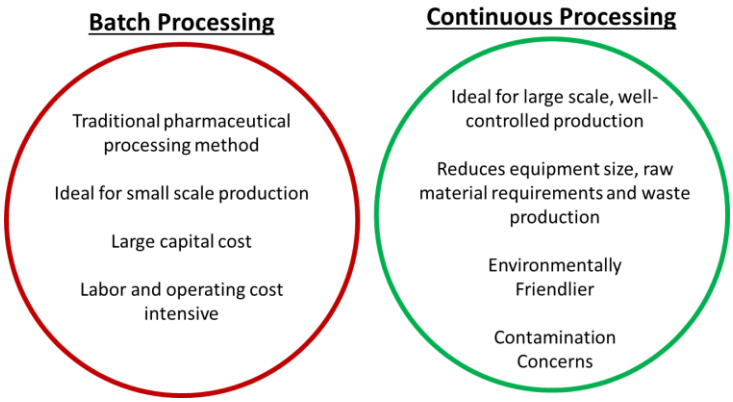

Figure 3.4. Comparison of batch processing and continuous processing methods for biochemical processes.

One of the challenges that continuous bioprocessing must overcome is the requirement of maintaining process productivity when compared to batch systems. The productivity of a biochemical system is defined as the amount of product made per unit volume per unit time and is affected by the dilution rate, as shown in Figure 3.5. At zero dilution rate, continuous processes are in fact batch operations that achieve a high titer, also known as concentration which is displayed as x from the top plot. However, this high titer comes at the cost of a low productivity, shown as xD on the bottom plot, due to long reaction times often associated with bioprocessing. As the dilution rate D increases the product titer slowly decreases while the reactor productivity increases. At a certain dilution rate the productivity reaches a maximum, and beyond that point both the titer and productivity decline. Eventually, both the titer and productivity reach a value of zero; this value of the dilution rate is referred to as the washout dilution rate, where the dilution rate is higher than the growth rate of the cells and growth cannot keep up with the removal of cells from the effluent. The goal of continuous processing is to operate at

the optimal dilution rate that leads to the maximum process productivity at the expense of product titer.

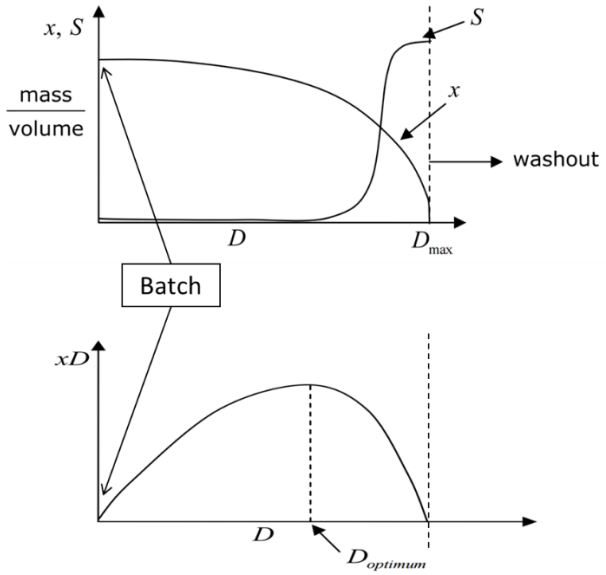

Figure 3.5. Effect of adding dilution rate on the titer and productivity of biochemical reactions.

The potential of continuous bioprocessing has opened the door for the development of novel processes. The ability to use process modeling strategies with the available first-principles models for cellular metabolism and enzyme kinetics suggest that a mathematical understanding of these biological processes can be reached. Process systems engineering strategies in process synthesis are uniquely designed to use these models to determine the optimal process topology from a variety of processing decisions; when used with the availability and versatility of genetic engineering, continuous processing strategies for a variety of products can be investigated. The existence of an optimal dilution rate that maximizes the productivity of these bio-systems invites the use of process control strategies to achieve and maintain operation at this critical point, reducing the stress on the downstream separation and reducing the

cost of producing critical products in the health care and alternative fuels industries. The combination of process systems engineering and biochemical engineering fundamentals provides an excellent opportunity to design and control large-scale, continuous systems to increase the availability and reduce the costs of biologically-derived products.

3.4. Conclusions

In this chapter, fundamentals of bioprocessing have been explored; specifically, how the discovery of genetic engineering introduced increased customizability to cellular strains and how continuous manufacturing practices can provide over batch processing. In Chapter II, the fundamentals of optimization were discussed as tools for the application of the process systems engineering strategies for the modeling, design, and controlled operation of chemical systems. *Using this knowledge, the remainder of this thesis is dedicated to discussing the use of process systems strategies in modeling, design, and control to develop novel methods for producing biological products via continuous processing methods.*

By combining the fundamentals of process systems engineering and the adaptability of bioprocessing, large-scale, well-controlled processes can be designed. Chapter IV and Chapter V utilize process modeling and optimal control techniques to model the production of nutraceuticals from recombinant yeast and subsequently use these models for the design a novel, multi-feed, continuous bioreactor that maximizes process productivity while decreasing process operating costs when compared to the traditional single-feed method. In Chapter VI and Chapter VII, process synthesis methodologies are used to design large-scale, continuous processes for the alternative

liquid fuels bioethanol and biobutanol from lignocellulosic biomass. By focusing on the areas of pharmaceuticals and biofuels, the impact of continuous processing can be exemplified.

CHAPTER IV

MODELING OF BATCH KINETICS OF AEROBIC CAROTENOID PRODUCTION USING *SACCHAROMYCES CEREVISIAE**

4.1. Introduction

Carotenoids, a diverse group of yellow-orange pigments found in many biological systems, are produced by diverse organisms such as plants, fungi, and bacteria [214-216]. Because of its colored characteristic, it has been extensively used in food pigmentation and as constituents in vitamins and dietary supplements [217-220]. Carotenes, such as β-carotene, have important biological roles as a precursor of vitamin A. They have been shown to have positive impacts in human health, having antioxidant effects and properties protective against cancer [221-224].

Presently, some carotenoids are industrially produced by synthetic chemical technology; however, some of the by-products have undesirable side effects when consumed and most of the carotenoids such as β-carotene have a structural complexity that makes the chemical synthesis an unviable option. For this reason, the production of carotenoids from microbial sources has been the focus of vast literature references [217, 225-228].

Overflow metabolism is the incomplete oxidation of the carbon source resulting in the excretion of generally inhibitory organic end products during aerobic operations. This phenomenon, also called the "short-term Crabtree effect", manifests usually at a high substrate concentration [229, 230]. Various microorganisms exhibit this

*Reprinted with permission from "Modelling of batch kinetics of aerobic carotenoid production using *Saccharomyces cerevisiae*" by Ordonez et al., 2016. *Biochemical Engineering Journal*, 114, 226-236, Copyright 2016 by Elsevier.

phenomenon, examples of which include *Saccharomyces cerevisiae* cultures with aerobic ethanol formation [231, 232], *Escherichia coli* cultures with aerobic acetate formation [233, 234], or mammalian cell cultures with aerobic lactate formation [235]. This inhibition effect is undesirable in many processes and leads to complications when trying to quantitatively describe these systems.

Mathematical models describing the kinetics of microbial growth, substrate uptake and product formation are very useful for optimization, process control, the reduction of process operating costs, and the increase of product quality of cultivation processes. In this context, kinetic models have been reported for yeast systems such as *Xanthophyllomyces dendrorhous* [236, 237] but there are very few studies dealing with appropriate modeling approach for engineered S. cerevisiae strains, and more specifically those engineered to produce β-carotene. The aim of this work is to develop a suitable and reliable kinetic model for the β-carotene production in batch cultures of an engineered S. cerevisiae strain using glucose as the main substrate. This model is also applied to predict cell growth, substrate consumption, ethanol and acetic acid formation and later assimilation. Furthermore, sensitivities of various model parameters were studied to elucidate the need to focus on key model parameters.

4.2. Materials and Methods

4.2.1. Microorganism and Culture Media

An *S. cerevisiae* strain mutant SM14 engineered to produced β-carotene was used in this study [238]. The yeast strain was stored in frozen vials at -80℃ and in plates at 4℃ which were subcultured every three weeks for maintenance. Experiments were

conducted to determine the optimal initial glucose loading for all experiments. A single flash of 50 mL YPD and 20 g/L glucose was inoculated with a single colony and grown at 200 rpm and 30℃ overnight until the exponential growth phase was reached. Using this culture, YPD media with varying concentrations of glucose (20 g/L, 40 g/L, 60 g/L, 80 g/L and 100 g/L) were inoculated in a 1:60 ratio that was intended for the bioreactor studies. These shake flask cultures were grown at 200 rpm and 30℃ for 110 hours after innoculation to allow for complete use of glucose. The beta-carotene concentrations were analyzed at the end of this time to determine the optimal initial glucose concentration that would yield the highest amount of β-carotene. Figure 4.1 below shows the productivity results of these experiments.

The results in Figure 4.1 show that an increase in initial glucose concentration shows an inverse relationship with the overall β-carotene titer on a gram-dry-cell-weight basis. Based on these results, it was determined that 20 g/L was optimal initial glucose loading. As such, all experiments were conducted with the cells grown in fresh Yeast Nitrogen Base (YNB) media supplemented with 20 g/L D-glucose.

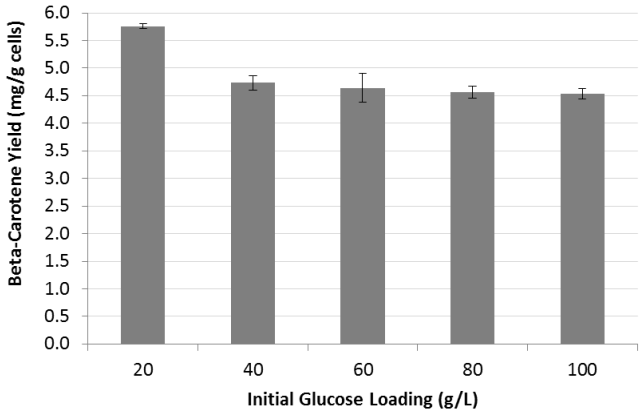

Figure 4.1. Effect of initial glucose concentration on overall carotenoid yield. Experiments were done in YPD media with varying concentrations of glucose (20 g/L, 40 g/L, 60 g/L, 80 g/L and 100 g/L), were inoculated in a 1:60 ratio, and were grown at 200 rpm and 30℃ for 110 hours after innoculation to allow for complete use of glucose.

4.2.2. Bioreactor Cultivation Studies

The bioreactor studies were carried out in a 7 L, glass, autoclavable bioreactor (Applikon®, Foster City, CA) with a 3 L working volume. The bioreactor was inoculated with the entire seed culture. The temperature was set at 30℃, pH was maintained at 4 by addition of 2 M HCL or 2 M NaOH as needed, the agitation speed was kept constant at 800 rpm and the bioreactor was supplied with a constant airflow of 6 L/min. These conditions were optimized experimentally by Olson [239]. All the bioreactor experiments were performed in batch mode carried out in duplicate with duration of 72 hours, the time at which the acetic acid produced was totally consumed. Substrate consumption and product formation and depletion were verified using the HPLC Agilent Technologies 1290 Infinity with the use of the Aminex HPX-87H HPLC column.

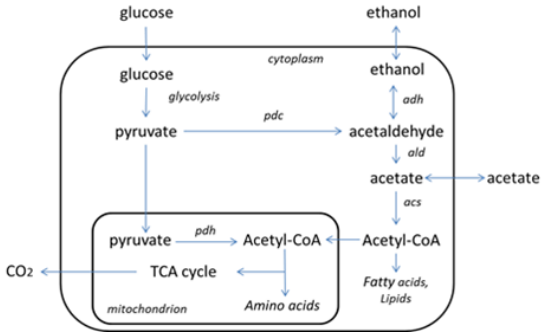

Figure 4.2. Overview of the metabolic pathways found in wild type *S. cerevisiea*. The key enzymes for overflow metabolism are: pdh: pyruvate dehydrogenase complex; pdc: pyruvate decarboxylase; adh: alcohol dehydrogenase; ald: acetaldehyde dehydrogenase; acs: acetyl

The *S. cerevisiae* SM14 culture presents the following characteristics in a stirred-tank bioreactor with 20 g/L initial glucose. Initially, ethanol and acetic acid are produced in the early stage of the batch culture due the overflow metabolism. The production of these metabolites starts with glycolysis, the metabolic pathway where the six-carbon glucose molecule is broken down into two three-carbon pyruvate molecules. At this level, cell respiration via the mitochondrial pyruvate dehydrogenase complex (*pdh*) producing Acetyl-CoA competes with cytosolic pyruvate decarboxylase (*pdc*) reducing it to acetaldehyde [240-243]. Once the acetaldehyde is formed through the pyruvate decarboxylase activity, it may feed the tricarboxylic acid cycle (TCA), reducing the acetaldehyde via acetaldehyde dehydrogenase (*ald*) to acetate and then producing acetyl coenzyme A (Acetyl-CoA) through acetyl-CoA synthetase (*acs*) [243, 244]. Alternatively, acetaldehyde is reduced to ethanol instead of being oxidized to carbon dioxide thru alcohol dehydrogenase (*adh*) [240], as shown in Figure 4.2. When the glucose is consumed, the ethanol present in the medium can be utilized via aerobic metabolism and converted to Acetyl-CoA via acetaldehyde and acetate, the pathway does not proceed via pyruvate. Similarly, once the ethanol becomes exhausted the acetate can be consumed [245].

The produced acetyl coenzyme A can be used in many different biochemical reactions such as the mevalonate pathway shown in Figure 4.3. Beta-carotene is derived from the isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) which can be condensed into geranyl pyrophosphate (GPP), a C_{10} molecule, and is then converted to the C_{15} molecule farnesyl pyrophosphate (FPP) with the addition of an

isoprene unit. The FPP is condensed into geranylgeranyl pyrophosphate (GGPP), a C_{20} compound, catalyzed by *crtE*. The β-carotene biosynthesis starts with the formation of the C_{40} carotenoid phytoene, catalyzed by phytoene synthase (*crtYB*). Isomerization and desaturation reactions allow the formation of lycopene, which forms β-carotene [246-248]. These pathways, highlighted in red in Figure 4.3, are not native and were recombinantly introduced to the genome.

4.2.3. Inhibition Studies

Experiments were performed to determine if the products inhibit the growth rate of the *S. cerevisiae* SM14. The experiments were done in a baffled Erlenmeyer flask containing with 50 mL of YNB media and incubated for 24 h at 30℃ in a shaker at 200 rpm. Cell growth was measured in terms of turbidity at 600 nm. To measure the inhibition effects of the products, e.g. ethanol and acetic acid, experiments were run with 20 g/L of glucose and various concentrations of the products. The ethanol concentration was varied from 0 to 70 g/L and the acetic acid from 0 to 10 g/L, as described in Maiorella *et al* [249]. The maximum growth rate was calculated for each experiment and plotted as a function of the ethanol or acetic acid initial concentration.

4.2.4. Analytical Methods

Cell growth was monitored based on an increase in the OD_{600} of the culture. For the β-carotene quantification the cells were disrupted using the method described by Reyes *et al* [238]. The D-glucose, ethanol and acetic acid concentrations in the media were analyzed by HPLC.

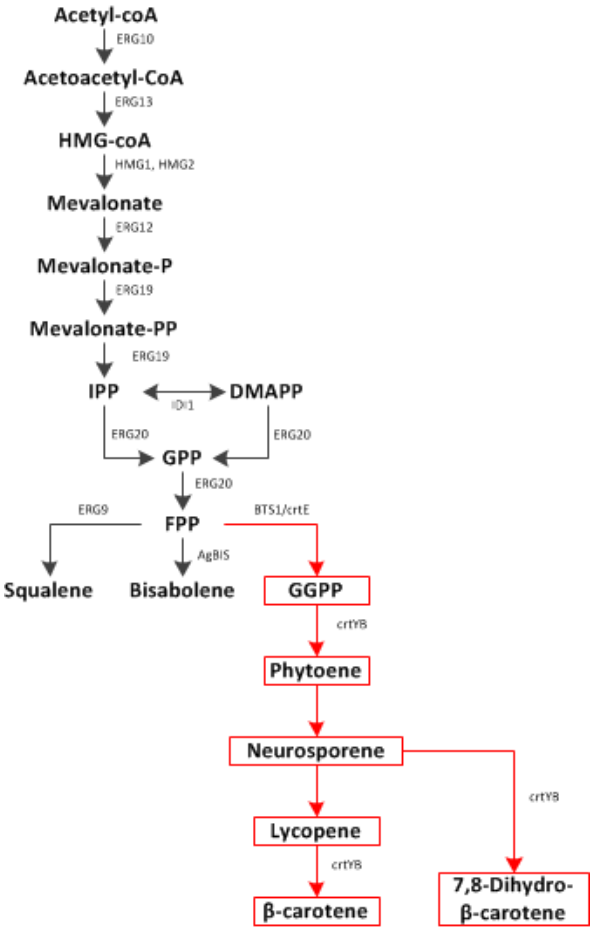

Figure 4.3. Mevalonate pathway genetically engineered via chromosomal integration into the S. cerevisiea SM14 strain. The β-carotene synthesis pathway introduced to the genome recombinantly is highlighted in red [246].

4.2.5. Kinetic Modeling Strategy

4.2.5.1. Kinetic Model

To account for the characteristics of the system and based on literature, a model was proposed to describe the overflow metabolism with the depletion of ethanol and acetic acid.

As glucose, ethanol and acetic acid can be utilized for biomass production, the cell growth rate can be represented by

$$\frac{dX}{dt} = (\mu_G + \mu_E + \mu_A)\, X \tag{4.1}$$

where μ_G, μ_E and μ_A represents the specific growth rate on glucose, ethanol and acetic acid respectively. The Monod equation is the most common unstructured kinetic model for microbial growth, which relates the microbial growth rate to a single limiting substrate [249]. When the medium contains more than one carbon source multiple lag phases may be observed, caused by a shift in metabolic pathways in the growth cycle, a phenomenon called "diauxic growth" by Monod [157]. For mixed substrates, the model requires a modification as described by Yoon et al [250].

$$\mu = \mu_G + \mu_E + \mu_A \tag{4.2}$$

$$\mu_G = \left(\frac{\mu_{max,G} \cdot \chi_E \cdot \chi_A \cdot G}{K_{SG} + G + a_{ge}\, E + a_{ga}\, A} \right) \tag{4.3a}$$

$$\mu_E = \left(\frac{\mu_{max,E}\, E}{K_{SE} + E + a_{eg}\, G + a_{ea}\, A} \right) \tag{4.3b}$$

$$\mu_A = \left(\frac{\mu_{max,A}\, A}{K_{SA} + A + a_{ag} G + a_{ae}\, E} \right) \tag{4.3c}$$

where a_{ij} represents the inhibition effect of the jth substrate on the utilization of the ith substrate by the organism [157]. If $a_{ij} = 1$, the jth substrate has the same inhibition effect as the ith substrate itself on the ith substrate utilization. If $a_{ij} > 1$, the jth substrate has inhibition on the ith substrate utilization. If a_{ij} is much greater than 1, the jth substrate has a repression effect on the utilization of the ith substrate by the microorganism. $\mu_{max,G}$, $\mu_{max,E}$, and $\mu_{max,A}$ are the maximum specific growth rates on glucose, ethanol and acetic acid, respectively. The variables G, E and A are the glucose,

ethanol and acetic acid concentrations. The variables χ_E and χ_A are added to $\mu_{max,G}$ to account for any effect of ethanol or acetic acid inhibition on the glucose growth rate. These variables are functions of the ethanol and acetic acid concentration, respectively, whose functionalities are determined using the inhibition studies discussed in Section 4.3.

The mass balance for substrate governs the production of biomass. The energy requirement for cellular maintenance is very small relative to that of growth [237], for that reason it is neglected in the mass balance calculations. Then, the glucose consumption rate is given by Equation 4.4, where $Y_{X/G}$ is the biomass yield coefficient on glucose.

$$\frac{dG}{dt} = -\frac{\mu_G X}{Y_{X/G}} \tag{4.4}$$

The ethanol production from fermentative catabolism of glucose occurs during the exponential growth phase and is growth associated [126, 251], hence ethanol formation is related to the yeast growth on glucose, shown by the first term of Equation 4.5. The second term represents the amount of ethanol that is consumed by the yeast cells, where $Y_{X/E}$is the biomass yield coefficient on ethanol.

$$\frac{dE}{dt} = k_1 \mu_G X - \frac{\mu_E X}{Y_{X/E}} \tag{4.5}$$

Similarly, the acetic acid model considers the amount produced by the consumption of glucose and ethanol as well as the acetic acid consumed by the yeast cells, where $Y_{X/A}$ is the biomass yield coefficient on acetic acid. We assumed that the acetic acid formation occurred in the growth phase of the cells.

$$\frac{dA}{dt} = (k_2\mu_G + k_3\mu_E)X - \frac{\mu_A X}{Y_{X/A}} \tag{4.6}$$

Carotenoid production in *X. dendrorhous* yeast cultures is partially growth associated, occurring in both the growth and stationary phase [236]. As the genes incorporated into the *S. cerevisiae* SM14 strain of interest came from an *X. dendrorhous* strain, we assumed that our recombinant yeast would follow the same behavior. Therefore, β-carotene production can be related to cell growth and biomass concentration by the Luedeking-Piret equation given in Equation 4.7, where α_i represents the coefficients for growth-associated product formation related to the yeast growth on each substrate, and β is the coefficient for non-growth-associated carotenoid production [252].

$$\frac{dP}{dt} = (\alpha_1\mu_G + \alpha_2\mu_E + \alpha_3\mu_A)X + \beta X \tag{4.7}$$

4.2.5.2. Parameter Estimation

In the previous section, five dynamic model equations for cell growth (Equation 4.1), substrate consumption (Equation 4.4), ethanol and acetic acid formation and depletion (Equations 4.5 and Equation 4.6) and β-carotene production (Equation 4.7) have been developed. This section will explain how the parameters of these model equations were estimated from experimental data. The parameter estimation is done in two steps, first to determine the optimal parameters in the growth rate equations and then using the determined parameters to find the optimal parameters in the dynamic model equations. Values used to initialize the two parameter estimation algorithms were

selected from literature for the growth of *S. cerevisiae* and *X. dendrorhous* [236, 237, 253].

The data used in estimating the parameters of the model were filtered using a cubic smoothing spline method, a piecewise function and has a high degree of "smoothness", available in Matlab [254]. From the batch cultivation data, we can distinguish three different growth rates related to the substrate that is being consumed at any moment in the cultivation. The specific growth rate expression is a generalized Monod model for the growth of an organism on mixed substrates (Equations 4.2 and 4.3). Specific growth rate μ was calculated as

$$\mu = \frac{\ln\left(\frac{x_2}{x_1}\right)}{t_2 - t_1} \tag{4.8}$$

With the collected data and the developed parameter estimation algorithm, the optimal parameters were estimated for the specific growth rate. The objective function used is the sum of squared errors (SSE) between the known data and model predictions and uses the Matlab *fmincon* function.

To estimate the parameters of the dynamic mass balance equations, we propose the use of an expression that minimizes the coefficient of determination, R^2, as the objective function. The R^2 is a value that ranges from zero to unity and indicates the "goodness" of the fit between the model and the data. Using this value, the objective function avoids the necessity of weighting the equations, a common method found in the literature due to the order of magnitude difference in the expected concentrations of carotenoids versus the other products [237].

Knowing that the best R^2 value for each curve to be unity, the objective function for the second optimization is shown in Equation 4.9, where the coefficient of determination was calculated for each variable of the system.

$$\min \ Z = 5 - (R_p^2 + R_x^2 + R_s^2 + R_e^2 + R_a^2) \tag{4.9}$$

Equation 4.10 shows how each R^2 value is calculated, where the sum of the square error (SSE) is given by Equation 4.11, Δ_i^2 being the squared difference between the data value and the predicted value at a given time point. Equation 4.12 represents the total of sum squares (SST) which is the difference between the value to be predicted and the arithmetic mean of the observed data.

$$R^2 = \frac{1}{m}\sum_{j=1}^{m}\left(1 - \frac{SSE}{SST}\right) \tag{4.10}$$

$$SSE = \sum_{i=1}^{n}\Delta_i^2 \tag{4.11}$$

$$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{4.12}$$

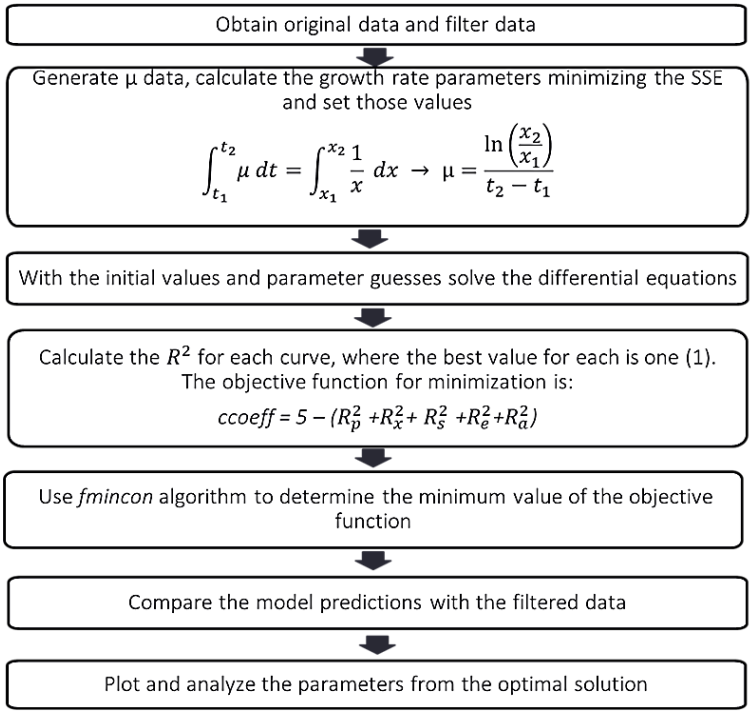

Figure 4.4. Optimization scheme used to determine the optimal parameters of the batch kinetic model.

The constraints of this optimization problem are the 5 dynamic mass balances describing the concentration of each component in the system. This programming problem is solved by setting the parameters optimized in the growth rate equations and using an iterative looping method that passes between the *fmincon* algorithm and *ode45* algorithm found in Matlab to estimate the remaining parameters. Figure 4.4 summarizes the steps that the parameter estimation algorithm takes to calculate the optimal value for the differential equations that describes our system.

4.2.5.3. Local Sensitivity Analysis – Direct Differential Method

Local sensitivity analysis is the analysis of the output of a set of models with respect to small perturbations in the model parameters or initial conditions. One of the

most widely used methods for doing this analysis is the calculation of sensitivity indices, *S*, via the direct differential method. Calculation of these sensitivity indices requires the development of differential equations of the sensitivity index, shown in Equation 4.13, which include calculations of the Jacobian matrix of the system as well as the derivative of each model with respect to the parameter or initial condition of interest [72]. The definition of the Jacobian matrix $\boldsymbol{J}$, sensitivity index vector $\boldsymbol{S}$, and vector of function-parameter derivatives $\boldsymbol{F_j}$ is shown in Equation 4.14 This results in a large-scale system of differential equations which need to be solved simultaneously with the initial conditions shown in Equation 4.15.

$$\frac{d}{dt}\frac{\partial \boldsymbol{y}}{\partial p_j} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\frac{\partial \boldsymbol{y}}{\partial p_j} + \frac{\partial f}{\partial p_j} = \boldsymbol{J} * \boldsymbol{S_j} + \boldsymbol{F_j} \tag{4.13}$$

$$J = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}} = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & & \frac{\partial f_2}{\partial y_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \frac{\partial f_n}{\partial y_2} & \cdots & \frac{\partial f_n}{\partial y_n} \end{bmatrix}, \qquad \boldsymbol{S} = \frac{\partial \boldsymbol{y}}{\partial\, p_j} = \begin{bmatrix} S_{1,j} \\ S_{2,j} \\ \cdots \\ S_{n,j} \end{bmatrix}, \qquad \boldsymbol{F}_j = \frac{\partial f}{\partial p_j} = \begin{bmatrix} \frac{\partial f_1}{\partial p_j} \\ \frac{\partial f_2}{\partial p_j} \\ \cdots \\ \frac{\partial f_n}{\partial p_j} \end{bmatrix} \tag{4.14}$$

$$\text{At } t = 0,\ s(y_n, y_n^i) = 1;\ s(y_n, p_j) = 0;\ s(y_n, y_{n\pm1}^i) = 0 \tag{4.15}$$

The kinetic models used to describe the carotenoid production by *S. cerevisiae SM14* above consist of 5 ordinary differential equations, 5 initial conditions, and 10 variables of interest. To implement the direct differential method described above, equations were developed for 75 total sensitivity indices. These equations were solved simultaneously and the results were analyzed to determine the parameters most important to each model equation when small perturbations are considered.

4.2.5.4. Global Sensitivity Analysis – ANOVA Method

While local sensitivity analysis looks at the sensitivity of a solution in a close proximity to the optimal solution, it is also worth analyzing the sensitivity of the models for the full range of possible parameter values. For this, global sensitivity analysis can be used, which instead of considering only small perturbations in one parameter at a time focuses on parameter variations over the entire possible parameter space. In addition, global sensitivity indices can examine the combined effect of multiple parameters on the model solutions. Unlike local sensitivity analysis, which considers the sensitivity of the model subject to the nominal optimal parameters, the global sensitivity analysis considers the sensitivity of the model over the entire feasible space of the parameters. Consequently, the calculated global sensitivity indices allow the study of the mathematical model rather than a specific solution.

The method we use for calculating the global sensitivity index relies upon stochastic variations of parameters. The method uses a normally distributed search of the parameter space and subsequent analysis of the variance (ANOVA) in the model outputs [73]. This is called the ANOVA method and utilizes a Monte Carlo search method to produce a large number ($2N > 10{,}000$) of parameter inputs (γ).

$$\xi_j = \left(\gamma_1^j, \dots, \gamma_{2N}^j\right)$$

The large number of sample points is then split into two matrices of equal size, denoted by ξ_j and ξ_j' below. These two matrices will be the basis for the calculation of the global sensitivity indices.

$$\xi_j = \left(\gamma_1^j, \dots, \gamma_N^j\right) = \left(\eta_j, \zeta_j\right)$$

$$\xi_j' = \left(\gamma_{N+1}^j, \ldots, \gamma_{2N}^j\right) = \left(\eta_j', \zeta_j'\right)$$

To begin the algorithm, the rows belonging to the individual variable or combination of variables of interest are swapped between the matrices ξ_j and ξ_j', as shown below.

$$\xi_j^* = \left(\eta_j, \zeta_j'\right) \qquad \xi_j'^* = \left(\eta_j', \zeta_j\right)$$

The equations are solved with each parameter set found in the matrices ξ_j and ξ_j^*, and are given by $f\left(\xi_j\right)$ and $f\left(\xi_j^*\right)$ in Equations 16 through 18. Using the model outputs, it is possible to calculate their variances with respect to the original data set, the set of variables of interest, denoted by y, and the set of all other parameters, denoted by z, using Equations 4.16 through 4.20.

$$\frac{1}{N}\sum_{j=1}^{N} f\left(\xi_j\right) \xrightarrow{P} f_0^2 \tag{4.16}$$

$$\frac{1}{N}\sum_{j=1}^{N} f^2\left(\xi_j\right) \xrightarrow{P} D + f_0^2 \tag{4.17}$$

$$\frac{1}{2N}\sum_{j=1}^{N}\left[f\left(\xi_j\right) - f\left(\xi_j^*\right)\right]^2 \xrightarrow{P} D_z^{tot} \tag{4.18}$$

$$D_y = D - D_z^{tot} \tag{4.19}$$

With the variances known it is possible to calculate two global sensitivity indices for the variable set of interest. This is done by using Equations 4.20 and 4.21. The first global sensitivity index, S_y, is a measure of the sensitivity of the model with respect to only the single variable of interest. For example, $S_{G,Y_{XG}}$ would denote the sensitivity of

the glucose model with respect to only the Y_{XG} parameter. The second sensitivity index, S_Y^{tot}, is the measure of the total sensitivity that a model has with respect to a parameter y. Unlike the first index, S_y, this second index also accounts for the variance associated with interactions between all parameters and the parameter of interest, thus accounting for the total variance associated with parameter y. To contrast with $S_{G,Y_{XG}}$, the sensitivity index $S_{G,Y_{XG}}^{tot}$ would account for the sensitivity of not only $S_{G,Y_{XG}}$ but also of the sensitivity of any combination of parameters combined with Y_{XG}.

$$S_y = \frac{D_y}{D} \tag{4.20}$$

$$S_y^{tot} = \frac{D_y^{tot}}{D} \tag{4.21}$$

These sensitivity indices can have values between zero and unity. While intermediate values of these are not necessarily informative, there are a few extreme cases which can divulge how the model and parameters interact [71]:

1. If $S_y = S_y^{tot}$, the variable of interest is not involved in any interaction with other input factors
2. If $S_y = S_y^{tot} = 0$, the model does not depend on the variable of interest.
3. If $S_y = S_y^{tot} = 1$, the model depends only on the variable of interest

4.3. Results and Discussion

4.3.1. Bioreactor Cultivation Results

Figure 4.5a shows the batch culture profiles of the cell growth, β-carotene production, ethanol and acetic acid production and subsequent consumption, and glucose

consumption of *S. cerevisiae* SM14 in a stirred-tank bioreactor with 20 g/L initial glucose. Figure 4.6 shows the time profile for the dissolved oxygen content in the bioreactor as a percentage of the saturation value, which remains above 70% for the duration of the experiments. The yeast exhibited an exponential growth period in the initial 24 hours, until the glucose was exhausted and the ethanol had reached its maximum concentration of 5.42 g/L. After glucose depletion, ethanol was utilized as a carbon source, resulting in ethanol depletion and an acetic acid peak concentration of 1.19 g/L at 50 hours. After this time, the acetic acid was consumed. After 24 hours, the yeast presented a slow growth period to reach a maximum biomass density of 7.9 g/L. Beta-carotene production started in the early exponential growth phase and continued throughout the cultivation period. The β-carotene production increased markedly when consuming the sub-products, reaching nearly 120 mg/L at the end of a 72-hour cultivation period.

Figure 4.5b shows the batch culture profiles with the initial glucose raised to 90 g/L. When compared to Figure 4.5a the same trends are seen regarding each component in the system. In addition, it can be noted that the β-carotene yield is comparable, 15.2 $mg \cdot L^{-1} \cdot h^{-1}$ when charged with 20 g/L versus 15.4 $mg \cdot L^{-1} \cdot h^{-1}$ when 90 g/L is used. These results are like those shown in Figure 4.1, where flask experiments show similar yields in the 20 g/L and 100 g/L initial glucose cases. However, when comparing the productivities of the two systems, an increase in initial substrate concentration leads to a longer initial lag phase, a result shown in literature for various biochemical systems [255, 256]. These longer cultivation times lead to a drop in the process productivity, 1.67

for the 20 g/L system versus 0.90 for the 90 g/L system, and thus an increase in initial substrate is not suitable for economically practical commercial implementation.

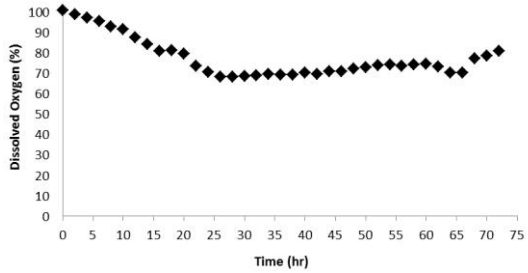

Figure 4.5. The dissolved oxygen profile for a batch culture of S. cerevisiae with 20 g/L initial glucose.

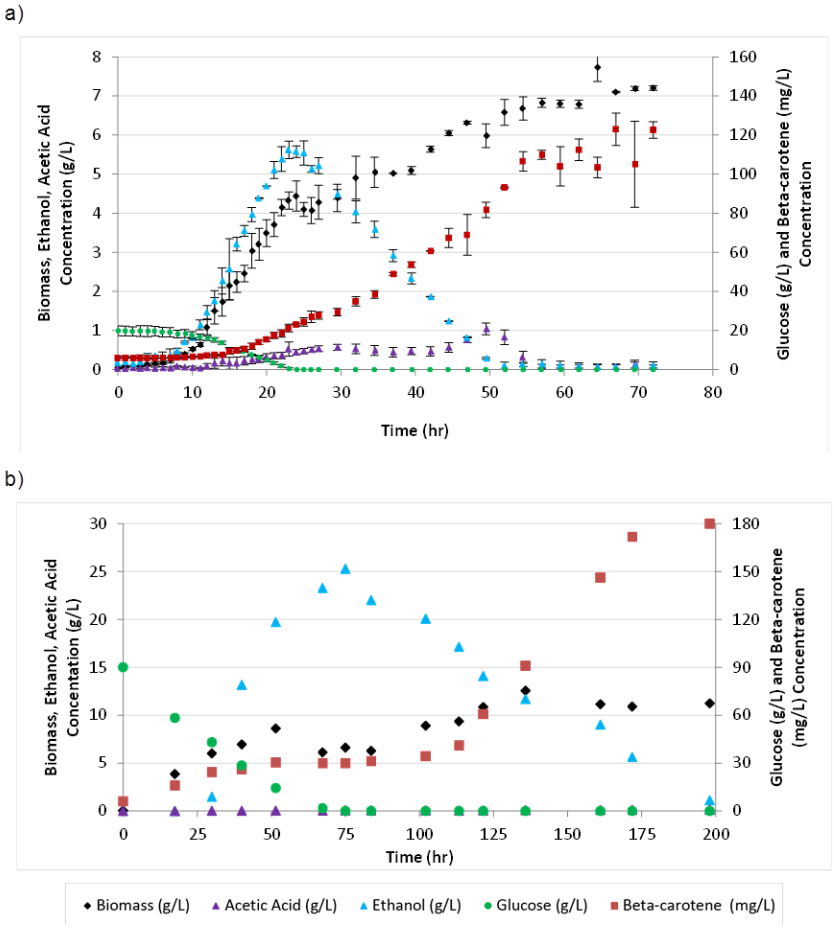

Figure 4.6. Time courses of cell growth, glucose consumption, ethanol and acetic acid concentration and carotenoids production in batch cultures of S. cerevisiae with (a) 20 g/L initial glucose and (b) 90 g/L glucose. Experiments with 20 g/L glucose were repeated

4.3.2. Inhibition Studies

The performed experiments reveal that the substrate concentrations studied did not present any inhibition effect in the strain as the maximum growth rate remained roughly constant up to glucose concentrations of 200 g/L (data not shown). The effect of ethanol inhibition in the *S. cerevisiae* SM14 can be observed in Fig 7a, where inhibition is seen for amounts of ethanol as low as 4.5 g/L. The growth rate is drastically affected with ethanol concentration higher than 30 g/L and completely inhibits the growth of the cells at 65 g/L. The acetic acid study is shown in Fig 7b, which illustrates that cell growth is affected at low concentrations of acetic acid, equivalent to 2 g/L. At 8.6 g/L cell growth is completely inhibited by acetic acid.

Using these data, it is possible to determine a normalized functional form for the inhibition effect of ethanol (χ_E) and acetic acid (χ_A) on the growth rate of *S. cerevisiae* SM14 on glucose. It has been shown in literature that equations such as that shown in Equation 4.22 have been used to fit cellular growth inhibition, where the toxicity parameter η is commonly found to be unity and the E_{crit} is the critical ethanol threshold beyond which there is no observable cell growth [257-259].

$$\mu_{obs} = \mu_{max}\left(1 - \frac{E}{E_{crit}}\right)^{\eta} \tag{4.22}$$

For this work, it was determined that this linear system would not accurately model the inhibition effects of ethanol and acetic acid due to their nonlinear behavior. The functions of ethanol and acetic acid inhibition used in this work were chosen as an extension of these linear models to higher order polynomial models of their respective

concentration, as shown by Equations 4.23 and 4.24, to try and capture the nonlinearity of the inhibition data. The fits given by these models are shown as the red lines in Figure 4.7a for ethanol and Figure 4.7b for acetic acid, respectively.

$$\chi_E = 1 - 4.1 \times 10^{-6} \cdot E^3 + 1.4 \times 10^{-4} \cdot E^2 - 9.1 \times 10^{-3} \cdot E \tag{4.23}$$

$$\chi_A = 1 - 0.011 \cdot A^2 - 0.021 \cdot A \tag{4.24}$$

The maximum concentration of ethanol and acetic acid found in the bioreactor during cultivation was 5.4 g/L and 1.2 g/L, respectively. Based on the inhibition studies, we can determine that the ethanol reduces the maximum glucose growth rate by approximately 6.5% and the acetic acid inhibition results in a decrease of about 5.7%.

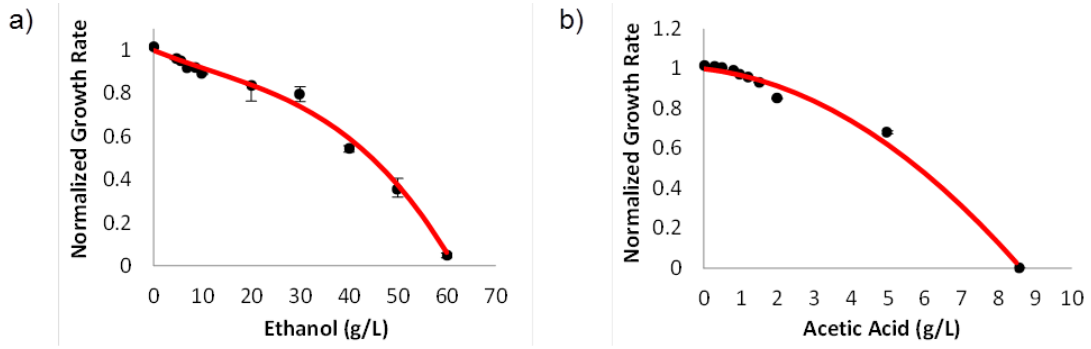

Figure 4.7. The inhibition effect of (a) ethanol and (b) acetic acid sub-products on the growth rate of *S. cerevisiae* SM14 on glucose. Experimental data is represented by the markers and the obtained fit is denoted by the continuous red line.

4.3.3. Parameter Estimation

The developed kinetic model equations (Equations 4.1 through 4.7) contain a total of 22 model parameters, all of which were estimated by using the previously described parameter estimation algorithm and the batch culture time profiles are shown in Figure 4.8. The values of the optimal parameters obtained from the parameter estimation are compiled in Table 4.1. The model simulation provided the coefficients of

determination for each variable of R_X^2 =0.989, R_G^2 =0.998, R_P^2 =0.993, R_E^2 =0.995 and R_A^2 =0.807; these R^2 values show that the predictive model can represent the experimental data with high accuracy. It is useful to note that the model predicts only production of carotenoids to be non-growth associated during glucose utilization and highly growth associated during ethanol utilization, as seen by the parameter α_1 being zero and α_2 being two orders of magnitude larger than β at the optimal solution.

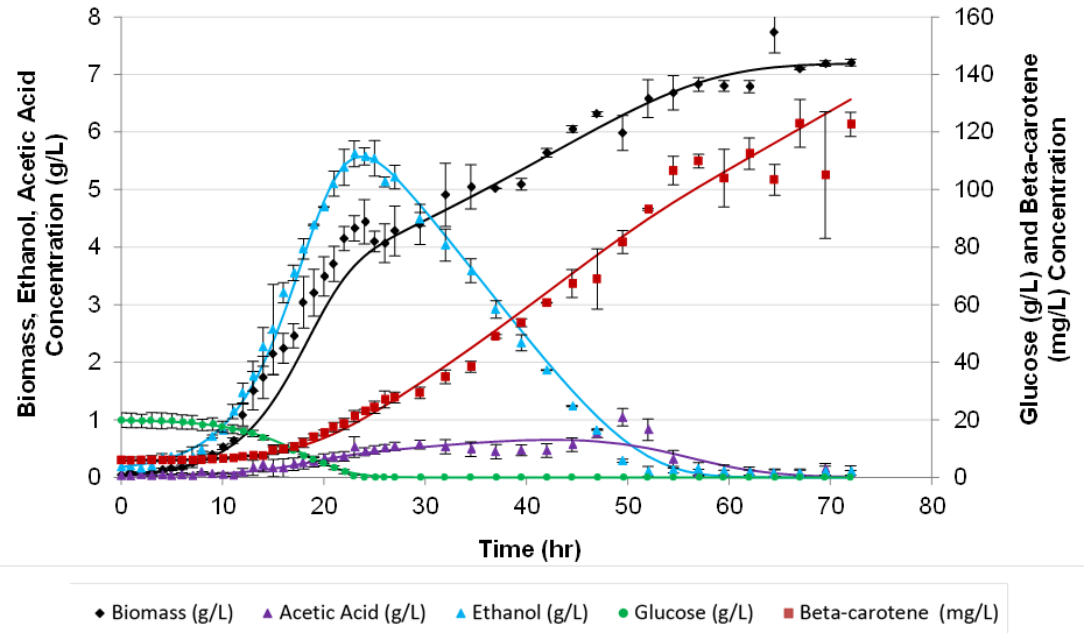

Figure 4.8. Time profiles and curves of best fit for the cell growth, glucose consumption, ethanol and acetic acid concentration and carotenoids production in batch cultures of *S. cerevisiae* with 20 g/L initial glucose. Cultivation data are represented by the markers and the optimal fit by the solid lines.

Table 4.1. Optimal parameter estimates for the batch kinetic models.

Parameter	Value	Units	Parameter	Value	Units
$\mu_{max,G}$	0.2516	hr^{-1}	a_{AE}	1.0031	$\frac{\text{g Cell}}{\text{g Ethanol}}$
K_{SG}	0.4137	$g\ Glucose\ L^{-1}$	Y_{XG}	0.1855	$\frac{\text{g Cell}}{\text{g Glucose}}$
$\mu_{max,E}$	0.0218	hr^{-1}	Y_{XE}	0.3637	$\frac{\text{g Cell}}{\text{g Ethanol}}$
K_{SE}	0.5618	$g\ Ethanol\ L^{-1}$	Y_{XA}	1.0163	$\frac{\text{g Cell}}{\text{g Acetic Acid}}$
$\mu_{max,A}$	0.0182	hr^{-1}	α_1	0.7545	$\frac{\text{mg Product}}{\text{g Glucose}}$
K_{SA}	0.4506	$g\ Acetic$ Acid L^{-1}	α_2	13.9280	$\frac{\text{mg Product}}{\text{g Ethanol}}$
a_{GE}	1.2964		α_3	1.1089	$\frac{\text{mgProduct}}{\text{g Acetic acid}}$
a_{GA}	1.0318		β	0.2804	$\frac{\text{mg Product}}{\text{g Cell} \cdot \text{hr}}$
a_{EG}	1.0636		k_1	1.7300	$\frac{\text{g Ethanol}}{\text{g Cell}}$
a_{EA}	1.0058		k_2	0.0936	$\frac{\text{g Acetic Acid}}{\text{g Cell}}$
a_{AG}	1.0000		k_3	0.2937	$\frac{\text{g Acetic Acid}}{\text{g Cell}}$

4.3.4. Local Sensitivity Analysis

The effect of each parameter and the initial values of the state variables was calculated and plotted as shown in the Figure 4.9. Figure 4.9a illustrates the sensitivity of the variables with respect to Y_{xg}. The glucose shows a high sensitivity during the time period it is used in the cultivation and then the sensitivity decreases to zero. The minimum and maximum values of glucose sensitivity are 0 and 60. Then, we can say that the local sensitivity of the glucose with respect to the parameter Y_{xg} is $0 < S^{local} < 60$. The same analysis was done for all the inputs and results are summarized in Table

4.2. Figure 4.9b presents the effect of the initial condition of the ethanol on the profiles of the state variables, where the negative sensitivity values represent an inhibitory effect.

In Table 4.2, we highlight the inhibitory effects captured by the model at the optimal values when changes in ethanol and acetic acid initial conditions are made. The predicted final value of the β-carotene shows a high sensitivity for small changes in the values of Y_{xg}, Y_{xe}, or initial biomass concentration G_i. Changes in the parameters governing the biomass production from glucose and ethanol would be expected to have an effect on the carotenoid production as it is an intracellular product (highlighted in pink in Table 4.2). The parameter β, symbolizing the non-growth associated carotenoid production, is a very influential variable in the prediction of β-carotene production, so small changes in this parameter lead to a larger effect on the carotenoid estimates, as seen by the elevated sensitivity index for the product. Other state variables are affected by parameter changes to a much smaller degree, most notably the sensitivity of biomass, glucose and ethanol on the Y_{XG} parameter (highlighted in orange in Table 4.2) and the biomass and ethanol on the initial biomass inoculum X_i (highlighted in green in Table 4.2).

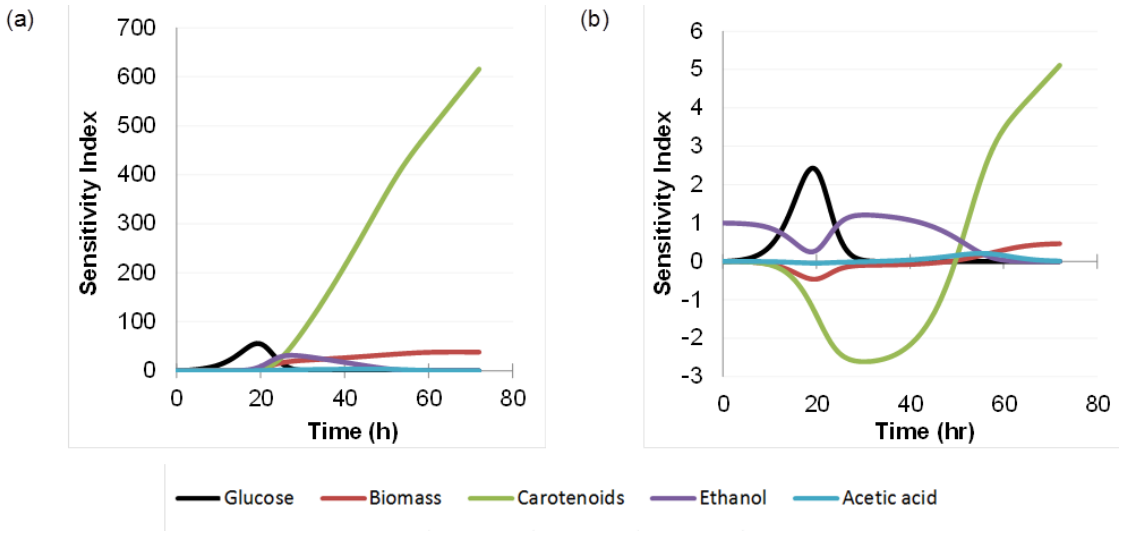

Figure 4.9. Local sensitivity analysis of the models at the optimal solution for the a) glucose yield with respect to biomass (Y_{XG}) and b) the initial concentration of ethanol (E_i).

Table 4.2. Minimum and maximum local sensitivity index values for each process variable with respect to all parameters and initial conditions

	Biomass	Glucose	Ethanol	Acetic Acid	Carotenoids
X_i	1.0 < S < 35.	-170. < S < 0.0	-23 < S < 52.	-4.7 < S < 3.4	0.0 < S < 280
G_i	0.0 < S < 0.4	0.0 < S < 1.0	0.0 < S < 0.3	-	0.0 < S < 6.0
E_i	-0.5 < S < 0.5	0.0 < S < 2.4	0.0 < S < 1.2	0.0 < S < 0.2	-2.6 < S < 5.1
A_i	-0.5 < S < 1.0	0.0 < S < 2.4	-0.7 < S < 1.0	0.0 < S < 1.0	-6.5 < S < 2.4
P_i	-	-	-	-	-
Y_{XG}	0.0 < S < 40.	0.0 < S < 60.	-0.1 < S < 30.	-0.3 < S < 3.4	0.0 < S < 620
Y_{XE}	-0.1 < S < 8.4	0.0 < S < 0.4	0.0 < S < 11	0.0 < S < 2.8	0.0 < S < 140
Y_{XA}	0.0 < S < 1.0	-	0.0 < S < 0.2	0.0 < S < 0.5	-0.6 < S < 4.0
k_1	-	-	-	-	-
k_2	-0.4 < S < 3.7	0.0 < S < 2.2	-0.6 < S < 3.0	0.0 < S < 3.4	-14 < S < 16
k_3	0.0 < S < 2.4	0.0 < S < 0.1	0.0 < S < 1.0	0.0 < S < 1.5	-4.2 < S < 11
α_1	-	-	-	-	0.0 < S < 4.0
α_2	-	-	-	-	0.0 < S < 2.4
α_3	-	-	-	-	0.0 < S < 1.1
β	-	-	-	-	0.0 < S < 320

4.3.5. Global Sensitivity Analysis

Figure 4.10 shows the relevant results from the global sensitivity analysis. In Figure 4.10a we show that the global (S_y) and total (S_y^{tot}) sensitivity indices of all variables concerning biomass production were identical, thus we conclude that the variance in this model is not associated with any interaction between variables at any time. During the first 9 hours, the biomass model is only influenced by the k_1 parameter representing the ethanol formation during glucose consumption. Between 9 and 25 hours, the time pertaining to glucose consumption and ethanol and biomass production, the variance in the model outputs are due to both k_1 and Y_{xg}. After 50 hours the Y_{xe} parameter begins to affect the variance in the biomass model as it is associated with ethanol consumption.

Ethanol formation and glucose consumption are related by the parameter k_1, leading to the only source of variance in the ethanol model during the time of ethanol production on glucose, as shown in Figure 4.10b. Later, the ethanol is being consumed to form biomass and acetic acid, and the model variance begins to depend on the yield of biomass on glucose, Y_{xg}, the yield of biomass on ethanol, Y_{xe}, as well as the coefficients k_1 and k_3. At the end of the cultivation, interaction effects between k_1 and Y_{xe} is present, as seen in the divergence of the total global sensitivity from the individual global sensitivity.

The main source of variance in the β-carotene model through the first 16 hours of the process was a result of the β parameter. As the cultivation progresses, the variance of the model shifts to depend on both β and Y_{xg}. The biomass formation is directly related

with the substrate consumption, which is expressed in the models by the parameter Y_{xg}. The amount of β-carotene is then related to the amount of biomass through the non-growth association parameter β. Due to these interactions, after 16 hours it is intuitive that these two parameters would play an important role in the model variance. Additionally, these two related parameters do not have any interaction effects on the model variance, as shown in Figure 4.10c.

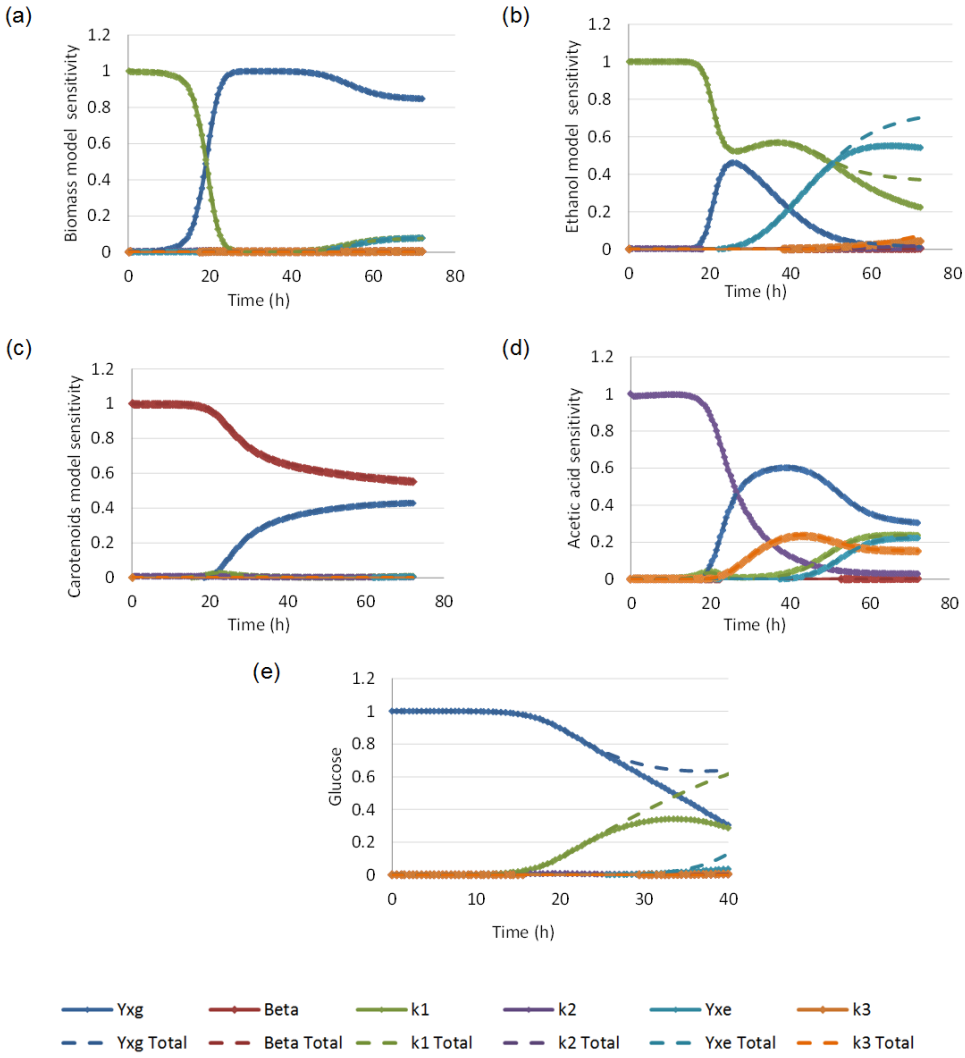

Figure 4.10. Global sensitivity analysis of the a) biomass, b) ethanol, c) carotenoids, d) acetic acid, and e) glucose models.

The global sensitivity analysis of acetic acid is shown in Figure 4.10d. The acetic acid is present in the cultivation in small amounts compared to other variables and it is also affected by glucose, ethanol and biomass; this is reflected on the global sensitivity analysis results for acetic acid. Initially the variance in the model is due only to parameter k_2, which represents the acetic acid formation by glucose consumption. Later, the model variance is affected by the parameters associated with glucose and ethanol yield with respect to biomass (Y_{xg}, Y_{xe}), ethanol formation (k_2), and acetic acid formation due the ethanol depletion (k_3). The total sensitivity index represents the same global sensitivity effect on the acetic acid as the variables did; therefore, we determine that the variance in the acetic acid model does not represent interaction with the related variables at any time.

The glucose model variance can be seen in Figure 4.10e and shows results from the variance in the first 9 hours depend on the yield of biomass on glucose, Y_{xg}, which is due the biomass formation during this period of the cultivation. Later, the variance of the model is due to both Y_{xg} and k_1; this is because of ethanol formation and during this time the global sensitivity shows the same effect as the total sensitivity. After the glucose is consumed, the total sensitivity differs from the global sensitivity, thus we conclude that the glucose model represents accurate information regarding interaction with the related variables.

The remaining variables Y_{xa}, α_1, α_2 and α_3 did not show any effect on the variance in the model predictions.

4.4. Conclusions

This study develops models and estimates the parameters necessary for accurately describing the growth of *S. cerevisiae* SM14 and the resulting product profiles, specifically glucose, ethanol, acetic acid, and β-carotene. Results indicate that the proposed model and estimated parameters are sufficient for describing the growth with glucose consumption and sub-products formation and depletion in batch cultures of *S. cerevisiae* SM14. Additionally, utilization of the coefficient of determination R^2 in the objective function provides a reliable optimal solution for parameter estimation of the model equations. It avoids the normalization of the equations or the weighting of them. The local sensitivity analysis reflects the inhibition effect of the ethanol and acetic acid concentration in the system and indicates for which parameters we must have accurate parameter estimation to develop a suitable description of the cultivation process. Global sensitivity analysis gives an understanding of the mathematical models in all the input space. The results show appropriate connections between the model variance, parameters of study, and characteristics of the cultivation process.

The methods presented here have a much wider application than the single case study of β-carotene production from genetically altered S. *cerevisiae*. Modified versions of these models and methods can be applied to any biological system for which there is a set of sub-products that experience generation and depletion through the course of a single batch fermentation. Additionally, the models described for this specific case study can be used in future optimization and controls studies, including applications in fed-batch and continuous production processes. Finally, knowledge of the *S. cerevisiae*

kinetics is the first step toward the analysis large-scale production for which the economic impact of biologically produced nutraceuticals can be ascertained.

CHAPTER V

OPTIMAL CONTROL OF A MULTI-FEED BIOREACTOR FOR THE CONTINUOUS PRODUCTION OF PHARMACEUTICALS†

5.1. Introduction

The production of pharmaceuticals is a multi-billion-dollar industry worldwide that is expected to see a dramatic increase in profit by 2019 [260]. These growth projections are particularly relevant in large markets such as the United States and Europe, accounting for 44.5% and 25.3% of the world's pharmaceutical sales in 2014, respectively, and high potential markets such as Brazil, China, India and Russia [261]. However, trends have shown stagnation in newly developed pharmaceutical products. Figure 5.1 depicts this trend through the number of approved new drug applications (NDAs) by the Food and Drug Administration (FDA) of the United States, showing a relatively constant number of approvals since 1980 [262]. To meet the economic growth projections set for 2019 despite this stagnant development of new products, new opportunities for economic improvement must be considered. A transition from batch to alternative processing methodologies in the pharmaceutical industry may present possible economic benefits despite introducing some additional complications.

†Reprinted with permission from "Economic Improvement of Continuous Pharmaceutical Production via the Optimal Control of a Multi-feed Bioreactor" by J.P. Raftery, M.R. Desessa and M.N. Karim, 2017. *Biotechnology Progress*, Copyright 2017 by John Wiley & Sons.

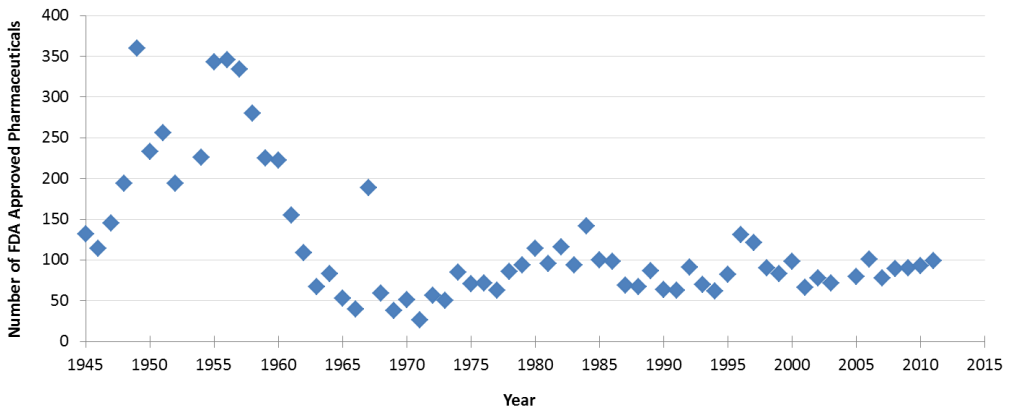

Figure 5.1. Number of approved new drug applications by the United States FDA from 1945 to 2011 [262].

Batch processing methods are the traditional production method for pharmaceuticals due to being ideal for small scale production. Batch processes in commodity chemical industries have been shown to incur large capital, operating and labor costs when compared their continuous counterparts. Batch systems in the pharmaceutical industry are subjected to intensive cleaning practices that utilize environmentally hazardous caustic solutions to ensure contamination is mitigated for each batch run [126]. Alternatively, continuous processing methods are easily controllable processes ideal for large scale implementation and result in a reduction of equipment size and minimal raw material requirements and waste production [263]. In recent years, the United States FDA has acknowledged the potential of continuous processing schemes in the pharmaceutical industry through the updating of their policies to accommodate continuous production strategies [263]. In addition, various process monitoring techniques are being developed to determine process contamination utilizing online data, allowing for quick process shutdowns that lessen the effect of contamination

downstream [205, 264, 265]. The main challenge of transitioning to continuous processing in the pharmaceutical industry is to design a reactor able to maintain or improve reactor productivity, and in doing so increasing profitability, while also maintaining strict standards of product quality when compared to traditional batch systems.

Many studies discussed in literature have shown the advantage of alternative processing methods for biological systems. Methods such as cell immobilization have been a staple of the bioprocessing community, with immobilization materials such as calcium alginate or silicon being readily available for use [266-268]. Researchers have also investigated reactor designs that can selectively remove the desired product from the reactor broth through selectively permeable membranes [269-271]. Specialized membrane reactors focused on the process of pervaporation, or the removal of a product via partial vaporization through a membrane in the bioreactor, have been considered for systems where volatile renewable liquid biofuels are produced [272-274]. San and Stephanopoulos investigated the use of fed-batch systems to maximize the productivity of penicillin production [206]. Warikoo et al. exemplified the use of continuous perfusion reactors for the production of therapeutic proteins from Chinese hamster ovary (CHO) cells [275]. The fed-batch production of the poly-3-hydroxybutyrate (PHB) has also been investigated by Yamane, Fukunaga and Lee to increase cell density and therefore improve productivity of the intracellular bio-plastic [276]. These bioreactor systems do not operate continuously or are limited to use with extracellular metabolites due to their necessity for cell immobilization. However, many important biological

products are intracellular and require cell harvesting and disruption of the biomass for product recovery. For intracellular products, the improvement of continuous bioprocessing systems can be achieved through the manipulation of the feed of the bioreactor to introduce or limit substrate in the system.

To contend with the challenges of fed-batch and continuous bioprocesses, many researchers have utilized the concepts of dynamic optimization, also known as optimal control, as a way of predicting the optimal operation of a bioreactor system *a priori*. Henson and Seborg have examined the use of nonlinear controllers based on exact linearization for continuous fermentation [277]. Work by Modak and Lim, as well as the aforementioned work by San and Stephanopoulos, has shown that the implementation and optimal control of biochemical processing methods beyond batch cultivation can result in improved productivity [206, 278]. Hodge and Karim have employed a nonlinear model predictive control algorithm that uses a sequential solution to predict solutions to the optimal control problem in real time and maximize fed-batch ethanol production by recombinant *Zymomonas mobilis* [279]. Additionally, work by Saucedo and Karim has shown that the use of input-output models for the real time optimization and control of fed-batch cultivation processes can be used to enhance ethanol production in *Escherichia coli* systems [280]. Further work by Sridhar and Saucedo has shown the existence of globally optimal solutions to the optimal control problem of continuous production of ethanol via *Saccharomyces cerevisiae* using the dilution rate or mass transfer coefficient as decision variables [281]. However, the use of a single feed at a set substrate concentration or mass transfer coefficient may result in a lack of controllability to

respond to the needs of biochemical systems while also trying to maintain the constant volume of a continuous reactor system.

This paper develops a novel strategy for continuous cultivation to enhance the economic viability of bioreactors when compared to batch cultivation methods. A three-feed bioreactor system is introduced that allows for the independent variation of the dilution rate and entering substrate concentrations. This reactor system is then applied to a model compound of beta-carotene from recombinant *S. cerevisiae*, where it is shown that by using optimal control methodologies the necessary control policy that governs its optimal production during continuous cultivation can be determined *a priori*. The optimal steady-state results of this control policy are used to compare the productivities of continuous operation of a bioreactor system to the batch cultivation of the same system. Finally, the design of the bioreactor will be optimized using the optimal control algorithm as constraints to further improve upon the productivity of continuous production.

5.2. Materials and Methods

5.2.1. Multi-feed Continuous Processing

Figure 5.2 illustrates the proposed set up of a multi-feed continuous bioreactor consisting of multiple independent feeds allowing for the variation in operation necessary to adapt to the complications of biological system. While any number of feeds can be utilized, Figure 5.2 depicts the configuration of a multiple feed bioreactor for an organism capable of assimilating two different substrates, denoted S_1 and S_2. The feed flow rates for these streams are represented by F_1 and F_2. A third feed devoid of any

substrate and with a flow rate F_3 is composed of the same media as the other two feeds and added for dilution effects. Pumps are added to each branching path as a way of manipulating the flow rate of these three feed streams. The feed streams are mixed together before being introduced by the reactor, leading to a feed flow rate described by Equation 5.1.

$$F_{in}(t) = F_{out}(t) = F_1(t) + F_2(t) + F_3(t) \tag{5.1}$$

The manipulation of the flow rates of the various feeds allows for the independent variation of reactor dilution rate as well as substrate concentrations. By restricting the flow rate of one of the substrate feeds and the media feed F_3 the reactor in Figure 5.2 becomes the typical single feed bioreactor capable of implementing these various single-feed feeding strategies. Many previous works have looked at the effect of single-feed strategies on the productivity of biological systems. Ho, Tam and Zhou utilized constant and exponential feeding to maximize carotenoid yield in *Phaffia rhodozyma* [282]. Ding and Tan showed the effects of pulse feeding as well as constant and exponential feeding strategies on the fed-batch L-lactic acid production of *Lactobacillus casei* [283]. However, while these strategies work to improve the economic viability via an increase in productivity of simple systems of a single feed, the addition of multiple feeds allows for more complex feeding strategies as well as the introduction of multiple potential substrates that may further improve the economic viability of more complex systems.

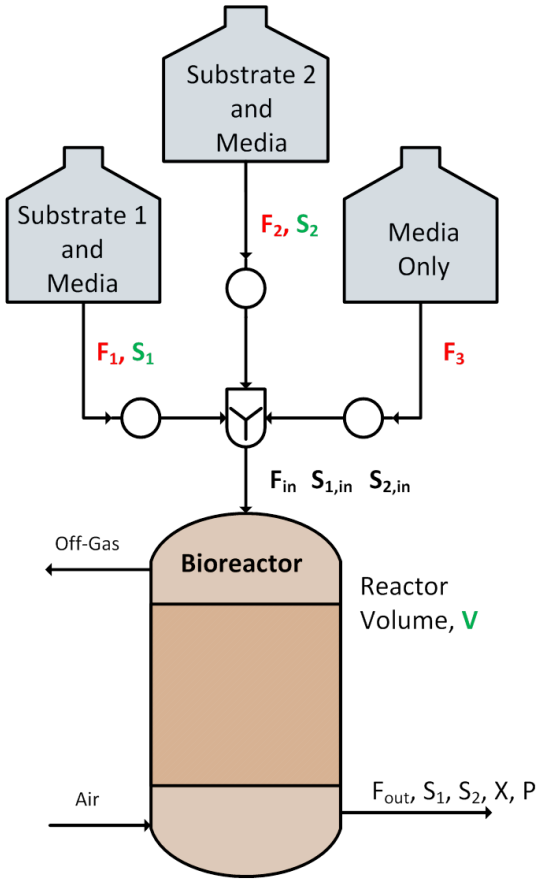

Figure 5.2. Diagram of the multi-feed system to improve the economic viability of continuous biochemical reactors

5.2.2. Optimal Control Problem Formulation and Solution Methodology

The development of an optimal control, also known as dynamic optimization, framework can be used to determine the optimal operation of the continuous multi-feed reactor system. The solution of the optimization problem determines the optimum manipulation of process variables as a function of time that maximizes some objective, such as product titer or productivity. In the bioreactor system shown in Figure 5.2, the manipulated variables are the flow rates dictated by the operation of the pumps for each feed. This methodology can be used to elucidate non-trivial feeding methods *a priori*

that can then be further examined in experimental studies. The general formulation of the dynamic optimization problem is shown in Equation 5.2. The objective is to maximize an index ϕ which takes the form of a function of both the set of system states $\boldsymbol{x}$ and the set of manipulated variables $\boldsymbol{u}$. Differential equations, denoted $\dot{\boldsymbol{x}}$ in Equation 5.2, are used as constraints to the optimization problem and usually govern the states of the systems as functions of either time, i.e. kinetic models, or space, i.e. heat or mass transport models. Algebraic equations describing physical constraints and physical properties can also be used as constraints to the optimization problem and are denoted as $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{u})$ in Equation 5.2. Strict limitations are also placed on the values allowed for the manipulated variables, for example setting the value of a pump flow rate to always be greater than or equal to zero. The generic form of the optimal control problem is shown in Equation 5.2.

$$\begin{aligned} \min_{\boldsymbol{u}} \phi &= \int_0^{t_f} f\big(\boldsymbol{x}(t),\boldsymbol{u}(t)\big) \\ s.t. \quad & \dot{\boldsymbol{x}} = \boldsymbol{h}(\boldsymbol{x},\boldsymbol{u}) \\ & \boldsymbol{g}(\boldsymbol{x},\boldsymbol{u}) = 0 \\ & u^{min} \le \boldsymbol{u} \le u^{max} \end{aligned} \tag{5.2}$$

Many solution methods have been proposed to solve the dynamic programming problems summarized in Equation 5.2. One subset of methods, deemed sequential methods, discretizes the dynamic programming problem into a multistage system where the operating conditions in each stage must be selected in sequence in an effort to optimize the entire process [284]. Another set of solution methods, termed pseudospectral methods, solves the entire time horizon simultaneously through the

conversion of the dynamic programming problem into a large-scale set of nonlinear equations that can be solved simultaneously with commercially available solvers. This relies on the conversion of the differential equation constraints to a set of algebraic constraints through the discretization along known points known as finite elements. Within each finite element, a set of points termed collocation points are chosen as the roots of a single Legendre polynomial or combinations of Legendre polynomials and their derivatives [285]. This work uses the method proposed by Flores-Tlacuahuac at al. which uses a three-point $(\text{ncol} = 3)$ Radau collocation method, with the exact point values defined by Hairer and Wanner, to converting the differential equations $\dot{\boldsymbol{x}}$ into the algebraic profiles shown in Equation 5.3 [286, 287]. Here, $\boldsymbol{x}_i$ denotes the value of the state variables at finite element i, h_i is the finite element length, t_i is the time at each finite element, and $\frac{dx}{dt_{i,q}}$ is the value of the differential equations describing state variables at finite element i and collocation point q. The Legendre polynomial describing the state variables within each finite element is denoted as Ω_q. Equation 5.4 is also included in the new optimization to guarantee the continuity of the state profile between finite elements.

$$\boldsymbol{x}(t) = \boldsymbol{x}_{i-1} + h_i \sum_{i=1}^{ncol} \Omega_q \left(\frac{t_i - t_{i-1}}{h_i}\right) \cdot \frac{d\boldsymbol{x}}{dt_{i,q}} \tag{5.3}$$

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} + h_i \sum_{i=1}^{ncol} \Omega_q \left(\frac{t_i - t_{i-1}}{h_i}\right) \cdot \frac{d\boldsymbol{x}}{dt_{i,q}} \tag{5.4}$$

With the conversion of the differential equations to a larger system of algebraic equations, the optimization program in Equation 5.2 becomes a nonlinear program that

can be solved using commercially available solvers. In this work these optimization problems developed for the multi-feed continuous bioreactor are modeled in the General Algebraic Modeling System (GAMS) and solved using the IPOPT solver developed by Wächter and Biegler [288].

5.2.3. Case Study: Production of β-carotene

In this work the efficacy of the multi-feed continuous bioreactor paradigm developed in Section 2.1 is exemplified through the production of beta-carotene. Beta-carotene is an orange pigment naturally produced by many plant species, including carrots, cantaloupe, and peppers, as well as some fungal and bacterial species. It has been used widely in many different industries like in the pharmaceutical industry as a vitamin A precursor and antioxidant, the paint and food industry as an orange coloring additive, and in the cosmetics industry to improve skin tone and reduce the cancer risk by resisting the damage by solar radiation, and also has other uses [217, 221, 223, 224]. Beta-carotene can be produced synthetically but this process is difficult due to the complicated structure of the molecule. One alternative method of producing biological beta-carotene is through its recovery from natural sources such as the plant species mentioned above. However, these methods require harsh chemicals and result in a devaluation of the product due to isomerization during the extraction process [289, 290]. Studies have suggested that biologically formed beta-carotene could exhibit enhanced antioxidant properties due to an increase in fat solubility when compared to its synthetic counterpart [291-293]. These enhanced properties in conjunction with the potential for higher profit margins due to the market price for natural beta-carotene, approximately

$1,000 to $2,000 per kilogram, being much higher than that of the synthetic product, approximately $400 to $800 per kilogram, make beta-carotene a model compound to be analyzed for biological production [292].

One way of increasing the availability of biologically produced beta-carotene is to use genetic engineering and biochemical processing techniques to develop a process capable of producing and purifying the recombinant product. Regarding beta-carotene production, literature shows many different species capable of producing the orange carotenoid. Luna-Flores et al. were successful in developing and modeling the beta-carotene production of a *Xanthophyllomyces dendrorhous* in both batch and fed-batch production modes [237]. Works by Reyes et al. and Olson et al. have shown that through chromosomal integration and adaptive evolution of a *Saccharomyces cerevisiae* SM14 strain it is possible to produce beta-carotene in batch culture [238, 294]. Figure 5.3 below shows in detail the path added to the SM14 strain, using a branching fatty-acid path from the production of acetyl-CoA in the cell to make the desired beta-carotene product. This strain developed by Reyes et al will be the focal point of this work.

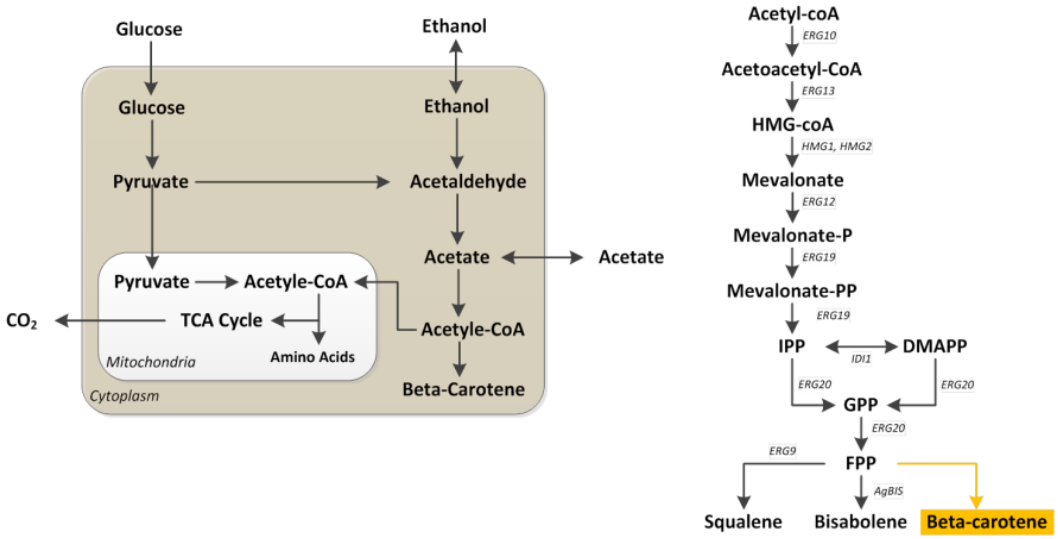

Figure 5.3. Modified biochemical pathway of S. cerevisiae to produce beta-carotene based on the work of Reyes et al [238]. The image on the left shows the location of the new branch pathway while the image on the right shows a more detailed version of the beta-carotene pathway.

Due to its intracellular nature, typical continuous production methodologies utilizing membrane reactors or other immobilization strategies are not applicable as the cell biomass needs to be harvested for product recovery. In addition, *S. cerevisiae* has been known to exhibit overflow metabolism, also known as the "short-term Crabtree effect", which results in the formation of inhibitory compounds during aerobic operation due to the incomplete oxidation of the carbon source [229, 230]. However, our previous work has also shown that the SM14 strain of *S. cerevisiae* utilizes these inhibitory compounds in a secondary growth phase once the initial carbon source has been completely utilized, during which there is an increase in the production rate of the beta-carotene product [295]. To incorporate this into the multi-feed bioreactor framework, the first feed, F_1, of the reactor is a mixture of ethanol and media at ethanol concentration E_1. This feed is used to promote additional beta-carotene production by feeding the

ethanol substrate most responsible for product formation in the cell. The second feed, F_2, is a mixture of glucose and media at a concentration G_2. This feed is used to provide the primary substrate necessary for biomass cultivation along with the generation of ethanol needed for beta-carotene production. Finally, a media only feed F_3 is used to provide dilution effects for the control of ethanol and/or glucose concentration as well as to maintain a constant dilution rate of the system, if needed. These feeds are combined before entering the reactor per Equation 5.1.

The differential growth models for continuous beta-carotene production developed in this section are based on the kinetic models described by the work of Ordonez et al [295]. The batch bioreactor studies used to develop the models in Equations 5.6 through 5.13 were carried out in a 7 L, autoclavable, glass, Applikon® bioreactor with a 3 L working volume. The temperature was controlled at 30℃ and the pH was maintained at 4 by addition of 2 M HCL or 2 M NaOH as needed. The bioreactor was supplied with a constant airflow of 6 L/min and the impeller speed was kept constant at 800 rpm. The models begin with an additive growth rate, as shown in Equation 5.6, which considers the multiple growth phases of the strain. The growth rate for each substrate is defined by a modified Monod kinetic equation, as seen in Equations 5.7a, 5.7b, and 5.7c, which is the most common unstructured kinetic model for microbial growth, and relates the microbial growth rate to a single limiting substrate [157, 249]. These equations are functions of the substrate of interest for each growth phase and include inhibition effects from the other substrates present. Modifications to the models

of this type were first utilized in the work of Yoon et al. to describe multiple substrate systems [250].

$$\mu = \mu_G + \mu_E + \mu_A \tag{5.6}$$

$$\mu_G = \left(\frac{\mu_{max,G} \cdot \chi_E \cdot \chi_A \cdot G}{K_{SG} + G + a_{ge} E + a_{ga} A} \right) \tag{5.7a}$$

$$\mu_E = \left(\frac{\mu_{max,E} E}{K_{SE} + E + a_{eg} G + a_{ea} A} \right) \tag{5.7b}$$

$$\mu_A = \left(\frac{\mu_{max,A} A}{K_{SA} + A + a_{ag} G + a_{ae} E} \right) \tag{5.7c}$$

While ethanol and acetic acid act as substrates to produce biomass and beta-carotene, they also work as inhibitors to the growth process. Equations 5.8a and 5.8b are taken from Ordonez et al. to mathematically describe the inhibition effects of each substrate [295].

$$\chi_E = 1 - 4.1 \times 10^{-6} \cdot E^3 + 1.4 \times 10^{-4} \cdot E^2 - 9.1 \times 10^{-3} \cdot E \tag{5.8a}$$

$$\chi_A = 1 - 0.011 \cdot A^2 - 0.021 \cdot A \tag{5.8b}$$

The growth rates described in Equations 5.6 and 5.7 are used to develop models for the five components present during the cultivation process, biomass (X), acetic acid (A), beta-carotene (P), glucose (G), and ethanol (E). Equations 5.9, 5.10, and 5.11 show the models developed for all species not associated with the feeding process. The change in the biomass concentration described by Equation 5.9 is the summation of two terms, the first of which is an autocatalytic first order reaction rate defined by the total growth

rate, μ. The second term of Equation 5.9 describes decrease in the biomass concentrations due to losses in the effluent and the dilution effects of the feed stream.

The change in acetic acid concentration, shown as Equation 5.10, is described by four terms, the first two denoting the production of acetic acid through the consumption of glucose or ethanol, respectively. Acetic acid is lost via reaction to form more biomass as well as by the dilution of the reactor inlet, as shown by the last two terms in Equation 5.10. The production of beta-carotene is denoted by Equation 5.11, where the kinetics are described by a Luedeking-Piret model that attributes production via growth association, a result of the utilization of any of the three substrates by the biomass, as well as non-growth association characterized by the parameter β [252]. The α_1, α_2, and α_3 parameters in this equation define the dependence for product formation on the biomass growth from the associated substrate utilization.

$$\frac{dX}{dt} = -\frac{F_{out}}{V}X + \mu_G X + \mu_E X + \mu_A X \tag{5.9}$$

$$\frac{dA}{dt} = -\frac{F_{out}}{V}A + \alpha_1 \mu_G X + \alpha_2 \mu_E X - \frac{\mu_A X}{Y_{X/A}} \tag{5.10}$$

$$\frac{dP}{dt} = -\frac{F_{out}}{V}P + \alpha_3 \mu_G X + \alpha_4 \mu_E X + \alpha_5 \mu_A X + \beta X \tag{5.11}$$

Equations 5.12 and 5.13 describe the time-dependent mass balance for the feed species glucose and ethanol. Each equation is composed of an inlet feed term based on the inlet concentrations G_f and E_f, outlet terms based on the outlet flowrate F_{out} and depletion terms based on the kinetic equations developed by Ordonez et al [295]. The second term in Equation 5.13 accounts for the formation of ethanol during the glucose utilization phase of *S. cerevisiae* SM14.

$$\frac{dG}{dt} = -\frac{F_{out}}{V}(G - G_{in}) - \frac{\mu_G X}{Y_{X/G}} \tag{5.12}$$

$$\frac{dE}{dt} = -\frac{F_{out}}{V}(E - E_{in}) + \alpha_6 \mu_G - \frac{\mu_E X}{Y_{X/E}} \tag{5.13}$$

To construct the dynamic programming problem described in Section 5.2.2, an objective function that combines the optimization criteria of maximizing the bioreactor productivity while also reaching steady state as quickly as possible is postulated. This objective function ϕ is shown as part of Equation 5.14, where the first term describes the average bioreactor productivity with respect to beta-carotene production. The productivity of a bioreactor is calculated by multiplying the average product concentration $\bar{P}$ by the dilution rate D. The dilution rate is calculated as the ratio of the average outlet flow rate $\bar{F}_{out}$ and the working volume V. The second term in the objective function is included to guarantee the bioreactor reaches a steady state as quickly as possible by minimizing the sum of square errors between the process states and their average state values. The algebraic constraints developed because of the discretization of Equations 5.6 through 5.13, described in Section 5.2.2, are used as the constraints to the optimal control problem; the values for the parameters in these equations are taken from Ordonez et al. [295].

The solution to the optimal control problem in Equation 5.14 is the time profiles for the manipulated variable(s) in the process that maximized the desired objective. These control actions of the manipulated variables are determined *a priori* but can be used experimentally to evaluate the bioreactor performance. In the beta-carotene case study proposed, the objective is to maximize bioreactor productivity and the manipulated

variables are the pump controls that determine the flow rate of each inlet feed. While not performed in this work, the feeding strategy defined by the determined pump control actions can be implemented experimentally through manual manipulation or a programmable controller. The necessary inlet concentrations of the substrates are calculated using simple mass balances based on the optimal flow rates of each inlet stream and the bulk concentration of the corresponding substrate. The predicted time profiles of the bioreactor concentrations are obtained by examining the values of $\boldsymbol{x}(t)$ serving as the constraints of the system. These results are discussed in the next section for each case of the multi-feed bioreactor used for beta-carotene production.

$$\min_{F_1,F_2,F_3} \phi = -\alpha \cdot \bar{P}D + \sum_{q=1}^{nfe}\sum_{i=1}^{ncol}\left[P_{i,q} - \bar{P}\right]^2 + \left[X_{i,q} - \bar{X}\right]^2$$

$s.t.$

$$\boldsymbol{x}(t) = \boldsymbol{x}_{i-1} + h_i \sum_{i=1}^{ncol} \Omega_q \left(\frac{t_i - t_{i-1}}{h_i}\right) \cdot \frac{d\boldsymbol{x}}{dt_{i,q}}$$

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} + h_i \sum_{i=1}^{ncol} \Omega_q \left(\frac{t_i - t_{i-1}}{h_i}\right) \cdot \frac{d\boldsymbol{x}}{dt_{i,q}} \tag{5.14}$$

$$F_{in}(t) = F_{out}(t) = F_1(t) + F_2(t) + F_3(t)$$

$$\chi_E = 1 - 4.1 \times 10^{-6} \cdot E^3 + 1.4 \times 10^{-4} \cdot E^2 - 9.1 \times 10^{-3} \cdot E$$

$$\chi_A = 1 - 0.011 \cdot A^2 - 0.021 \cdot A$$

$$0 \le F_1(t), F_2(t), F_3(t) \le F^{max}$$

$$\boldsymbol{x}(t) = [G(t), X(t), P(t), E(t), A(t)]$$

5.3. Results and Discussion

5.3.1. Optimal Control a Multi-feed β-carotene Bioreactor

One of the benefits of the continuous cultivation process shown in Figure 5.2 is its adaptability to different operating scenarios. As mentioned previously, *S. cerevisiae* SM14 exhibits multiple growth phases; the first growth phase utilizes the main carbon source glucose, resulting in the formation of produces ethanol, biomass and some product, and then a second growth phase utilizes the ethanol and exhibits a higher beta-carotene formation rate. Through restriction of the pump operation in the system $(F_i(t) = 0)$ it is possible to operate the reactor using different feeding strategies and determine which feeds are necessary to achieve the optimal economic performance. Three feeding strategies are tested here, and include (1) the use of a single glucose inlet to simulate the current feeding strategies seen in literature $(F_1(t) = F_3(t) = 0)$, (2) the use of the glucose and media inlets to examine the effects of the independent variation of both dilution rate and glucose inlet concentration $(F_2(t) = 0)$, and (3) the use of the a glucose, ethanol and media inlets to target both growth phases of *S. cerevisiae* to enhance beta-carotene production.

All strategies are tested using a working volume of 3 liters to ensure the minimal amount of change to the conditions of the bioreactor whose effects are not captured by the kinetic models by Ordonez et al., i.e. nutrient levels, dissolved oxygen, etc. The bulk concentration of the glucose and ethanol feeds are 20 g/L and 5 g/L, respectively, and are chosen as these were the highest concentrations of each substrate seen during the work of Ordonez et al. The bioreactor is initially charged with 20 g/L of glucose and

0.06 g/L of biomass containing 6 mg/L of beta-carotene. The ethanol and acetic acid initial conditions are zero. The final constraint to the program is the restriction of possible flow rates for the pumps for the multiple feeds. The value of the maximum flow rate of each pump is determined based on the number of feeds used in the reactor; this sets a standard maximum dilution rate D possible for the reactor, allowing for comparable operation regardless of the number of feeds. For example, assuming a total flow rate of 12 L/s into the reactor is allowed, the maximum pump flow rate would result 12 L/s for a single pump configuration, 6 L/s each for a dual feed configuration having two pumps, and 4 L/s each for a three-feed configuration having three pumps.

The results from the simulation of the optimal control of a continuous bioreactor utilizing a single glucose feed are shown in Figure 5.4 and will be used as a guideline to compare the improved performance of the multi-feed bioreactor. From the dynamic state profiles shown in Figure 5.4a the process comes to an oscillatory steady state after approximately 25 hours of operation. Table 5.1 gives the average concentrations for all five reactor components. Though reactor feeding is allowable in the optimal control program from the start, the reactor operates in batch mode for the first 20 hours to establish an adequate biomass culture. After that initial 20 hours, the glucose pump begins to operate, reaching a steady state control policy as shown in Figure 5.4b. The corresponding feed concentration of glucose is shown in Figure 5.4c. The optimal control policy for the single glucose feed shows the pump exhibits cyclic, pulsing operation over the course of each hour of operation, starting at the maximum possible flow rate of 1.2 L/hr, decreasing to 0.4 L/hr in the second part of the hour, and then

closing the pump completely in the final third of the hour. The cyclic nature of the optimal feeding profile is independent of the finite element length, though this value does affect the final steady state values of the state variables and as such should be optimized to achieve the maximum profitability in future work.

Table 5.1. Average steady state bioreactor compositions for the single glucose feed, glucose and media feed, and glucose, ethanol and media feed reactors.

Concentration (units)	Glucose Only Feed	Glucose and Media Feed	Glucose, Ethanol and Media Feed
Beta-carotene (mg/L)	20.56	18.62	17.35
Biomass (g/L)	3.62	2.79	2.36
Glucose (g/L)	2.98	1.53	1.38
Ethanol (g/L)	4.19	2.84	3.14
Acetic Acid (g/L)	0.39	0.31	0.28

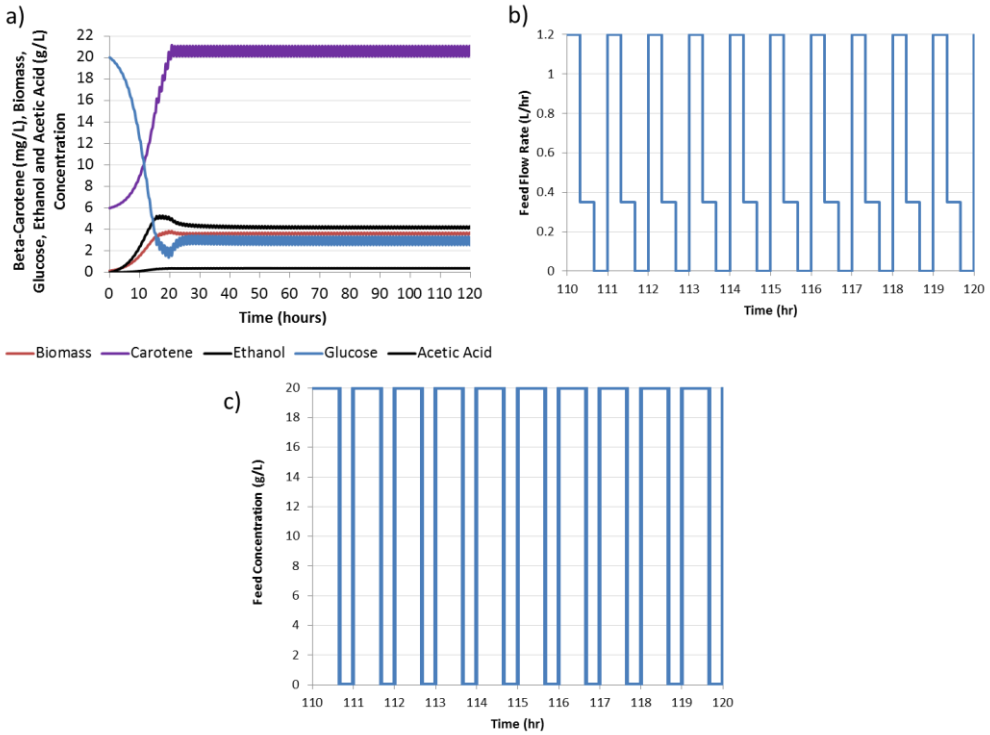

Figure 5.4. (a) State profiles, (b) steady state feed flow rates and (c) steady state feed composition for a single feed bioreactor (glucose only) for beta-carotene production.

Figure 5.5 displays predicted control policy and resulting state profiles of adding a secondary controllable media feed as a diluent along with the initial glucose feed. The optimal control policies for the glucose pump shown in Figure 5.5b follows the same hourly, oscillatory trend as the single glucose feed control policy. The flow rate begins at the maximum glucose flow rate of 0.4 L/hr, dropping to 0.2 L/hr and then finally closing for the final 20 minutes of each hour of operation. The media feed is also introduced in an oscillatory fashion, spiking to the maximum possible flow rate of 0.6 L/hr for the first one third of the hour and then closing the pump for the remaining 40 minutes. The glucose concentration shown in Figure 5.5c depicts the effect of the multiple feeds on the inlet glucose concentration. While both the media and glucose feeds are active we can see a reduction in the optimal glucose concentration being fed into the reactor. This value then increases to the maximum possible concentration of 20 g/L until both pumps are closed. The result of this oscillatory optimal control policy on the bioreactor states are shown in Figure 5.5a. Like those of Figure 5.4a for a single glucose feed, the bioreactor states for the duel feed reactor show oscillatory behavior after reaching an oscillatory steady state at around 60 hours. This bioreactor also sees an initial period of 20 hours where the reactor is operated in batch mode with no inlet or effluent flow. The inclusion of a second dilution feed does decrease the steady state concentrations of the key reactor products; the beta-carotene titer shows a slight decrease from 20.56 mg/L to 18.62 mg/L.

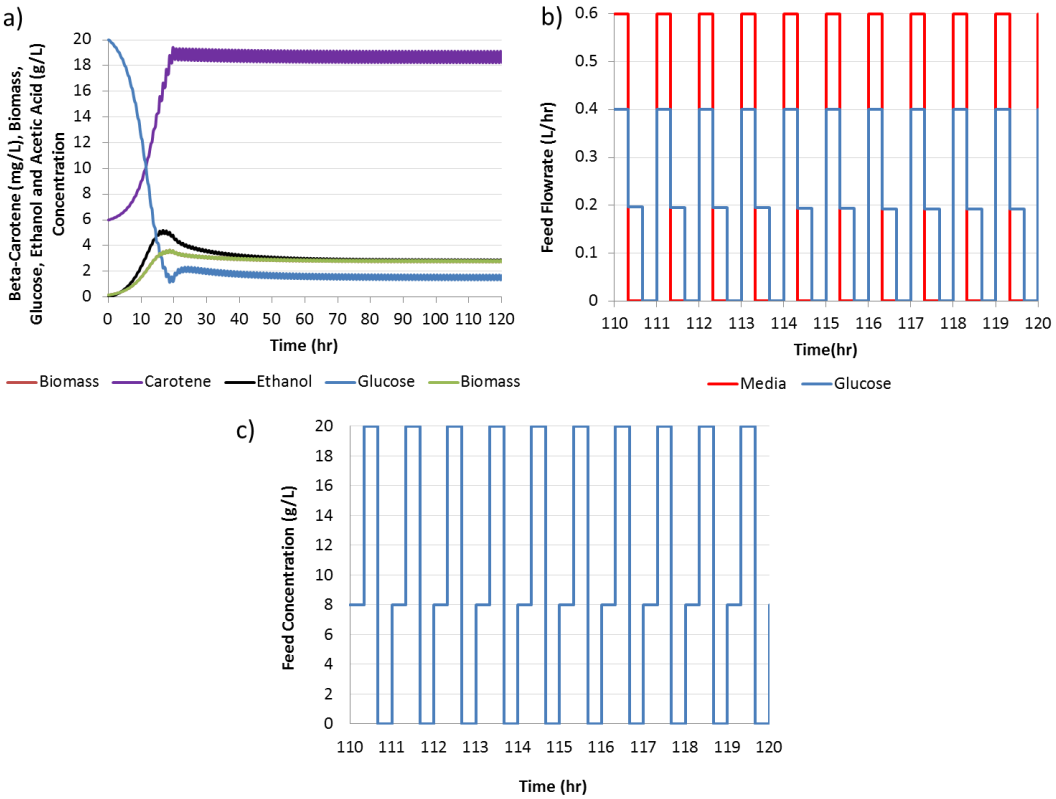

Figure 5.5. (a) State profiles, (b) steady state feed flow rates and (c) steady state feed composition for a two-feed bioreactor (glucose and media) for beta-carotene production.

The predicted control of a three-feed bioreactor utilizing independent glucose, ethanol, and media feeds is shown in Figure 5.6. Figure 5.6b shows that this system with three independent, controllable pumps again displays a pulsing feeding strategy. In this case, the ethanol, glucose and media are introduced to the bioreactor together at the maximum allowable flow rate of 0.4 L/hr each. During the middle one third of the hour the ethanol and media pumps are closed for the remainder of the hour while the glucose pump flow rate is halved to approximately 0.2 L/hr. During the final third of each hour all three pumps are closed completely. This feeding pattern gives the feed concentration profile shown in Figure 5.6c, where the initial feeding of glucose and ethanol are done well below the maximum concentration of 20 g/L and 5 g/L, respectively. The glucose

inlet concentration does reach the maximum during the period at which the ethanol and media pumps are switched off. The optimal feeding policy again results in an oscillatory steady state exhibiting the same characteristics of the previous two bioreactor configurations. However, the introduction of the third feed again reduced the beta-carotene titer when compared to the single glucose feed and glucose and media feed reactor arrangements.

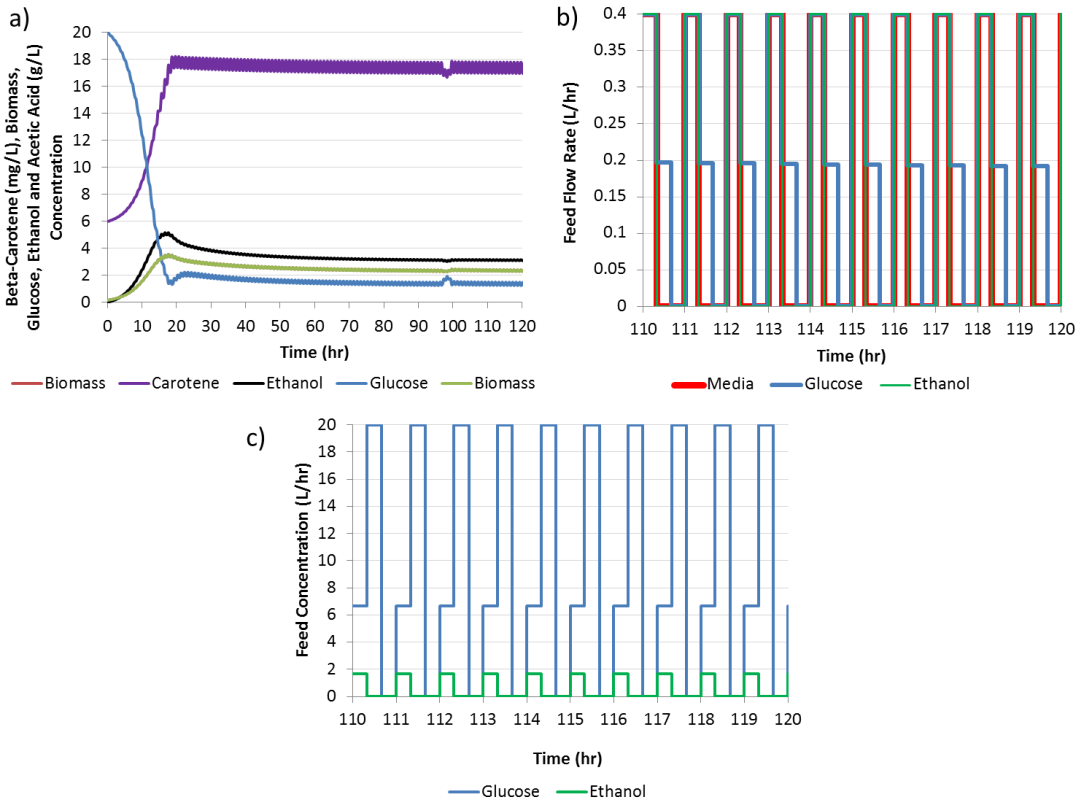

Figure 5.6. (a) State profiles, (b) steady state feed flow rates and (c) steady state feed composition for a three-feed bioreactor (glucose, ethanol and media) for beta-carotene production.

Due to the similarity of the optimal control policies for reactor flow rates and the resulting state profiles systems shown in Figures 5.4 through 5.6, a more in-depth analysis of the improvement through the introduction of multiple feeds is required. Table

5.2 compiles these more advanced analytics beginning with the calculation of the average bioreactor productivity. This is calculated as follows for continuous systems, where $\overline{F_i}$ indicates the average flow rate for feed i:

$$\text{Average Productivity} = \overline{P} \cdot \overline{D} = \overline{P} \cdot \frac{\sum_i \overline{F_i}}{V} \tag{5.15}$$

In addition to its effect on the beta-carotene titer ($\overline{P}$ in Equation 5.15), the multiple feeds affect the optimal average dilution rate ($\overline{D}$ in Equation 5.15). These variations in titer and dilution rate lead to differences in productivity. The single glucose feed results in the highest beta-carotene productivity, as it shows the largest titer and dilution rate among all the reactor options at 3.61 $mg \cdot L^{-1} \cdot hr^{-1}$. The addition of more feeds to the system results in a 29.75% and 23.53% decrease in productivity for the two-feed and three-feed configurations, respectively. This difference, while partially attributed to the decrease in product titer, is also a result of a lower average dilution rate, a result that suggests that the multi-feed bioreactors require longer residence times. However, when compared to the results of Ordonez et al, all three bioreactor configurations reach a higher beta-carotene productivity than the 1.67 $mg \cdot L^{-1} \cdot hr^{-1}$ achieved during batch production [295]. Based on these results it seems that the traditional single glucose feed is the ideal method for the continuous production of beta-carotene.

Table 5.2. Comparison of the single, double, and three-feed reactors via key operation metrics, i.e. product titer, productivity, cellular and substrate yields, etc.

	Glucose Only Feed	Glucose and Media Feed	Glucose, Ethanol and Media Feed
Average Beta-carotene Titer (mg/L)	20.57	18.62	17.35
Average Dilution Rate (hr^{-1})	0.176	0.136	0.159
Average Productivity (mg/L·hr)	3.61	2.54	2.76
Breakeven β-Carotene Price ($/kg)	$1,258.58	$688.27	$873.47
Cellular Yield (mg/g biomass)	5.69	6.67	7.33
Substrate Yield (mg/g substrate fed)	1.31	3.09	2.16
Productivity Difference	0.00%	-29.75%	-23.53%
Price Difference	0.00%	-45.31%	-30.60%

While the multi-feed bioreactors underperform compared to their single-feed counterpart in terms of productivity (labeled in green in Table 5.2), they show vast improvement when it comes to economic viability (labeled in orange in Table 5.2). To determine the economics of these reactors, and knowing that the capital cost of these systems were similar, the breakeven cost for the beta-carotene product was calculated based on their operating costs. Equation 5.16 was used to calculate the breakeven cost of the average amount of beta-carotene produced C_P based on the average amount of each substrate necessary for its production. For these calculations the price of the ethanol C_E was set to be approximately $3 per liter ($1 per gram) [296]. The price of the glucose C_G was set to be $0.40 per pound ($0.88 per kilogram) based on the reports available from the United States Department of Agriculture for the commodity cost of glycose syrup based on dry weight [297].

$$C_P \bar{P} \bar{F}_{out} = C_G G_f \bar{F}_1 + C_E E_f \bar{F}_2 \tag{5.16}$$

The benefits of the multi-feed approach to the production of biologics can be seen through the calculation of this breakeven analysis. While the single feed excelled when comparing the productivity of the reactor systems, it falls far behind in the economics. The addition of both the ethanol and media feeds reduces this breakeven cost by over 30%. This level of price reduction outpaces the decrease in reactor productivity and is a result of an improvement in both the cellular and substrate yield of the bioreactors with multiple feeds. As expected, the supplementation of ethanol targets the secondary growth phase of the *S. cerevisiae* and results in a 64.86% increase in substrate yield, allowing for the decrease in operating cost despite the decrease in overall reactor productivity.

5.3.2. Optimal Solution Verification and Investigating Increased Operability

Additional analysis was performed to verify the optimality of the obtained solutions and to examine the effect of additional operability constraints on the system. With respect to optimality, it was important to determine whether the parameters of the discretization method, specifically the chosen finite element length, were influential on the optimal exit concentrations of the system. Figure 5.7 below shows the effect on the average exit concentrations of each of the five state variables for a three-feed reactor with an ethanol, glucose, and media feed. Changes in the finite element length vary from 0.1 hours to 2 hours. Three internal collocation points were used for each finite element length, and the percent change in the optimal steady state exit concentrations were calculated using the one-hour finite element length as a basis. From Figure 5.7 below it can be seen that the optimal steady-state concentrations vary slightly based on finite

element length, with the highest variation of ±1.5% shown in the glucose concentration. All other changes showed sensitivities at or below ±0.5%. The steady state optimal control policies (not shown) were all similar to the cyclic optimal control policy shown for the 1 hour case in Figures 5.4, 5.5, and 5.6. These results suggest that the cyclic optimal control policy for the three independent flowrates is actually a set of cyclic solutions based on the chosen finite element length that result in the same optimal state profiles for the multi-feed bioreactor system.

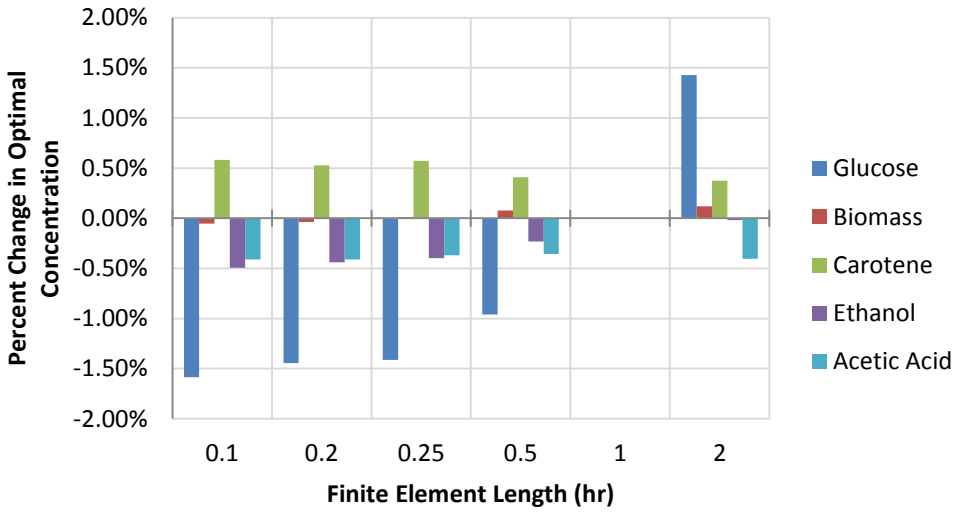

Figure 5.7. The effect (calculated in percent change) of the finite element length on the optimal solution to the optimal control problem for the three-feed bioreactor.

With the optimality of the solution determined, it is also important to investigate the role operability plays in the optimal solution of the multi-feed bioreactor. Bang-bang control schemes like those seen in the optimal control policies for the multi-feed bioreactor often lead to faster degradation of equipment due to the constantly changing values of the manipulated variables between their maximum and minimum values. Often it is more amenable to operate within small windows of change, resulting in less wear on

equipment such as pumps and valves. One method to determine the impact of operability on the optimal control is to penalize large changes in the manipulated variable values in the objective function of the optimal control problem. This can be seen specifically applied to the previously determined optimal two-feed bioreactor problem in Equation 5.17 below, where the ϕ_{free} term denotes the original objective function from Equation 5.14 and the second term is used to penalize large, quick changes in the flowrates of the glucose feed and media feed, respectively. The parameter ξ is a chosen scaling parameter to adjust the importance of the operability constraint; a value of $\xi = 0$ would result in the original free problem.

$$\min_{F_1,F_2,F_3} \phi = \phi_{free} + \xi \cdot \int_0^{t_f} \left[\left(\frac{dF_1}{dt} \right)^2 + \left(\frac{dF_3}{dt} \right)^2 \right] dt \tag{5.16}$$

Once discretized, the objective function in Equation 5.16 can be used to limit the allowable solution space to optimal control policies to those that decrease the wear on the flow valves. The resulting optimal state profiles can then be used to determine the effect of increased operability. Figure 5.8 below shows the optimal solution for the original "free" problem discussed in the previous section, the problem that considers operability with changes in the manipulated variables at each collocation point using Equation 5.16 as the objective function, and the problem that considers operability while also limiting changes in the manipulated variable to each finite element and remaining constant at each collocation point. Comparing the optimal control policy in Figure 5.8a for the free problem to Figure 5.8b denoting the operability problem it is easy to see that the cyclic nature remains in the problem where operability is considered. However, the

magnitude of the changes has been limited when operability is considered, and the resulting bang-bang control seen in the free problem is no longer encountered. If Figure 5.8c is considered, when operability is taken into account with fewer opportunities to change the manipulated variable, it can be seen that a constant glucose flowrate is used to operate the reactor without the presence of a media stream for dilution effects.

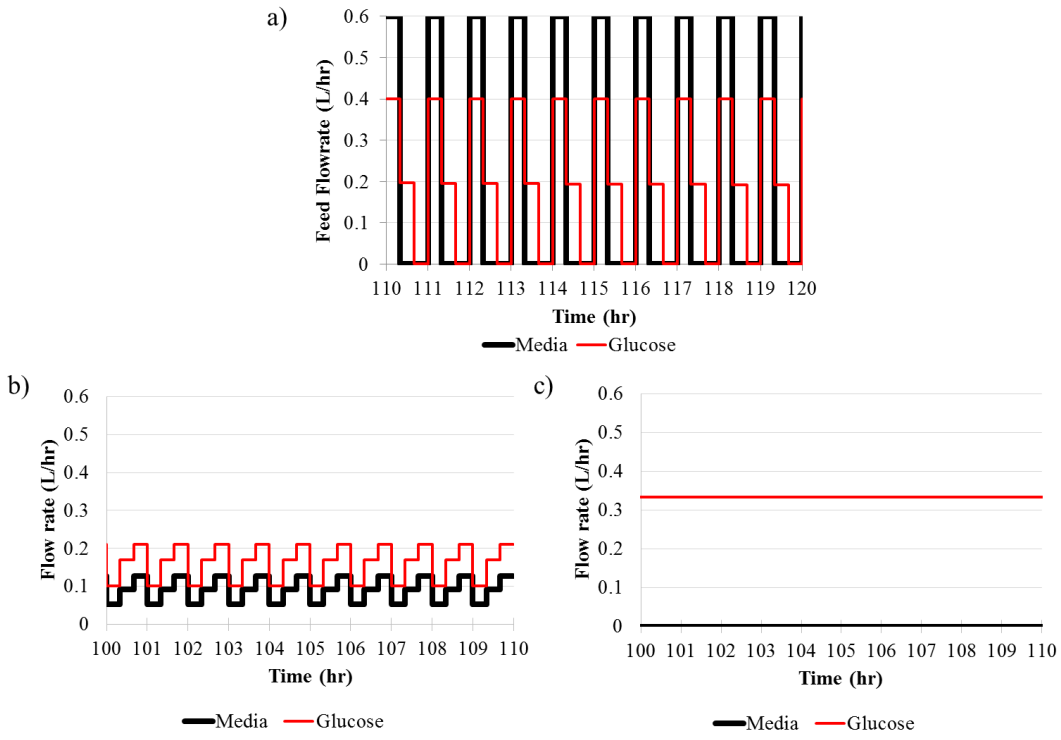

Figure 5.8. Effect of adding operability penalties for the allowable flow rate changes to the objective function of the optimal control problem. Plots correspond to: (a) free manipulation of independent variables every 20 minutes, (b) penalties placed on large changes in the manipulated variables to improve controllability, and (c) penalties placed on changes in manipulated variables and only allowing for infrequent changes corresponding to once every hour.

Table 5.3. Effect of penalizing and limiting the number of changes in the independent, manipulated variables (MV). Columns correspond to freely manipulated independent variables, penalizing large changes in the manipulated variables to improve controllability, and penalties placed on changes in manipulated variables and only allowing for infrequent changes corresponding to once every hour

	Infrequent & Penalized	Penalized	Free
Average Exit Concentrations			
Glucose (g/L)	4.648	1.584	1.530
Biomass(g/L)	3.090	2.466	2.790
Beta-carotene (mg/L)	13.461	15.660	18.622
Ethanol (g/L)	4.230	2.640	2.836
Acetic Acid (g/L)	0.324	0.272	0.309
Average Dilution Rate (hr^{-1})	0.112	0.084	0.136
Beta-Carotene Productivity (mg/L-hr)	1.501	1.315	2.539
Breakeven Beta-carotene Price ($/kg)	$1,292.61	$708.64	$688.27

While visible effects can be seen in the optimal control policies when operability is considered, it is important to quantify their effects on the operation of the bioreactor. These effects, including the average exit concentration of the five state variables, average dilution rate, β-carotene productivity, and calculated breakeven β-carotene price are shown in Table 5.3 below. When comparing the free manipulation of independent variables to the more constrained operability problem, it can be seen that the calculated breakeven β-carotene price is increased by approximately $20 per kilogram when operability is considered. This a result of a lower productivity of the reactor, a decrease from 2.539 hr^{-1} in the free case to 1.315 hr^{-1} in the case where the manipulated variables are penalized for operability. It can also be seen that the optimal average exit concentration of the β-carotene product is reduced by about 3 mg/L, as highlighted in green. However, if the ability to manipulate the flowrates is reduced to once per hour, then the system reverts to a single, constant glucose flowrate and the breakeven β-

carotene price rises substantially to a value of almost $1300 per kilogram. This is a direct result of both a further decrease in β-carotene production and a rise in the exit glucose concentration, highlighted in orange, demonstrating that a constant glucose feed does not allow for adequate residence time for the glucose to be consumed by the *S. cerevisiae*. The need for more glucose that cannot be used in the single pass bioreactor results in higher operating cost and thus a higher breakeven β-carotene price. This outcome demonstrates the need for a multi-feed approach that can be used to manipulate the glucose concentration based on the metabolic cycle of the organism.

5.3.3. Summary of Future Analysis

Ethanol substrate is inhibitory to cellular growth and its introduction to the reactor, as well as being a by-product of glucose metabolism, results in the subpar performance of the three-feed reactor when compared to the dual feed system. By eliminating the ethanol feed and using the media feed to manipulate the inlet glucose concentration and control the amount of produced ethanol while still maintaining the necessary dilution rate, the dual feed reactor can maximize its cellular yield of beta-carotene and achieve a breakeven price that is 45.31% lower than the single feed reactor despite having the lowest productivity of the three reactor systems. While further analysis of the long-term economics of these reactors is needed, the results discussed here show great promise for stimulating economic improvement of biochemical systems through the implementation of multiple feed bioreactors.

The results of this work have direct applicability to the control of many biochemical reactor systems. The optimal feeding profiles developed here can be used

directly with cultivation experiments using a programmable controller or manual manipulation of the pump flow rates. If online feedback control is desired, the optimal control models discussed in Equation 5.14 can serve as the beginnings of a model predictive control (MPC) framework that can be developed for direct online control of a bioreactor system. In this framework, the bioreactor concentrations can be directly measured in the instance of extracellular products or estimated from online data using different state estimation techniques, i.e. Kalman filters, in the instance of intracellular products. This concentration can be directly related to the concentration predicted by the controller, thereby giving an error signal that can be used to adjust the multiple pumps controlling the various substrate flow rates. Finally, the general nature of the control framework can be adjusted to any biochemical system, regardless of the intracellular or extracellular nature of the product, if kinetic equations can be developed for the system of interest.

5.4. Conclusions

This work looks to develop a novel multi-feed paradigm for the economic improvement of biochemical processes. This new biochemical processing strategy aims to maintain process productivity while decreasing operating costs through the introduction of multiple independent feeds with associated pumps that allow for the manipulation of substrates and diluent flow rates and substrate concentrations. A beta-carotene production process is utilized as a case study for the new paradigm utilizing a traditional single glucose feed, two separate glucose and media feeds, or three feeds consisting of glucose, ethanol, and media as substrates. Kinetic models taken from

literature are utilized as constraints to the optimal control problem looking to determine the individual feed flow rate control policies necessary to maximize reactor productivity as well as minimize the response time to steady state. Results of this study show that while the introduction of multiple feeds reduces the bioreactor productivity through decreases in product titer and dilution rate. However, the implementation of these feeds reduces the breakeven beta-carotene price due to an increase in the cellular and substrate yields of the product. This result suggests that the use of a multi-feed bioreactor system can be used to insight cost savings in bioreactor system through the reduction of operating costs associated with the production of biologics.

CHAPTER VI

PROCESS SYNTHESIS FOR COMMERCIAL-SCALE CELLULOSIC BIOETHANOL PRODUCTION[‡]

6.1. Introduction

Even with the recent discoveries of shale gas, concerns over the future availability of petroleum-based liquid fuels exist due to the high demands of the transportation industry. Renewable liquid fuel sources such as ethanol will play a key role in filling the gap left by declining fossil-fuel-based options. At present, popular methods of producing ethanol rely upon the conversion of food sources with high starch content, such as corn and sweet potatoes, resulting in the debate of the necessary sustainability if used in both the food and fuel markets [298]. An alternative method is the thermochemical conversion of lignocellulosic biomass to ethanol, which has been shown to be competitive with current petroleum prices, but these processes are associated with high energy requirements for operation [299]. The production of bioethanol from lignocellulosic sources can also be done biochemically, but this method of production introduces operational and other, e.g. mass transfer, issues. For example, to produce high concentrations of bioethanol through fermentation one needs to have high concentrations of glucose and xylose, which are only obtainable using a high concentration (e.g. 25-30 wt%) of pretreated lignocellulosic biomass that is processed with the appropriate enzymes [300]. However, the biochemical production of ethanol

[‡]Reprinted with permission from "Economic viability of consolidated bioprocessing utilizing multiple biomass substrates for commercial-scale cellulosic bioethanol production" by Raftery and Karim, 2017. *Journal of Biomass and Bioenergy*, 103, 35-46, Copyright 2017 by Elsevier.

requiring high solid loadings has the potential to utilize many different cellulosic biomass substrates and pretreatment technology options without the use of food resources, giving it the versatility to replace many liquid fuels produced from crude oil.

Research in the techno-economic analysis of biochemical ethanol production has focused mainly on processes involving conversion technologies featuring three main steps: biomass pretreatment, enzyme hydrolysis, and fermentation. The effect of many different pretreatment technologies, including ammonia fiber expansion (AFEX), dilute sulfuric acid (DSA) and liquid hot water (LHW), on the economic viability of lignocellulosic ethanol production when using a switchgrass feedstock has been demonstrated by the Consortium for Applied Fundamentals and Innovation (CAFI) [301]. Dutta et al. have demonstrated the economic viability of biochemical ethanol production in various process configurations using corn stover and a Zymamonas mobilis bacterium to facilitate the fermentation [302]. In 2011, the National Renewable Energy Laboratories (NREL) showed that the conversion of corn stover to ethanol using a dilute acid pretreatment can result in an attractive ethanol selling price [303]. Zondervan et al. studied the optimal plant topology to produce ethanol alone as well as with coproduction of butanol and succinic acid [304]. Gabriel et al. used mixed-integer nonlinear programming methods to determine the optimal biomass choice and pretreatment necessary for a single-feed process from a process superstructure to analyze the viability of bioethanol production [96]. Furthermore, Martín and Grossmann have studied the production of ethanol from a switchgrass feedstock, first by minimizing energy needs of the process and then simultaneously producing i-butene [305, 306]. Da

Silva et al. have investigated the techno-economics of different pretreatment processes for cellulosic ethanol production at the industrial scale [307]. Additional work has been done regarding the selection of technologies specific to bioethanol production [308, 309]. While the studies described above have all shown promising results to produce bioethanol via biochemical pathways, in most cases they are limited by their focus on separate hydrolysis and fermentation (SHF) as a processing methodology.

The SHF process for the biochemical conversion of lignocellulosic biomass requires an enzymatic saccharification step to convert the cellulose chains to simple sugars, a costly venture and a common bottleneck in their implementation [310]. An alternative method discussed in the literature aims to circumvent this costly approach in favor of using a single organism or consortium of organisms capable of hydrolysis via enzyme production and simultaneous fermentation of the cellulose polymer to ethanol in a single process step; this process is known as consolidated bioprocessing or CBP [311]. This process has been analyzed at the laboratory scale using many different types of organisms, including the genetically engineered bacteria *Caldicellulosiruptor besci*i, *Clostridium phytofermentans,* and *Saccharomyces cerevisiae,* to name a few [312-314]. Ryu and Karim have also analyzed the use of genetically engineered E. coli for creating a whole cell biocatalyst for the single-step conversion of lignocellulose to ethanol [315]. However, this alternate conversion route has the potential of further increasing the economic potential of ethanol production via biochemical means, but no large-scale techno-economic analysis of consolidated bioprocessing has yet been reported in the literature.

In addition to their emphasis on SHF as a processing methodology, much of the literature on biochemical ethanol production focuses on the use of a single biomass feedstock for operation. However, work by Quaglia et al. has shown that a mixed biomass feed of cassava rhizome, corn stover, and sugarcane is economically favorable to many single biomass feeds [316]. Marvin et al. have also investigated the use of multiple biomass feedstocks to determine the optimal ethanol supply chain in Iowa and the surrounding states [317]. The development of a single biorefinery capable of utilizing multiple feedstock options simultaneously would facilitate its application in many more areas of the United States than just those rich in a single biomass type, e.g. corn in the Midwest.

To further improve the economic viability for biochemically produced ethanol, the otherwise unprofitable waste product and key component in the cell walls of the lignocellulosic biomass, lignin, is used to generate cogenerate energy in the form of steam utilities and salable electricity. As much as 25% of the biomass used in these processes, regardless of specific type, is composed of this highly complex polymer that cannot be converted to the desired ethanol product. Repurposing this lignin for energy cogeneration has been shown to be feasible, resulting in the reduction of plant operating costs for utilities [6, 20]. Steam produced through the burning of this otherwise unusable lignin can also be used for electricity generation, and has been shown to benefit second generation biofuels profitability by Dias et al [318]. The implementation of cogeneration has the capability of reducing the energy requirements of the plant while increasing the

potential of providing green electricity to the grid from any excess energy produced during the lignin combustion process.

This work looks to be the first to analyze the techno-economics of large scale consolidated bioprocessing and compare it to the more conventional separate hydrolysis and fermentation pathway. A process superstructure of biomass alternatives and pretreatment technologies is considered for each conversion pathway, allowing for the selection of an optimal feed consisting of a single biomass or multiple biomasses with corresponding optimal pretreatments. Steady state mathematical models are utilized to formulate a mixed-integer nonlinear optimization (MINLP) problem whose solutions will be compared to determine the economically superior option based on minimum ethanol selling price. The long term economic implications of the superior processing technology as well as the impact of utilizing multiple biomass feedstocks simultaneously are examined. Additionally, the optimal topology of the ethanol production process is used to elucidate the ideal location for biorefinery infrastructure based on the selected biomass types necessary for economically optimal ethanol production.

6.2. Materials and Methods

6.2.1. Biomass Selection and Geographic Implications

This study focuses on the development of consolidated bioprocess (CBP) as a method of producing ethanol from lignocellulosic biomass and its techno-economic comparison to the existing separate hydrolysis and fermentation (SHF) methodology. Each process incorporates the options of processing multiple biomass feeds and the cogeneration of utilities and electricity through the use of the unreacted lignin product.

The processes are designed to handle a total of 2,000 to 3,500 U.S. tons (1,800 to 3,200 metric tons) per day of biomass and operates for 350 days per year. Biomass options focus on prevalent biomass options available in the United States, including sugarcane bagasse (SB), sorghum (SOR), hybrid poplar (HP), switchgrass (SW) and corn stover (CS).

Table 6.1. Oven-dried composition of each biomass type considered in the proposed separate hydrolysis and fermentation and consolidated bioprocessing bioethanol facilities [96, 319-321].

Biomass component	Corn stover	Switchgrass	Poplar	Sorghum	Bagasse
Glucan	0.3546	0.3325	0.4373	0.2825	0.3911
Mannan	0.0047	0.0032	0.0281	0.0016	0.0032
Galactan	0.0087	0.0107	0.0070	0.0046	0.0045
Xylan	0.1919	0.2186	0.1720	0.1306	0.2023
Arabinan	0.0231	0.0291	0.0055	0.0148	0.0158
Water	0.0271	0.0215	0.0318	0.0365	0.0286
Lignin	0.1869	0.1806	0.2355	0.1372	0.2430
Other	0.2032	0.2039	0.0827	0.3923	0.1115
Price per tonne	$61.60	$81.50	$101.60	$173.90	$40.00

While it is acknowledged that the process of drying the biomass would add both capital and operating costs to the process designs, thereby increasing the minimum selling price of the ethanol produced, these factors are not considered in this work. Biomass moisture content at harvest is extremely variable due to factors such as growth region, the weather of a harvest year, soil type and fertilization, as well as harvesting and storage practices [303]. This work assumes the off-site drying of biomass before delivery to the processes developed here, focusing on the use of oven-dried biomass as a raw material. The biomass prices used in this work are chosen to reflect the additional preprocessing of the biomass. In this way, the uncertainty associated with biomass moisture and the drying process is eliminated and can be considered in more rigorous

designs of the optimal biorefinery developed here. Table 6.1 outlines the average oven-dried composition and purchase cost of each type of biomass considered in this work [320]. Prices (in $ per tonne) are taken from Gabriel et al. for corn stover, switchgrass and poplar, NREL for sugarcane bagasse, and the World Bank Group for sorgum [96, 319, 321].

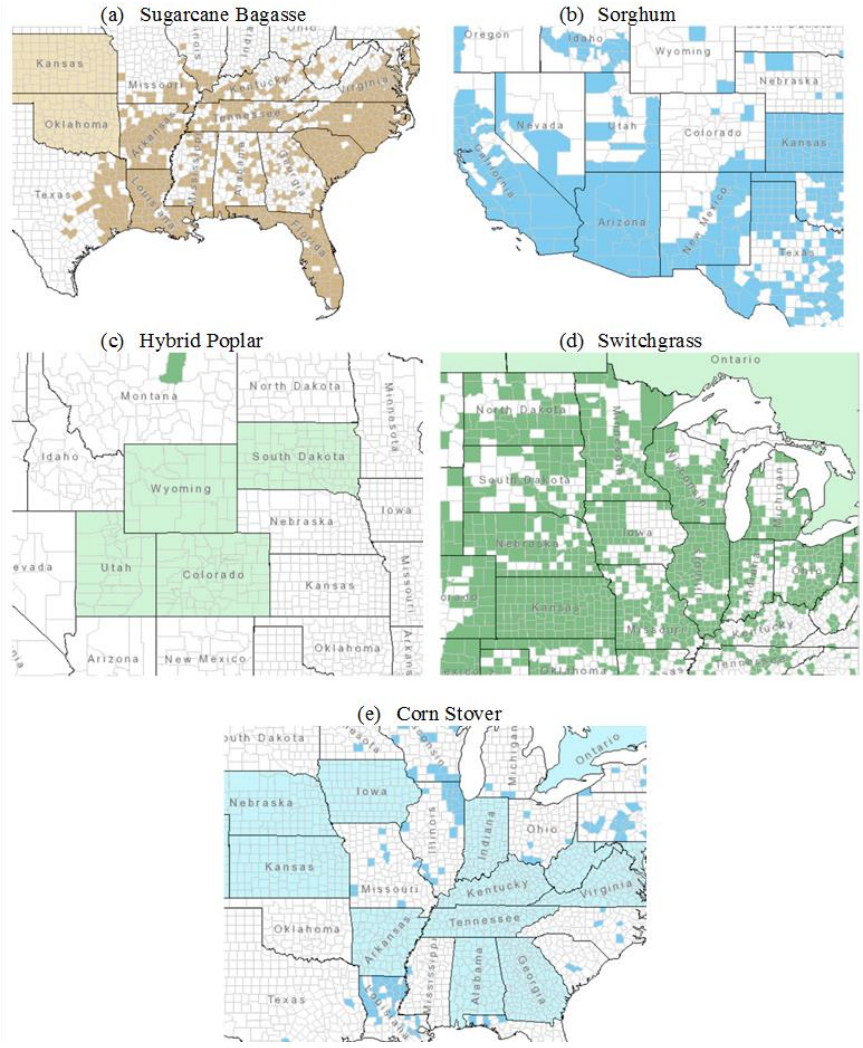

Figure 6.1. Predominant location in the United States to produce each biomass option considered as a raw material for ethanol production [322].

The economics of the biochemical conversion of biomass to ethanol is largely linked to the type(s) of biomass used in the optimal ethanol process. Figure 6.1 shows

the main regions of the United States in which each of these biomass types is predominantly found. The types of biomass available for processing is highly dependent on the location at which the plant is constructed, e.g. sorghum is likely to be the main feedstock for a bioethanol facility constructed in southern California or Arizona, while a plant constructed in the Midwest would be more apt to use switchgrass or corn stover as its raw material [322]. Likewise, for facilities that utilize multiple biomass options it would be beneficial to be built in an area such as Kansas, where multiple biomass types are readily available. As such, the elucidation of the optimal biomass selection is directly representative of the optimal construction location for the facility. Thus, the choice of biomass feedstock(s) in the optimal process topology can be used to clarify the optimal location of biorefinery construction.

6.2.2. Pretreatment Selection

Pretreatment is essential for any process involving the biochemical conversion of biomass to commodity chemicals. Various chemical and physiochemical types of pretreatment can be used to weaken the cell walls and facilitate the hydrolysis and fermentation steps downstream. Pretreatment with alkali chemicals, e.g. sodium hydroxide, causes a swelling of the cell walls that leads to an increase in the internal surface area, a decrease in the degree of polymerization of the cell wall, and the disruption of the lignin structure responsible for the rigidity of plant cell walls [323]. Treatment with dilute acid mixtures using sulfuric or hydrochloric acid is also common; these treatments hydrolyze the hemicellulose to its monomeric sugar units and make the cellulosic portion of the biomass more accessible to enzymatic action [323].

Physiochemical pretreatments, typically using of steam, liquid hot water, or aqueous ammonia, rely on more mechanical methods for cell wall disruption. During these processes the cell walls are subject to the working chemical under high pressures, facilitating hemicellulose hydrolysis, and then the pressure is released to cause a disruption in the cell walls, making the cellulose portion of the biomass available for further enzyme hydrolysis [323].

Specific pretreatment options considered in this work include sodium hydroxide ($NaOH$), liquid hot water (LHW), dilute sulfuric acid (DSA) and ammonia fiber explosion ($AFEX$), and at least two of these options are considered for each type of biomass feed. All combinations of the biomass and pretreatment combinations considered in this work, as well as their operating temperature, conversion of xylan during pretreatment, and conversion of glucan during subsequent enzyme hydrolysis steps, are shown in Table 6.2.

6.2.3. Process Design Superstructures

Figure 6.2a shows the process design superstructure used in the formulation of the design models for the consolidated bioprocessing process. While a feed of multiple biomass types to the process is possible, each biomass is treated separately through its respective pretreatment and hydrolysis pathway. The use of independent processing streams before fermentation allows for an accurate depiction of the pretreatment and hydrolysis steps of each biomass type, whereas the use of a single mixed biomass stream would introduce complications (e.g. kinetic and mass transfer issues) into the process modelling not accounted for in reported literature data.

Table 6.2. Biomass and corresponding pretreatment options for the proposed bioethanol plant design. Xylan conversion refers to the pretreatment process and glucan conversion refers to the hydrolysis process.

Biomass	Pretreatment	Conversion		Temperature (°C)	Reference
		Xylan	Glucan		
Sugarcane bagasse	NaOH	96%	87%	110	[324]
	LHW	44%	42%	190	
Sorghum	DSA	25%	54%	100	[325]
	NaOH	68%	85%	50	[326]
Hybrid poplar	DSA	90%	82%	160	[327]
	LHW	22%	65%	200	[328]
Switchgrass	AFEX	92%	70%	100	[329]
	DSA	85%	76%	160	[327]
	NaOH	84%	68%	50	[330]
Corn stover	AFEX	99%	80%	90	[331]
	DSA	91%	90%	160	[327]

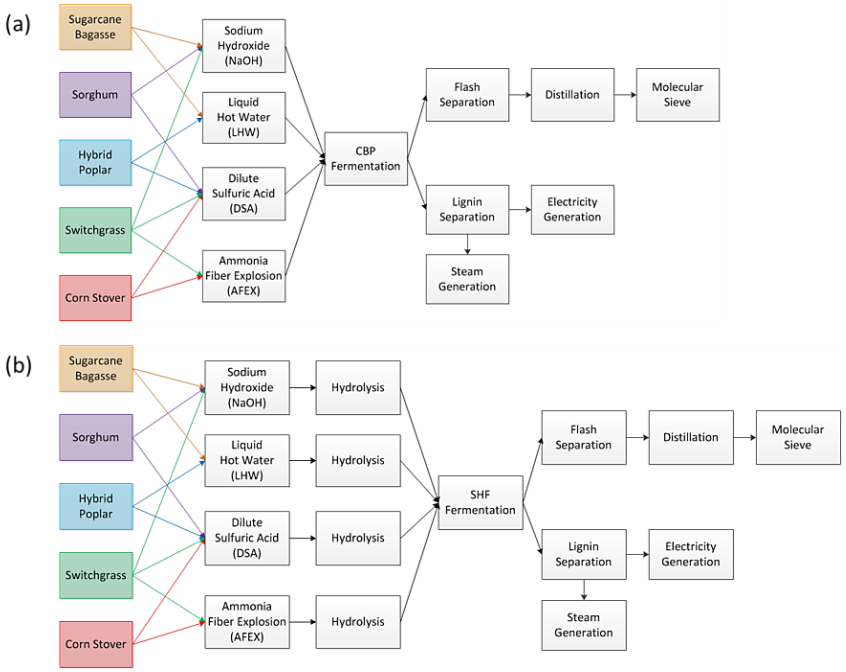

Figure 6.2. Biomass and associated pretreatment options considered in plant superstructures for (a) consolidated bioprocessing (CBP) and (b) separate hydrolysis and fermentation (SHF).

As no single organism has shown the capabilities of utilizing cellulose and hemicellulose as well as the fermenting the resulting monomer sugars, it is suggested

that a consortium of organisms that collectively can complete these tasks are used [332, 333] Literature has shown that *Thermoanaerobacterium saccharolyticum*, a bacterium capable of hydrolyzing the hemicellulose and fermenting the five-carbon sugar monomers, and *Clostridium thermocellum*, an organism capable of hydrolyzing and then fermenting the monomers from the cellulose polymers, show potential as a consortium for CBP fermentation due to their conditions of pH and temperature under which both experience optimum growth, cellulose and hemicellulose decomposition, and ethanol production [334-336]. In 2011, work done by Argyros et al**.** demonstrated that CBP fermentation can be done with a co-culture of *T. saccharolyticum* and *C. thermocellum*, and their results and methods will serve as a basis for the analysis done here [336].

The removal of the carbon dioxide and excess nitrogen from the fermenter, though usually built into the vessel, is modeled in this work using a flash tank. A distillation column is used to concentrate the flash tank effluent to the azeotropic level of ethanol and water as well as remove the solid portion of the beer. Finally, a molecular sieve is used to complete the dehydration of the ethanol to the desired 99.5% product. The biological waste found in the bottoms of the beer column is used in the production of energy. This stream is first centrifuged to separate the solid portion of the stream from the aqueous phase. The solids are then separated with a portion going to the boiler to provide energy to produce steam from the aqueous phase. The rest of the solids are used as fuel for a separate steam generation cycle that generates electricity to be sold as an additional product.

Figure 6.2b shows the process design superstructure used for the separate hydrolysis and fermentation process used to compare to the consolidated bioprocessing. This process begins with the same biomass and pretreatment options used in the consolidated bioprocessing superstructure as shown in Figure 6.2a. An additional hydrolysis step is facilitated using cellulase enzymes capable of converting the long sugar polymers to their simple sugar monomers. The streams that exit the separate hydrolysis units are combined before the fermentation process. A strain of *Zymomonas mobilis*, genetically engineered for high ethanol tolerance and titers, is used for the mixed-sugar, anaerobic fermentation [303]. After fermentation, the reactor effluent undergoes the same ethanol recovery and energy generation procedure as described previously.

6.2.4. Process Modeling

6.2.4.1. Objective Function and Cost Correlations

The objective of this work is to minimize ethanol selling price (MESP) determined as part of the total cost function shown in Equation 6.1, which must offset the capital cost (CC), operating cost (OC) and labor cost (LC) of the plant. Labor costs are scaled linearly with ethanol production, as shown in Equation 6.1b, at a rate of about \$0.17 per gallon of ethanol as given by Humbird et al [303]. Additional revenue from electricity generation ($\mathrm{G_{elec}}$) by the plant is factored in to the overall economics of the process through the last term in Equation 1a where C_{Elec} is \$0.0665 per kilowatt-hour, the average industrial electricity price for the United States in May 2015 [337].

$$\mathrm{MESP} = \frac{\mathrm{CC} + \mathrm{OC} + \mathrm{LC} - \mathrm{G_{elec}} \cdot \mathrm{C_{Elec}}}{\text{Ethanol Produced}} \tag{6.1a}$$

$$\mathrm{LC} = \begin{pmatrix} \text{Labor Cost} \\ \text{Per Gallon Ethanol} \end{pmatrix} \begin{pmatrix} \text{Ethanol} \\ \text{Produced} \end{pmatrix} \tag{6.1b}$$

The capital cost (CC) in Equation 6.1a is calculated as the annualized cost contribution of all unit operations for the process. Contributions to the capital cost come from the reactor costs (RC), separator costs (SC) and heat exchanger costs (HEC) and are annualized using Equation 6.2 to adjust for an interest rate q applied over a certain payback period of years n [338]. For these processes a 5% interest rate is used, a slight overestimate of industry standards reported by the United States Federal Reserve System [339].

$$CC - \frac{q \cdot (RC + SC + HEC)}{1 - (1 + q)^{-n}} = 0 \tag{6.2}$$

Table 6.3. Base size (flowrate) and base cost to calculate the cost (Equation 6.3) of each piece of process equipment.

Process Unit	Material of Construction	Base Size (kg/hr)	Base Cost (USD)	Ref.
Pretreatment	Incoloy Clad Carbon Steel	83,333.00	9,972,958	[303]
Hydrolysis	304 Stainless Steel	52,722.00	1,000,847	[303]
Fermentation	304 Stainless Steel	133,689.84	1,274,490	[303]
Flash Separation	316 Stainless Steel	264,116.00	1,103,895	[303]
Distillation	316 Stainless Steel	30,379.00	8,040,557	[303]
Molecular Sieve	Stainless Steel	22,687.00	2,540,057	[303]
Heat Exchanger	Stainless Steel	10,000.00	81,402	[211]

To calculate the installed cost of each unit operation (UOC), i.e. heat exchangers, reactors, distillation columns, etc., a six-tenths scaling factor is used [340-342]. Shown in Equation 6.3, this model adjusts the cost of each unit operation based on the costs of unit operations of a known size. Base size and base cost parameters are taken from

Humbird et al. or calculated using Guthrie correlations [211, 303]. In this work, all unit sizes correspond to the hourly capacity (or flow rate, measured in kg per hour) processed by the unit. The corresponding base size and cost for each piece of equipment is shown in Table 6.3.

$$UOC - Base\ Cost \cdot \left(\frac{Unit\ Size}{Base\ Size}\right)^{0.6} = 0 \tag{6.3}$$

Operating costs (OC) of the plant include the cost of biomass (BC), pretreatment chemicals (PC), enzymes (EC), process water (PWC), biomaterial costs which include enzymes and fermentation organisms (BMC), cooling water and steam utilities (UC), and waste treatment (WC) needed for the daily operation of the plant, and is calculated as shown in Equation 6.4. All operating costs are calculated based on the yearly usage of the material in question. Unit costs for each of the non-biomass raw materials used in the calculation of the operating cost, e.g. process water, pretreatment chemicals, utilities, etc., are shown in Table 6.4.

$$\mathrm{OC} = \mathrm{BC} + \mathrm{PC} + \mathrm{EC} + \mathrm{PWC} + \mathrm{BMC} + \mathrm{UC} + \mathrm{WC} \tag{6.4}$$

Table 6.4. Cost of each non-biomass raw material used in the consolidated bioprocessing and separate hydrolysis and fermentation processes.

Raw Material	Cost Value	Units	Reference
Sodium Hydroxide	430.000	$/1000 kg	[343]
Dilute Sulfuric Acid	55.110	$/1000 kg	[96]
Ammonia	500.000	$/1000 kg	[96]
Water			
Process	0.001	$/1000 kg	[96]
Cooling	0.065	$/1000 kg	[96]
Steam	6.610	$/1000 kg	[96]
Waste Treatment	36.000	$/1000 kg	[96]
Electricity Price	66.500	$/1000 kWh	[337]

6.2.4.2. Biomass and Pretreatment Selection

As stated in Section 4.2.1, both the CBP and SHF process superstructures allow for the use of multiple biomass feeds. In the models developed in this section, index *b* will be used for the available biomass types, index *i* will be used for the individual components of each stream (with a complete list shown in Table 6.1, and index *j* will be used for the available pretreatment options. Equation 6.5 is used to maintain the limit on the total biomass introduced to the system while allowing for multiple biomass types to be selected. For this work, a maximum of 3,500 U.S. tons of total biomass feed, regardless of type, is allowed into the biorefinery. Each biomass type selected is then broken down into the flow rates of its individual components using Equation 6.6, with $x_{i,b}$ denoting the mass fraction of each component (e.g. glucan, xylan, lignin, etc.) in the specific biomass option shown in Table 6.1.

$$m_{bio} - \sum_{b} m_b = 0 \tag{6.5}$$

$$m_i \ - m_b \cdot x_{i,b} = 0 \tag{6.6}$$

For each biomass type fed into the system only a single pretreatment technology can be selected. This is controlled through the introduction of binary decision variables $y_{b,j}$, where the subscript b denotes the five biomass options available and j denotes a specific pretreatment option. Equation 6.7 shows how the selection of each technology is limited to only one option per biomass type. There exists one integer constraint like that in Equation 6.7 for each biomass type available.

$$1 - \sum_{j} y_{b,j} = 0 \tag{6.7}$$

The amount of pretreatment chemical added to each branch is dependent on the biomass flow rate, m_b, in that branch system. Based on the literature sources outlined in Table 6.2 for each pretreatment-biomass combination, the amount of pretreatment chemical (m_{pre}) is added in proportion to the biomass loading, as shown in Equation 6.8. The parameter $\xi_{b,j}$ denotes the necessary amount of pretreatment chemicals added to the system for pretreatment type j. Water is added as a diluent based on the necessary pretreatment concentration needed in the pretreatment feed, the proportion of which is denoted as $\beta_{b,j}$. For cases where liquid hot water is a pretreatment option, which requires no additional pretreatment chemicals, an extra term is added to account for the water addition as exemplified in Equation 6.9a for sugarcane bagasse. Where liquid hot water is not an option, this extra term is omitted, as shown in Equation 6.9b for the sorghum feedstock pretreatment. There exists one constraint like that in Equation 6.8 and either Equation 6.9a or 6.9b for each biomass type available.

$$m_{pre} - m_b \cdot \sum_{j} y_{b,j} \cdot \xi_{b,j} = 0 \tag{6.8}$$

$$m_{water} - m_{pre} \cdot \sum_{j} y_{b,j} \cdot \beta_{b,j} - y_{b,LHW} \cdot m_b \cdot \beta_{b,LHW} = 0 \tag{6.9a}$$

$$m_{water} - m_{pre} \cdot \sum_{j} y_{b,\mathrm{j}} \cdot \beta_{b,j} = 0 \tag{6.9b}$$

6.2.4.3. Hydrolysis and Fermentation

Reactor mass balance models were developed for the pretreatment, hydrolysis and fermentation reactors. Pretreatment reactors involve the breakdown of xylan based on the pretreatment type selected and biomass type used. Hydrolysis reactors utilize cellulase enzymes to break down the glucan made accessible via pretreatment; its conversion is therefore also dependent on pretreatment type. These values are shown in Table 6.2 for each biomass-pretreatment combination. Yields for each reaction were calculated based on the stoichiometry of each reaction, as shown below [303]. The polymer chain length for the glucan and xylan are given by *n* and *m* and were taken to be 10,000 units and 1,000 units, respectively. While it is understood that these chain length values vary and will influence the amount of product ethanol, the numbers used in this study reflect an average length of these polymer chains [344].

$$1.0\ (\text{Glucan})_{\text{n}} + \text{n}\ H_2O \rightarrow \text{n Glucose}$$

$$1.0\ (\text{Xylan})_{\text{m}} + \text{m}\ H_2O \rightarrow \text{m Xylose}$$

All reactor mass balance models are component balances that use a constant conversion $(\chi_{b,i})$ assumption. For consolidated bioprocessing a conversion of 90% and an ethanol yield of 0.41 grams of ethanol per gram of glucan are used for the fermentation step [336]. For the xylose fermentation, a conversion of 85% and an ethanol yield of 0.38 grams of ethanol per gram of xylose are used [345]. For the SHF process, fermentation conversions of 90% and 80% are used for glucose and xylose, respectively. Yields for separate hydrolysis and fermentation were based on the stoichiometry of ethanol fermentation shown below [303]:

$$
\begin{aligned}
&1.0\ \text{Glucose} \rightarrow 2.0\ \text{Ethanol} + 2.0\ CO_2 \\
&1.0\ \text{Glucose} \rightarrow 6.0\ \text{Z. mobilis} + 2.4\ H_2O \\
&3.0\ \text{Xylose} \rightarrow 5.0\ \text{Ethanol} + 5.0\ CO_2 \\
&1.0\ \text{Xylose} \rightarrow 5.0\ \text{Z. mobilis} + 2.0\ H_2O
\end{aligned}
$$

An example of a reactor mass balance model for the pretreatment and hydrolysis is shown in Equation 6.10, specifically for the pretreatment of sugarcane bagasse. The conversion during the pretreatment (glucan to glucose) or hydrolysis step for that specific pretreatment pathway is given as $\chi_{SB,i}$. Values of this $\chi_{b,j}$ parameter for the conversion of xylan to xylose during pretreatment and glucan to glucose during hydrolysis can be found in Table 6.2. The value of $\chi_{b,i}$ for all other components i is zero. The mass flow rates of the inlet and outlet flow for each component in the reactor is given as m_i^{in} and m_i^{out}, respectively, while the molecular weight and reaction stoichiometry of each component is given as MW_i and a_i.

$$m_i^{out} - m_i^{in} - m_{xylan}^{in} \cdot \frac{MW_i}{MW_{xylan}} \cdot a_i \cdot \sum_j \left(y_{b,j} \cdot \chi_{b,i}\right) = 0 \tag{6.10}$$

6.2.4.4. Separation Processes

Identical separation models are developed for the two processes to characterize the flash tank, distillation column and molecular sieve, respectively. A flash tank is used to model the capability of the fermenter to remove the gaseous components, such as carbon dioxide produced from the anaerobic respiration during fermentation, from the final ethanol product. Component balance models are developed to separate gasses from the liquid and solid components, which continue to the distillation process.

The feed to the distillation process is assumed to be a binary mixture of ethanol and water and is modeled as such. The distillate of the column produces the product

ethanol to be sent for further refining while the bottoms stream consists of a water-rich solution that contains the unreacted solids from the pretreatment, hydrolysis, and fermentation. The separation of ethanol and water via distillation is modeled using a fractional recovery method shown in Equation 6.11 [346]. The fractional recovery, f_l, is determined by the desired compositions of the ethanol in the distillate and bottoms streams, z_d and z_b, as well as the ethanol concentration in the distillation feed, z_{eth}^f. A 93 wt% stream of ethanol is desired as the distillate of the column and we look to restrict the bottoms of the column to less than 1 wt% ethanol. The fractional recovery can then be used to determine the amount of ethanol and water found in the distillate stream [346].

$$f_l - \frac{z_d \cdot \left(z_{eth}^f - z_b\right)}{z_{eth}^f \cdot (z_d - z_b)} = 0 \tag{6.11}$$

The final step toward 99.5% pure ethanol product involves separation via molecular sieve. The molecular sieve is modeled using individual component balances with a split fraction approach as shown in Equation 6.12 and Equation 6.13. Equation 6.14 shows the overall component mass balance of the sieve, with the feed flow rate m_i^{in} split into a pure product stream with flow rate m_i^p and a waste stream with flow rate m_i^w. Here, the split fraction is denoted as ϕ.

$$m_{eth}^p - \phi \cdot m_{eth}^{in} = 0 \tag{6.12}$$

$$m_{water}^p - (1 - \phi) \cdot m_{water}^{in} = 0 \tag{6.13}$$

$$m_i^{in} - m_i^w - m_i^p = 0 \tag{6.14}$$

6.2.4.5. Energy Cogeneration

Energy considerations for this plant design include the energy requirement needed to heat and cool process streams and reactors. Heat exchangers are found before the pretreatment reactors for each biomass feed branch to adjust the temperature of the stream to that of the reactor feed; this allows effective pretreatment of the biomass feed. Additional heat exchangers include the reboiler and the condenser of the distillation column. All heating is done using low pressure steam utility and all cooling is done via cooling water. Equation 6.15 is used to determine the amount of energy transfer in each heat exchanger (Q_{ex}), with Equation 6.16 being used to calculate the heat capacity for the multicomponent streams. The constants A_i, B_i, C_i, and D_i in Equation 6.16 are found in literature [347]. For lignin and other biological solids, C_i and D_i are assumed to be zero while the constants A_i and B_i were fit to data provided by Voitkevich et al. [348]. Additional thermodynamic model parameters were taken from Sandler [349].

$$Q_{ex} - m_b \cdot \sum_i \left(x_i C_{P,i}\right) \cdot \Delta T_j = 0 \tag{6.15}$$

$$C_{p,i} - (A_i + B_i T + C_i T^2 + D_i T^3) = 0 \tag{6.16}$$

Energy requirements for each reactor must also be calculated. Each reactor is designed to remain isothermal to allow for the optimal performance of all biological reactions. As such, the heat generated or consumed by the reactions must be removed or replaced. Equation 6.17 uses the heats of formation $\left(\Delta H_i^f\right)$ to calculate the heats of reaction (ΔH_{rxn}) for each reaction, and Equation 6.18 uses this calculation to determine the heating or cooling requirement of each reactor ($Q_{reactor}$) based on the change in

flow rate of the limiting reactant (Δm_{rxn}). The heats of formation for glucan, xylan, glucose, xylose and biomass are determined from literature [256].

$$\Delta H_{rxn} - \sum_i \nu_i \Delta H_i^f = 0 \tag{6.17}$$

$$Q_{reactor} - \sum_{rxn} (\Delta m_{rxn} \cdot \Delta H_{rxn}) = 0 \tag{6.18}$$

From this energy requirement, the required amount of utility, either low pressure steam or cooling water, can be ascertained. Equation 6.19 shows the determination of the steam flow rate needed for a typical heat exchanger using the heat of vaporization of water ΔH_{water}^{vap}, while Equation 6.20 shows the cooling water flow rate required for a typical reactor using the sensible heat of water.

$$m_{steam} - \frac{Q_b}{\Delta H_{water}^{vap}} = 0 \tag{6.19}$$

$$m_{cw} - \frac{Q_{reactor}}{C_p^{water} \cdot \Delta T_{water}} = 0 \tag{6.20}$$

Steam and electricity generation using the lignin content of the distillate bottoms stream are considered to offset the operating costs of the processes. First, the bottoms effluent is sent through a centrifuge that separates the solid phase into a pellet containing 5% water. The aqueous phase continues to a boiler, where a portion of the solid pellet is used to generate low pressure steam that can be used to supplement or replace the energy needs of the heat exchangers in the process. The boiler operates at an efficiency of 80%. The remaining solids are sent to a separate steam cycle for electricity generation. Equation 6.21 is used to calculate the electricity generation of the steam cycle, where

η_{SC} denotes the steam cycle efficiency, F_{Lignin} denotes the flow of lignin to the steam cycle, and ΔH^{c}_{Lignin} denotes the heat of combustion of lignin given as 22,870 kJ per kilogram [256]. It is assumed that the steam cycle operates at 33% efficiency [350]. The cost of this steam cycle ($Cost_{SC}$) is calculated based on the amount of electricity generated, as shown in Equation 6.22. A base cost of $1000 per kilowatt of electricity generated $\left(Cost^{base}_{SC}\right)$ is used for the steam cycle [351].

$$\binom{Electricity}{Generation} - \eta_{SC} \cdot F_{Lignin} \cdot \Delta H^{c}_{Lignin} = 0 \tag{6.21}$$

$$Cost_{SC} - Cost^{base}_{SC} \cdot \binom{Electricity}{Generation} = 0 \tag{6.22}$$

6.3. Results and Discussion

6.3.1. Economic Viability of Consolidated Bioprocessing

Table 6.5 shows the minimum ethanol price for the CBP facility with energy cogeneration. This plant produces ethanol at a breakeven ethanol selling price of $1.31 per gallon ($0.35 per liter). The solver BARON, developed for the solution of non-convex, mixed-integer nonlinear programs via a branch-and-reduce framework, was used to solve the optimization problem described by the models in Section 2.4 to a 5% optimality gap [65]. The optimal topology foregoes the selection of multiple biomass feed stocks, instead selecting only the least expensive feedstock, sugarcane bagasse. This optimal topology is depicted in Figure 6.3, where a single biomass feed train of sugarcane bagasse is selected and the energy cogeneration is used only for the initial pretreatment heater. A pretreatment method of sodium hydroxide was chosen for this sugarcane bagasse feed; this pretreatment leads to an increased ethanol production and a

decrease in energy requirements due to its higher conversions and lower operating temperature when compared to the alternately available liquid hot water pretreatment, as shown in Table 6.2.

Table 6.5. Results for the minimum ethanol selling price, optimal topology, and optimal ethanol and electricity production rates for the separate hydrolysis and fermentation (SHF) and consolidated bioprocessing (CBP) facilities.

	CBP	SHF
Ethanol price ($/gal)	$1.31	$1.64
Ethanol production (MM gal/yr)	66.0	76.2
Total biomass feed (U.S. ton/day)	3,500	3,500
Feed composition	Sugarcane bagasse	Sugarcane bagasse
Pretreatment selection	NaOH	NaOH
Electricity generation (MW·hr)	67	23

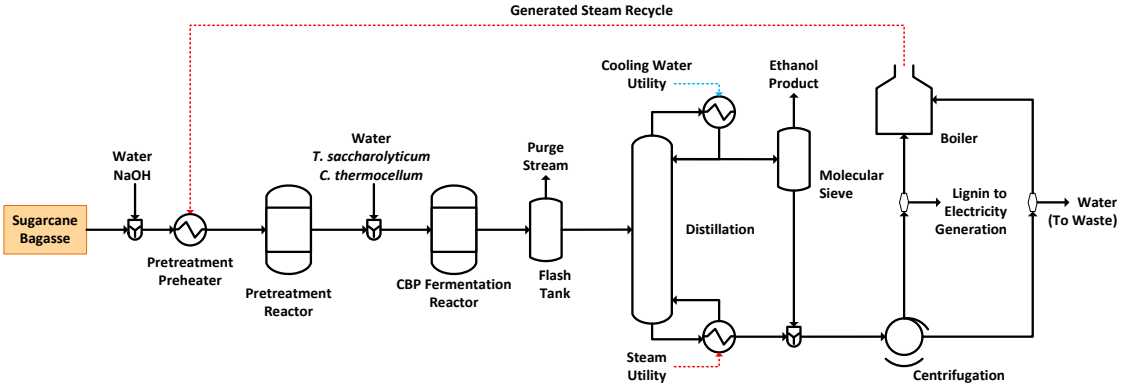

Figure 6.3. Optimal topology of the consolidated bioprocessing pathway for biological conversion and subsequent separation of lignocellulosic biomass to ethanol.

6.3.2. Comparison of the Optimal CBP and SHF Processes

In addition to those of the optimal CBP process, the optimal SHF process topology with respect to the selected biomass option(s), pretreatment(s), production rates of ethanol and electricity, and the resulting minimum ethanol selling price are shown in Table 6.5. Like the CBP process, the selection of sugarcane bagasse at the maximum allowable feed rate with a sodium hydroxide pretreatment is also featured in the optimal

SHF solution. This result follows that of Michailos et al., who demonstrated that for a sugarcane bagasse feed the optimal pretreatment was independent of fermentation strategy [352]. However, the difference seen in the ethanol production and minimum ethanol price exemplifies the impact of the removal of the enzyme hydrolysis step. While the removal of a hydrolysis step does lead to a 13.3% decrease in ethanol production for the CBP process, the removal of operating costs associated with enzyme addition results in a 20.2% decrease in the breakeven minimum ethanol selling price.

Table 6.6. Breakdown of the costs (in $MM per year) associated with the optimal consolidated bioprocessing (CBP) and separate hydrolysis and fermentation (SHF) facilities with energy cogeneration.

	CBP	SHF
Annualized capital costs	23.66	13.44
Fermentation cost	2.97	0.66
Enzyme cost	0.00	25.91
Total biomass cost	44.45	44.45
Total pretreatment cost	1.98	1.98
Process water cost	6.67	9.82
Waste water treatment cost	12.84	11.29
Utilities cost	2.48	0.07
Labor cost	34.47	39.77
Total annual cost	123.89	137.58

Table 6.6 shows a breakdown of the major cost contributors for both the CBP and SHF processes and gives more insight into the differences between the optimal CBP and SHF processes. Many cost contributions vary slightly between the two technologies, but consolidated bioprocessing results in a cheaper process annually. This is a direct result of the removal of the enzyme cost needed for enzyme hydrolysis, an operation accounting for $25.91 million dollars in operating cost for the SHF process. This 9.95% reduction in yearly cost, coupled with a 197% increase in electricity generation, is large

enough to overcome the decrease in ethanol production and leads to the decrease in the minimum ethanol selling price.

While the knowledge of the commercial scale yields and consumptions of the SHF process have a very high degree of reliability based on its industrial scale performance, the analysis for CBP performed here is based on lab-scale experiments and assumptions that the scaled processes would perform identically to the experimental scale at the commercial level. Bioprocessing in general is subject to a high level of cellular variability that often results in uncertainty regarding bioreactor conversion; the scale up of these systems see a further increase in uncertainty due to an increase in mass transfer limitations with respect necessary nutrients and substrates. This uncertainty in bioreactor conversion could be detrimental to the production of ethanol, which could result in a higher selling price. To mitigate these concerns, traditional scale up techniques using pilot plant analysis need to be used to determine the extent of the uncertainty and level of deviation from lab-scale results.

In addition to uncertainty regarding the scale-up of CBP, this work utilizes many constant values that are a subject of uncertainty due to seasonal variability. For example, the sales price of electricity is variable with respect to time of the year and location in the United States. Biomass prices are also variable, reaching a low point when in peak harvesting season and reaching a maximum during the non-growth seasons. The aim of this work is to compare the two processes at the commercial scale in a way that determines if consolidated bioprocessing is a technology that can potentially provide economic improvement when compared to the well-known separate hydrolysis and

fermentation. As such, uncertainties in the variables listed above are not addressed here but would be necessary considerations for the optimal topology determined here. Optimization methods that consider the uncertainty of various model parameters have been developed and used for a variety of other systems [353-355]. However, this aim of this work is to compare the two processes at the commercial scale in a way that determines if consolidated bioprocessing is a technology worth further investigation due to providing economic improvement when compared to the well-known separate hydrolysis and fermentation. The results demonstrated here show the use of consolidated bioprocessing with energy cogeneration is an economically favorable option for liquid transportation fuel production that should be investigated in more detail.

6.3.3. Long Term Economics of Consolidated Bioprocessing

As the leading strategy for the biological conversion of lignocellulosic biomass to ethanol, the long term economic implications of the consolidated bioprocessing plant are studied further. While analysis to this point had determined the breakeven selling price for the ethanol, this analysis aims to determine the price markup for the sale of ethanol that would ensure a long-term profitable process, indicated by achieving a discounted cash flow return on investment (DCFROI) of greater than 15% [211]. This study assumes a process lifetime of 20 years at a tax rate of 9% based on data collected across the entire chemical industry [356]. Additional assumptions include an engineering and contingency cost of 20% of the inside boundary limit (ISBL) costs each, a startup cost of 10% of the total capital investment (TCI), and straight-line depreciation for the lifetime of the plant. The plant is constructed in 2 years, the initial capital investment of

the plant is paid over a course of the following 10 years, and the sale price of the electricity was not adjusted in this analysis. The payout time of the project is also examined.

The long term economic analysis shows that to achieve the economic indicators previously described that the ethanol price must be marked up by about 78%, leading to an ethanol selling price of $2.32 per gallon ($0.61 per liter). The cumulative cash flow for a discount rate of 10% is shown in Figure 6.4, resulting in a net present value of $135.4 million. This would result in a payout time of 5.80 years. The ethanol selling price of $2.32 per gallon ($0.61 per liter) is comparable to today's selling prices for petroleum products used for transportation fuel and does not involve the need of any subsidy that would only serve to drive the price for profitability down.

An additional sensitivity analysis on the DCFROI was performed for the instance where the ISBL costs, raw material costs, and product prices determined here are inaccurate. These parameters were chosen as then have the most associated uncertainty in their estimation. While the input-output models used in this work capture the general flows of mass, energy, and cost in the facility, they are simple balances that do not include a high degree of detail. For example, the use of the six-tenths rule (Equation 6.3) for cost estimation gives a general idea of the cost of equipment but will not be as accurate as a direct vendor quote. Increasing the level of detail in the design will increase the accuracy of the capital expenditures and as such it is important to look at how the long-term economics will be affected by more detailed design process. For this

work, it is assumed that the economic estimations for the capital cost are within 50% of the more detailed design.

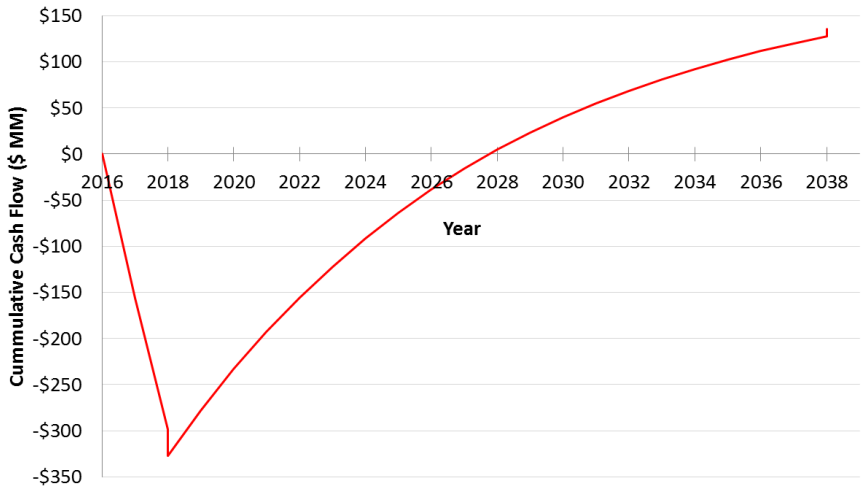

Figure 6.4. Cumulative cash flow for the consolidated bioprocessing plant with cogeneration, ethanol sold at $2.32 to ensure at least a 15% DCFROI.

A similar statement can be made about both the prices of the raw material and the selling price of the ethanol and electricity products. The prices used here are static prices while these are variable prices dictated by their respective markets that tend to deviate both across the United States and throughout the year. Examining the resilience of the designs discussed in this work to these deviations is vital to making any conclusions about the economic feasibility of the CBP biorefinery. To ensure that most of the variability in the cost of these materials is utilized, a deviation of $\pm 10\%$ is used in the sensitivity analysis for the raw material and product prices.

The results of this sensitivity analysis are shown in Table 6.7. The consolidated bioprocessing plant with cogeneration was found to be economically favorable for up to 50% disturbances in the total capital investment (TCI) and 10% disturbances in the

product prices or raw material costs as all values of the DCFROI were close or greater than 10%. It is interesting to note that an underestimate of 50% in the TCI results in an almost 6% drop in the DCFROI, whereas an overestimate of 50% to the TCI results in a 14% increase in the DCFROI. Additionally, perturbations in the price of raw materials shows minor changes in the economic potential of the process, as a 10% increase or decrease in this input only exhibits a positive of negative 1.5% change in the DCFROI. The ability of the process to handle disturbances in raw material costs is imperative to its long term financial success, as the price of sugarcane bagasse is subject to variation depending on that growth season.

Table 6.7. Sensitivity analysis on the discounted cash flow return on investment (DCFROI) economic indicator for the consolidated bioprocessing plant with energy cogeneration.

Deviated Economic Factor	Variation	DCFROI	Change in DCFROI
Base case	-	15.17%	
Change in total capital investment	+50%	9.38%	-5.79%
	-50%	29.51%	14.34%
Change in product prices	+10%	18.96%	3.79%
	-10%	11.02%	-4.15%
Change in raw material prices	+10%	13.62%	-1.55%
	-10%	16.67%	1.50%

6.3.4. Implications of Using Multiple Feeds with Consolidated Bioprocessing

While the proposed CBP superstructure allows for the use of multiple biomass options, only the sugarcane bagasse feedstock with sodium hydroxide pretreatment was selected in the optimal topology. This can be attributed the necessity of reactors dedicated to pretreatment if any additional biomass types are selected, leading to an increase in capital and operating expenses. As discussed in Section 1, the biomass selection can elucidate the geographical location for each plant to be built. As Figure 6.1

shows, sugarcane bagasse is grown in the southeastern part of the United States, primarily in the states of Virginia, the Carolinas, Florida, and Louisiana. The optimal CBP process calls for approximately 1.2 million U.S. tons of biomass to be processed per year; this amount of biomass must be harvested from the surrounding area and available for processing for these designs to be feasible.

However, the use of multiple biomass types can be used to effectively reduce the necessary radius used for harvesting the process feed. Figure 6.5 shows the effect of using a mixture of sugarcane bagasse and corn stover on the minimum ethanol selling price (MESP) and the radius of influence of the processes. Harvesting radius was calculated for each biomass type using the method described by Overend while assuming a biomass productivity of 10 oven dry metric tons per hectare per year ($\mathrm{ODt} \cdot \mathrm{ha}^{-1} \cdot \mathrm{yr}^{-1}$) and a complete circular geometry with 10% devoted to the growth of each crop [357]. As it is assumed that each biomass can be found around the process the largest harvest radius for each type of biomass is used as the overall harvesting radius.

As the percentage of bagasse in the feed is reduced, while maintaining a similar total biomass flow rate, the MESP is increased to account for the additional capital and operating cost necessary. However, a minimum in the harvest radius can be seen around an even split in the biomass feed. This minimum of approximately 20 km is due to the equal use of bagasse and corn stover, leading to a similar harvesting radius. The tradeoff between the economic and environmental implications of bioethanol production via biochemical means can be seen clearly here, as a pure feed of sugarcane bagasse allows for the minimum ethanol selling price but utilizes the maximum amount of land while a

mixture of biomass types necessitates an increased price of bioethanol but reduces the strain on the surrounding agricultural community. It can also be seen that with an increase in corn stover usage comes a decrease in the electricity that can be generated by the process, a direct result of the significant difference in lignin composition between the two feed stocks, as can be seen in Table 6.1 in Section 6.2.1. When coupled with the decrease in ethanol production, this decrease in electricity production leads to an increase in the minimum ethanol selling price, as seen in Figure 6.5.

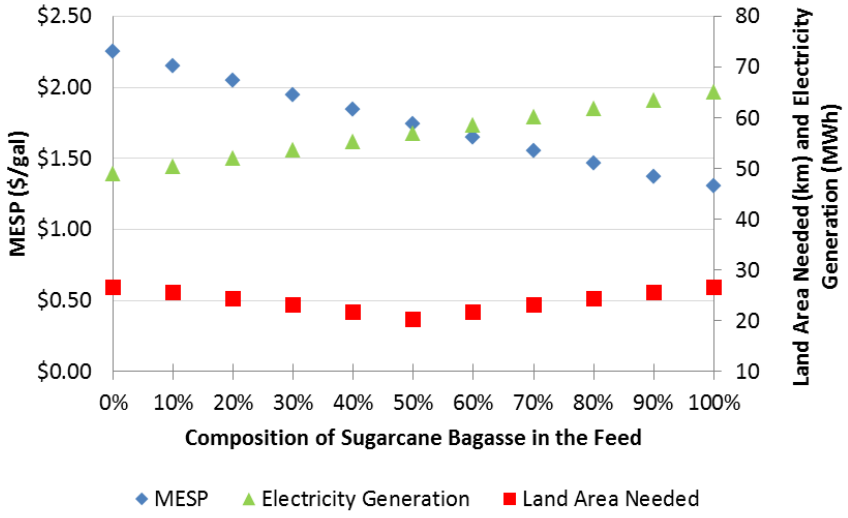

Figure 6.5. Minimum ethanol selling price (MESP), required land area (expressed as a radius of a circle) and electricity generation as a function of sugarcane bagasse fed to the CBP process.

6.4. Conclusions

The economic and environmental impact of two different methods to produce ethanol from lignocellulosic biomass using biochemical conversion methods has been analyzed. It has been shown that a consolidated bioprocessing plant utilizing a sugarcane bagasse feedstock, sodium hydroxide pretreatment and energy cogeneration can be used to provide ethanol at a breakeven selling price of $1.31 per gallon. Additionally, the long

term economic outlook for this plant has been shown to be attractive for an ethanol selling price of $2.32 per gallon, a markup of 78%. Finally, the use of sugarcane bagasse as a feedstock suggests the construction of these facilities in the southeastern United States, where sugarcane is plentiful, while of a split feed of sugarcane and bagasse and corn stover can be used to lower the environmental impact of the plant by reducing the harvest area required for its feedstock supply to be met.

CHAPTER VII

PROCESS INTENSIFICATION OF CONTINUOUS BIOBUTANOL PRODUCTION VIA A MULTI-FEED BIOREACTOR WITH *IN SITU* GAS STRIPPING

7.1. Introduction

New methods to produce alternative energy sources are of great interest to increase sustainability, combat global warming, and achieve energy independence. Many researchers are interested in the harvesting energy from wind and solar resources; however, while these energy sources are critical for securing a renewable energy future, their direct application is limited in the transportation industry without improvements in battery technology and the distribution infrastructure [358-361]. The use of these technologies can be used indirectly to produce hydrogen as a fuel source, but questions remain about the ability to transport and store this volatile product for wide distribution [362, 363]. Biofuels offer an alternative to petroleum products for energizing the transportation industry using a liquid fuel that can utilize the current distribution network. While distribution may not present a challenge, methods are still needed for producing these alternative fuels in quantities necessary to meet the demand of current fossil fuel sources.

One of the promising alternative liquid fuels is ethanol. Though currently made industrial using a corn feedstock, many researchers have investigated using novel technologies to process the sugars found in lignocellulosic biomass to avoid using food products as a supply for energy needs. Many process have been designed on the premise of first pretreating the biomass to delignify the biomass and allow access to the sugars

during a subsequent enzyme hydrolysis step [96, 301, 303, 305-307, 317, 364]. This sugar mixture, containing both six-carbon glucose and five-carbon xylose, can then be fermented using a variety of different organisms [365-368]. As an alternative, consolidated bioprocessing has been developed to remove the costly enzyme hydrolysis step by using organisms capable of the direct fermentation of the sugar polymer cellulose [168, 334, 369, 370]. Despite the inroads in process development, ethanol suffers from a few disadvantages when compared to traditional petroleum-derived fuels, most notably a higher volatility, a higher corrosiveness, and a high miscibility with water, making pipeline transportation challenging, and a lower energy density than gasoline or diesel fuel [371, 372].

Another alcohol product, butanol, provides a promising alternative to bioethanol as a renewable liquid fuel. With a higher energy density that rivals that of gasoline, a lower volatility, and negligible miscibility with water, butanol has many favorable qualities over its ethanol counterpart [373]. Traditionally, biobutanol is produced simultaneous with ethanol and acetone through the fermentation of glucose by the bacteria *Clostridium acetobutylicum*, a process called acetone-butanol-ethanol (ABE) fermentation. However, the production of butanol during the ABE fermentation process increases the toxicity of the broth to the cells [374, 375]. Overcoming the butanol toxicity limitations, which in turn limits the productivity and potential profitability, is necessary for ABE to be produced at levels that meet current gasoline demands.

Many methods focusing on coupling reaction and separation processes, a field known as process intensification, have been developed to selectively remove the desired

ABE products and combat the toxic nature of butanol. Selective removal of the ABE species using adsorption in an *ex situ* column has been investigated using cyclic filling, emptying, and regeneration to continuously remove the ABE products from a batch fermentation process [376-379]. Pervaporation, the process of selectively removing volatile compounds such as ethanol, butanol, and acetone via partial vaporization and diffusion through a nonporous, selective membrane, has also been studied for improving ABE fermentation productivity [380-383]. Additionally, both *in situ* and *ex situ* gas stripping using inert gases such as nitrogen, hydrogen, or carbon dioxide has been considered to actively remove the volatile alcohol components without affecting the growth of the *Clostridium* cells [377, 384-387]. Many of these technologies have been investigated experimentally and computationally and have been shown to be effective at the small, batch scale for increasing reactor productivity, but little work has been done on investigating its efficacy and economic feasibility in a large, continuous setting.

In this work, the economics of biobutanol production are examined through the optimal control and subsequent process synthesis of an intensified ABE bioreactor. First, the optimal control of a multi-feed, continuous ABE fermenter with *in situ* gas stripping is analyzed to maximize the reactor profitability using the dilution rate, feed glucose composition, and stripping gas flowrate as manipulated variables. The optimally controlled, intensified bioreactor is compared to an optimally controlled, multi-feed bioreactor in the absence of gas stripping to demonstrate the improved performance. After determining the optimal control policy for a profitable reactor, the long-term economic prospect of large-scale, continuous biobutanol production from lignocellulosic

biomass is considered. By considering a process superstructure of potential biomass feedstocks and pretreatment methods, process synthesis techniques are used to determine the necessary butanol selling price to guarantee a 15% discounted cash flow return on investment (DCFROI). The sensitivity of the selling price to process scale is investigated and compared to the current butanol market, and potential improvements to the reactor and overall process to increase economic viability are discussed.

7.2. Materials and Methods

7.2.1. Butanol Production from *Clostridium acetobutylicum*

The production of biobutanol via fermentation is a complex set of metabolic reactions typically done using a strain of *Clostridium* bacteria [388]. In the first part of the fermentative pathway, the substrate is converted to acetyl-CoA, a key intermediate that can be utilized to produce biomass or the production of the desired solvent products. While glucose is typically used as the substrate, work has been done to develop recombinant strains of Clostridium that can simultaneously ferment xylose to allow the efficient use of lignocellulosic biomass feedstocks [389]. The subsequent reactions necessary to produce ABE solvents are completed in two steps. The first step involves the conversion of acetyl-CoA through three separate and simultaneous pathways to generate butyric acid and acetic acid in a process called acidogenesis. By changing the pH of the fermentation both, these intermediates initiate the second step called solventogenesis, during which the acids are converted to the final solvent products butanol, ethanol, and acetone. A generalized diagram for the metabolic pathways for acidogenesis and solventogenesis are shown in Figure 7.1.

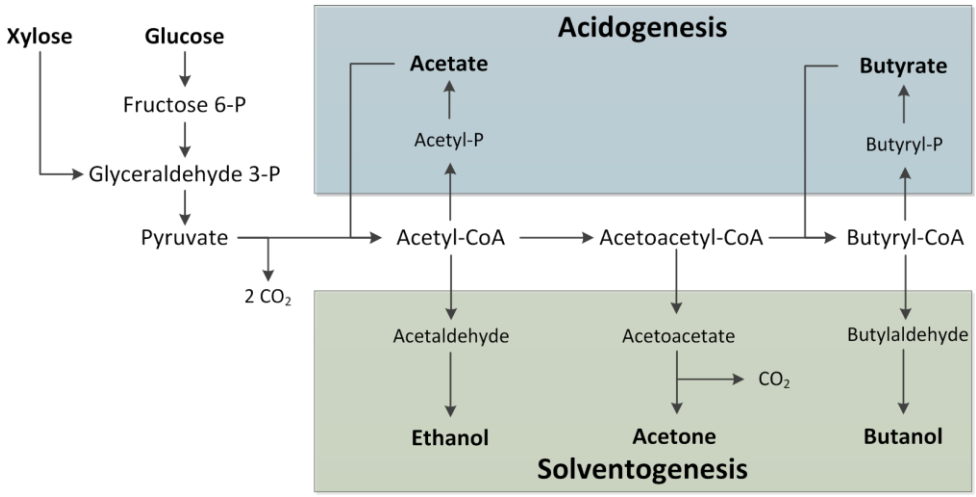

Figure 7.1. Generalized metabolic pathways of *Clostridium acetobutylicum* to produce acetone, butanol, and ethanol.

Many researchers have focused on developing mathematical equations necessary to design chemical processes that harness the complexity of the ABE fermentation pathways [390-394]. One model of interest was proposed by Votruba et al. and assumes the media is glucose limited and there is no limitation on the nitrogen source [395]. By using a state variable y to define the ratio of the RNA concentration and the critical RNA concentration needed for growth, this model uses the RNA concentration as a physiological state marker to account for all phases of batch culture growth and production. The equation for the normalized RNA concentration is shown below in Equation 7.1, with the 0.56 constant defining the relationship between the growth rate μ and the RNA concentration as defined by Harder and Roels [396]. Equation 7.2 and Equation 7.3 below describe the time-dependence of the biomass growth (X) and substrate utilization (S). In Equation 7.2, the first term is the growth equation based on the RNA ratio y, where cell growth occurs if $y > 1$ and cell death occurs if $0 \leq y < 1$.

The second term in that equation is used to model the inhibition effects of the butanol product.

$$\frac{dy}{dt} = k_1 X \left(\frac{K_I}{K_I + B}\right) y - 0.56(y-1)y \tag{7.1}$$

$$\frac{dX}{dt} = 0.56(y-1)X - k_2 BX \tag{7.2}$$

$$\frac{dS}{dt} = -\left[k_3 S + k_4 \left(\frac{S}{K_S + S}\right)\right] X \tag{7.3}$$

Models for the products of the ABE pathway, including the butyrate (R) and acetate (T) products of the acidogenesis pathway and acetone (A), butanol (B), and ethanol (E) of the solventogenesis pathway, are shown as Equations 7.4 through 7.8. Apart from the ethanol production equation, which is only dependent on the growth of the *C. acetobutylicum*, the equation for each metabolite is composed of two terms; the first term denotes the increase in concentration due to substrate utilization and biomass growth and the second term couples the butyrate and butanol products as well as the acetate and acetone. In Equations 7.4 and 7.5, the concentration of the acid products is increased with the conversion of glucose, subject to the inhibition effects of butanol, and decreased as these components are used to produce acetone and butanol. In Equations 7.6 and 7.7, butanol and acetone production are modeled with a first term that increases concentration when glucose is utilized and a second term that is linearly dependent on the rate of change of the corresponding acid; the stoichiometric coefficient of the acid to alcohol reactions are used as the linear constants.

$$\frac{dR}{dt} = \left[k_5 S\left(\frac{K_I}{K_I + B}\right) - k_6\left(\frac{R}{K_R + R}\right)\right] X \tag{7.4}$$

$$\frac{dT}{dt} = k_8\left(\frac{S}{K_S + S}\right)\left(\frac{K_I}{K_I + B}\right) X - k_9\left(\frac{T}{K_T + T}\right)\left(\frac{S}{K_S + S}\right) X \tag{7.5}$$

$$\frac{dB}{dt} = k_7 SX - 0.841\frac{dR}{dt} \tag{7.6}$$

$$\frac{dA}{dt} = k_{10}\left(\frac{S}{K_S + S}\right) X - 0.484\frac{dT}{dt} \tag{7.7}$$

$$\frac{dE}{dt} = k_{11}\left(\frac{S}{K_S + S}\right) S \tag{7.8}$$

7.2.2. Optimal Control of a Single, Intensified ABE Bioreactor

7.2.2.1. The continuous, intensified multi-feed bioreactor

With knowledge of the ABE fermentation kinetics, bioreactors can be designed with the goal of maximizing profitability and productivity. In this work, we propose the intensification of a continuous, multi-feed ABE fermenter with in situ gas stripping to selectively remove the solvent products. The configuration of the reactor is shown in Figure 7.2. The traditional method for continuous operation of bioreactors uses a single substrate feed that allows for the modulation of the dilution rate at a set glucose concentration. However, recent studies have shown that using a multi-feed configuration allows for the independent control of the dilution rate and glucose concentration and can improve process controllability, productivity, and profitability of continuous bioreactors [397]. In addition to the multiple feeds for substrate additional, the inclusion of gas stripping directly to the reactor allows for the selective removal of the desired ABE products. Studies have also shown that the direct inclusion of gas stripping has no

negative effects on the cell culture [398]. An optimal control framework can be developed that determines the optimal control policy for the two liquid feed flow rates and a gas flow rate that maximize reactor profitability. The development of this optimal control problem, including the development of a properly formulated objective function and continuous liquid and gas mass balances, will now be discussed.

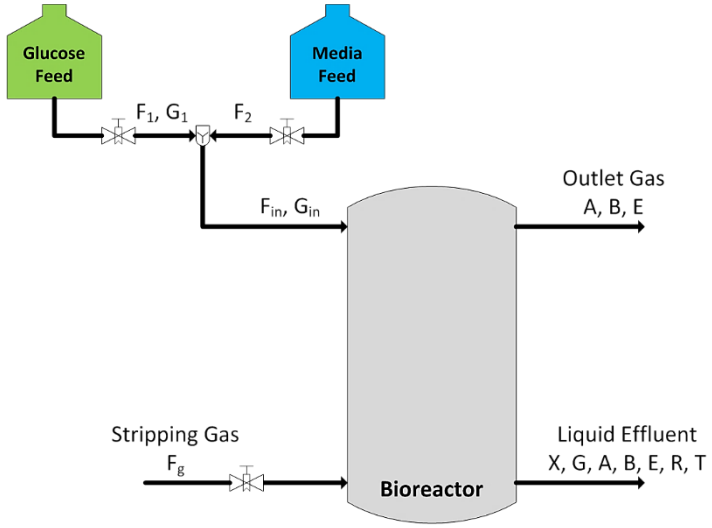

Figure 7.2. Multi-feed bioreactor with *in situ* gas stripping. The glucose flow rate F_1, media flow rate F_2, and gas flow rate F_g can be modulated to maximize the conversion of glucose (G) to biomass (X), acetone (A), ethanol (E), butanol (B), butyrate (R), and acetate (T).

7.2.2.2. Optimal control framework for the intensified reactor

The framework of the optimal control problem for the multi-feed, intensified bioreactor begins with the development of continuous mass balance models for the gas and liquid phase concentrations. Equation 7.9 and 7.10 shows the mass balance model for a generic component in the liquid (P) and gas phases (P_g). In the liquid phase, component P can be added to the system at concentration P_{in} through the liquid phase inlet and removed by the liquid effluent, respectively, at a rate equal to the total flowrate

$F_1 + F_2$ normalized by the reactor volume V_r, also known as the dilution rate. Alternatively, component P can be produced or consumed by reaction, with the reaction kinetics for each component denoted as r_p. In this work, the reaction kinetics and corresponding kinetic parameters are described by the models of Votruba et al. discussed previously [395]. In the gas phase, products are only removed from the reactor, at a rate proportional to the gas phase concentration P_g. The rate of product removal is determined by the gas dilution rate, defined by the gas flowrate F_g normalized by the liquid reactor volume V_r, and the gas hold-up ε, or volume fraction of gas in the fermenter.

$$\frac{dP}{dt} = \frac{F_1 + F_2}{V_r}(P_{in} - P) + r_p - r_e - r_c \tag{7.9}$$

$$\frac{dP_g}{dt} = -\frac{F_g}{V_R}\left(\frac{1-\varepsilon}{\varepsilon}\right)P_g + \left(\frac{1-\varepsilon}{\varepsilon}\right)r_e + \left(\frac{1-\varepsilon}{\varepsilon}\right)r_c \tag{7.10}$$

As the reaction progresses, the volatile compounds will be transferred from the liquid phase to the gas phase, either through diffusive or convective transport. These phenomena couple the gas and liquid phase concentrations and are described by the rates r_e and r_c, respectively. Definitions for the mass transfer rates are shown in Equation 7.11. The rate of diffusive transport r_e is proportional to the difference between the current concentration P_g and the equilibrium concentration P_g^* defined by Henry's law. The Henry's law constant H is defined for each species and is taken from literature [399]. The convective mass transfer rate r_c is calculated similarly, using a difference of the current concentration P_g and the dispersion concentration P_{disp} that denotes the concentration of product in the entering dispersed gas phase; this value is necessarily

zero to denote a clean gas feed. The respective mass transfer coefficient (K_g) values are determined from literature while the convective mass transfer coefficients h and the bubble size a are calculated using the correlation in Equation 7.12 [400-402]. The correlations for h and a require knowledge of the Schmidt number Sc, which describes the interaction between viscosity and mass diffusivity, and the power-to-volume ratio that describes the mass transfer properties imparted by the impeller speed. Correlations for these values, along with the gas hold-up parameter ε used in the gas mass balance, are shown as Equation 7.13 [400]. Other constants and physical properties are taken from literature [400].

$$r_e = K_g a\left(P_g^* - P_g\right); \qquad r_c = ha\left(P_g - P_{disp}\right) \tag{7.11}$$

$$h = 0.31 \cdot Sc^{-0.66}\left(\frac{\Delta\rho\mu_l g}{\rho_l}\right)^{0.33}; \qquad a = 593\left(\frac{P}{V_r}\right)^{0.25} U^{0.75} \tag{7.12}$$

$$Sc = \frac{\mu_l}{\rho_l D_l}; \qquad \varepsilon = 0.34\left(\frac{P}{V_r}\right)^{0.25} U^{0.75} \qquad \frac{P}{V_r} = 0.354\rho_l g U \tag{7.13}$$

With the understanding of how the composition of the gas and liquid phases will change as a function of time and feed flowrates, an optimal control or dynamic optimization problem can be formulated that maximizes a process objective. In this work, we aim to maximize the reactor profitability, defined as how much money the reactor makes over a certain operating time, by manipulating the two liquid feed flow rates $F_1(t)$ and $F_2(t)$ and one gas flow rate $F_g(t)$. The mathematical structure of the objective function in shown in Equation 7.14, calculated as the difference between the profit made from selling i desired products acetone, butanol, and ethanol and the expense of the glucose added to the reactor. In this equation, the mass flow rate of the

liquid, products, gas products, and glucose addition is defined as $\dot{m}_i^l(t)$, $\dot{m}_i^g(t)$, and $\dot{m}_G^{in}(t)$, respectively, the cost of each solvent product is C_i, and the cost of glucose is C_g. This objective function is maximized subject to the dynamic mass balance constraints defined by Equations 7.9 through 7.13 and constraints on the allowable values of the manipulated variables $F_1(t)$, $F_2(t)$, and $F_g(t)$. Values for the substrate and product costs are given in Table 7.1 below.

$$\begin{gathered}\max_{F_1,F_2,F_g} \phi = \int_0^{t_f} \left[\sum_{i=A,B,E} C_i \left(\dot{m}_i^l(\mathrm{t}) + \dot{m}_i^g(t) \right) - C_G \dot{m}_G^{in}(t) \right] dt \\ \text{Equations } 9-13 \\ \dot{m}_i^l = P_i\left(F_1(t) + F_2(t)\right) \\ \dot{m}_i^g = P_{g,i} F_g(\mathrm{t}) \\ s.t. \quad 0 \le F_1(t), F_2(t) \le 0.15 \cdot V_r \\ 0 \le F_g(t) \le 60 \cdot V_r \\ P = \{G, X, A, B, E, R, T\} \\ P_g = \{A, B, E\}\end{gathered} \tag{7.14}$$

Table 7.1. Raw material and product prices for the optimal control of the multi-feed, intensified bioreactor.

Price, $/kg	Value	Units	Reference
Butanol	1.49	$/g	[296]
Ethanol	1.00	$/g	[296]
Acetone	1.03	$/g	[296]
Glucose	0.88	$/kg	[297]

Obtaining solutions to the optimal control problem described in Equation 7.14 require the discretization of the differential and integral terms to form a large system of coupled nonlinear equations capable of being solved by commercially available solvers. While many methods for discretization have been proposed, the method used here is the method of collocation of finite elements. In this method, the time horizon t_f is

partitioned into segments called finite elements. Within each finite element, a set of points called collocation points are chosen that satisfy the roots of an interpolating polynomial and their derivatives. In this work, the method proposed by Flores-Tlacuahuac at al. is used, which employs a three-point Radau collocation method ($ncol = 3$) with exact values of the collocation points defined by Hairer and Wanner [286, 403]. For a more detailed explanation for the application of collocation of finite elements to a multi-feed bioreactor problem, the reader is referred to our previous work on the production of continuous pharmaceuticals [397].

Applying this collocation method on the dynamic optimization problem posed in Equation 7.14 results in the formulation of large system of algebraic equations as shown in Equation 7.19. First, Equations 7.15 and 7.16 describe the component concentrations in the liquid and gas phases by discretizing Equations 7.9 and 7.10 within each finite element j. Here, h_j is the finite element length and Ω_j is the Legendre polynomials describing the system states within a finite element. Equations 7.17 and 7.18 are included in the new problem formulation to guarantee continuity in the final optimal control policy solution. Additionally, the objective function is reformulated as a summation of profit and operating cost for each collocation point with in each finite element, as shown in Equation 7.19. This newly reformed system, now strictly a nonlinear programming (NLP) problem, can be solved to local optimality using the commercially available solver IPOPT developed by Wächter and Biegler [404].

$$P(t) = P_{q-1} + h_j \sum_{q=1}^{ncol} \Omega_j \left(\frac{t_q - t_{q-1}}{h_j} \right) \cdot \frac{dP}{dt_{j,q}} \tag{7.15}$$

$$P_g(t) = P_{g,q-1} + h_j \sum_{q=1}^{ncol} \Omega_j \left(\frac{t_q - t_{q-1}}{h_j}\right) \cdot \frac{dP_g}{dt_{j,q}} \tag{7.16}$$

$$P_q = P_{q-1} + h_j \sum_{q=1}^{ncol} \Omega_j \left(\frac{t_q - t_{q-1}}{h_j}\right) \cdot \frac{dP}{dt_{j,q}} \tag{7.17}$$

$$P_{g,q} = P_{g,q-1} + h_j \sum_{q=1}^{ncol} \Omega_j \left(\frac{t_q - t_{q-1}}{h_j}\right) \cdot \frac{dP_g}{dt_{j,q}} \tag{7.18}$$

$$\max_{F_1,F_2,F_g} \phi = \sum_{j=1}^{nfe} \sum_{q=1}^{ncol} \left[\sum_{i=A,B,E} C_i\left(\dot{m}^l_{i,j,q} + \dot{m}^g_{i,j,q}\right) - C_G \dot{m}^{in}_{G,j,q} \right]$$

$$\begin{aligned} &\text{Equations } 11-13 \\ &\text{Equations } 15-18 \\ &\dot{m}^l_i = P_i\left(F_1(t) + F_2(t)\right) \\ &\dot{m}^g_i = P_{g,i} F_g(\mathrm{t}) \\ s.t. \quad &0 \le F_1(t), F_2(t) \le 0.15 \cdot V_r \\ &0 \le F_g(t) \le 60 \cdot V_r \\ &P = \{G, X, A, B, E, R, T\} \\ &P_g = \{A, B, E\} \end{aligned} \tag{7.19}$$

To assess the impact of gas stripping on continuous ABE production, the optimal control problem will be solved for an industrially-sized bioreactor ($\mathrm{V_r} = 400\ m^3$) to maximize profitability and determine the control policy for the glucose feed F_1, media feed F_2, and gas feed F_g. The results of this optimization will be compared to the solution of the same optimal control problem in the absence of gas stripping $\left(F_g(t) = 0\right)$. While glucose is the main component of the operating cost, the cost of compressing the stripping gas and the condensation cost to recover the ABE product after leaving the reactor will be evaluated for each case.

7.2.3. Large-scale Process Synthesis for ABE Production

While maximizing the economic potential of the intensified multi-feed bioreactor is the first step toward developing an economical feasible process for acetone, butanol, and ethanol, the cost of product recovery and source of feedstock must be considered. This section discusses the development of a process synthesis framework to design a sustainable process to produce ABE using lignocellulosic biomass as the primary feedstock. A general process flow diagram is shown in Figure 7.3. The lignocellulosic biomass will enter the process and undergo pretreatment and hydrolysis to convert the cellulose and hemi-cellulose to the six-carbon glucose and five-carbon xylose monomers. These sugars are then converted using the multi-feed, intensified bioreactor developed in this work. Multiple bioreactors are considered to employ economy-of-scale and increase process profitability. The reactor effluent is processed using a system of distillation columns and decanters to produce pure acetone, butanol, and ethanol products. The resulting solid waste and water are sent to an energy generation section that produces steam for internal heating requirements and electricity as a fourth salable product. The general framework emulates that of our previous work for the process synthesis of cellulosic ethanol production [405]. The remainder of this section will briefly discuss these process subsections in more detail and outline the process synthesis framework that will be used to design an economically viable ABE production process.

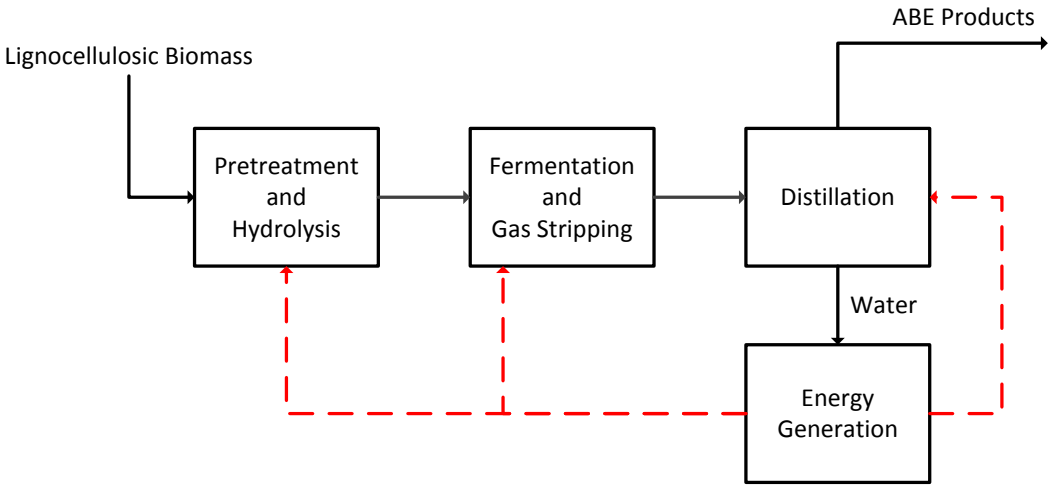

Figure 7.3. Simplified process flow diagram for the continuous production of acetone, butanol, and ethanol.

7.2.3.1. Upstream biomass processing and ABE production

The first step of the ABE production process is the selection and treatment of the biomass feedstock to produce simple sugars. Figure 7.4 shows the selection of prevalent biomass options across the United States, including agricultural residues such as corn stover (CS), sorghum (SOR), and sugar cane bagasse (SB), energy crops such as switchgrass (SW), and forest residues like hybrid poplar (HP). Each biomass type has a different composition of the cellulose, hemicellulose, and lignin polymers, resulting in a different sugar availability and different delignification requirements. Table 7.2 shows the over-dry composition, taken from the United States Department of Energy PLANTS database, and price point of the five biomass options considered in this work; prices were taken from Gabriel at al. for corn stover, Bain for switchgrass, sugarcane bagasse and hybrid poplar, and the World Bank Group for sorghum [96, 319-321]. As biomass moisture content is largely variable due to growth conditions such as growth region,

weather, soil type and fertilization, and harvesting and storage practices, the price point provided considers the off-site drying of the biomass used in the process [303].

When using lignocellulosic biomass as a primary feedstock, pretreatment is required to delignify and weaken the cell walls and allow for downstream hydrolysis of the sugar polymers to monomers. Many different pretreatment options are available that either chemically or physiochemically alter the structure of the cell wall. Chemical methods such as dilute alkali or acidic treatments, e.g. sodium hydroxide or sulfuric acid, cause swelling or decompose the hemicellulose to increase the internal surface area and reduce the rigidity of the cell wall, allowing enzymes to more freely attach the cellulose polymers [323]. Physiochemical methods typically use steam, liquid hot water, or aqueous ammonia utilize the chemical to hydrolyze the hemicellulose portion of the cell and then employ a quick pressure change to vaporize the active chemical and expand the cells [323].

Table 7.2. Oven-dried composition of each biomass type considered in the proposed separate hydrolysis and fermentation and consolidated bioprocessing bioethanol facilities [96, 319-321].

Biomass component	Corn stover	Switchgrass	Poplar	Sorghum	Bagasse
Glucan	0.3680	0.3464	0.4724	0.2887	0.3988
Xylan	0.2150	0.2477	0.1775	0.1454	0.2181
Water	0.0271	0.0215	0.0318	0.0365	0.0286
Lignin	0.1869	0.1806	0.2355	0.1372	0.2430
Other	0.2032	0.2039	0.0827	0.3923	0.1115
Price per tonne	$61.60	$81.50	$101.60	$173.90	$40.00

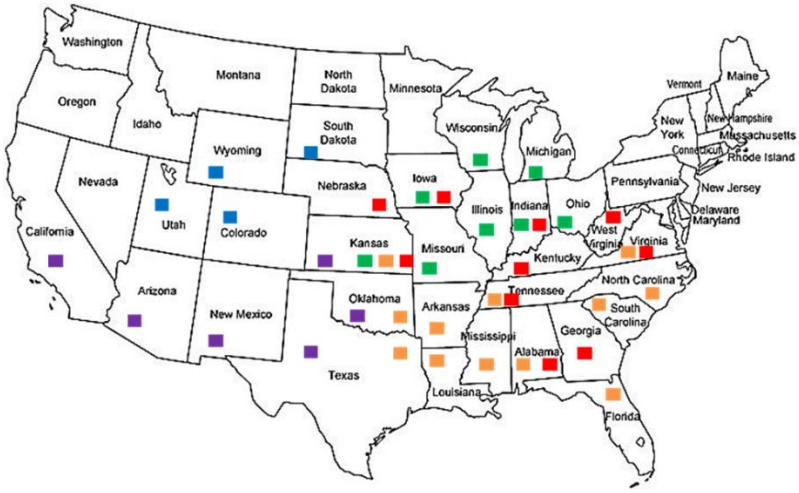

Figure 7.4. General availability of certain biomass types throughout the United States. Colors denote the specific types of biomass: blue → hybrid poplar, purple → sorghum, orange → sugarcane bagasse, red → corn stover, and green → switchgrass.

The process developed in this work looks to utilize one or multiple biomass options and a corresponding pretreatment for each biomass option selected to select the most profitable ABE process. The process is designed to utilize a maximum of 3,500 U.S. tons per day of lignocellulosic biomass and operates for 350 days per year. The general superstructure for upstream processing is shown in Figure 7.5, where five biomass types are coupled with four pretreatment methods based on data available from literature for their hemi-cellulose and cellulose conversions [324-331]. The data, including the biomass-pretreatment pairings, conversion of glucose and xylose, and operating temperatures can be found in our previous work [405]. Models based on the conversion of hemicellulose and cellulose are used to determine the availability of sugar monomers after pretreatment and enzyme hydrolysis, respectively. After the production of sugars, the hydrolysis effluent will be processed by a n multi-feed, intensified ABE

bioreactors discussed in the previous section to demonstrate the impact of economy-of-scale.

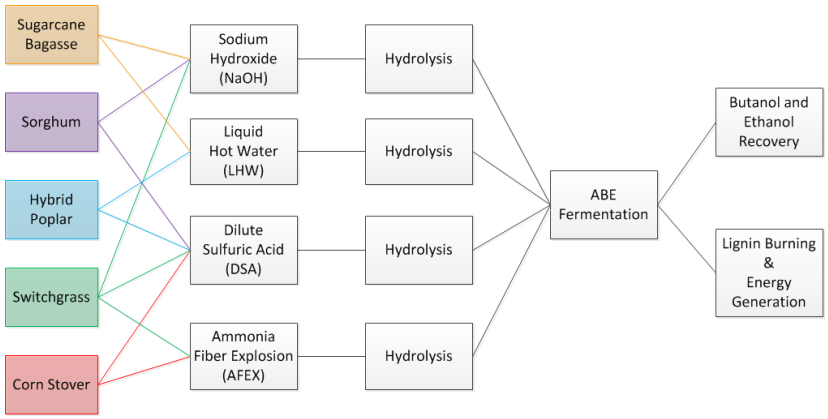

Figure 7.5. Process superstructure for the biomass and pretreatment method options.

7.2.3.2. Downstream separation and ABE recovery

Upstream production of ABE is followed by the downstream separation of the reactor effluent. This purification of the ABE mixture is complex due to the liquid-liquid and vapor-liquid azeotropes between the alcohol products with water. Methods to overcome these challenges using the well-known methods of distillation have been studied by many researchers, with the consensus process design using a series of distillation columns and decanters necessary to overcome the azeotropes and produce pure streams of acetone, ethanol, and butanol [406-408]. In this work, the downstream separation scheme proposed by Strobel and optimized by Wali will be used as it provides two advantages: (1) the predominant component, water, and the major solvent component, butanol, are removed as early as possible to reduce flow rates in later units and (2) one less column is required to effect the butanol separation [409, 410].

The process flow diagram for this separation train is shown in Figure 7.6. The first column utilized steam stripping with the reactor effluent to produce a bottoms product containing the solid components and most of the water. The distillate stream consists of the remaining water, acetone, butanol, and ethanol at a concentration close to the water/butanol azeotropic composition. This distillate stream is mixed with the bottoms stream of the third column and sent to a decanter where an aqueous-rich phase which is refluxed to the first column and a solvent-rich phase that is used as the feed to column two. The pure butanol product is removed as the bottoms stream of the second column and the remaining components exit the column as the distillate product and feed to the third column. This third column separated the mixture of the more volatile components, ethanol and acetone, from any remaining water and butanol. The bottoms stream of column three is mixed with the distillate from column one and decanted, which the acetone is separated from the ethanol in column four. Finally, the remaining ethanol and water mixture is sent to a molecular sieve to overcome the vapor-liquid azeotrope between the two components, producing a product stream of ethanol. Finally, the water outlet of the molecular sieve and the first column are combined and sent to the energy generation section of the process.

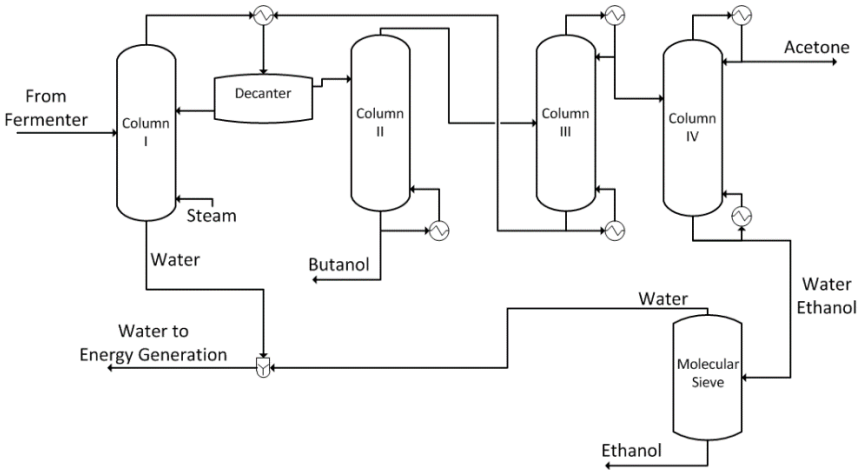

Figure 7.6. Downstream separation system for the recovery of the acetone, butanol, and ethanol products.

With a goal of developing the optimal process topology, models are necessary to describe the mass and energy balances of the separation process. Using data from the optimized system described by Wali, linear split fraction models are developed to determine the component mass balances as functions of the flowrate entering each column [410]. Additionally, linear energy balances for the reboilers and condensers are utilized by determining the heating and cooling requirements of the columns per kilogram of flow processed. These mass and energy balance models are shown in Equations 7.20 through 7.23, where m_b^i and m_d^i are the bottoms and distillate flowrates of column i, Q_C^i and Q_R^i are the energy requirements of the condenser and reboiler, respectively. The parameters ϕ^i, H_C^i and H_R^i are the split fraction, reboiler energy requirement per kilogram of feed, and condenser energy requirement per kilogram of feed for column i.

$$m_b^i = \phi \cdot m_f^i \tag{7.20}$$

$$m_d^i = (1 - \phi) \cdot m_f^i \tag{7.21}$$

$$Q_C^i = H_C^i \cdot m_f^i \tag{7.22}$$

$$Q_R^i = H_R^i \cdot m_f^i \tag{7.23}$$

7.2.3.3. Steam production and energy cogeneration

The last section of the ABE process topology uses the combustion of the solid residues of the fermentation process to produce low pressure steam for use in the heat exchangers found throughout the process as well as salable electricity. The models used for both steam generation and electricity production are taken from our previous work and are not discussed here for brevity [405].

7.2.3.4. Process synthesis framework for determining the optimal ABE production process

A process synthesis framework can be developed to determine the topology of the optimal, most economically attractive ABE production process. By combining the mass and energy balance models discussed in the previous three sections with cost models that determine the fixed capital cost, operating cost, labor cost of the process, a mixed-integer nonlinear program (MINLP) can be formulated and solved using commercially available solvers.

Equations 7.24 through 7.27 depict the models for the process expenses. First, Equation 7.24 calculated the combined purchase and installation cost of each piece of equipment by applying a sixth-tenths factor scaling based on a known size-cost pair [211, 303]. The size-cost pairs are taken from our previous work [405]. The fixed capital investment is then calculated using Equation 7.25, which sums up all the individual unit operation costs. The gas cost (GC) for the stripping gas is also included as it is assumed

that this gas will be recycled after the condensation step of the reactors. The total capital investment (TCI) can be calculated, of which it is assumed 85% is the FCI [342]. The labor cost (LC) is calculated using Equation 7.26 and are assumed to be proportional to the amount of ABE produced. The operating costs (OC) are calculated as the sum of all the required material costs, including the biomass cost (BC), pretreatment cost (PC), enzyme cost (EC), process water cost (PWC), the cost of the *C. acetobutylicum* fermentation organisms (CAC), utility cost (UC), and waste cost (WC).

$$C_{eq}^{j} = C_{eq}^{j,base}\left(\frac{S_{eq}}{S_{eq}^{j,base}}\right)^{0.6} \tag{7.24}$$

$$FCI = \sum_{j=1}^{n} C_{eq}^{j} + GC \tag{7.25}$$

$$LC = \begin{pmatrix} Labor\ per \\ Kilogram\ ABE \end{pmatrix}\begin{pmatrix} ABE \\ Produced \end{pmatrix} \tag{7.26}$$

$$OC = BC + PC + EC + PWC + CAC + UC + WC \tag{7.27}$$

With a set of models for the process economics, a complete formulation of the process synthesis problem can be developed. This work looks to minimize the cost of the butanol product to allow for a direct comparison with currently used fossil fuels and bioethanol processes. To guarantee profitability of the process, this objective will be calculated subject to a required discounted cash flow return on investment ($DCFROI$) of 15% over a time span (t) of 20 years. This is done by setting the net present value (NPV) of the process to zero with a value of the discount factor (q) equal to 15%. The profit P is calculated at a tax rate of 35% based on the gross profit (GP) of selling the butanol, ethanol, acetone, and electricity produced. The cost of the ethanol and acetone

are set at the cost values listed previously in Table 7.2. When the constraints of the mass balances, energy balances, and cost correlations are added, the full MINLP formulation is postulated. This optimization problem, shown as Equation 2.28, is solved using the BARON solver available in the GAMS software package to 1% optimality gap for different numbers of bioreactors n in the process to determine the economy-of-scale of cellulosic butanol production [65].

$$\begin{aligned} &\min C_B \\ &NPV = 0 = -TCI + P \cdot \left(\frac{(1+q)^t - 1}{q(1+q)^t} \right) \\ &P = (GP - OC - LC - D)(1-r) + D \\ s.t. \quad &GP = C_B \dot{m}_B + C_E \dot{m}_E + C_A \dot{m}_A + C_{El} \dot{G}_{El} \\ &\text{Mass Balance Equations} \\ &\text{Energy Balance Equations} \\ &\text{Cost Correlations} \end{aligned} \tag{7.28}$$

7.3. Results and Discussion

7.3.1. Optimal Process Intensification of the Multi-feed ABE Bioreactor

The solution to the optimal control problem defined as Equation 2.19 will now be discussed, beginning with the case of continuous processing in the absence of gas stripping, or $F_g(t) = 0$. The liquid concentration profiles for this case are shown in Figure 7.7, with steady-state operation being achieved after approximately 100 hours of operation. During start-up, the glucose concentration drops over the first 25 hours as the substrate is converted to butyrate and acetate. This sharp spike in acid concentration results in a shift to solventogenesis, corresponding to an increase and eventual maximum in the butanol, acetone, and ethanol concentrations. After achieving their peaks, the concentrations of all components settle into their steady state values. The specific steady

state concentrations for each the glucose, biomass, acetone, ethanol, and butanol component are listed in Table 7.2.

To achieve the profiles shown in Figure 7.7, the optimal control policy shown in Figure 7.8 for the glucose and media feeds is used. Over the first 50 hours of operation these flowrates are slowly elevated until reaching a cyclic steady-state around 20,000 L/hr and 10,000 L/hr, respectively. A more definitive look at the cycle steady-state glucose and media flowrates is shown in Figure 7.8b. Here, the glucose flow rate are cyclic within each hour, beginning at a high flowrate, decreasing to an intermediate flowrate, and finishing at a low flowrate before returning to the first high flowrate. The glucose flowrate varies between values of 25,000 L/hr, decreases by a little over 5,000 L/hr, and then decreases by another 3,000 L/hr before returning to 25,000 L/hr. The media flow rate follows a similar trend, beginning at a value of 14,000 L/hr, decreasing to a value of about 6,000 L/hr, and finishing at a value of about 4,000 L/hr before returning. These cyclic transitions result in the glucose inlet concentration to have a similar cyclic pattern as shown in Figure 7.8c. While the flowrates decrease over the course of each hour of operation, the glucose concentration increases due to the relative change in each flowrate. The glucose concentration begins at its lowest value of 64 g/L, rises to 76 g/L, and the rises again to 80 g/L before returning to the initial value.

In contrast, the concentration profiles for the multi-feed bioreactor intensified with *in situ* gas stripping. The inclusion of this separation technique complicates the operation of the bioreactor, resulting in the pseudo-batch operation that result in the addition of substrate, conversion to the desired products, and removal of the products via

dilution effects and gas stripping. The optimal concentration profiles for this pseudo-batch operation are shown in Figure 7.9. Figure 7.9a shows the liquid concentration profiles of the glucose, biomass, and normalized RNA concentration, where glucose is consumed and added in a cyclic nature that resemble batch operation. Each "batch" is approximately 100 hours long, during which the biomass and RNA concentration increase with the addition of glucose, decreases as the inhibitory butanol is produced, and continues to decrease during the dilution step as the ABE products are recovered.

Figure 7.9b, 7.9c, and 7.9d show the effects of this pseudo-batch operation on the acid concentrations, liquid ABE concentrations, and gas ABE concentrations, respectively. As expected, all concentrations follow the same cyclic nature of the glucose and biomass concentration. In the liquid phase, the acids acetate and butyrate reach a maximum by about 20 hours into the 100-hour "batch", resulting in a shift toward solventogenesis. This results in the production of the butanol, ethanol, and acetone until the 50-hour mark, marking the peak of the "batch" and the beginning of the removal step. This removal step continuous for the remaining half of the batch time, and at the 100-hour mark the solvent concentrations are at their minimum values to minimize inhibition and allow for cell growth to occur again. The deviation in the ABE concentrations in the gas phase mirror their liquid phase counterparts a result of the equilibrium relationships between the products in the two phases.

Figure 7.10 shows the optimal control policies of the glucose flow, media flow, and gas flow that result in the pseudo-batch operation of the multi-feed bioreactor. Like the case study without gas stripping, the media flowrate follows a cyclic nature within

each 100 hours, beginning the period with little flow and increasing over the first 30 hours to a flowrate between 50,000 and 60,000 L/hr as shown in Figure 7.10b. The gas flowrate shown in Figure 7.10c and Figure 7.10d shows an increase during the end of each of these media flow cycles, increasing the ability to remove the ABE productions and conditioning the media for the beginning of another 100 hour "batch". However, the optimal control policy for the glucose feed is less intuitive. This result is more complex than the typical constant or pulsing flowrate used in current biochemical practices, a direct result of the optimal control framework used to limit the glucose addition to only that which is required and increasing the profitability bioreactor process.

The complete comparison of the multi-feed bioreactor with and without the inclusion of gas stripping, including the process economics, is shown in Table 7.2. With respect to the liquid concentrations, the inclusion of gas stripping has no real effect on the average butanol or biomass concentrations; however, it should be noted that a higher ethanol, higher acetone, and lower glucose concentration is achieved when selective removal is used. A notable difference can also be seen when comparing the average media and glucose flowrates, where the inclusion of gas stripping and the need for pseudo-batch operation decreases the flowrates and therefore lowers the dilution rate of the overall system. The intensification of the continuous bioreactor has a negative effect on the process productivity, decreasing the ABE productivity from 0.55 g/L-hr in the absence of stripping to a value 0.37 g/L-hr. However, gas stripping results in a higher overall glucose conversion, resulting in a value of 72% for the intensified process and 23% otherwise.

As the objective of utilizing the multi-feed bioreactor with gas stripping was to maximize process profitability, the gross profit and raw material costs were analyzed. Costs of the glucose substrate and ABE products are listed in Table 7.1. In the case of gas stripping, the average glucose flowrate added to the bioreactor was 190.51 L/hr. At a concentration of 100 g/L, this results in an hourly glucose cost of $15.24 per hour. The resulting ABE components produce a gross profit of $192.42 per hour, resulting in a net gain of $177.18 per hour or $1,552,049.90 per year. Alternatively, the process without gas stripping has an average glucose flowrate of 19,706.24 L/hr at the same concentration of 100 g/L. This, coupled with the gross profit made from the resulting ABE products, leads to a yearly loss of $11,162,830.36. Without the use of the selective removal of ABE products via gas stripping, the continuous production of acetone, butanol, and ethanol is not profitable. However, due to the profitability of the reactor with gas stripping, the additional cost of a separation system and the use of lignocellulosic biomass as a carbon source can be investigated to determine the economic viability of ABE production.

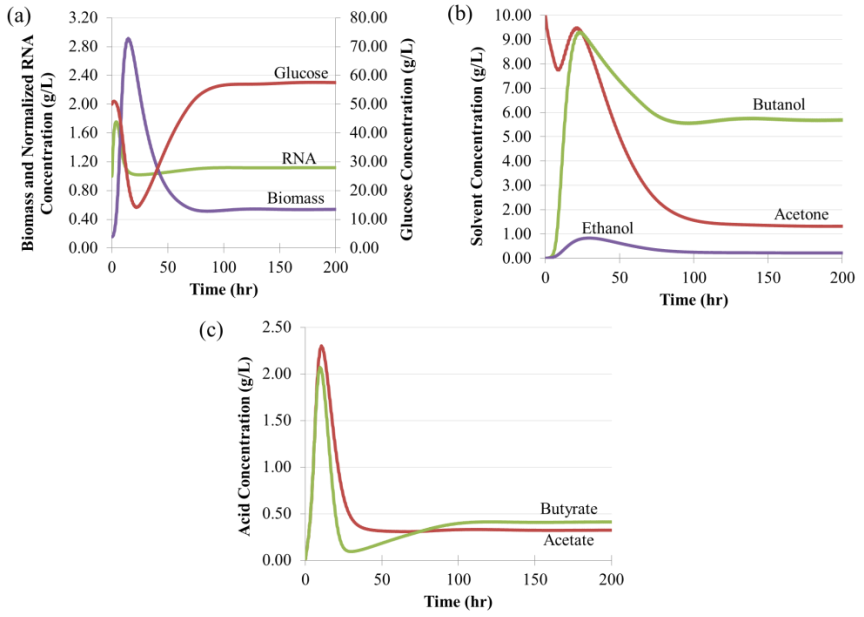

Figure 7.7. Concentration profiles for the (a) substrate, biomass, and RNA, (b) solvent, and (c) acid components of ABE fermentation in the absence of gas stripping.

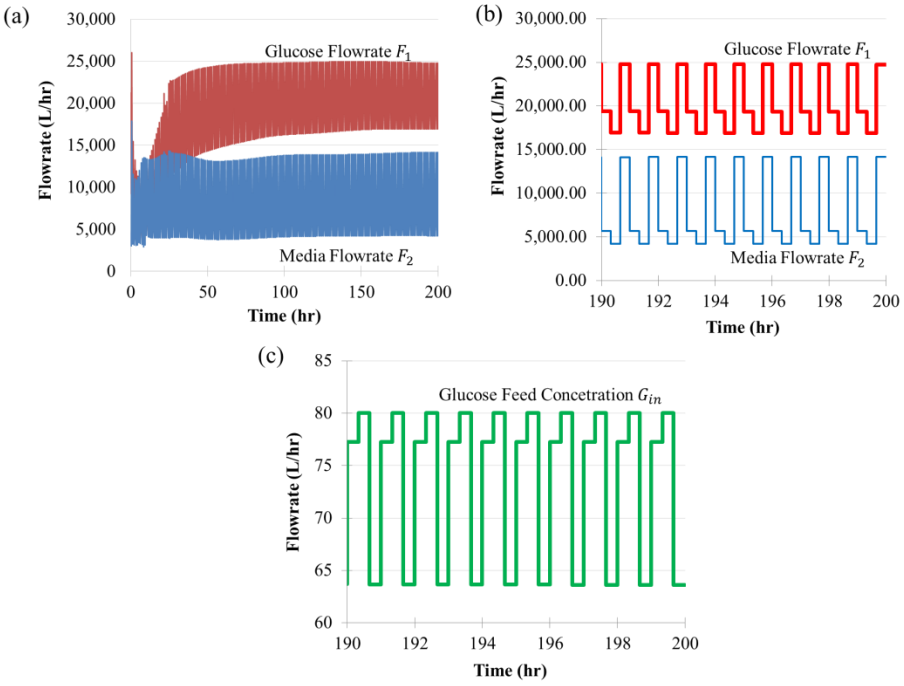

Figure 7.8. Optimal control policies for the glucose and media flowrates in the absence of gas stripping. Figures correspond to (a) the full optimal control profile, (b) the optimal control profile during steady state, and (c) the corresponding glucose inlet concentration at steady state.

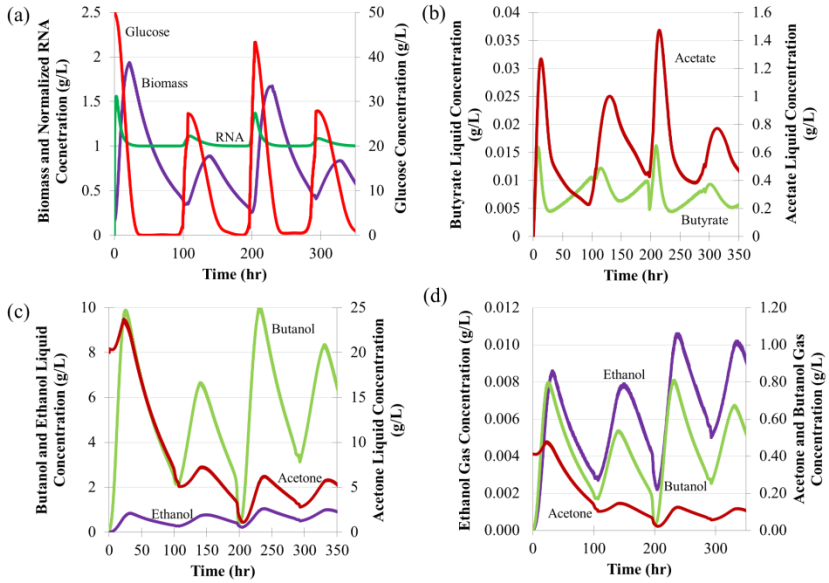

Figure 7.9. Concentration profiles for the liquid and gas phases in the multi-feed bioreactor with in situ gas stripping. Figures correspond to (a) liquid substrate, biomass, and RNA concentration, (b) liquid ABE solvent concentrations, (d) liquid acid concentrations, and (d) gas ABE solvent concentrations.

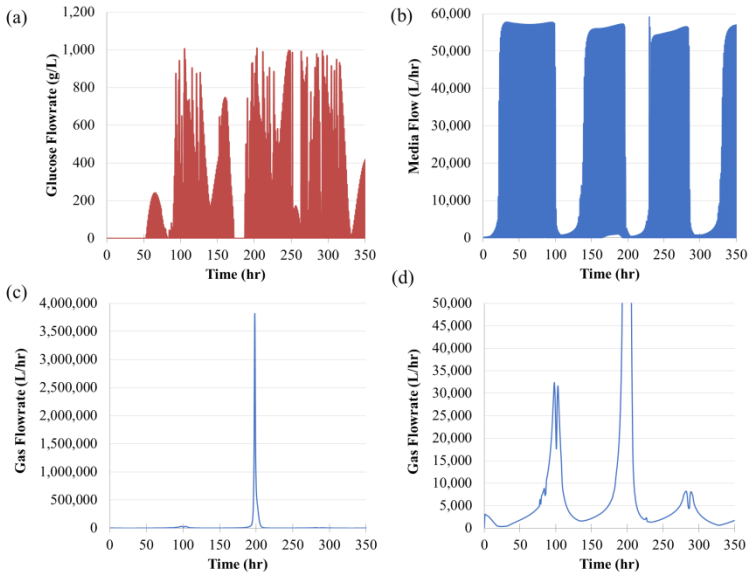

Figure 7.10. Concentration profiles for the liquid and gas phases in the multi-feed bioreactor with in situ gas stripping. Plots correspond to (a) the optimal control profile for the glucose flowrate, (b) the optimal control profile for the media flowrate, and (c) the optimal control profile for the gas flowrate, and (d) a more detailed plot of the gas flowrate optimal control profile.

Table 7.3. Comparison of the operation parameters as a result of the optimal control of the multi-feed bioreactor with and without gas stripping.

	No Stripping	Gas Stripping
Average Liquid Concentration (g/L)		
Glucose	55.21	9.56
Biomass	0.62	0.75
Acetone	1.86	4.09
Butanol	5.91	5.60
Ethanol	0.26	0.71
Average Gas Concentration (g/L)		
Acetone	-	0.15
Butanol	-	0.42
Ethanol	-	0.01
Average Flowrate (L/hr)		
Glucose Flowrate	19,706.24	190.51
Media Flowrate	7,751.77	12,485.34
Total Liquid Flowrate	27,458.01	12,675.85
Gas Flowrate	-	29,366.76
Average Dilution Rate (1/hr)		
Liquid	0.07	0.03
Gas	-	0.07
ABE Productivity (L/hr)		
	0.55	0.37
Total ABE Production (g/hr)		
Acetone	50,944.89	56,210.41
Butanol	162,266.79	83,367.26
Ethanol	7,022.87	9,234.36
Total Gross Profit per Year		
Hourly	$302.20	$192.42
Yearly	$2,647,303.78	$1,685,559.85
Raw Material Expenses		
Average Glucose Inlet Conc. (g/L)	73.52	33.96
Hourly Glucose Cost	$1,576.50	$15.24
Yearly Total Cost	$13,810,134.13	$133,509.94
Net Profit		
	-11,162,830.36	1,552,049.90

7.3.2. Process Synthesis for the Economically Viable Production of ABE

The process synthesis framework previously discussed is solved to determine the choice of lignocellulosic biomass, pretreatment method, and effect of an added separation system on the profitability of the intensified bioreactor developed in the previous section. Table 7.4 shows the results of this optimization. The optimal topology of the ABE process uses sugarcane bagasse with a sodium hydroxide pretreatment in lieu of any other biomass or pretreatment combination. This choice maximized the amount of glucose available to the process per dollar of operating cost. The process uses 283 U.S. tons of biomass to produce a total of 460,000 gallons of ABE product per year in a ratio of 4.6:8:1 of acetone to butanol to ethanol. The capital cost of the process is approximately $12 million dollars and costs approximately $1.3 million to operate. At the butanol price of $14.87 per gallon ($4.96 per kilogram), this process is not competitive with current fossil fuel prices or production rates; however, with the addition of more reactors the effects of economy-of-scale can be obtained.

Figure 7.11 shows the effects of multiple bioreactors on the butanol price, ABE production, capital cost, operating cost, and the resulting total cost of the ABE production process. From Figure 7.11a it can be seen that the addition of more reactors, up to a total of fifteen 400,000 L bioreactors, causes a linear increase of the ABE produced to a total of seven million gallons of alternative fuel. The butanol price, however, decays exponentially toward a lower limit of approximately $7 per gallon ($2 per kilogram) in response to this increase in product. Figure 7.11b shows the increase in the capital, operating, and total cost with an increase in reactor number. The operating

cost and capital cost increase linearly and exponentially, respectively, resulting in a total capital investment that also increases exponentially at a faster rate than the capital investment. While there is no theoretical limit to the number of reactors that can be employed, the land area that would be required for the implementation of this process would be immense and would restrict the application of this technology.

Table 7.4. Optimal topology and process economics for the complete ABE production process.

	Value	Units
Butanol Price and Profits		
Butanol Price	$14.87	per gallon
Gross ABE Profit	$4,363,606.20	per year
Gross Electricity Profit	$329,722.06	per year
Production Rates		
Acetone	156,124.77	gallons/year
Butanol	270,887.73	gallons/year
Ethanol	33,935.14	gallons/year
Total ABE	460,947.63	gallons/year
Electricity Generation	0.59	MW hr
Optimal Process Topology		
Total Biomass Feed	283.83	U.S. tons/year
Feed Composition	Sugarcane Bagasse	
Pretreatment Selection	Sodium Hydroxide	
Process Economics		
Capital Cost	$12,076,970.00	
Operating Cost	$1,297,275.60	per year

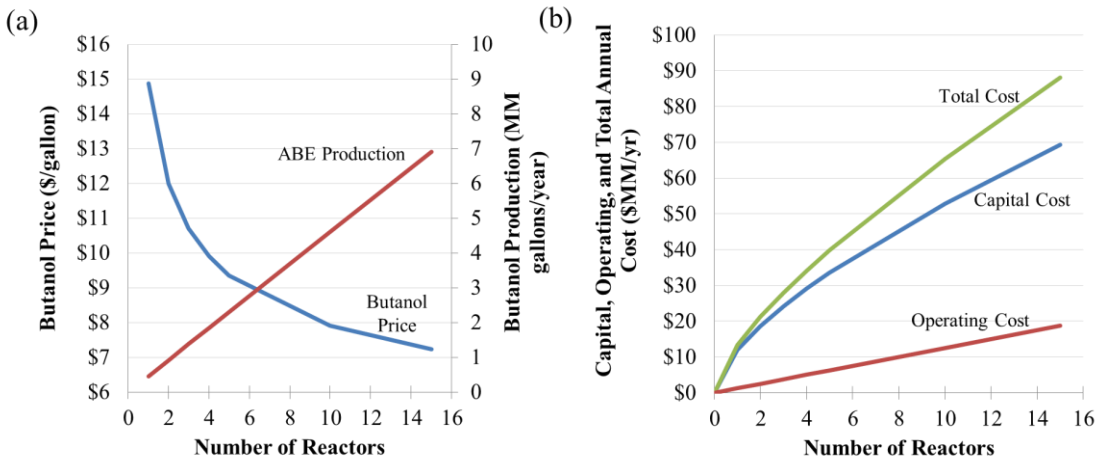

Figure 7.11. Economy-of-scale analysis for the complete ABE production process. Figures show the effect of additional bioreactors on (a) butanol production and butanol price and (b) the operating, capital, and total cost of the process.

7.3.3. Other Considerations to Improve Process Productivity and Economics

While the technology considered in this work is not economically competitive with current fossil fuels, there are many cellular and process considerations that can be made to enhance the process economics of ABE production. The models used in this work are for the wild type *C. acetobutylicum*, and in recent years genetic engineering has allowed for increases butanol tolerance and flux of carbon toward the desired butanol product [411-413]. However, even with these advancements in ABE fermentation at the microorganism level, mathematical models describing the changes in the cellular metabolism and product formation have yet to be developed. By exploiting the adaptability of the optimal control strategy and process synthesis framework proposed in this work, the updated cellular kinetics can be used to examine the improved profitability of ABE production at both the unit operation and process level.

At the process level, many researchers have considered alternatives to the enzymatic hydrolysis step. Consolidated bioprocessing, or CBP, is the use of genetically engineered microorganisms capable of fermenting cellulose directly to alcoholic products, forgoing the need for expensive enzymatic hydrolysis methods to produce simple sugars. This technology has already been shown to greatly improve the process economics of bioethanol production systems from lignocellulosic biomass, and organisms capable of consolidated bioprocessing for ABE fermentation are being developed [414-416]. Other separation methods have also been considered for butanol recovery as alternatives to the gas stripping and distillation used in this work. Two possible separation systems employing pervaporation membranes or selective adsorption have been shown to be profitable for ABE production on a variety of carbon sources [378, 380-382, 386]. The optimal control strategy proposed in this work can also be used as a tool for analyzing the operability and profitability of these alternative technologies.

7.4. Conclusions

Novel methods for the continuous production of alternative liquid fuels are needed to compete with currently employed petroleum processes. While certain fermentation products, specifically butanol and ethanol, are considered potential options, their inhibitory effects on the fermentation process pose challenges to their economic viability. To this end, the intensification of a continuous, multi-feed acetone-butanol-ethanol bioreactor with in situ gas stripping is proposed and optimized to maximize process profitability. The optimal intensified bioreactor is compared to a continuous ABE bioreactor absent of gas stripping, and results show that gas stripping is necessary

for a profitable process due to an increase in glucose utilization despite a loss of process productivity. With the determination of a profitable bioreactor, the economic viability of a complete ABE process using lignocellulosic biomass as a carbon source is determined by calculating the butanol price necessary to achieve a DCFROI of 15%, guaranteeing process profitability. Results show that this while not competitive with present petrochemical prices, the butanol cost can be reduced through the use of additional reactors, achieving the effects of economy-of-scale and opening the door for further improvements from genetic engineering, alternative selective separation systems, and consolidated bioprocessing. Overall, this work shows that the production of ABE should be considered a potential replacement to fossil fuels and requires further investigation.

CHAPTER VIII

A HOLISTIC APPROACH TO THE INSTRUCTION OF PROCESS INTEGRATION AND ECONOMICS

8.1. Introduction

Learning basic process systems engineering (PSE) methodologies such as heat integration [86-88, 417-420], mass integration [421-429], cost estimation, and financial integration [430] at the undergraduate level is critical for chemical engineering students to be able to design economically viable and environmentally sustainable chemical processes. In fact, a sound knowledge on process integration and economics comprises a prerequisite for capstone plant design in most chemical engineering curriculum. Recognizing the impact that process integration and economic analysis can have on profit, environmental footprint, raw material usage, energy needs and capital investment risks, many textbooks and teaching materials have been developed in the past on these topics [21, 211, 340, 342, 431-438]. Perkins (2002) detailed two complementary approaches to teaching PSE concepts in chemical engineering curricula. The first approach provides students with the knowledge of PSE tools as they are developed, while the second approach looks to develop a general perspective based on "systems" thinking that can be applied to a broader range of problems [438]. Seider *et al.* (2016) and Biegler *et al.* (1997, 2010) described systematic approaches for the design and optimization of chemical processes and included product design and life cycle analysis in process design [431, 436, 439]. Klemes *et al.* (2013) discussed visualization techniques and interactive multimedia for enhanced learning [437]. Kiss *et al*. discussed

process integration for the biochemical engineering field, remarking that any biotechnology based processes exhibit potential for integration to improve performance [440].

The purpose of this chapter is not to introduce new concepts in PSE literature, but to elucidate a holistic framework for teaching process integration and process economics. Specifically, an approach is proposed that generalizes the common ideas and underlying concepts from direct recycling, mass integration, heat integration and financial integration. Of particular interest is the pinch method [419, 424] that cuts across mass and heat integration in the domains of process integration [342, 434], and recently introduced financial integration [430] in the domain of process economics. By incorporating the ideas of financial integration with process integration, the generalized pinch method is shown to serve as the lynchpin between the core fundamentals of chemical engineering, *i.e.*, mass and energy balances, mass transfer, heat transfer, and process economics, and their applications in both traditional chemical engineering and biochemical processing.

In this chapter, the instruction of *financial integration* as a capstone to teaching process economics is proposed that can serve as the missing link between topics. The concept of financial integration can be considered in the same vein as mass and energy integration, used to mitigate the risk associated with the allocation of finances to process improvement expenditures. These four integration techniques for seemingly different applications represent the quadrants of the holistic instructional framework posed in this work. This is shown in detail in Figure 8.1., with definitions given for the typical sources

and sinks of each integration type. By introducing the general graphical pinch strategy for integration, and developing a novel algebraic approach that emulates that of the other integration methods, the connection between process economics and process integration will be made apparent to students and help smooth the transition.

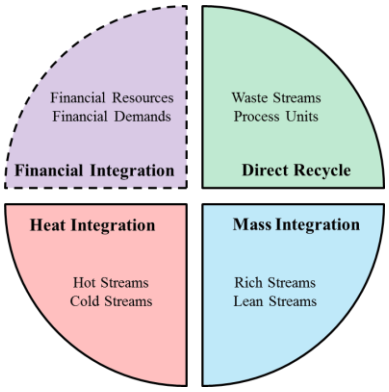

Figure 8.1. Four quadrants of the general process integration framework. Definitions of the sources (top) and sinks (bottom) for each integration method are listed in each quadrant. The dotted line depicts the missing quadrant of the holistic framework.

We then propose a holistic framework to the instruction of process economics and integration based on the collective understanding of pinch analysis for various integration strategies, including the newly developed methods for financial integration discussed in the next section. This framework uses the instruction of more general pinch analysis methods to develop a core understanding on the overall methodology, which allows for a more in-depth examination of the key subtleties and differences between the applications of direct recycle, mass integration, heat integration, and financial integration. By separating the methodology from the application, a larger emphasis can be placed on the physical meaning behind these analyses as well as the implementation of key findings. Additionally, the intent of this work is to seamlessly bridge the gap

between process economics and process integration by focusing on the similarities between the pinch analyses methods of each discipline. This method will be exemplified using a classical chemical engineering application as well as a biochemical application.

The chapter is organized as follows. First, a brief review of the general graphical and algebraic pinch analysis methods will be given for direct recycle, heat integration, and mass integration applications. Second, the idea of a graphical method for financial integration will be introduced and the methodology for a new algebraic method will be developed. Third, the current linear approach to teaching process economics and integration will be introduced and compared to a newly proposed holistic approach that looks to use financial integration and generalized methods for process integration to bridge the gap between core chemical engineering fundamentals and their integration applications. Finally, two case studies are presented to demonstrate the use of the generalized pinch methodology. Graphical methods for integration are exemplified for a generic chemical process and an algebraic approach including the newly developed algebraic financial integration method will be demonstrated using a bioprocessing example.

8.2. A Brief Review of Pinch Analysis for Process Integration

Pinch analysis determines the maximum possible use of currently existing process resources to minimize external requirements. These resources, termed as *sources*, provide opportunities for the recovery, recycling and/or reuse of mass and energy. Examples of sources in mass integration methodologies include waste streams leaving the process. In heat integration, sources primarily include hot process streams

with excess heat. Locations that accept these otherwise rejected resources are termed *sinks* and are often constrained by operating conditions such as impurity level or temperature. Sinks for mass integration applications are usually process units that utilize a raw material. For energy integration, they primarily include cold streams requiring the use of a heating utility. Process integration strategies look to match the available resources in the source streams with the process sinks to reduce the required external sources to their minimum values. Successful integration yields an efficient use of process resources and minimizes its financial and environmental impact. Based on the laws of conservation of mass and energy, strategies have been developed for *direct recycle* of process streams to recover lost raw materials, for *mass integration* to reduce the use of external solvents for removing stream impurities, and for *heat integration* to lessen external utility usage.

8.2.1. Graphical Approach

A general pinch diagram is shown in Figure 8.2 to help illustrate the graphical pinch analysis technique. Pinch analysis begins with the development of a source composite curve (shown in red in Figure 8.2) that contains the information for available resources in the nominal process. A sink curve (shown in blue in Figure 8.2) is also developed illustrating the information for the available locations within a plant that these excess resources may be utilized. The solution to a pinch diagram is found by shifting either the source or sink curve, depending on application, until the pinch point is found. The *pinch point* is defined by the point at which the sink and source curves touch at a single or multiple location(s), with the source curve lying solely underneath the sink

curve. Physically, the pinch signifies the location below which no waste should be discharged from the sources and above which no fresh resources should be used in the sinks [342]. The solution of the pinch analysis provides three key values for the integration implementation: the minimum external resource to be used (region 1 in Figure 8.2), the optimal amount of integration between sinks and sources (region 2 in Figure 8.2), and the minimum excess of resources that cannot be used due to operational constraints (region 3 in Figure 8.2). Using these values, a strategy for minimum targets can be developed.

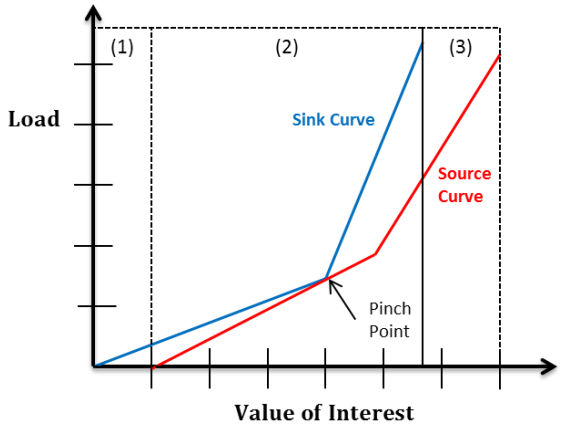

Figure 8.2. A generalized pinch diagram for process integration applications including the source curve (red), sink curve (blue), and pinch point location. Regions denote the minimum external resource to be used (region 1), the optimal amount of integration between sinks and sources (region 2), and the minimum excess of resources that cannot be used due to operational constraints (region 3). The value of interest and load quantity varies with each integration application.

The difficulty in understanding when using pinch analysis for process integration is the small subtleties of each application. These differences are mainly two-fold: (a) the important values to be plotted on the axes of the pinch diagram and (b) how to shift the graph to locate the correct pinch point. Regarding the axis values, the vertical axis for the pinch integration relates to a transferrable *load* important to the specific integration

strategy; these loads are to be passed directly from the sources to the sinks. The load is the impurity load in direct recycle, the mass exchanged in mass integration, and the heat load (Q) in heat integration. The corresponding horizontal axis values in a pinch diagram relates to the unit of interest driving the transfer of that load. This includes the cumulative flowrate (F) in direct recycle, rich stream mass fraction (y) in mass integration, and hot stream temperature (T) in heat integration.

The method to determine the location of the pinch point also differs between specific applications of these process integration strategies. For a strategy like direct recycle, where the horizontal axis is based on the cumulative flow of the sources and sinks, horizontal translation of the source curve indicates the addition of a pure source stream, typically an external fresh resource. However, horizontal translation in heat or mass transfer would require changing defined compositions or temperatures, inherently changing the integration problem. In these methods, the pinch is found by vertical translation of the sink curve, which represents the use of an external mass exchange agent to help remove the impurity load from the rich streams or the addition of an external source of heating to provide energy to the cold streams. Figure 8.3 depicts the methods for translating the source or sink curve for the three integration methods.

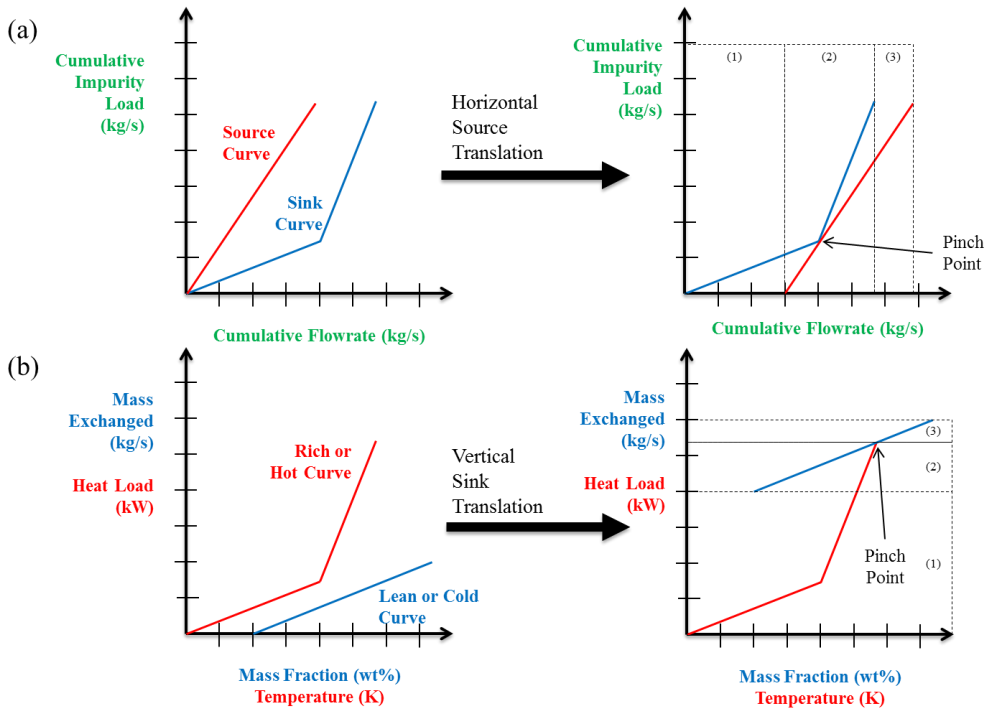

Figure 8.3. Translation methods for the determination of the pinch point in (a) direct recycle and (b) heat and mass integration. The resulting regions correspond to (1) external resource requirement, (2) integration of resources within the process, and (3) excess process resources.

The correct translation of the pinch diagram results in three distinct regions that correspond to physical properties of the integrated system. The integration region, labeled (2) in both Figure 8.3a and Figure 8.3b, denotes the overlapping region of the source and sink curves that indicate the ability to transfer mass or energy between process streams. The region labeled (1) represents the need for the addition of resources to compensate for the lack of complete integration. More specifically, in Figure 8.3a region (1) spans from the vertical axis to the beginning of the source curve and depicts the amount of fresh raw materials needed to supplement the recycled raw materials. In Figure 8.3b, this same region stretches from the horizontal axis to the bottom of the lean

or cold curve and signifies the external mass exchange agent or cooling utility needed in the system. Similarly, the region labeled (3) in Figure 8.3 signifies excess resources in the process. In the direct recycle case of Figure 8.3a, this value is the excess source streams that have a level of impurity too high to segregate and mix to reutilize per the sink specifications. In Figure 8.3b, these excess resources come in the form of excess mass exchange agent due to thermodynamic equilibrium limitations (mass integration) or excess cold utility that needs to be heated using an external heating utility (heat integration). Due to the significance of these three regions of the pinch diagram, understanding the correct way to translate the pinch diagrams is vital to determining the maximum amount of allowable integration and the minimum resource requirements within a process.

8.2.2. Algebraic Approach

Like the graphical methods previously discussed, algebraic methods for pinch analysis have been developed for heat, mass, and direct recycle applications to determine the maximum integration of process resources, thereby minimizing external resource usage [342, 441-443]. These methods focus on the idea of an *interval diagram*, where the two-dimensional nature of the graphical pinch method is converted to a one-dimensional visualization to determine intervals where resources can be exchanged. The specific representation of data in an interval diagram is unique to the application of the graphical pinch method. The implementation for direct recycle is called the *load interval diagram* (LID) and is depicted in Figure 8.4. Similar to the graphical method, the *NI* available process sources are arranged in order of increasing impurity level ($y_1 < y_2 <$

$\cdots < y_{k-1} < y_k$) while the *NO* available process sinks are arranged in order of the increasing maximum allowable impurity level ($z_1^{max} < z_2^{max} < \cdots < z_{k-1}^{max} < z_k^{max}$). Using this arrangement, arrows representing the source streams are aligned tip-to-tail vertically beginning at the top of the diagram where the impurity load is defined as zero; similarly, the process sinks are separately stacked starting from the zero point.

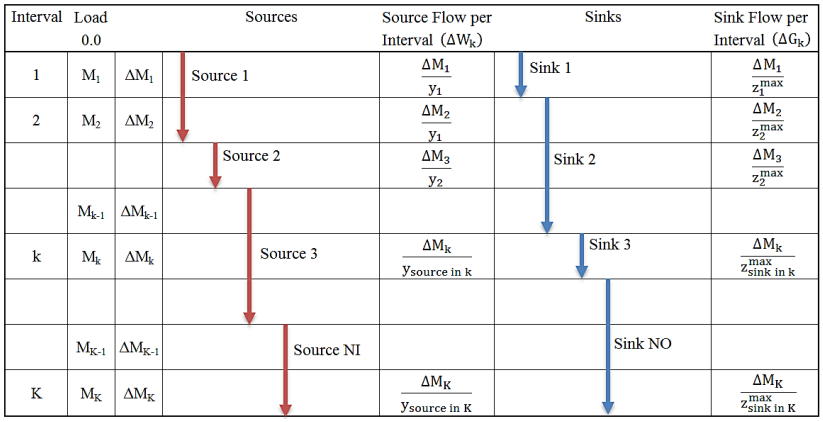

Figure 8.4. Load-interval diagram (LID) for the algebraic approach to direct cycle. Sources and sinks are ordered based on increasing impurity level ($y_1 < y_2 < \cdots < y_N$ for sources and $z_1^{max} < z_2^{max} < \cdots < z_N^{max}$ for sinks) and arranged tip-to-tail to determine the cumulative load [443].

Horizontal lines are then drawn across the diagram at each tip and tail of the source and sink arrows; the horizontal regions between these lines are known as *intervals*, in this case load intervals, where the flow from the source curves can be administered to the sinks. Finally, the source flow per interval, denoted ΔW_k, is determined by Equation 8.1, where the load range in interval k ΔM_k is divided by the impurity load of the source in that interval. A similar procedure is done to calculate the sink flow per interval, or ΔG_k, by dividing ΔM_k by the maximum allowable impurity load for the sink in interval k. This is also shown in mathematical form as Equation 8.2.

$$\Delta W_k = \frac{\Delta M_k}{y_{\text{source in interval k}}} \tag{8.1}$$

$$\Delta G_k = \frac{\Delta M_k}{y_{\text{sink in interval k}}} \tag{8.2}$$

Algebraic pinch methods can also be used for the integration of heat and mass. These methods also begin with the development of an interval diagram, specifically the composition interval diagram (CID) and the temperature interval diagram (TID) for heat integration. Examples of these diagrams are shown in Figure 8.5 and Figure 8.6, respectively. While like the implementation of the LID, these two interval diagrams are more complex as they use multiple vertical axes that allow for more than a single source or sink to exist in any given interval. For the CID, rich streams are plotted on a single axis based on their supply (y^s) and target compositions (y^t). To ensure feasibility, each lean stream has its own scale that converts the rich compositions y to the corresponding lean compositions *x* for every given lean stream by using the equilibrium data ϵ_i and b_i. The supply composition (x_i^s) and target composition (x_i^t) for each lean stream are plotted against their respective *x*-scale. For example, Figure 8.5 uses a single scale for all process rich streams, but employs a separate scale for each of the three lean streams in the process based on mass transfer equilibrium. Similarly, the TID uses a multi-scale system; the first scale represents hot stream temperatures (T) while the second represents cold stream temperatures (t). These two scales are separated by a minimum temperature difference (ΔT_{min}) that acts as the driving force for heat transfer and ensure feasibility. The *NH* hot streams are plotted against one scale based on their supply (T_i^s) and target (T_i^t) temperatures, whereas the *NC* cold streams are plotted against the

alternate scale based on their own supply (t_i^s) and target (t_i^t) temperatures. Figure 8.6 exemplifies the complete TID for a process that has two hot streams and two cold streams available for heat integration.

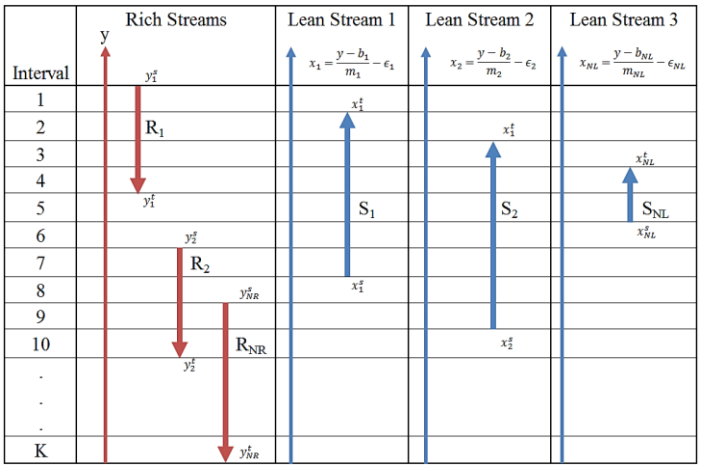

Figure 8.5. Composition-interval diagram (CID) for the algebraic approach to mass integration. Rich stream arrows are arranged based on their supply (y_i^s) and target (y_i^t) compositions. Supply streams are arranged by their supply $\left(x_j^s\right)$ and target $\left(x_j^t\right)$ compositions after converting to their equilibrium *y* values [342].

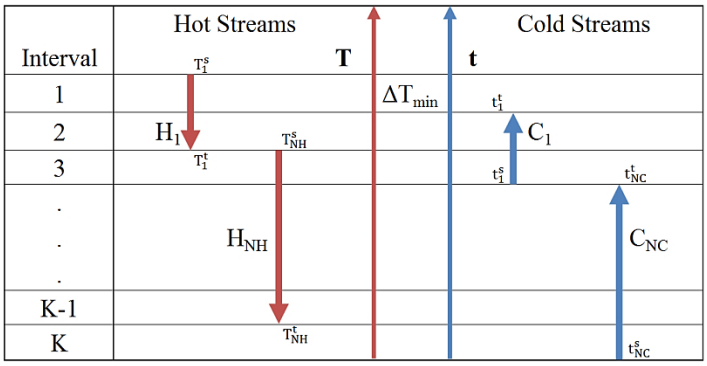

Figure 8.6. Temperature-interval diagram (TID) for the algebraic approach to heat integration. Hot and cold streams are arranged by descending temperature against their respective temperature scales that are separated by an amount of ΔT_{min} [342].

Once the sink and source streams have been placed on the CID or TID, intervals can be determined by placing horizontal lines across the tips and tails of each source and

sink arrow. Within each interval, the available resource load and resource needs can be calculated for each stream *i*. For mass integration, this corresponds to the available impurity load $\left(\Delta R_k^i\right)$ and the amount of impurity that can be removed by the lean stream $\left(\Delta L_k^i\right)$. For heat integration, this is the energy available from hot streams $\left(\Delta H_k^i\right)$ and the energy needed to heat cold streams $\left(\Delta C_k^i\right)$. Equations for calculating these values is shown in Equation 8.3 and Equation 8.4, where F_i is the flowrate of stream *i*, Δy^k is the rich stream composition range for interval *k*, Δx_i^k is the lean stream composition range for stream *i* in interval k, the quantity $\left(FC_p\right)_i$ is the product of the flowrate and heat capacity for a hot stream *i*, the quantity $\left(fc_p\right)_i$ is the product of the flowrate and heat capacity for a cold stream *i*, and ΔT_k is the temperature range in interval *k*. Finally, the total amount of available resources $\left(\Delta R_k^T \text{ or } \Delta H_k^T\right)$ and total amount of needed resources $\left(\Delta L_k^T \text{ or } \Delta C_k^T\right)$ in each interval can be calculated using Equation 8.5 and Equation 8.6. Here, the total amount of available resources is simply the sum of all ΔR_k^i or ΔH_k^i for source streams in interval *k* and the total amount of needed resources is the sum of all ΔL_k^i or ΔC_k^i for sink streams in interval *k*.

$$\Delta R_k^i = F_i \Delta y^k \;\forall\; NR \in k \qquad \text{or} \qquad \Delta H_k^i = \left(FC_p\right)_i \Delta T_k \;\forall\; NH \in k \tag{8.3}$$

$$\Delta L_k^i = F_i \Delta x_i^k \;\forall\; NL \in k \qquad \text{or} \qquad \Delta C_k^i = \left(fc_p\right)_i \Delta T_k \;\forall\; NC \in k \tag{8.4}$$

$$\Delta R_k^T = \sum_{i=1}^{NR} \Delta R_k^i \qquad \text{or} \qquad \Delta H_k^T = \sum_{i=1}^{NH} \Delta H_k^i \tag{8.5}$$

$$\Delta L_k^T = \sum_{i=1}^{NL} \Delta L_k^i \qquad \text{or} \qquad \Delta C_k^T = \sum_{i=1}^{NC} \Delta C_k^i \tag{8.6}$$

Once the important resource loads and resource needs for each interval have been calculated for any integration strategy, balances around the interval are calculated as shown in Figure 8.7. Process resources, including the flow of raw materials (direct recycle), energy (heat integration), and stream impurities (mass integration), are added to each interval based on the source process streams in interval *k*. Additionally, excess process resources from the previous interval *k*-1 can be passed down to interval *k*. Process sinks in interval *k* utilize these resources, and any excess resources are provided to the following interval *k*+1. These balances are done for all intervals *K* in the interval diagram, and the set of interval balances are stacked upon each other into a *cascade diagram* as seen in Figure 8.8a.

Each δ_k between the intervals is the amount of resources passed between interval *k* and interval *k*-1, and the pinch point is defined as the δ_k value(s) that are equivalently zero. To determine this location, the cascade diagram is *revised* by adding absolute value of the most negative δ_k value to the top of the cascade diagram (δ_0). This results in modified δ_k' values in the *revised cascade diagram* shown in Figure 8.8b; at least one of these values will correspond to the pinch point. The value of δ_0' at the top of the revised cascade diagram corresponds to the minimum amount of resources that need to be added to the process, i.e. flow of fresh raw materials, impurity load to be removed using external lean process streams, or hot utility to be used in a process. Alternatively, the δ_K' value at the bottom of the revised cascade diagram represents the minimum amount of

excess process resources that cannot be integrated in the process, i.e. flow that must be discarded as waste, internal mass exchange agents with equilibrium compositions too large, or cold utility needed to cool the remaining hot stream requirements.

The understanding of process economics is fundamental in understanding the impact of process integration strategies on cost savings and sustainability, either in the form of reducing raw material requirements, mitigating energy needs, or eliminating process units. Topics in process economics include the estimation of capital and operating costs, the evaluation of the impact of time on the value of money, and the development of economic indicators, e.g. net present value (NPV), return on investment (ROI), payback period (PBP) and discounted cash flow return on investment (DCFROI), that demonstrate the economic viability of a given process design and act as a springboard to introduce integration as a method for improving this viability.

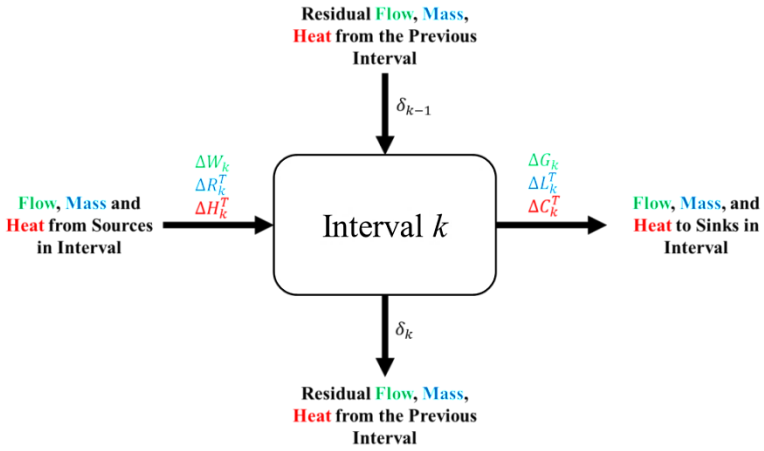

Figure 8.7. Generalized resource balance around a single interval. Resources from process sources introduced at interval k or excess resources from previous intervals (δ_{k-1}) can be utilized for sinks introduced at interval n or sent to subsequent intervals (δ_k).

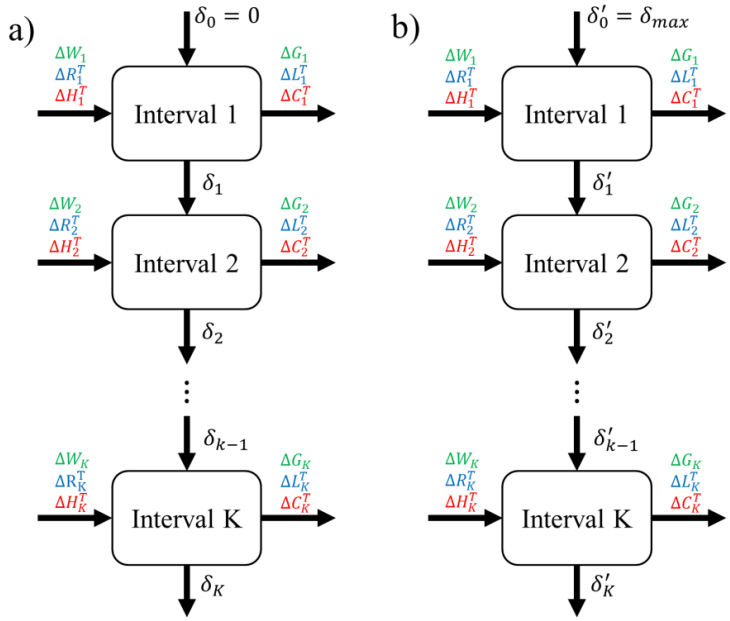

Figure 8.8. (a) Cascade and (b) revised cascade diagrams for the generalized algebraic pinch method. The initial cascade diagram is revised by equating δ_0 to the absolute value of the most negative δ_k value (δ_{max}) at the top of the cascade diagram and recalculating all δ'_k. The pinch point is recognized as the δ'_k value which is equal to zero. The δ'_0 value is the minimum amount of external resources that need to be added and the δ'_N value is the excess amount of process resources.

8.3. Closing the Loop with Financial Integration

While the synergy between process economics and process integration cannot be refuted, in our experience the currently disjointed teaching methodologies for these topics often leave students lost and confused during the transition between topics. The association of numerical calculations for process economics with graphical methods often comes across as jarring. One of the ways this can be overcome is through a project that connects the two topics seamlessly. Students are asked to choose a process from the process flow diagrams made available from SRI International and utilize their knowledge of process economics and process simulation, e.g. ASPEN Plus and ASPEN Economic Analyzer, to evaluate the existing process design. They are then asked to

suggest, implement, and evaluate integration strategies to improve the process economics. Typically, the evaluation of the integration strategies follows a sequential approach – first mass flows are addressed, via direct recycle or mass integration, then energy flows are integrated via heat integration. The economic improvement offered by these integration strategies is then evaluated. In this way, students can see the direct connection between process design, integration, and economics when applied to real chemical processes.

While this approach is effective for exemplifying the connections between topics, a project that combines topics in the way described must be done after the necessary material is covered during instruction periods. Like the methods of direct recycle, mass integration, and heat integration, financial integration relies on pinch analysis to maximize the use of monetary resources and minimize spending waste. The specific details of financial integration, while described here and depicted in Figure 8.9, were developed by Bandyopadhyay *et al* [430]. Financial integration examines the optimal distribution of monetary *sources*, e.g. grants, company budgets, etc., toward various monetary *sinks,* e.g. capital projects, to maximize budgetary use and mitigate the risk associated with the allocation of finances to process improvement expenditures. The transferable *load*, equivalent to the heat load in heat integration or impurity load in mass integration, is the annual returns (in $/year) of each capital project. These are plotted as a function of the investment cost for each capital project, with the slope of each line segment denoting the payback period of the project. Determining the optimum level of financial integration is done via pinch analysis, albeit differently than any other

integration method; the pinch point is located by the shifting of the source curve along each vertex of the sink composite curve until it lies completely underneath. Locating the pinch point results in three distinct methods as depicted in Figure 8.9. Region (1) signifies all projects that are not able to be funded with their current investment cost and annual return. Projects that can be funded are shown by the overlap of the sink and source curves, highlighted by region (2). Finally, excess financial resources that cannot be used in the current projects being considered are shown in region (3).

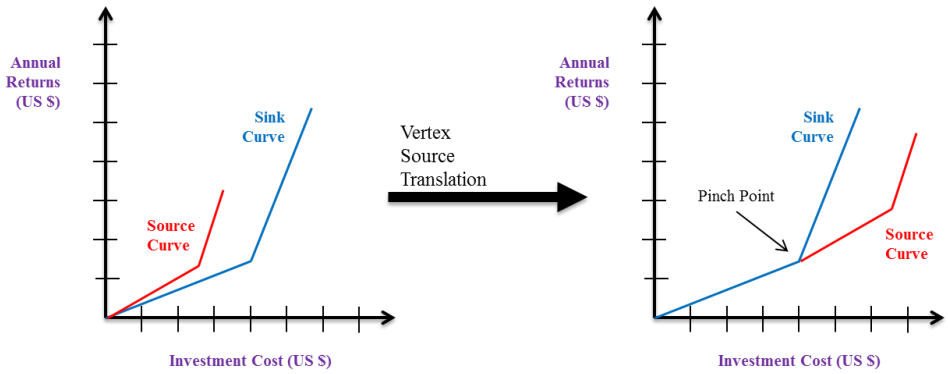

Figure 8.9. Economic equivalent pinch point method for the determination of funding allocations to capital projects. The resulting regions correspond to (1) projects without funding resources, (2) projects that can be funded, and (3) excess financial resources. [430].

Although it has been demonstrated that a graphical pinch approach to financial integration is like those for the other integration applications, an algebraic pinch approach has not been developed until now. Figure 8.10 depicts the novel fiscal-interval diagram (FID) to be used similarly to the load-interval diagram (LID) in direct recycle integration strategies. The algebraic method begins by calculating the maximum allowable payback period $(MAPP)$ for all available funding sources and the expected

payback period (EPP) for projects that need funding, defined as the ratio of the investment cost (IC) and the annual returns (AR). These are shown as Equation 8.7.

$$\mathrm{MAPP_i} = \frac{\mathrm{IC_i}}{\mathrm{AR_i}} \qquad \text{or} \qquad \mathrm{EPP_i} = \frac{\mathrm{IC_i}}{\mathrm{AR_i}} \tag{8.7}$$

To begin the FID, the *NS* available funding sources are arranged in order of decreasing MAPP ($MAPP_1 > MAPP_2 > \cdots > MAPP_{NS-1} > MAPP_{NS}$) while the *NP* available projects are arranged in order of the decreasing EPP ($EPP_1 > EPP_2 > \cdots > EPP_{NP-1} > EPP_{NP}$). Arrows representing the annual returns of each funding source are aligned tip-to-tail beginning at the top of the diagram that indicated zero annual returns; arrows representing the annual returns for the projects are aligned in a similar representation also beginning at zero annual returns. Intervals are determined by drawing horizontal lines across the diagram at the tips and tails of all source and project arrows. Once the intervals have been determined, the investment cost in each interval can be calculated. For the funding sources the investment cost (ΔI_k) is calculated as the product of the annual returns in each interval (ΔF_k) and the MAPP of the funding source in interval *k*, as shown in Equation 8.8. A similar calculation is used to determine the investment cost of the projects (ΔP_k), where ΔF_k is multiplied by the EPP of the funding source in interval *k* as in Equation 8.9.

$$\Delta I_k = \Delta F_k \cdot MAPP_k \tag{8.8}$$

$$\Delta P_k = \Delta F_k \cdot EPP_k \tag{8.9}$$

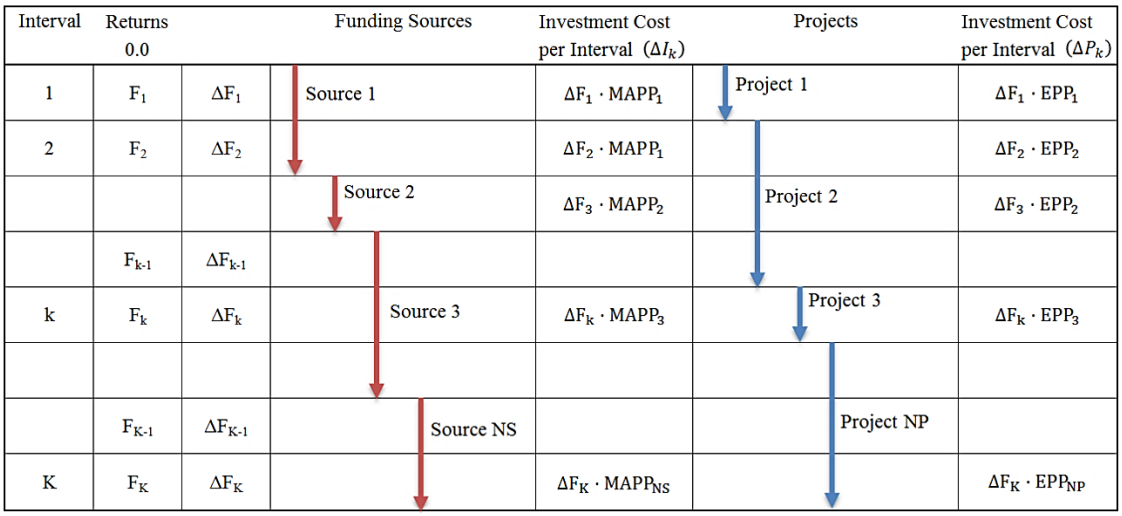

Figure 8.10. Fiscal-interval diagram (FID) for the algebraic approach to financial integration. Sources and sinks are ordered based on decreasing payback period ($MAPP_1 \geq MAPP_2 \geq \cdots \geq MAPP_{NS}$ for funding sources and $EPP_1 \geq EPP_2 \geq \cdots \geq EPP_{NP}$ for projects) and arranged tip-to-tail to determine the cumulative investment cost.

The ΔI_k and ΔP_k values in the algebraic method for financial integration are synonymous with the ΔG_k and ΔW_k values for direct recycle and will be used in the cascade diagram in the same way in the interval resource balances shown in Figure 8.7. A cascade diagram can be developed using these interval balances like the one shown in Figure 8.11a. However, financial integration is performed under the restriction that a project is either completely funded or not funded at all. This restriction makes determining the pinch point in a revised cascade diagram more challenging than adding the most negative δ_k value to the δ_0 value as done in other integration methods; the pinch point is not defined by have one or more δ'_k values of zero. To determine the pinch point for the investment cost, the cumulative sum of the investment costs of the projects is added to the top of the cascade diagram (δ_0) until all δ_k values are greater than or equal to zero. This is defined as a simple optimization program in Equation 8.10. The

value at the top of the revised cascade diagram (δ'_k) and is the minimum amount of additional external finances needed to fund the projects. The δ'_K value at the bottom of the cascade diagram will represent the excess funding available from the current funding sources. The complete revised cascade diagram is shown in Figure 8.11b

$$\delta'_0 = \mathrm{IC}_{\min} = \min_{\mathrm{k} \in 1,2,\ldots,\mathrm{NP}} \sum_{\mathrm{i}=1}^{k} \mathrm{IC}_k \qquad (8.10)$$

$$\text{s.t. } \delta'_{\mathrm{k}} \geq 0 \; \forall \; \mathrm{k} \in \mathrm{K}$$

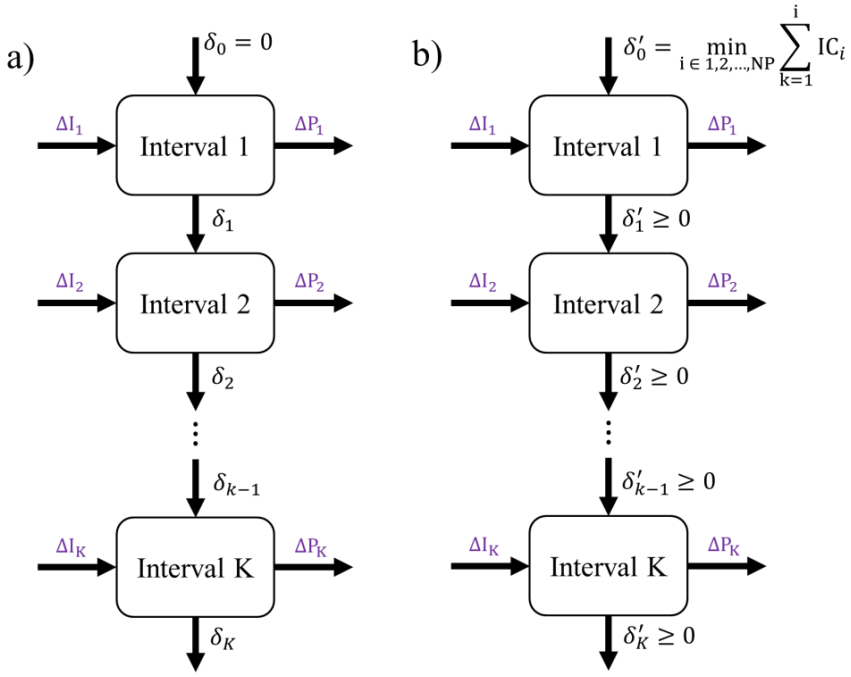

Figure 8.11. (a) Cascade and (b) revised cascade diagrams for the algebraic pinch method for financial integration. The initial cascade diagram is revised by equating δ_0 to the cumulative sum of the investment costs (IC) for the project streams and recalculating until all δ'_k values are positive. Unlike the other algebraic methods, the pinch point is not defined by δ'_k of zero due to the all-or-nothing restriction of funding capital projects. The δ'_0 value in the revised cascade diagram is the minimum investment cost that needs to be added to fully fund all projects and the δ'_K value is the excess amount of currently available funding resources.

8.4. A General Pinch Approach to Process Economics and Integration

Due to the similarities of both the graphical and algebraic methods of financial integration with the other pinch methods, we suggest their inclusion within the course materials of process economics to seamlessly bridge the gap between process economics and process integration. The traditional framework for teaching process economics and integration is illustrated in Figure 8.12a. In this *sequential* approach, the instruction is linear and may lack any feed-back necessary to solidify student understanding of both the similarities of the integration methods as well as how they relate to chemical engineering fundamentals. The inclusion of financial integration, as shown in dotted outline in Figure 8.12a, does little to help these shortcomings, since the fundamentals of process economics are coupled nicely with the financial integration method but do not exemplify the same connections with its heat and mass counterparts. While viable, this approach does not take the full advantage of the similarities between the fundamental concepts or the integration methods they utilize.

Contrary to the linear instructional framework, a *holistic* framework can be developed for the instruction of process economics and integration due to the otherwise similar nature of these strategies. This method, depicted in Figure 8.12b, stresses the interconnectivity between the integration methods and their underlying fundamentals by using the general integration strategy as the lynchpin. This framework looks to develop an understanding that allows the instructional focus to be placed on the concepts of process economics and a general pinch analysis strategy, allowing for an increased emphasis on understanding the key differences that make these analyses unique.

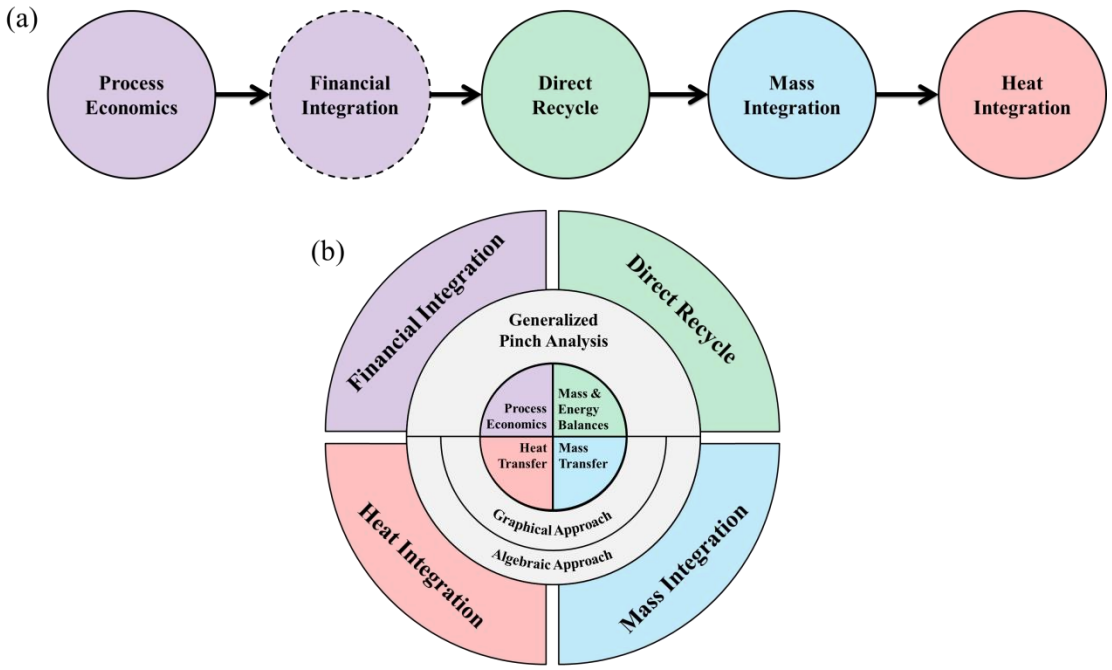

Figure 8.12. (a) Sequential and (b) holistic frameworks for the instruction of process economics and integration. In the sequential method, integration applications are taught in a specified order with no feedback or connectivity between non-sequential topics. In the holistic framework, a generalized pinch concept serves as the bridge between the innermost level of fundamental chemical engineering concepts and the outermost level of integration applications. The sequential method is not influenced by the inclusion of financial integration, whereas financial integration is required in the holistic method for completeness.

The development of a general integration methodology will bridge the gap between chemical engineering fundamentals and the applications of the integration strategies discussed. This would include the generalized algebraic (interval and cascade diagrams) methods and optimization-based methods (transshipment models). The focus of this work will be the instruction of the general graphical and algebraic pinch methods, beginning with the defining of the typical sources and sinks of each integration type. These generalized methods are applicable to all four integration strategies (mass, heat, direct recycle, and financial integration) and focuses on the identification of the pinch

point to determine the optimal opportunities for integration within a process. This analysis also results in the determination of the minimum level of external resources required in a process and the minimal waste of available process sources, leading to an increase in process sustainability.

Applications of the general integration strategy are then introduced for each specific application, including the translation method to locate the pinch point and the meaning of the resulting regions of the pinch diagram. Specific applications are financial integration, direct recycle, mass integration, and heat integration. This way, it is possible to provide a complete picture of integration strategies and how they can be used to influence decision-making, promote economic viability, and ensure sustainability during process design.

Based on our experience in the instruction of a senior-level *Process Integration, Simulation, and Economics* course at Texas A&M University, we propose the holistic framework to be included in the course following the suggested weekly itinerary shown in Table 8.1, beginning with an introduction to process economics. This would typically include four to five weeks of a general fourteen-week semester. After a sufficient treatment of process economics, an introduction of the general integration strategy would be conducted, specifically teaching the pinch analysis that is central to all four integration methods. This section of the course would constitute one or two weeks of the semester, during which a strong understanding of the method can be built. After the introduction of the general pinch method, each specific application can be the focus of the instruction. The focus on the central ideology of the pinch method allows for any of

the integration applications to be taught in any order, providing adaptability to how the course is instructed. It is suggested that each integration application be given about one week of treatment to dissect the nuances of each, specifically the use of thermodynamic equilibrium in heat and mass transfer, the appropriate translation methods, and the meaning of the regions of the successful integration pinch diagram. Any additional time in the course can be allotted to teaching a brief introduction to the optimization based methods of process integration. The holistic method outlined in this work utilizes the student's knowledge of the specific chemical engineering fundamentals in combination with the pinch strategy to bridge the gap of knowledge and increase understanding of integration. By predicating the teaching on an understanding of chemical engineering fundamentals, including process economics, teaching process integration can focus on similarities and interchangeability between the integration methodologies.

Table 8.1. Suggested incorporation of General Pinch Analysis and financial integration within a semester schedule for the holistic instruction of process economics and integration.

Weeks	Topics
1	*Introduction to Process Economics and Integration*
2-6	*Process Economics* (Capital and Operating Cost, Depreciation, Time-value of Money, Net Present Value, Payback Period, Return on Investment, etc.)
7	*Introduction to the Generalized Pinch Concept (Graphical Method)*
8	*Introduction to the Generalized Pinch Concept (Algebraic Method)*
9	*Financial Integration*
10	*Direct Recycle*
11	*Mass Integration*
12	*Heat Integration*
13-14	*Process Integration (Optimization Based Methods)*

8.5. Case Studies

This section outlines two case studies that exemplify the benefits of the proposed holistic teaching method in this work. The first case study applies the concepts of mass, heat, and financial integration to a general chemical process to demonstrate the ability to teach the applications of process integration in any order once the understanding of the generalized pinch method has been established. This case study depicts the adaptability of the holistic method for integration applications while also showing how financial integration can be incorporated with heat and mass integration to suggest priority for the integration projects. The second case study demonstrates the use of pinch methodologies for a biochemical process, giving insight into applying the use of conventional chemical engineering strategies to non-traditional processing methods. Explanations and solutions to these case studies are kept short to allow the focus to be placed on the implementation of the holistic teaching framework; additional information and full solutions can be found in the Supplemental Materials.

8.5.1. Case Study I: Chemical Processes

This illustrative example follows the structure of analyzing the opportunities of heat integration and mass integration before using financial integration to determine the economic feasibility of implementing these improvements. This case study is designed for instruction after the general pinch methodology has been established, beginning approximately Week 7 per the suggested schedule in Table 8.1. This example problem was designed to mimic the role of a term project: first the students are asked to investigate integration strategies for sustainable process improvement, and then students

are asked to determine if these proposed improvements make are feasible based on their knowledge of process economics. However, due to the necessity of teaching this problem over a multi-week span as more integration applications are covered, this problem should be used to demonstrate the merits, requirements, and expectations of a term project and not as a direct replacement. Additionally, this example problem and the solution that follows illustrates the advantages of using a holistic method in lieu of a more linear approach; by using a holistic method, the order of instruction for the integration applications does not matter, whereas instruction of an example problem of this nature would be required to be taught in the final weeks of the semester per the linear teaching method in Figure 8.12a.

More specifically, this case study applies the pinch method to the generalized chemical process shown in Figure 8.13. The process consists of two reactors, a separation network, a scrubbing unit and a flash tank. Four heat exchangers exist within the system for heating (labeled in blue) and cooling (labeled in orange), providing opportunities for heat integration. Data for these exchangers, including inlet temperatures, outlet temperatures, heat duties, and FC_P values, are shown in Table 8.2 below. A minimum driving force of 10 K is desired. Additionally, it is desired to clean the wastewater stream (labeled in green) by using the off-gas stream (labeled in red). Table 8.3 and Table 8.4 give the data for the mass integration problem, including supply and target compositions, mass flowrates, and equilibrium data. Using the data provided, it is desired to use the pinch methodology to maximize heat and mass integration using available internal resources; doing so will determine the minimum amount of heating

and cooling utilities and the minimum amount of external mass separation agent necessary for this process. Additionally, due to the limited monetary resources available to finance these projects, it is desired use financial integration methods to determine which of these projects should be prioritized for implementation.

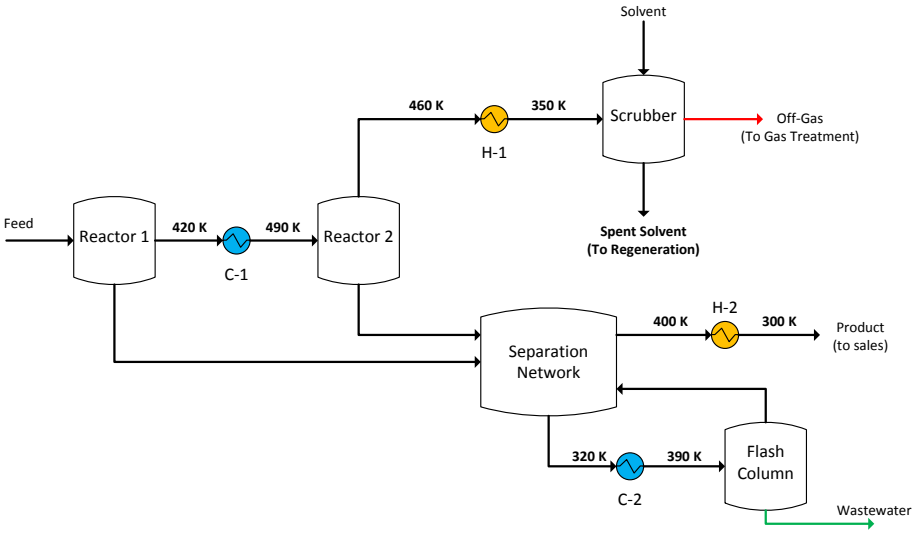

Figure 8.13. Generalized chemical process to exemplify the methods of process integration. Adapted from textbook by El-Halwagi [342].

Table 8.2. Heat exchanger data for the generalized chemical process example.

Stream	FC_P (kW/K)	Supply Temperature (K)	Target Temperature (K)	Enthalpy Change (kW)
H_1	300	460	350	-33,000
H_2	500	400	300	-50,000
C_1	600	420	490	42,000
C_2	200	320	390	14,000

Table 8.3. Data for the wastewater stream for the generalized chemical process example.

Stream	Flowrate (kg/s)	Supply Composition y_i^s (ppmw)	Target Composition y_i^t (ppmw)
Process wastewater	0.2	350	200

Table 8.4. Data for the process mass separation agent (MSA) for the generalized chemical process example.

Stream	Flowrate (kg/s)	Supply Composition x_i^s (ppmw)	Target Composition x_i^t (ppmw)	m_j	ϵ $(ppwm)$
Off Gas	0.012	300	1000	0.2	50

Figure 8.14 below shows the pinch analysis for the heat and mass integration of the generalized chemical process. In Figure 8.14a, the composite curves are determined using the temperature and enthalpy data provided in Table 8.2. The determination of the cold composite curve also considers the necessary temperature difference (ΔT_{min}) of 10 K by adding this amount to all the cold stream temperatures. Once these composite curves are plotted, the cold composite curve can be shifted vertically to locate the pinch point of the system. As described in Section 8.2.1, we see three distinct regions by locating the correct pinch point, as shown in Figure 8.14b. Regions (1) and (3) denote the minimum amount of cooling and heating utility needed for the process, respectively; the minimum utility requirements amount to 60,000 kW of heating and 33,000 kW of cooling. These minimum values are due to the integration opportunities within the process, resulting in a total amount of 46,000 kW of saved utilities as shown by region (2) in Figure 8.14b. By using pinch methodologies to minimize the excess heating and cooling in the process, the total external energy requirement can be reduced from 139,000 kW to 93,000 kW.

Similar observations can be made about the mass integration by using the same pinch methodology. In this case the rich and lean composite curves are plotted using the rich compositions y, after adjusting the lean stream compositions x to y using the

equilibrium data, and load of impurities. This is shown in Figure 8.14c. After shifting the lean composite curve vertically to find the pinch point of the system, three distinct regions can be observed in Figure 8.14d. Region (1) denotes the amount of external resources needed to remove the impurity load, in this case a load of 23 mg/s. This is reduced from the initial 30 mg/s due to the ability for the off-gas to accept 7 mg/s of impurity, as shown by region (2) in Figure 8.14d. Finally, the process mass separation agent is not fully utilized due to thermodynamic limitations, and 3 mg/s of potential is not utilized as denoted by the existence of region (3) in Figure 8.14d.

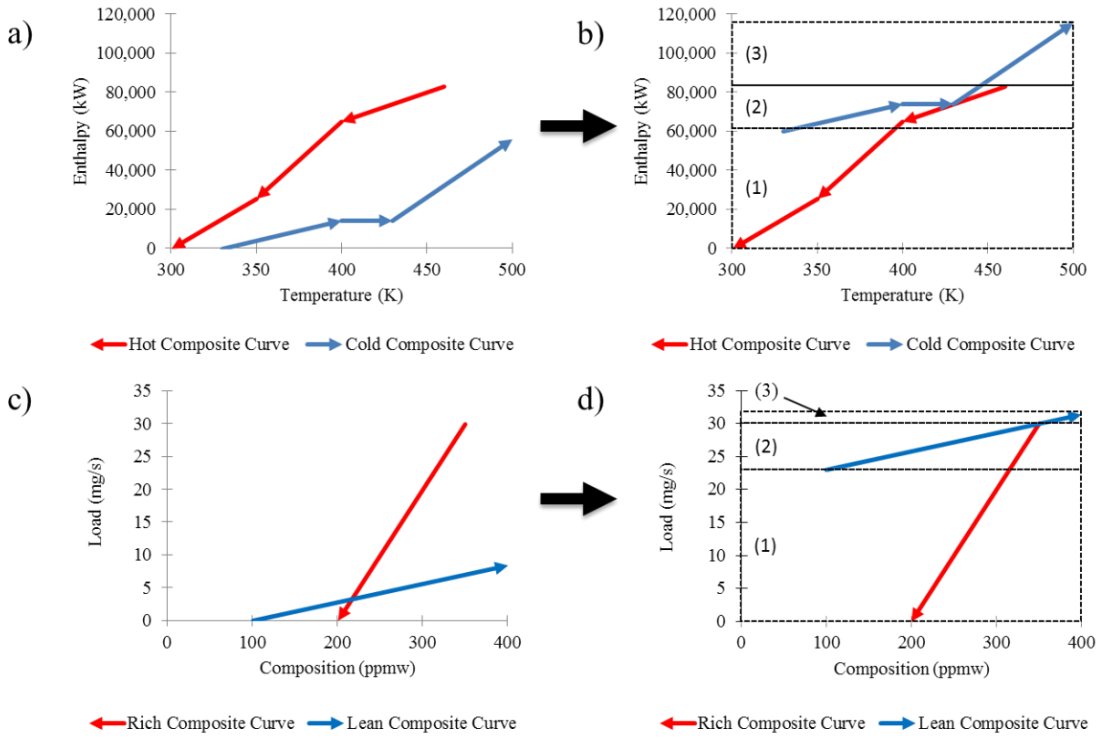

Figure 8.14. (a) Heat integration and (b) mass integration solution for the classical chemical engineering example. The resulting regions correspond to (1) external resource requirement, (2) integration of resources within the process, and (3) excess process resources.

After successfully analyzing the benefits of both heat and mass integration, the economic feasibility of implementing these strategies can be considered. In addition to the annual returns these projects provide, they are also subject to an investment cost for implementation. By using another pinch analysis, this time for financial integration, it is possible to determine which of these projects should be implemented based on the financial resources available. Equations 8.11 and 8.12 outline simple functions relating the minimum external resource requirements with the investment cost and annual returns, respectively, of the heat and mass integration projects. These equations follow the typical form of the six-tenths scaling for capital costs and linear scaling for annual returns that are demonstrated in the economics portion of the class. The values for parameters A and B in each equation are given in Table 8.5 and Table 8.6. Additionally, funding is available for these projects from two different sources, and the investment cost, annual returns, and maximum allowable payback period data are outlined in Table 8.7.

$$\begin{pmatrix}\text{Investment}\\ \text{Cost}\end{pmatrix} = \text{A} \cdot \begin{pmatrix}\text{Minimum}\\ \text{External}\\ \text{Requirement}\end{pmatrix}^{\text{B}} \tag{8.11}$$

$$\begin{pmatrix}\text{Annual}\\ \text{Returns}\end{pmatrix} = \text{A} \cdot \begin{pmatrix}\text{Minimum}\\ \text{External}\\ \text{Requirement}\end{pmatrix} + \text{B} \tag{8.12}$$

Table 8.5. Parameters for the estimation of investment cost based on external resources using Equation 8.11.

Project	A	B
Mass Integration	$250,000	0.279
Heat Integration	$100,000	0.141

Table 8.6. Parameters for the estimation of annual returns based on external resources using Equation 8.12.

Project	A	B
Mass Integration	2.0	$19,507
Heat Integration	50.0	$73,850

Table 8.7. Data for the available funding sources. MAPP is the maximum allowable payback period.

	Investment Costs	Annual Returns	MAPP (yr)
Source 1	$400,000	$20,000	20
Source 2	$750,000	$250,000	3

Using the investment costs from Equations 8.11 and 8.12 and the data for available funding sources found in Table 8.7, we can perform a financial integration analysis using the pinch method to determine which of the projects we have considered, i.e. heat and mass integration, are worth investing monetary resources into. The results for that pinch method are shown in Figure 8.15, first starting with the plotting of the composite curves in Figure 8.15a and then demonstrating the translation of the funding source curve to find the pinch point. Like heat and mass integration, three distinct regions are recognizable in the pinch analysis once the correct pinch point is determined; region 1 denotes projects that cannot be funded, region 2 shows projects for which funding is available, and region 3 shows how much of your financial resources would be in excess. In this example, funding is available for the heat integration project in region 2, but should not be used for the mass integration project in region 1 due to its longer payback period, 8 years, when compared to the heat integration project's 2.86 years.

Ultimately, it can be seen from region 3 that $0.65 million is excess funding that should not be utilized for these projects.

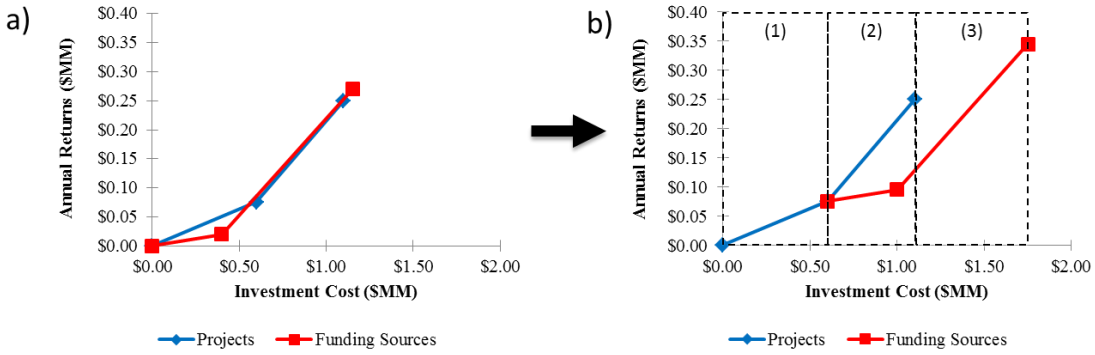

Figure 8.15. Financial integration solution for the classical chemical engineering example. The resulting regions correspond to (1) additional external funding required, (2) use of funding sources for integration projects, and (3) excess financial resources.

As expressed earlier, the development of this problem is not possible until the end of the semester based on the linear approach to teaching process economics and process integration. Even with the addition of financial integration into the curriculum, the requirement that each integration strategy be discussed in a specific order requires that all three integration strategies be discussed before this example problem can be examine, making it no more effective than a term project that relies on the same strategy. However, under the holistic approach proposed in this work, the focus on a generalized integration method means the order of instruction of each integration application does not matter. This problem can then be used through a semester to full effect as a way of serving as an illustrative example to how students should approach their term project, adding a new layer of understanding for students.

8.5.2. Case Study II: Biochemical Conversion of Corn Stover to Ethanol

While often considered a method for reducing the operating cost of a chemical process, process integration is also useful for increasing the sustainability of a process. This characteristic of process integration is of interest in biotechnology applications; these processes often involve solids handling that require the use of substantial amounts of fresh water to maintain continuous flow. One important example is the biochemical production of biofuels from lignocellulosic biomass. Though fuel-grade ethanol is currently produced using corn as a raw material, the food versus fuel debate has sparked an increased interest in its production from energy crop or agricultural waste sources [96, 298, 300, 301, 308, 311, 444]. Though not considered a main contributor to the operating cost for these lignocellulosic systems, the large water requirements for biofuel production processes puts a strain on the surrounding environment [445-447]. In these instances, integration methods can be utilized to reduce the fresh water consumption and increase the sustainability of the existing facility. While financial factors will always be considered, sustainability will ultimately play a significant role in determining the feasibility of biofuels as a main alternative energy source.

The example problem discussed in this section examines the use of the direct recycle methodology for water recycle in the biochemical production of ethanol from pretreated corn stover. The simplified process flow diagram, shown in Figure 8.16, consists of three reactors in series, a distillation column, and a molecular sieve. The first two reactors, responsible for biomass pretreatment and enzyme hydrolysis to liberate the sugar monomers from the solid corn stover, require a specific solid loading achieved by

the addition of water. These units are considered the process sinks. After the fermentation step, the distillation column and molecular sieve are used to concentrate the ethanol product, resulting in two reject streams that contain mostly water and can be used as process sources for recycle. The data for the process sinks and sources are shown in Table 8.8 and

Table **8.9**, respectively. In this example, it is desired to know what the minimum fresh water usage is after applying direct recycle to the biofuels process.

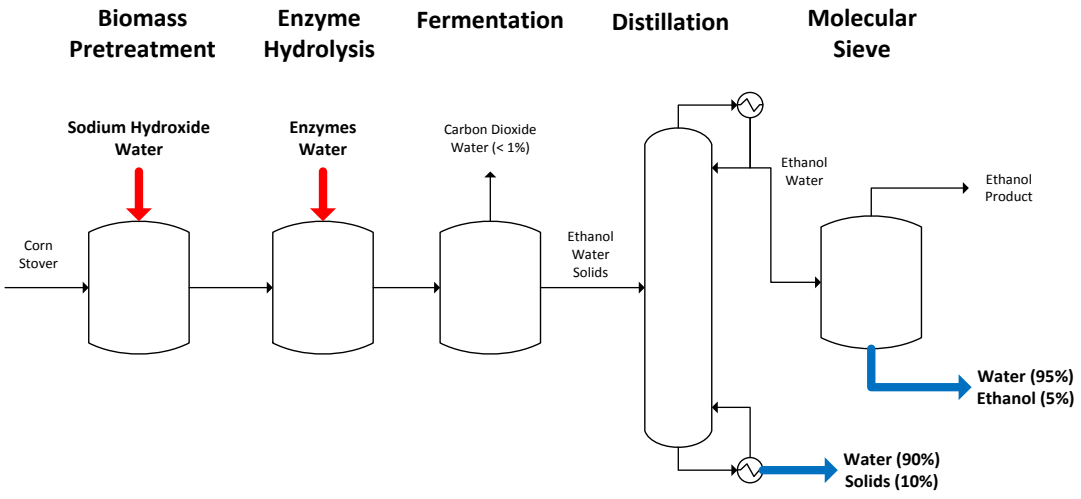

Figure 8.16. Simplified process flow diagram for the biochemical production of ethanol from corn stover. Process sources for the direct recycle are shown in blue, and process sinks are shown in red.

Table 8.8. Data for the water sources for the biomass to ethanol example.

Stream	Flowrate (kg/hr)	Impurity Concentration (wt%)	Load (kg/hr)
Molecular Sieve	500	5	25
Distillation Column	5500	11	605

Table 8.9. Data for the water sinks for the biomass to ethanol example.

Stream	Flowrate (kg/hr)	Impurity Concentration (wt%)	Load (kg/hr)
Pretreatment Reactor	3000	2	210
Hydrolysis Reactor	4000	7	80

The generalized algebraic pinch method is utilized for the direct recycle of water, with the load-interval diagram (LID) shown in Figure 8.17. The integration of two process sources and two process sinks results in a total of four intervals where integration can occur. The source (ΔW_k) and sink flows (ΔG_k) are calculated for each interval and used in the cascade diagram in Figure 8.18a. The pinch point is determined by the most negative δ_k value and is in red between intervals 3 and 4 in Figure 8.18a. This value is added to the top of the revised cascade diagram in Figure 8.18b to determine the minimum resource requirements of the biochemical process, which are marked in green. Based on the operating parameters of the bioethanol production plant, the minimum amount of external fresh water that would need to be supplied to the biochemical process is given as the δ_0' value in the revised cascade diagram and is 4090.9 kg/hr. A total of 3,090.9 kg/hr of process water cannot be recycled, the value which is denoted by δ_4', and will need to be sent to wastewater treatment. Overall, process integration results in a 50% reduction in wastewater in the biochemical process.

Interval	Load 0.0	ΔM_k	Sources	Source Flow per Interval (ΔW_k)	Sinks	Sink Flow per Interval (ΔG_k)
1	500	500	Source 1	25	Sink 1	10
2	4000	3500	Source 2	385		70
3	6000	2000		220	Sink 2	140
4	7000	1000				70

Figure 8.17. Load-interval diagram (LID) for the direct recycle of water in the biochemical engineering example. The integration of the two sources and two sinks results in four intervals.

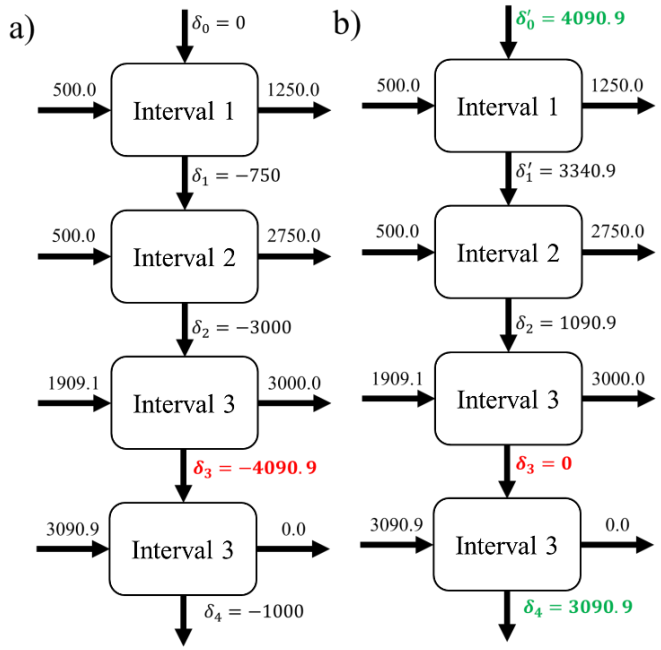

Figure 8.18. (a) Cascade diagram and (b) revised cascade diagram for the direct recycle of water in the biochemical engineering example. The pinch point is indicated in red between intervals 3 and 4 and the external fresh water requirement (δ'_0) and excess internal water resources (δ'_4) are denoted in green.

After determining the base case scenario, two water treatment methods are proposed to improve the effectiveness of a direct recycle strategy: enzyme digestion and filtering. These purification techniques will be used to reduce the impurity level of the

distillation bottoms stream from 10% impurities to 6% for enzyme digestion and 8% for filtration; these data are also shown in Table 8.10. Direct recycle techniques can again be used to understand the benefits of adding additional purification techniques to the bioethanol process. Each technology reduces the impurity content of the bottoms stream of the distillation column, reducing this concentration to levels closer to the maximum allowable concentrations of the process sinks. This, in turn, will increase the maximum water recycle and decrease the amount of fresh water needed and wastewater sent to treatment. Figure 8.19 shows the results of direct recycle algebraic pinch point analysis for both (a) enzyme digestion and (b) filtration. The application of both methods results in a reduction of fresh water usage when compared to the base case scenario in Figure 8.17, from 4,090.9 kg/hr to 2,583.3 kg/hr for enzyme digestion and 3,187.5 kg/hr for filtration. While these results suggest that enzyme digestion should be implemented over filtration to achieve a larger level of integration, each technology is associated with a capital cost that needs to be considered. These capital costs, including the investment costs, annual returns based on the freshwater savings, and expected payback period (EPP) are shown in Table 8.10.

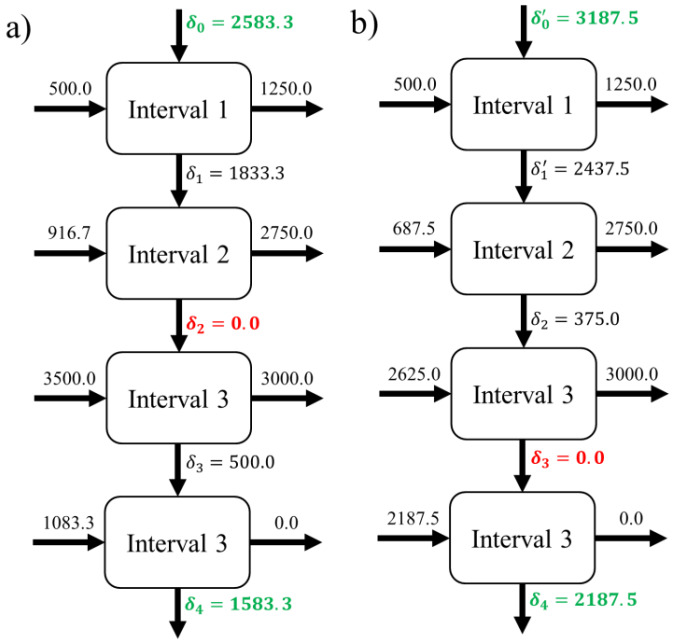

Figure 8.19. Revised cascade diagrams for the improved direct recycle solution with the implementation of (a) enzyme digestion and (b) filtering to improve the impurity concentration of the distillation column bottoms. The impurity reduction reduces the necessary fresh water usage from 4,090.9 kg/hr in the base case to 2,583.3 kg/hr and 3,187.5 kg/hr for the digestion and filtration, respectively.

Table 8.10. Cost data for the available wastewater treatment methods. EPP is the expected payback period.

Unit	New Distillation Waste Impurity Concentration (wt%)	Investment Cost	Annual Returns	EPP (yr)
Enzyme Digestion	6%	$250,000	$140,000	1.79
Filtering	8%	$500,000	$80,000	6.25

Table 8.11. Funding data available to be applied to wastewater treatment in the biomass to ethanol example. MAPP is the maximum allowable payback period.

	Investment Costs	Annual Returns	MAPP (yr)
Source 1	$100,000	$100,000	1
Source 2	$300,000	$120,000	2.5

Two funding sources are available for the implementation of these new water treatment methods; the investment costs, annual returns, and maximum allowable payback period (MAPP) details for each of these sources are shown in Table 8.11. Using this data with the project data in Table 8.10, the new algebraic method for financial integration is used to determine which treatment strategy is financially feasible to implement to improve the efficacy of the direct recycle implementation and further reduce the demand on natural resources. The implementation of the fiscal-interval diagram is shown in Figure 8.20. The two funding sources and two projects result in three intervals. Within these intervals, the values of the investment cost of the funding sources (ΔI_k) and investment costs of the projects (ΔP_k) can be calculated. These values are used to develop the cascade diagram in Figure 8.21a for the financial integration problem. The pinch point can be determined using a guess-and-check methodology beginning with adding the investment cost of the filtering project to δ_0 and checking to see if all δ_k' values are greater than zero. If they are not all zero, the investment cost of the enzyme digestion project is also added to δ_0 and the δ_k' values are checked again. Alternatively, Equation 8.10 can be solved for the number of projects *NP* that need to be added before all the δ_k' values are greater than zero. Either method will result in a solution of only the investment cost of the filtering, a value of \$0.50 MM, being needed, as shown in green at the top of the revised cascade diagram in Figure 8.21b. This results in the value of δ_3', and therefore the amount excess funding, being \$0.15 MM as shown in green at the bottom of the cascade diagram. This result is a result of a complete funding of the enzyme digestion project that was also selected when performing the

direct recycle analysis. The filtration project cannot be funded due to a restrictively high payback period and requires external outside funding to finance.

Interval	Returns 0.0	ΔF_k	Funding Sources	Investment Cost per Interval (ΔI_k)	Projects	Investment Cost per Interval (ΔP_k)
1	0.08	0.08	Source 1	0.20	Filtering	0.50
2	0.12	0.04		0.10		0.07
3	0.22	0.10	Source 2	0.10	Enzyme Digestion	0.18

Figure 8.20. Fiscal-interval diagram (FID) for the biochemical engineering example. The integration of the two funding sources and two projects results in three intervals.

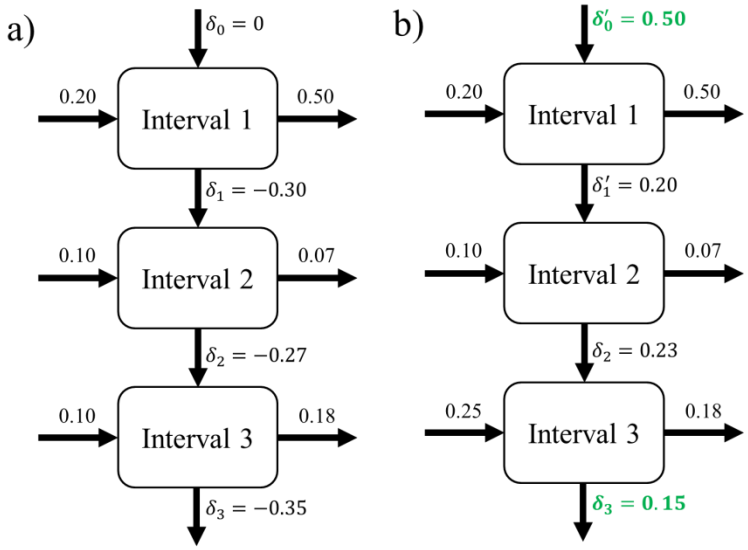

Figure 8.21. (a) Cascade diagram and (b) revised cascade diagram for the direct recycle of water in the biochemical engineering example. The pinch point is found by adding the investment cost of the filtering project (IC_1) to the top of the cascade diagram (δ_0). The external investment cost requirement (δ_0') and excess funding resources (δ_4') are denoted in green.

This example problem has exemplified the use of the algebraic pinch technique for both direct recycle and financial integration in biochemical processing problems. By using a generalized framework for teaching pinch-based integration strategies, it is

possible to focus on the similarities and implications of the financial and recycle applications of the framework. This problem specifically has placed the focus on understanding the implications of direct recycle, how the method can be used to evaluate the benefits of adding of new process units, and how financial integration can be another tool for evaluating processing options and aiding decision making. It also exemplifies how these methods can be used for bioprocessing as a way of improving both profitability and sustainability.

8.6. Conclusions

This chapter develops a holistic framework for the instruction of process economics and integration strategies in chemical engineering. The strategy of this framework is to use general pinch analysis as the lynchpin between fundamental chemical engineering concepts, i.e. heat and mass balances, heat transfer, mass transfer, and process economics, and their respective integration applications. This framework is more advantageous than the currently employed linear framework for teaching process integration that is sequential in nature and lacks the ability to accentuate the similarities between the applications of process integration. This holistic framework is also readily applicable to the algebraic and optimization methods of process integration. The framework developed in this work was exemplified with two case studies. The first case study investigated the cost savings via heat and mass integration for a traditional chemical engineering process and evaluated using financial integration the economic feasibility of implementing these strategies. The second case study utilized direct recycle for a biomass-to-bioethanol production facility to minimize the need for fresh process

water. Financial integration was used to evaluate the feasibility of two process technologies that when implemented increased the ability for water recycle and increased sustainability. Overall, this holistic framework outlines an improved approach for teaching process economics and integration that introduces a generalized approach that relies on the similarities of each integration application to shift the focus of teaching to their implications in process design.

CHAPTER IX

CONCLUSIONS AND FUTURE WORK

9.1. Summary of thesis

In this thesis, continuous methods for the biochemical production of fuels and pharmaceuticals are explored by coupling process systems engineering techniques with biochemical engineering applications. Process modeling strategies are used to gain insight into the biochemical kinetics of products in the metabolic mevalonate pathway. Optimal control methods are applied to a novel, multi-feed continuous bioprocessing paradigm to demonstrate an improved productivity and profitability over batch production methods. A process synthesis framework is developed to determine the optimal process economics of bioethanol and biobutanol production from consolidated bioprocessing and acetone-butanol-ethanol processing strategies, respectively. More detailed results of these studies discussed in Chapters IV-VIII are summarized as follows.

9.1.1. Kinetic Modeling of β-carotene from recombinant *S. cerevisiae* SM14

Reliable mathematical models for batch cultivation are developed to describe the glucose consumption, product formation and depletion, and β-carotene production of the Saccharomyces cerevisiae strain mutant SM14 with 20 g/L glucose as the carbon source. Parameter estimation for the models employ an objective function minimizing an expression that uses the coefficient of determination R^2, which avoids the necessity of weighting or normalizing the equations. Comparison with experimental data shows that

the developed models show a satisfactory prediction of the overflow metabolism that is happening in the cell. Additionally, local and global sensitivity analysis of the models with respect to the optimal parameters is also studied and further show that the models developed here accurately describe trends in the dynamic states of the bioreactor during batch fermentation.

9.1.2. Continuous pharmaceutical development via a multi-feed bioreactor

This chapter proposes a novel multi-feed bioreactor system composed of multiple independently controlled feeds for substrate(s) and media used as a diluent and allows for the free manipulation of the bioreactor dilution rate and substrate concentrations to maximize bioreactor productivity and substrate utilization while reducing operating costs. These optimal flow rates of the bioreactor are determined through the solution to an optimal control problem where the kinetic models describing the time-variant system states are used as constraints. This new bioreactor paradigm is exemplified through the intracellular production of beta-carotene using a three-feed bioreactor consisting of separate glucose, ethanol and media feeds. The traditional single feed process is compared to a two and three feed process featuring glucose, ethanol and media as feed options. Results show up to a 30% reduction in the productivity with the addition of multiple feeds, though all three systems show an improvement in productivity when compared to batch production. Economic viability is considerably improved through this new paradigm as the addition of multiple feeds results in a 45% and 30% reduction of the breakeven product selling price for the dual feed and three feed bioreactors, respectively, when compared the single feed counterpart.

9.1.3. Process synthesis for an optimal lignocellulosic ethanol refinery

The large-scale production of ethanol via a method of converting lignocellulosic material without the need for enzyme hydrolysis, termed consolidated bioprocessing (CBP), is analyzed and compared to traditional separate hydrolysis and fermentation (SHF) methods to determine the profitability of biochemical ethanol production. Optimal process topologies for both the CBP and SHF technologies are selected from superstructures considering multiple biomass feeds, chosen from those available across the United States, and multiple prospective pretreatment technologies. Steam cogeneration is also considered to produce the necessary hot utilities as well as electricity as a salable byproduct. Results show the optimal consolidated bioprocessing plant produces about 66 million gallons of ethanol to be sold at a breakeven price of $1.31 per gallon ($0.35 per liter) from a pure sugarcane bagasse feed and sodium hydroxide pretreatment method in addition to 67 MW of electricity per year. These results compare favorably to SHF production methods, showing a 20% decrease in the minimum ethanol selling price (MESP) and 197.5% increase in electricity generation despite a 13.3% drop in ethanol production. A long term economic evaluation is performed to determine an attractively profitable selling price for the ethanol produced. Additional analysis is done on the effect of a mixed biomass feed on process profitability and impact on biomass harvesting radius.

9.1.4. Continuous biobutanol production using an intensified, multi-feed bioreactor

The multi-feed bioreactor developed in Chapter V is intensified with gas stripping to produce biobutanol via acetone-butanol-ethanol fermentation. Application of

an optimal control framework to the reactor, either with or without the presence of gas stripping, show that stripping increases glucose conversion to a level necessary for ABE production to be economically profitable. When implemented in a complete ABE process using lignocellulosic biomass as a carbon source, the single multi-feed reactor does not produce enough ABE to be economically competitive with current fossil fuel technologies. However, the addition of more bioreactors allows for the benefits of economy-of-scale, reducing the butanol price to guarantee profitability and increasing the viability of continuous biobutanol production.

9.1.5. Holistic pedagogy for the teaching of process integration

We explore the similarities between pinch-based mass, heat and financial integration methods for chemical and biochemical process industries, and propose a holistic approach to teaching these methods using a 'generalized pinch' framework. This holistic approach introduces the idea of using a graphical approach and a novel algebraic pinch method for financial integration into the semester curriculum to close the loop between process economics and process integration applications. In a classroom setting, the framework allows the instructor to teach the fundamental aspects of general integration at the onset, and then use different mass, heat and financial integration problems as applications, as the course progresses. To this end, the generalized pinch serves as a lynchpin between different process integration and economic analyses. This is exemplified using two case studies, first for a traditional chemical engineering process and then in the emerging field of biotechnology for biofuels production.

9.2. Conclusions

The purpose of this research was to investigate novel methods for the continuous, sustainable production of fuels and chemical using the marriage of process systems engineering and biochemical engineering disciplines. From the results of this work, we can conclude that:

1. The development of reaction kinetics for biochemical pathways is vital to the development and understanding of biochemical production processes. Process modeling and optimization strategies can be used to determine the optimal kinetic parameters that describe the production of products in the mevalonate pathway, such as β-carotene, bisabolene, and taxadiene, with the accuracy necessary for future process development purposes.
2. The multi-feed reactor capable of independently manipulating the dilution rate and inlet glucose concentration of a continuous bioreactor provides the adaptability necessary to improve process productivity and profitability. The use of optimal control method to determine the optimal control policy for the manipulated variables stimulates and facilitates the transition from traditional batch processing to more economically efficient continuous strategies.
3. The use of consolidated bioprocessing is shown to be more economically attractive that separate hydrolysis and fermentation to produce bioethanol. By eliminating the need for enzymatic hydrolysis in favor of the direct conversion of cellulose to ethanol, a significant improvement can be made to the

economic feasibility of renewable fuel production from lignocellulosic biomass.

4. Process intensification of the multi-feed bioreactor with *in situ* gas stripping is necessary for the economical production of biobutanol. The production of biobutanol from lignocellulosic biomass using a single intensified bioreactor suffers from a low productivity that leads to a high economic butanol sale price. However, the biobutanol process is aided with the inclusion of more bioreactors, achieving an economy-of-scale that increases the profitability exponentially with a linear increase in butanol production.

9.3. Future Research Directions

In Chapter IV of this thesis we describe the modeling and parameter estimation for β-carotene from recombinant *S. cerevisiae*. While the models described are specific to the β-carotene production process, the parameter estimation techniques defined can be applied to many other products both in the mevalonate pathway and from other metabolic pathways. Through additional genetic modifications, the kinetic models can be used to describe other metabolic products, including bisabolene and taxadiene. Additionally, the inclusion of other dependencies can be explored to better predict the dynamics of cellular metabolism from online measurements including dissolved oxygen content, off-gas oxygen and carbon dioxide measurements, temperature, and acid and base addition.

With more accurate models, the design of model predictive controllers can be performed. These controllers combine the predictive nature of state observers with the

optimal control framework discussed in Chapter V for the multi-feed bioreactor. By integrating the prediction and control capabilities with real time data acquisition the bioreactor can be controlled and optimized in real time toward the maximum possible productivity. The model predictive controller can also be coupled with state-of-the-art fault detection methods for contamination detection to mitigate the concerns over continuous bioprocessing.

In addition to focus placed on the bioreactor, the impact of the multi-feed paradigm on the downstream separation processes should be explored. The cyclic nature of the optimal control policies seen in Chapters VI and VII will result in a cyclic nature in the outlet reactor effluent. Most separation systems operate with an assumed constant flowrate from the upstream process, and the variable flowrate exiting the reactor will put additional processing constraints on the operation and control of the process downstream. This is especially true for intracellular products, which require additional downstream separation for cell harvesting and disruption.

Finally, it is worth mentioning that the methods described in this work is generic and can be used for the investigation of the continuous production of a variety of other biochemical products. All process systems engineering strategies used in this work only rely on knowledge of the biochemical kinetics of the reaction of interest. With this knowledge, the models used in the optimal control and process synthesis frameworks can be updated and utilized for any system of interest.

REFERENCES

1. Grossmann, I.E. and A.W. Westerberg, Research challenges in Process Systems Engineering. Aiche Journal, 2000. 46(9): p. 1700-1703.

2. Douglas, J.M., A Hierarchical Decision Procedure for Process Synthesis. Aiche Journal, 1985. 31(3): p. 353-362.

3. Eiben, A.E., P.E. Raue, and Z. Ruttkay, Genetic algorithms with multi-parent recombination. Parallel Problem Solving from Nature - Ppsn Iii - International Conference on Evolutionary Computation, Proceedings, 1994. 866: p. 78-87.

4. Khachaturyan, A.G., S.V. Semenovskaya, and B.K. Vainshtein, Statistical Thermodynamical Approach to the Problem of Determination of Phases of Structure Amplitudes. Kristallografiya, 1978. 24(5): p. 905-916.

5. Spall, J.C., Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control. 2003, Hoboken, NJ: Wiley.

6. Griva, I., S.G. Nash, and A. Sofer, Linear and Nonlinear Optimization. 2nd ed. 2009, Philadelphia, PA: Society for Industrial and Applied Mathematics.

7. Cauchy, A., Méthode générale pour la résolution des systemes d'équations simultanées. Comp. Rend. Sci. Paris, 1847. 25: p. 536–538.

8. Fletcher, R. and M.J.D. Powell, A Rapidly Convergent Descent Method for Minimization. Computer Journal, 1963. 6(2): p. 163-&.

9. Shanno, D.F., Conditioning of quasi-Newton methods for function minimization. Mathematics of computation, 1970. 24: p. 647–656.

10. Dennis, J., J. and J. Moré, Quasi-Newton Methods, Motivation and Theory. SIAM Review, 1977. 19: p. 46-89.

11. Byrd, R.H., J. Nocedal, and Y.X. Yuan, Global Convergence of a Class of Quasi-Newton Methods on Convex Problems. Siam Journal on Numerical Analysis, 1987. 24(5): p. 1171-1190.

12. Martínez, J.M., Practical quasi-Newton methods for solving nonlinear systems. Journal of Computational and Applied Mathematics, 2000. 124: p. 97-121.

13. Kolda, T.G., R.M. Lewis, and V. Torczon, Optimization by direct search: New perspectives on some classical and modern methods. Siam Review, 2003. 45(3): p. 385-482.

14. Nelder, J.A. and R. Mead, A Simplex-Method for Function Minimization. Computer Journal, 1965. 7(4): p. 308-313.

15. Felippa, C.A., Variational-Methods for the Solution of Problems of Equilibrium and Vibrations (Reprinted from Bull Am Math Soc, Vol 49, Pg 1, 1943). International Journal for Numerical Methods in Engineering, 1994. 37(13): p. 2159-&.

16. Frisch, K., The logarithmic potential method of convex programming. 1955, University Institute of Economics: Oslo, Norway.

17. Fiacco, A. and G. McCormick, Nonlinear programming: Sequential Unconstrained Minimization Techniques. 1968, New York, New York: John Wiley & Sons.

18. Lagrange, J.L., A new method for determining the maxima and minima of the infinite integral formulas [Essai sur une nouvelle methode pour determiner les maxima et minima des formules integrales indefinies], in Complete Work of Joseph Louis de Lagrange. 1761, Gallica. p. 319-332.

19. Karush, W., Minima of functions of several variables with inequalities as side conditions, in Mathematics. 1939, University of Chicago.

20. Kuhn, H.W. and A.W. Tucker, Nonlinear Programming. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. J. Neyman. 1951, Berkeley, CA: University of California Press. 481-492.

21. Floudas, C.A., Nonlinear and mixed-integer optimization: Fundamentals and applications. 1995, New York, NY: Oxford University Press.

22. Gomory, R., Integer Programming. Communications of the Acm, 1960. 3(7): p. 395-395.

23. Gomory, R.E., Outline of an Algorithm for Integer Solutions to Linear Programs. Bulletin of the American Mathematical Society, 1958. 64(5): p. 275-278.

24. Vanroy, T.J. and L.A. Wolsey, Valid Inequalities for Mixed 0-1 Programs. Discrete Applied Mathematics, 1986. 14(2): p. 199-213.

25. Vanroy, T.J. and L.A. Wolsey, Solving Mixed Integer Programming-Problems Using Automatic Reformulation. Operations Research, 1987. 35(1): p. 45-57.

26. Benders, J.F., Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik, 1962. 4(1): p. 238-252.

27. Fisher, M.L., The Lagrangian-Relaxation Method for Solving Integer Programming-Problems. Management Science, 1981. 27(1): p. 1-18.

28. Magnanti, T.L. and R.T. Wong, Accelerating Benders Decomposition: Algorithmic Enhancements and Model Selection Criteria. Operations Research Letters, 1979. 29(3): p. 464.

29. Balas, E., Disjunctive programming: Properties of the convex hull of feasible points. Discrete Applied Mathematics, 1998. 89(1-3): p. 3-44.

30. Balas, E., Nonconvex Quadratic Programming Via Generalized Polars. Siam Journal on Applied Mathematics, 1975. 28(2): p. 335-349.

31. Hooker, J.N., Resolution Vs Cutting Plane Solution of Inference Problems - Some Computational Experience. Operations Research Letters, 1988. 7(1): p. 1-7.

32. Raman, R. and I.E. Grossmann, Integration of Logic and Heuristic Knowledge in Minlp Optimization for Process Synthesis. Computers & Chemical Engineering, 1992. 16(3): p. 155-171.

33. Land, A.H. and A.G. Doig, An Automatic Method of Solving Discrete Programming-Problems. Econometrica, 1960. 28(3): p. 497-520.

34. Dakin, R.J., A Tree-Search Algorithm for Mixed Integer Programming-Problems. Computer Journal, 1965. 8(3): p. 250-253.

35. Beale, E.M.L. and J.J.H. Forrest, Global Optimization Using Special Ordered Sets. Mathematical Programming, 1976. 10(1): p. 52-69.

36. Duran, M.A. and I.E. Grossmann, An Outer-Approximation Algorithm for a Class of Mixed-Integer Nonlinear Programs. Mathematical Programming, 1986. 36(3): p. 307-339.

37. Varvarezos, D.K., I.E. Grossmann, and L.T. Biegler, An Outer-Approximation Method for Multiperiod Design Optimization. Industrial & Engineering Chemistry Research, 1992. 31(6): p. 1466-1477.

38. Rahmaniani, R., et al., The Benders decomposition algorithm: A literature review. European Journal of Operational Research, 2017. 259(3): p. 801-817.

39. Geoffrion, A.M., Generalized Benders Decomposition. Journal of Optimization Theory and Applications, 1972. 10(4): p. 237.

40. Bertolazzi, E., F. Biral, and M. Da Lio, Symbolic–Numeric Indirect Method for Solving Optimal Control Problems for Large Multibody Systems. Multibody System Dynamics, 2005. 13(2): p. 233-252.

41. von Stryk, O. and R. Bulirsch, Direct and indirect methods for trajectory optimization. Annals of Operations Research, 1992. 37(1): p. 357-373.

42. Hartl, R.F., S.P. Sethi, and R.G. Vickson, A Survey of the Maximum Principles for Optimal Control Problems with State Constraints. SIAM Review, 1995. 37(2): p. 181-218.

43. Hager, W.W., Runge-Kutta methods in optimal control and the transformed adjoint system. Numerische Mathematik, 2000. 87(2): p. 247-282.

44. Dontchev, A.L. and W.W. Hager, The Euler approximation in state constrained optimal control. Mathematics of Computation, 2001. 70(233): p. 173-203.

45. Rao, A.V., A Survey of Numerical Methods for Optimal Control. Astrodynamics 2009, Vol 135, Pts 1-3, 2010. 135: p. 497-528.

46. Grimm, W. and A. Markl, Adjoint estimation from a direct multiple shooting method. Journal of Optimization Theory and Applications, 1997. 92(2): p. 263-283.

47. Logsdon, J.S. and L.T. Biegler, Accurate Solution of Differential Algebraic Optimization Problems. Industrial & Engineering Chemistry Research, 1989. 28(11): p. 1628-1639.

48. Schwartz, A. and E. Polak, Consistent approximations for optimal control problems based on Runge-Kutta integration. Siam Journal on Control and Optimization, 1996. 34(4): p. 1235-1269.

49. Reddien, G.W., Collocation at Gauss Points as a Discretization in Optimal-Control. Siam Journal on Control and Optimization, 1979. 17(2): p. 298-306.

50. Enright, P.J. and B.A. Conway, Discrete Approximations to Optimal Trajectories Using Direct Transcription and Nonlinear-Programming. Journal of Guidance Control and Dynamics, 1992. 15(4): p. 994-1002.

51. Herman, A.L. and B.A. Conway, Direct optimization using collocation based on high-order Gauss-Lobatto quadrature rules. Journal of Guidance Control and Dynamics, 1996. 19(3): p. 592-599.

52. Cuthrell, J.E. and L.T. Biegler, Simultaneous-Optimization and Solution Methods for Batch Reactor Control Profiles. Computers & Chemical Engineering, 1989. 13(1-2): p. 49-62.

53. Cuthrell, J.E. and L.T. Biegler, On the Optimization of Differential-Algebraic Process Systems. Aiche Journal, 1987. 33(8): p. 1257-1270.

54. Williams, P., Jacobi pseudospectral method for solving optimal control problems. Journal of Guidance Control and Dynamics, 2004. 27(2): p. 293-297.

55. Elnagar, G.N. and M. Razzaghi, A collocation-type method for linear quadratic optimal control problems. Optimal Control Applications & Methods, 1997. 18(3): p. 227-235.

56. Elnagar, G., M.A. Kazemi, and M. Razzaghi, The Pseudospectral Legendre Method for Discretizing Optimal-Control Problems. Ieee Transactions on Automatic Control, 1995. 40(10): p. 1793-1796.

57. Benson, D.A., et al., Direct trajectory optimization and costate estimation via an orthogonal collocation method. Journal of Guidance Control and Dynamics, 2006. 29(6): p. 1435-1440.

58. Garg, D., et al., Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a Radau pseudospectral method. Computational Optimization and Applications, 2011. 49(2): p. 335-358.

59. Fahroo, F. and I.M. Ross, Pseudospectral methods for infinite-horizon optimal control problems. Journal of Guidance Control and Dynamics, 2008. 31(4): p. 927-936.

60. Mccormick, G.P., Computability of Global Solutions to Factorable Nonconvex Programs .1. Convex Underestimating Problems. Mathematical Programming, 1976. 10(2): p. 147-175.

61. Kesavan, P., et al., Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Mathematical Programming, 2004. 100(3): p. 517-535.

62. Adjiman, C.S., I.P. Androulakis, and C.A. Floudas, Global optimization of mixed-integer nonlinear problems. Aiche Journal, 2000. 46(9): p. 1769-1797.

63. Androulakis, I.P., C.D. Maranas, and C.A. Floudas, alpha BB: A global optimization method for general constrained nonconvex problems. Journal of Global Optimization, 1995. 7(4): p. 337-363.

64. Scott, J.K., M.D. Stuber, and P.I. Barton, Generalized McCormick relaxations. Journal of Global Optimization, 2011. 51(4): p. 569-606.

65. Tawarmalani, M. and N.V. Sahinidis, A polyhedral branch-and-cut approach to global optimization. Mathematical Programming, 2005. 103(2): p. 225-249.

66. Misener, R. and C.A. Floudas, A Framework for Globally Optimizing Mixed-Integer Signomial Programs. Journal of Optimization Theory and Applications, 2014. 161(3): p. 905-932.

67. Gounaris, C.E., R. Misener, and C.A. Floudas, Computational Comparison of Piecewise-Linear Relaxations for Pooling Problems. Industrial & Engineering Chemistry Research, 2009. 48(12): p. 5742-5766.

68. Misener, R. and C.A. Floudas, ANTIGONE: Algorithms for coNTinuous/Integer Global Optimization of Nonlinear Equations. Journal of Global Optimization, 2014. 59(2-3): p. 503-526.

69. Harrell, F.E., Jr., et al., Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treat Rep, 1985. 69(10): p. 1071-77.

70. Babyak, M.A., What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosomatic medicine, 2004. 66(3): p. 411-421.

71. Saltelli, A., et al., Global Sensitivity Analysis. 2008, West Sussex, UK: John Wiley & Sons Ltd.

72. Varma, A., M. Morbidelli, and H. Wu, Parametric Sensitivity in Chemical Systems. 1999, Cambridge, UK: Cambridge University Press.

73. Sobol, I.M., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation, 2001. 55(1-3): p. 271-280.

74. Solomatine, D., L.M. See, and R.J. Abrahart, Data-Driven Modelling: Concepts, Approaches and Experiences, in Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications, R.J. Abrahart, L.M. See, and D.P. Solomatine, Editors. 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 17-30.

75. Hodge, D., L. Simon, and M.N. Karim, Data driven approaches to modeling and analysis of bioprocesses: Some industrial examples. Proceedings of the 2003 American Control Conference, Vols 1-6, 2003: p. 2062-2076.

76. Eikens, B., M.N. Karim, and L. Simon, Combining neural networks and first principle models for bioprocess modeling. Application of Neural Networks and Other Learning Technologies in Process Engineering, 2001: p. 121-148.

77. Eikens, B., M.N. Karim, and L. Simon, Process identification with self-organizing networks. Application of Neural Networks and Other Learning Technologies in Process Engineering, 2001: p. 49-75.

78. Eikens, B. and M.N. Karim, Process identification with multiple neural network models. International Journal of Control, 1999. 72(7-8): p. 576-590.

79. Simon, L., M.N. Karim, and A. Schreiweis, Prediction and classification of different phases in a fermentation using neural networks. Biotechnology Techniques, 1998. 12(4): p. 301-304.

80. Karim, M.N., et al., Global and local neural network models in biotechnology: Application to different cultivation processes. Journal of Fermentation and Bioengineering, 1997. 83(1): p. 1-11.

81. Eikens, B. and M.N. Karim, Identification of a Waste-Water Neutralization Process Using Neural Networks. World Congress on Neural Networks-San Diego - 1994 International Neural Network Society Annual Meeting, Vol 1, 1994: p. A185-A190.

82. Karim, M.N. and S.L. Rivera, Comparison of Feedforward and Recurrent Neural Networks for Bioprocess State Estimation. Computers & Chemical Engineering, 1992. 16: p. S369-S377.

83. Rivera, S.L. and M.N. Karim, On-Line Estimation of Bioreactors Using Recurrent Neural Networks. Modeling and Control of Biotechnical Processes 1992, 1992. 1992(10): p. 159-162.

84. Karim, M.N., D. Hodge, and L. Simon, Data-based modeling and analysis of bioprocesses: Some real experiences. Biotechnology Progress, 2003. 19(5): p. 1591-1605.

85. Cheng, C. and M.S. Chiu, A new data-based methodology for nonlinear process modeling. Chemical Engineering Science, 2004. 59(13): p. 2801-2810.

86. Yee, T.F. and I.E. Grossmann, Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis. Computers & Chemical Engineering, 1990. 14(10): p. 1165-1184.

87. Floudas, C.A., A.R. Ciric, and I.E. Grossmann, Automatic Synthesis of Optimum Heat-Exchanger Network Configurations. Aiche Journal, 1986. 32(2): p. 276-290.

88. Papoulias, S.A. and I.E. Grossmann, A structural optimization approach in process synthesis—II. Computers & Chemical Engineering, 1983. 7(6): p. 707-721.

89. Aggarwal, A. and C.A. Floudas, Synthesis of General Distillation Sequences - Nonsharp Separations. Computers & Chemical Engineering, 1990. 14(6): p. 631-653.

90. Andrecovich, M.J. and A.W. Westerberg, A Simple Synthesis Method Based on Utility Bounding for Heat-Integrated Distillation Sequences. Aiche Journal, 1985. 31(3): p. 363-375.

91. Agrawal, R., Synthesis of distillation column configurations for a multicomponent separation. Industrial & Engineering Chemistry Research, 1996. 35(4): p. 1059-1071.

92. Luyben, M.L. and C.A. Floudas, Analyzing the Interaction of Design and Control .1. A Multiobjective Framework and Application to Binary Distillation Synthesis. Computers & Chemical Engineering, 1994. 18(10): p. 933-969.

93. Brignole, E.A., S. Bottini, and R. Gani, A Strategy for the Design and Selection of Solvents for Separation Processes. Fluid Phase Equilibria, 1986. 29: p. 125-132.

94. Hasan, M.M.F., et al., A multi-scale framework for CO2 capture, utilization, and sequestration: CCUS and CCUM. Computers & Chemical Engineering, 2015. 81: p. 2-21.

95. Kongpanna, P., et al., Systematic methods and tools for design of sustainable chemical processes for CO2 utilization. Computers & Chemical Engineering, 2016. 87: p. 125-144.

96. Gabriel, K.J. and M.M. El-Halwagi, Modeling and optimization of a bioethanol production facility. Clean Technologies and Environmental Policy, 2013. 15: p. 931-944.

97. Prasad, P. and C.T. Maravelias, Batch selection, assignment and sequencing in multi-stage multi-product processes. Computers & Chemical Engineering, 2008. 32(6): p. 1106-1119.

98. Herron, J.A., et al., A general framework for the assessment of solar fuel technologies. Energy & Environmental Science, 2015. 8(1): p. 126-157.

99. Sen, S.M., et al., Catalytic conversion of lignocellulosic biomass to fuels: Process development and technoeconomic evaluation. Chemical Engineering Science, 2012. 67(1): p. 57-67.

100. Bond, J.Q., et al., Production of renewable jet fuel range alkanes and commodity chemicals from integrated catalytic processing of biomass. Energy & Environmental Science, 2014. 7(4): p. 1500-1523.

101. Eden, M.R., et al., A novel framework for simultaneous separation process and product design. Chemical Engineering and Processing, 2004. 43(5): p. 595-608.

102. Babi, D.K., J.M. Woodley, and R. Gani, Achieving More Sustainable Designs through a Process Synthesis-Intensification Framework. 24th European Symposium on Computer Aided Process Engineering, Pts a and B, 2014. 33: p. 31-36.

103. Babi, D.K., et al., A process synthesis-intensification framework for the development of sustainable membrane-based operations. Chemical Engineering and Processing, 2014. 86: p. 173-195.

104. Babi, D.K., et al., Sustainable process synthesis intensification. Computers & Chemical Engineering, 2015. 81: p. 218-244.

105. Riggs, J.B. and M.N. Karim, Chemical and Bio-process Control. 2006: Ferret.

106. Friedland, B., Control system design: an introduction to state-space methods. 1986: McGraw-Hill.

107. Chen, C.T., Linear System Theory and Design. 1999: Oxford University Press.

108. Biegler, L.T., A.M. Cervantes, and A. Wachter, Advances in simultaneous strategies for dynamic process optimization. Chemical Engineering Science, 2002. 57(4): p. 575-593.

109. Bhatia, T. and L.T. Biegler, Dynamic optimization in the design and scheduling of multiproduct batch plants. Industrial & Engineering Chemistry Research, 1996. 35(7): p. 2234-2246.

110. Lima, N.M.N., et al., Novel two-steps optimal control of batch polymerization reactors and application to PMMA production for the fabrication of artificial bone tissue. 23 European Symposium on Computer Aided Process Engineering, 2013. 32: p. 163-168.

111. Srinivasan, B., S. Palanki, and D. Bonvin, Dynamic optimization of batch processes - I. Characterization of the nominal solution. Computers & Chemical Engineering, 2003. 27(1): p. 1-26.

112. Fikar, M., et al., Application of iterative dynamic programming to optimal control of a distillation column. Canadian Journal of Chemical Engineering, 1998. 76(6): p. 1110-1117.

113. Hu, Q., et al., Optimal control of a batch cooling seeded crystallizer. Powder Technology, 2005. 156(2-3): p. 170-176.

114. Lang, Y.D., A.M. Cervantes, and L.T. Biegler, Dynamic optimization of a batch cooling crystallization process. Industrial & Engineering Chemistry Research, 1999. 38(4): p. 1469-1477.

115. Luus, R., Application of Iterative Dynamic Programming to Establish Optimal Control of a L-lysine Fermentation Process. Revista De Chimie, 2010. 61(8): p. 713-717.

116. Balsa-Canto, E., et al., Dynamic optimization of chemical and biochemical processes using restricted second-order information. Computers & Chemical Engineering, 2001. 25(4-6): p. 539-546.

117. Logist, F., P.M.M. Van Erdeghem, and J.F. Van Impe, Efficient deterministic multiple objective optimal control of (bio)chemical processes. Chemical Engineering Science, 2009. 64(11): p. 2527-2538.

118. Hugh, T.B., Howard Florey, Alexander Fleming and the fairy tale of penicillin. Medical Journal of Australia, 2002. 177(1): p. 52-53.

119. Coates, T., et al., Hepatitis B vaccines: Assessment of the seroprotective efficacy of two recombinant DNA vaccines. Clinical Therapeutics, 2001. 23(3): p. 392-403.

120. Scolnick, E.M., et al., Clinical-Evaluation in Healthy-Adults of a Hepatitis-B Vaccine Made by Recombinant DNA. Jama-Journal of the American Medical Association, 1984. 251(21): p. 2812-2815.

121. Ballou, W.R., et al., Safety and Efficacy of a Recombinant-DNA Plasmodium-Falciparum Sporozoite Vaccine. Lancet, 1987. 1(8545): p. 1277-1281.

122. Sriraman, K. and G. Jayaraman, Enhancement of recombinant streptokinase production in Lactococcus lactis by suppression of acid tolerance response. Applied Microbiology and Biotechnology, 2006. 72(6): p. 1202-1209.

123. Goyal, D., G. Sahni, and D.K. Sahoo, Enhanced production of recombinant streptokinase in Escherichia coli using fed-batch culture. Bioresource Technology, 2009. 100(19): p. 4468-4474.

124. Raskin, P. and R.S. Clements, The Use of Human Insulin Derived from Bakers-Yeast by Recombinant-DNA Technology. Clinical Therapeutics, 1991. 13(5): p. 569-578.

125. Johnson, I.S., Human Insulin from Recombinant DNA Technology. Science, 1983. 219(4585): p. 632-637.

126. Shuler, M.L. and F. Kargi, Bioprocess Engineering: Basic Concepts. 2002, New Jersey: Prentice Hall.

127. Roberts, R.J., How restriction enzymes became the workhorses of molecular biology. Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(17): p. 5905-5908.

128. Alberts, B., A. Johnson, and J. Lewis, Isolating, Cloning, and Sequencing DNA., in Molecular Biology of the Cell. 2002, Garland Science: New York.

129. Bartlett, J.M. and D. Stirling, A short history of the polymerase chain reaction. Methods Mol Biol, 2003. 226: p. 3-6.

130. Kaufman, R.I. and B.T. Nixon, Use of PCR to isolate genes encoding sigma54-dependent activators from diverse bacteria. J Bacteriol, 1996. 178(13): p. 3967-70.

131. Berg, P. and J.E. Mertz, Personal Reflections on the Origins and Emergence of Recombinant DNA Technology. Genetics, 2010. 184(1): p. 9-17.

132. Lederberg, J., Cell Genetics and Hereditary Symbiosis. Physiological Reviews, 1952. 32(4): p. 403-430.

133. Baudin, A., et al., A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. Nucleic Acids Res, 1993. 21(14): p. 3329-30.

134. Hartman, S.C. and R.C. Mulligan, Two dominant-acting selectable markers for gene transfer studies in mammalian cells. Proc Natl Acad Sci U S A, 1988. 85(21): p. 8047-51.

135. Herreraestrella, L., et al., Chimeric Genes as Dominant Selectable Markers in Plant-Cells. Embo Journal, 1983. 2(6): p. 987-995.

136. Jang, C.W. and T. Magnuson, A Novel Selection Marker for Efficient DNA Cloning and Recombineering in E. coli. Plos One, 2013. 8(2).

137. Weiss, B. and C.C. Richardson, Enzymatic Breakage and Joining of Deoxyribonucleic Acid .I. Repair of Single-Strand Breaks in DNA by an Enzyme System from Escherichia Coli Infected with T4 Bacteriophage. Proceedings of the National Academy of Sciences of the United States of America, 1967. 57(4): p. 1021-+.

138. Avery, O.T., C.M. MacLeod, and M. McCarty, Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. Journal of Experimental Medicine, 1944. 79(2): p. 137-158.

139. Griffith, F., The Significance of Pneumococcal Types. Journal of Hygiene, 1928. 27(2): p. 113-159.

140. Hanahan, D., Studies on Transformation of Escherichia-Coli with Plasmids. Journal of Molecular Biology, 1983. 166(4): p. 557-580.

141. Mandel, M. and A. Higa, Calcium-Dependent Bacteriophage DNA Infection. Journal of Molecular Biology, 1970. 53(1): p. 159-&.

142. Wirth, R., A. Friesenegger, and S. Fiedler, Transformation of Various Species of Gram-Negative Bacteria Belonging to 11 Different Genera by Electroporation. Molecular & General Genetics, 1989. 216(1): p. 175-177.

143. Doudna, J.A. and E. Charpentier, The new frontier of genome engineering with CRISPR-Cas9. Science, 2014. 346(6213): p. 1077-+.

144. Hsu, P.D., et al., DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology, 2013. 31(9): p. 827-+.

145. Ran, F.A., et al., Genome engineering using the CRISPR-Cas9 system. Nature Protocols, 2013. 8(11): p. 2281-2308.

146. Cong, L., et al., Multiplex Genome Engineering Using CRISPR/Cas Systems. Science, 2013. 339(6121): p. 819-823.

147. Sakuma, T., et al., Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Scientific Reports, 2014. 4.

148. Lu, X.J., et al., CRISPR-Cas9: a new and promising player in gene therapy. Journal of Medical Genetics, 2015. 52(5): p. 289-296.

149. Liao, H.K., et al., Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. Nature Communications, 2015. 6.

150. Long, C.Z., et al., Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science, 2014. 345(6201): p. 1184-1188.

151. Wang, T., et al., Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science, 2014. 343(6166): p. 80-84.

152. Hsu, P.D., E.S. Lander, and F. Zhang, Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell, 2014. 157(6): p. 1262-1278.

153. Jeong, H., et al., The large-scale organization of metabolic networks. Nature, 2000. 407(6804): p. 651-654.

154. Kauffman, K.J., P. Prakash, and J.S. Edwards, Advances in flux balance analysis. Current Opinion in Biotechnology, 2003. 14(5): p. 491-496.

155. Schellenberger, J., et al., Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nature Protocols, 2011. 6(9): p. 1290-1307.

156. Krieger, C.J., et al., MetaCyc: a multiorganism database of metabolic pathways and enzymes. Nucleic Acids Research, 2004. 32: p. D438-D442.

157. Monod, J., The Growth of Bacterial Cultures. Annual Review of Microbiology, 1949. 3: p. 371-394.

158. Segel, I.H., Enzyme-Kinetics. Bioscience, 1976. 26(7): p. 425-426.

159. Johnson, K.A. and R.S. Goody, The Original Michaelis Constant: Translation of the 1913 Michaelis-Menten Paper. Biochemistry, 2011. 50(39): p. 8264-8269.

160. Briggs, G.E. and J.B.S. Haldane, A note on the kinetics of enzyme action. Biochemical Journal, 1925. 19(2): p. 338-339.

161. Lineweaver, H. and D. Burk, The determination of enzyme dissociation constants. Journal of the American Chemical Society, 1934. 56: p. 658-666.

162. Eadie, G.S., The inhibition of cholinesterase by physostigmine and prostigmine. Journal of Biological Chemistry, 1942. 146(1): p. 85-93.

163. Hofstee, B.H.J., Non-Inverted Versus Inverted Plots in Enzyme Kinetics. Nature, 1959. 184(4695): p. 1296-1298.

164. Hanes, C.S., Studies on plant amylases. I. The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. Biochemical Journal, 1932. 26: p. 1406-1421.

165. Haldane, J.B.S., Graphical Methods in Enzyme Chemistry. Nature, 1957. 179(4564): p. 832-832.

166. Binder, J.B. and R.T. Raines, Fermentable sugars by chemical hydrolysis of biomass. Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(10): p. 4516-4521.

167. Peri, S., et al., Dynamics of cello-oligosaccharides on a cellulose crystal surface. Cellulose, 2012. 19(6): p. 1791-1806.

168. Ryu, S. and M.N. Karim, A whole cell biocatalyst for cellulosic ethanol production from dilute acid-pretreated corn stover hydrolyzates. Applied Microbiology and Biotechnology, 2011. 91(3): p. 529-542.

169. Hodge, D.B., et al., Model-Based Fed-Batch for High-Solids Enzymatic Cellulose Hydrolysis. Applied Biochemistry and Biotechnology, 2009. 152(1): p. 88-107.

170. Hodge, D.B., et al., Soluble and insoluble solids contributions to high-solids enzymatic hydrolysis of lignocellulose. Bioresource Technology, 2008. 99(18): p. 8940-8948.

171. Peri, S., et al., Modeling intrinsic kinetics of enzymatic cellulose hydrolysis. Biotechnology Progress, 2007. 23(3): p. 626-637.

172. Ito, S., Alkaline cellulases from alkaliphilic Bacillus: Enzymatic properties, genetics, and application to detergents. Extremophiles, 1997. 1(2): p. 61-66.

173. Schmid, R.D. and R. Verger, Lipases: Interfacial enzymes with attractive applications. Angewandte Chemie-International Edition, 1998. 37(12): p. 1609-1633.

174. Bhat, M.K., Cellulases and related enzymes in biotechnology. Biotechnology Advances, 2000. 18(5): p. 355-383.

175. Hasan, F., A.A. Shah, and A. Hameed, Industrial applications of microbial lipases. Enzyme and Microbial Technology, 2006. 39(2): p. 235-251.

176. Patel, A.K., R.R. Singhania, and A. Pandey, Novel enzymatic processes applied to the food industry. Current Opinion in Food Science, 2016. 7: p. 64-72.

177. Minussi, R.C., G.M. Pastore, and N. Duran, Potential applications of laccase in the food industry. Trends in Food Science & Technology, 2002. 13(6-7): p. 205-216.

178. Shahidi, F. and Y.V.A.J. Kamil, Enzymes from fish and aquatic invertebrates and their application in the food industry. Trends in Food Science & Technology, 2001. 12(12): p. 435-464.

179. Carpio, C., et al., Bone-bound enzymes for food industry application. Food Chemistry, 2000. 68(4): p. 403-409.

180. Fujiwara, N. and K. Yamamoto, Production of Alkaline Protease in a Low-Cost Medium by Alkalophilic Bacillus-Sp and Properties of the Enzyme. Journal of Fermentation Technology, 1987. 65(3): p. 345-348.

181. Phadatare, S.U., V.V. Deshpande, and M.C. Srinivasan, High-Activity Alkaline Protease from Conidiobolus-Coronatus (Ncl 86.8.20) - Enzyme-Production and Compatibility with Commercial Detergents. Enzyme and Microbial Technology, 1993. 15(1): p. 72-76.

182. Zheng, G.W. and J.H. Xu, New opportunities for biocatalysis: driving the synthesis of chiral chemicals. Current Opinion in Biotechnology, 2011. 22(6): p. 784-792.

183. Woodley, J.M., New opportunities for biocatalysis: making pharmaceutical processes greener. Trends in Biotechnology, 2008. 26(6): p. 321-327.

184. Shewale, J.G. and H. Sivaraman, Penicillin Acylase - Enzyme-Production and Its Application in the Manufacture of 6-Apa. Process Biochemistry, 1989. 24(4): p. 146-154.

185. Kim, S., C. Jimenez-Gonzalez, and B.E. Dale, Enzymes for pharmaceutical applications-a cradle-to-gate life cycle assessment. International Journal of Life Cycle Assessment, 2009. 14(5): p. 392-400.

186. Sanchez, S. and A.L. Demain, Enzymes and Bioconversions of Industrial, Pharmaceutical, and Biotechnological Significance. Organic Process Research & Development, 2011. 15(1): p. 224-230.

187. Simaria, A.S., et al., Allogeneic Cell Therapy Bioprocess Economics and Optimization: Single-Use Cell Expansion Technologies. Biotechnology and Bioengineering, 2014. 111(1): p. 69-83.

188. Polak, J.M. and S. Mantalaris, Stem cells bioprocessing: An important milestone to move regenerative medicine research into the clinical arena. Pediatric Research, 2008. 63(5): p. 461-466.

189. Kirouac, D.C. and P.W. Zandstra, The systematic production of cells for cell therapies. Cell Stem Cell, 2008. 3(4): p. 369-81.

190. Yan, Y., L. Song, and Y. Li, Toward biomanufacturing of pluripotent stem cell derived products: scale out and scale up. Pharmaceutical Bioprocessing, 2015. 3: p. 25-33.

191. Schnitzler, A.C., et al., Bioprocessing of human mesenchymal stem/stromal cells for therapeutic use: Current technologies and challenges. Biochemical Engineering Journal, 2016. 108: p. 3-13.

192. Sart, S., S.N. Agathos, and Y. Li, Process engineering of stem cell metabolism for large scale expansion and differentiation in bioreactors. Biochemical Engineering Journal, 2014. 84: p. 74-82.

193. Misener, R., et al., Global superstructure optimisation of red blood cell production in a parallelised hollow fibre bioreactor. Computers & Chemical Engineering, 2014. 71: p. 532-553.

194. Sueblinvong, V. and D.J. Weiss, Stem cells and cell therapy approaches in lung biology and diseases. Translational Research, 2010. 156(3): p. 188-205.

195. Applegate, L.A., et al., Whole-Cell Bioprocessing of Human Fetal Cells for Tissue Engineering of Skin. Skin Pharmacology and Physiology, 2009. 22(2): p. 63-73.

196. Ladisch, M.R., Bioseparations scale-up principles, practice and economics. Abstracts of Papers of the American Chemical Society, 2002. 224: p. U218-U218.

197. Farid, S.S., Economic Drivers and Trade-Offs in Antibody Purification Processes. Biopharm International, 2009: p. 38-42.

198. Harrison, R.G., et al., Bioseparations Science and Engineering. 2003: Oxford University Press.

199. Sadama, A. and A.M. Beelaram, Efficiency and Economics of Bioseparation - Some Case-Studies. Abstracts of Papers of the American Chemical Society, 1995. 209: p. 56-Biot.

200. Ramakrishnan, A. and A. Sadana, Economics of bioseparation processes. Separation Science and Technology, 2000. 2: p. 667-685.

201. Kim, J.J. and K. Park, Smart hydrogels for bioseparation. Bioseparation, 1998. 7(4-5): p. 177-184.

202. Przybycien, T.M., N.S. Pujar, and L.M. Steele, Alternative bioseparation operations: life beyond packed-bed chromatography. Current Opinion in Biotechnology, 2004. 15(5): p. 469-478.

203. Wang, J.L., T.Y. Gu, and J.J. Zhong, Enhanced recovery of antitumor ganoderic acid T from Ganoderma lucidum mycelia by novel chemical conversion strategy. Biotechnology and Bioengineering, 2012. 109(3): p. 754-762.

204. Plumb, K., Continuous processing in the pharmaceutical industry - Changing the mind set. Chemical Engineering Research & Design, 2005. 83(A6): p. 730-738.

205. Yang, C. and J. Hou, Fed-batch fermentation penicillin process fault diagnosis and detection based on support vector machine. Neurocomputing, 2016. 190: p. 117-123.

206. San, K.-Y. and G. Stephanopoulos, Optimization of fed-batch penicillin fermentation: A case of singular optimal control with state constraints. Biotechnology and Bioengineering, 1989. 34: p. 72-78.

207. Mendoza-Vega, O., C. Hebert, and S.W. Brown, Production of recombinant hirudin by high cell density fed-batch cultivations of a Saccharomyces cerevisiae strain: physiological considerations during the bioprocess design. Journal of Biotechnology, 1994. 32: p. 249-259.

208. Dhir, S., et al., Dynamic optimization of hybridoma growth in a fed-batch bioreactor. Biotechnology and Bioengineering, 2000. 67: p. 197-205.

209. Ierapetritou, M., F. Muzzio, and G. Reklaitis, Perspectives on the Continuous Manufacturing of Powder-Based Pharmaceutical Processes. Aiche Journal, 2016. 62(6): p. 1846-1862.

210. Leuenberger, H., New trends in the production of pharmaceutical granules: batch versus continuous processing. European Journal of Pharmaceutics and Biopharmaceutics, 2001. 52(3): p. 289-296.

211. Douglas, J.M., Conceptual Design of Chemical Processes. 1988, New York, NY: McGraw-Hill.

212. Fogler, H.S., Elements of chemical reaction engineering. Fourth edition. ed. Prentice Hall PTR international series in the physical and chemical engineering sciences; Prentice-Hall international series in the physical and chemical engineering sciences. 2006, Upper Saddle River, NJ :: Prentice Hall PTR.

213. Lee, S.L., et al., Modernizing Pharmaceutical Manufacturing: from Batch to Continuous Production. Journal of Pharmaceutical Innovation, 2015. 10(3): p. 191-199.

214. Armstrong, G.A. and J.E. Hearst, Carotenoids .2. Genetics and molecular biology of carotenoid pigment biosynthesis. Faseb Journal, 1996. 10(2): p. 228-237.

215. Bender, D.A., Vitamin A: Retinoids and Carotenoids, in Nutritional Biochemistry of the Vitamins. 2003, Cambridge University Press: New York, NY. p. 30-76.

216. Raja, R., S. Hemaiswarya, and R. Rengasamy, Exploitation of Dunaliella for beta-carotene production. Applied Microbiology and Biotechnology, 2007. 74(3): p. 517-523.

217. Ausich, R.L., Commercial opportunities for carotenoid production by biotechnology. Pure and Applied Chemistry, 1997. 69(10): p. 2169-2173.

218. Marusich, W.L. and J.C. Bauernfeind, Oxycarotenoids in Poultry Pigmentation .1. Yolk Studies. Poultry Science, 1970. 49(5): p. 1411-+.

219. Nelis, H.J. and A.P. Deleenheer, Microbial Sources of Carotenoid-Pigments Used in Foods and Feeds. Journal of Applied Bacteriology, 1991. 70(3): p. 181-191.

220. Simpson, K.L., T. Katayama, and C.O. Chichester, Carotenoids in Fish Feeds, in Carotenoids as Colorants and Vitamin A Precursors: Technological and Nutritional Applications. 1981, Academic Press: London, UK.

221. Edge, R., D.J. McGarvey, and T.G. Truscott, The carotenoids as anti-oxidants - a review. Journal of Photochemistry and Photobiology B-Biology, 1997. 41(3): p. 189-200.

222. Hulshof, P.J.M., et al., Quick screening of maize kernels for provitamin A content. Journal of Food Composition and Analysis, 2007. 20(8): p. 655-661.

223. Palozza, P. and N.I. Krinsky, Antioxidant Effects of Carotenoids Invivo and Invitro - an Overview. Methods in Enzymology, 1992. 213: p. 403-420.

224. Vanpoppel, G. and R.A. Goldbohm, Epidemiologic Evidence for Beta-Carotene and Cancer Prevention. American Journal of Clinical Nutrition, 1995. 62(6): p. S1393-S1402.

225. Frengova, G.I. and D.M. Beshkova, Carotenoids from Rhodotorula and Phaffia: yeasts of biotechnological importance. Journal of Industrial Microbiology & Biotechnology, 2009. 36(2): p. 163-180.

226. Vachali, P., P. Bhosale, and P.S. Bernstein, Microbial carotenoids. Methods Mol Biol, 2012. 898: p. 41-59.

227. Johnson, E.A. and W.A. Schroeder, Microbial carotenoids. Adv Biochem Eng Biotechnol, 1996. 53: p. 119-78.

228. Lee, P.C. and C. Schmidt-Dannert, Metabolic engineering towards biotechnological production of carotenoids in microorganisms. Applied Microbiology and Biotechnology, 2002. 60(1-2): p. 1-11.

229. Crabtree, H.G., Observations on the carbohydrate metabolism of tumours. Biochem J, 1929. 23(3): p. 536-45.

230. De Deken, R.H., The Crabtree effect: a regulatory system in yeast. J Gen Microbiol, 1966. 44(2): p. 149-56.

231. Fiechter, A. and F.K. Gmunder, Metabolic control of glucose degradation in yeast and tumor cells. Adv Biochem Eng Biotechnol, 1989. 39: p. 1-28.

232. Sonnleitner, B. and O. Kappeli, Growth of Saccharomyces-Cerevisiae Is Controlled by Its Limited Respiratory Capacity - Formulation and Verification of a Hypothesis. Biotechnology and Bioengineering, 1986. 28(6): p. 927-937.

233. Paczia, N., et al., Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms. Microbial Cell Factories, 2012. 11.

234. Xu, B., M. Jahic, and S.O. Enfors, Modeling of overflow metabolism in batch and fed-batch cultures of Escherichia coli. Biotechnology Progress, 1999. 15(1): p. 81-90.

235. Amribt, Z., H.X. Niu, and P. Bogaerts, Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures. Biochemical Engineering Journal, 2013. 70: p. 196-209.

236. Liu, Y.S. and J.Y. Wu, Modeling of Xanthophyllomyces dendrorhous Growth on Glucose and Overflow Metabolism in Batch and Fed-Batch Cultures for Astaxanthin Production. Biotechnology and Bioengineering, 2008. 101(5): p. 996-1004.

237. Luna-Flores, C.H., et al., Batch and fed-batch modeling of carotenoids production by Xanthophyllomyces dendrorhous using Yucca fillifera date juice as substrate. Biochemical Engineering Journal, 2010. 53(1): p. 131-136.

238. Reyes, L.H., J.M. Gomez, and K.C. Kao, Improving carotenoids production in yeast via adaptive laboratory evolution. Metabolic Engineering, 2014. 21: p. 26-33.

239. Olson, M.L., Metabolic Engineering of S. cerevisiae for Carotenoid Production Optimization, in Chemical Engineering. 2014, Texas A&M University.

240. Holzer, H., Regulation of carbohydrate metabolism by enzyme competition. Cold Spring Harb Symp Quant Biol, 1961. 26: p. 277-88.

241. Holzer, H. and H.W. Goedde, [Two ways from pyruvate to acetyl-coenzyme A in yeast]. Biochem Z, 1957. 329(3): p. 175-91.

242. Kappeli, O., Regulation of carbon metabolism in Saccharomyces cerevisiae and related yeasts. Adv Microb Physiol, 1986. 28: p. 181-209.

243. Postma, E., et al., Enzymic Analysis of the Crabtree Effect in Glucose-Limited Chemostat Cultures of Saccharomyces-Cerevisiae. Applied and Environmental Microbiology, 1989. 55(2): p. 468-477.

244. Jacobson, M.K. and C. Bernofsky, Mitochondrial Acetaldehyde Dehydrogenase from Saccharomyces-Cerevisiae. Biochimica Et Biophysica Acta, 1974. 350(2): p. 277-291.

245. Wills, C., Regulation of Sugar and Ethanol-Metabolism in Saccharomyces-Cerevisiae. Critical Reviews in Biochemistry and Molecular Biology, 1990. 25(4): p. 245-280.

246. Ozaydin, B., et al., Carotenoid-based phenotypic screen of the yeast deletion collection reveals new genes with roles in isoprenoid production. Metabolic Engineering, 2013. 15: p. 174-183.

247. Verwaal, R., et al., High-level production of beta-carotene in Saccharomyces cerevisiae by successive transformation with carotenogenic genes from Xanthophyllomyces dendrorhous. Applied and Environmental Microbiology, 2007. 73(13): p. 4342-4350.

248. Ye, V.M. and S.K. Bhatia, Pathway engineering strategies for production of beneficial carotenoids in microbial hosts. Biotechnology Letters, 2012. 34(8): p. 1405-1414.

249. Maiorella, B., H.W. Blanch, and C.R. Wilke, By-Product Inhibition Effects on Ethanolic Fermentation by Saccharomyces-Cerevisiae. Biotechnology and Bioengineering, 1983. 25(1): p. 103-121.

250. Yoon, H., G. Klinzing, and H.W. Blanch, Competition for Mixed Substrates by Microbial-Populations. Biotechnology and Bioengineering, 1977. 19(8): p. 1193-1210.

251. Ge, X.M. and F.W. Bai, Intrinsic kinetics of continuous growth and ethanol production of a flocculating fusant yeast strain SPSC01. Journal of Biotechnology, 2006. 124(2): p. 363-372.

252. Luedeking, R. and E.L. Piret, A Kinetic Study of the Lactic Acid Fermentation - Batch Process at Controlled Ph. Journal of Biochemical and Microbiological Technology and Engineering, 1959. 1(4): p. 393-412.

253. Krishnan, M.S., N.W.Y. Ho, and G.T. Tsao, Fermentation kinetics of ethanol production from glucose and xylose by recombinant Saccharomyces 1400(pLNH33). Applied Biochemistry and Biotechnology, 1999. 77-9: p. 373-388.

254. Reinsch, C.H., Smoothing by Spline Functions. Numerische Mathematik, 1967. 10(3): p. 177-&.

255. Swinnen, I.A.M., et al., Predictive modelling of the microbial lag phase: a review. International Journal of Food Microbiology, 2004. 94(2): p. 137-159.

256. Atkinson, B.M., F, Biochemical Engineering and Biotechnology Handbook. 1983, Surrey, UK: Macmillan.

257. Zeng, A.P., et al., Multiple Product Inhibition and Growth Modeling of Clostridium-Butyricum and Klebsiella-Pneumoniae in Glycerol Fermentation. Biotechnology and Bioengineering, 1994. 44(8): p. 902-911.

258. Cheng, K.K., H.J. Liu, and D.H. Liu, Multiple growth inhibition of Klebsiella pneumoniae in 1,3-propanediol fermentation. Biotechnology Letters, 2005. 27(1): p. 19-22.

259. Khan, N.S., et al., Modeling the growth of Corynebacterium glutamicum under product inhibition in L-glutamic acid fermentation. Biochemical Engineering Journal, 2005. 25(2): p. 173-178.

260. PriceWaterhouseCoopers, From Vision to Decision: Pharma 2020. 2016.

261. The Pharmaceutical Industry in Figures. 2015; Available from: http://eee.efpia.eu/uploads/Figures_2015_Key_data.pdf.

262. Summary of NDA Approvals & Receipts. 2014; Available from: http://www.fda.gov/AboutFDA/WhatWeDo/ProductRegulation/SummaryofNDAApprovalsReceipts1938tothepresent/default.htm.

263. Jungbauer, A. and J. Peng, Continous bioprocessing: An interview with Konstantin Konstantinov from Genzyme. Biotechnology Journal, 2011. 6: p. 1431-1433.

264. Gnoth, S., R. Simutis, and A. Lübbert, Fermentation process supervision and strategies for fail-safe operation: A practical approach. Engineering in Life Sciences, 2011. 11: p. 94-106.

265. Höpfner, T., et al., A review of non-invasive optical-based image analysis systems for continuous bioprocess monitoring. Bioprocess and Biosystems Engineering, 2009. 33: p. 247-256.

266. Wu, S.-Y., et al., Fermentative hydrogen production and bacterial community structure in high-rate anaerobic bioreactors containing silicone-immobilized and self-flocculated sludge. Biotechnology and Bioengineering, 2006. 93: p. 934-946.

267. González, G., et al., Biodegradation of phenolic industrial wastewater in a fluidized bed bioreactor with immobilized cells of Pseudomonas putida. Bioresource Technology, 2001. 80: p. 137-142.

268. Kheyrandish, M., et al., Direct production of acetone–butanol–ethanol from waste starch by free and immobilized Clostridium acetobutylicum. Fuel, 2015. 142: p. 129-133.

269. Belfort, G., Membranes and bioreactors: A technical challenge in biotechnology. Biotechnology and Bioengineering, 1989. 33: p. 1047-1066.

270. Taheran, M., et al., Membrane processes for removal of pharmaceutically active compounds (PhACs) from water and wastewaters. Science of The Total Environment, 2016. 547: p. 60-77.

271. Mannina, G., et al., Sequential batch membrane bio-reactor for wastewater treatment: The effect of increased salinity. Bioresource Technology, 2016. 209: p. 205-212.

272. Yao, P., et al., Cell growth behaviors of Clostridium acetobutylicum in a pervaporation membrane bioreactor for butanol fermentation. Biotechnology and Applied Biochemistry, 2016. 63: p. 101-105.

273. Fan, S., et al., Kinetic model of continuous ethanol fermentation in closed-circulating process with pervaporation membrane bioreactor by Saccharomyces cerevisiae. Bioresource Technology, 2015. 177: p. 169-175.

274. Staniszewski, M., W. Kujawski, and M. Lewandowska, Ethanol production from whey in bioreactor with co-immobilized enzyme and yeast cells followed by pervaporative recovery of product – Kinetic model predictions. Journal of Food Engineering, 2007. 82: p. 618-625.

275. Warikoo, V., et al., Integrated continuous production of recombinant therapeutic proteins. Biotechnology and Bioengineering, 2012. 109: p. 3018-3029.

276. Yamane, T., M. Fukunaga, and Y.W. Lee, Increased PHB productivity by high-cell-density fed-batch culture of Alcaligenes latus, a growth-associated PHB producer. Biotechnology and Bioengineering, 1996. 50: p. 197-202.

277. Henson, M.A. and D.E. Seborg, Nonlinear control strategies for continuous fermenters. Chemical Engineering Science, 1992. 47: p. 821-835.

278. Modak, J.M. and H.C. Lim, Simple nonsingular control approach to fed-batch fermentation optimization. Biotechnology and Bioengineering, 1989. 33: p. 11-15.

279. Hodge, D.B. and M.N. Karim, Modeling and Advanced Control of Recombinant Zymomonasmobilis Fed-Batch Fermentation. Biotechnology Progress, 2002. 18: p. 572-579.

280. Saucedo, V.M. and M.N. Karim, Experimental optimization of a real time fed-batch fermentation process using Markov decision process. Biotechnology and Bioengineering, 1997. 55: p. 317-327.

281. Sridhar, L.N. and E.S.L. Saucedo, Optimal Control of Saccharomyces cerevisiae. Fermentation Process. Chemical Engineering Communications, 2016. 203: p. 318-325.

282. Ho, K., C. Tam, and B. Zhou, Growth and carotenoid production of Phaffia Rhodozyma in fed-batch cultures with different feeding methods. Biotechnology Letters, 1999. 21: p. 175-178.

283. Ding, S. and T. Tan, l-lactic acid production by Lactobacillus casei fermentation using different fed-batch feeding strategies. Process Biochemistry, 2006. 41: p. 1451-1454.

284. Vassiliadis, V.S., R.W.H. Sargent, and C.C. Pantelides, Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints. Industrial & Engineering Chemistry Research, 1994. 33: p. 2111–2122.

285. Garg, D., et al., A unified framework for the numerical solution of optimal control problems using pseudospectral methods. Automatica, 2010. 46: p. 1843-1851.

286. Flores-Tlacuahuac, A., S.T. Moreno, and L.T. Biegler, Global optimization of highly nonlinear dynamic systems. Industrial & Engineering Chemistry Research, 2008. 47(8): p. 2643-2655.

287. Hairer, E. and G. Wanner, Stiff differential equations solved by Radau methods. Journal of Computational and Applied Mathematics, 1999. 111: p. 93-111.

288. Wächter, A. and L.T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Mathematical Programming, 2006. 106: p. 25-57.

289. Chandler, L.A. and S.J. Schwartz, Isomerization and Losses of Trans-Beta-Carotene in Sweet-Potatoes as Affected by Processing Treatments. Journal of Agricultural and Food Chemistry, 1988. 36(1): p. 129-133.

290. Desobry, S.A., F.M. Netto, and T.P. Labuza, Preservation of beta-carotene from carrots. Critical Reviews in Food Science and Nutrition, 1998. 38(5): p. 381-396.

291. Britton, G., Structure and properties of carotenoids in relation to function. The FASEB Journal, 1995. 9: p. 1551-1558.

292. Caswell, M. and D. Zilberman, Algolculture. 2001.

293. Stahl, W. and H. Sies, Bioactivity and protective effects of natural carotenoids. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2005. 1740: p. 101-107.

294. Olson, M.L., et al., Characterization of an evolved carotenoids hyper-producer of Saccharomyces cerevisiae through bioreactor parameter optimization and Raman spectroscopy. Journal of Industrial Microbiology & Biotechnology, 2016: p. 1-9.

295. Ordoñez, M.C., et al., Modelling of batch kinetics of aerobic carotenoid production using Saccharomyces cerevisiae. Biochemical Engineering Journal, 2016. 114: p. 226-236.

296. Chang, J. ICIS Chemical Pricing. [cited 2017; Available from: http://www.icis.com/chemicals/channel-info-chemicals-a-z/.

297. U.S. wholesale list price for glucose syrup, Midwest markets, monthly, quarterly, and by calendar and fiscal year. Sugars and Sweeteners Yearbook Tables 2016; Available from: http://www.ers.usda.gov/data-products/sugar-and-sweeteners-yearbook-tables.aspx#25442.

298. Thompson, P., The Agricultural Ethics of Biofuels: The Food vs. Fuel Debate. Agriculture, 2012. 2: p. 339-358.

299. Niziolek, A.M., et al., Coal and Biomass to Liquid Transportation Fuels: Process Synthesis and Global Optimization Strategies. Industrial & Engineering Chemistry Research, 2014. 53: p. 17002-17025.

300. Hodge, D.B., et al., Soluble and insoluble solids contributions to high-solids enzymatic hydrolysis of lignocellulose. Bioresource Technology, 2008. 99: p. 8940-8948.

301. Tao, L., et al., Process and technoeconomic analysis of leading pretreatment technologies for lignocellulosic ethanol production using switchgrass. Bioresource Technology, 2011. 102: p. 11105-11114.

302. Dutta, A., et al., An economic comparison of different fermentation configurations to convert corn stover to ethanol using Z. mobilis and Saccharomyces. Biotechnology Progress, 2010. 26: p. 64–72.

303. Humbird, D., et al., Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol. Technical Report, 2011: p. 1-147.

304. Zondervan, E., et al., Optimal design of a multi-product biorefinery system. Computers & Chemical Engineering, 2011. 35: p. 1752-1766.

305. Martín, M. and I.E. Grossmann, Energy optimization of bioethanol production via hydrolysis of switchgrass. AIChE Journal, 2012. 58: p. 1538–1549.

306. Martín, M. and I.E. Grossmann, Optimal simultaneous production of i-butene and ethanol from switchgrass. Biomass and Bioenergy, 2014. 61: p. 93-103.

307. da Silva, A.R.G., C.E. Torres Ortega, and B.-G. Rong, Techno-economic analysis of different pretreatment processes for lignocellulosic-based bioethanol production. Bioresource Technology, 2016. 218: p. 561-570.

308. Piccolo, C. and F. Bezzo, A techno-economic comparison between two technologies for bioethanol production from lignocellulose. Biomass and Bioenergy, 2009. 33: p. 478-491.

309. Scott, F., et al., Selection of process alternatives for lignocellulosic bioethanol production using a MILP approach. Bioresource Technology, 2013. 148: p. 525-534.

310. Wyman, C.E., What is (and is not) vital to advancing cellulosic ethanol. Trends in Biotechnology, 2007. 25: p. 153-157.

311. Lynd, L.R., Overview and Evaluation of Fuel Ethanol from Cellulosic Biomass: Technology, Economics, the Environment, and Policy. Annual Review of Energy and the Environment, 1996. 21: p. 403-465.

312. Chung, D., et al., Direct conversion of plant biomass to ethanol by engineered Caldicellulosiruptor bescii. Proceedings of the National Academy of Sciences, 2014. 111: p. 8931-8936.

313. Den Haan, R., et al., Hydrolysis and fermentation of amorphous cellulose by recombinant Saccharomyces cerevisiae. Metabolic Engineering, 2007. 9: p. 87-94.

314. Jin, M., et al., Consolidated bioprocessing (CBP) of AFEX™-pretreated corn stover for ethanol production using Clostridium phytofermentans at a high solids loading. Biotechnology and Bioengineering, 2012. 109: p. 1929-1936.

315. Ryu, S. and M.N. Karim, A whole cell biocatalyst for cellulosic ethanol production from dilute acid-pretreated corn stover hydrolyzates. Applied Microbiology and Biotechnology, 2011. 91: p. 529-542.

316. Quaglia, A., et al., Systematic network synthesis and design: Problem formulation, superstructure generation, data management and solution. Computers & Chemical Engineering, 2015. 72: p. 68-86.

317. Alex Marvin, W., et al., Economic Optimization of a Lignocellulosic Biomass-to-Ethanol Supply Chain. Chemical Engineering Science, 2012. 67: p. 68-79.

318. Dias, M.O.S., et al., Cogeneration in integrated first and second generation ethanol from sugarcane. Chemical Engineering Research and Design, 2013. 91: p. 1411-1417.

319. Bain, R.L., World Biofuels Assessment; Worldwide Biomass Potential: Technology Characterizations (Milestone Report). 2007, National Renewable Energy Laboratory (NREL), Golden, CO.

320. Biomass Feedstock Composition and Property Database. [cited 2016; Available from: http://www.afdc.energy.gov/biomass/prog/search1.cgi.

321. Sorghum Monthly Prices - US Dollars per Metric Ton. 2016; Available from: http://ycharts.com/indicators/us_sorghum_price_gulf_ports.

322. PLANTS Database. [cited 2016; Available from: http://plants.usda.gov/java/.

323. Agbor, V.B., et al., Biomass pretreatment: Fundamentals toward application. Biotechnology Advances, 2011. 29: p. 675-685.

324. Yu, Q., et al., Liquid hot water pretreatment of sugarcane bagasse and its comparison with chemical pretreatment methods for the sugar recovery and structural changes. Bioresource Technology, 2013. 129: p. 592-598.

325. Banerji, A., M. Balakrishnan, and V.V.N. Kishore, Low severity dilute-acid hydrolysis of sweet sorghum bagasse. Applied Energy, 2013. 104: p. 197-206.

326. Wu, L., et al., Low temperature alkali pretreatment for improving enzymatic digestibility of sweet sorghum bagasse for ethanol production. Bioresource Technology, 2011. 102: p. 4793-4799.

327. Esteghlalian, A., et al., Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. Bioresource Technology, 1997. 59: p. 129-136.

328. Kim, Y., N.S. Mosier, and M.R. Ladisch, Enzymatic digestion of liquid hot water pretreated hybrid poplar. Biotechnology Progress, 2009. 25: p. 340-348.

329. Alizadeh, H., et al., Pretreatment of switchgrass by ammonia fiber explosion (AFEX). Applied Biochemistry and Biotechnology, 2005. 124: p. 1133-1141.

330. Xu, J., et al., Sodium Hydroxide Pretreatment of Switchgrass for Ethanol Production. Energy & Fuels, 2010. 24: p. 2113-2119.

331. Teymouri, F., et al., Optimization of the ammonia fiber explosion (AFEX) treatment parameters for enzymatic hydrolysis of corn stover. Bioresource Technology, 2005. 96: p. 2014-2018.

332. Currie, D.H., et al., Functional heterologous expression of an engineered full length CipA from Clostridium thermocellum in Thermoanaerobacterium saccharolyticum. Biotechnology for Biofuels, 2013. 6: p. 32.

333. Groposo, C., A.M.d. Castro, and N. Pereira Jr., Effects of agitation and exogenous H2 on bioconversion of sugarcane bagasse into ethanol by Clostridium thermocellum ATCC 27405. Electronic Journal of Biotechnology, 2013. 16.

334. Lin, P.P., et al., Consolidated bioprocessing of cellulose to isobutanol using Clostridium thermocellum. Metabolic Engineering, 2015. 31: p. 44-52.

335. Shaw, A.J., et al., Urease expression in a Thermoanaerobacterium saccharolyticum ethanologen allows high titer ethanol production. Metabolic Engineering, 2012. 14: p. 528-532.

336. Argyros, D.A., et al., High Ethanol Titers from Cellulose by Using Metabolically Engineered Thermophilic, Anaerobic Microbes. Applied and Environmental Microbiology, 2011. 77: p. 8288-8294.

337. Hankey, R., Electric Power Monthly (May 2015), in Electric Power Monthly. 2015, U.S. Energy Information Administration: U.S. Department of Energy. p. 222.

338. Osmani, A. and J. Zhang, Economic and environmental optimization of a large scale sustainable dual feedstock lignocellulosic-based bioethanol supply chain in a stochastic environment. Applied Energy, 2014. 114: p. 572-587.

339. System, B.o.G.o.t.F.R., Survey of Terms of Business Lending, in Survey of Terms and Lending. 2015, U.S. Federal Reseve System.

340. Towler, G.P. and R.K. Sinnott, Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design. 2nd ed. 2013, Waltham, MA: Butterworth-Heinemann.

341. Avis, K.E. and V.L. Wu, Biotechnology and Biopharmaceutical Manufacturing, Processing, and Preservation. Drug Manufacturing Technology. Vol. 2. 2000, New York, NY: Interpharm/CRC.

342. El-Halwagi, M.M., Sustainable Design through Process Integration. 2011, Boston, MA: Butterworth-Heinemann.

343. Cheng, Y.S., et al., Evaluation of High Solids Alkaline Pretreatment of Rice Straw. Applied Biochemistry and Biotechnology, 2010. 162(6): p. 1768-1784.

344. Gibson, L.J., The hierarchical structure and mechanics of plant materials. Journal of The Royal Society Interface, 2012. 9: p. 2749-2766.

345. Gonçalves, D.L., et al., Xylose and xylose/glucose co-fermentation by recombinant Saccharomyces cerevisiae strains expressing individual hexose transporters. Enzyme and Microbial Technology, 2014. 63: p. 13-20.

346. Doherty, M.F. and M.F. Malone, Conceptual Design of Distillation Systems. 2001, Boston, MA: McGraw-Hill.

347. Miller Jr., J.W., G.R. Shorr, and C.L. Yaws, Heat Capacity of Liquids. Chemical Engineering, 1976. 83: p. 129-131.

348. Voitkevich, O.V., et al., Thermodynamic Properties of Plant Biomass Components. Heat Capacity, Combustion Energy, and Gasification Equilibria of Lignin. Journal of Chemical & Engineering Data, 2012. 57: p. 1903-1909.

349. Sandler, S., Chemical, Biochemical, and Engineering Thermodynamics. 2006, New Jersey: Wiley.

350. Battery and Energy Technologies. Steam Turbine Electricity Generation Plants [cited 2016; Available from: http://www.mpoweruk.com/steam_turbines_html.

351. World Alliance for Decentralized Energy. 2006; Available from: http://www.localpower.org/deb_tech_gt.html.

352. Michailos, S., D. Parker, and C. Webb, Simulation Studies on Ethanol Production from Sugar Cane Residues. Industrial & Engineering Chemistry Research, 2016. 55: p. 5173-5179.

353. Sahinidis, N.V., Optimization under uncertainty: state-of-the-art and opportunities. Computers & Chemical Engineering, 2004. 28: p. 971-983.

354. Cheali, P., et al., Effect of Market Price Uncertainties on the Design of Optimal Biorefinery Systems—A Systematic Approach. Industrial & Engineering Chemistry Research, 2014. 53: p. 6021-6032.

355. Kim, J., M.J. Realff, and J.H. Lee, Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty. Computers & Chemical Engineering, 2011. 35: p. 1738-1751.

356. Damodaran, A. Tax Rates by Sector (U.S.). Debt Fundamentals 2017; Available from: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/taxrate.html.

357. Overend, R.P., The Average Haul Distance and Transportation Work Factors for Biomass Delivered to a Central Plant. Biomass, 1982. 2(1): p. 75-79.

358. Azar, C., K. Lindgren, and B.A. Andersson, Global energy scenarios meeting stringent CO2 constraints - cost-effective fuel choices in the transportation sector. Energy Policy, 2003. 31(10): p. 961-976.

359. Fthenakis, V., J.E. Mason, and K. Zweibel, The technical, geographical, and economic feasibility for solar energy to supply the energy needs of the US. Energy Policy, 2009. 37(2): p. 387-399.

360. Saber, A.Y. and G.K. Venayagamoorthy, Plug-in Vehicles and Renewable Energy Sources for Cost and Emission Reductions. Ieee Transactions on Industrial Electronics, 2011. 58(4): p. 1229-1238.

361. Thackeray, M.M., C. Wolverton, and E.D. Isaacs, Electrical energy storage for transportation-approaching the limits of, and going beyond, lithium-ion batteries. Energy & Environmental Science, 2012. 5(7): p. 7854-7863.

362. Oney, F., T.N. Veziroglu, and Z. Dulger, Evaluation of Pipeline Transportation of Hydrogen and Natural-Gas Mixtures. International Journal of Hydrogen Energy, 1994. 19(10): p. 813-822.

363. Sharma, S. and S.K. Ghoshal, Hydrogen the future transportation fuel: From production to applications. Renewable & Sustainable Energy Reviews, 2015. 43: p. 1151-1158.

364. Gonzalez, R., et al., Economics of cellulosic ethanol production in a thermochemical pathway for softwood, hardwood, corn stover and switchgrass. Fuel Processing Technology, 2012. 94(1): p. 113-122.

365. Katahira, S., et al., Ethanol fermentation from lignocellulosic hydrolysate by a recombinant xylose- and cellooligosaccharide-assimilating yeast strain. Applied Microbiology and Biotechnology, 2006. 72(6): p. 1136-1143.

366. Madhavan, A., et al., Alcoholic fermentation of xylose and mixed sugars using recombinant Saccharomyces cerevisiae engineered for xylose utilization. Applied Microbiology and Biotechnology, 2009. 82(6): p. 1037-1047.

367. Lindsay, S.E., R.J. Bothast, and L.O. Ingram, Improved strains of recombinant Escherichia coli for ethanol production from sugar mixtures. Appl Microbiol Biotechnol, 1995. 43(1): p. 70-5.

368. Beall, D.S., K. Ohta, and L.O. Ingram, Parametric studies of ethanol production form xylose and other sugars by recombinant Escherichia coli. Biotechnol Bioeng, 1991. 38(3): p. 296-303.

369. Olson, D.G., et al., Recent progress in consolidated bioprocessing. Curr Opin Biotechnol, 2012. 23(3): p. 396-405.

370. Lynd, L.R., et al., Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol, 2005. 16(5): p. 577-83.

371. Maclean, H.L. and L.B. Lave, Life cycle assessment of automobile/fuel options. Environmental Science & Technology, 2003. 37(23): p. 5445-5452.

372. Adekunle, A., V. Orsat, and V. Raghavan, Lignocellulosic bioethanol: A review and design conceptualization study of production from cassava peels. Renewable & Sustainable Energy Reviews, 2016. 64: p. 518-530.

373. Kumar, M., et al., Comparative economic assessment of ABE fermentation based on cellulosic and non-cellulosic feedstocks. Applied Energy, 2012. 93: p. 193-204.

374. Liu, X.-B., et al., Enhancement of butanol tolerance and butanol yield in Clostridium acetobutylicum mutant NT642 obtained by nitrogen ion beam implantation. Journal of Microbiology, 2012. 50: p. 1024-1028.

375. Bowles, L.K. and W.L. Ellefson, Effects of butanol on Clostridium acetobutylicum. Applied and Environmental Microbiology, 1985. 50: p. 1165-1170.

376. Kim, B., et al. Optimal design and operation of an extractive fermentation process for continuous biobutanol production. in 2014 14th International Conference on Control, Automation and Systems (ICCAS 2014). 2014.

377. Felgueira, D.M.S., Downstream processing and in situ product recovery techniques of isopropanol and butanol production from biomass. 2015.

378. Eom, M.-H., et al., Modeling of a Biobutanol Adsorption Process for Designing an Extractive Fermentor. Industrial & Engineering Chemistry Research, 2013. 52: p. 603-611.

379. Eom, M.-H., et al., Dynamic Modeling of a Fermentation Process with Ex situ Butanol Recovery (ESBR) for Continuous Biobutanol Production. Energy & Fuels, 2015. 29: p. 7254-7265.

380. Izak, P., et al., Increased productivity of Clostridium acetobutylicum fermentation of acetone, butanol, and ethanol by pervaporation through supported ionic liquid membrane. Applied Microbiology and Biotechnology, 2008. 78(4): p. 597-602.

381. Qureshi, N. and H.P. Blaschek, Butanol recovery from model solution/fermentation broth by pervaporation: evaluation of membrane performance. Biomass & Bioenergy, 1999. 17(2): p. 175-184.

382. Favre, E., Q.T. Nguyen, and S. Bruneau, Extraction of 1-butanol from aqueous solutions by pervaporation. Journal of Chemical Technology and Biotechnology, 1996. 65(3): p. 221-228.

383. Qureshi, N. and H.P. Blaschek, Production of acetone butanol ethanol (ABE) by a hyper-producing mutant strain of Clostridium beijierinckii BA101 and recovery by pervaporation. Biotechnology Progress, 1999. 15(4): p. 594-602.

384. Qureshi, N. and H.P. Blaschek, Recovery of butanol from fermentation broth by gas stripping. Renewable Energy, 2001. 22: p. 557-564.

385. Ezeji, T.C., et al., Improving performance of a gas stripping-based recovery system to remove butanol from Clostridium beijerinckii fermentation. Bioprocess and Biosystems Engineering, 2005. 27: p. 207-214.

386. Cai, D., et al., Gas stripping–pervaporation hybrid process for energy-saving product recovery from acetone–butanol–ethanol (ABE) fermentation broth. Chemical Engineering Journal, 2016. 287: p. 1-10.

387. Wei, C.J. and L.S. Chen, Dynamic analysis of mathematical model of ethanol fermentation with gas stripping. Nonlinear Dynamics, 2009. 57(1-2): p. 13-23.

388. Mayank, R., A. Ranjan, and V.S. Moholkar, Mathematical models of ABE fermentation: review and analysis. Critical Reviews in Biotechnology, 2013. 33(4): p. 419-447.

389. Xiao, H., et al., Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient Clostridium acetobutylicum for simultaneous utilization of glucose, xylose, and arabinose. Appl Environ Microbiol, 2011. 77(22): p. 7886-95.

390. Srivastava, A.K. and B. Volesky, Updated Model of the Batch Acetone-Butanol Fermentation. Biotechnology Letters, 1990. 12(9): p. 693-698.

391. Desai, R.P., L.K. Nielsen, and E.T. Papoutsakis, Stoichiometric modeling of Clostridium acetobutylicum fermentations with non-linear constraints. Journal of Biotechnology, 1999. 71(1-3): p. 191-205.

392. Yerushalmi, L., B. Volesky, and J. Votruba, Modeling of Culture Kinetics and Physiology for Clostridium-Acetobutylicum. Canadian Journal of Chemical Engineering, 1986. 64(4): p. 607-616.

393. Shinto, H., et al., Kinetic modeling and sensitivity analysis of acetone-butanol-ethanol production. Journal of Biotechnology, 2007. 131(1): p. 45-56.

394. Shinto, H., et al., Kinetic study of substrate dependency for higher butanol production in acetone-butanol-ethanol fermentation. Process Biochemistry, 2008. 43(12): p. 1452-1461.

395. Votruba, J., B. Volesky, and L. Yerushalmi, Mathematical-Model of a Batch Acetone Butanol Fermentation. Biotechnology and Bioengineering, 1986. 28(2): p. 247-255.

396. Harder, A. and J.A. Roels, Application of simple structured mode in bioengineering, in Microbes and Engineering Aspencts. 1982, Springer: Berlin, Heidelberg. p. 55-107.

397. Raftery, J.P., M.R. DeSessa, and M.N. Karim, Economic improvement of continuous pharmaceutical production via the optimal control of a multifeed bioreactor. Biotechnol Prog, 2017.

398. Ezeji, T.C., N. Qureshi, and H.P. Blaschek, Production of acetone butanol (AB) from liquefied corn starch, a commercial substrate, using Clostridium beijerinckii coupled with product recovery by gas stripping. Journal of Industrial Microbiology & Biotechnology, 2007. 34(12): p. 771-777.

399. Sander, R., Compilation of Henry's law constants (version 4.0) for water as solvent. Atmospheric Chemistry and Physics, 2015. 15(8): p. 4399-4981.

400. Andrew, A.N., Evaluation of Ethanol Stripping from Fermenters, in Department of Chemical Engineering and Chemical Technology. 2010, Imperial College London.

401. Nedeltchev, S. and A. Schumpe, New approaches for theoretical estimation of mass transfer parameters in both gas-liquid and slurry bubble columns. 2011: InTech Open Access Publisher.

402. Guo, Z. and N.F. Roache, Overall Mass Transfer Coefficient for Pollutant Emissions from Small Water Pools under Simulated Indoor Environmental Conditions. The Annals of Occupational Hygiene, 2003. 47(4): p. 279-286.

403. Hairer, E. and G. Wanner, Stiff differential equations solved by Radau methods. Journal of Computational and Applied Mathematics, 1999. 111(1-2): p. 93-111.

404. Wachter, A. and L.T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Mathematical Programming, 2006. 106(1): p. 25-57.

405. Raftery, J.P. and M.N. Karim, Economic viability of consolidated bioprocessing utilizing multiple biomass substrates for commercial-scale cellulosic bioethanol production. Biomass and Bioenergy, 2017. 103: p. 35-46.

406. Kujawska, A., et al., ABE fermentation products recovery methods-A review. Renewable & Sustainable Energy Reviews, 2015. 48: p. 648-661.

407. van der Merwe, A.B., et al., Comparison of energy efficiency and economics of process designs for biobutanol production from sugarcane molasses. Fuel, 2013. 105: p. 451-458.

408. Sanchez-Ramirez, E., et al., Process Alternatives for Biobutanol Purification: Design and Optimization. Industrial & Engineering Chemistry Research, 2015. 54(1): p. 351-358.

409. Strobel, M.K. and J.B. Bader, Economic evaluation of neutral-solvents fermentation product separation. 1981, ; Oak Ridge National Lab., TN (USA); Massachusetts Inst. of Tech., Oak Ridge, TN (USA). School of Chemical Engineering Practice. p. Medium: X; Size: Pages: 63.

410. Wali, M.M.N., Optimum Separation of Strongly Non-ideal Fermentation Products, in Department of Agricultural and Chemical Engineering. 1986, Colorado State University.

411. Green, E.M., et al., Genetic manipulation of acid formation pathways by gene inactivation in Clostridium acetobutylicum ATCC 824. Microbiology-Sgm, 1996. 142: p. 2079-2086.

412. Lutke-Eversloh, T. and H. Bahl, Metabolic engineering of Clostridium acetobutylicum: recent advances to improve butanol production. Current Opinion in Biotechnology, 2011. 22(5): p. 634-647.

413. Ezeji, T.C., N. Qureshi, and H.P. Blaschek, Bioproduction of butanol from biomass: from genes to bioreactors. Current Opinion in Biotechnology, 2007. 18(3): p. 220-227.

414. Lin, P.P., et al., Consolidated bioprocessing of cellulose to isobutanol using Clostridium thermocellum. Metabolic Engineering, 2015. 31: p. 44-52.

415. Salimi, F., K. Zhuang, and R. Mahadevan, Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. Biotechnology Journal, 2010. 5(7): p. 726-738.

416. Jouzani, G.S. and M.J. Taherzadeh, Advances in consolidated bioprocessing systems for bioethanol and butanol production from biomass: a comprehensive review. Biofuel Research Journal-Brj, 2015. 2(1): p. 152-195.

417. Gundersen, T. and L. Naess, The Synthesis of Cost Optimal Heat-Exchanger Networks - an Industrial Review of the State of the Art. Computers & Chemical Engineering, 1988. 12(6): p. 503-530.

418. Hohmann, E.C., Optimum Networks for Heat Exchange. 1971, University of Southern California. p. 206.

419. Linnhoff, B. and J.R. Flower, Synthesis of heat exchanger networks: II. Evolutionary generation of networks with various criteria of optimality. AIChE Journal, 1978. 24(4): p. 642-654.

420. Hasan, M.M.F., et al., Synthesis of heat exchanger networks with nonisothermal phase changes. AIChE Journal, 2010. 56(4): p. 930-945.

421. Almutlaq, A.M., V. Kazantzi, and M.M. El-Halwagi, An algebraic approach to targeting waste discharge and impure fresh usage via material recycle/reuse networks. Clean Technologies and Environmental Policy, 2005. 7(4): p. 294-305.

422. El-Halwagi, M.M., Pollution prevention through process integration. Clean Products and Processes, 1998. 1: p. 5–19.

423. El-Halwagi, M.M., F. Gabriel, and D. Harell, Rigorous graphical targeting for resource conservation via material recycle/reuse networks. Industrial & Engineering Chemistry Research, 2003. 42(19): p. 4319-4328.

424. El-Halwagi, M.M. and V. Manousiouthakis, Synthesis of Mass Exchange Networks. Aiche Journal, 1989. 35(8): p. 1233-1244.

425. El-Halwagi, M.M. and V. Manousiouthakis, Automatic Synthesis of Mass-Exchange Networks with Single-Component Targets. Chemical Engineering Science, 1990. 45(9): p. 2813-2831.

426. Hallale, N. and F. Liu, Refinery hydrogen management for clean fuels production. Advances in Environmental Research, 2001. 6(1): p. 81-98.

427. Petela, E., R. Smith, and Y.P. Wang, Defining a Strategy for Fresh Water and Wastewater Minimization Using Water-Pinch Analysis. Water Use and Reuse, 1994. 49: p. 39-46.

428. Salary, R., et al., Design of Oil Refineries Hydrogen Network Using Process Integration Principles. Iranian Journal of Chemistry & Chemical Engineering-International English Edition, 2008. 27(4): p. 49-64.

429. Hasan, M.M.F., I.A. Karimi, and C.M. Avison, Preliminary Synthesis of Fuel Gas Networks to Conserve Energy and Preserve the Environment. Industrial & Engineering Chemistry Research, 2011. 50(12): p. 7414-7427.

430. Bandyopadhyay, S.F., Dominic C. Y.; Tan, Raymond R., Feeling the Pinch?, in Chemical Engineering Progress. 2016, AIChE. p. 46-49.

431. Biegler, L.T., I.E. Grossmann, and A.W. Westerberg, Systematic Methods of Chemical Engineering and Process Design. 1997, Upper Saddle River, NJ: Prentice Hall.

432. Peters, M.S. and K.D. Timmerhaus, Plant Design and Economics for Chemical Engineers. 5th ed. 2003, New York, NY: McGraw-Hill.

433. Ravindran, A., G.V. Reklaitis, and K.M. Ragsdell, Engineering optimization: methods and applications. 2006: John Wiley & Sons.

434. El-Halwagi, M.M., Process Integration, Volume 7 (Process Systems Engineering). 2006, Orlando, FL: Academic Press, Inc.

435. Edgar, T., D. Himmelblau, and L. Lasdon, Optimization of Chemical Processes. 2001. McGraw-Hill Book Co.

436. Biegler, L.T., I.E. Grossmann, and A.W. Westerberg, Issues and Trends in the Teaching of Process and Product Design. Aiche Journal, 2010. 56(5): p. 1120-1125.

437. Klemes, J.J., et al., Advanced multimedia engineering education in energy, process integration and optimisation. Applied Energy, 2013. 101: p. 33-40.

438. Perkins, J., Education in process systems engineering: past, present and future. Computers & Chemical Engineering, 2002. 26(2): p. 283-293.

439. Seider, W.D.L., Daniel R.; Seader, J.D.; Widagdo, Soemantri; Gani, Rafiqul; Ng, Ka Ming, Product and Process Design Principles: Synthesis, Analysis and Design. 4th ed. 2016: Wiley.

440. Kiss, A.A., J. Grievink, and M. Rito-Palomares, A systems engineering perspective on process integration in industrial biotechnology. Journal of Chemical Technology and Biotechnology, 2015. 90(3): p. 349-355.

441. Manan, Z.A., Y.L. Tan, and D.C.Y. Foo, Targeting the minimum water flow rate using water cascade analysis technique. AIChE Journal, 2004. 50: p. 3169-3183.

442. Almutlaq, A.M. and M.M. El-Halwagi, An algebraic targeting approach to resource conservation via material recycle/reuse. International Journal of Environment and Pollution, 2007. 29(1-3): p. 4-18.

443. Almutlaq, A.M., Algebraic Approaches to Resource Conservation via Process Integration, in Chemical Engineering. 2005, Texas A&M University. p. 124.

444. Dogaris, I., D. Mamma, and D. Kekos, Biotechnological production of ethanol from renewable resources by Neurospora crassa: an alternative to conventional yeast fermentations? Appl Microbiol Biotechnol, 2013. 97(4): p. 1457-73.

445. Dias, M.O.S., et al., Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to conventional bioethanol production process. Chemical Engineering Research and Design, 2009. 87(9): p. 1206-1216.

446. Cardona, C.A. and O.J. Sanchez, Fuel ethanol production: process design trends and integration opportunities. Bioresour Technol, 2007. 98(12): p. 2415-57.

447. Čuček, L., et al., Energy, water and process technologies integration for the simultaneous production of ethanol and food from the entire corn plant. Computers & Chemical Engineering, 2011. 35(8): p. 1547-1557.