The accuracy of linear flux models in predicting reaction rate profiles in a model batch biochemical reaction system

By:

Alistair Paul Hughes

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Engineering (Chemical)

February 2014

Department of Chemical Engineering

University of Cape Town

Rondebosch, 7701

Republic of South Africa
The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.
Abstract

Metabolic flux analysis is commonly used in the modelling of biochemical reactions. The use of MFA models has gained large amounts of interest due to the simplicity of the computational procedures required for the model, and the exclusion of difficult to measure intracellular reaction data. There are many examples of the use of MFA models in literature studies in a number of applications, ranging from the medical industry through to the development of novel biochemical processes. Little to no mention is provided in literature studies regarding the applicability of the MFA model to a specified set of reaction data. Furthermore, the techniques and routines used to compute the flux models are not well described in these studies. The objectives of this research were to determine the sensitivity of the MFA models to various operating and kinetic parameters and to highlight the considerations required when setting up the computational routine used to solve the flux balances.

The study was conducted using a model pathway populated with a set of hypothetical elemental reactions and branch points. The model pathway was used in this study to negate the affects of complex regulatory biochemical architectures which are not well described in literature. The use of the model pathway ensured that the reaction system was thermodynamically feasible and there was consistency in the mass balances. The exclusion of the complex regulatory reactions did not affect the accuracy of the results generated in this study. A set of reaction mechanisms were used to describe each reaction step and were populated with parameters reference from literature. The cellular and reactor mass balances were generated using correlations presented in literature.

The results generated by the model system were validated using experimental data obtained from literature. The results clearly showed that the use of linear flux models can provide accurate estimations of the reaction fluxes, regardless of the non-linearity of the reaction data. However, the accuracy of the estimated reaction fluxes was dependent on how the flux model was populated with reaction data. The
results clearly showed that the design of experimental MFA studies should be conducted in such a way as to ensure that computational errors are minimised. The use of poorly considered flux constraints resulted in large degrees of linear dependency in the reaction matrices and high MFA model errors. Evidence was provided to show that many literature studies have not considered the conditioning of the reaction matrix in the design of the experiments or the computational procedure used. The results from this study also clearly showed that the change out of the linear solver cannot overcome inherent computational errors for a poorly conditioned system. Furthermore, the approach used in setting up the inequality and equality constraint matrices for a linear programming routines had a significant impact on the overall computational error and the subsequent accuracy of the estimated reaction fluxes.

Results were generated to illustrate the sensitivity of the model errors to various operating and kinetic parameters. It was found that the MFA model errors were very sensitive to the perturbation of the initial substrate concentrations, the relative substrate concentrations, the equilibrium constant ($K_{Eq}$), the inhibition constant ($K_i$) and the MM constant ($K_M$). The results showed that there was no sensitivity of the model errors to the reaction rate constant ($k$). The form of the kinetic mechanism also affected the degree of sensitivity of the model error to the parameter perturbations. The presence of inhibition terms in the reaction mechanism resulted in lower model errors, whereas the presence of the saturation terms in the reaction mechanisms increased the model errors.

The MFA model suffered from severe model error at low substrate conversions. The results indicated that this was true under all conditions. No accurate estimation of the reaction rates were obtained at substrate conversions less than 5%. A set of qualitative guidelines have been provided to guide the design of experimental studies for the population of an MFA model, and the procedure for the computation of the model. The results concluded that linear flux models can provide accurate estimations of the reaction fluxes using a simplified linear model. The results do however show that the accuracy of the MFA model was not guaranteed under all conditions and that in some instances, the simplicity of the MFA model was its own downfall. It has been recommended that a structured kinetic model be used in instances where large perturbations in the parameter space are required, and the uptake of metabolites is to be maximised.
Acknowledgements

I would like to thank SASOL for their financial support during my full-time studies and motivation to complete the thesis part-time whilst working permanently in Secunda. I would also like to thank the C-Change research centre for their financial support during my studies at UCT. Thanks also goes to my supervisors for their financial support and the opportunities to attend and present at many great conferences. I would like to acknowledge the administrative and personal support I have been given at UCT, especially from Sue Jobson and Fran Pocock. Their support and guidance was greatly appreciated.

I would like to thank my friends and family for their continued support and motivation to complete the thesis. It has been a comfort to know that my decisions and actions have always been well supported. I would also like to thank my fiancé for her continued support and love. It has been a long and difficult journey, but your support has always been resolute and very encouraging.
Declaration

I know the meaning of plagiarism and declare that all the work in the document, save for that which is properly acknowledged, is my own.
Contents

List of Tables ix

List of Figures xii

I Introduction, Literature Review and Research Methodology 1

1 Review of literature and research studies 2

1.1 Introduction 2

1.2 Review of MFA models 6

1.2.1 Intrinsic flux balances and model development 6

1.2.2 The use of pseudo-steady state approximations in reaction rate analyses 8

1.3 Computation of MFA models 14

1.3.1 Review of standard linear algebra routines 15

1.3.2 Review of linear programming routines 19

1.4 Scope and objectives 23
2 Research Methodology

2.1 Determination of cellular reaction rates

2.1.1 Elemental reactions

2.1.1.1 Reaction pathways

2.1.1.2 Kinetic reaction mechanisms

2.2 Cellular and reactor mass balances

2.3 Sensitivity analysis of kinetic and operating parameters

2.4 Metabolic flux simulations

2.4.1 Constraint flux calculation

2.4.2 Linearised reaction flux models

2.4.3 Implementation of numerical linear solvers for flux simulations

2.4.4 Error analysis for flux simulations

2.5 Computation details

2.6 Research approach

2.6.1 Batch flux simulations

2.6.2 Steady state flux simulations

II Results and Discussion

3 Batch reaction profiles

3.1 Introduction

3.2 Methodology
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1</td>
<td>Computation of the kinetic reaction profiles and estimated reaction fluxes</td>
<td>57</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Computation of the reaction fluxes using various constraint sets</td>
<td>58</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Computation of the reaction fluxes using various linear solvers</td>
<td>60</td>
</tr>
<tr>
<td>3.3</td>
<td>Results and Discussion</td>
<td>64</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Reaction rate profiles for reaction pathway DF\textsubscript{1}</td>
<td>64</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Reaction rate profiles for reaction pathway DF\textsubscript{2}</td>
<td>77</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Reaction rate profiles for reaction pathway DF\textsubscript{3}</td>
<td>86</td>
</tr>
<tr>
<td>3.4</td>
<td>Concluding Remarks</td>
<td>91</td>
</tr>
<tr>
<td>4</td>
<td>Sensitivity of cellular dilution rates in the cellular mass balances</td>
<td>93</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>93</td>
</tr>
<tr>
<td>4.2</td>
<td>Methodology</td>
<td>95</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and Discussion</td>
<td>97</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Kinetic mechanism Kin\textsubscript{1} (Michaelis-Menten with non-competitive inhibition)</td>
<td>97</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Kinetic mechanism Kin\textsubscript{2} (Michaelis-Menten with competitive inhibition)</td>
<td>105</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Kinetic mechanism Kin\textsubscript{3} (Michaelis-Menten with no inhibition)</td>
<td>110</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Kinetic mechanism Kin\textsubscript{4} (First order)</td>
<td>114</td>
</tr>
<tr>
<td>4.4</td>
<td>Concluding Remarks</td>
<td>119</td>
</tr>
<tr>
<td>5</td>
<td>MFA model errors</td>
<td>121</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>121</td>
</tr>
<tr>
<td>5.2</td>
<td>Methodology</td>
<td>123</td>
</tr>
</tbody>
</table>
5.3 Results and Discussion

5.3.1 Kinetic mechanism Kin_1 (Michaelis-Menten with non-competitive inhibition) 125

5.3.2 Kinetic mechanism Kin_2 (Michaelis-Menten with competitive inhibition) 136

5.3.3 Kinetic mechanism Kin_3 (Michaelis-Menten with no inhibition) 140

5.3.4 Kinetic mechanism Kin_4 (First order) 143

5.3.5 Steady state MFA models 145

5.4 Concluding Remarks 149

6 Conclusions 151

6.1 Computation of MFA models 151

6.2 Effect of the cellular dilution rates on the cellular mass balances 152

6.3 MFA model errors 152

Bibliography 154

III Appendices 161

A Reaction Rates 162

B Reaction Matrices 165

C Reactor and Cellular Mass Balances 169

D MATLAB Source Code 176
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Reaction mechanisms used to describe all reaction rates for all reaction pathways</td>
<td>35</td>
</tr>
<tr>
<td>2.2</td>
<td>Standard reaction kinetic parameters used with all reaction mechanisms for all reaction pathways</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>Standard average specific growth rates and molar fluxes obtained for reaction pathway $DF_1$ with the various kinetic mechanisms</td>
<td>36</td>
</tr>
<tr>
<td>2.4</td>
<td>Reactor mass balances for all modes of operation with all kinetic mechanism for all reaction pathways</td>
<td>38</td>
</tr>
<tr>
<td>2.5</td>
<td>Standard operating parameters used for various modes of operation for all reaction pathways</td>
<td>40</td>
</tr>
<tr>
<td>2.6</td>
<td>Range of values used for the sensitivity analysis on the kinetic parameters for all reaction pathways</td>
<td>41</td>
</tr>
<tr>
<td>2.7</td>
<td>Range of values used for the sensitivity analysis on the operating parameters for all reaction pathways</td>
<td>41</td>
</tr>
<tr>
<td>2.8</td>
<td>Example of the formulae used to calculate the extracellular flux constraints for all modes of operation for all reaction pathways</td>
<td>43</td>
</tr>
<tr>
<td>2.9</td>
<td>Selection of data analysis and the corresponding integration times for the computation of the extracellular reaction fluxes for all reaction pathways</td>
<td>43</td>
</tr>
<tr>
<td>2.10</td>
<td>Linear programming model parameters for all reaction pathways</td>
<td>48</td>
</tr>
</tbody>
</table>
### List of Tables

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.11</td>
<td>Movement of variables in the computational model</td>
<td>51</td>
</tr>
<tr>
<td>2.12</td>
<td>References for all procedures used to obtain the results for the batch flux simulations</td>
<td>52</td>
</tr>
<tr>
<td>2.13</td>
<td>References for all procedures used to obtain the results for the steady state flux simulations</td>
<td>53</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of possible constraint sets for each reaction pathway</td>
<td>58</td>
</tr>
<tr>
<td>3.2</td>
<td>Constraint sets used for pathway DF₂</td>
<td>59</td>
</tr>
<tr>
<td>3.3</td>
<td>Equations used for the reaction pathway DF₂ with the constraint sets 3 and 5</td>
<td>59</td>
</tr>
<tr>
<td>3.4</td>
<td>Equations used for linear algebra routines with reaction pathway DF₃ and constraint set 5</td>
<td>60</td>
</tr>
<tr>
<td>3.5</td>
<td>Equations used for the linear programming routine with reaction pathway DF₃ and constraint set 5</td>
<td>61</td>
</tr>
<tr>
<td>3.6</td>
<td>Constraint sets used for pathway DF₃</td>
<td>62</td>
</tr>
<tr>
<td>3.7</td>
<td>Summary of reaction data obtained at standard conditions</td>
<td>64</td>
</tr>
<tr>
<td>3.8</td>
<td>Summary of reaction data obtained at standard conditions for each reaction pathway with the Kin₁ mechanism</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>Computation of the relative cellular dilution rates for each intracellular metabolite in reaction pathway DF₁</td>
<td>95</td>
</tr>
<tr>
<td>A.1</td>
<td>Reaction rates used in pathway DF₁ (Kin₁ and Kin₂ mechanisms)</td>
<td>162</td>
</tr>
<tr>
<td>A.2</td>
<td>Reaction rates used in pathway DF₁ (Kin₃ and Kin₄ mechanisms)</td>
<td>163</td>
</tr>
<tr>
<td>A.3</td>
<td>Reaction rates used in pathway DF₂ and DF₃</td>
<td>164</td>
</tr>
<tr>
<td>B.1</td>
<td>Equations used for the reaction pathway DF₁</td>
<td>166</td>
</tr>
<tr>
<td>B.2</td>
<td>Equations used for the reaction pathway DF₂ (Constr sets 1 - 5)</td>
<td>167</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>B.3</td>
<td>Equations used for the reaction pathway DF₂ (Constr sets 6 - 10)</td>
<td>168</td>
</tr>
<tr>
<td>C.1</td>
<td>Extracellular mass balances for the batch process with reaction pathway DF₁</td>
<td>171</td>
</tr>
<tr>
<td>C.2</td>
<td>Extracellular mass balances for the continuous process with reaction pathway DF₁</td>
<td>171</td>
</tr>
<tr>
<td>C.3</td>
<td>Intracellular mass balances for reaction pathway DF₁</td>
<td>172</td>
</tr>
<tr>
<td>C.4</td>
<td>Extracellular mass balances for the batch process with reaction pathway DF₂</td>
<td>173</td>
</tr>
<tr>
<td>C.5</td>
<td>Intracellular mass balances for reaction pathway DF₂</td>
<td>173</td>
</tr>
<tr>
<td>C.6</td>
<td>Extracellular mass balances for the batch process with reaction pathway DF₃</td>
<td>174</td>
</tr>
<tr>
<td>C.7</td>
<td>Intracellular mass balances for reaction pathway DF₃</td>
<td>175</td>
</tr>
</tbody>
</table>
# List of Figures

2.1 Set of hypothetical reaction pathways used ........................................... 34

2.2 Computational model flow diagram ......................................................... 50

3.1 Graphical representation of the methodologies .......................................... 63

3.2 Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for models A, B and C .......................................................... 69

3.3 Reaction rate profiles obtained with the Kin\textsubscript{2} mechanism under standard conditions for models A, B and C .......................................................... 71

3.4 Reaction rate profiles obtained with the Kin\textsubscript{3} mechanism under standard conditions for models A, B and C .......................................................... 74

3.5 Reaction rate profiles obtained with the Kin\textsubscript{4} mechanism under standard conditions for models A, B and C .......................................................... 76

3.6 Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for model C using constraint sets 1 and 2 ........................................ 79

3.7 Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for model C using constraint sets 3 and 4 ........................................ 80

3.8 Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for model C using constraint sets 5 and 6 ........................................ 81
3.9 Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using constraint sets 7 and 8 .................................................. 82

3.10 Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using constraint sets 9 and 10 .................................................. 83

3.11 Condition numbers for all reaction matrices for reaction pathways DF$_1$, DF$_2$ and DF$_3$ . . . 85

3.12 Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using linear algebra routines .................................................. 87

3.13 Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using linear programming routines (LP$_1$ and LP$_2$) ......................... 89

3.14 Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using linear programming routines (LP$_3$ and LP$_4$) ......................... 90

4.1 Methodology used for the sensitivity analysis of the relative cellular dilution rates 96

4.2 Relative cellular dilution terms obtained as a function of the operating parameters with the Kin$_1$ mechanism ................................................................. 100

4.3 Relative cellular dilution terms obtained as a function of the kinetic parameters with the Kin$_1$ mechanism ................................................................. 103

4.4 Summary of the sensitivities of the relative cellular dilution rates with the Kin$_1$ mechanism ................................................................. 105

4.5 Relative cellular dilution terms obtained as a function of the operating parameters with the Kin$_2$ mechanism ................................................................. 106

4.6 Relative cellular dilution terms obtained as a function of the kinetic parameters with the Kin$_2$ mechanism ................................................................. 108

4.7 Summary of the sensitivities of the relative cellular dilution rates with the Kin$_2$ mechanism ................................................................. 109
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8</td>
<td>Relative cellular dilution terms obtained as a function of the operating parameters with the Kin(3) mechanism</td>
<td>111</td>
</tr>
<tr>
<td>4.9</td>
<td>Relative cellular dilution terms obtained as a function of the kinetic parameters with the Kin(3) mechanism</td>
<td>112</td>
</tr>
<tr>
<td>4.10</td>
<td>Summary of the sensitivities of the relative cellular dilution rates with the Kin(3) mechanism</td>
<td>114</td>
</tr>
<tr>
<td>4.11</td>
<td>Relative cellular dilution terms obtained as a function of the operating and kinetic parameters with the Kin(4) mechanism</td>
<td>116</td>
</tr>
<tr>
<td>4.12</td>
<td>Summary of the sensitivities of the relative cellular dilution rates with the Kin(4) mechanism</td>
<td>118</td>
</tr>
<tr>
<td>5.1</td>
<td>Methodology used for the evaluation of the MFA model errors under various conditions for the DF(_1) reaction pathway</td>
<td>124</td>
</tr>
<tr>
<td>5.2</td>
<td>Relative errors for model C with the Kin(_1) mechanism as a function of (C_B(0))</td>
<td>128</td>
</tr>
<tr>
<td>5.3</td>
<td>Relative accumulation rate profiles with the Kin(_1) mechanism</td>
<td>130</td>
</tr>
<tr>
<td>5.4</td>
<td>Absolute accumulation rate profiles with the Kin(_1) mechanism</td>
<td>131</td>
</tr>
<tr>
<td>5.5</td>
<td>Relative errors for model C with the Kin(<em>1) mechanism as a function of (K</em>{Eq}) and (K_i)</td>
<td>134</td>
</tr>
<tr>
<td>5.6</td>
<td>Relative errors for model C with the Kin(_1) mechanism as a function of (K_M) and (k)</td>
<td>135</td>
</tr>
<tr>
<td>5.7</td>
<td>Relative errors for model C with the Kin(<em>2) mechanism as a function of (C_B(0)) and (K</em>{Eq})</td>
<td>137</td>
</tr>
<tr>
<td>5.8</td>
<td>Relative errors for model C with the Kin(_2) mechanism as a function of (K_i) and (K_M)</td>
<td>139</td>
</tr>
<tr>
<td>5.9</td>
<td>Relative errors for model C with the Kin(_2) mechanism as a function of (k)</td>
<td>140</td>
</tr>
<tr>
<td>5.10</td>
<td>Relative errors for model C with the Kin(<em>3) mechanism as a function of (C_B(0)) and (K</em>{Eq})</td>
<td>141</td>
</tr>
<tr>
<td>5.11</td>
<td>Relative errors for model C with the Kin(_3) mechanism as a function of (K_M) and (k)</td>
<td>142</td>
</tr>
<tr>
<td>5.12</td>
<td>Relative errors for model C with the Kin(<em>4) mechanism as a function of (C_B(0)) and (K</em>{Eq})</td>
<td>144</td>
</tr>
</tbody>
</table>
5.13 Relative errors for model C with the Kin4 mechanism as a function of k . . . . . . . . 145

5.14 Steady state errors as a function of the dilution rate with all kinetic mechanisms . 147

5.15 Steady state cellular dilution rates as a function of the reactor dilution rate with all kinetic mechanisms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 148
Nomenclature

Abbreviations

CSTR  Contious Stirred Tank Reactor
FBA   Flux Balance Analysis
GC-MS  Gas Chromatography - Mass Spectroscopy
LP    Linear programming
MFA   Metabolic Flux Analysis
MILP  Linear mixed-integer programming
NMR   Nuclear Magnetic Resonance
ODE   Ordinary Differential Equation
PSS   Pseudo-steady state
PSSH  Pseudo-steady state hypothesis
SVD   Singular value decomposition

Greek Symbols

$\mu$  Specific growth rate
$\nu$  Reaction Flux Vector
$\rho$  Cellular biomass density
$\varphi$  Chemical Potential
in vivo  Biochemical process which takes place inside of the cellular membrane

in vitro  Biochemical process which takes place outside of the cellular membrane

Mathematical Symbols

\( \hat{V} \)  Specific cellular volume [mL\(_{\text{cell}}\)/gDW]

\( \mu C_i \)  Cellular dilution rate of metabolite i [mmol/L\(_{\text{cells}}\).h]

\( \frac{dC_i}{dt} \)  Accumulation rate of metabolite i [mmol/L\(_{\text{cells}}\).h]

\( C^\text{ext}_i \)  Extracellular metabolite concentration of metabolite i [mmol/L]

\( C^\text{int}_i \)  Intracellular metabolite concentration of metabolite i [mmol/L\(_{\text{cell}}\)]

\( C_X \)  Biomass concentration [gDW biomass/L]

\( F_i \)  Metabolic flux i [mmol/gDW.h]

\( G^T \)  Stoichiometric matrix [-]

\( K_{Eq} \)  Equilibrium constant [L/L\(_{\text{cells}}\)]

\( K_i \)  Inhibition constant [mmol/L]

\( K_M \)  MM constant [mmol/L]

\( k_s \)  Reaction rate constant for first order mechanism [hr\(^{-1}\)]

\( n_i \)  Moles of component i [mmol]

\( \text{Rate}_i \)  Reaction rate i [mmol/L\(_{\text{cell}}\).hr]

\( \sum_{i} r_{ij} \)  Net rate of generation of the intracellular metabolite j by rate i [mmol/L\(_{\text{cell}}\).hr]

\( gDW \)  Grams of dry weight biomass [g]

\( k \)  Reaction rate constant [s\(^{-1}\)]

Roman Symbols

\( A \)  Matrix of variable co-efficients
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>Constraint vector</td>
</tr>
<tr>
<td>L</td>
<td>Lower triangular matrix</td>
</tr>
<tr>
<td>P</td>
<td>Permutation matrix</td>
</tr>
<tr>
<td>S</td>
<td>System Entorpy</td>
</tr>
<tr>
<td>T</td>
<td>System Temperature</td>
</tr>
<tr>
<td>U</td>
<td>Upper triangular matrix</td>
</tr>
<tr>
<td>x</td>
<td>Solution vector</td>
</tr>
</tbody>
</table>
Part I

Introduction, Literature Review and Research Methodology
Chapter 1

Review of literature and research studies

1.1 Introduction

For a biotechnology process to be commercially successful, it must be able to out compete traditional chemical technologies for the production of existing products or provide feasible technologies for the production of new products (Kim and Dale, 2004; Hatzimanikatis et al., 1996 and Schulze, 1995). Schulze (1995) states that the advent of techniques for in vitro DNA recombination has provided researchers with the tools to fully control and manipulate the biological world to achieve these goals. The author suggests that the positive impact of biotechnology on the production of commercial chemicals and agricultural products has not been as significant as it was predicted to be after the “euphoria” of early successes. A number of studies identified that performing modifications to a single reaction in a metabolic pathway or ad hoc manipulations to a cell’s genome did not have the desired affects on productivity and yields (Forster, 2003; Hatzimanikatis et al., 1998; Stephanopoulos et al., 1998; Schulze, 1995 and Bailey, 1991). It was discovered that an integrated approach was the solution, where systemic characteristics of the metabolic network are required to predict the cell’s behaviour to specified perturbations to its genome or physical environment (Bailey, 1991). The formulation of an optimal biochemical system therefore requires a large amount of biological data from all levels of metabolism which needs to be integrated and compiled using mathematical models (Forster, 2003; Hatzimanikatis et al., 1998 and Voit, 1992). Stephanopoulos and Vallino (1991) describe the concept of metabolic engineering as a combination of genetic engineering, microbial
physiology and mathematical models to obtain the optimal distribution of metabolite fluxes through a metabolic pathway to increase both the productivity and yield of desired compounds.

A large variety of mathematical models have been used for the simulation and optimisation of biochemical reaction pathways. These range from complex and data intensive non-linear kinetic models (Hynne et al., 2001 and Peretti and Bailey, 1986) to the simpler linearised flux balance models (Edwards et al., 2001 and Stephanopoulos et al., 1998). High level non-linear kinetic models require detailed reaction data and knowledge of the reaction mechanisms for the models to be accurately and fully populated (van Riel, 2006; Bornholdt, 2005 and Hatzimanikatis et al., 1996). It has been discovered that obtaining accurate *in vivo* kinetic parameters for enzyme catalysed reactions and information about the complex regulatory mechanisms in a metabolic pathway involves complex and exhaustive data generation and analyses (Lee et al., 2008; Bornholdt, 2005; Endy and Brent, 2001; Gombert and Nielsen, 2000 and Liao and Oh, 1999). To simplify the required analyses, linear models are used to describe the same biochemical pathway without the need for detailed kinetic and reaction regulation data (Hatzimanikatis et al., 1996). The linear mathematical models only require knowledge of the reaction stoichiometry as well as analysis of the input and output fluxes for the cells under set experimental conditions (Stephanopoulos et al., 1998 and Bailey, 1991). Computation of the branching of metabolite fluxes through a metabolic pathway under a set of experimental conditions can be achieved using Metabolic Flux Analysis (MFA).

Metabolic Flux Analysis or Flux Balance Analysis (FBA) has found considerable application as a modelling technique in Metabolic Engineering. Stephanopoulos et al. (1998) describe it as a powerful methodology for the determination of pathway fluxes. Yang et al. (2010); Lee et al. (2005); Stephanopoulos (1999); Nielsen (1998); Rizzi et al. (1997) and Vallino and Stephanopoulos (1994a) provide examples of practical MFA applications where it is used as the primary pathway analysis tool or forms part thereof. Kauffman et al. (2003) and Gombert and Nielsen (2000) provide summaries of a number of MFA examples along with key successes achieved in metabolic engineering by predicting constrained reaction fluxes using this methodology. Reed and Palsson (2003) go further to discuss the principle achievements of constraint based models in helping to predict the phenotypic behaviour of *Escherichia coli* through a decade of dedicated research. These are just a few examples that provide clear evidence of the importance of MFA in the field of metabolic engineering.
Although more sophisticated experimental procedures are becoming more common in the metabolic engineering research environment, they are often used in conjunction with MFA. The combined analysis can be used for cross validation of experimental data and metabolic models. In these instances, the simplicity of MFA is complemented by the accuracy of the experimental techniques to form a powerful research tool. Jamshidi and Palsson (2008); Lee et al. (2005) and Edwards et al. (2001) provide insight into the incorporation of MFA in higher level analyses of biochemical pathways and their resulting accuracies. This is achieved through the inclusion of well populated metabolic reaction, thermodynamic and metabolomic databases generated using either nuclear magnetic resonance (Ghosh et al. 2005) (NMR) or gas chromatographic mass spectroscopy (Dauner 2010 and Zamboni 2010) (GC-MS). Werf (2005) goes on to discuss the advantages and disadvantages associated with performing either a completely computational or experimental analysis of a metabolic reaction network. The author indicates that a purely computational analysis of a metabolic reaction network can be very useful in the pre-screening of biochemical hypotheses and can also be very beneficial in complementing experimental data sets for the generation of comprehensive network and regulatory models. However, Werf (2005) states that the this approach has many disadvantages as it is based on a closed data analysis approach where biological interactions are not taken into account. Furthermore it is claimed that the generation of metabolic models is laborious and can become very complex if the system is not well known and a high level of detail is required to describe the system.

MFA models can also be used to perform an “extreme” pathway analysis to generate feasible operating spaces of a given reaction pathway. Famili et al. (2005) and Schilling et al. (2001) discuss the use of carefully selected combinations of a number of stationary point pathway flux analyses to predict feasible and optimal operating spaces known as k-cones. These spaces are created by imposing constraints on the system, thereby leading to constraint based models. These constraints can be thermodynamic (Jamshidi and Palsson 2008), phenotypic (Edwards et al. 2001) or even regulatory (Covert and Palsson 2003). From these models, sets of feasible reaction regions can be predicted.

MFA has been used in a number of different contexts for an ever increasing range of research applications. The wide range of applications where these techniques have been used to study the reaction pathways in microbes stretch from industrial (Yang et al. 2010; Lee et al. 2008; Stephanopoulos 2007; Hua et al. 2001; Vallino and Stephanopoulos 1994a and Vallino and Stephanopoulos 1994b) through to medical applications (Henry et al. 2007; Balcarcel and Clark...
In many cases, the systems being studied have very different metabolic pathways and many are used with a range of substrates. However the generalised MFA hypotheses were applied across all systems with very little to no mention of their applicability to the given system being studied. Furthermore, the choice of numerical routines used to solve the systems were also not discussed in detail within the context of the reaction systems themselves. An assumption is therefore made that the linearised models used to perform the MFA are not affected by the system parameters (kinetic or operating) or the computational structure of the model. The objective of the research conducted in this study is to simulate hypothetical biochemical reaction pathways to determine the effects of the kinetic and operating parameters on the accuracies of the linearised reaction models. The hypothetical system was also used to deconstruct the elements of the MFA model to determine the sensitivity of the model to various operating and kinetic parameters. Insight has also been provided by the study into the selection and use of suitable numerical routines for the computation of MFA models.

Reviews have been conducted on the fundamentals of MFA models as well as the computational routines used to compute the reaction fluxes with the MFA models. These have been structured as critical reviews of the available literature to highlight the approaches used in the development of MFA models as well as the variations between the various numerical routines which can be used to solve the models. A critical review of the fundamentals of the MFA models was also required to illustrate the specificity of the models and the data used for their validation.
1.2 Review of MFA models

1.2.1 Intrinsic flux balances and model development

The fundamental formalism of MFA models is based on mass balances developed for all metabolites in the pathway of interest. All balances include generation and consumption terms which create a well described network of interconnected metabolite distributions within a metabolic pathway (See Eq. 1.1) (Stephanopoulos et al., 1998). Rizzi et al. (1997) discuss the development of such models for the study of glycolysis in *Saccharomyces cerevisiae*. The formulation described relates to previous research published by Fredrickson (1976). The focus of this research is on the development of intrinsic reaction rates that are controlled by the intracellular metabolite concentrations. This is important as it is the intracellular metabolite concentrations which drive the enzyme autocatalyzed reactions. This is a result of the separation of all reacting bio-materials by cellular membranes into compartments (Fredrickson, 1976). In the case of eukaryotes, the distribution of the reacting bio-materials can be between discrete membrane enclosed compartments within the cell (Stephanopoulos et al., 1998 and Bailey and Ollis, 1986). However, whether the biological system is eukaryotic or prokaryotic, (Fredrickson, 1976) emphasises that the reacting bio-material must always be compartmentalised into the biological reaction volume and not the extracellular reaction media. The author describes the accumulation of mass within a metabolic reaction pathway mathematically as:

\[
\frac{d}{dt}(\hat{V} C_j) = \hat{V} \sum_i r_{ij} 
\]  

(1.1)

which can be simplified to

\[
\frac{dC_j}{dt} = \sum_i r_{ij} - \mu C_j
\]

where

\[
\mu = \frac{1}{m} \frac{dm}{dt}
\]

represents the specific growth rate. \(\hat{V}\) represents the volume of bio-material per unit of biomass (described as constant for a given microbial system), \(m\) represents the biomass of the system, \(C_j\) represents the mass of the \(j\)th metabolite per unit of volume of bio-material, which is intrinsic to the intracellular reaction volume, and \(\sum_i r_{ij}\) represents the net rate of generation of metabolite \(j\) per unit
1.2. REVIEW OF MFA MODELS

Fredrickson (1976) states that the inclusion of the cellular dilution rate ($\mu c_i$) is very important as it takes into account the effect that cellular growth has on depleting the pool of intracellular metabolites. Rizzi et al. (1997) utilise the models described by Fredrickson (1976) and make clear reference to this fact. Their balances include all the intrinsic reaction and dilution terms described by Fredrickson (1976), and they go even further to separate the cellular reaction volume into different compartments. The inclusion of reaction terms, intrinsic to the biological reaction volume in the continuous stirred tank reactor (CSTR) (intrinsic to the extracellular reaction media), by Rizzi et al. (1997) was made possible by multiplying the reaction terms by $(C_x/\rho)$. The authors define $C_x$ as the biomass concentration in the reactor [gDW/L reactor] and $\rho$ as the specific volume [L cell/gDW]. However, upon evaluation of the units of this equation, the $\rho$ term is more likely representative of the cellular biomass density, which is simply the inverse of the specific volume described, giving the units [gDW/L cell]. The inclusion of this term by Rizzi et al. (1997) allows for the incorporation of the cellular reaction rate within the total reactor balances. This conforms to the approach stipulated by Fredrickson (1976) to avoid inconsistencies which arose from reactor balances in previously published research studies. These inconsistencies occurred when the intrinsic nature of the cellular rates were not properly accounted for in the reactor balances. The importance of the intrinsic nature of the reaction terms is also described by Stephanopoulos et al. (1998). The authors clearly state that the incorrect use of the reaction rates in either a cellular or reactor balance will result in significant inconsistencies. They however advocate the use of specific reaction rates based on the biomass concentration rather than the cellular volumetric rates, as it allows for easier comparison across experiments. This conversion can be easily achieved through the use of the cellular biomass density (Stephanopoulos et al. 1998). Stephanopoulos et al. (1998) indicate that this can be achieved when the cellular biomass density is in the range of 1g cell/mL cell and the water content of the cell 0.67 mL water/mL cell.

By converting the intracellular concentrations and rates obtained for the glycolysis in S. cerevisiae by Theobald et al. (1997) and Rizzi et al. (1997), to specific variables intrinsic to the grams of dry weight biomass (gDW), Stephanopoulos et al. (1998) were able to determine that the dilution rate ($\mu c_i$) present in the intracellular mass balance can be considered to be negligible. They state:

*Because the intracellular level of most pathway metabolites is very low, the dilution effect*

---

7 February 2014
is generally small, especially when compared with the other fluxes affecting the same metabolite.

The data provided by Stephanopoulos et al. (1998) indicates that the dilution rates for the intracellular metabolites are 10,000 times smaller than the reaction fluxes through these pools. In the research conducted by Theobald et al. (1997) and Rizzi et al. (1997), the cellular specific volume used is stated to be 2.38 mL cell/gDW. This corresponds to a dry weight cellular density (ca. 3mL cell/gDW) which Stephanopoulos et al. (1998) state is acceptable to allow for easy conversion of cellular reactions and concentrations to specific quantities.

Nielsen and Villadsen (1994) warn about applying the assumption that for all systems the dilution rate is negligible. They state that the metabolite pools in some systems are of a size which is not insignificant compared with their formation and consumption rates, and therefore careful consideration must be made for each case. Furthermore, Stephanopoulos et al. (1998) also indicate that the levels of some cellular dilution terms may have small effects on the model errors for some systems, and can be accounted for in the metabolite balances as a separate rate. Furthermore, research conducted by Fredrickson (1976) suggests that the inclusion of the dilution term in the intrinsic mass balances is significant and should not be neglected without consideration of the specific system.

1.2.2 The use of pseudo-steady state approximations in reaction rate analyses

A crucial requirement for the simplification of a set of metabolic balances to the conventional MFA balances, described by Stephanopoulos et al. (1998) and Bailey (1991), is that the reaction system is at steady state. It is noted by Stephanopoulos et al. (1998) that if the system is not under steady state conditions, the pseudo-steady state hypothesis (PSSH) can be applied. The authors indicate that the unsteady state operation may be due to a perturbation to the steady state system or from the operation of batch or fed-batch systems. Nielsen and Villadsen (1994) indicate that the time scale for a change in the size of metabolite pools due to an extracellular perturbation is in fact significant. They indicate that these changes are only considered insignificant if compared to the time scale for cellular growth. Stephanopoulos et al. (1998) state:
1.2. REVIEW OF MFA MODELS

It is generally accepted that there is very high turnover of the pools of most metabolites. As a result, the concentrations of the different metabolite pools rapidly adjust to new levels, even after large perturbations in the environment experienced by the cells. It is therefore reasonable to assume that the pathway metabolites are at a pseudo-steady state.

This is due to the fact that a set of enzymatic reactions will provide faster response to new environmental conditions than that of the cellular growth rates. This allows one to simplify the system dramatically, providing a set of linear balances with a zero vector to describe the combined net accumulation and cellular dilution terms for each metabolite in the reaction pathway. Stephanopoulos et al. [1998] provide the following steps in the formulation of the simplified MFA balances which have been utilised in a number of metabolic engineering studies.

\[
\frac{dX_{\text{met}}}{dt} = r_{\text{met}} - \mu X_{\text{met}} \tag{1.2}
\]

Eq. 1.2 represents the system in the same form as that described by Eq. [1.1] only differing by the nomenclature used. When evaluating the system at steady state or by applying the pseudo-steady state hypothesis the following simplification can be made:

\[
0 = r_{\text{met}} - \mu X_{\text{met}} \tag{1.3}
\]

Further simplifications are made by neglecting the dilution rate \( \mu X_{\text{met}} \), yielding

\[
0 = r_{\text{met}} = G^T \nu \tag{1.4}
\]

which can be used to calculate the reaction flux vector \( \nu \), using the stoichiometric matrix which is represented by \( G^T \). A significant benefit identified by Stephanopoulos et al. [1998] is that simple linear algebra can be used to solve the flux vector. This is provided that sufficient constraints are imposed to make the set of equations determined. They go on to indicate that linear programming routines can also be employed to solve for the flux vector for underdetermined systems, where a sufficient number of constraints are not available. It is therefore fairly simple to simulate the reaction flux profiles of a given metabolic pathway, whether the system is determined or even underdetermined. The only requirement is a balanced reaction stoichiometry and some flux constraints which can be
determined experimentally as indicated in research conducted by Price et al. (2003); Schilling et al. (2001); Stephanopoulos (1999); Nielsen (1998); Vallino and Stephanopoulos (1994a,b) among many others.

If the generalised application of the PSSH used in MFA is analysed from a process engineering perspective, its validation and applicability can be viewed using a fundamental reaction engineering approach. If one was to view a metabolic reaction pathway from the perspective of process dynamics, it can be argued that it is very difficult to assume that the turnover rates for all metabolite pools, which can be compared to a process response or relaxation time, result in no observable system dynamics. Seborg et al. (2004) describe how the response time of a first-order process under feedback control can be significantly affected by the systems dynamics and the specific controlled response to this change. Furthermore, if the process has second-order dynamics, overshoot of the system response can have a significant effect on the ability of the process to achieve a rapid and stable shift to a new steady state. This may seem to be an irrelevant scenario when it comes to analysing metabolic reaction pathways, but the similarities are quite distinct. Like any given process, a biological reaction network will have its own dynamics. These dynamics are governed by the kinetic capabilities of a network to metabolise certain metabolites. Furthermore the feedback control loop of a process has similar analogues to that of the pheno- and genotypic control of a metabolic reaction pathway discussed in detail by Stephanopoulos et al. (1998). Therefore, in a similar manner, the dynamics of shifts in the steady state operation, as a result of any perturbation to the system, can not be simply neglected without a good understanding of how these are affected by the reaction kinetics and regulatory networks.

The use of first order response models for biological systems is discussed by Stephanopoulos et al. (1998). The authors indicate that deviations around a given steady state operating point can be modelled with linearised deviation models which are dependent on the relaxation times for the reactions. However, it is important to note that the dynamics described are a result of shifts between two steady state operations and not that obtained for a batch reaction system. Stephanopoulos et al. (1998) discuss how the lower relaxation times of the intracellular reactions result in a faster shift to a new steady state within the cell as a result of perturbations to the external environment. They indicate that if the relaxation time of the cellular reactions is three times lower than that of the extracellular reactions, then the dynamics of the cellular reactions can be considered negligible with respect to that of the cellular growth.
Van Santen and Niemantsverdriet (1995) discuss this principle of rate limiting analysis and the application of PSS to chemical kinetic mechanisms in detail. They indicate that if all reactions in an observed network move within a similar time scale, the PSSH is difficult to apply because the rate determining step no longer controls the observed shift in reaction intermediates through the reaction mechanisms. They describe the use of the PSSH as being valid only for cases where the observed rates in consecutive reactions differ by a factor of 100. Under these conditions, the faster reactions in a given mechanism are said to be in pseudo-steady state. Stephanopoulos et al. (1998) indicate however that the pseudo-steady state assumptions hold if the relaxation time of a set of reactions is 10 times higher than the relaxation time of the system.

The application of the PSSH allows for the computation of the effective reactant concentrations in terms of measurable variables. However once these have been obtained, the overall reaction rate is still viewed within the context of dynamic mass balances (van Santen and Niemantsverdriet, 1995). In other words, the accumulation terms of all the species are not neglected. Furthermore, the authors state that each application of the PSS must be carefully considered. If not, this could lead to inconsistent mass balances, due to the accumulation of mass within the dynamic intermediate pools. Theobald et al. (1997) observed from laboratory studies that relative changes in the extra- and intracellular concentrations of pyruvate, due to a pulse in the steady state glucose concentration, occurred within the same time scales. These observations suggest that the application of the PSSH to the intracellular pyruvate balance under these conditions may have led to an inconsistent mass balance as the accumulation terms for pyruvate were not taken into account.

Nielsen and Villadsen (1994) indicate that the rate of biomass production in a biochemical reaction pathway is slow enough to view all the other species as being at pseudo-steady state. However, computation of the extracellular rates with the pseudo-steady state balances will require the use of some dynamics to account for the accumulation of biomass and its constituents in the system. If this is not done carefully, there can be a significant propagation of model error through the reaction pathway and lead to low accuracies for the estimation of the reaction rates. Because of these inconsistencies, the MFA of a given batch reaction system is usually applied to only a few of the data points obtained from an entire reaction data set (Vallino and Stephanopoulos, [1994a,b]). The choice of suitable regions within the reaction profile where the MFA assumptions are appropriate is highly dependent on the dynamics of the given system and determining these regions is not always a trivial task. Stephanopoulos et al.
describe the comparison of the relaxation times for reactions with simple first order mechanisms in their development of the MFA model. However, the kinetic mechanisms described by Hynne et al. (2001), Rizzi et al. (1997), Bailey and Ollis (1986) and Peretti and Bailey (1986) are significantly more complex. Most of the mechanisms described in these texts are highly non-linear.

It is important to note that due to the high degree of regulation in metabolic pathways, the preceding catabolic reactions in a pathway can be either up or down regulated. Stephanopoulos et al. (1998) describe the significant role of feedback inhibition in maintaining the levels of metabolite pools to prevent toxic or osmotic shocks and even inefficient energy utilisation. The authors indicate that the timescales for these regulatory reactions are very fast. However, the ability for the cellular reactions to respond rapidly to any changes depends on the whether the required enzymes are available in high concentrations. Therefore, the rapid shifts in the reaction state can be limited by low enzyme concentrations which may occur at the start of an experiment or where substrates become exhausted. Bailey and Ollis (1986) describe the rate constant of enzymatic reactions as being intrinsic to the enzyme concentration present in the reaction volume. This suggests that the ability of the reaction system to shift to a new steady state may be highly dependent on the dynamics of that system.

Taking a classical thermodynamic approach, the disruption of any steady state or stationary point close to equilibrium can result in an oscillatory response. Van Santen and Niemantsverdriet (1995) describe how oscillations can occur due to perturbations around a stationary point for autocatalytic reactions. They state that for a stationary point to exist thermodynamically, the entropy of the system must be maximised. They go onto to describe this criterion mathematically by:

\[ \delta^2 S = \sum_{ij} \frac{\partial^2 S}{\partial n_i \partial n_j} = -\frac{1}{T} \sum_{ij} \frac{\partial \phi_i}{\partial n_j} \delta n_i \delta n_j < 0 \]  

where \( S \) represents the system entropy, \( n_i \) the moles of component \( i \), \( \phi \) the chemical potential and \( T \) the temperature of the system. To satisfy this thermodynamic criterion, Eq. 1.5 must be minimised. As a result, an oscillatory system response can occur for even the simplest three step sequential reaction pathways (van Santen and Niemantsverdriet 1995). The period and amplitude of the oscillations is dependent on the level of perturbation of each species as well as the kinetics of each step in the reaction. This phenomenon was also observed by Hynne et al. (2001) for their experimental and computational analyses of the glycolysis in S. cerevisiae. They describe the oscillatory nature of the stationary points.
of the system as a response to mixed flow concentrations of glucose and cyanide in CSTR experiments. They also provide experimental and computational evidence that oscillations occurred for up to 100 minutes as a result of moderate perturbations to the mixed flow concentration of glucose fed to the system. They do however also show response times and oscillations to be much lower for a number of other metabolites as a result of this perturbation. The results presented by Hynne et al. (2001) suggest that the shift to a new steady state was very dependent on the system dynamics. They have shown clearly that a given metabolite pool can shift rapidly to a new stationary point due to certain perturbations, but not all of them. Even more significant is that these observations were made for glycolysis in *S. cerevisiae*, which is considered to be a well understood reaction pathway.
1.3 Computation of MFA models

The simplifications described by Stephanopoulos et al. (1998) in developing sets of metabolic reaction balances has allowed for the simulation of internal metabolic reaction fluxes using linear numerical routines. Stephanopoulos et al. (1998) state that standard linear algebra is commonly used to solve systems that are determined. They also state that linear programming can be used to minimise a specified objective function to obtain a set of predicted flux profiles for under-determined systems. Covert and Palsson (2003); Kauffman et al. (2003); Reed and Palsson (2003) and Schilling et al. (2001) discuss the use of linear programming to minimise objective functions described by the system constraints. These models are used to generate simulation or feasible reaction spaces for the metabolism of a given micro organism. Lee et al. (2005) provide details of the use of linear programming to identify gene knockout targets in *E. coli*. The authors also provide a list of on-line software tools that are available for the computation of the metabolic fluxes. The high level front end graphics user interfaces (GUIs) provided by these software tools limits the user’s ability to access the source code used to perform the numerical analysis of the metabolic system. Therefore computations can be carried out without the user even being aware of which numerical routine has been used and how the models have been populated with the various experimental or optimisation constraints.

Standard linear algebra routines, such as LU factorisation and Singular Value Decomposition, and linear programming routines use a number of different procedures for the factorisation of a matrix and the subsequent computation of the solution vector (Press et al., 1996). The authors provide a number of examples to illustrate the differences between the linear algebra and linear programming routines and how the structure and characteristics of the linear models can affect the accuracy of the solution vector obtained from each routine. A critical review of the LU factorisation and Singular Value Decomposition linear algebra routines and the linear programming routines has been conducted. A critical review has been included to highlight the difference in the computation procedures between these routines and how they can affect the accuracy of the solution vector. Furthermore, the critical review aims to highlight how the characteristics of a given reaction model can affect the accuracy of the solution vector with each of the numerical routines.
1.3. COMPUTATION OF MFA MODELS

1.3.1 Review of standard linear algebra routines

There are a large number of numerical routines that can be used to solve linear systems (Rice and Do, 1995). A number of these routines can be used to solve square and non-square systems. The routines range from using various elimination techniques, orthogonal decompositions and iterative procedures to solve the linear system (Heath, 2001; Rice and Do, 1995 and Hager, 1988). Hager (1988) indicates that the well studied Gaussian elimination routines, and their adaptations, are commonly used to generate L (lower triangular matrix) and U (upper triangular matrix) decomposition factors for a given square matrix A. These factors can be used very simply to solve a system of linear equations using forward- and backward-substitution. The method used to solve the linear system \((Ax = b)\) using LU factorisation has been described algebraically by Heath (2001) as:

\[
Ax = b \quad A = LU
\]

where \(A\) represents the matrix of variable co-efficients, \(x\) the solution vector and \(b\) the vector of system constraints. Heath (2001) states that the major benefit of using this LU factorisation routine is that higher computational efficiency can be achieved for large systems. This becomes especially evident when performing a large number of computations on updated systems. The simplicity of the LU factorisation procedure is accompanied with an inherent disadvantage. Heath (2001) and Hager (1988) discuss the algorithm stability issues that are associated with the elimination steps of the factorisation. In some cases the leading diagonal entry of the remaining unreduced portion of the matrix can be zero. This results in an inability to compute all the multipliers and eventually leads to a break down of the Gaussian elimination process (Heath, 2001). This can be overcome through the use of pivot elements which lead to an interchange of rows in the matrix before the Gaussian elimination takes place. The use of these elements results in a pivoting of the matrix and is carried out using a permutation matrix \((P)\). The technique of partial pivoting for the linear system \((Ax = b)\) is described by Heath (2001) algebraically as:

\[
PAX = Pb \quad PA = LU
\]

\[
Ly = Pb \quad Ux = y
\]
where $P$ represents the permutation matrix which has only one non-zero element in each row. Hager (1988) indicates that all non-zero elements are suitable as diagonal elements during the factorisation routine. The author states that the magnitude of the pivot element should not be too large in an attempt to minimise the the propagation of numerical error. The limitation of all pivot multipliers to values that never exceed 1 is required if for each column of the matrix $A$, the choice of the diagonal element is made for that which has the largest magnitude. Heath (2001) describes this practise as partial pivoting. Rice and Do (1995) discuss the use of matrix scaling and the use of a diagonal matrices to provide further numerical stability to the elimination procedure.

Hager (1988) describes the use of Cholesky factorisation as an alternative to LU factorisation through the incorporation of a diagonal matrix. Heath (2001) indicates for symmetric and positive definite matrices, the use of Cholesky factorisation requires half as much work and storage space as the LU factorisation by Gaussian elimination. Furthermore, no pivoting is required for numerical stability. Heath (2001); Rice and Do (1995) and Hager (1988) describe a number of iterative techniques that can be advantageous over using the standard direct factorisation methods. Heath (2001) indicates that the main advantages of these iterative techniques are in the saving of the required computational storage memory. Furthermore, although solution accuracy is dependent on the availability of a good initial estimate, the benefits of using such methods to solve large sparse systems becomes very significant.

Improvement of the numerical stability and the ultimate solution accuracy is limited to non-singular matrices. According to Heath (2001), a matrix ($A\{n \times n\}$) satisfies the following equivalent non-singular conditions :

- $A$ has a computable inverse
- $\det(A) \neq 0$
- $\text{rank}(A) = n$
- For any vector $z \neq 0$, $Az \neq 0$

Heath (2001) describes the measure of a matrix’s singularity through determination of the matrix condition number. It is indicated that the condition number is a quantitative measure of the relative maximum stretching that the matrix does to any non-zero vector. Graphically it can be viewed as
the distortion of the unit sphere under transformation by a matrix. Therefore due to the invertability of a non-singular matrix $A$, there will always be a unique solution to $Ax = b$ regardless of the value of $b$. [Heath] (2001) indicates that if $A$ is singular, there can be infinitely many solutions as long as $b \in A$ (consistent system). However, if $b \notin A$ (inconsistent system), there is no solution to the system.

Stephanopoulos et al. (1998) states that matrix singularity arises from the presence of dependent reaction balances in the stoichiometric matrix. The presence of dependent rows in a matrix leads to the generation of a row with zero elements during the factorisation routine [Heath] (2001). The author indicates that the parallel nature of straight lines generated by the planes for dependent rows in the matrix lead to poorly defined points of intersection. Geometrically this can lead to a very large solution space which is further expanded by the error bounds attributed to rounding or measurement errors. This large solution space generated by the rank deficient matrix results in very poor solution accuracy (Hager] (1988). This leads to the generation of infeasible parameter spaces for estimation of intrinsic reaction parameters for a set of reaction pathways and reaction conditions.

In an attempt to mitigate the effects of using poorly conditioned matrices, [Stephanopoulos et al.] (1998) suggest that the biochemistry of the reaction matrix needs to be analysed to allow for the removal dependent reactions. They cite an example of the existence of singularity in prokaryotic reaction pathways when anaplerotic pathways were analysed along with the TCA and glyoxylate cycles. It is stated that biochemical analysis of the system allows for simplification due to the non-concurrent nature of the glyoxylate and TCA cycles, necessitating the inclusion of only one of the cycles. However, simplifications of this nature must be done with extreme caution. [Stephanopoulos et al.] (1998) indicate that the exclusion of either of the two cycles can lead to radically different flux distributions. Furthermore, these simplifications immediately draw attention towards the resulting consistency of the adapted mass balances. The ad-hoc removal of certain reactions from the balances only to obtain a set of independent reactions can result in large amounts of mass being incorrectly distributed across the metabolites. This can lead to poor intracellular reaction flux prediction even in cases where the extracellular fluxes are completely constrained.

To prevent uncertainties due to incorrect dependent reaction removal, the matrix can be left as is and alternative computational procedures can be applied. In this way, all the information can be retained, even the reactions that occur to very small extents, whilst obtaining suitable reaction flux simulation accuracy. This is possible provided there are enough constraints available to prevent the system from
becoming rank deficient. [Heath(2001)] describes singular value decomposition (SVD) as one such numerical routine that is capable of solving poorly conditioned systems. The author indicates that SVD is a diagonal factorisation technique which generates a diagonal matrix in which the elements are the singular values of the original matrix. The other two matrices have columns that correspond to the corresponding left and right singular vectors. [Heath(2001)] describes the algebraic factorisation via SVD as follows:

\[ A = U\Sigma V^T \]  \hspace{1cm} (1.8)

where \( \Sigma \) represents the diagonal matrix and \( U \) and \( V \) represent the orthogonal matrices populated by the left and right singular vectors respectively. According to [Heath(2001)], the singular values and vectors can be used directly to solve a linear least squares system as follows:

\[ x = \sum_{\sigma_i \neq 0} \frac{u_i^T b}{\sigma_i} v_i \]  \hspace{1cm} (1.9)

where \( \sigma \) represents the singular value, \( u \) the singular vector from matrix \( U \) and \( v \) the singular vector from matrix \( V \). The author goes on to mention that SVD has superior robustness and reliability in comparison to all the factorisation techniques. However, this comes at a very high computational cost in comparison to the commonly used LU factorisation routine. It is also important to note that solution inaccuracies can result from the use of a robust SVD numerical routine, as the routine does not guarantee the generation of a mechanistically unique solution. The linear least squares routine can be affected by inherent non-linearities that can arise in dependent reaction networks.

For non-square systems, QR factorisation can be used to determine the unitary matrix \( Q \) and upper triangular matrix \( R \) through orthogonal transformations rather than elementary eliminations ([Demmel, 1997]). [Heath(2001)] and [Demmel(1997)] discuss the computation of over- or underdetermined systems using the \( Q \) and \( R \) matrices to determine the linear least squares solution for the set of linear equations. [Heath(2001)] illustrates the usefulness of QR factorisation in that the method is always capable of obtaining consistent factors, but goes on to indicate that this does not necessarily mean that a unique solution can always be obtained. The author indicates that for rank deficient systems, the obtained upper triangular factor (\( R \)) is singular. Therefore in these instances, many solution vectors can provide the same minimum residual norm. In these instances, [Heath(2001)] indicates that the model should be reformulated or if possible, the author advocates the use of SVD to determine the solution that gives
1.3. COMPUTATION OF MFA MODELS

the minimum Euclidean norm.

1.3.2 Review of linear programming routines

[Edwards et al. (2001); Stephanopoulos et al. (1998); Hatzimanikatis et al. (1996) and Voit (1992)] discuss the use of linear programming (LP) routines to determine intracellular flux distributions in under-determined systems. They indicate that maximisation of the biomass growth rate can be used as a suitable objective function subject to the constraints imposed by the metabolic reaction pathways. [Beasley et al. (1996); Dantzig (1965) and Gass (1964)] provide a large number of examples where linear programming routines have been used to optimise a wide range of objective functions. These examples focus on the application of linear programming routines for the optimisation of transportation, matrix games, scheduling and financial algorithms. However, the authors clearly state that the range of application of linear programming routines is much broader. [Edgar and Himmelblau (1988)] provide a number of practical examples where linear programming has been used in the optimisation of chemical processes.

[Dantzig (1965)] describes the development of the subject of linear algebra through a number of key optimisation projects. The author describes the formulation of the simplex algorithm to solve linear systems with inequality constraints. The simplex algorithm described is only initiated once the equations are in a canonical form. [Beasley et al. (1996)] state that since the initial formulation of the simplex method in the tableau form in 1947, a number of subsequent contributions have been made to improve the routine. [Gass (1964)] states that in the development of the simplex algorithm, it is assumed that the problem is feasible, every feasible solution is non-degenerate and that a basic feasible solution is given. [Beasley et al. (1996)] provide an algebraic description for the determination of the solution of the following primal linear programming problem:

\[
\begin{align*}
\text{minimise} & \quad c^T x, \\
\text{subject to} & \quad Ax = b, \\
& \quad l \leq x \leq u,
\end{align*}
\]

where \( c \) represents a vector of objective function weighting coefficients, \( x \) represents the feasible
solution vector, $A$ represents the matrix containing the coefficients for the linear equations, $b$ represents the solution constraint vector and $l$ and $u$ represent the lower and upper bounds on the elements of the solution vector. These inequality constraints provide a specified degree of slack to the system. This slack results in a number of solutions to the system which correspond to the bounded solution space. However, solutions only become feasible if they satisfy both the equality and inequality constraints. By partitioning the matrix $A$ into basic and non-basic variables, Beasley et al. (1996) state that the solution can be obtained algebraically by:

$$A_{B}x_{B} + A_{N}x_{N} = b \quad \text{or} \quad Bx_{B} + Nx_{N} = b \quad (1.11)$$

where $B$ represents the basis of matrix $A$, and represents any square non-singular sub-matrix selected from the columns of $A$. $N$ represents the corresponding non-basic sub-matrix of $A$ and $x_{B}$ and $x_{N}$ represent the basic and non-basic solution respectively. Beasley et al. (1996) describe the iterative nature of the simplex algorithm and how there is a continual update of $x_{B}$ with each iteration until a feasible ($l \leq x \leq u$) solution is obtained. They describe the development of the dual problem and the link between the solutions of primal and dual problems. The dual pair for the problem summarised by Eq. 1.10 is described algebraically by Beasley et al. (1996) as:

$$\text{maximise } b^{T}y, \quad (1.12)$$

subject to $A^{T}y \leq c$,

where $y$ is a vector whose elements are unrestricted in their values. This can be written as:

$$\text{maximise } b^{T}y, \quad (1.13)$$

subject to $A^{T}y + z = c$,

and $z \geq 0$,

where $z$ is a vector of dual slack variables. There are a number of steps that have been added to the simplex algorithm to improve efficiency, ultimately resulting in the development of the revised simplex algorithm (Beasley et al., 1996; Gass, 1964). These include, among others, variations to the matrix
1.3. COMPUTATION OF MFA MODELS

factorisation, forward and backward transformations and pivoting procedures. Gass (1964) states that essentially the main difference between the original and revised simplex algorithm is that only the elements of the inverse matrix are transformed in the latter. These have improved computational suitability and efficiency. Edgar and Himmelblau (1988) describe the use of slack or surplus variables in the inequality constraint equations. They indicate that these are non negative variables used to transform the inequality constraints into equality constraints and enable the selection of basic and nonbasic variables for the augmented optimisation problem. The combination of the basic and nonbasic variables allows the generation of the solution space. Beasley et al. (1996) and Dantzig (1965) indicate that the generation of these solution spaces allows the solution vector more mobility as solutions are no longer constrained to the intersection vectors generated by equality constraint planes. This makes the linear programming routines better suited for use in optimisation procedures than the linear algebra routines.

Edgar and Himmelblau (1988) discuss how unique solutions are not always obtained for linear programming routines and the manner in which the solution space is generated can result in a large number of local minima for the objective function. The authors, do however, indicate that the generation of a number of local minima for the objective function can be very useful when optimisation is required for a system where all the fundamental data is not readily available or when alternate regions of optimal process operation are required. They indicate that the number of local minima which are obtained for a given optimisation are dependent on the size and shape of the feasible solution space generated by the inequality constraints and the intersections of multidimensional surfaces or curves generated by the equality constraints. Edgar and Himmelblau (1988) provide details around how the form of the objective functions can generate non-convex regions of search for the routine. Furthermore, the authors also indicate that feasible regions can be generated by inequality constraints that do not include the extreme maximum or minimum for the objective function. The authors clearly state that the nature of the search region generated by the constraints have an important bearing on the potential for obtaining suitable results in optimisation.

Beasley et al. (1996) describe the development of the interior point methods which allow the solving of linear programming problems with polynomial functions inherent in the upper bounds. They indicate that this has enabled the implementation of barrier functions to penalise any violation of the system constraints by forcing the system to diverge exponentially. These barrier functions are used to help
develop a convex set for the solution spaces. The major focus of these methods is on the concept of the central path. Beasley et al. (1996) describe this as a privileged continuous curve that is interior to the feasible set and converges to an optimal point. An optimal solution is obtained by traversing the interior of the feasible solution with this routine (Beasley et al., 1996). Although these methods become very useful in solving large and convex optimisations, Beasley et al. (1996) indicate that the convergence to an optimal solution may not be asymptotic for some algorithms. Therefore the time complexity of a given problem may not represent the true computational effort required, if the reference point is not an interior feasible solution. Beasley et al. (1996) also provide insight into the implementation of linear mixed-integer linear programming (MILP). They state that these methods provide significant advantages in solving difficult problems by dividing up the feasible space into a finite number of smaller regions. However, the individual optimisation of each of the regions can result in very large computational effort and therefore these methods are generally used on smaller problems.

It should be appreciated that linear algebra and linear programming techniques can be used either in conjunction with or independently of each other. The choice will ultimately be made depending on the complexity of the system and required solution accuracy. The literature suggests that an LU factorisation routine provides the most accurate solution (unique solution) as long as the data error is low and the system is well conditioned. In the case where the system is poorly conditioned, SVD can be applied as it does not suffer from the same effects due to matrix singularity. However, the robust nature of this routine comes at a significant computational cost compared to that of the LU factorisation routine. This limitation becomes especially severe when solving large systems. In situations where the system is poorly conditioned and has inherent error in the data, it may in fact be advantageous to utilise linear programming routines to solve the system. The slack available and the fact that the factorisation is performed only on the basis of the matrix can allow for the computation of suitable solutions for large square or non-square singular systems. However, the routine’s ability to adequately differentiate between dependent reactions is not guaranteed. This is especially true when the system itself has a high level on inherent non-linearity. It is also important to note from the literature review conducted for non-linear programming routines, that the selection of reaction constraints can severely hamper the ability of the optimisation routine to reach a feasible solution that represents the global maximum or minimum.
1.4 Scope and objectives

Considering the theory discussed, uncertainty can arise about the validity of the broad application of linear models for the computation of the reaction rates for any biochemical reaction system under any conditions. Although a large amount of evidence is available to support the use of MFA models in many cases, some evidence provided suggests that it may not be true across all systems. In other words, the suitability of the linear models may be dependent of the characteristics of a given reaction system. These may include the kinetic and operating parameters as well as the form of the kinetic mechanisms themselves. Furthermore, the perturbation of the system, either through external or internal manipulations to the reaction system, does not guarantee that an accurate steady state solution will be obtained.

The scope of the study was limited primarily to batch systems to highlight the potential inconsistencies that can arise from using the generalised formulation of MFA that the pseudo-steady state hypothesis applies under all conditions, regardless of the system dynamics and the design of the experimental system. It has been noted that the majority of MFA studies are conducted on continuous cultures and therefore the affect of system dynamics can be ignored. However, the generalised statement that the PSSH is valid for all biochemical reaction systems can result in some confusion and inconsistency when dynamic systems are modelled.

The focus of this research is to provide insight into the effect of the characteristics of each metabolic system on the computation of reaction fluxes within a given pathway using linearised reaction models. This is necessary to determine the validity of the linearised models for the simulation of reaction pathways under various conditions. The rigorous computational analyses provided specific details about the validity of the linearised models under a wide range of conditions. A set of hypothetical reaction pathways has been developed to investigate the accuracy of metabolite flux simulations under a large number of conditions. The case studies were developed to keep the reaction system as simple as possible but also to maintain the thermodynamic and kinetic feasibility of the system.

The reaction system was chosen to create a network of chain and branched reactions, providing a good representation of many biochemical reaction pathways. The glycolytic pathway in *S. cerevisiae* is one such pathway which involves the cleaving and isomerisation of various metabolites for the
1.4. SCOPE AND OBJECTIVES

generation of energy for biomass growth from glucose (Bailey and Ollis, 1986). The simplicity of the developed network allows for the investigation of any numerical inconsistencies obtained in this regard, without the added complication of considering biologically specific regeneration balances. Such balances include those for the energy carrier ATP and the redox reagents NADH and NADPH (Forster, 2003 and Schulze, 1995). Furthermore the lack of comprehensive sets of in vivo kinetic and system data limited the ability of the numerical analyses to be performed rigorously across a wide range of parameter sets. It is important to note that it has been recognised that, although every effort was made to make the systems representative of biological systems, there may still be some inconsistencies that were obtained due to the exclusion of suitable control architectures and biologically specific reaction balances. However, the time scales and magnitudes of the reaction data was matched to experimental data to ensure that there was consistency in the macro-scale metabolic flux analysis. Furthermore, the reaction data was validated using experimental flux data obtained from literature to ensure that the magnitudes of the reaction terms were representative of that commonly obtained in practical biochemical systems.

The systems used in these investigations therefore merely highlight the computational and numerical inaccuracies that can be obtained when performing flux computations using linearised reaction models, and how the accuracies of the models change when the system is altered either kinetically or through changes to the experimental protocol. The observed trends in the errors as a result of these changes are of primary importance in this study. It is hypothesised that, due to the focus of the analyses on the computation of the reaction pathways under the various conditions, the observed sensitivities in the errors will be similar to practical biochemical systems. Therefore, the magnitudes of the errors obtained for these analyses should not be considered to be directly applicable to actual practical biological systems.

The primary aim of the research conducted in this study is to determine the sensitivity of the computational accuracy of the linearised models to various parameters and properties of the numerical solvers and not to provide a set of results that either refutes or supports the use of MFA based models for the computation of biological reaction fluxes. The parameters used in the simulations are commonly altered within a given reaction study either through manipulation of the reaction pathways (Forster, 2003) or by changing the reaction parameters (Schulze, 1995) with little or no consideration of these effects on the consistencies of the linearised reaction models. It is intended that the observations made
from this research will guide further studies to investigate these effects within the context of practical biological systems.

It is very common for researchers to use MFA to elucidate the biochemistry of the system based on the inconsistency of the models (Vallino and Stephanopoulos [1994a,b]), but very seldom applied to test the inconsistency of the actual flux models themselves. This research study will focus on the later as the “biochemistry” of the reaction pathways has been fully defined.

The objectives for this research are summarised below:

- To determine the effect of batch kinetic and operating parameters on the computational accuracies of the linearised reaction models
- To determine the effect of the numerical structure of the model on the computational accuracies of the linearised reaction models
- To identify regions in the reaction space where the linearised reaction models are most accurate

A detailed research methodology (Chapter 2) has been included to describe how the “experimental” reaction data was generated using well described kinetic models, and the subsequent computation of the reaction fluxes using this data. The reaction data was generated using a number of kinetic mechanisms and reactor operations (batch and steady state). The results obtained for the batch reaction models have been deconstructed to provide a detailed analysis of how the inherent model components are affected by the kinetic and operating parameters. The results obtained for the batch reactions have been analysed in Chapters 3, 4 and 5. The steady state results have been analysed in Chapter 5 to provide a more detailed deconstruction of how the PSSH is coupled in the model data between the two modes of operation.

A detailed analysis of the suitability of the various numerical routines for the computation of the linearised flux models has been included in Chapter 3. This analysis provides much needed insight into how the commonly used numerical routines can fall foul, due to inconsistencies obtained from poorly designed experimental and computational studies.
Chapter 2

Research Methodology

As stated in Chapter 1, the aim of the research conducted in this study is to determine the sensitivity of linearised reaction models to various kinetic and operating parameters. To achieve this, hypothetical reaction pathways were developed based on a set of elementary reactions. The reactions in these pathways were populated with a range of kinetic mechanisms referenced from literature. The hypothetical reaction pathways used in the research study were made simple to allow for detailed analyses on the computational accuracies obtained under various conditions. The data generated from the hypothetical reaction pathways was used to generate concentration data under various conditions. The concentration data were then used to calculate the constrained extracellular reaction fluxes which were used in the subsequent flux calculations. The procedure used for the flux calculations was identical to those used in conventional experimental studies (Huang [2003], Rizzi et al., 1997, Vallino and Stephanopoulos, 1994a and Vallino and Stephanopoulos, 1994b). The concentration data was obtained from the computation of the reaction balances using the various kinetic mechanisms and the reactor balances under various modes of operation for each condition. The errors obtained for the estimated fluxes under each condition were obtained from a relative error analysis using the actual kinetic rates used to generate the reaction data. The error analysis was used to determine the ability of the models to back calculate the actual reaction rates. A general outline of the methodology used to conduct the research has been included below.

To test the sensitivity of the linearised reaction models to the form of the kinetic mechanisms used to populate the reaction data, a range of mechanisms were used in the study. The form of the mechanisms
range from linear to highly non-linear to obtain a wide range of reaction spaces. The mechanisms used were both enzymatic and chemical in origin. The following sets of kinetic expressions were used to describe the reaction pathways:

- Monod kinetics with non-competitive (Kin$_1$) and competitive (Kin$_2$) substrate inhibition (Bailey and Ollis [1986])
- Simple Monod kinetics (Kin$_3$) (Bailey and Ollis [1986])
- First-order kinetics (Kin$_4$) (van Santen and Niemantsverdriet [1995])

All reactions within the pathway were assigned with the same mechanisms. It has been noted that this is not truly representative of an actual biochemical reaction pathway, as in most cases various mechanisms have been used to describe the reactions in a single pathway (Hynne et al. [2001], Rizzi et al. [1997] and Peretti and Bailey [1986]). However, for this study, the same mechanisms were used for all reactions in a single pathway to illustrate the sensitivity of the reaction models to the form of the reaction mechanism. This would have been difficult to achieve with a mixture of mechanisms as the effects of each reaction would need to be isolated from the observed model errors.

The reactions were modelled under batch and continuous operation to determine the sensitivity of the linearised reaction models under the different modes of operation. The operating parameters varied for the batch and continuous operations were:

- Initial substrate concentrations
- Relative substrate concentrations
- Reactor dilution rates

The relative substrate concentrations were used to simulate the effects of limiting and excess substrates on the model error. The reactor dilution rates were used to simulate the transition between continuous and batch cultures.

The sensitivity of the linearised reaction models to perturbations in the reaction pathway and kinetic parameters was conducted to simulated changes in the reactions either through genetic or physical
manipulation of the reaction pathway. The change in the structure of the reaction pathway was achieved through the addition of reactions. This was done to mimic systems where extra reactions are removed or added from a biochemical pathway through the manipulation of the cells genome. Variation in the reaction parameters was used to simulate the genetic and physical effects on the enzyme activity.

Various numerical routines were used to solve the linearised reaction models to determine the sensitivity of the models to the various characteristics of each solver. The following linear routines were used to solve the models. The suitability of these routines to perform the computations was discussed in Section 1.3 of Chapter 1. The numerical routines used to solve the system were:

- LU factorisation
- Singular value decomposition (SVD)
- Linear programming (LP)

Variation of the computational structure of the model was achieved by changing the combination of fluxes used to constrain the models. The different constraint sets resulted in variation of the geometry and size of the feasible solution spaces generated (See Section 1.3.2 of Chapter 1).

To perform the sensitivity analyses described above, over 100,000 simulations were conducted. The high number of simulations conducted is further indication of the importance of the simplicity of the reaction pathways. The details for the methodology used to conduct the sensitivity analyses described above have been described below.

### 2.1 Determination of cellular reaction rates

#### 2.1.1 Elemental reactions

The first step in the generation of reaction rate data was to select a set of elementary reactions to describe the hypothetical reaction system. To maintain the simplicity of the analyses, a well described reaction pathway was required with no cycling reactions and closed reaction loops. Although there are
many well described biochemical reaction pathways, glycolysis in *S. cerevisiae* for example (Forster, 2003; Hynne et al., 2001; Stephanopoulos et al., 1998; Rizzi et al., 1997 and Schulze, 1995), modelling of this pathway can become cumbersome and complex when trying to account for all the redox and energy metabolites. Forster (2003) and Schulze (1995) provide long lists of possible reactions that can occur in glycolysis in *S. cerevisiae*. It was decided to remove the complexity of the models away from the “biochemistry” of the system and to rather focus on the computational aspects of the models. The analysis of a hypothetical pathway using a simplified reaction scheme is possible when using the linearised reaction models as they only take the reaction stoichiometry into account. Therefore specifics of the reactions being modelled, such as the thermodynamics or presence of cycling redox and energy reactants, do not affect the computation of the linearised models. However, to ensure that the reaction system was representative of a biochemical system, the structure of the reaction pathway was made similar to that of a conventional biochemical pathway.

The reactions used to describe the reaction pathway were a set of cleaving and oxidation reactions which mimic those in the glycolysis in *S. cerevisiae* described by Hynne et al. (2001) and Rizzi et al. (1997). To obtain a small reaction pathway, large conventional biochemical substrates could not be used. The simplified pathway described by Hynne et al. (2001) for the conversion of glucose to ethanol involved 24 reactions and and over 20 metabolites. The full mechanistic model for the metabolism of glucose in *E. coli* described by Peretti and Bailey (1986) involved over 50 reactions and 40 metabolites. Therefore a smaller molecule (n-butane) which is easy to cleave and oxidise in only a few reaction steps was chosen as the primary “carbon source” for the reactions. It must be noted that the choice of molecules to generate the reactions was completely random and that the selection of any molecule could be possible as long as the reactions are thermodynamically feasible and that balanced chemical equations can be developed. Therefore the “metabolites” used in these reactions are purely theoretical and as a result no solubility criterion or mass transfer limitations were considered in the reactions. Again, these were not necessary in this study as the linearised reaction models do not require any thermodynamic or mass transfer relations, only the reaction stoichiometry.

The kinetic reactions and metabolic pathways described by Hynne et al. (2001) and Rizzi et al. (1997) were compiled individually in an attempt to validate the reaction pathways used in this study. However the complexity of the models and their high degree of specificity to the given systems prevented the generation of suitable reaction data for validation of the reaction pathways. Hynne et al. (2001) are
very clear about the large number of limitations which are associated with their kinetic model, even though it is described as a full-scale model of glycolysis in *S. cerevisiae*. The authors state clearly that the reaction pathway described, lacks conclusive biochemical justification. Furthermore, the authors also state that the magnitude of the computational problem significantly impeded the simulation of the reaction pathway and the subsequent estimation of the system parameters. To overcome these limitations, the authors made a number of simplifications to the model to prevent the exponential growth of the parameter space.

The instability of the kinetic model described by Hynne et al. (2001) was evident in the oscillatory nature of the concentration profiles obtained for some of the cometabolites. The presence of unstable regions in the parameter space was evident from the bifurcation diagrams generated for the model. Apart from the complexity of the kinetic models described by Hynne et al. (2001) and Rizzi et al. (1997), the inconsistencies in the models made validation of the model used in this study even more difficult. The exclusion of important cellular mass balances and the inconsistencies in the described reaction mechanisms made the simulations very unstable. Furthermore, the lack of mass balances for the biomass prevented accurate simulation of biomass growth and the closure of the overall mass balance for the system. Due to the complexity and instability of the two models, suitable sensitivity analyses were not possible. The instability of the model described by Hynne et al. (2001) during the sensitivity analysis on the kinetic and operating parameters was identified by the authors. They indicate that the model is only valid over small regions around the neighbourhood of the single stationary point considered during the optimisation. This resulted in the generation of smaller reaction pathways which were well described and computationally stable. This allowed the study to focus on the evaluation of the computational aspects of the model rather than getting weighed down by discussions around complex biochemical reaction systems. However, validation of the reaction models was conducted using data obtained from experimental and computational flux studies reported in literature. The model validation has been included in the results section of Chapter 3.

Theoretical metabolites such as A, B, C etc. could have also been used in the development of the elemental reactions. However, molecular formulae were required for the reactions to ensure that there was consistency in the mass balance for each reaction. The reactions are described as irreversible to allow for the determination of the net rate of reaction. The set of elemental reactions used were:
2.1. DETERMINATION OF CELLULAR REACTION RATES

(1) \( C_4H_{10} \rightarrow C_2H_4 + C_2H_6 \)

(2) \( C_2H_6 \rightarrow C_2H_4 + H_2 \)

(3) \( C_4H_{10} + 6.5O_2 \rightarrow 4CO_2 + 5H_2O \)

(4) \( C_2H_6 + 3.5O_2 \rightarrow 2CO_2 + 3H_2O \)

(5) \( 2C_2H_4 + 4O_2 \rightarrow CH_{1.5}O_{0.5} + 3CO_2 + 1.5H_2O + 1.75H_2 \)

To ensure that the system was representative of a biochemical system, a hypothetical biomass reaction was developed (Rxn 5). The formula for the biomass was obtained from standard molecular formulae provided by Schuler and Kargi (2002) and Bailey and Ollis (1986). From the reactions shown above it can be seen that there is an overall “bioconversion” of n-butane for the generation of biomass (\( CH_{1.5}O_{0.5} \)), with the production of \( H_2 \), \( CO_2 \) and \( H_2O \) as by-products. The molar mass of the model biomass constituents was 22 g/mol, which is equivalent to the molar masses provided for a number of model micro organisms (Schuler and Kargi, 2002; Stephanopoulos et al., 1998 and Bailey and Ollis, 1986). The biomass constituents produced in these reactions are exported directly out of the reaction volume for the formation of biomass in the extracellular mass balances. Internal biomass constituents could be produced and exported out of the reaction volume via another reaction (Forster, 2003 and Huang, 2003). However, in this investigation, the intracellular rate of production and transportation rate of the constituents was combined into a single reaction. This does not have any effect on the reaction balances as the biomass production rate is always constrained in MFA computations for the determination of extracellular fluxes [mmol/g\(_{\text{biomass}}\).hr] (Stephanopoulos et al., 1998; Rizzi et al., 1997; Vallino and Stephanopoulos, 1994a and Vallino and Stephanopoulos, 1994b).

It must be noted that the presence of very large concentrations of extracellular \( O_2 \), \( H_2 \) or \( CO_2 \) in the liquid phase is very unlikely due to low saturation concentrations of these gasses in aqueous media (Perry and Green, 1997). However, the hypothetical reactions used in the simulations are not controlled by any thermodynamic mixing or interaction rules. Furthermore, the generation of \( H_2 \) as a likely product from a practical biochemical pathway is not common and the study only references \( H_2 \) as a product to ensure consistency in the mass balances.
The extracellular H$_2$O formed in the reactions was represented as a separate concentration for that of the H$_2$O making up the aqueous reaction media. The water composition of the cells was estimated to be 70% (w/w), based on information provided by Schuler and Kargi (2002), Stephanopoulos et al. (1998) and Bailey and Ollis (1986). This value was used to determine the dry weight of the biomaterial for use in the flux calculations. The elementary reactions described above were used to generate the reaction pathways. Although the hypothetical elemental reactions used in the compilation of the model pathways would seem very strange with respect to a practical biochemical system, it should be noted that the system presented in the study is purely hypothetical. The only criteria required to perform the numerical analyses was that the reactions resulted in consistent mass balances and that there was a set of branching reactions which could be used to simulate the linear dependency that can be generated in conventional biochemical systems.

2.1.1.1 Reaction pathways

Once the elementary reactions were generated, they were used in various combinations to generate 3 reaction pathways. The reaction pathways differed only through the addition of an elemental reaction to pathway. The reactions used in the pathways are described in Section 2.1.1. The reaction pathways obtained from the elementary reactions have been shown in Figure 2.1. The pathways differ in the number of reaction fluxes that occur but have the same number of intracellular metabolites. Due to the increase in reaction fluxes within the different reaction pathways, the degrees of freedom of the reaction pathway increased accordingly. The labels provided to each pathway indicate the degrees of freedom that exist for that pathway. i.e. For pathway DF$_1$ (Figure 2.1a), there are 8 fluxes and 7 intracellular metabolites resulting in a single degree of freedom reaction pathway denoted by DF$_1$. The degrees of freedom of each reaction pathway indicate the number of reactions which need to be constrained to generate a square or fully determined reaction matrix. The numbers which precede each reaction step in the pathways indicate which elemental reaction is described by each rate. For reaction pathway DF$_1$, there are 8 rates in the pathway and only 7 metabolite balances resulting in 1 degree of freedom. For the reaction pathway DF$_3$, there are 10 rates in the pathway and only 7 metabolite balances resulting in 3 degrees of freedom.

The reaction pathways illustrated in Figure 2.1 show some similarity to those of a conventional
2.1. DETERMINATION OF CELLULAR REACTION RATES

biochemical pathway, albeit much smaller and simpler. The similarity in the reactions and the resulting pathways are only representative of a biochemical system in that there are a set of branched pathways that are generated as the metabolites are broken down into their individual constituents for use in the catabolic and anabolic pathways. The study does not in anyway claim that this hypothetical reaction model pathway is an actual biochemical network that exists in practise.

The simplified reaction pathways were used in the study as the reactions were well described and the size of the computational system was reduced significantly. This allowed for more comprehensive sensitivity analyses and evaluation of the computational accuracies of the linearised models under various conditions.

The reaction pathways have been annotated with the number used to represent each of the rates. The elementary reactions used to describe each of the rates have been labelled with the circled number. The internal elemental rates have also been colour coded for easier reference. Figure 2.1 illustrates the extracellular and intracellular reaction volume. In the extracellular reaction volume, the metabolite concentrations are intrinsic to the volume of media in the reactor. In the intracellular reaction volume, the metabolite concentrations are intrinsic to the reaction volume inside the cell membrane where the intracellular reactions take place.
2.1. DETERMINATION OF CELLULAR REACTION RATES

(a) Reaction pathway 1 (DF\textsubscript{1})

(b) Reaction pathway 2 (DF\textsubscript{2})

(c) Reaction pathway 3 (DF\textsubscript{3})

Figure 2.1: Set of hypothetical reaction pathways used

2.1.1.2 Kinetic reaction mechanisms

Once the reaction pathways were obtained (Figure 2.1), each reaction rate was populated using various reaction mechanisms referenced from literature. Reaction data is generated for all the kinetic mechanisms Kin\textsubscript{1} - Kin\textsubscript{4}. The kinetic parameters used in the reaction rate terms were chosen to be in the same order of magnitude as those estimated by Rizzi et al. (1997). Table 2.1 provides an example to illustrate the form used to describe the reaction rates based on each kinetic mechanism. The rate described is Rate\textsubscript{2} from reaction pathway DF\textsubscript{1} (Figure 2.1a). The full set of reactions used to describe the reactions in all reaction pathways have been shown in Appendix A. It must be noted that the product export rates from the cellular volume was model used the Kin\textsubscript{3} mechanism for the Kin\textsubscript{1} and Kin\textsubscript{2} mechanisms. This was due to the fact that the presence of feed-back inhibition on the product
export rates is not likely.

Table 2.1: Reaction mechanisms used to describe all reaction rates for all reaction pathways

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Form of the mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Michaelis-Menten with non competitive inhibition (Kin\textsubscript{1})</td>
<td>( \text{Rate}<em>2 = \frac{[kC</em>{int}C_{4H10}/(1 + C_{int}^H/K_i + C_{int}^H/K_i)]}{C_{4H10}^H + K_M} )</td>
</tr>
<tr>
<td>b) Michaelis-Menten with competitive inhibition (Kin\textsubscript{2})</td>
<td>( \text{Rate}<em>2 = \frac{kC</em>{int}^H}{C_{4H10}^H + K_M(1 + C_{int}^H/K_i + C_{int}^H/K_i)} )</td>
</tr>
<tr>
<td>c) Michaelis-Menten with no inhibition (Kin\textsubscript{3})</td>
<td>( \text{Rate}<em>2 = \frac{kC</em>{int}^H}{C_{4H10}^H + K_M} )</td>
</tr>
<tr>
<td>d) First-order (Kin\textsubscript{4})</td>
<td>( \text{Rate}<em>2 = k_s C</em>{int}^H )</td>
</tr>
</tbody>
</table>

**Constants**

\( k, k_s = \) Rate constants [hr\textsuperscript{-1}]

\( K_M = \) Michaelis-Menten constant [mmol/L]

\( K_i = \) Inhibition constant [mmol/L]

The first-order consumption term \((k \times C_{int}^H)\) in each mechanism was used to describe the consumption of \(C_{4H10}\) in each reaction. The saturation term \( \left( C_{int}^H/C_{4H10} + K_M \right) \) was used to describe the saturation of the active site of the enzyme by \(C_{4H10}\) in the Kin\textsubscript{1}, Kin\textsubscript{2} and Kin\textsubscript{3} mechanisms. The inhibition term \(1 + C_{int}^H/K_i\) was used to describe the product inhibition on the reaction rates for the Kin\textsubscript{1} and Kin\textsubscript{2} mechanisms. The reactions were all simulated as isothermal at constant pH. This allowed for the exclusion of the temperature and pH correlations from the reaction balances (Bailey and Ollis, 1986).

The standard values used for each constant are shown in Table 2.2. The values were chosen from the values estimated by Rizzi et al. (1997) from their kinetic model of glycolysis in \textit{S. cerevisiae}. The magnitude of the values reported by Rizzi et al. (1997) corresponded well with those reported by Hynne et al. (2001) for their kinetic study of glycolysis in \textit{S. cerevisiae}. The units of the parameters obtained from Rizzi et al. (1997) were converted to appropriate units before they were used in the reaction mechanisms. The following unit conversions were used to ensure consistency between the reaction mechanisms and the reaction balances:
2.1. DETERMINATION OF CELLULAR REACTION RATES

- \( k \ [s^{-1} \times 3600 \text{ s/hr} \rightarrow \text{hr}^{-1}] \)

- \( K_M \) and \( K_i \ [\text{mmol/L}_{\text{reactor}} \times 80 \frac{\text{L}_{\text{reactor}}}{\text{L}_{\text{cells}}} \rightarrow \text{mmol/L}_{\text{cells}}] \)

The conversion factor \( (\frac{\text{L}_{\text{reactor}}}{\text{L}_{\text{cells}}}) \) was determined using the specific volume \( (\hat{V}) \) and the initial biomass concentration \( (C_x(t = 0)) \) in the reaction media. The value obtained for \( (\frac{\text{L}_{\text{reactor}}}{\text{L}_{\text{cells}}}) \) was of a similar order of magnitude to that reported by (Hynne et al., 2001) for the volume factor \( (y_{vol}) \). The use of the exponent term \( E \) denotes the power of the number \((0.01 \rightarrow 1E-2)\).

Table 2.2: Standard reaction kinetic parameters used with all reaction mechanisms for all reaction pathways

<table>
<thead>
<tr>
<th>Kinetic Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k \ [s^{-1}] )</td>
<td>5</td>
</tr>
<tr>
<td>( k_s \ [\text{hr}^{-1}] )</td>
<td>2.75</td>
</tr>
<tr>
<td>( K_{Eq}[\text{mmol/L}] )</td>
<td>0.12</td>
</tr>
<tr>
<td>( K_M [\text{mmol/L}] )</td>
<td>15</td>
</tr>
<tr>
<td>( K_i [\text{mmol/L}] )</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The rate constant \( (k_s) \) used to describe the first order reaction for the Kin4 mechanism had to be scaled to ensure that the mechanisms generated reaction times of a similar order of magnitude. This was due to the lack of inhibition or saturation terms present in the Kin4 mechanism which meant the overall observed reaction rate was significantly higher.

Table 2.3: Standard average specific growth rates and molar fluxes obtained for reaction pathway DF1 with the various kinetic mechanisms

<table>
<thead>
<tr>
<th>Kinetic mechanism</th>
<th>( \bar{\mu} \ [\text{hr}^{-1}] )</th>
<th>( \bar{F} \ [\text{mmol/g.DW.hr}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kin1</td>
<td>0.045</td>
<td>2.1</td>
</tr>
<tr>
<td>Kin2</td>
<td>0.052</td>
<td>3.0</td>
</tr>
<tr>
<td>Kin3</td>
<td>0.064</td>
<td>3.3</td>
</tr>
<tr>
<td>Kin4</td>
<td>0.085</td>
<td>4.5</td>
</tr>
</tbody>
</table>

The average specific growth rates \( (\bar{\mu}) \) obtained using the standard kinetic parameters for each kinetic mechanism are shown in Table 2.3. The data shows that a wide range of specific growth rates were obtained for the various mechanisms. Data describing the average molar fluxes \( (\bar{F}) \) for the uptake of \( C_4H_{10} \) obtained using each kinetic mechanism has also been included. The average specific growth
rates and fluxes obtained with all mechanisms were in range of experimental fluxes reported by Forster (2003); Huang (2003); Stephanopoulos et al. (1998); Rizzi et al. (1997) and Schulze (1995).

2.2 Cellular and reactor mass balances

Once the reaction rates were determined by the form of the selected reaction mechanism, cellular and reactor mass balances were required to relate the reaction rates to the concentrations of the metabolites.

The cellular mass balances were intrinsic to the reaction volume which is defined as the intracellular unit volume where all the reactions take place. The reactor mass balances were intrinsic to the volume of the reaction media and were used to relate the intracellular consumption and production of the metabolites to the volume of the reaction media.

The intracellular mass balances for each metabolite was modelled using the following balance described by Bailey and Ollis (1986) and Stephanopoulos et al. (1998), which was an identical form to Eq. 1.1 in Section 1.2 of Chapter 1:

\[
\frac{dC_{i}^{\text{int}}}{dt} = \sum_{k} \text{Rate}_{ik} - \mu C_{i}^{\text{int}}
\]  

(2.1)

The variable \(C_{i}^{\text{int}}\) represents the intracellular concentration of metabolite \(i\) [mmol/L\(_{\text{cells}}\)], \(\text{Rate}_{ik}\) represents the intracellular rate of generation or consumption of component \(i\) in reaction \(k\) [mmol/L\(_{\text{cells}}\).hr] and \(\mu\) represents the specific cellular growth rate [hr\(^{-1}\)]. The term \(\mu C_{i}^{\text{int}}\) is the cellular dilution term which represents the dilution of the intracellular metabolites due to cellular growth (Stephanopoulos et al., 1998). Cellular mass balances were generated for all intracellular metabolites which appeared within the selected reaction pathway. Because the cellular mass balances are intrinsic to the cellular reaction volume, the form of the cellular mass balances were independent of the mode of operation of the reactor.

The effective intracellular concentration of the substrates \(C_{4}H_{10}\) and \(O_{2}\) (Eq. 2.2) was modelled using the extracellular species concentration \(C_{i}^{\text{ext}}\) and an equilibrium constant \(K_{\text{Eq}}\). The equilibrium constant represents the relative intracellular concentration of the metabolites based on the extracellular metabolite concentration across the cellular membrane. This model used to describe the relationship
between the intracellular and extracellular concentrations was referenced from Hynne et al. (2001) and Theobald et al. (1997). The standard value used for the equilibrium constant was selected based on those presented by Theobald et al. (1997) from their experimental study of glycolysis in *S. cerevisiae*. This was to ensure that the system used in this study did not provide any bias in the data generated for the intracellular concentrations under the various conditions. This is a result of the fundamental parameters that were used to describe the properties of the reaction pathways and mass balances.

\[ C_{\text{intr}}^{\text{ext}} = \frac{C_{\text{extr}}^{\text{ext}}}{K_{\text{Eq}}} \]  \hspace{1cm} (2.2)

The reactor mass balances for ideally mixed systems with constant density described by Rizzi et al. (1997) and Stephanopoulos et al. (1998) were used to generate the extracellular concentration profiles as functions of the intracellular reaction rates (Table 2.4). The extracellular concentration profiles were generated using the biomass concentration \( C_x \) and the specific cellular volume \( \hat{V} \).

**Table 2.4: Reactor mass balances for all modes of operation with all kinetic mechanism for all reaction pathways**

<table>
<thead>
<tr>
<th>Operation</th>
<th>Reactor Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Batch</td>
<td>[ \frac{dC_{\text{extr}}}{dt} = \sum_k \text{Rate}_{ik} C_x \hat{V} ]</td>
</tr>
<tr>
<td>b) Continuous</td>
<td>[ \frac{dC_{\text{extr}}}{dt} = \sum_k \text{Rate}<em>{ik} C_x \hat{V} + D(C</em>{\text{feed}} - C_{\text{extr}}) ]</td>
</tr>
</tbody>
</table>

**Variables and Parameters**

- \( C_{\text{extr}}^{\text{ext}} \) = Extracellular metabolite concentration [mmol/L reactor]
- \( C_x \) = Extracellular biomass concentration [g DW/L reactor]
- \( \hat{V} \) = Specific volume [L cells/g DW]
- \( D \) = Reactor dilution rate [hr\(^{-1}\)]

The reactor dilution rate \( (DC_{\text{extr}}^{\text{ext}}) \) in the continuous reactor represents the change in extracellular metabolite concentration of component \( j \) due to flow of media out of the reactor. The reactor dilution rate is equivalent to the inverse of the reactor space time.
2.2. CELLULAR AND REACTOR MASS BALANCES

Under steady state operation, the dynamic terms in the cellular and reactor mass balances Under steady state operation with a sterile feed \( C_{feed}^x = 0 \), the continuous reactor balance for the biomass becomes:

\[
0 = \mu C_x + D(0 - C_x)
\]

\[
\mu = D
\] (2.3)

Therefore under steady state operation, the specific growth rate \( \mu \) is determined by the reactor dilution rate. The formula used to determine the specific growth rate in the batch operation is described by Eqs. 2.4.

\[
\mu = \frac{dC_x}{dt} \times \left( \frac{1}{C_x} \right)
\] (2.4)

The reactor balances were used to simulate the extracellular concentrations of the substrates, products and biomass. The term \( \sum_k \text{Rate}_{ik} C_x V \) represents the net rate of generation of component \( i \), intrinsic to the volume of the extracellular reaction media. This term takes into account all intracellular reactions which are occurring within the extracellular reaction media at a given time, which is analogous to the reaction rates described by [Fredrickson (1976)] (Eq. 1.1). The cellular rates [mmol/L\_cells/hr] were converted to rates based on the volume of the reaction media [mmol/L\_reactor/hr] with the term \( (C_x V) \) [g\_DW\_biomass/L\_reactor \times L\_cells/g\_DW\_biomass]. It must be noted that the reactor volume was assumed to be a constant in the reactor mass balances. This is not necessarily true for batch experimental studies as the reactor volume can vary due to evaporation of the solution and sampling.

The specific cellular volume was estimated based on data provided by [Biswas et al. (2003)] and [Arnold (1972)] for the cellular reaction volume and the specific cell number respectively. The value determined for \( V \) (2.5 mL/gDW) was very similar to that used in the mass balance calculations by [Rizzi et al. (1997)] (2.38 mL/gDW). The Biomass concentration was determined from the extracellular mass balances through the conversion of the molar concentrations to mass concentrations using the molar mass calculated for the model biomass component (22 g/mol). A full set of reactor and cellular mass balances have been shown in Appendix C.
The standard values used to define the operating parameters in the reactor balances have been shown in Table 2.5.

Table 2.5: Standard operating parameters used for various modes of operation for all reaction pathways

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Std. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{\text{C,H}_10}^{\text{ext}}(t = 0)$ [mmol/L&lt;sub&gt;reactor&lt;/sub&gt;]</td>
<td>80</td>
</tr>
<tr>
<td>$C_{\text{O}_2}^{\text{ext}}(t = 0)$ [mmol/L&lt;sub&gt;reactor&lt;/sub&gt;]</td>
<td>320</td>
</tr>
<tr>
<td>$C_{\chi}(t = 0)$ [gDW/L&lt;sub&gt;reactor&lt;/sub&gt;]</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Std. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{\text{feed}}^{\text{fed}}$ [mmol/L&lt;sub&gt;feed&lt;/sub&gt;]</td>
<td>$6 \times C_{\text{feed}}^{\text{ext}}(t = 0)$</td>
</tr>
<tr>
<td>$D$ [hr&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>$\mu_{\text{max}}(\text{Kin}_i)$</td>
</tr>
</tbody>
</table>

The ratio between the standard initial and feed concentrations of the substrates was determined by the overall stoichiometry of the reactions. This was done to prevent the addition of excess substrate to the reactor.

The reactor dilution rate used under standard conditions was set to the maximum specific growth rates obtained under the batch operations with the specific kinetic mechanism ($\mu_{\text{max}}(\text{Kin}_i)$). This was done to ensure that there was no washout of biomass from the system due to an excessive dilution rate. Schuler and Kargi (2002) and Bailey and Ollis (1986) report that biomass washout occurs in continuous cultures when the reactor dilution rate is higher than the maximum specific growth rate that can be obtained under those conditions.

### 2.3 Sensitivity analysis of kinetic and operating parameters

Once the cellular mass balances were generated for each pathway and the reactor mass balances were selected for the mode of operation, the sensitivity analysis was conducted using the operating and kinetic parameters. The range of values used for the kinetic parameters and operating parameters to conduct the sensitivity analysis have been shown in Table 2.6 and 2.7 respectively. The upper and
lower bound of these parameters were chosen to be an order of magnitude larger and smaller than the standard values. The units of the kinetic parameters were converted using the same conversion described in Section 2.1.1.2. The reaction times obtained from the parameter sensitivities ranged from 1 to 750 hrs, and resultant biomass concentrations ranging from 5 to 20 gDW/L.

**Table 2.6: Range of values used for the sensitivity analysis on the kinetic parameters for all reaction pathways**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min. Value</th>
<th>Std. Value</th>
<th>Max. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>k [s(^{-1})]</td>
<td>0.5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>k(_s) [hr(^{-1})]</td>
<td>0.275</td>
<td>2.75</td>
<td>27.5</td>
</tr>
<tr>
<td>K(_{Eq}) [mmol/L]</td>
<td>0.012</td>
<td>0.12</td>
<td>1.2</td>
</tr>
<tr>
<td>K(_M) [mmol/L]</td>
<td>1.5</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td>K(_i) [mmol/L]</td>
<td>0.05</td>
<td>0.5</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 2.7: Range of values used for the sensitivity analysis on the operating parameters for all reaction pathways**

**Batch and Continuous**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min. Value</th>
<th>Std. Value</th>
<th>Max. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(<em>{C_4H</em>{10}}(t = 0)) [mmol/L(_{reactor})]</td>
<td>25</td>
<td>250</td>
<td>2500</td>
</tr>
<tr>
<td>C(<em>{O_2}(t = 0)) [mmol/L(</em>{reactor})]</td>
<td>500</td>
<td>1000</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Continuous**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min. Value</th>
<th>Std. Value</th>
<th>Max. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D [hr(^{-1})]</td>
<td>0.25 \times \bar{\mu} (Kin)</td>
<td>\bar{\mu} (Kin)</td>
<td>4 \times \bar{\mu} (Kin)</td>
</tr>
</tbody>
</table>

For the variation in the initial substrate concentrations, the ratio between the C\(_4\)H\(_{10}\) and O\(_2\) concentrations was kept constant. For variations in the concentration of O\(_2\), the standard value for the initial concentration of C\(_4\)H\(_{10}\) was used (250 mmol/L) and the initial concentration of O\(_2\) was varied between 500 and 2000 mmol/L. This sensitivity analysis resulted in a variation in the relative substrate concentrations and hence the presence of excess and limiting substrates.

For each flux analysis point obtained, two sample points were used to determine the averaged net accumulation and cellular dilution terms along with the extracellular reaction rates. The sample time around these analysis points was kept very small as to remove any bias due to the inaccurate determination of these rates over the analysis time.
2.4 Metabolic flux simulations

Once the reaction data was generated under the various conditions, the data was used to populate the
flux models using constraints for the extracellular fluxes. The number of constraints used for each
simulation was dependent on the degrees of freedom for each reaction pathway. Choosing different
extracellular fluxes to constrain the reaction matrix can result in changes to the condition number of
the reaction matrix. This is due to the change in the linear dependency of the cellular mass balances
in the reaction matrix. A detailed analysis of these affects on the model accuracy was conducted in
Chapter 3.

Once the extracellular reaction flux constraints were determined from the reaction data and the
extracellular reactor balances, the linearised flux models were used to simulate the metabolic fluxes for
each reaction in the selected reaction pathway. The computed metabolic fluxes were then compared
to the actual reaction rates determined from the reaction kinetic data obtained from the integration
routine. The flux simulation procedure has been discussed in more detail in this section.

2.4.1 Constraint flux calculation

The extracellular constraint fluxes were calculated using the extracellular metabolite balances. The
fluxes were determined using only the time derivatives of the extracellular metabolite concentrations
and the reactor operating parameters. This methodology has been used extensively by many researchers
to determine the extracellular reaction fluxes (Gombert and Nielsen 2000, Stephanopoulos et al. 1998;
Vallino and Stephanopoulos 1994a,b and Bailey 1991). The formulae used to compute the constraint
fluxes under the various modes of operation has been summarised using the \( \text{C}_4\text{H}_{10} \) metabolite (Table
2.8). The same formulae was used for all other extracellular metabolites. The only difference was the
sign of the time derivative \( \left( \frac{\Delta C_{\text{ext}}^i}{\Delta t} \right) \) for the extracellular products (H\(_2\), CO\(_2\) and H\(_2\)O). The correlation used to compare the reaction fluxes and the cellular reaction rates was obtained using Eq. \ref{eq:2.5} where \( \hat{V} \) is the specific intracellular volume (see section 2.2) and \( \text{Rate}_i \) is the intracellular reaction rate computed with the kinetic data from the ODE routine.

\[
F_i = \text{Rate}_i \times \hat{V}
\]  

(2.5)

Table 2.8: Example of the formulae used to calculate the extracellular flux constraints for all modes of operation for all reaction pathways

<table>
<thead>
<tr>
<th>Mode of Operation</th>
<th>Flux Calculation [mmol/gDW.hr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Batch</td>
<td>( F_{C_4H_{10}} = \left[ -\frac{\Delta C_{\text{ext}}^{C_4H_{10}}}{\Delta t} \right] \left( \frac{1}{\hat{C}_v} \right) )</td>
</tr>
<tr>
<td>c) Continuous</td>
<td>( F_{C_4H_{10}} = \left[ -\frac{\Delta C_{\text{ext}}^{C_4H_{10}}}{\Delta t} \right] + D \left( C_{C_4H_{10}}^{\text{feed}} - C_{C_4H_{10}}^{\text{ext}} \right) \left( \frac{1}{\hat{C}_v} \right) )</td>
</tr>
</tbody>
</table>

The value of \( \Delta t \) was kept constant for all data points. This was achieved using specified integration times (See Table 2.9). This was done to keep the relative slopes of the concentration profiles constant throughout the reaction time thereby removing any bias of the gradients from the flux computations. Therefore the reaction flux calculated at time of 1 hr was determined using the concentration and reactor data obtained at 0.943 and 1.057 Hrs.

Table 2.9: Selection of data analysis and the corresponding integration times for the computation of the extracellular reaction fluxes for all reaction pathways

<table>
<thead>
<tr>
<th>Data Analysis Time [Hrs]</th>
<th>Integration Time 1 [Hrs]</th>
<th>Integration Time 2 [Hrs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.943</td>
<td>1.057</td>
</tr>
<tr>
<td>5</td>
<td>4.943</td>
<td>5.057</td>
</tr>
<tr>
<td>10</td>
<td>9.943</td>
<td>10.057</td>
</tr>
<tr>
<td>100</td>
<td>99.943</td>
<td>100.057</td>
</tr>
<tr>
<td>1000</td>
<td>999.943</td>
<td>1000.057</td>
</tr>
</tbody>
</table>

2.4.2 Linearised reaction flux models

Once the extracellular flux constraints were calculated, the flux model was used to simulate the fluxes for the entire reaction pathway. The flux models were all based on the reaction stoichiometry of the
internal cellular mass balances for the various reaction pathways. Three different models were used for the subsequent calculation of the reaction fluxes. For the first model (model A), the fluxes were calculated using a constraint vector $b$ populated with the intracellular net accumulation rate ($dC_i^{\text{int}}/dt$) and dilution rates ($\mu C_i^{\text{int}}$) for each intracellular metabolite balance. These terms have been denoted as $\text{DynT}$ and $\text{DilT}$ respectively in the models. Because model A was fully populated with the intracellular reaction data, it was the most accurate model that could be used for the linear computation of the reaction pathway. This model is not used in any practical systems as it requires the intracellular concentration data to populate the constraint vector. Stephanopoulos et al. (1998) and Bailey (1991) indicate that to obtain accurate intracellular concentration data can be very difficult in many reaction systems. However, results from model A were included in the study to depict the computational inaccuracies of the MFA model. An example of how model A was populated for reaction pathway $DF_1$ has been shown in Eq. 2.6.

The vector was populated with the accumulation and cellular dilution rate data obtained at each integration time interval discussed in section 2.4.1.

$$b_{\text{int}}(t) = \hat{V} (\text{DynT}(t) + \text{DilT}(t))$$

$$\begin{bmatrix}
    b_1 \\
    b_2 \\
    b_3 \\
    b_4 \\
    b_5 \\
    b_6 \\
    b_7 \\
\end{bmatrix} = \hat{V} \begin{bmatrix}
    \frac{\Delta C_{C_4H_{10}}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{O_2}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{C_2H_6}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{C_2H_4}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{H_2}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{CO_2}^{\text{int}}}{\Delta t} \\
    \frac{\Delta C_{H_2O}^{\text{int}}}{\Delta t}
\end{bmatrix} + \begin{bmatrix}
    \mu C_{C_4H_{10}}^{\text{int}} \\
    \mu C_{O_2}^{\text{int}} \\
    \mu C_{C_2H_6}^{\text{int}} \\
    \mu C_{C_2H_4}^{\text{int}} \\
    \mu C_{H_2}^{\text{int}} \\
    \mu C_{CO_2}^{\text{int}} \\
    \mu C_{H_2O}^{\text{int}}
\end{bmatrix}$$  \hspace{1cm} (2.6)

**Model A (DF$_1$)**

For model B, the constraint vector $b$ was populated with the intracellular net accumulation rate ($dC_i^{\text{int}}/dt$) only. This model was used to determine the effect of the cellular dilution rate on the computational accuracy of the flux model. This model is not used in any practical systems as it requires the intracellular concentration data to populate the constraint vector. An example of how model A was populated for reaction pathway DF$_1$ has been shown in Eq. 2.7.
2.4. METABOLIC FLUX SIMULATIONS

\[
b_{\text{int}}(t) = \hat{V} \left( D_{\text{DynT}}(t) \right)
\]

\[
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3 \\
  b_4 \\
  b_5 \\
  b_6 \\
  b_7
\end{bmatrix} = \hat{V} \begin{bmatrix}
  \frac{\Delta C_{\text{C$_2$H$_{10}$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{D$_2$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{C$_2$H$_6$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{C$_5$H$_4$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{H$_2$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{CO$_2$}}^{\text{int}}}{\Delta t} \\
  \frac{\Delta C_{\text{H$_2$O}}^{\text{int}}}{\Delta t}
\end{bmatrix}
\]

(2.7)

**Model B (DF$_1$)**

For model C, the constraint vector was populated with zero elements only. Model C provided a flux model that was analogous to those obtained using conventional MFA balances. Therefore Model C is an MFA model. Eqs. 2.6 to 2.8 illustrate the models which were used to solve the flux network for DF$_1$.

\[
b_{\text{int}}(t) = 0
\]

\[
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3 \\
  b_4 \\
  b_5 \\
  b_6 \\
  b_7
\end{bmatrix} = \begin{bmatrix}
  0 \\
  0 \\
  0 \\
  0 \\
  0 \\
  0 \\
  0
\end{bmatrix}
\]

(2.8)

**Model C (DF$_1$)**

Eq. 2.9 summarises the mathematical description of the equation sets used to calculate the reaction vector \( \mathbf{v} \) using the reaction stoichiometric matrix \( \mathbf{S} \) and the constraint vector \( b_{\text{int}} \) (described above). For the computation of Models A, B and C, only the constraint vector \( b_{\text{int}} \) was changed in the linear
2.4. METABOLIC FLUX SIMULATIONS

balance described by Eq 2.9. This formula has been well described in literature and used in a number of flux simulations (Stephanopoulos et al., 1998). The full set of reaction stoichiometric matrices and constraint vectors have been included in Appendix B. Eq. 2.10 is an example of the model used to simulate the fluxes for reaction pathway DF₁. In this example, model C has been used to simulate the fluxes and \( F_{C_2H_10} \) or \( F_1 \) has been used to constrain the flux model.

\[
S \nu = b_{\text{int}} \tag{2.9}
\]

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -4 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & -2 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 1.75 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 3 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1.5 & 0 & -1 & 0
\end{bmatrix}
\begin{bmatrix}
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_7 \\
F_8
\end{bmatrix}
= \begin{bmatrix}
-F_1 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix} \tag{2.10}
\]

2.4.3 Implementation of numerical linear solvers for flux simulations

The numerical linear solvers that were used for the flux simulations were initiated once the flux models were developed. Two linear algebra routines were used to solve the reaction flux networks. LU factorisation (LUf) and SVD were chosen as linear algebra routines. These routines were chosen based on the information provided in the review of the different numerical routines (Section 1.3.1 of Chapter 1). For the linear programming simulations, there was no obvious objective function which needed to minimised for the flux simulations. In standard MFA studies that use linear programming routines, either maximisation of the biomass or product formation rate is used define the objective function (Famili et al., 2005 and Hatzimanikatis et al., 1996). Therefore a generalised objective function was developed across all the systems for the minimisation of the overall cell mass balance to ensure consistency in the flux models. The objective function \( g \) has been summarised below:

\[
g = \sum_{j=1}^{n} \Theta_j M_j F_j + \sum_{i=1}^{m} \hat{V} M_i (DynT_i) \tag{2.11}
\]
### 2.4. METABOLIC FLUX SIMULATIONS

$\Theta_j$ is used to differentiate between extracellular fluxes in and out of the cell, $M_j$ represents the molar mass of component $j$ and $F_j$ represents the extracellular flux of that component. $\sum_{i=1}^{m} M_i (D_{y_i} T_i)$ represents the mass balances for all intracellular metabolites. $n$ is set to 5, to sum through the 5 extracellular fluxes, and $m$ is set to 7 to sum through the 7 intracellular metabolites.

When the linear programming routines were used, a number of equation forms were developed for each simulation. The different equation forms were obtained through the combination of equality and inequality constraints for the extracellular fluxes and the internal cellular mass balances.

The four equation forms have been described below (Eq. 2.12 - 2.15). The parameters used in these equations have been summarised in Table 2.10.

**EqForm$_1$ (LP$_1$):** Extracellular fluxes and all internal cellular mass balances were constrained with a set of inequality constraints

\[
S_v \leq b_{int} \tag{2.12}
\]

\[
W_{ext} v_{ext} \leq b_{ext}
\]

\[
v \geq lb
\]

**EqForm$_2$ (LP$_2$):** External fluxes and all internal cellular mass balances were constrained with a set of equality constraints

\[
S_v = b_{int} \tag{2.13}
\]

\[
W_{ext} v_{ext} = b_{ext}
\]

\[
v_{all} \geq lb
\]

**EqForm$_3$ (LP$_3$):** External fluxes were constrained with a set of equality constraints and the internal cellular mass balances were constrained with a set of inequality constraints

\[
S_v \leq b_{int} \tag{2.14}
\]

\[
W_{ext} v_{ext} = b_{ext}
\]

\[
v_{all} \geq lb
\]
2.4. METABOLIC FLUX SIMULATIONS

EqForm$_4$ (LP$_4$): External fluxes were constrained with a set of inequality constraints and the internal cellular mass balances were constrained with a set of equality constraints

\[
S_v = b_{int} \\
W_{ext}v_{ext} \leq b_{ext} \\
v_{all} \geq lb
\]  \hspace{1cm} (2.15)

Table 2.10: Linear programming model parameters for all reaction pathways

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{int}$</td>
<td>Internal cellular mass balance constraint vector</td>
</tr>
<tr>
<td>$b_{ext}$</td>
<td>Extracellular flux constraint vector</td>
</tr>
<tr>
<td>$lb$</td>
<td>Lower bound flux constraint vector</td>
</tr>
<tr>
<td>$S$</td>
<td>Stoichiometric matrix</td>
</tr>
<tr>
<td>$v$</td>
<td>Reaction flux vector</td>
</tr>
<tr>
<td>$v_{ext}$</td>
<td>Extracellular reaction flux vector</td>
</tr>
<tr>
<td>$W_{ext}$</td>
<td>Extracellular reaction flux constraint matrix</td>
</tr>
</tbody>
</table>

The vector $b_{ext}$ contains the external flux constraints obtained from external concentration data and $b_{int}$ contains the constraints for the internal metabolite balances only (See Eq. 2.6 - 2.8). The lower bounds ($lb$) for all the equation forms were set to 0. The absolute values of the consumption rates for the extracellular substrates, which come out as negative fluxes when computed with the concentration data, were used in the vector $b_{ext}$. $W_{ext}$ describes the matrix used to determine which extracellular fluxes were constrained for each computation. The elements of this matrix were altered to obtain various constraint sets for each routine. The vector $v_{ext}$ contains all the extracellular fluxes for each system. A more detailed description of how the constraint sets were developed for each routine has been included in Chapter 3. The LP routines conducted using the equation forms described by Eqs. 2.12 to 2.15 have been labelled as LP$_1$, LP$_2$, LP$_3$ and LP$_4$ respectively.

2.4.4 Error analysis for flux simulations

The accuracy of the simulated flux profiles were determined through comparison with the actual kinetic reaction rates. The kinetic reaction rate data was determined from the concentration data at each of
the integration time steps provided to the ordinary differential equation (ODE) routine. The fluxes determined from the flux balance were first converted to the intrinsic rates using the specific volume ($\hat{V}$). A relative error analysis was performed to determine the simulation accuracies obtained for each model under various conditions. The formula used to calculate the errors is described by Eq. 2.16.

$$\text{Err}_{\text{Rate}_i} = 100 \left( \frac{\| \text{Rate}_{i,\text{actual}} - \left( \frac{F_i,\text{predicted}}{\hat{V}} \right) \|}{\text{Rate}_{i,\text{actual}}} \right)$$ (2.16)

The $\left( \frac{F_i,\text{predicted}}{\hat{V}} \right)$ term represents the predicted reaction rate and is directly proportional to the predicted flux as the specific volume of the cells remained constant under all conditions. All references that have been made to the accuracy of the estimated reaction rates in the results are therefore directly comparable to the accuracy of the estimated fluxes. The term reaction volume has been used to describe the intracellular volume where all reaction takes place. This is equivalent to the volume inside the biomass where the enzyme catalysed reactions are taking place in a biological process (See Figure 2.1 in Section 2.1.1.1).

### 2.5 Computation details

Simulations were conducted using a Lenovo® T61 ThinkPad notebook with an Intel® Core™ 2 Duo CPU T8300 @ 2.40GHz and 3GB of RAM. The Linux based Ubuntu release 10.04 (Lucid) operating system was used. All numerical simulations were conducted using the Unix version of Matlab 7.6.0 (R2008a). The concentration profiles were generated using the built in stiff solver ODE15s. This is a stiff ODE solver of variable order that uses numerical differentiation formulae. Because of the extrapolative nature of the solver, only the data at end points of the integration spans are outputted to the main function file. Therefore specified time points were supplied to the integration routine and all data obtained at each point was used as the “experimental” concentration data obtained at the specified reaction time.

A stiff solver was used due to the stiffness between the balance equations created by the variations in the kinetic and operating parameters in the kinetic reaction rates and the high intracellular concentrations. All profiles for both the extracellular and intracellular metabolites were generated and saved along with
the raw data which was written as delimited text to text files. The MFA balances were solved using MATLAB’s LINSOLVE routine which uses an LU factorisation with built in partial pivoting and the SVD routine which computes the SVD matrices for the system which were used to solve the system using Eq. 1.9. The LINPROG routine was used to solve the flux systems using linear programming. The routine uses a projection algorithm derived from the standard simplex algorithm.

Figure 2.2: Computational model flow diagram

A graphical representation of the computational methodology used has been shown in Figure 2.2. The numbered arrows indicate the passing of data between the main function file and the calling functions for the built-in numerical routines. The solid horizontal lines in the main function file represent the positions from which external calls were made to the external functions. The bullets shown in the main function file represent the parameters which were set within nested loops. For each simulation, a single value was selected for each of the kinetic parameters and operating conditions with a single reaction mechanism. The selection of the parameters in the main function file took place within the nested loops resulting in the sensitivity analysis for the various linearised reaction models.

The numbered bullets shown in the external function files represent the variables that were calculated, with the last variable representing the variable which was passed back into the main function file.
for further computation. The arrow numbered 4 represents the final outputs of the computational routine which were written to delimited text files and various graphs. Table 2.11 provides a summary describing the movement of variables between the main function file and the secondary function files.

<table>
<thead>
<tr>
<th>Arrow</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Kinetic &amp; Operating parameters</td>
</tr>
<tr>
<td>1b</td>
<td>Dynamic concentration profiles</td>
</tr>
<tr>
<td>2a</td>
<td>1a &amp; 1b</td>
</tr>
<tr>
<td>2b</td>
<td>Kinetic reaction rates</td>
</tr>
<tr>
<td>3a</td>
<td>Flux balances</td>
</tr>
<tr>
<td>3b</td>
<td>Calculated flux profile</td>
</tr>
<tr>
<td>4</td>
<td>Conc. &amp; reaction profiles, prediction errors</td>
</tr>
</tbody>
</table>

The kinetic reaction rates illustrated by arrow 2b represent the actual rates which were obtained from the concentration data generated by the ODE routine. These rates were used in the error analysis with the reaction flux vector (arrow 3b) to determine the accuracy of each model under the various conditions. The source code used in Matlab 7.6.0 (R2008a) for the computation of the results has been included in Appendix D.

### 2.6 Research approach

A brief description of the approach used in each investigation has been included in this section. The description makes reference to the various methodologies which were used in each section of the study. This has been done to provide easier reference between the results included in each chapter and the methodology which was used to generate the reaction data, calculate the reaction fluxes and determine the accuracy of each flux model. The descriptions of each methodology included in this section refer to the general methodologies which were used. More specific methodologies that pertain to the results generated in Chapters 3, 4 and 5 have been included in the short methodology sections of those chapters. This was done to keep the specific methodologies separate from the overall methodologies discussed in this chapter.
2.6. RESEARCH APPROACH

2.6.1 Batch flux simulations

The following approach was taken to obtain the results for the batch flux simulations which have been shown in Chapters 3, 4 and 5. A summary of the procedures used to obtain the results have been shown in Table 2.12 along with the references for each procedure and the page number for each of these references.

Table 2.12: References for all procedures used to obtain the results for the batch flux simulations

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Reference</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction Pathway</td>
<td>Figure 2.1</td>
<td>34</td>
</tr>
<tr>
<td>Reaction stoichiometry</td>
<td>Section 2.1.1</td>
<td>28</td>
</tr>
<tr>
<td>Kinetic Mechanism</td>
<td>Table 2.1</td>
<td>35</td>
</tr>
<tr>
<td>Kinetic Parameter</td>
<td>Table 2.6</td>
<td>41</td>
</tr>
<tr>
<td>Operating Parameter</td>
<td>Table 2.7</td>
<td>41</td>
</tr>
<tr>
<td>Cellular Mass Balance</td>
<td>Eq. 2.1</td>
<td>37</td>
</tr>
<tr>
<td>Reactor Mass Balance</td>
<td>Table 2.4a</td>
<td>38</td>
</tr>
<tr>
<td>Constraint Fluxes</td>
<td>Table 2.8a</td>
<td>43</td>
</tr>
<tr>
<td>Flux Models</td>
<td>Eqs. 2.6-2.8</td>
<td>44</td>
</tr>
<tr>
<td>Linear Solvers</td>
<td>Section 2.4.5</td>
<td>46</td>
</tr>
<tr>
<td>Error Analysis</td>
<td>Eq. 2.16</td>
<td>49</td>
</tr>
</tbody>
</table>

2.6.2 Steady state flux simulations

The following approach was taken to obtain the results for the continuous flux simulations which have been shown in Chapter 5. A summary of the procedures used to obtain the results have been shown in Table 2.13 along with the references for each procedure and the page number for each of these
Table 2.13: References for all procedures used to obtain the results for the steady state flux simulations

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Reference</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction Pathway</td>
<td>Figure 2.1</td>
<td>34</td>
</tr>
<tr>
<td>Reaction stoichiometry</td>
<td>Section 2.1.1</td>
<td>28</td>
</tr>
<tr>
<td>Kinetic Mechanism</td>
<td>Table 2.1</td>
<td>35</td>
</tr>
<tr>
<td>Kinetic Parameter</td>
<td>Table 2.6</td>
<td>41</td>
</tr>
<tr>
<td>Operating Parameter</td>
<td>Table 2.7</td>
<td>41</td>
</tr>
<tr>
<td>Cellular Mass Balance</td>
<td>Eq. 2.1</td>
<td>37</td>
</tr>
<tr>
<td>Reactor Mass Balance</td>
<td>Table 2.4b</td>
<td>38</td>
</tr>
<tr>
<td>Constraint Fluxes</td>
<td>Table 2.8b</td>
<td>43</td>
</tr>
<tr>
<td>Flux Models</td>
<td>Eqs. 2.6 - 2.8</td>
<td>44</td>
</tr>
<tr>
<td>Linear Solvers</td>
<td>Section 2.4.3</td>
<td>46</td>
</tr>
<tr>
<td>Error Analysis</td>
<td>Eq. 2.16</td>
<td>49</td>
</tr>
</tbody>
</table>
Part II

Results and Discussion
Chapter 3

Batch reaction profiles

3.1 Introduction

The use of linearised models for the computation of biochemical reaction fluxes has been well described in Chapters 1 and 2. These linearised reaction models require the use of extracellular concentration data and knowledge of the reaction stoichiometry only. Therefore no consideration needs to be given to the kinetic structure of the reaction pathway nor the distribution of intermediates within the intracellular reaction volume. This allows for the simplification of the computation procedures and reduces the need for large sets of kinetic and thermodynamic data to populate the reaction models (Stephanopoulos et al., 1998). These simplified models are therefore very beneficial as Bornholdt (2005) indicates that it can be very difficult to obtain accurate intrinsic kinetic data sets for the estimation of the parameters and the forms of the kinetic mechanisms, even for well understood biochemical reaction pathways.

The primary objectives of this chapter are to validate the model data generated from the batch reaction simulations using equivalent data obtained from literature studies and to perform a qualitative review of the accuracy of the estimated reaction flux models for the various kinetic mechanisms. The secondary objectives of this chapter are to highlight the considerations required when computing the reactions fluxes in a highly branched reaction pathway. Under these situations, the linear dependency of the reaction matrix becomes an important factor when evaluating the accuracy of linearised flux models.
3.1. INTRODUCTION

The generation of kinetic reaction data and subsequent flux estimations obtained for each flux model have been described in detail in Sections 2.1 to 2.4 of Chapter 2. The methodology specific to the generation of the kinetic reaction and estimated flux profiles for each analysis performed in this chapter have been discussed in the methodology section below.
3.2 Methodology

The first set of results provide a graphical representation of how the linearity of each kinetic mechanism affected the accuracy of each reaction flux model under standard conditions. The analysis was conducted using the kinetic reaction profiles and the estimated reaction fluxes obtained for each flux model.

The second set of results provide a graphical representation of how the conditioning of the reaction matrix affected the accuracy of flux model C (model equivalent to MFA implemented in literature). Computation of the reaction flux profile can be done using various constraint sets due to the degrees of freedom available in the reaction pathways. The constraint sets are merely combinations of different extracellular fluxes which were used to constrain the reaction flux model. This is equivalent to a real life biochemical reaction system which has a number of extracellular reaction fluxes that can be measured independently and relatively simply using extracellular metabolite concentrations as well as by-product evolution rates such as CO$_2$ and biomass production using well established laboratory techniques (Yang et al. (2010) and Huang (2003)).

The final set of results provide a graphical representation of how various linear solvers handle highly dependent reaction matrices and the implications on the accuracy of estimating reaction fluxes for highly branched reaction pathways which have poorly conditioned reaction matrices.

3.2.1 Computation of the kinetic reaction profiles and estimated reaction fluxes

The reaction profiles were calculated for each kinetic mechanism using the concentration data generated by the ODE routine under standard conditions. This was done to overcome the random specification of the reaction time steps generated by the stiff ODE solver. It must be noted that these reaction profiles are considered to be accurate representations of the actual reaction data as they were generated using the actual kinetic mechanisms, parameters and concentration data extracted from the ODE sub-routine used to compute the reactor and cellular mass balances.

The first step in estimation of the reaction flux profiles was to compute the extracellular fluxes used
to constrain the model from the simulated extracellular metabolite concentration data generated by the ODE routine. The computed extracellular flux along with the reaction matrix was used to estimate the reaction flux profiles for each model. The variability in the accuracy of the reaction flux models (models A, B and C) was dependent only on the assumptions used in the formulation of each model as the same concentration data and reaction matrices were used for each model. The reaction fluxes for models A, B and C were computed using the linear equations shown in Section 2.4.2 of Chapter 2. The estimated reaction fluxes were plotted against the actual reaction profiles calculated from the concentration data to determine the accuracy of the fit obtained for each model with the various kinetic mechanisms.

3.2.2 Computation of the reaction fluxes using various constraint sets

The reaction fluxes for model C were computed using various flux constraint sets. To illustrate the effect of the reaction flux constraint sets on the accuracy of the flux model, the reaction fluxes for model C were computed for reaction pathway DF_2. This pathway was used as the model pathway for this analysis because of the larger number of constraint sets that can be obtained compared to reaction pathway DF_1 (Table 3.1). This is due to the higher number of degrees of freedom in constraining the reaction pathway DF_2.

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Possible Constraint Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF_1</td>
<td>5</td>
</tr>
<tr>
<td>DF_2</td>
<td>10</td>
</tr>
<tr>
<td>DF_3</td>
<td>10</td>
</tr>
</tbody>
</table>

The data shown in Table 3.1 illustrates the combination of extracellular reaction fluxes used to constrain the reaction flux matrix. From the different constraint sets, reaction matrices were generated to mathematically describe the reaction pathway. The use of 1 indicates that the flux was constrained in that set and a 0 indicates that the flux was an unknown that had to be solved for. An example of two reaction matrices generated using constraint sets 3 and 5 for reaction pathway DF_2 have been illustrated in Table 3.3. The significance of the choice of these two constraint sets as the examples is clearly illustrated by the results shown in section 3.3.2.
3.2. METHODOLOGY

Table 3.2: Constraint sets used for pathway DF$_2$

<table>
<thead>
<tr>
<th>Constr. Set</th>
<th>$F_{C_4H_{10}}^{ext}$</th>
<th>$F_{H_2}^{ext}$</th>
<th>$F_{O_2}^{ext}$</th>
<th>$F_{CO_2}^{ext}$</th>
<th>$F_{H_2O}^{ext}$</th>
<th>Constr. Set</th>
<th>$F_{C_4H_{10}}^{ext}$</th>
<th>$F_{H_2}^{ext}$</th>
<th>$F_{O_2}^{ext}$</th>
<th>$F_{CO_2}^{ext}$</th>
<th>$F_{H_2O}^{ext}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

It is understood that in practise, the choice of constraint sets in a practical system will be based on the ease at which each extracellular metabolic concentration and flux can be obtained from current laboratory techniques and the inherent experimental error of each of these techniques.

Table 3.3: Equations used for the reaction pathway DF$_2$ with the constraint sets 3 and 5

<table>
<thead>
<tr>
<th>Constraint Set</th>
<th>Solution Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$\begin{bmatrix} -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 1 &amp; -3.5 &amp; -4 &amp; 0 &amp; 0 \ 1 &amp; -1 &amp; 0 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 \ 1 &amp; 1 &amp; 0 &amp; 0 &amp; 0 &amp; -2 &amp; 0 &amp; 0 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; 0 &amp; 1.75 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 &amp; 0 \end{bmatrix} \begin{bmatrix} F_2 \ F_3 \ F_4 \ F_5 \ F_6 \ F_7 \ F_8 \ F_9 \end{bmatrix} = \begin{bmatrix} -F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \ F_{C_4H_{10}}^{ext} \end{bmatrix}$</td>
</tr>
<tr>
<td>5</td>
<td>$\begin{bmatrix} 1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 1 &amp; -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; -2 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; 0 &amp; 1.75 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; -1 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; 0 &amp; -1 &amp; 0 \end{bmatrix} \begin{bmatrix} F_1 \ F_2 \ F_3 \ F_4 \ F_5 \ F_6 \ F_7 \ F_8 \end{bmatrix} = \begin{bmatrix} 0 \ -F_{O_2}^{ext} \ F_{O_2}^{ext} \ F_{O_2}^{ext} \ F_{O_2}^{ext} \ F_{O_2}^{ext} \ F_{O_2}^{ext} \ F_{O_2}^{ext} \end{bmatrix}$</td>
</tr>
</tbody>
</table>

To maintain uniformity of the reaction data in this analysis, only kinetic mechanism Kin$_1$ was used to generate the concentration data. The same methodology used to compute the reaction data and reaction fluxes discussed in Section 3.2.1 was used to generate the results for this analysis.
3.2. METHODOLOGY

3.2.3 Computation of the reaction fluxes using various linear solvers

The reaction fluxes for model C were computed using various linear routines for reaction pathway DF3. Reaction pathway DF3 has a higher condition number than the other reaction pathways due to the higher degree of branching in the cellular mass balance. As was the case with the DF2 reaction pathway, the DF3 reaction pathway also has a larger number of constraint sets than pathway DF1. The constraint set for this analysis was fixed to ensure that the variability observed in the results was only due to the variation in the linear routine used to compute the reaction fluxes. The linear algebra routines, LU factorisation and singular value decomposition (SVD), along with the linear programming routine were used to compute the reaction fluxes with the same concentration data and reaction parameters. Sections 2.4.2 and 2.4.3 of Chapter 2 provide the details of each linear routine used to compute the reaction fluxes. The reaction matrices generated for the linear algebra routines have been shown in Table 3.4. The reaction matrices and the constraint vectors generated for the linear programming routine have been shown in Table 3.5.

Table 3.4: Equations used for linear algebra routines with reaction pathway DF3 and constraint set 5

<table>
<thead>
<tr>
<th>Routine</th>
<th>Solution Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUF &amp; SVD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\begin{bmatrix} -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ -6.5 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; 0 &amp; -2 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; -1 &amp; 0 &amp; 1.75 &amp; 0 \ 4 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 \ 5 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 \end{bmatrix} \begin{bmatrix} F_2 \ F_3 \ F_4 \ F_5 \ F_6 \ F_8 \ F_9 \end{bmatrix} = \begin{bmatrix} -F_{\text{ext}}^{\text{C}<em>{4}H</em>{10}} \ -F_{\text{ext}}^{\text{CO}<em>2} \ -F</em>{\text{ext}}^{\text{O}_2} \end{bmatrix}$</td>
</tr>
</tbody>
</table>
Table 3.5: Equations used for the linear programming routine with reaction pathway DF$_3$ and constraint set 5

<table>
<thead>
<tr>
<th>Routine</th>
<th>Solution Matrix</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP$_1$</td>
<td>$\begin{bmatrix} -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \negF_2 &amp; \negF_1 \ -6.5 &amp; 0 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 \negF_3 &amp; \negF_7 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \negF_4 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; 0 &amp; -2 &amp; 0 \negF_5 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; -1 &amp; 1 &amp; 0 &amp; 0 \negF_6 &amp; 1.75 \ 4 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 \negF_8 &amp; \negF_{10} \ 5 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 \negF_9 &amp; 0 \end{bmatrix}$</td>
<td>$\begin{bmatrix} F_1 \ F_7 \ F_{10} \end{bmatrix} \leq \begin{bmatrix} F_{CO_{2}}^{ext} \ F_{O_{2}}^{ext} \end{bmatrix}$</td>
</tr>
<tr>
<td>LP$_2$</td>
<td>$\begin{bmatrix} -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \negF_2 &amp; \negF_1 \ -6.5 &amp; 0 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 \negF_3 &amp; \negF_7 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \negF_4 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; 1 &amp; 2 &amp; 0 \negF_5 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; -1 &amp; 0 &amp; 0 &amp; 1.75 \negF_6 &amp; 0 \ 4 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 \negF_8 &amp; \negF_{10} \ 5 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 \negF_9 &amp; 0 \end{bmatrix}$</td>
<td>$\begin{bmatrix} F_1 \ F_7 \ F_{10} \end{bmatrix} = \begin{bmatrix} F_{CO_{2}}^{ext} \ F_{O_{2}}^{ext} \end{bmatrix}$</td>
</tr>
<tr>
<td>LP$_3$</td>
<td>$\begin{bmatrix} -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \negF_2 &amp; \negF_1 \ -6.5 &amp; 0 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 \negF_3 &amp; \negF_7 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \negF_4 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; 0 &amp; -2 &amp; 0 \negF_5 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; -1 &amp; 0 &amp; 0 &amp; 1.75 \negF_6 &amp; 0 \ 4 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 \negF_8 &amp; \negF_{10} \ 5 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 \negF_9 &amp; 0 \end{bmatrix}$</td>
<td>$\begin{bmatrix} F_1 \ F_7 \ F_{10} \end{bmatrix} \leq \begin{bmatrix} F_{CO_{2}}^{ext} \ F_{O_{2}}^{ext} \end{bmatrix}$</td>
</tr>
<tr>
<td>LP$_4$</td>
<td>$\begin{bmatrix} -1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \negF_2 &amp; \negF_1 \ -6.5 &amp; 0 &amp; 0 &amp; 0 &amp; -3.5 &amp; -4 &amp; 0 \negF_3 &amp; \negF_7 \ 0 &amp; 1 &amp; -1 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \negF_4 &amp; 0 \ 0 &amp; 1 &amp; 1 &amp; 0 &amp; 0 &amp; -2 &amp; 0 \negF_5 &amp; 0 \ 0 &amp; 0 &amp; 1 &amp; -1 &amp; 0 &amp; 0 &amp; 1.75 \negF_6 &amp; 0 \ 4 &amp; 0 &amp; 0 &amp; 0 &amp; 2 &amp; 3 &amp; 0 \negF_8 &amp; \negF_{10} \ 5 &amp; 0 &amp; 0 &amp; 0 &amp; 3 &amp; 1.5 &amp; -1 \negF_9 &amp; 0 \end{bmatrix}$</td>
<td>$\begin{bmatrix} F_1 \ F_7 \ F_{10} \end{bmatrix} \leq \begin{bmatrix} F_{CO_{2}}^{ext} \ F_{O_{2}}^{ext} \end{bmatrix}$</td>
</tr>
</tbody>
</table>

Constraint set 5 was selected for this analysis because this set generated the lowest condition number of all constraint sets for reaction pathway DF$_3$ (Results in Figure 5.11 of Section 5.3.2). This provided the best possibility of obtaining more accurate estimations of the reaction fluxes with the linear algebra routines as these routines become less accurate at higher conditions numbers (Heath, 2001). A summary of the constraint sets available for reaction pathway DF$_3$ has been shown in Table 3.6.
3.2. METHODOLOGY

The same methodology used to compute the reaction rates and fluxes from the concentration data discussed in Section 3.2.1 was used to generate the results for this analysis. A summary of the methodologies used to generate the reaction profiles in each results section has been shown in Figure 5.1. In the graphics, the dashed lines represent variables that were changed to generate the reaction rate data in each section. The graphics show the inputs and outputs from the computational routines and the final results as the reaction rate profiles generated.

In Figure 3.1a it can be seen that the only variable used to generate the first set of reaction profiles was the choice of kinetic mechanism and the flux model used to compute the estimated reaction rates. Data A obtained from the ODE solver refers to the concentration data which was used for the determination of the actual reaction rate data and the extracellular fluxes for the flux computation. Data B refers to the cellular dilution and accumulation terms computed from the cellular mass balances in the ODE routine which were used in the constraint vectors for flux models A and B. The methodologies illustrated in Figure 5.1 are a more detailed description of the methodology described in Figure 2.2 in Section 2.5 of Chapter 2. The details shown in Figure 5.1 illustrate the selection of parameters in the main function file and the transfer of data to the numerical routines.

From Figure 3.1b it can be seen that the only variable used for the generation of the reaction rate profiles in section 3.3.2 was the constraint set for the flux computation. Reaction pathway DF2 was used to generate the reaction rate data but was not varied in each simulation and therefore the input line is marked as a solid line in the graphic. The same was true for the data generated for reaction pathway DF3. Only the choice of linear solver was varied (3.1c). It is important to note from Figures 3.1b and 3.1c that the concentration data generated by the kinetic mechanisms did not change between any of the simulations. Therefore data sets A and B generated from the ODE routine remained constant. The only variation was the manner in which flux model C was computed using this data.

### Table 3.6: Constraint sets used for pathway DF3

<table>
<thead>
<tr>
<th>Constr. Set</th>
<th>$F_{C_4H_{10}}^{ext}$</th>
<th>$F_H^2$</th>
<th>$F_{O_2}$</th>
<th>$F_{CO_2}$</th>
<th>$F_{H_2O}$</th>
<th>Constr. Set</th>
<th>$F_{C_4H_{10}}^{ext}$</th>
<th>$F_H^2$</th>
<th>$F_{O_2}$</th>
<th>$F_{CO_2}$</th>
<th>$F_{H_2O}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2. METHODOLOGY

(a) Methodology used for reaction pathway DF₁ (Section 3.3.1)

(b) Methodology used for reaction pathway DF₂ (Section 3.3.2)

(c) Methodology used for reaction pathway DF₃ (Section 3.3.3)

Figure 3.1: Graphical representation of the methodologies
3.3 Results and Discussion

The results discussed in this chapter provide an overview of reaction rate data obtained for various simulations. The reaction rate data has been generated using the model concentration generated by the ODE routine for a given set of conditions. The use of reaction profiles obtained under standard conditions with all kinetic mechanisms provided a good high level description of the model accuracy for each kinetic mechanism. The reaction profiles represent the data that is commonly used from batch metabolic models for parameter estimation (Voit (2003); Hynne et al. (2001) and Rizzi et al. (1997)) or elucidation of flux distribution in the reaction pathway (Jamshidi and Palsson (2008); Lee et al. (2008) and Gombert and Nielsen (2000)). Although these profiles form the backbone of a conventional reaction modelling study, these profiles are not as useful in meeting the overall objectives of this study. They do however provide the necessary graphical overview of how each flux model was capable of predicting the reaction rates obtained using data from various kinetic systems.

The concentration data used in the computation of the reaction rates and fluxes was generated using the standard parameters listed in Tables 2.6 and 2.7 of Section 2.3 in Chapter 2.

3.3.1 Reaction rate profiles for reaction pathway DF

A summary of the reaction data obtained with each mechanism at standard conditions has been included in Table 3.7. This data has been included in this section to provide a reference to the intrinsic reaction data which is not provided in the reaction rate profiles. The data has also been used for the validation of the reaction rate data generated using the hypothetical reaction pathway and simulated reaction kinetics.

<table>
<thead>
<tr>
<th>Kin</th>
<th>$R_1$ [mmol/L.hr]</th>
<th>$R_5$ [mmol/L.hr]</th>
<th>$\mu$ [hr$^{-1}$]</th>
<th>$\left(\mu\bar{C}^\text{i}_{\text{in}}\right)$ [mmol/L.hr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kin$_1$</td>
<td>1080</td>
<td>7500</td>
<td>0.045</td>
<td>0.017</td>
</tr>
<tr>
<td>Kin$_2$</td>
<td>1130</td>
<td>7980</td>
<td>0.052</td>
<td>0.019</td>
</tr>
<tr>
<td>Kin$_3$</td>
<td>1270</td>
<td>8420</td>
<td>0.064</td>
<td>0.055</td>
</tr>
<tr>
<td>Kin$_4$</td>
<td>1350</td>
<td>9200</td>
<td>0.085</td>
<td>0.011</td>
</tr>
</tbody>
</table>
3.3. RESULTS AND DISCUSSION

The reaction profiles obtained with the Kin$_1$ mechanism have been shown in Figure 3.2. Included in the figure are the reaction profiles estimated using models A, B and C from the same set of batch reaction data and reaction stoichiometry. The reaction rates plotted in the results represent the rate of each reaction step described by the elemental reactions in Section 2.1.1 of Chapter 2. Each rate has been illustrated graphically in the reaction pathways shown in Figure 2.1 in Section 2.1.1 of Chapter 2. A logarithmic time scale was used to populate the reaction time. A reaction time of 0.1 Hrs represents a reaction time of $t \approx 0$. The limits for the reaction rate axes were standardised for all kinetic mechanisms to provide ease of comparison of the rates obtained with each flux model.

As expected, the reaction profiles obtained from model A are the most accurate estimation of the actual reaction profiles generated by the structured kinetic model (Figures 3.2a and 3.2b). This observation clearly indicates that the accuracy of the flux model was not biased by the linearisation of the reaction data nor the computational accuracy of the SVD routine used to obtain the solution for the linear flux model. It is important to note that the reaction rates plotted in Figures 3.2a and 3.2b represent the observed cellular rates, intrinsic to the intracellular reaction volume. The reaction fluxes are calculated based on the extracellular metabolite concentrations (Section 2.4.1 of Chapter 2). The plotted estimated reaction fluxes were therefore converted to an intrinsic rate using the conversion factor described in Section 2.4 of Chapter 2 for the determination of the accuracy of the cellular based rate estimation. The larger rates observed for the extracellular rates are due to the reaction stoichiometry only. The elemental reactions shown in Section 2.1.1 of Chapter 2 clearly show that the rate of CO$_2$ evolution into the extracellular reaction volume (Rate 7), is 3 times higher than the rate of biomass formation (Rate 6).

Due to the consistency between the extracellular and intracellular reaction rate data shown in Figure 3.2, a determination of the validity of the pseudo-steady state hypothesis cannot be made using the data as it is presented. For PSSH to apply in the sense of MFA models, the intracellular reaction rates or relaxation time for cellular reactions should be at least 10 times higher than the extracellular fluxes (Stephanopoulos et al., 1998). In reference to a chemical reaction system, the total mass of intermediates should be negligible compared to that of the total mass of all reacting species (van Santen and Niemantsverdriet, 1995). This data is not readily available from the reaction rate data shown in this chapter. A review of this data and its impact on the accuracy of model C will be discussed in greater detail in Chapter 5.
3.3. RESULTS AND DISCUSSION

It can be seen from the estimated reaction profiles obtained for model B (Figures 3.2c and 3.2d) that the exclusion of the cellular dilution terms from the flux model resulted in a small decrease in the accuracy of the estimated reaction profiles. However, it was evident from the review of the computational data that the errors obtained from model B were not significant. The relative error obtained for the estimation of the reaction fluxes using model B was less than 5%. The form of the reaction flux profile generated with model B is representative of that generated by the structured kinetic model. This indicates that the exclusion of the cellular dilution rates from the model only resulted in a shift in the estimated rate data around the actual reaction profile, and the overall form of the profiles was maintained by the accumulation rates which were included in the cellular mass balances with model B.

The observed shift in reaction profiles with model B was more prominent during a reaction time of between 2 to 5 Hrs. During this reaction time, the cellular growth rate ($\mu$) reached a maximum value of 0.045 hr$^{-1}$ resulting in an overall maximum cellular dilution rate ($\mu C_i^{\text{in}}$) of 0.017 mmol/gDW.hr. Experimental data presented by Canelas et al. (2009); Li et al. (2006); Siddique et al. (2004a,b) and Buchholz et al. (2001) indicate that the sum of the intracellular dilution terms obtained for an entire reaction network ranged between 0.005 - 0.04 mmol/gDW.hr. This data shows good correlation between the model pathway using structured kinetic models developed with the formulations described by Rizzi et al. (1997) and that of experimental literature studies. The results therefore indicate that the estimated reaction profiles obtained for model B were a good representation of the errors expected if model B would be applied to literature based experimental data. The data shown in Table 3.7 for all kinetic mechanisms were within the range of experimental data provided in literature. Furthermore, the cellular growth rate and average metabolic flux data was validated against experimental data in Section 2.1.2 of Chapter 2.

The low cellular dilution terms calculated from the model data and the literature data are due to the comparatively low cellular growth rate with respect to the internal reaction fluxes recorded. Therefore the assumption of constant cell mass during each calculation interval is a good approximation as the cellular dilution rate only represents a depletion of the metabolite pool due to cellular growth (Stephanopoulos et al. [1998]), rather than a deviation in the time dependency and magnitude of the metabolite pool. The depletion of metabolite pools accounts for the overall change in reaction volume due to cellular growth to maintain consistency in the cellular mass balances. A detailed sensitivity
3.3. RESULTS AND DISCUSSION

Analysis has been conducted in Chapter 4 to determine the effect of operating conditions and kinetic parameters on the overall cellular dilution rates for the intracellular metabolites.

The reaction rates predicted using flux model C were less accurate than those obtained for model B (Figures 3.2e and 3.2f). The data clearly indicates that the exclusion of the accumulation term \( \frac{dC_i}{dt} \) from the flux estimation model resulted in a fundamental shift in the form of the estimated reaction profile. The PSSH resulted in an instantaneous jump in the estimated reaction rates at a reaction time of \( t \approx 0 \). The observed jump in the estimated reaction profiles is a result of the linear scaling of each reaction to the constrained flux (Rate 1) via the stoichiometry of the reaction pathway. Rate 1 was constrained in the reaction flux matrix to obtain a determined reaction matrix. The form of the reaction profile predicted by model C is typical for a system where the PSSH is applied to estimate reaction intermediate concentrations (Fogler (1999) and van Santen and Niemantsverdriet (1995)).

The observations made for model C in Figures 3.2e and 3.2f are not consistent with the reaction profiles generated in batch experimental studies conducted by Li et al. (2006); Aminul Hoque et al. (2005); Siddique et al. (2004b); Huang (2003); Vallino and Stephanopoulos (1994a) and Vallino and Stephanopoulos (1994b). In these studies, it is clearly evident that all extracellular fluxes calculated using the extracellular concentration data do not follow the reaction form generated by the pseudo-state assumption, as observed in Figures 3.2e and 3.2f. It must be noted however that in these experimental studies, all extracellular fluxes are constrained in the flux model to obtain a determined reaction matrix.

In practice, the metabolic flux models applied to experimental data, using a metabolic reaction pathway for a given microorganism, will be under determined and will require the extracellular fluxes to be constrained. Velagapudi et al. (2007); Famili et al. (2005) and Stephanopoulos et al. (1998) describe the details around constraining the metabolic flux model using extracellular fluxes as well as the distribution of fluxes around a branch point to obtain a determined reaction matrix. The authors indicate that in many instances, linear programming techniques are used to compute an under determined reaction matrix when sufficient constraints are not available. Therefore, the model accuracy for extracellular reaction flux estimation shown in Figure 3.2c is not considered to be a direct correlation to observations made in experimental based metabolic flux analyses. The data does however provide a unique view into the overall accuracy of the pseudo-steady state assumption in predicting the reaction profiles when not all of the extracellular flux are constrained, and when a forward computation of all reaction fluxes is made using only the reaction stoichiometry. The results also show the expected
reaction profiles that will be obtained for the intracellular reactions where no rates are constrained
due to a lack of available intracellular reaction data. In Sections 3.3.2 and 3.3.3, the use of extra flux
constraints were required to constrain the pathways with higher degrees of freedom.

It is evident from the results obtained for model C that the PSSH was not valid for the estimation of
the reaction data obtained with the Kin1 mechanism at reaction times of less than 2 Hrs. During this
reaction time, a substrate conversion of approximately 15% was recorded. The intracellular reaction
rates show that the accuracy of the estimated fluxes improved as the intracellular reactions reached their
maximum levels. When the intracellular reaction rates reach the maximum values, the consumption
rate of the intracellular metabolites will be at its highest. During this period, the relative mass of
reaction intermediates decreases and the accuracy of the PSSH improves. The data indicates that the
accumulation term in the intracellular mass balances could not be excluded from the linear model as the
magnitude of the terms, relative to the fluxes in the mass balances, was significant at low conversions.
The relative errors obtained for model C as a function of the substrate conversions have been discussed
in detail in Chapter 5. The results were generated as a function of the substrate conversions to provide
a basis for the comparison of the error profiles generated under various conditions.
3.3. RESULTS AND DISCUSSION

(a) Model A (Extracellular Rates)

(b) Model A (Intracellular Rates)

(c) Model B (Extracellular Rates)

(d) Model B (Intracellular Rates)

(e) Model C (Extracellular Rates)

(f) Model C (Intracellular Rates)

Figure 3.2: Reaction rate profiles obtained with the Kin1 mechanism under standard conditions for models A, B and C

The reaction profiles obtained using the kinetic data generated with the Kin2 mechanism have been shown in Figure 3.3. The reaction rates generated using the Kin2 mechanism were higher than those obtained with the Kin1 mechanism. A maximum reaction rate of 7500 mmol/L.hr was obtained for Rate 5 with the Kin1 mechanism compared to 7980 mmol/L.hr for the same rate with the Kin2 mechanism. This indicates that the competitive inhibition term \( K_M (1 + C_i / K_i) \), did not have the same “dampening” effect on the kinetic model and the resulting reaction rates obtained with the Kin2
mechanism as the non-competitive inhibition term \( \left( k_i C_i / (1 + C_i/K_i) \right) \) present in the Kin1 mechanism. This was expected as the presence of a non-competitive inhibition term reduced the overall reaction rate constant where the competitive inhibition term will only affect the substrate affinity \( [\text{Bailey and Ollis, 1986}] \).

As was the case with the Kin1 mechanism, there was a decrease in the accuracy of model B compared to model A in the estimation of the reaction profiles with the Kin2 mechanism (Figures 3.3c and 3.3d). The relative errors were of the same magnitude as those obtained for model B with the Kin1 mechanism (< 5%). The higher \( \mu \) obtained with the Kin2 mechanism resulted in a higher cellular dilution term. However, the relative errors of below 5% indicate that this did not affect the accuracy of model B.

The data indicates that under standard conditions, the form of the inhibition term in the kinetic mechanism did not affect the accuracy of the flux simulations for models A and B. This was due to the fact that the bias associated with linearisation of the reaction data was specifically removed from the computations using very short time differentials in the computation of the reaction fluxes. It is recognised that the use of these short time intervals is not realistic in actual experimental flux studies. \( [\text{Schwender (2008); Li et al. (2006) and Huang (2003)}] \) discuss the use of time intervals that are up to one tenth of the overall reaction time, which are significantly larger than the 0.114 Hrs used in this study. It must be noted however that the use of very short time intervals was done merely to provide as accurate a representation of the reaction rates as possible with the chosen flux analysis technique.

The results obtained for model C with the Kin2 mechanism (Figures 3.3e and 3.3f) indicate that the assumption of pseudo-steady state was not suitable for the entire reaction data set generated at standard conditions. The accuracy of model C improved as the conversion increased and the PSSH became more valid. The observed sudden increase in the estimated reaction rates at reaction time \( t \approx 0 \) once again illustrated the dependence of the estimated fluxes to the constrained reaction rate. The form of the rates estimated with model C were identical to those estimated with the Kin1 mechanism and mimicked the form of Rate 1 in both cases. This shows that the validity of the pseudo-steady state assumption was not affected by the form of the inhibition term present in the kinetic mechanism, or in other words, the structure of the kinetic model.

It can be observed graphically from the results shown in Figures 3.2e, 3.2f, 3.3e and 3.3f that estimation of the reaction rates using model C would result in an inaccurate estimation of the kinetic parameters...
over the entire reaction time. This is because the model simplifies the form of the actual reaction rates and some model parameters may become redundant. However, a more detailed analysis of the affect of the kinetic parameters in each kinetic mechanism on the relative intracellular metabolite concentrations, and hence the validity of the PSSH, has been conducted in Chapter 5.

![Reaction rate profiles](image)

**Figure 3.3:** Reaction rate profiles obtained with the Kin₂ mechanism under standard conditions for models A, B and C

The data obtained with the Kin₃ mechanism under standard conditions has been shown in Figure 3.4. It is evident from the data that the accuracy of the models was not affected by the linearisation of
3.3. RESULTS AND DISCUSSION

the reaction pathway to generate the flux constraints, regardless of the form of the kinetic mechanism (Figures 3.4a and 3.4b). There is a distinct difference in the form of the reaction profiles obtained with the Kin$_3$ mechanism, and yet there was no observed effect on the accuracy of model A. Relative errors of below 1% were obtained for all rates with the Kin$_1$, Kin$_2$ and Kin$_3$ mechanisms for model A. The deviation in the form of the reaction rates obtained with the Kin$_3$ mechanism is due to the increased relevance of the saturation term $(K_M + C_i)$ in the structured kinetic model, and the lack of feed-back inhibition of the substrate consumption rate.

The recorded rate for Rate$_5$ increased to 8420 mmol/L.hr with the Kin$_3$ mechanism due to the removal of the inhibition on the reaction rates. The intracellular reaction rates increased to 900 and 850 mmol/L.hr for Rate$_2$ and Rate$_3$ respectively with the Kin$_3$ mechanism. It is interesting to note from the data that the rate of biomass formation (Rate$_6$) did not increase to the same extent with the Kin$_3$ mechanism. The rate recorded for Rate$_6$ increased 36% compared to the 50% and 54% increase in Rate$_2$ and Rate$_3$ respectively with the Kin$_3$ mechanism. The smaller increase in Rate$_6$ was due to the presence of the saturation terms of both intracellular substrates (C$_2$H$_4$ and O$_2$) in the reaction. There is only 1 substrate saturation term for Rate$_2$ and Rate$_3$. With the Kin$_1$ and Kin$_2$ mechanisms, the actual intracellular rates were almost identical. The intracellular metabolite C$_2$H$_4$ was formed by both Rate$_2$ and Rate$_3$, but only consumed by Rate$_6$ (Figure 2.1 in Section 1.1.1 of Chapter 2). With the Kin$_1$ and Kin$_2$ mechanisms, the overall effect of the saturation term on Rate$_6$ was reduced by the inhibition terms $(1 + C_{int}^{sat}/K_i)$ due to the feed back inhibition of Rate$_2$ and Rate$_3$. Therefore the rate of accumulation of the intermediate was reduced and hence the saturation of its consumption rate.

The same behaviour observed with the Kin$_1$ and Kin$_2$ mechanisms for the estimated reactions obtained with model B was observed with those estimated with the Kin$_3$ mechanism (Figures 3.4c and 3.4d). The maximum cellular dilution terms increased to 0.055 mmol/L.hr with the Kin$_3$ mechanism. The higher cellular growth rate obtained with the Kin$_3$ mechanism is attributed to the higher cellular rates due to the lack of feed-back inhibition of the cellular reactions. The impact of the higher $\mu C_{int}^{sat}$ term on model B was not very significant as a review of the model data indicated relative errors of below 5% were still obtained with the Kin$_3$ mechanism. The sensitivity of the cellular dilution terms to the operating and kinetic parameters has been discussed in greater detail in Chapter 4.

The data obtained for model C with the Kin$_3$ mechanism (Figures 3.4e and 3.4f) indicates that the
presence of the saturation term \((K_M + C_i)\) without an inhibition term (present in the Kin1 and Kin2 mechanisms) did not improve the overall accuracy of the reaction rates generated using model C. As was the case with the Kin1 and Kin2 mechanisms, the model accuracy improved as the conversion increased. The accumulation rate for the intracellular metabolite \(\text{C}_2\text{H}_4\) increased by 58% with the Kin3 mechanism. This was due to the “imbalance” described above with regard to the relative change in \(\text{Rate}_2\) and \(\text{Rate}_3\) compared to \(\text{Rate}_6\). The higher accumulation term resulted in a more inaccurate estimation of \(\text{Rate}_6\) with the Kin3 mechanism. However, this accumulation term decreased as the substrate conversion increased above 30%. Accurate rate estimations (relative errors below 10%) were obtained with the Kin1 and Kin2 mechanisms when the substrate conversion increased above 20%. These results indicate that the form of the kinetic mechanism affected the rate of intracellular metabolite accumulation and hence the validity of the PSSH. A more detailed analysis of these observations has been conducted in Chapter 5.
3.3. RESULTS AND DISCUSSION

(a) Model A (Extracellular Rates)  
(b) Model A (Intracellular Rates)  
(c) Model B (Extracellular Rates)  
(d) Model B (Intracellular Rates)  
(e) Model C (Extracellular Rates)  
(f) Model C (Intracellular Rates)  

Figure 3.4: Reaction rate profiles obtained with the Kin\textsubscript{3} mechanism under standard conditions for models A, B and C

The reaction data obtained with the Kin\textsubscript{4} mechanism under standard conditions has been shown in Figure 3.5. It can be seen clearly from the results that model A was accurate in predicting the reaction profiles generated with the first order kinetic mechanism using a linearised flux model (Figures 3.5a and 3.5b). The data once again confirms that the accuracy of the flux model was not affected by the linearisation of the reaction data. This observation is significant if one considers the variation in the shape of the reaction profiles generated with the different mechanisms. Therefore, the data clearly
indicates that if the flux model is populated with sufficient reaction data, linearised over short time intervals, there is no inherent inaccuracy in using a flux model to estimate the reaction profiles.

The reaction rate obtained with the Kin$_4$ mechanism for Rate$_5$ increased to 9200 mmol/L.hr. This was due to the lack of saturation terms in the reaction mechanisms which allowed the reactions to proceed without being slowed by the accumulation of intracellular metabolites. The deviation in the reaction rates obtained with the Kin$_4$ mechanism from a classic first order reaction rate (Levenspiel [1999]) was due the presence of the equilibrium constant ($K_{Eq}$) in the reaction mechanism for the extracellular rates.

As was the case with the Kin$_1$, Kin$_2$ and Kin$_3$ mechanisms, the accuracy of model B was at its lowest when the intracellular reactions reached their maximum values (Figures 3.5c and 3.5d). However, the relative errors with the Kin$_4$ mechanism decreased to below 1.5% for all rates with model B. The improved model accuracy was due to the $\mu C^\text{int}_i$ term which decreased to 0.011 mmol/L.hr. A lower $\mu C^\text{int}_i$ term was obtained with the Kin$_4$ mechanism despite the increase in $\mu$ to 0.085 hr$^{-1}$. Therefore the lower relative errors obtained for model B with the Kin$_4$ mechanism were due to the higher overall rates which lowered the accumulation of intermediates within the reaction volume.

Although there was an improvement in the accuracy for model B with the Kin$_4$ mechanism, there was no improvement in the accuracy of model C at a reaction time of $t \approx 0$ (Figures 3.5e and 3.5f). The results do however suggest that the accuracy of model C improved over a shorter reaction time with the Kin$_4$ mechanism. Accurate predictions of the reaction rates (relative errors < 10%) were obtained at conversions above 10% with the Kin$_4$ mechanism. This data indicates that the relative consumption rate of the intracellular metabolites was faster due to lack of saturation or inhibition terms present in the kinetic mechanism. Therefore the observed dynamics of the mass of intracellular metabolites remained low in comparison to that of the total mass of the system. This resulted in an improvement in the validity of the PSSH with the Kin$_4$ mechanism.
3.3. RESULTS AND DISCUSSION

From the data shown in this section, it is evident that the accuracy of model A was not affected by the form of the kinetic mechanism used to generate the kinetic data under standard conditions. These observations clearly indicate that the linearisation of the reaction data by the flux model, had no effect on the accuracy of the estimated reaction rates. Furthermore, the computational procedure used to solve the flux models did not have any inherent errors associated with it. Therefore, the variation in accuracy of the reaction profiles generated with models B and C, with the various kinetic mechanisms,
was as a result of inherent errors associated with each model and not the computation of the model. Furthermore, the observed variation in the accuracies for models B and C for each mechanism, was not a result of the linear flux model being unable to simulate the non-linear behaviour of the reaction profiles over discreet analysis points.

The deviation in the model accuracy, observed with the various kinetic mechanisms, indicates that the suitability of the reaction data for use in the flux model was not universal. What is also important to note is that these observations were made under standard conditions with all mechanisms. It is expected that variation in the operating conditions will also influence the accuracy of the flux models and the affects of each change will be different with each mechanism. It must be noted that the purpose of the research conducted in this study is to evaluate the accuracy of the linearised flux model and not the kinetic reaction models. The fully defined kinetic reaction models were only used to generate reaction data sets for the computation of the simplified linearised reaction fluxes, dependent only on the reaction stoichiometry.

The preliminary data collected from the investigations discussed in this section, clearly show that the assumptions made in the development of the MFA models are valid for the majority of the reaction data collected with the various mechanisms. However, the observed inaccuracies of model B, when the reactions are at a maximum, and model C, during the initial stages of the reaction time, cannot be ignored. To study these observations further, a detailed analysis has been conducted on the models in the following chapters. This has been done to determine how these models respond to simulated changes in the operating and kinetic parameters.

### 3.3.2 Reaction rate profiles for reaction pathway DF$_2$

The results shown in this section were obtained for the DF$_2$ reaction pathway with the Kin$_1$ mechanism under standard conditions. The only difference in the simulation data generated in this section was the selection of reaction pathway. All other parameters remained constant (Figure 3.1b). The results in this section have been included to show how the algebraic solution for the linear flux model can be affected by altering the set of fluxes used to constrain the reaction matrix. This data is useful as it can help guide the design of experiments to ensure that suitable reaction flux data is collected to constrain the reaction matrix correctly. The estimated reaction rates obtained for model C have been included in
3.3. **RESULTS AND DISCUSSION**

This section as this model relates directly to the approach used in MFA studies. The same linear solver, singular value decomposition, was also used to generate the estimated reaction rates in this section.

**Table 3.8: Summary of reaction data obtained at standard conditions for each reaction pathway with the Kin1 mechanism**

|        | \(\mu C_i^{\text{pm}}\) [mmol/L.hr] | \(dC_i^{\text{pm}}/dt\) [mmol/L.hr] | Flux [mmol/L.hr] | \(X_{C_4H_{10}}\) [-] *
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DF1</td>
<td>0.017</td>
<td>0.16</td>
<td>1.07</td>
<td>0.20</td>
</tr>
<tr>
<td>DF2</td>
<td>0.015</td>
<td>0.13</td>
<td>0.97</td>
<td>0.44</td>
</tr>
<tr>
<td>DF3</td>
<td>0.012</td>
<td>0.10</td>
<td>1.23</td>
<td>0.73</td>
</tr>
</tbody>
</table>

*\(X_{C_4H_{10}}\) refers to the conversion of extracellular \(C_4H_{10}\) where accurate estimations (relative errors < 10%) were obtained for model C.

The results obtained for model C with constraint sets 1 and 2 have been shown in Figure [3.6]. The data shows a significant deterioration in the accuracy of model C for constraint set 1 when the DF2 reaction pathway was simulated (Figures [3.6a] and [3.6b]). It is important to note that the results shown in these figures only differed from those shown in Figures [3.2e] and [3.2f] due to the change in reaction pathway. The negative rates predicted for Rate3 are a clear indication of the large inaccuracies obtained for pathway DF2 with constraint set 1. The estimated profiles obtained for model C exhibit the same phenomenon of rate scaling due to the PSSH. However, the inaccuracies in the estimated rates show some significant error was present in the scaling of the rates. This resulted in a significant propagation of the model error in the linear transformations.

Model C was more inaccurate with constraint set 2 (Figures [3.6c] and [3.6d]). The errors in the scaling of the constrained rates were significantly higher and resulted in unrealistic solutions for some of the extracellular and intracellular rates. The accuracy of model C became even worse with constraint set 3 (Figures [3.7a] and [3.7b]). The data shows that the solution was completely unrealistic over the entire simulated reaction time. There was an improvement in the estimated extracellular rates obtained with constraint set 4 (Figure [3.7c]). However, unrealistic reaction rates were estimated for the intracellular rate of \(C_2H_6\) consumption (Rate6) with constraint set 4 (Figure [3.7d]). The negative rate estimated for Rate6 was however 10 times lower than that estimated with constraint set 3, indicating some improvement in the model accuracy.
3.3. RESULTS AND DISCUSSION

Figure 3.6: Reaction rate profiles obtained with the Kin1 mechanism under standard conditions for model C using constraint sets 1 and 2
3.3. RESULTS AND DISCUSSION

There was an observed improvement in model C with constraint sets 5 to 7 in the estimation of the extracellular rates (Figures 3.8a, 3.8c and 3.9a). There was however an unrealistic solution generated for some of the intracellular rates at low reaction times (< 1 Hr) with constraint sets 5 to 7 (Figures 3.8b, 3.8d and 3.9b). As was the case with the results obtained for the reaction pathway DF1, there was an observed improvement in the estimation of the rates as the reaction time proceeded. However, accurate estimations of the reaction rates (relative errors < 10%) were only obtained when the substrate conversion exceeded 44%. This is more than double that observed with reaction pathway DF1 (Table 3.8).

The inaccurate estimation of the reaction rates with model C can only be due to a higher significance of the cellular dilution terms or metabolite accumulation terms in the cellular mass balances. A review of the reaction data generated for the DF2 reaction pathway revealed that the cellular dilution rate for pathway DF2 was in fact lower than that recorded with the DF1 pathway (Table 3.8). The maximum accumulation rate obtained for the DF2 pathway (0.13 mmol/L.hr) was also lower than that calculated...
for the DF\textsubscript{1} pathway (0.16 mmol/L.hr). The cellular dilution rates, accumulation rates and reaction fluxes obtained for the DF\textsubscript{2} pathway were of the same order of magnitude if not less than those obtained for the DF\textsubscript{1} pathway (Table 3.8), and yet model C was more inaccurate for the DF\textsubscript{2} pathway. Therefore the decrease in the accuracy of model C for pathway DF\textsubscript{2} was a result of computational errors. The computational errors resulted in a more pronounced propagation of the model error resulting in the very inaccurate estimated reaction rates. The cause of the higher computational error is discussed in detail below.

It is important to note that the reaction data shown in Table 3.8 was the same for all constraint sets as this data was generated from the kinetic reactions and cellular mass balances which were generic for each constraint set. Only the manner in which the reaction matrix was constrained to solve the linear equations changed with each constraint set.

![Figure 3.8: Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for model C using constraint sets 5 and 6](image)
3.3. RESULTS AND DISCUSSION

(a) Constraint Set 7 (Extracellular Rates)
(b) Constraint Set 7 (Intracellular Rates)
(c) Constraint Set 8 (Extracellular Rates)
(d) Constraint Set 8 (Intracellular Rates)

Figure 3.9: Reaction rate profiles obtained with the Kin1 mechanism under standard conditions for model C using constraint sets 7 and 8

The estimation of the reaction rates with constraint set 8 was more inaccurate than that with constraint sets 5, 6 and 7 (Figures 3.9c and 3.9d). Unrealistic reaction rates were obtained for 3 of the reaction rates and the rates predicted for Rate_6 and Rate_9 were 7 and 10 times larger than those estimated with constraint set 7 respectively. There was an improvement in the estimation of the rates with constraint sets 9 and 10 (Figures 3.10a, 3.10b, 3.10c and 3.10d). The results obtained with constraint sets 9 and 10 were similar to those obtained with constraint sets 5, 6 and 7. The similarity in the results obtained with these constraint sets is not related to the fluxes used to constrain the reaction matrices. Table 3.2 shows that no same flux was used in constraint sets 5, 6, 7, 9 and 10. This further validates the conclusion that the inaccuracy observed in each computation with the different constraint sets was as a result of computational error, and not the model constraints.
3.3. RESULTS AND DISCUSSION

Figure 3.10: Reaction rate profiles obtained with the Kin\textsubscript{1} mechanism under standard conditions for model C using constraint sets 9 and 10

The results shown above indicate graphically that the worst estimation of the reaction rates for pathway DF\textsubscript{2} with model C was obtained with constraint set 3 and the best estimation was obtained with constraint set 5. To investigate the cause of the computational error associated with model C for the DF\textsubscript{2} pathway, the condition numbers for the reaction matrices generated with each constraint set were estimated for pathways DF\textsubscript{1}, DF\textsubscript{2} and DF\textsubscript{3} (Figure 3.11). It can be seen from the data that the condition numbers estimated for reaction pathway DF\textsubscript{1} were dependent on the constraint set (Figure 3.11a). This data shows that the choice of flux used to constrain the reaction matrix affected the singularity of the reaction matrix. Heath (2001) states that as the condition number for a matrix increases, the linear system becomes ill-conditioned and the uncertainty in the solution increases due to the formation of a larger geometric solution space. The formation of this larger geometric space makes the accuracy of linear solution significantly more sensitive to perturbations or errors in the linear transformations. In the case of model C, these errors are the exclusion of the accumulation and cellular dilution rates from the cellular mass balances. Therefore, the errors inherent with the exclusion of these rates from the
cellular mass balances are amplified by the condition number resulting in higher model errors as the reaction matrix becomes linearly dependent.

The condition numbers obtained for the DF$_2$ pathway were significantly larger than those obtained for the DF$_1$ pathway (Figure 3.11b). The lowest condition number obtained for the DF$_2$ pathway was approximately 41 with constraint set 5 compared to the largest condition number of 19 for pathway DF$_1$. The larger condition numbers obtained for pathway DF$_2$ clearly correlate with the more inaccurate estimation of the rates with model C. Furthermore, as the condition number decreased for pathway DF$_2$ for a given constraint set, the accuracy of model C improved. The higher condition numbers obtained with the DF$_2$ pathway was due to the increase in the linear dependency of the reaction matrices. This was a result of an increase in the degree of branching in the reaction pathway (Figure 2.1 in Section 2.1.1 of Chapter 2). The increased branching resulted from the inclusion of the elemental reaction for the direct combustion of C$_2$H$_6$ (Elemental reaction 4, Section 2.1.1 of Chapter 2).

The low condition numbers obtained with constraint sets 5, 6, 7, 9 and 10 correlated to the most accurate estimation of the reaction rates for pathway DF$_2$. Therefore, as the condition number decreased, the propagation of the inherent model C errors was reduced in the linear transformations associated with a reduction in the geometric solution space. Although the model errors decreased with a decrease in the condition number, the model was only accurate at substrate conversions in excess of 44% for constraint set 5. Therefore the uncertainty in the solution was only reduced when the accumulation rates decreased sufficiently to decrease the bound for the solution space at higher substrate conversions.

It must be noted that the lowest condition number estimated for the DF$_1$ pathway was obtained with constraint set 3 and not constraint set 1 which was used in the estimation of the reaction rates in section 3.3.1. However, this did not affect the results presented in this section as the model errors were only compared with the different kinetic mechanisms. In Chapter 5, constraint set 3 was used in the sensitivity analysis performed for model C to reduce the computational error.

The results shown in this section clearly illustrate how the accuracy of the estimated rates can be negatively affected by computational errors as a result of using ill-condition reaction matrices. Although the results shown for the DF$_2$ pathway may be exaggerated to some extent, they do show
3.3. RESULTS AND DISCUSSION

how the accuracy of the model can be lost even when the PSSH is applied to a poorly designed reaction matrix. The data clearly shows that the design of a set of experiments must be done in such a manner as to minimise the condition number for the reaction matrix. This can be achieved by measuring the specific extracellular fluxes that ensure that the matrix does not become ill-conditioned when they are constrained to obtain a determined flux model. This is not always done in practise. In work conducted by Forster (2003) and Huang (2003), the condition numbers estimated for the reaction matrices used in the estimation of the reaction fluxes exceeded 100. This clearly shows that the results generated from the hypothetical pathways have significance to practical MFA applications.

![Condition numbers for all reaction matrices for reaction pathways DF1, DF2 and DF3](image)

Figure 3.11: Condition numbers for all reaction matrices for reaction pathways DF1, DF2 and DF3

It is noted that the use of the SVD routine for the computation of the linear equation set should result in the model being less sensitive to the presence of a singular reaction matrix. Heath (2001) states that the use of the SVD routine results in the removal of singular values from the diagonal matrix which allows for the estimation of a solution which is overdetermined. However, the reaction systems in pathway DF2 were not overdetermined and therefore the ill-conditioned matrices became rank-deficient when
3.3. RESULTS AND DISCUSSION

the singular values were removed. Unlike the severe error encountered with a linear routine, such as LU factorisation, in the computation of ill-conditioned systems, the rank-deficiency of the matrix makes the least squares solution sensitive to perturbations or errors in the input data [Heath, 2001]. In the context of this model, these errors are the inherent errors of model C. It must be noted that the reaction system was not setup as a least squares computation as this is not commonly done in the computation of reaction fluxes using the conventional MFA methodology outlined in literature.

The use of linear programming routines are sometimes used to solve underdetermined reaction systems (Stephanopoulos et al., 1998). Therefore different computation routines have been used to solve the reaction matrix obtained for reaction pathway DF$_3$ in section 3.3.3. The results shown in Figure 3.11c clearly indicate that the reaction matrices obtained for pathway DF$_3$ are poorly conditioned. Constraint set 5 was selected for this analysis as it generated the lowest condition number, albeit 10 times higher than the lowest condition number obtained for pathway DF$_2$. The near 10 fold increase in the condition number for pathway DF$_3$ is a clear indication of the affect of increased branching in the pathway. For pathway DF$_3$, the branching was increased through the inclusion of the direct combustion of C$_4$H$_{10}$ (Elemental reaction 3, Section 2.1.1 of Chapter 2). The significance of the observed increase in the condition numbers and computational errors is that this becomes relevant in actual systems where reaction branching is purposefully altered through metabolic engineering techniques. In fact, this is a significant research focus in many literature studies in an attempt to enhance product yields by altering the flux through branch points or nodes (Lee et al., 2008; Jamshidi and Palsson, 2008; Peng et al., 2004 and Stafford et al., 2002).

3.3.3 Reaction rate profiles for reaction pathway DF$_3$

The results obtained for the computation of model C for the DF$_3$ reaction pathway using linear algebra routines have been shown in Figure 3.12. The results show that the same qualitative model accuracy was obtained using the LU factorisation and SVD routines for both the extracellular and intracellular rates (Figures 3.12a, 3.12b, 3.12c and 3.12d). This result indicates that even though the SVD routine is more robust in handling ill-conditioned systems, the rank-deficiency resulted in the same level of inaccuracy of the final solution. It must be noted that computational warnings were generated by the MATLAB® LU factorisation routine when the model was solved using constraint set 7. Inconsistent
solutions were estimated under these conditions with the LU factorisation routine. Although the SVD routine was less sensitive to the conditioning of the reaction matrix, there was little or no improvement in the accuracy of the final solution obtained with this routine.

Figure 3.12: Reaction rate profiles obtained with the Kin1 mechanism under standard conditions for model C using linear algebra routines

There was an observed improvement in the accuracy of the rates estimated using the LP1 routine (Figures 3.13a and 3.13b). The observed improvement was both in the magnitude of the predicted rates as well as the their form from those obtained with the linear algebra routines. It must be noted however that even though there was an improvement in the accuracy of the solution, relative errors of below 10% were not obtained for the estimated reaction rates at any of the computation points in the reaction time. It can be seen from Figure 3.13b that the estimated rate obtained for Rate2 was an unrealistic solution. The most feasible solution obtained for the linear programming routine was when Rate2 was zero. This occurred when the branch node for the intracellular metabolite C4H10 was deactivated. In essence, the most feasible solution for the LP1 routine was obtained when reaction pathway DF2 was used to estimate the reaction rates for pathway DF3. The inclusion of Rate2 in
pathway DF₃ is the only additional parallel reaction that was added to pathway DF₂. These results clearly show how the inequality constraints tried to minimise the feasible solution space. To achieve this, Rate₂ was set to zero and the linear dependency of the reactions decreased.

The data in Table [3.8] indicates that the accuracy of the estimated reaction rates only improved once the substrate conversion exceeded 74%. The data shows that only once the substrate was almost completely exhausted were the accumulation rates low enough not too be perturbed by the large computational errors. Therefore meaningful reaction data can only be obtained at high conversions in these instances which will make accurate kinetic parameter estimation very difficult. High model sensitivities are required to elucidate the kinetic parameters from a set of reaction data [van Santen and Niemantsverdriet (1995)]. This becomes more unlikely as the substrates become exhausted and the observed reaction rates become insignificant.

The accuracy of the estimated rates obtained with the LP₂ routine were worse than those obtained with the LP₁ routine (Figures [3.13c] and [3.13d]). The results indicate that the use of equality constraints to constrain the reaction matrix and extracellular fluxes prevented the solver from obtaining a realistic solution as it was prevented from setting Rate₂ to 0. The ability of the inequality constraints in LP₁ to achieve this allowed for a more feasible solution space as it was not limited by the rigidity set by the equality constraint in the LP₂ solver. Edgar and Himmelblau (1988) state that the use of equality constraints in a linear programming routine constrain the solution vector to a multidimensional surface or a single point. The inequality constraints used in the LP₁ routine specify a feasible region where the solution vector can exist. Therefore the use of the LP₂ routine meant that the ability to set Rate₂ to 0 mmol/L.hr was not feasible, as this would have violated the points on the constrained surface.
3.3. RESULTS AND DISCUSSION

The accuracy of the estimated rates obtained with the LP\textsubscript{3} and LP\textsubscript{4} routines were better than those obtained with the LP\textsubscript{2} routine (Figures 3.14a, 3.14b, 3.14c and 3.14d). The improvement in the model accuracy however did not result in accurate estimations of the reaction rates with either of the routines. The use of equality constraints to constrain the reaction fluxes resulted in a significant decrease in the accuracy of the estimation of Rate\textsubscript{6} and Rate\textsubscript{9} with the LP\textsubscript{3} routine compared to the LP\textsubscript{1} routine. The use of inequality constraints to constrain the reaction fluxes in the LP\textsubscript{4} routine resulted in an observed improvement in the accuracy of the estimated rates. However, the overall accuracy of the estimated rates was still lower than that obtained with the LP\textsubscript{1} routine. This data shows that the use of equality constraints for either the extracellular reaction fluxes or the reaction matrix decreased the accuracy of the estimated rates. This is significant in the context of the DF\textsubscript{3} pathway as it clearly illustrates the underlying linear dependency which cannot be overcome by merely swapping out the numerical routines.
3.3. RESULTS AND DISCUSSION

The results obtained in the section clearly indicated that linear programming routines should be used with caution when computing reaction fluxes. The choice of constraints and constraint types can significantly affect the accuracy of the results obtained. Edgar and Himmelblau (1988) clearly state that the choice of ill considered constraints can negatively affect the character of the solution space which can prevent the solution vector from moving easily between local and global maxima or minima.

Figure 3.14: Reaction rate profiles obtained with the Kin$_1$ mechanism under standard conditions for model C using linear programming routines (LP$_3$ and LP$_4$)
3.4 Concluding Remarks

This chapter has achieved its primary objective of validating the reaction data under standard conditions to experimental literature data. There was good correlation between the simulated cellular mass balance data to that reported in literature. The validation of the simulated reaction data showed that the trends observed in the reaction profiles are applicable to those which can be expected from practical MFA studies.

The high level of accuracy obtained with model A in estimating the reaction rates with all kinetic mechanisms showed that the linearised reaction flux model was capable of predicting the reaction rates, regardless of the non-linearity in the reaction mechanisms used to describe the reaction pathway. The short time intervals used in the estimation of the batch reaction fluxes were sufficient to remove any errors associated with the linearisation of the reaction data. The accuracy of model A clearly indicating the ability of the linear model to predict the reaction fluxes, provided that it was adequately populated with the intrinsic reaction data.

Model errors of less than 5% were obtained for model B with all kinetic mechanisms. The magnitude of the errors obtained for model B decreased with the Kin4 mechanism due to a decrease in the relative effect of the cellular dilution term in the cellular mass balances. The low errors obtained for model B validate the assumption made about the insignificance of the cellular dilution terms in the cellular mass balances under standard conditions.

Model C provided a poor estimation of the reaction rates at low substrate conversions with all kinetic mechanisms. The substrate conversions where model C became accurate varied with each kinetic mechanism. The mechanism where model C became accurate at the lowest conversion was the Kin4 mechanism. The results confirmed that the PSSH was not valid over the entire reaction time and its validity was dependent on the form of the reaction data generated with each kinetic mechanism. A qualitative analysis of the accuracy of model C indicated that the model was not successful in accurately predicting the rates generated with the Kin3 mechanism.

The condition number of the reaction matrix was affected by the manner in which the matrix was constrained by the extracellular fluxes. The lowest condition number obtained for reaction pathway DF1 was when constraint set 3 was used to constrain the reaction matrix. An increase in the condition
number of the reaction matrix negatively affected the accuracy of model C. Inaccurate estimations of
the reaction rates were obtained even in conditions where the PSSH was shown to be valid, due to the
increased computational error associated with high condition numbers.

The design of experiments should be done in such a manner as to ensure that the measured extracellular
fluxes are suitable for constraining the reaction matrix so that high computational errors are not
generated. The use of the SVD routine over the LU factorisation routine did not result in any significant
improvement in the final solution accuracy for the ill-conditioned reaction matrices generated for the
DF$_3$ pathway.

The character of the solution space generated with the constraints affected the accuracy of solutions
obtained with the linear programming routine. Careful selection of the constraints must be made
when generating the optimisation routine and they should be factored into the design of experiments to
ensure the best chance of reaching a global maximum or minimum for the objective function. A linear
programming routine cannot guarantee that inherent computational errors of a given reaction network
can be overcome if linear algebra routines fail to produce a realistic solution. Swapping out a linear
algebra routine with a linear programming routine will not result in any significant improvement in the
model accuracies if the system is linearly dependent. The inherent model errors are propagated to such
an extent that accurate rate estimation will not be obtained for poorly conditioned systems.
Chapter 4

Sensitivity of cellular dilution rates in the cellular mass balances

4.1 Introduction

The results shown in Chapter 3 clearly indicated that low errors were obtained for model B with each kinetic mechanism under standard conditions. The results were validated with literature based experimental data and validity of the MFA model assumption regarding the cellular dilution rates. Stephanopoulos et al. (1998) indicate that metabolic reaction fluxes were of an order 10,000 times larger than that of the cellular dilution terms recorded in numerous studies. The results presented in Chapter 3 indicated that the maximum cellular dilution rate obtained for reaction pathway DF$_1$ was only 30 times smaller than the reaction fluxes for the intracellular intermediate C$_4$H$_{10}$. However, the reaction data also showed that the cellular dilution rates obtained for some of the other intermediate metabolites were up to 2000 times smaller than the reaction flux through the metabolite pool at the same reaction time. Furthermore, the cellular dilution rate for the intracellular intermediate C$_4$H$_{10}$ became 680 times smaller the reaction fluxes through the intermediate as the extracellular substrate conversion increased above 60%.

Although the results obtained in Chapter 3 do not contradict the assumption made in the development of the MFA model, the data did indicate that the magnitude of the cellular dilution rates can vary significantly, even within the same set of reaction data. The data provided by Stephanopoulos et al.
for the validation of this assumption was obtained from various continuous cultures where the substrate conversions would be very high and the residual intermediate concentrations would be low. There is no specific reference of the validity of this assumption to batch reactions where the conversions and intermediate concentrations are variable, but it is inferred that the validity of this assumption is universal. Stephanopoulos et al. (1998) state that the dilution term can be considered to be negligible if it is 50 times less than the reaction flux. However, in the same text the authors mention that the cellular dilution rate for the amino acid pools in \textit{P. chrysogenum} were found to be only 10 times smaller than the metabolic fluxes. It was mentioned by Stephanopoulos et al. (1998) that the cellular dilution terms should be included in the flux model in this instance to avoid small errors.

The objective of this chapter is to perform a detailed analysis of the results obtained for model B, with the aim of determining the validity of the assumption made that the cellular dilution terms are negligible for all reaction data sets generated for batch MFA models. The objective will be achieved by performing a sensitivity analysis on the reaction data generated using each kinetic model for the DF\textsubscript{1} reaction pathway.

The generation of kinetic reaction data and subsequent flux estimations obtained for each flux model have been described in detail in Sections 2.1 to 2.4 of Chapter 2. The methodology specific to the generation of the sensitivity data and relative cellular dilution rate trends has been discussed in detail below.
4.2 Methodology

The reaction rate data was generated using the same methodology outlined in Section 3.2.1 of Chapter 3. To analyse the validity of the assumption that the cellular dilution rate was negligible under all conditions, the cellular dilution rate data generated from the kinetic reaction data was compared directly with the actual reaction flux data computed with the fully populated model A. The use of the metabolic fluxes from model A ensured that the flux data was accurate.

The relative cellular dilution rates were determined for each intracellular metabolite under all simulated reaction conditions. The formulae shown in Table 4.1 show how the relative cellular dilution rates were computed for each metabolite in reaction pathway DF1. \( [F_i; F_j] \) describes the average value of fluxes \( i \) and \( j \) that are associated with the specified metabolite pool. In other words, the cellular dilution rate computed in the cellular mass balances was calculated relative to the metabolic fluxes which produce and consume the same metabolite.

Table 4.1: Computation of the relative cellular dilution rates for each intracellular metabolite in reaction pathway DF1

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Relative Cellular Dilution Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₄H₁₀</td>
<td>( \frac{\mu C_{\text{int}}^{\text{C}_4\text{H}_10}}{[F_1;F_2]} )</td>
</tr>
<tr>
<td>O₂</td>
<td>( \frac{\mu C_{\text{int}}^{\text{O}_2}}{[F_5;F_6]} )</td>
</tr>
<tr>
<td>C₂H₆</td>
<td>( \frac{\mu C_{\text{int}}^{\text{C}_2\text{H}_6}}{[F_2;F_3]} )</td>
</tr>
<tr>
<td>C₂H₄</td>
<td>( \frac{\mu C_{\text{int}}^{\text{C}_2\text{H}_4}}{[F_2;F_3;F_6]} )</td>
</tr>
<tr>
<td>H₂</td>
<td>( \frac{\mu C_{\text{int}}^{\text{H}_2}}{[F_5;F_6;F_8]} )</td>
</tr>
<tr>
<td>CO₂</td>
<td>( \frac{\mu C_{\text{int}}^{\text{CO}_2}}{[F_6;F_7]} )</td>
</tr>
<tr>
<td>H₂O</td>
<td>( \frac{\mu C_{\text{int}}^{\text{H}_2\text{O}}}{[F_6;F_8]} )</td>
</tr>
</tbody>
</table>

The results shown in this chapter were obtained directly from the relative cellular dilution rate data
without the relative errors for each simulation. The effect of the cellular dilution rates on the relative errors for model B have already been discussed in the results presented in Chapter 3. The maximum relative cellular dilution rate obtained from the mean of the rates for all metabolites at each sample point was plotted as a function of each operating and kinetic parameter used in the sensitivity analysis. This data provided an indication of how much the cellular dilution rate increased or decreased relative to the reaction fluxes through a range of operating spaces. The effect of numerical error on the model computation was mitigated by ignoring data which was below the machine precision ($x < \sqrt{\text{eps}}$). This data was generated in the reaction time after the reactions ran to completion.

A graphical representation of the methodology used to generate the results shown in this chapter has been included in Figure 5.1. The data labelled as A, which was obtained from the ODE routine, represents the cellular dilution rates generated using the cellular concentration data and cellular growth rates. The data labelled B represents the extracellular concentration data which was used to constrain the reaction flux model and the dilution and accumulation rates which were used to populate model A. The linear solver used was the SVD routine with constraint set 3 as this generated the lowest computational error according to the data shown in Section 3.3.2 of Chapter 3. The sensitivity analysis conducted in this chapter was a relative sensitivity analysis. This allowed for direct comparison of the sensitivity data generated.
4.3 Results and Discussion

The results shown in the chapter were generated using surface plots to illustrate the sensitivity of the cellular dilution rates to the operating and kinetic parameters. The surface plots for each parameter were generated as a function of the initial concentration of the extracellular metabolite $C_4H_{10}$. This was done because the variation in the initial substrate concentrations is the most commonly manipulated parameter in batch MFA studies and the easiest parameter to manipulate when performing a large number of experimental studies. The axes for the surface plots have been plotted on a logarithmic scale.

A colour-bar has been included in each surface plot as a legend to illustrate the magnitude of the cellular dilution rates plotted in the surface plots. The upper and lower range for the colour-bar has been fixed in each plot. The maximum value illustrated on the colour-bar (0.05) represents a cellular dilution rate that is 20 times lower than the average metabolic flux through the metabolite pool. It was observed in the results shown in Chapter 3 that relative errors obtained for model B reached 5% when the metabolic flux reached between 20 and 30 times the value of the cellular dilution rates. It must be noted that the colour included as a legend does not represent a linear scale for the relative cellular dilution rates. The setup of the colour-bar was done in such a manner as to emphasise the relative cellular dilution rate trends observed between 0.01 and 0.05. The relative sensitivity analysis for the relative cellular dilution rates has been computed as the relative change in the relative cellular dilution rate for a 10 fold increase in the specified parameter.

4.3.1 Kinetic mechanism $Kin_1$ (Michaelis-Menten with non-competitive inhibition)

The results obtained for the variation of the limiting substrate concentration with the $Kin_1$ mechanism have been shown in Figure 4.2a. The results were generated using various limiting concentrations of $O_2$ ($C_B(0)$) for each initial substrate concentration of $C_4H_{10}$ ($C_A(0)$). An increase in $C_B(0)$ illustrated an increase in the relative concentration of $O_2$ for the same concentration of $C_4H_{10}$.

The results show that an increase in the initial substrate concentrations resulted in an increase in the relative cellular dilution rates for each metabolite. However, the results show that the fluxes were
at least 50 times larger than the cellular dilution rates for each simulation. Therefore it can be said that over the range of simulated initial substrate concentrations with the $\text{Kin}_1$ mechanism, the cellular dilution rates remained insignificant in the cellular mass balances. Therefore the exclusion of the cellular dilution rates was a valid assumption in the development of the MFA model for this data set.

The increase in the relative cellular dilution rates with the increase in the initial substrate concentrations was expected. An increase in the initial substrate concentration resulted in an increase in the intermediate concentrations as well as in increase in the cellular growth rate. The increase in the relative cellular dilution rates show that the reaction fluxes did not increase at the same rate as the cellular dilution rates at the higher substrate concentrations. The linear trend observed for the relative cellular dilution rates as a function of $C_A(0)$ represents a logarithmic increase in the relative cellular dilution rates with the initial substrate concentrations. The results indicated a 4 fold increase in the relative cellular dilution rates were recorded with a 10 fold increase in the initial concentration of $\text{C}_4\text{H}_{10}$ at low relative O$_2$ concentrations [$C_B(0) = 2C_A(0)$]. This sensitivity decreased to 2.2 at the higher relative O$_2$ concentrations [$C_B(0) = 8C_A(0)$].

The relatively insignificant change in the relative cellular dilution rates as a function of the limiting substrate shows that the intermediate concentrations were not affected by the presence of an excess or limiting substrate in the system. The maximum sensitivity recorded for the relative cellular dilution rates was 1.3. This confirmed that the relative cellular dilution rate was insensitive to the relative concentration of the secondary substrate (O$_2$). Therefore the assumption that the cellular dilution rates remain negligible in MFA studies has been shown to be valid regardless of how the secondary metabolite concentration is controlled in the experiment.

The relative cellular dilution rates recorded as a function of the equilibrium constant ($K_{\text{Eq}}$) have been shown in Figure 4.2b. The equilibrium constant was used to describe the relative intracellular concentration of the intermediates as a function of the extracellular concentrations. See Section 2.2 of Chapter 2 for a detailed description of how the equilibrium constant was used in the development of the cellular and reactor mass balances. As the equilibrium constant increased, the relative cellular dilution rate decreased. This shows that a higher equilibrium constant resulted in lower concentrations for the intracellular metabolites with respect to the extracellular metabolites and extracellular metabolite uptake or extracellular fluxes. This resulted in an overall decrease in the the relative cellular dilution
rate. The maximum sensitivity for the relative cellular dilution rate to $K_{Eq}$ was recorded to be 2.2. Therefore the results show that the relative cellular dilution rate was more sensitive to $K_{Eq}$ than the limiting substrate concentration $[C_B(0)]$. The results indicate that the sensitivity of the relative cellular dilution rates to the initial substrate concentrations was very similar for all values of $K_{Eq}$. The maximum sensitivity obtained as a function of the initial substrate concentration for various equilibrium constants was 2.3.

It is noted that the equilibrium constant is an inherent property for a given reaction system that cannot be easily manipulated. However, the significance of this result is that it illustrates how the distribution of metabolites between the intracellular reaction volume and the extracellular reactor volume can influence the validity of the assumption made in the development of an MFA model. This may become significant when the metabolite transport mechanisms across the cellular membrane are altered through metabolic engineering techniques thereby affecting the observed equilibrium.

It can be seen from the results shown in Figure 4.2b that the same relationship between the initial substrate concentration and the relative cellular dilution rates was observed, regardless of the equilibrium constant. Therefore any attempt to manipulate the equilibrium constant will not remove the sensitivity of the relative cellular dilution rate to the initial substrate concentration. However, operation at conditions that favour higher equilibrium constants will result in the ability to run the reactions at higher initial substrate concentrations whilst maintaining the accuracy of model B.
4.3. RESULTS AND DISCUSSION

(a) Effect of the limiting substrate concentration \([C_B(0)]\)

(b) Effect of the equilibrium constant \([K_{Eq}]\)

Figure 4.2: Relative cellular dilution terms obtained as a function of the operating parameters with the Kin_1 mechanism

It can be seen from the results shown in Figure 4.3a that the relative cellular dilution rate was more sensitive to the inhibition constant \((K_i)\) than the equilibrium constant \((K_{Eq})\). A sensitivity of 2.45 was recorded at higher initial substrate concentrations. The increase in the relative cellular dilution rate as the inhibition constant increased confirmed that there was an increased accumulation of intermediates within the reaction volume as a result of the decrease in the feed-back inhibition \((k_i C_i / (1 + C_i / K_i))\). A decrease in the inhibition meant that the substrate uptake rates were not limited to the same extent as the intermediate concentrations increased. These observations are significant as any manipulation of the inhibition rate, through metabolic engineering of the feed-back mechanisms, can negatively affect the validity of the MFA based model. The sensitivity of the relative cellular dilution rate to the
initial substrate concentration increased as the inhibition constant increased. The effect of the substrate concentrations on the cellular dilution rates can be reduced with lower inhibition constants. This allows for a wider range of valid experimental data for MFA models at high initial substrate concentrations.

The relative cellular dilution rate was more sensitive to the Michaelis-Menten constant ($K_M$) than the inhibition constant ($K_i$) (Figure 4.3b). The sensitivity recorded for $K_M$ was 2.8 at high initial substrate concentrations. The higher sensitivity of the relative cellular dilution rate to $K_M$ is due to the saturation ($K_M (1 + C_i)$) of the consumption rates of the intermediates. Unlike with the increase in the levels of inhibition of the reaction rate, the increase in the saturation of the reaction rates with an increase in $K_M$ resulted in a decrease in the consumption rates of the intermediates. An increase in the level of inhibition in the rates resulted in a decrease in the production rates of the intermediates. This is due to the feed-back mechanism used to describe the inhibition term. Therefore the increase in $K_M$ resulted in a higher sensitivity of the relative cellular dilution rates. These results clearly indicate that any attempt to manipulate the observed $K_M$ constant of a rate in a specific branch point can lead to the decrease in the model accuracy for the same reaction pathway.

The results shown in Figure 4.3c clearly indicate that the sensitivity of the relative cellular dilution rate to the rate constant ($k$) was very low. A sensitivity of 0.98 was obtained for the relative cellular dilution rate as a function of $k$ for all initial substrate concentrations. The insensitivity of the relative cellular dilution term to $k$ is due to the linear increase that was obtained in the intermediate concentrations and metabolic fluxes. An increase in the rate constant for all rates resulted in a scaling of each rate, including the biomass formation rate. It must be noted that in practise, manipulation of the observed rate constant would most likely not be done to all rates in a given reaction pathway. Usually an observed rate constant for a specific rate at a branch point would be increased through metabolic engineering to re-distribute the fluxes to favour the production of more valued products. However, the results shown in Figure 4.3c show how the sensitivity of the relative cellular dilution rates can be decoupled from the reaction rate constant. The results show that this was not possible with the other kinetic and operating parameters and may be significant in the design of experiments for the purpose of estimating kinetic parameters.

The results shown in Figures 4.3 and 4.2 indicated that the relative cellular dilution rates remained below 0.05 for all simulated conditions with the Kin$_1$ mechanism. Therefore within the simulated...
operating spaces, the exclusion of the cellular dilution rates from the cellular mass balances can be considered to be a valid assumption as the reaction fluxes remained at least 20 times higher than the cellular dilution rates.
4.3. RESULTS AND DISCUSSION

Figure 4.3: Relative cellular dilution terms obtained as a function of the kinetic parameters with the Kin1 mechanism

The results shown in Figure 4.4 are a summary of the sensitivities obtained for the relative cellular rates...
as a function of each parameter. The results shown in Figure 4.4a illustrate the sensitivities as a function of $C_B(0)$, $K_{Eq}$, $K_i$, $K_M$ and $k$ for 3 different initial concentrations. For example, the three sensitivities illustrated for $C_B(0)$ in Figure 4.4a illustrate the sensitivity of the relative cellular dilution rate to $C_B(0)$ at initial substrate concentrations of 8, 80 and 800 mmol/L respectively for the substrate C$_4$H$_{10}$ [C$_A(0)$]. The results shown in Figure 4.4b represent the sensitivities of the relative cellular dilution rates as a function of $C_A(0)$ at various parameter extremes. The three sensitivities illustrated for $C_B(0)$ in Figure 4.4a illustrate the sensitivity of the relative cellular dilution rate to $C_A(0)$ at a relative O$_2$ concentration of $C_B(0) = 2C_A(0)$, $C_B(0) = 4C_A(0)$ and $C_B(0) = 8C_A(0)$ respectively. It is important to note that the sensitivities plotted for $K_{Eq}$ in Figure 4.4a represent the inverse of the sensitivities calculated. This was done to provide a direct comparison to the other parameter sensitivities. This has been done for the sensitivity results shown for all mechanisms and was necessary as the sensitivity results were fractions. The magnitudal effect of the sensitivity results for $K_{Eq}$ did not change with the inverted data.

The results in Figure 4.4a show clearly that the relative cellular dilution rates were most sensitive to the MM constant at high initial substrate concentrations. This confirmed the observations made from the results shown in Figure 4.3b. The sensitivity data also shows that the relative cellular dilution rates were more sensitive to $K_M$ at all initial substrate concentrations. The results confirm that there was no observed sensitivity of the cellular dilution rates to the reaction rate constant. The decrease in the sensitivity of the cellular dilution rates at high relative secondary substrate concentrations indicated that under conditions of excess secondary substrates, the overall concentration of intermediates decreased, making the assumption that the cellular dilution terms are insignificant more valid in the flux model. This observation is discussed in more detail in the results shown for model C in Chapter 5.

The results in Figure 4.4b show clearly that the cellular dilution rates were more sensitive to the initial substrate concentrations at low relative substrate concentrations. Therefore in the presence of a limiting secondary substrate, the intermediate concentrations increased at a faster rate with an increase in the initial substrate concentrations. This is due to the accumulation of unreacted intermediates in the reaction volume which accounted for higher relative cellular dilution rates. The presence of unreacted intermediates within the reaction volume was due to the lack of O$_2$ available to drive the consumption rate of the carbon based metabolites. The results showed that the relative cellular dilution rates were more sensitive to the initial substrate concentrations at higher values of $K_M$. This data further validates the results shown in Figure 4.3b and the observations regarding the effect of the MM constant on the
4.3. RESULTS AND DISCUSSION

Overall validity of the MFA model.

![Graph](image)

**Figure 4.4: Summary of the sensitivities of the relative cellular dilution rates with the Kin_2 mechanism**

4.3.2 Kinetic mechanism Kin_2 (Michaelis-Menten with competitive inhibition)

The results obtained with the Kin_2 mechanism indicate that the same observations were made regarding the sensitivity of the relative cellular dilution rates to the relative secondary substrate concentration (Figure 4.5a). The results showed that the relative cellular dilution rates were more sensitive to the relative secondary substrate concentration with the Kin_2 mechanism. The increase in the sensitivity indicated that the form of the inhibition term in the Kin_2 mechanism \( K_M \left( 1 + C_i/K_i \right) \) did not result in the same inhibition of the substrate uptake rates into the reaction volume. This resulted in higher levels...
of unreacted intermediates with the depletion of the substrates.

The results in Figures 4.5b and 4.7a show a small increase in the sensitivity of the cellular dilution rates to $K_{Eq}$ with the Kin\textsubscript{2} mechanism. The sensitivity increased from 2.2 with the Kin\textsubscript{1} mechanism to 2.3 with the Kin\textsubscript{2} mechanism at high initial substrate concentrations. The increase in the sensitivity shows that the Kin\textsubscript{2} mechanism was more sensitive to the intracellular concentrations. The inhibition term in the Kin\textsubscript{1} mechanism was more effective in limiting the accumulation of intracellular intermediates as the equilibrium constant decreased and the equivalent intracellular substrate concentrations increased. The higher sensitivity to the equilibrium constant with the Kin\textsubscript{2} mechanism resulted in the cellular dilution rates to increase to within 20 times that of the metabolic fluxes.

![Figure 4.5](image-url)
The results in Figures 4.6a and 4.7a further illustrate the relative affect of the inhibition terms of the two mechanisms. The inhibition term in the Kin\(_1\) mechanism resulted in a lower overall intermediate concentration compared to that in the Kin\(_2\) mechanism. This was due to the reduction in the observed reaction rate constant in the Kin\(_1\) mechanism \((k_iC_i/(1+C_i/K_i))\) compared to the reduction in the substrate affinity in the Kin\(_2\) mechanism \((K_M(1+C_i/K_i))\) with an increase in \(K_i\). This resulted in the higher sensitivity of the cellular dilution rates to the value of the inhibition constant with the Kin\(_2\) mechanism. The significance of the form of the inhibition term in the Kin\(_2\) mechanism is clearly evident from the higher sensitivities to the MM constant.

A higher sensitivity was also observed for the cellular dilution rates to \(K_M\) with the Kin\(_2\) mechanism (Figures 4.6b and 4.7a). The sensitivity increased from 2.8 with the Kin\(_1\) mechanism to 3 with the Kin\(_3\) mechanism. The results show once again that the form of the inhibition term in the kinetic mechanism played a significant role in determining the validity of the assumptions made in the MFA model. This becomes even more significant when the kinetic parameters and the operating parameters are varied either through metabolic engineering techniques or experimental design. The results in Figures 4.6c and 4.7a show that the relative cellular dilution rates were not sensitive to the reaction rate constant under any initial substrate concentrations.

The results shown in Figures 4.3 and 4.2 indicated that the relative cellular dilution rates remained below 0.05 for all simulated conditions with the Kin\(_2\) mechanism. The only exception was observed when the equilibrium constant was at its lowest value. The results indicate that in this parameter space, the exclusion of the cellular dilution rates from the cellular mass balances is not a valid assumption as this will result in model errors in excess of 5%.
4.3. RESULTS AND DISCUSSION

The results in Figure 4.7b show clearly that the cellular dilution rates were more sensitive to the
initial substrate concentrations at low relative substrate concentrations. There was an increase in the sensitivity obtained at low substrate concentrations from 3 with the Kin$_1$ mechanism to 4.1 with the Kin$_2$ mechanism. The higher sensitivity observed with the Kin$_2$ mechanism indicated that the design of any experiment used to generate data for MFA studies should be done while considering the affects of the form of the kinetic mechanisms. Applying an MFA analysis to an ad hoc set of reaction data generated from a range of operating and kinetic parameters can affect the validity of the assumptions made in the model development.

![Figure 4.7: Summary of the sensitivities of the relative cellular dilution rates with the Kin$_2$ mechanism](image)

(a) Sensitivity as a function of $C_B(0)$; $K_{eq}$; $K_i$; $K_M$ and $k$

(b) Sensitivity as a function of $C_A(0)$

Figure 4.7: Summary of the sensitivities of the relative cellular dilution rates with the Kin$_2$ mechanism
4.3.3 Kinetic mechanism Kin$_3$ (Michaelis-Menten with no inhibition)

The results shown in Figure 4.8 indicate that the sensitivities of the relative cellular dilution rates were higher with the Kin$_3$ mechanism. The relative cellular dilution rate obtained under all relative substrate concentrations reached 0.035 at an initial substrate concentration of 800 mmol/L (Figure 4.8a). The higher relative cellular dilution rates obtained with the Kin$_3$ mechanism illustrates the effect of reducing the feed-back inhibition on the accumulation of intracellular intermediates within the reaction volume. The lack of feed-back inhibition in the reaction mechanism meant that the substrate uptake into the reaction volume was not reduced as the intracellular metabolites accumulated. The results show that the reaction fluxes obtained with the Kin$_3$ mechanism at high initial substrate concentrations was only 28 times higher than the cellular dilution rates. However, the results shown in Chapter 3 indicate that this increase in the relative cellular dilution rates would not result in model B errors to exceed 5%.

The results in Figure 4.8b show that the relative cellular dilution terms increased to 0.05 when the equilibrium constant decreased to 0.012 mmol/L. The results indicate that as the equilibrium constant decreased, the accumulation of the intracellular metabolites increased resulting in higher relative cellular dilution rates. Under these conditions, relative errors of over 5% were obtained for flux simulations conducted within this parameter space. The sensitivity of the relative cellular dilution rate to the equilibrium constant increased to 2.4 with the Kin$_3$ mechanism at high initial substrate concentrations. This result further highlighted the effect of no inhibition in the kinetic mechanism on the overall accumulation of metabolites within the reaction volume.
The results shown in Figure 4.9a illustrate the increase in sensitivity observed for the relative cellular dilution rates as a function of $K_M$ with the Kin$_3$ mechanism. At high initial substrate concentrations and high values of $K_M$, the relative cellular dilution rate increased above 0.05. In this parameter space, model errors in excess of 5% were obtained and the assumption made regarding the insignificance of the cellular dilution rates became invalid. The sensitivity of the relative cellular dilution rates increased to above 2.6 at high initial substrate concentrations. The results clearly show that model B became inaccurate very quickly as the MM constant is increased at high initial substrate concentrations.

The results shown in Figure 4.9b indicate that the relative cellular dilution rates were not affected by the rate constant ($k$). This confirms the results obtained with the Kin$_1$ and Kin$_2$ mechanisms that the
variation in the kinetic rate constant resulted in a linear scaling of both the reaction fluxes and cellular growth rates.

![Diagram showing the effect of kinetic parameters on relative dilution terms](image)

The results shown in Figure 4.10 indicated that the highest sensitivity for the relative cellular dilution rate was obtained for \( K_M \) with the Kin3 mechanism. The same results were obtained with the Kin1 and Kin2 mechanisms. These results highlight the significance of the saturation term \((K_M + C_i)\) in the overall accumulation of intracellular metabolites. As the relative magnitude of the saturation term increases, the accumulation of metabolites in the intracellular volume increases. This results in an increase in the relative cellular dilution rates and a decrease in the accuracy of model B. The presence of inhibition terms in the Kin1 and Kin2 mechanisms managed to reduce the substrate uptake rate due...
to feed-back inhibition at high intracellular metabolite levels. This helped off set the observed effects of $K_M$ on the sensitivity of the relative cellular dilution rates with the Kin1 and Kin2 mechanisms.

The sensitivity of the relative cellular dilution rate to the initial substrate concentrations were similar at all values for the relative substrate concentrations (Figure 4.10b). The relative insignificance of the relative substrate concentrations with the Kin3 mechanism indicates that the lack of inhibition terms made a significant change in how the intracellular metabolite concentrations responded to the presence of excess substrates. This is a significant observation as it has implications on the design of experimental procedures used to generate reaction data for MFA models. The results show that the use of excess or limiting substrate concentrations will have a more significant impact on the model errors with a reaction pathway that is populated with high levels of feed-back inhibition. Similar observations were made with the sensitivities obtained as a function of the initial substrate concentration with for all values of $K_{Eq}$ and $K_M$ with the Kin3 mechanism.
4.3. RESULTS AND DISCUSSION

(a) Sensitivity as a function of $C_B(0)$; $K_{eq}$; $K_M$ and $k$

(b) Sensitivity as a function of $C_A(0)$

Figure 4.10: Summary of the sensitivities of the relative cellular dilution rates with the Kin_3 mechanism

4.3.4 Kinetic mechanism Kin_4 (First order)

The results obtained for the sensitivity analysis using the relative substrate concentration with the Kin_4 mechanism have been shown in Figure 4.11a. The results indicate the relative cellular dilution rates obtained with the Kin_4 mechanism were low for all initial substrate concentrations and relative substrate concentrations. The sensitivity of the relative cellular dilution rate as a function of the relative substrate concentration was also low with the Kin_4 mechanism. The results show that the computed reaction fluxes were at least 40 times larger than the cellular dilution rates.

The lower relative cellular dilution rates obtained with the Kin_4 mechanism were somewhat unexpected
due to the lack of inhibition terms in the mechanism. It was shown from the results obtained with the Kin$_3$ mechanism that the lack of inhibition terms present in the mechanism resulted in an overall increase in the relative cellular dilution rates. The cellular dilution rates obtained with the Kin$_4$ mechanism decreased to 0.011 mmol/L.hr from 0.054 mmol/L.hr with the Kin$_3$ mechanism under standard conditions (See Table 3.7 in Section 3.3.1 of Chapter 3). The reaction fluxes recorded through the metabolite pool under these conditions increased to 2.12 mmol/h.gDW with the Kin$_4$ mechanism from 1.87 mmol/h.gDW with the Kin$_3$ mechanism. The higher reaction fluxes coupled with the lower cellular dilution terms obtained with the Kin$_4$ mechanism is a clear indication of low levels of accumulated intermediates within the reaction volume. The results clearly show that even though there was no inhibition in the reaction rates with the Kin$_4$ mechanism, the lack of saturation (as with the Kin$_3$ mechanism) of the rates resulted in a faster overall consumption of intermediates within the reaction volume. This resulted in low levels of accumulation of intermediates and higher reaction fluxes which resulted in an improved accuracy for model B due to the lower relative cellular dilution rates.

Similar observations were made for the results obtained for the equilibrium constant and the reaction rate constant (Figures 4.11b and 4.11c). With the decrease in the equilibrium constant, the intracellular concentrations in the reaction volume increased resulting in higher levels of accumulation of the reaction intermediates. However, the ability of the Kin$_4$ mechanism to rapidly consume these intermediates resulted in lower overall relative cellular dilution rates. As a result, the sensitivity of the relative cellular dilution rates to K$_{Eq}$ decreased to 1.4 with the Kin$_4$ mechanism from 2.2 with the Kin$_3$ mechanism with K$_{Eq}$ at 0.012 mmol/L.
4.3. RESULTS AND DISCUSSION

Figure 4.11: Relative cellular dilution terms obtained as a function of the operating and kinetic parameters with the Kin4 mechanism

The sensitivity of the relative cellular dilution rates to $C_B(0)$ at low substrate concentrations increased
with the Kin₄ mechanism (Figure 4.12a). The results show that the cellular rates obtained at low substrate concentrations were not high enough to consume the excess secondary substrates on the accumulation of intermediates within the reaction volume. At higher initial substrate concentrations, the cellular rates were high enough to ensure that any excess substrates were distributed through a larger intracellular volume due to the higher growth rates obtained under these conditions. The cellular growth rate obtained with the Kin₄ mechanism increased from 0.0053 hr⁻¹ to 0.18 hr⁻¹ with the increase in the initial concentration of C₄H₁₀ from 8 to 800 mmol/L. It must be noted however that the higher sensitivity of the relative cellular dilution rate to C_B(0) at low initial substrate concentrations did not mean that the relative cellular dilution rates did not increase with an increase in the initial substrate concentrations (Figure 4.12b). Higher relative cellular dilution rates were obtained at higher initial substrate concentrations, but these rates were not sensitive to the presence of excess secondary metabolites.

The sensitivity of the relative cellular dilution rates to the initial substrate concentrations was lower than that obtained with the other reaction mechanisms (Figure 4.12b). The lower sensitivity was expected due to the higher reaction fluxes and lower accumulation of intermediates within the reaction volume. The results obtained with the Kin₄ mechanism were unexpected at first. The progression of the results shown in this chapter indicated an increasing trend in the overall effect of the cellular dilution rates in the cellular mass balances as the reaction rates along with the cellular growth rates increased. However, the results are a clear indication of how sensitive the MFA model can become to variations in the both the structure of the kinetic mechanisms as well as the parameters used to populate them.
4.3. RESULTS AND DISCUSSION

(a) Sensitivity as a function of \( C_B(0); K_{Eq} \) and \( k \)

(b) Sensitivity as a function of \( C_A(0) \)

Figure 4.12: Summary of the sensitivities of the relative cellular dilution rates with the Kin4 mechanism
4.4 Concluding Remarks

This chapter has achieved its primary objective of determining the sensitivity of the relative cellular dilution rates to various operating and kinetic parameters for all kinetic mechanisms. The results obtained in this chapter clearly illustrated the importance of considering the experimental design when attempting to generate experimental data for MFA based studies. The results showed that a simple action of increasing the initial substrate concentration can significantly affect the accuracy of a flux model. The magnitude of the cellular dilution rates increased to within 20 times of the reaction fluxes under conditions of high initial substrate concentrations and where the observed MM constant ($K_M$) was high.

The relative cellular dilution rates were less sensitive to the presence of a limiting substrate at higher initial substrate concentrations. The overall relative cellular dilution rate decreased as the secondary metabolite ($O_2$) concentration increased from a limiting to an excess substrate. The presence of inhibition terms in the kinetic mechanism ($K_{n1}$ and $K_{n2}$ mechanisms) resulted in a higher sensitivity of the relative cellular dilution rates to the presence of a limiting substrate.

Higher equilibrium constants resulted in lower relative cellular dilution rates. The sensitivity of the relative cellular dilution rates to the initial substrate concentrations was lower at higher equilibrium constants. A decrease in the inhibition terms in the $K_{n1}$ and $K_{n2}$ mechanisms resulted in an increase in the relative cellular dilution rates. The sensitivity of the relative cellular dilution rates to the initial substrate concentrations was lower when the inhibition constant decreased. The relative cellular dilution rates were more sensitive to $K_i$ with the $K_{n2}$ mechanism than with the $K_{n1}$ mechanism.

An increase in $K_M$ resulted in an increase in the relative cellular dilution rates with the $K_{n1}$, $K_{n2}$ and $K_{n3}$ mechanisms. The cellular dilution rates were more sensitive to $K_M$ than $K_i$ and $K_{Eq}$ with all 3 mechanisms. There was no observed sensitivity of the cellular dilution rates to the reaction rate constant $k$ with any of the kinetic mechanisms.

The $K_{n4}$ mechanism showed the lowest sensitivity of the relative cellular dilution rates to the equilibrium constant and the initial substrate concentrations. The lack of inhibition terms in the $K_{n4}$ mechanism did not result in a higher sensitivity of the relative cellular dilution rates as the consumption rates of the intermediates were high enough to ensure that the accumulation of intermediates remained
4.4. CONCLUDING REMARKS

low. Therefore the most accurate estimations of the reaction fluxes with model B were obtained with the Kin\textsubscript{4} mechanism. The relative cellular dilution rates showed the highest level of sensitivity to the reaction data generated with the Kin\textsubscript{3} mechanism.

Low relative cellular dilution rates were not generated by the observed reaction fluxes and cellular growth rates alone. The intrinsic properties of the kinetic mechanism which generated these higher rates, ultimately determined the accumulation of intermediates within the reaction volume. The results presented in this chapter were significant in highlighting the effects of the form of the kinetic mechanism and magnitude of the parameters on the cellular dilution rates.

It is noted that model B is not considered to be relevant in practical MFA studies as this model still requires the inclusion of accumulation terms of intracellular metabolites which is usually difficult and cumbersome to determine, especially for large experimental studies. These results become more significant when they are considered along with those obtained for model C which is the commonly used model in MFA studies. A review of the results obtained for model C under various conditions has been conducted in the following chapter.
Chapter 5

MFA model errors

5.1 Introduction

The results shown in Chapter 4 illustrated the sensitivity of the relative cellular dilution rates to the operating and kinetic parameters with all mechanisms. The results presented in Chapter 4 were significant as they illustrated how experimental design or manipulation of the kinetic mechanisms, through metabolic engineering techniques, can affect the validity of the assumptions made in the development of the MFA model. In practical applications of MFA, model B would not be used to compute the reaction fluxes as it still requires information about the intracellular metabolite concentrations to determine the accumulation rates \( \frac{dC_i}{dt} \). Therefore the results presented in this chapter discuss the relative errors obtained for model C with all kinetic mechanisms for the same conditions used to populate the results in Chapter 4. The results included in this chapter provide a better indication of how experimental design and metabolic engineering of kinetic pathways should be conducted when utilising a conventional MFA analysis to compute the reaction fluxes.

In Chapter 3, the results clearly indicated that the use of the pseudo-steady state hypothesis resulted in inaccurate predictions of the reaction rates at low conversions. The results also indicated that there was some sensitivity in the accuracy of model C to the form of the kinetic mechanism used to populate the reaction data. The results presented in Chapter 4 illustrated the relative cellular dilution rates under each condition. The results presented in this chapter illustrate the sensitivity of the relative errors that can be expected from a conventional MFA analysis. The choice of relative errors to present the results
in this chapter was made due to its relevance to practical reaction model analyses. Although the relative cellular dilution rate data presented in Chapter 4 provided a deeper insight into the formulation of the MFA models, the presentation of the accumulation rates would not be as significant. This is because any difference in the sensitivity of the errors for models B and C would be inherent to the application of the PSSH in model C and the interaction between the cellular dilution and accumulation rates.

The objective of this chapter is to determine the accuracy of the MFA model in predicting the reaction rates under various conditions with all kinetic mechanisms. The objective will be achieved by comparing the relative errors obtained under each condition for the DF₁ reaction pathway.

The generation of kinetic reaction data and subsequent flux estimations obtained for model C has been described in detail in Sections 2.2 to 2.4 of Chapter 2. The methodology specific to the generation of the errors for the MFA model has been discussed below.
5.2 Methodology

The reaction rate data was generated using the same methodology outlined in Section 3.2.1 of Chapter 3. The fluxes for reaction pathway DF\textsubscript{1} were calculated using model C. The error profiles were generated using the sum of the relative errors for all rates at each computation point. This provided the best representation of the accuracy of the MFA model under various conditions.

The results presented in Sections 3.3.2 and 3.3.3 of Chapter 3 indicated that the presence of numerical error in the flux computations can be minimised through the correct selection of linear solver and constraint set. To ensure that numerical error was minimised, the SVD numerical routine was used along with constraint set 3 to compute the reaction fluxes with model C.

The results shown in Chapter 3 clearly indicated that the reaction data obtained with the different kinetic mechanisms varied as a function of the reaction time and hence the comparison of the accuracy of model C was not very meaningful when the data was presented in this manner. To provide uniformity in the reaction rate data, the error analysis has been conducted as a function of the substrate conversion ($X_{C_4H_{10}}$).

$$X_{C_4H_{10}}(t) = \frac{C_{C_4H_{10}}^{\text{ext}}(t = 0) - C_{C_4H_{10}}^{\text{ext}}(t)}{C_{C_4H_{10}}^{\text{ext}}(t = 0)}$$

A graphical representation of the methodology used to generate the results shown in this chapter has been included in Figure [5.1]. The data labelled as A, obtained from the ODE routine, represents the concentration data used to compute the actual reaction rates. The data labelled B represents the extracellular concentration data which was used to constrain the reaction flux model.
5.2. METHODOLOGY

Figure 5.1: Methodology used for the evaluation of the MFA model errors under various conditions for the DF$_1$ reaction pathway
5.3 Results and Discussion

The results shown in the chapter were generated using 2D scatter plots to illustrate the effect of the operating and kinetic parameters on the sum of the relative errors. The scatter plots were used to depict the error profiles because they provided a better representation of the error space when there was discontinuity due to the variable spacing required to fit the uniform substrate conversion profiles. The scatter plots for each parameter were generated as a function of the initial concentration of the extracellular metabolite \( \text{C}_4\text{H}_{10} \). The relative errors are plotted on a logarithmic scale. The maximum error range has been capped at 100% and the minimum error range has been capped at 1%. A legend has been provided at the bottom of each figure to describe the conditions used in the generation of each error profile.

5.3.1 Kinetic mechanism \( \text{Kin}_1 \) (Michaelis-Menten with non-competitive inhibition)

The results obtained for model C errors as a function of \( C_B(0) \) at low and high initial substrate concentrations have been shown in Figure 5.2. The results obtained at low and high initial substrate concentrations have been presented to illustrate the error profiles as a function of both \( C_B(0) \) and the initial substrate concentrations. This allowed for comparison of the results with those shown in Chapter 4. It is important to note once again that the variation of the initial concentration of \( \text{C}_4\text{H}_{10} \) refers to an equivalent variation of the initial concentration of \( \text{O}_2 \). The extent to which the relative substrate concentrations were maintained was determined by \( C_B(0) \).

The results show that under conditions of limited secondary substrate concentrations \( (C_B(0) = 2x C_A(0)) \), the relative errors obtained for the MFA model remained above 30% at low initial substrate concentrations (Figure 5.2a). The high errors are a clear indication that the PSSH was not valid under these conditions. The results show that the accumulation of unreacted intermediates within the reaction volume was too high to allow for the accurate implementation of the PSSH. With an increase in secondary substrate \( (\text{O}_2) \) concentration to the stoichiometric requirement \( (C_B(0) = 4x C_A(0)) \), the relative errors decreased to below 10%. The high relative errors obtained at a conversion of less than 15% were representative of the form of the estimated reaction profiles obtained...
for model C in Chapter 3. The results confirm that the PSSH is a poor assumption at low conversions
due to the rapid initial increase in intermediates within the reaction volume during the early stages of
the reaction time.

The presence of an excess concentration of the secondary substrate (O\textsubscript{2}) resulted in a decrease in
the overall accuracy of the MFA model \((C_B(0) = 8xC_A(0))\). The results do however indicate that the
overall error was lower than that obtained when O\textsubscript{2} was a limiting substrate \((C_B(0) = 2xC_A(0))\). This
is due to the fact that under conditions where O\textsubscript{2} was a limiting substrate, there was an accumulation
of unreacted carbon based metabolites (C\textsubscript{4}H\textsubscript{10}, C\textsubscript{2}H\textsubscript{6} and C\textsubscript{2}H\textsubscript{4}) within the reaction volume. Due to
the lack of sufficient O\textsubscript{2}, C\textsubscript{2}H\textsubscript{4} will remain unreacted. Due the feed-back inhibition present in the
Kin\textsubscript{1} mechanism, this will then slow the consumption rate of C\textsubscript{2}H\textsubscript{6} and then C\textsubscript{4}H\textsubscript{10} resulting in a
higher accumulation rate for all carbon based metabolites. When C\textsubscript{4}H\textsubscript{10} was the limiting substrate,
only excess amounts of O\textsubscript{2} would accumulate within the reaction volume resulting in a lower overall
model error.

The variation in the overall relative error shown in Figure 5.2a as a function of C\textsubscript{B}(0) was not expected
after the results in Chapter 4 were analysed. The results showed that the relative cellular dilution
rates were not very sensitive to C\textsubscript{B}(0). However, the results obtained for model C clearly illustrate the
sensitivity of the accumulation rates to C\textsubscript{B}(0).

The increase in the initial substrate concentrations resulted in a decrease in the overall relative
errors obtained for the MFA model (Figure 5.2b). The relative error obtained when O\textsubscript{2} was a
limiting substrate \((C_B(0) = 2xC_A(0))\) tracked the relative errors obtained using high relative substrate
concentrations \((C_B(0) = 8xC_A(0))\) until the conversion reached 5%. As the conversion increased, the
relative error increased rapidly indicating that the consumption rate of carbon based metabolites was
being “starved” at very low conversions. The “starvation” of the reaction was due to the rapid depletion
of O\textsubscript{2}. The results obtained at low initial substrate concentrations showed that the same rates only
became “starved” at a conversion in excess of 15%. It must be noted that the conversion of C\textsubscript{4}H\textsubscript{10} did
not reach 100% under conditions where O\textsubscript{2} was limiting with both initial substrate concentrations. The
conversion reached approximately 35% under these conditions. It can be seen that the relative errors
levelled off as the conversion increased beyond this. The data plotted at higher conversions represents
the continued uptake of extracellular C\textsubscript{4}H\textsubscript{10} into the reaction volume even when the intracellular levels
of O\textsubscript{2} were depleted. In practise this may be unlikely as the microorganism will most likely limit the uptake of substrates into the intracellular volume to prevent a build of possible toxins (Stephanopoulos et al., [1998]). However, this regulatory architecture has not be included in the model system to maintain simplicity of the MFA based model to represent those commonly reference in literature. The exclusion of this architecture will not changed the observations made regarding the overall accumulation of intermediates under the different conditions.

The decrease in the overall error obtained at high initial substrate concentrations was unexpected, based on the results shown in Chapter 4. It was clear from those results that the relative cellular dilution rate increased as the substrate concentration increased. The results show in Figure 5.2 therefore indicate that the accumulation rate increased at high initial substrate concentrations and the combined affect of the two rates \((dC_i/dt + \mu C_i)\) was lower because the two rates cancelled each other out. This observation is concerning as it indicates that the observations made for the MFA model errors are merely a coincidence at a specified set of conditions. A detailed analysis of this observation and its implications has been conducted later on in this section.

The results shown in Figure 5.2b indicate that the presence of an excess concentration of O\textsubscript{2} had a minimal affect on the overall model error. The relative error increased to 13% from 4% at conversions in excess of 90%. The increase in the model error with \((C_B(0) = 8xC_A(0))\) at high conversions was due to the presence of small amounts of unreacted O\textsubscript{2} present in the reaction volume. The model errors were lower under these conditions than when O\textsubscript{2} was a limiting substrate as the amount of unreacted mass in the reaction volume was lower and therefore the accumulation rate was lower. This allowed for a more accurate implementation of the PSSH.

The increase in the relative errors obtained at higher conversions (> 80%) with both initial substrate concentrations is an indication of an increase in numerical error due to the low values of the fluxes. A more detailed analysis of the relative accumulation rate profiles has been provided later on in this section.
5.3. RESULTS AND DISCUSSION

(a) Low initial substrate concentrations \([C_{C_{4}H_{10}(0)} = 8 \text{mmol/L}]\)

(b) High initial substrate concentrations \([C_{C_{4}H_{10}(0)} = 800 \text{mmol/L}]\)

Figure 5.2: Relative errors for model C with the Kin_1 mechanism as a function of \(C_B(0)\)

As discussed previously, the decrease in the MFA model errors with an increase in the initial substrate concentrations was contradictory to the observations made in Chapter 4 with regard to the sensitivity of the cellular dilution rates. To investigate this further, the relative combined accumulation and cellular dilution rates \((dC/dt + \mu C_i)\) have been plotted under both initial substrate concentrations with \((C_B(0) = 4xC_A(0))\) (Figure 5.3). The rates have been calculated relative to the reaction fluxes through the metabolite pool using the same approach used in Chapter 4 (See Table 4.1 in Section 4.2). It can be seen from the profiles that the relative rates for all metabolites were lower under higher initial substrate concentrations.
concentrations. The rapid decrease observed at high initial substrate concentrations occurred before the substrate conversion reached 5% and hence the reason for the low relative errors obtained under these conditions (Figure 5.2b).

The decrease in the relative rates with an increase in the initial substrate concentrations was an intuitive result based on the error profiles shown in Figure 5.2. An improvement in the accuracy of model C implies that the relative rates must have decreased making the assumptions used in the development of the model valid. However, the decrease in the relative rates observed in Figure 5.3 can be due to higher reaction fluxes or a decrease in the absolute value of the rates or both. The distinction between them is significant, as it would imply that by purely increasing the reaction fluxes, lower model errors would always be achieved. This would be regardless of how the intrinsic reaction mass balance data responds to any changes. This would make the analysis conducted in this chapter irrelevant because the sensitivity of the model errors would not be dependent on the fundamental formulation of the PSSH under various conditions.

The absolute values of the rates \( \left( \frac{dC_i}{dt} + \mu C_i \right) \) obtained under the conditions highlighted above have been shown in Figure 5.4. If the decrease in the relative rates observed in Figure 5.3 with an increase in the initial substrate concentrations was purely due to higher reaction fluxes, it would be expected that the absolute value of the rates would be higher under these conditions. However, the results show the absolute value of the rates obtained at higher initial substrate concentrations was of the same order of magnitude if not lower than those obtained at lower initial substrate concentrations. For the rates to be lower at the higher initial substrate concentrations, the accumulation rate \( \left( \frac{dC_i}{dt} \right) \) must be higher to counter the higher cellular dilution rate observed under these conditions (Figure 4.2a of Chapter 4). This would then contradict the PSSH which states that the accumulation rate is negligible with respect to the extracellular accumulation rates. But, because the errors for model C decreased, the PSSH must be valid at higher initial substrate concentrations. For this to be possible, \( \left( \frac{dC_i}{dt} + \mu C_i \right) \) must represent the overall accumulation rate of intracellular metabolites within the reaction volume and when this rate approaches zero, the PSSH becomes valid. The cellular dilution rate therefore represents the accumulation rate of the mass of metabolite \( i \) with respect to cellular growth. The proof of this is the accumulation rate profiles shown in Figure 5.4. It can be seen from the data that the form of these profiles matches those of the relative errors obtained under the different conditions. Therefore the fluxes in the relative accumulation rates merely scale the relative rates under the different conditions.
and the absolute rates ultimately determine the suitability of the reaction data for use in the MFA model.

![Graph showing relative accumulation rate profiles](image)

**Figure 5.3: Relative accumulation rate profiles with the Kin₁ mechanism**

<table>
<thead>
<tr>
<th>$C_{C_4H_{10}}(0) = 8 \text{ mmol/L}$</th>
<th>$C_{C_4H_{10}}(0) = 800 \text{ mmol/L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_4H_{10}$ -- $H_2$ -- $C_4H_{10}$</td>
<td>$C_4H_{10}$ -- $H_2$ -- $C_4H_{10}$</td>
</tr>
<tr>
<td>$O_2$ -- $CO_2$ -- $O_2$ -- $\Delta$</td>
<td>$CO_2$ -- $O_2$ -- $\Delta$</td>
</tr>
<tr>
<td>$C_2H_6$ -- $H_2O$ -- $C_2H_6$</td>
<td>$C_2H_6$ -- $H_2O$ -- $C_2H_6$</td>
</tr>
<tr>
<td>$C_2H_4$ -- $C_2H_4$</td>
<td>$C_2H_4$ -- $O_-$</td>
</tr>
</tbody>
</table>
5.3. RESULTS AND DISCUSSION

(a) Profiles for the intracellular metabolites \( \text{C}_4\text{H}_{10}, \text{O}_2, \text{C}_2\text{H}_6 \) and \( \text{C}_2\text{H}_4 \)

(b) Profiles for the intracellular metabolites \( \text{H}_2, \text{CO}_2, \) and \( \text{H}_2\text{O} \)

Figure 5.4: Absolute accumulation rate profiles with the \( \text{Kin}_1 \) mechanism

<table>
<thead>
<tr>
<th></th>
<th>( \text{C}<em>4\text{H}</em>{10}(0) = 8\text{mmol/L} )</th>
<th>( \text{C}<em>4\text{H}</em>{10}(0) = 800\text{mmol/L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{C}<em>4\text{H}</em>{10} )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{H}_2 )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{O}_2 )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{CO}_2 )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{H}_2\text{O} )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{C}_2\text{H}_6 )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{C}_2\text{H}_4 )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( \text{O} )</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The relative error profiles obtained as a function of \( K_{\text{Eq}} \) and \( K_i \) have been shown in Figure 5.5. It can be seen from the results that the an increase in the equilibrium constant resulted in a decrease in the overall relative errors obtained with the \( \text{Kin}_1 \) mechanism (Figures 5.5a and 5.5b). An increase in the equilibrium constant resulted in a decrease in the intermediate concentrations within the reaction volume. This resulted in a reduction in the accumulation rates of the intermediates and hence an
improvement in the model C error. The decrease in the errors for model C with an increase in $K_{\text{Eq}}$ corresponded with an observed decrease in the relative cellular dilution rates (Figure 4.2b in Chapter 4). This resulted in a lower overall accumulation rate of intracellular metabolites. The results in Figure 5.5a indicate that with $K_{\text{Eq}}$ set at 1.2 mmol/L, the relative errors increased above 10%. The higher relative errors clearly indicate that the PSSH was not valid under these conditions.

The error profiles obtained at low initial substrate concentrations all indicate that there was an increase in the errors at high substrate conversions (Figure 5.5a). The increase in the errors was due to an increase in numerical error that was propagated through the reaction pathway due to the exclusion of the accumulation rates from the cellular mass balances. The presence of numerical error is clearly evident from the accumulation rate profiles shown in Figures 5.3 and 5.4. It can be seen from the absolute rate profiles at low substrate concentrations that there was no significant increase in the absolute value of the accumulation rates. This means that the validity of the PSSH did not change as the substrate conversion increased above 60%. In fact, the absolute value of the accumulation rates for most metabolites actually decreased. However, evaluation of the relative accumulation rate profiles (Figure 5.3) clearly shows an increase in the relative rates as the conversion increased above 60%. This was due to the rapid decrease in the reaction fluxes which resulted in an increase in the numerical errors. These errors were propagated by the error which is inherent to model C. It can be seen at high initial substrate concentrations, the relative accumulation rates for all profiles remained low (< 0.01) as the conversion increased above 60%. This resulted in the relatively insignificant increase in the relative errors at high conversions with high initial substrate concentrations.

The relative errors decreased with an increase in the initial substrate concentration with all values of $K_{\text{Eq}}$ (Figure 5.5b). The lower errors are a clear indication of an improvement in the validity of the model. The increase in the initial substrate concentrations resulted in an improvement in the model errors due to an overall decrease in the accumulation rate of intermediates within the reaction volume. The lower errors were due to the establishment of a pseudo-steady state as the intracellular consumption rate of the intermediates increased rapidly at low conversions. Therefore the relaxation time of the intracellular reactions decreased with respect to the extracellular rates, resulting in the observed PSS and improved model errors. This is clearly evident from the accumulation rate profiles shown in Figure 5.4. It can be seen at the higher initial substrate concentrations, the accumulation rates decreased at a faster rate than those obtained at low initial substrate concentrations. Furthermore,
the higher reaction fluxes obtained under these conditions made the overall effect of the PSS at higher initial substrate concentrations more significant (Figure 5.3). The results obtained with $K_{\text{Eq}}$ for model C were similar to those obtained for model B (Chapter 4). The results clearly showed that the decrease in the relative cellular dilution rate as $K_{\text{Eq}}$ increased, indicated a decrease in the accumulation rate.

The relative errors for model C increased with an increase in the value of $K_i$ at low and high initial substrate concentrations (Figures 5.5c and 5.5d). The increase in the model errors was more significant at the lower initial substrate concentrations. This confirmed the observations made in Figures 5.2b and 5.5b that the model error decreased as the initial substrate concentration increased under all conditions. The observed increase in the model errors with an increase in $K_i$ was as a result of a decrease in the effect of the inhibition term $(k_iC_i/(1 + C_i/K_i))$ on the reduction of the intermediate concentrations and hence the accumulation rate for intracellular metabolites. As $K_i$ decreased, the feedback inhibition increased and the rate of metabolite uptake into the reaction volume decreased, thereby reducing the accumulation rate of the intermediates. The results clearly show that the manipulation of the inhibition constant in the kinetic mechanism can have a negative impact on the MFA model errors.

The observed decrease in the model error with a decrease in $K_i$ is particularly significant in the context of the discussion provided earlier in relation to the results shown in Figures 5.3 and 5.4. It was mentioned that the decrease in model error may have only been due to an increase in the reaction fluxes which reduced the effect of the overall accumulation rate on the cellular mass balances. The results shown in Figures 5.5c and 5.5d validate the conclusion that this was not the case. The increase in the inhibition term resulted in a decrease in the reaction fluxes but lower model errors were obtained under these conditions. The reaction fluxes decreased from 3.65 to 0.75 mmol/g.DW.h with an increase in $K_i$. This confirmed that the MFA model was only valid when the overall accumulation rates were minimised through the intrinsic properties of the kinetic mechanisms, and not by forcing the data to suit the MFA model by merely maximising the reaction fluxes.
5.3. RESULTS AND DISCUSSION

(a) Error profile as a function of $K_{Eq} [C_{C_4H_{10}} (0) = 8\text{mmol/L}]$

(b) Error profile as a function of $K_{Eq} [C_{C_4H_{10}} (0) = 800\text{mmol/L}]$

(c) Error profile as a function of $K_i [C_{C_4H_{10}} (0) = 8\text{mmol/L}]$

(d) Error profile as a function of $K_i [C_{C_4H_{10}} (0) = 800\text{mmol/L}]$

Figure 5.5: Relative errors for model C with the Kin_1 mechanism as a function of $K_{Eq}$ and $K_i$

<table>
<thead>
<tr>
<th>$K_{Eq}$ (mmol/L)</th>
<th>$K_i$ (mmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.012</td>
<td>0.05</td>
</tr>
<tr>
<td>0.12</td>
<td>0.5</td>
</tr>
<tr>
<td>1.2</td>
<td>5</td>
</tr>
</tbody>
</table>

The relative model errors increased with an increase in the value of $K_M$ (Figures 5.6a and 5.6b). The increase in the model errors with $K_M$ correlated with an increase in the relative cellular dilution rates obtained with the Kin_1 mechanism (Figure 4.3b in Chapter 4). The increase in the accumulation rates at higher values of $K_M$ is a result of an increase in the saturation term $(K_M + C_I)$ in the kinetic mechanism. An increase in the saturation term meant that the consumption rate of the intermediates decreased and therefore the accumulation of intermediates increased. Unlike with the increase in the inhibition constant, the increase in the MM constant resulted in an increase in the intracellular metabolite concentrations along with a decrease in the overall reaction fluxes. This resulted in higher relative accumulation rates and hence higher model errors due to the invalid application of the PSSH under these conditions.
5.3. RESULTS AND DISCUSSION

The model errors were not affected by the value of the reaction rate constant (k) in the reaction mechanism (Figures 5.6c and 5.6d). These results correlate with those shown in Chapter 4 for the relative cellular dilution rates under the same conditions. The results clearly show a linear scaling of the reaction rate data with the variation in the reaction rate constant. The results once again show an observed decrease in the model errors with an increase in the initial substrate concentrations. These results confirm the observations made in this section regarding the affect of the initial substrate concentrations on the validity of the MFA model assumptions.

![Graphs showing relative errors for model C with the Kin1 mechanism as a function of KM and k](image)

**Figure 5.6: Relative errors for model C with the Kin1 mechanism as a function of KM and k**

- KM = 1.5 mmol/L
- KM = 15 mmol/L
- KM = 150 mmol/L
- k = 0.5 s\(^{-1}\)
- k = 5 s\(^{-1}\)
- k = 50 s\(^{-1}\)

---

February 2014
5.3. RESULTS AND DISCUSSION

5.3.2 Kinetic mechanism Kin\(_2\) (Michaelis-Menten with competitive inhibition)

The results obtained for the model C errors as a function of the relative initial substrate concentrations (\(C_B(0)\)) and the equilibrium constant (\(K_{Eq}\)) with the Kin\(_2\) mechanism have been shown in Figure 5.7. The data shows that the presence of O\(_2\) as a limiting substrate resulted in reduced model accuracy both at low and high initial substrate concentrations (Figures 5.7a and 5.7b). These results confirm the observations made with the Kin\(_1\) mechanism regarding the higher accumulation rate of unreacted carbon based metabolites within the reaction volume under these conditions. It is evident from the data that although the model errors decreased as the relative substrate concentration increased, the lowest errors were obtained when the O\(_2\) was supplied to meet stoichiometric requirement.

The results in Figures 5.7c and 5.7d clearly indicate that an increase in the equilibrium constant resulted in an improvement in the model errors. The improvement in the model errors was significant at both low and high initial substrate concentrations. The results obtained with the Kin\(_2\) mechanism indicated that the value of the equilibrium constant had a larger effect on the model errors than those obtained with the Kin\(_1\) mechanism (Figures 5.5a and 5.5b). With the Kin\(_1\) mechanism, model errors of below 5% were obtained with \(K_{Eq}\) set to 0.012 mmol/L at high initial substrate concentrations. Under the same conditions with the Kin\(_2\) mechanism, the errors remained above 5% for the majority of the data points. These results correlated with the relative cellular dilution rate data shown for the Kin\(_2\) mechanism as a function of \(K_{Eq}\) (Figure 4.5b in Chapter 4), where an increase in the sensitivity of the relative cellular dilution rates to \(K_{Eq}\) was observed with the Kin\(_2\) mechanism. The results shown in Figures 5.7c and 5.7d confirmed that the form of inhibition term \((K_M(1+C_i/K_i))\) present in the Kin\(_2\) mechanism resulted in higher overall accumulation rates for the intermediates within the reaction volume compared with the Kin\(_1\) mechanism under the same conditions. The higher accumulation rates obtained with the Kin\(_2\) mechanism resulted in the increase in the MFA model error due to decrease in the validity of the PSSH to the reaction data.

The results shown in Figure 5.7 confirmed that the model errors decreased as the initial substrate concentrations increased. The decrease in the model errors clearly indicated that the MFA model assumptions were valid under high initial substrate concentrations with both the Kin\(_1\) and Kin\(_2\) mechanisms.
The results obtained for model C with the Kin$_2$ mechanism as a function of the inhibition constant ($K_i$) and the MM constant ($K_M$) have been shown in Figure 5.8. The results show that the model errors decreased with a decrease in $K_i$ with the Kin$_2$ mechanism under both low and high initial substrate concentrations (Figures 5.8a and 5.8b). The results confirmed that a decrease in $K_i$ resulted in a decrease in the overall accumulation rate of intermediates within the reaction volume. This was confirmed by the cellular dilution rate data obtained with the Kin$_2$ mechanism as a function of $K_i$ (Figure 4.6a of Chapter 4). The observed model error was more sensitive to $K_i$ with the Kin$_2$ mechanism than with the Kin$_1$ mechanism. Similar observations were made between the results obtained for the relative cellular dilution rates with the Kin$_1$ and Kin$_2$ mechanisms as a function of $K_i$ (Figures 4.4 and 4.7 in Chapter 4). The results indicated that the sensitivity of the inhibition term to $K_i$ in the two mechanisms affected the validity of the PSSH for a set of reaction conditions.
The results obtained with the $\text{Kin}_2$ mechanism indicated that the model errors increased with an increase in $K_M$ (Figures 5.8c and 5.8d). The increase in $K_M$ resulted in higher overall accumulation rates of the intracellular intermediates due to the reduction in the consumption rate of the metabolites without an equivalent reduction in the production rate. This was due to the increased affect of the saturation term $(K_M + C_i)$ in the mechanism. The model errors obtained with the $\text{Kin}_2$ mechanism were more sensitive to the value of $K_M$ under low and high substrate concentrations compared to those obtained with the $\text{Kin}_1$ mechanism. The model errors obtained with the $\text{Kin}_2$ mechanism only decreased below 5% when the conversion increased above 20%. The same errors were obtained at a conversion of 10% with the $\text{Kin}_1$ mechanism. The higher sensitivity of the model errors to the value of $K_M$ with the $\text{Kin}_2$ mechanism was confirmed by the relative cellular dilution rate data shown in Figure 4.6b in Chapter 4. The results clearly showed a higher sensitivity to $K_M$ with the $\text{Kin}_2$ mechanism. This was due to the form of the $\text{Kin}_2$ mechanism where the effect of MM constant on the kinetic mechanism is amplified by the inhibition constant $(K_M(1 + C_i/K_i))$. 


5.3. RESULTS AND DISCUSSION

(a) Error profile as a function of $K_i$ $[C_{C_4H_{10}}(0) = 8 \text{mmol/L}]$

(b) Error profile as a function of $K_i$ $[C_{C_4H_{10}}(0) = 800 \text{mmol/L}]$

(c) Error profile as a function of $K_M$ $[C_{C_4H_{10}}(0) = 8 \text{mmol/L}]$

(d) Error profile as a function of $K_M$ $[C_{C_4H_{10}}(0) = 800 \text{mmol/L}]$

Figure 5.8: Relative errors for model C with the Kin$_2$ mechanism as a function of $K_i$ and $K_M$

The MFA model errors obtained as a function of the reaction rate constant (k) have been shown in Figure 5.9. The results confirmed that an increase in the initial substrate concentrations resulted in a decrease in the overall accumulation rates and hence made the MFA model assumptions more valid (Figures 5.9a and 5.9b). The model errors obtained with the Kin$_2$ mechanism confirmed that the variation in the reaction rate constant did not affect the overall accumulation rate of the intermediates within the reaction volume.
5.3. RESULTS AND DISCUSSION

(a) Error profile as a function of $k$ $[C_{C,H_{10}}(0) = 8\text{mmol/L}]$

(b) Error profile as a function of $k$ $[C_{C,H_{10}}(0) = 800\text{mmol/L}]$

Figure 5.9: Relative errors for model C with the Kin$_{2}$ mechanism as a function of $k$

\[
\begin{align*}
  k &= 0.5s^{-1} & + + \\
  k &= 5s^{-1} & - o - \\
  k &= 50s^{-1} & - \square - 
\end{align*}
\]

5.3.3 Kinetic mechanism Kin$_{3}$ (Michaelis-Menten with no inhibition)

The MFA model errors obtained with the Kin$_{3}$ mechanism as a function of $C_{B}(0)$ and $K_{Eq}$ have been shown in Figure 5.10. The results clearly showed that the lowest model errors were obtained when $O_{2}$ was supplied in the stoichiometric amount at both low and high initial substrate concentrations (Figures 5.10a and 5.10b). The model errors obtained with the Kin$_{3}$ mechanism were very similar to those obtained with the Kin$_{1}$ mechanism at high initial substrate concentrations. The increase in the model errors observed at high conversions with $C_{B}(0) = 8xC_{A}(0)$ was a result of the presence of unreacted $O_{2}$ in the reaction volume. This increased the overall accumulation rates resulting in higher model errors.

The model errors decreased as $K_{Eq}$ increased with the Kin$_{3}$ mechanism (Figures 5.10c and 5.10d). The decrease in the model errors confirmed the observations made with the Kin$_{1}$ and Kin$_{2}$ mechanisms that an increase in $K_{Eq}$ resulted in a decrease in the overall accumulation rate of intermediates within the reaction volume. The sensitivity of the model errors to $K_{Eq}$ increased with the Kin$_{3}$ mechanism. The model errors did not decrease below 5% with the Kin$_{3}$ mechanism at high initial substrate concentrations with $K_{Eq}$ set to 0.012 mmol/L. The higher sensitivity of the model errors was also confirmed by the sensitivity analysis conducted for the cellular dilution rates (Figure 4.8b in Chapter 140 February 2014.
4). The results clearly showed that the relative cellular dilution rates were most sensitive to $K_{\text{Eq}}$ with the Kin$_3$ mechanism. This was due to the form of the mechanism which resulted in a saturation of the consumption rates of the intermediates at a high value of $K_{\text{Eq}}$, without the corresponding reduction in the production rate achieved with the presence of feedback inhibition in the Kin$_1$ and Kin$_2$ mechanisms.

![Figure 5.10: Relative errors for model C with the Kin$_3$ mechanism as a function of $C_B(0)$ and $K_{\text{Eq}}$](image)

The results clearly showed that an increase in $K_M$ resulted in an increase in the model errors with the Kin$_3$ mechanism at both low and high initial substrate concentrations (Figures 5.11a and 5.11b). These results confirmed the observations made regarding the effect of the saturation term on the overall accumulation rates of intermediates within the reaction volume. The sensitivity of the model errors to $K_M$ was also higher with the Kin$_3$ mechanism. This was due to the lack of feedback inhibition in the mechanism which could reduce the accumulation rate of the intermediates. The results clearly show
that any manipulation of the MM constant in a reaction pathway, either through metabolic engineering
techniques or co-feeding a metabolite to suppress gene expression, can have a significant impact on
the validity of the MFA model. The results highlight the sensitivity of the PSS model to the reaction
data generated by pre-determined reaction mechanisms and parameters which can be manipulated
experimentally.

The results obtained for the variation in the reaction rate constant showed once again that there was no
observed sensitivity of the model errors to the reaction rate constant (Figures 5.11c and 5.11d).

Figure 5.11: Relative errors for model C with the Kin_3 mechanism as a function of K_M and k

- K_M = 1.5 mmol/L  - k = 0.5 s^{-1}  
- K_M = 15 mmol/L  - k = 5 s^{-1}  
- K_M = 150 mmol/L  - k = 50 s^{-1}
5.3. RESULTS AND DISCUSSION

5.3.4 Kinetic mechanism Kin4 (First order)

The results obtained with the Kin4 mechanism as a function of \(C_B(0)\) and \(K_{Eq}\) have been shown in Figure 5.12. The results clearly indicate that the model errors decreased with an increase in the initial substrate concentrations. These results confirmed that increasing the initial substrate concentrations had a positive impact on the MFA model errors due to lower overall accumulation rates, regardless of the form of the kinetic mechanisms in the reaction pathway.

The results showed that the model errors were more sensitive to \(C_B(0)\) at both low and high initial substrate concentrations with the Kin4 mechanism (Figures 5.12a and 5.12b). This was confirmed by the sensitivity analysis conducted in Chapter 4. The increase in the sensitivity of the model errors was due to the rapid consumption of the intermediates with the Kin4 mechanism. It can be seen that the model errors obtained at \(C_B(0) = 8xC_A(0)\) at high initial substrate concentrations increased rapidly as the conversion reached 30%. With the other mechanisms, the model errors only increased when the conversions reached 80%. This shows that the accumulation rate for \(O_2\) began to increase at a lower conversion. This was due to the complete consumption of the co-metabolite \(C_2H_4\) in the reaction volume at a lower conversion due to the lack of feedback inhibition or saturation terms in the reaction mechanism.

The model errors decreased with an increase in \(K_{Eq}\) with the Kin4 mechanism (Figures 5.12c and 5.12d). These results confirmed that the increase in \(K_{Eq}\) resulted in a decrease in the overall accumulation rate of intermediates within the reaction volume regardless of the form of the kinetic mechanism. The sensitivity of the model errors to \(K_{Eq}\) decreased with the Kin4 mechanism compared to the Kin3 mechanism. This was confirmed by the results for the sensitivity of the relative cellular dilution rate data as a function of \(K_{Eq}\) shown in Figure 4.11b of Chapter 4. The lack of saturation terms in the Kin4 mechanism meant that the consumption rates of the intermediates were not reduced as the intermediate concentrations increased. This resulted in the lower overall accumulation rates and hence the lower model errors.

The results obtained with the Kin4 mechanism as a function of the reaction rate constant confirmed that the model errors were not sensitive to \(k\) regardless of the form of the reaction mechanism (Figure 5.13). The results obtained for the Kin4 mechanism, as with the other mechanisms, confirmed the observations made in Chapter 4 from the parameter sensitivity analysis. These observations provided
further confirmation for the conclusion that the \( \frac{dC_i}{dt} + \mu C_i \) term excluded from the cellular mass balances in model C represented an overall accumulation rate of intermediates within the reaction volume. Therefore minimising this overall accumulation rate is required to ensure that the PSSH used in the MFA model assumptions is valid.

Figure 5.12: Relative errors for model C with the Kin4 mechanism as a function of \( C_B(0) \) and \( K_{Eq} \)

- \( C_B(0) = 2xC_A(0) \) - + - \( K_{Eq} = 0.012 \text{ mmol/L} \) - + -
- \( C_B(0) = 4xC_A(0) \) - o - \( K_{Eq} = 0.12 \text{ mmol/L} \) - o -
- \( C_B(0) = 8xC_A(0) \) - □ - \( K_{Eq} = 1.2 \text{ mmol/L} \) - □ -
5.3. RESULTS AND DISCUSSION

(a) Error profile as a function of k \([C_{\text{C4H10}}(0) = 8\text{mmol/L}]\)

(b) Error profile as a function of k \([C_{\text{C4H10}}(0) = 800\text{mmol/L}]\)

Figure 5.13: Relative errors for model C with the Kin4 mechanism as a function of k

\[
\begin{align*}
    k &= 0.275\text{hr}^{-1} & - + - \\
    k &= 2.75\text{hr}^{-1} & - o - \\
    k &= 27.5\text{hr}^{-1} & - □ - 
\end{align*}
\]

5.3.5 Steady state MFA models

The results included in this chapter indicated that the MFA model was only accurate when the total accumulation rate \((dC_i/dt + \mu C_i)\) for the intracellular metabolites was negligible compared to the metabolic fluxes. This was feasible for the batch reaction as the cellular dilution rate represented a dilution of the cellular intermediates with respect to cellular growth.

For continuous or steady state cultures, the accumulation rate \((dC_i/dt)\) for the intracellular metabolites will inherently be zero, but the cellular dilution rate \((\mu C_i)\) will not be zero due to the steady state growth that takes place. Bailey and Ollis (1986) state that the mass balance for biomass formation can be simplified according to Eq. [5.1] for continuous cultures. Therefore in a continuous culture, the dilution rate will ultimately control the cellular growth rate and therefore the cellular dilution rate.
5.3. RESULTS AND DISCUSSION

\[ \frac{dC_x}{dt} = C_{x,in}V_{in} - C_{x,out}V_{out} + \mu C_x \] (5.1)

\[ 0 = -\frac{V_{out}}{V}C_x + \mu C_x \]

\[ DC_x = \mu C_x \]

\[ D = \mu \]

The formulation shown above indicates that the validity of the steady state model is dependent on the magnitude of the reactor dilution rate \((D)\). To verify this, the effect of the dilution rate on the MFA model error for a continuous culture was determined. The continuous culture was simulated using standard conditions whilst varying the reactor dilution rate around the maximum cellular growth rate obtained with each mechanism. The maximum growth rate values obtained for the batch systems are shown in Table 3.7 in Section 3.3.1 of Chapter 3.

The model error results obtained for the continuous cultures with each kinetic mechanism as a function of the reactor dilution rate have been shown in Figure 5.14. The x-axis is a logarithmic scale representing the variation in the reactor dilution rate.

It is clearly evident from the results shown in Figure 5.14 that the relative errors obtained for the MFA model increased as the dilution rate was increased with all mechanisms. There was an exponential increase in the model errors as the dilution rate increased beyond the maximum cellular growth rate for each mechanism. This exponential increase in model error represents a move towards “washout” conditions. Bailey and Ollis (1986) characterise “washout” as a rapid loss of cell mass within the reactor volume. The rapid loss of cell mass results in a reduction in the overall product formation rate and an increase in the extracellular substrate concentrations. A move towards “washout” conditions represents a move towards non-steady state behaviour. In the context of this study, this was represented by a rapid increase in the cellular dilution rate and the model errors. Therefore the overall accumulation rate of intracellular metabolites increased as the reactor dilution rate increased, and the MFA model became invalid. The increase in the overall accumulation rate as a function of the reactor dilution rate is illustrated in the results shown in Figure 5.15.

The results generated for the continuous culture clearly illustrate how the increase in the cellular...
dilution rate through increasing the reactor dilution rate can significantly affect the MFA model errors. These observations confirm that the overall accumulation rate \( \left( \frac{dC_i}{dt} + \mu C_i \right) \) affects the validity of the PSSH in both batch and continuous cultures and ultimately the accuracy of the MFA models.

**Figure 5.14: Steady state errors as a function of the dilution rate with all kinetic mechanisms**

- Kin_1 mechanism: \(-+\)
- Kin_2 mechanism: \(-o-\)
- Kin_3 mechanism: \(-□-\)
- Kin_4 mechanism: \(-Δ-\)
Figure 5.15: Steady state cellular dilution rates as a function of the reactor dilution rate with all kinetic mechanisms
5.4 Concluding Remarks

This chapter has achieved its primary objective of determining the accuracy of the MFA model in predicting the reaction rates under various conditions with all kinetic mechanisms. The data showed that the model error was only minimised when the overall accumulation rate \( \frac{dC_i}{dt} + \mu C_i \) was negligible in the cellular mass balance. The results clearly showed that the accumulation rate of reaction intermediates due to both the intracellular reactions and cellular growth were significant in determining the model error. The results shown in this chapter also clearly illustrated the importance of a structured approach to experimental design to ensure the most accurate model can be obtained for the reaction fluxes.

The results showed that the estimated reaction rates were very inaccurate at low conversions. No accurate estimation of the reaction rates were obtained with any kinetic mechanism at conversions below 5%. The model errors increased at high conversions \( (X > 90\%) \). Analysis of the accumulation data clearly showed that this was due primarily to numerical error.

An increase in the initial substrate concentration resulted in lower model errors under all conditions with all kinetic mechanisms. This was due to the minimisation of the overall accumulation rate \( \frac{dC_i}{dt} + \mu C_i \). The lowest model errors were obtained when the secondary substrate was supplied in the correct stoichiometric amount. In practise this may not be possible and therefore the excess metabolite should be chosen such that the total accumulation of mass within the reaction volume will remain low when the excess substrate remains unreacted.

An increase in \( K_{Eq} \) resulted in a decrease in the model errors under all conditions with all kinetic mechanisms. An increase in \( K_i \) resulted in an increase in the model errors under all conditions with the Kin1 and Kin2 mechanisms. The effect of \( K_i \) on the model error was more significant with the Kin2 mechanism due to the form of the inhibition term present in the mechanism.

An increase in \( K_M \) resulted in an increase in the model errors under all conditions with the Kin1, Kin2 and Kin3 mechanisms. The effect of \( K_M \) on the model error was more significant with the Kin3 mechanism due to the lack of inhibition terms present in the mechanism. An increase in \( k \) had no effect on the model error under all condition with all kinetic mechanisms. The MFA model was the most inaccurate with the Kin3 mechanism. The presence of inhibition terms in the kinetic mechanisms
improved the model accuracy. The lack of saturation terms in the Kin₄ mechanism resulted in lower model errors.

The results obtained for the continuous culture simulations as a function of the reactor dilution rate showed an exponential increase in the model errors as the reactor dilution rate approached and exceeded the maximum cellular growth rate. The rapid increase in the cellular dilution rates under these conditions resulted in the low model accuracies and therefore an observed failure of the MFA model. The failure of the MFA model was due to dynamics which were not taken into account in the model.

The model errors presented in this chapter clearly illustrate the effect of both the form of the kinetic mechanism and the parameters used to populate the model on the validity of the PSSH and ultimately the accuracy of the MFA models. The results clearly showed that any “ad-hoc” manipulation of the operating parameters and the mechanics of the cellular reactions can affect the accuracy of the MFA model.
Chapter 6

Conclusions

The research conducted in this study has achieved the objectives set out in Chapter 1. The model data generated with the hypothetical reaction system was validated with experimental data reported in literature. The linear flux models were found to be an accurate estimation of the reaction rates, provided the model was adequately populated with the mass balance data. The models did not suffer from any numerical inaccuracies associated with the linearisation of the reaction rate data.

6.1 Computation of MFA models

The experimental design used to generate data for MFA models should take into account how the model will be constrained with the experimental data. A pre-analysis should be conducted to determine which data will provide the lowest condition number for the reaction matrix. This will provide the best chance of reducing the computational error associated with the linear dependency of a reaction network. The experimental design should also take note of how the manipulation of the reaction pathway, through metabolic engineering techniques, can affect the computational error.

The results from this study clearly showed that severe computational errors cannot be overcome by simply changing the type of numerical routine used to compute the reaction fluxes. It is also important that the choice of flux constraints be made in conjunction with the type of linear solver, especially when a linear programming routine is used. The ill considered use of the inequality and equality constraints in a linear programming routine can result in the generation of an infeasible solution space.
6.2 Effect of the cellular dilution rates on the cellular mass balances

The results generated in this study confirmed that the cellular dilution rates can be considered to be negligible in a cellular mass balance if the cellular dilution rate is at least 20 times less than the magnitude of the reaction fluxes in the pathway. The form of the kinetic mechanism affected the sensitivity of the cellular dilution rate to perturbations of the operating and kinetic parameters. An increase in the inhibition of the rates resulted in a decrease in the sensitivity of the cellular dilution rates while as an increase in the saturation of the rates resulted in an increase in the sensitivity.

Manipulation of the equilibrium constant should only be done if it results in a decrease in the relative intracellular metabolite concentrations. The sensitivity of the cellular dilution rates could not be reduced by only increasing the reaction fluxes. Depending on the form of the kinetic mechanism, increasing the reaction fluxes only decreased the sensitivity of the cellular dilution rates when the overall accumulation of intracellular intermediates was reduced.

6.3 MFA model errors

The results generated in this study clearly showed that the MFA model was valid only when the overall accumulation of intracellular intermediates was minimised. Any ad hoc manipulation of the operating and kinetic parameters can have a significant impact on the model errors. The MFA model had high errors at low substrate conversions. Using the MFA model to simulate the reaction fluxes during the initial stages of a reaction can result in poor model accuracy. This will be true even when the computational errors are low. Attempts to model the reaction fluxes should not be done with substrate conversions below 5%. The majority of the results indicated that the MFA model is most accurate at conversions in excess of 20%. However, this will not guarantee that low model errors will be obtained for all systems. Results obtained with the Kin3 mechanism indicated high model errors were obtained at substrate conversions as high as 40%.

To ensure the best chance of obtaining low MFA model errors, experiments should be designed taking the following qualitative guidelines into account:
• Use high initial substrate concentrations

• Maintain a high equilibrium constant

• Maintain the substrate concentrations as close to the stoichiometric ratio as possible

• If stoichiometric amounts of the substrates cannot be maintained, ensure that the substrate in excess generates the lowest mass of intermediates within the reaction volume once the other substrate is exhausted

• Increase the inhibition for the uptake and accumulation of intermediates

• Decrease the saturation of the primary reaction pathway

It is noted that some of the recommendations provided above may be counter-intuitive, especially when it can result in the decrease in the uptake of substrate into the reaction volume. However, the model data generated in this study clearly shows that the accumulation of intermediates can negatively affect the accuracy of the MFA model. Under these circumstances, it is recommended that a structured model be used rather than the linear flux model.

The results of the study indicate that the simplicity of the MFA model can result in its own downfall, if the model is not correctly designed and matched to the experimental system. In fact, it can be argued that the simplification of the flux model can result in complication of the analyses and the evaluation of the suitability of the model for a set of reaction data.
Bibliography

Aminul Hoque, M., Ushiyama, H., Tomita, M., Shimizu, K., 2005. Dynamic responses of the intracellular metabolite concentrations of the wild type and pykA mutant Escherichia coli against pulse addition of glucose or NH₃ under those limiting continuous cultures. Biochemical Engineering Journal 26, 38 – 49.


Li, M., Yee Ho, P., Yao, S., Shimizu, K., 2006. Effect of lpdA gene knockout on the metabolism in Escherichia coli based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by $^{13}$C-labeling experiments. Journal of Biotechnology 122, 254 – 266.


Part III

Appendices
Appendix A

Reaction Rates

Reaction Pathway DF1

Table A.1: Reaction rates used in pathway DF1 (Kin1 and Kin2 mechanisms)

<table>
<thead>
<tr>
<th>Rate</th>
<th>Kin1</th>
<th>Kin2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Rate_1$</td>
<td>$\frac{C_{2}H_{10}^{*}}{K_{Eq}}/(1 + \frac{C_{2}H_{4}}{K_{1}})$</td>
<td>$\frac{C_{2}H_{10}^{*}}{K_{Eq}} + K_{M}$</td>
</tr>
<tr>
<td>$Rate_2$</td>
<td>$\frac{k C_{2}H_{10}^{<em>} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}{C_{2}H_{10}^{</em>} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{10}^{<em>}}{C_{2}H_{10}^{</em>} + K_{M} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}$</td>
</tr>
<tr>
<td>$Rate_3$</td>
<td>$\frac{C_{2}H_{4}/K_{Eq}}{C_{2}H_{4} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M}}$</td>
</tr>
<tr>
<td>$Rate_4$</td>
<td>$\frac{k C_{2}H_{4}/K_{Eq}}{C_{2}H_{4} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}$</td>
</tr>
<tr>
<td>$Rate_5$</td>
<td>$\frac{k C_{2}H_{4}/K_{Eq} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}{C_{2}H_{4} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}$</td>
</tr>
<tr>
<td>$Rate_6$</td>
<td>$\frac{k C_{2}H_{4}/K_{Eq} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2}) + C_{2}H_{4}/K_{2}}{(C_{2}H_{4} + K_{M}) (C_{2}H_{4} + K_{M} + C_{2}H_{4}/K_{2})}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M} (1 + C_{2}H_{4}/K_{1} + C_{2}H_{4}/K_{2})}$</td>
</tr>
<tr>
<td>$Rate_7$</td>
<td>$\frac{k C_{2}H_{4}/K_{Eq}}{C_{2}H_{4} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M}}$</td>
</tr>
<tr>
<td>$Rate_8$</td>
<td>$\frac{k C_{2}H_{4}/K_{Eq}}{C_{2}H_{4} + K_{M}}$</td>
<td>$\frac{k C_{2}H_{6}}{C_{2}H_{6} + K_{M}}$</td>
</tr>
</tbody>
</table>
Table A.2: Reaction rates used in pathway DF (Kin3 and Kin4 mechanisms)

<table>
<thead>
<tr>
<th>Kin3</th>
<th>Kin4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Rate_1 = \left[ C_{C_2H_10}^{int} \right] / \left[ C_{C_2H_10}^{eq} + K_M \right]$</td>
<td>$Rate_1 = k_0 C_{C_2H_10}^{eq}$</td>
</tr>
<tr>
<td>$Rate_2 = \left[ C_{C_2H_10}^{int} \right] / \left[ C_{C_2H_10}^{eq} + K_M \right]$</td>
<td>$Rate_2 = k_0 C_{C_2H_10}^{int}$</td>
</tr>
<tr>
<td>$Rate_3 = \left[ C_{C_2H_6}^{int} \right] / \left[ C_{C_2H_6}^{eq} + K_M \right]$</td>
<td>$Rate_3 = k_0 C_{C_2H_6}^{int}$</td>
</tr>
<tr>
<td>$Rate_4 = \left[ C_{C_2H_4}^{int} \right] / \left[ C_{C_2H_4}^{eq} + K_M \right]$</td>
<td>$Rate_4 = k_0 C_{C_2H_4}^{int}$</td>
</tr>
<tr>
<td>$Rate_5 = \left[ C_{C_2H_4}^{int} / K_M \right] / \left[ C_{C_2H_4}^{eq} + K_M \right]$</td>
<td>$Rate_5 = k_0 C_{C_2H_4}^{int}$</td>
</tr>
<tr>
<td>$Rate_6 = \left[ C_{C_2H_4}^{int} \right] / \left[ C_{C_2H_4}^{eq} + K_M \right]$</td>
<td>$Rate_6 = k_0 C_{C_2H_4}^{int}$</td>
</tr>
<tr>
<td>$Rate_7 = \left[ C_{CO_2}^{int} \right] / \left[ C_{CO_2}^{eq} + K_M \right]$</td>
<td>$Rate_7 = k_0 C_{CO_2}^{int}$</td>
</tr>
<tr>
<td>$Rate_8 = \left[ C_{H_2O}^{int} \right] / \left[ C_{H_2O}^{eq} + K_M \right]$</td>
<td>$Rate_8 = k_0 C_{H_2O}^{int}$</td>
</tr>
</tbody>
</table>
### Reaction Pathways DF2 and DF3

#### Table A.3: Reaction rates used in pathway DF2 and DF3

<table>
<thead>
<tr>
<th>Kin1</th>
<th>Kin1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Rate_1 = \frac{[C_{\text{H}<em>2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_4 = \frac{[C_{\text{H}<em>2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_2 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_5 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_3 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_6 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_4 = \frac{[C_{\text{H}<em>2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_7 = \frac{[C_{\text{H}<em>2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_5 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_8 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_6 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_9 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
<tr>
<td>$Rate_7 = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
<td>$Rate_{10} = \frac{[C_{\text{C}<em>2H_2}]^{\text{ref}}}{k</em>{\text{C}<em>2H_4} + C</em>{\text{H}_2}}$</td>
</tr>
</tbody>
</table>
Appendix B

Reaction Matrices

The reaction matrices generated for reaction pathways DF₁ and DF₂ have been shown below. The reaction matrices for pathway DF₃ have not been included. The only reaction matrix which used to compute the DF₃ reaction pathway has been shown in Section 3.2.3 of Chapter 3. The set of linear equations shown below illustrate the reaction matrices used to compute model C, where the constraint vector was set to zero. Only the extracellular fluxes used to constrain the matrices are included in the constraint vector. No accumulation or cellular dilution rates were used in model C.
Table B.1: Equations used for the reaction pathway DF\textsubscript{1}

<table>
<thead>
<tr>
<th>Constraint Set</th>
<th>Solution Matrix</th>
</tr>
</thead>
</table>
| 1              | \[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -4 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & -2 & 0 & 0 \\
0 & 1 & -1 & 0 & 1.75 & 0 & 0 \\
0 & 0 & 0 & 0 & 3 & -1 & 0 \\
0 & 0 & 0 & 0 & 1.5 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_7 \\
F_8
\end{bmatrix} = \begin{bmatrix}
F_{\text{ext}}^{\text{CH}_4}\text{H}_10 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\] |
| 2              | \[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -4 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & -2 & 0 & 0 \\
0 & 0 & 1 & 1 & 1.75 & 0 & 0 \\
0 & 0 & 0 & 0 & 3 & -1 & 0 \\
0 & 0 & 0 & 0 & 1.5 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_8
\end{bmatrix} = \begin{bmatrix}
F_{\text{H}_2}^{\text{ext}} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\] |
| 3              | \[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -4 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & -2 & 0 & 0 \\
0 & 0 & 1 & -1 & 1.75 & 0 & 0 \\
0 & 0 & 0 & 0 & 3 & -1 & 0 \\
0 & 0 & 0 & 0 & 1.5 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_8
\end{bmatrix} = \begin{bmatrix}
F_{\text{O}_2}^{\text{ext}} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\] |
| 4              | \[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -4 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & -2 & 0 \\
0 & 0 & 1 & -1 & 0 & 1.75 & 0 \\
0 & 0 & 0 & 0 & 0 & 3 & 0 \\
0 & 0 & 0 & 0 & 0 & 1.5 & -1
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_8
\end{bmatrix} = \begin{bmatrix}
F_{\text{CO}_2}^{\text{ext}} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\] |
| 5              | \[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -4 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & -2 & 0 \\
0 & 0 & 1 & -1 & 0 & 1.75 & 0 \\
0 & 0 & 0 & 0 & 0 & 3 & -1 \\
0 & 0 & 0 & 0 & 0 & 1.5 & 0
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2 \\
F_3 \\
F_4 \\
F_5 \\
F_6 \\
F_7
\end{bmatrix} = \begin{bmatrix}
F_{\text{H}_2}\text{O}^{\text{ext}} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\] |
### Reaction pathway DF$_2$

Table B.2: Equations used for the reaction pathway DF$_2$ (Constr sets 1 - 5)

<table>
<thead>
<tr>
<th>Constraint Set</th>
<th>Solution Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\begin{bmatrix} -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{bmatrix}$</td>
</tr>
<tr>
<td>2</td>
<td>$\begin{bmatrix} -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{bmatrix}$</td>
</tr>
<tr>
<td>3</td>
<td>$\begin{bmatrix} -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{bmatrix}$</td>
</tr>
<tr>
<td>4</td>
<td>$\begin{bmatrix} 1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{bmatrix}$</td>
</tr>
<tr>
<td>5</td>
<td>$\begin{bmatrix} 1 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{bmatrix}$</td>
</tr>
</tbody>
</table>
Table B.3: Equations used for the reaction pathway DF$_2$ (Constr sets 6 - 10)

<table>
<thead>
<tr>
<th>Constraint Set</th>
<th>Solution Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>$F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_9 F_{H_2} F_{CO_2}$</td>
</tr>
<tr>
<td>7</td>
<td>$F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_9 F_{H_2O}$</td>
</tr>
<tr>
<td>8</td>
<td>$F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_9 F_{CO_2}$</td>
</tr>
<tr>
<td>9</td>
<td>$F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_{H_2O}$</td>
</tr>
<tr>
<td>10</td>
<td>$F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_{CO_2}$</td>
</tr>
</tbody>
</table>
Appendix C

Reactors and Cellular Mass Balances

The mass balances for the extracellular and intracellular metabolites have been shown for each reaction pathway. The mass balances for continuous cultures have also been included for reaction pathway DF 1.

In the continuous mass balances, the accumulation rates have been included. These were necessary in the mass balances to simulate the pre-steady state behaviour. Once the system reached steady state, the accumulation rates are automatically set to zero once the concentration data reached steady state.

The accumulation and cellular dilution rate data was computed outside of the ODE routine using the concentration data as follows:

\[
\mu(i) = \frac{1}{C_x(i)} \times \left( \frac{C_x(i+1) - C_x(i-1)}{t(i+1) - t(i-1)} \right)
\]

\[
\mu C_i(i) = \mu(i) \times C_i(i) \times \hat{V}
\]

\[
\frac{dC_i}{dt}(i) = \left( \frac{C_i(i+1) - C_i(i-1)}{t(i+1) - t(i-1)} \right) \times \hat{V}
\]

169
For the cellular dilution rates and accumulation rates to be used in the flux model by populating the constraint vector $b$ (See Section 2.4.1 of Chapter 2), the rates were made intrinsic to the gDW biomass to match the unit of the reaction fluxes [mmol/gDW.hr]. This was achieved using the specific cellular volume ($\hat{V}$).

The $\mu C$ and $\frac{dC}{dt}$ terms shown in the mass balances in this section were obtained directly from the ODE routine and were not computed using the formulae shown above. The formulae shown above illustrate the computation of these terms for the population of the flux constraint model using the concentration obtained as an output from the ODE routine.
**Reaction pathway DF**

Table C.1: Extracellular mass balances for the batch process with reaction pathway DF

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
</table>
| $\text{C}_4\text{H}_{10}$ | \[
\frac{dC_{\text{ext}}^{\text{C}_4\text{H}_{10}}}{dt} = -\text{Rate}_1 C_x \hat{V}
\] |
| $\text{O}_2$ | \[
\frac{dC_{\text{ext}}^{\text{O}_2}}{dt} = -\text{Rate}_5 C_x \hat{V}
\] |
| $\text{H}_2$ | \[
\frac{dC_{\text{ext}}^{\text{H}_2}}{dt} = \text{Rate}_4 C_x \hat{V}
\] |
| $\text{CO}_2$ | \[
\frac{dC_{\text{ext}}^{\text{CO}_2}}{dt} = \text{Rate}_7 C_x \hat{V}
\] |
| $\text{H}_2\text{O}$ | \[
\frac{dC_{\text{ext}}^{\text{H}_2\text{O}}}{dt} = \text{Rate}_8 C_x \hat{V}
\] |
| Biomass | \[
\frac{dC_x}{dt} = \text{Rate}_6 C_x \hat{V}
\] |

Table C.2: Extracellular mass balances for the continuous process with reaction pathway DF

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
</table>
| $\text{C}_4\text{H}_{10}$ | \[
\frac{dC_{\text{ext}}^{\text{C}_4\text{H}_{10}}}{dt} = -\text{Rate}_1 C_x \hat{V} + (V_{\text{in}} C_{\text{ext}}^{\text{feed}} - V_{\text{out}} C_{\text{ext}}^{\text{C}_4\text{H}_{10}})
\] |
| $\text{O}_2$ | \[
\frac{dC_{\text{ext}}^{\text{O}_2}}{dt} = -\text{Rate}_5 C_x \hat{V} + (V_{\text{in}} C_{\text{ext}}^{\text{O}_2} - V_{\text{out}} C_{\text{ext}}^{\text{O}_2})
\] |
| $\text{H}_2$ | \[
\frac{dC_{\text{ext}}^{\text{H}_2}}{dt} = \text{Rate}_4 C_x \hat{V} - V_{\text{out}} C_{\text{ext}}^{\text{H}_2}
\] |
| $\text{CO}_2$ | \[
\frac{dC_{\text{ext}}^{\text{CO}_2}}{dt} = \text{Rate}_7 C_x \hat{V} - V_{\text{out}} C_{\text{ext}}^{\text{CO}_2}
\] |
| $\text{H}_2\text{O}$ | \[
\frac{dC_{\text{ext}}^{\text{H}_2\text{O}}}{dt} = \text{Rate}_8 C_x \hat{V} - V_{\text{out}} C_{\text{ext}}^{\text{H}_2\text{O}}
\] |
| Biomass | \[
\frac{dC_x}{dt} = \text{Rate}_6 C_x \hat{V} - V_{\text{out}} C_x
\] |
Table C.3: Intracellular mass balances for reaction pathway DF

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C}<em>4\text{H}</em>{10}$</td>
<td>$\frac{dC_{\text{C}<em>4\text{H}</em>{10}}}{dt} = \text{Rate}_1 - \text{Rate}<em>2 - \mu C</em>{\text{C}<em>4\text{H}</em>{10}}$</td>
</tr>
<tr>
<td>$\text{O}_2$</td>
<td>$\frac{dC_{\text{O}_2}}{dt} = \text{Rate}_5 - 4\text{Rate}<em>6 - \mu C</em>{\text{O}_2}$</td>
</tr>
<tr>
<td>$\text{C}_2\text{H}_6$</td>
<td>$\frac{dC_{\text{C}_2\text{H}_6}}{dt} = \text{Rate}_2 - \text{Rate}<em>3 - \mu C</em>{\text{C}_2\text{H}_6}$</td>
</tr>
<tr>
<td>$\text{C}_2\text{H}_4$</td>
<td>$\frac{dC_{\text{C}_2\text{H}_4}}{dt} = \text{Rate}_2 + \text{Rate}_3 - 2\text{Rate}<em>6 - \mu C</em>{\text{C}_2\text{H}_4}$</td>
</tr>
<tr>
<td>$\text{H}_2$</td>
<td>$\frac{dC_{\text{H}_2}}{dt} = \text{Rate}_3 + 1.75\text{Rate}_6 - \text{Rate}<em>4 - \mu C</em>{\text{H}_2}$</td>
</tr>
<tr>
<td>$\text{CO}_2$</td>
<td>$\frac{dC_{\text{CO}_2}}{dt} = 3\text{Rate}_6 - \text{Rate}<em>7 - \mu C</em>{\text{CO}_2}$</td>
</tr>
<tr>
<td>$\text{H}_2\text{O}$</td>
<td>$\frac{dC_{\text{H}_2\text{O}}}{dt} = 1.5\text{Rate}_6 - \text{Rate}<em>8 - \mu C</em>{\text{H}_2\text{O}}$</td>
</tr>
</tbody>
</table>
### Table C.4: Extracellular mass balances for the batch process with reaction pathway DF$_2$

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$<em>4$H$</em>{10}$</td>
<td>$\frac{dC_{\text{extr}}}{dt} = -\text{Rate}_1 \dot{C}_x \hat{V}$</td>
</tr>
<tr>
<td>O$_2$</td>
<td>$\frac{dC_{O_2}}{dt} = -\text{Rate}_2 \dot{C}_x \hat{V}$</td>
</tr>
<tr>
<td>H$_2$</td>
<td>$\frac{dC_{\text{H}_2}}{dt} = \text{Rate}_4 \dot{C}_x \hat{V}$</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>$\frac{dC_{CO_2}}{dt} = \text{Rate}_8 \dot{C}_x \hat{V}$</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>$\frac{dC_{\text{H}_2O}}{dt} = \text{Rate}_9 \dot{C}_x \hat{V}$</td>
</tr>
<tr>
<td>Biomass</td>
<td>$\frac{dC_x}{dt} = \text{Rate}_7 \dot{C}_x \hat{V}$</td>
</tr>
</tbody>
</table>

### Table C.5: Intracellular mass balances for reaction pathway DF$_2$

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$<em>4$H$</em>{10}$</td>
<td>$\frac{dC_{\text{int}}}{dt} = \text{Rate}<em>1 - \text{Rate}<em>2 - \mu C</em>{\text{int}}^{C_4H</em>{10}}$</td>
</tr>
<tr>
<td>O$_2$</td>
<td>$\frac{dC_{O_2}}{dt} = \text{Rate}_5 - 3.5 \text{Rate}_6 - 4 \text{Rate}<em>7 - \mu C</em>{\text{int}}^{O_2}$</td>
</tr>
<tr>
<td>C$_2$H$_6$</td>
<td>$\frac{dC_{\text{int}}}{dt} = \text{Rate}_2 - \text{Rate}_3 - \text{Rate}<em>6 - \mu C</em>{\text{int}}^{C_2H_6}$</td>
</tr>
<tr>
<td>C$_2$H$_4$</td>
<td>$\frac{dC_{\text{int}}}{dt} = \text{Rate}_2 + \text{Rate}_3 - 2 \text{Rate}<em>7 - \mu C</em>{\text{int}}^{C_2H_4}$</td>
</tr>
<tr>
<td>H$_2$</td>
<td>$\frac{dC_{\text{int}}}{dt} = \text{Rate}_3 + 1.75 \text{Rate}_7 - \text{Rate}<em>4 - \mu C</em>{\text{int}}^{H_2}$</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>$\frac{dC_{\text{int}}}{dt} = 2 \text{Rate}_6 + 3 \text{Rate}_7 - \text{Rate}<em>8 - \mu C</em>{\text{int}}^{CO_2}$</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>$\frac{dC_{\text{int}}}{dt} = 3 \text{Rate}_6 + 1.5 \text{Rate}_7 - \text{Rate}<em>9 - \mu C</em>{\text{int}}^{H_2O}$</td>
</tr>
</tbody>
</table>
### Reaction pathway DF₃

#### Table C.6: Extracellular mass balances for the batch process with reaction pathway DF₃

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₄H₁₀</td>
<td>[ \frac{dC_{\text{C}_4\text{H}_10}^{\text{ext}}}{dt} = -\text{Rate}_1 C_x \dot{V} ]</td>
</tr>
<tr>
<td>O₂</td>
<td>[ \frac{dC_{\text{O}_2}^{\text{ext}}}{dt} = -\text{Rate}_7 C_x \dot{V} ]</td>
</tr>
<tr>
<td>H₂</td>
<td>[ \frac{dC_{\text{H}_2}^{\text{ext}}}{dt} = \text{Rate}_5 C_x \dot{V} ]</td>
</tr>
<tr>
<td>CO₂</td>
<td>[ \frac{dC_{\text{CO}_2}^{\text{ext}}}{dt} = \text{Rate}_10 C_x \dot{V} ]</td>
</tr>
<tr>
<td>H₂O</td>
<td>[ \frac{dC_{\text{H}_2\text{O}}^{\text{ext}}}{dt} = \text{Rate}_9 C_x \dot{V} ]</td>
</tr>
<tr>
<td>Biomass</td>
<td>[ \frac{dC_x}{dt} = \text{Rate}_8 C_x \dot{V} ]</td>
</tr>
</tbody>
</table>
Table C.7: Intracellular mass balances for reaction pathway DF3

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C}<em>4\text{H}</em>{10}$</td>
<td>[ \frac{dC^\text{int}_{\text{C}<em>4\text{H}</em>{10}}}{dt} = \text{Rate}_1 - \text{Rate}_2 - \text{Rate}<em>3 - \mu C^\text{int}</em>{\text{C}<em>4\text{H}</em>{10}} ]</td>
</tr>
<tr>
<td>$\text{O}_2$</td>
<td>[ \frac{dC^\text{int}_{\text{O}_2}}{dt} = \text{Rate}_7 - 6.5\text{Rate}_2 - 3.5\text{Rate}_6 - 4\text{Rate}<em>8 - \mu C^\text{int}</em>{\text{CO}_2} ]</td>
</tr>
<tr>
<td>$\text{C}_2\text{H}_6$</td>
<td>[ \frac{dC^\text{int}_{\text{C}_2\text{H}_6}}{dt} = \text{Rate}_3 - \text{Rate}_4 - \text{Rate}<em>6 - \mu C^\text{int}</em>{\text{C}_2\text{H}_6} ]</td>
</tr>
<tr>
<td>$\text{C}_2\text{H}_4$</td>
<td>[ \frac{dC^\text{int}_{\text{C}_2\text{H}_4}}{dt} = \text{Rate}_3 + \text{Rate}_4 - 2\text{Rate}<em>8 - \mu C^\text{int}</em>{\text{C}_2\text{H}_4} ]</td>
</tr>
<tr>
<td>$\text{H}_2$</td>
<td>[ \frac{dC^\text{int}_{\text{H}_2}}{dt} = \text{Rate}_4 + 1.75\text{Rate}_8 - \text{Rate}<em>5 - \mu C^\text{int}</em>{\text{H}_2} ]</td>
</tr>
<tr>
<td>$\text{CO}_2$</td>
<td>[ \frac{dC^\text{int}_{\text{CO}_2}}{dt} = 4\text{Rate}_2 + 2\text{Rate}_6 + 3\text{Rate}_8 - \text{Rate}<em>10 - \mu C^\text{int}</em>{\text{CO}_2} ]</td>
</tr>
<tr>
<td>$\text{H}_2\text{O}$</td>
<td>[ \frac{dC^\text{int}_{\text{H}_2\text{O}}}{dt} = 2\text{Rate}_2 + 3\text{Rate}_6 + 1.5\text{Rate}_8 - \text{Rate}<em>9 - \mu C^\text{int}</em>{\text{H}_2\text{O}} ]</td>
</tr>
</tbody>
</table>
Appendix D

MATLAB Source Code

The raw source code has been included as data files on the data disc for the thesis. The raw source code could not be included in the thesis due to the size of the document.