UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS AND COPYRIGHT

Authors full name : **MOHAMMED ADAM KUNNA AZRAG**

Date of birth : **22nd MARCH 1987**

Title : **KINETIC PARAMETERS IDENTIFICATION FOR LARGE-SCALE METABOLIC MODEL OF *ESCHERICHA COLI***

Academic Session : **2014/2015**

I declare that this thesis is classified as:

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*

☑ **OPEN ACCESS** I agree that my thesis to be published as online open access (Full text)

I acknowledge that Universiti Malaysia Pahang reserve the right as follows:

1. The Thesis is the Property of Universiti Malaysia Pahang.
2. The Library of Universiti Malaysia Pahang has the right to make copies for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

Certified By:

_____ _____

(Students Signature) (Signature of Supervisor)

P00254313 DR. TUTY ASMAWATY ABDUL KADIR

Num IC/Passport Number Name of Supervisor

Date: Date:

KINETIC PARAMATERS IDENTIFICATION FOR LARGE-SCLE METABOLIC MODEL OF *ESCHERICHIA COLI*

MOHAMMED ADAM KUNNA AZRAG

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Master of Computer Science

Faculty of Computer System & Software Engineering
UNIVERSITI MALAYSIA PAHANG

2015

SUPERVISORS DECLARATION

I hereby declare that I have checked this thesis and any opinion, this thesis is adequate in terms of scope and quality of the award of the degree of Master of Science in (Computer).

Signature : .....

Name of Supervisor : DR. TUTY ASMAWATY ABDUL KADIR

Position :

Date :

STUDENT DECLARATION

I hereby declare that the work in this thesis is my on except for quotations and summaries which have been duly acknowledged. The thesis has not been accepted for any degree and is not concurrently submitted for award of another degree.

Signature : ……

Name : MOHAMMED ADAM KUNNA AZRAG

ID Number : MCC12002

Date :

DEDICATED TO

My

Parents;

Brothers and sister

Friends

.

ACKNOWLEDGEMENT

First and for most, I want to give all the praise and glory to my Almighty Allah. I am greatly grateful for all the difficulties and testing He put upon me for my own sake in the future.

I would like especially to thank my supervisor Doctor Tuty Asmawaty Abdul Kadir for her advice and guidance in the development of this research. It has truly been a pleasure to work with you, and I appreciate the supervision you have given me to accomplish this work.

I would like to express my sincere gratitude to Universiti Malaysia Pahang for granting me a master scholarship, without their support, my ambition to study abroad can hardly be realized. It was a wonderful place to work, and they are very dedicated people. I gained so much from it. Furthermore, special thanks to the academic, management and technical staff in Faculty of Computer System and Software Engineering and the staff of Institute of Postgraduate Studies (IPS) in UMP. I thank all my friends and colleagues for every bit of support; I thank to all Malaysian people whom I met, for their openness, friendship and hospitality.

At last and most important, I would like to thank my parents and all family members for their open-mindedness and endless support. They are in my heart. Without the help and support of Allah, supervisors and all these people, this thesis would not be completed.

ABSTRACT

One of the biggest challenging in metabolic engineering is to design an accurate model of large-scale of metabolic network in metabolic engineering field; which require an appropriate sensitivity analysis and optimization techniques. This research focusing on identifying the optimize values of large-scale kinetic parameters of *E. coli* model. The model under study consist of five metabolic pathways which are Glycolysis, Pentose Phosphate, TCA cycle, Gluconegenesis and Glycoxylate; which contain 194 kinetic parameters to be optimize. This model also includes PTS system in addition to Acetate formation, 23 metabolites, 28 enzymatic reactions and 10 co-factors. The experimental data were run in 0.1 and 0.2 dilution rates at continuous culture on steady-state condition. The One-At-A-Time Sensitivity Measure and Particle Swarm Optimization (PSO) techniques was applied to the model under study in order to identify the optimum values of the kinetics. The result stated from the One-At-A-Time Sensitivity Measure shows that there are 7 kinetics affecting highly in the model response under 0.1 dilution rate, while in 0.2 there are 8 kinetics affecting highly in the model response also. The result stated from PSO shows that, this technique can minimize the errors of our simulation result by % as compare to (Ishii et al., 2007) and % as compare to (Hoque et al., 2005). Based on the results found by the techniques, these tichniques can be applied to correct the model response through large-scale kinetic parameters.

ABSTRAK

Menghasilkan model yang tepat bagi rangkaian metabolik berskala besar merupakan satu cabaran besar dalam bidang kejuruteraan metabolik yang mana ianya memerlukan penggunaan teknik analisa sensitif dan pengoptimuman yang baik. Kajian ini memfokus kepada pencarian nilai bagi parameter kinetik dalam model rangkaian metabolik *Escherichia coli* berskala besar. Model ini terdiri daripada 194 pembolehubah kinetik dari lima laluan metabolic; *glycolysis*, *pentose phosphate*, *TCA cycle*, *Gluconeogenesis* dan *Glycolysis*. Model ini juga mengandungi laluan metabolik bagi pembentukan Acetate dan sistem PTS meliputi 23 metabolite, 28 tindakbalas enzim dan 10 co-factor. Data bagi eksperimen di laksanakan pada kondisi keadaan tetap dan kultur berterusan dalam kadar pencairan 0.1D dan 0.2D. Teknik *One at a Time Sensitivity Measure* dan *Particle Swarm Optimization* di cadang untuk digunakan dalam mengenalpasti nilai optimum bagi parameter kinetik agar hasil simulasi adalah selari dengan nilai eksperimen. Hasil analisa sensitif menggunakan *One at a Time Sensitivity Measure* terhadap 194 parameter kinetik mendapati, tujuh pembolehubah yang sangat terjejas pada kadar peningkatan konsentrasi 40% pada kadar pencairan 0.1D, manakala 8 parameter kinetik sangat terjejas pada kadar pencairan 0.2D. Penggunaan teknik *Particle Swarm Optimization* pula dilihat terbukti dapat mengurangkan ralat sebanyak 294% hasil simulasi berbanding dengan data ekperimen yang ekperimen yang dibuat oleh Ishii, 2007, 11% dari eksperimen data oleh Hoque, 2005 berbanding hasil simulasi asal (Kadir, 2010).

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Pts	Phosphotransferase system
Pgi	Phosphoglucose isomerase / Glucosephosphate isomerase
Pfk	Phosphofructokinase-1
Aldo	Aldolase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
Pyk	Pyruvate kinase
Pdh	Pyruvate dehydrogenase
Aces	Acetylcoenzyme A synthetase
Pta	Phosphotransacetylase
Ack	Acetate kinase
cs	Citrate synthase
ICDH	Isocitrate dehydrogenase
2KGDH	2-Keto-D-gluconate Dehydrogenase
SDH	Succinate dehydrogenase
Fum	Fumarase
MDH	Malate dehydrogenase
Mez	Malic enzyme
Pck	Phosphoenolpyruvate carboxykinase
Ppc	PEP carboxylase
ICL	Isocitrate lyase
Ms	Malate synthase
G6pdh	Glucose-6-phosphate dehydrogenase
6Pgdh	6Phsophogluconate dehydrogenase
Rpi	Ribulose 5phosphate 3-isomerase
Rpe	Ribulose phosphate 3epimerase
Tkta	TransketolaseI
Tktb	TransketolaseII
Tal	Transaldolase

GLCex	Glucose
G6P	Glucose-6-phosphate
F6P	Fructose-6-phosphate
FDP	Fructose 1,6-bisphosphate
GAP	Glyceraldehyde 3-phosphate
DHAP	Dihydroxyacetone phosphate
PEP	Phosphoenolpyruvate
PYR	Pyruvate
AcCOA	Acetyl-CoA
AcP	Acetylphosphate
ACE	Acetate
ICIT	Isocitrate
2KG	2-Keto-Dgluconate
SUC	Succinate
FUM	Fumarate
MAL	Malate
OAA	Oxaloacetate
6PG	6-Phosphogluconolactone
Ru5P	Ribose 5-phosphate
Xu5P	Xylulose 5-phosphate
R5P	Ribulose 5-phosphate
S7P	Sedoheptulose 7-phosphate
E4P	Erythrose 4-phosphate
Gox	Glyxoylate
D	Dilution rate
mM	Milli-molar
g/l	Gram/ Liter
ADP	Adenosine diphosphate
ATP	Adenosine-5-triphosphte
AMP	Dihydroxyacetone phosphate
NAD/NADH	Nicotinamide adenine dinucleotide

NADP/NADPH	Nicotinamide adenine dinucleotide phosphate
P	Phosphate
H	Histidine
COA	Coenzyme A

CHAPTER 1

INTRODUCTION

1.1 Introduction

Metabolic engineering has become very important in the production of a new scientific endeavor in *E. coli* (Edwards and Palsson, 1997). Metabolic engineering based on genetic engineering which is the targeted manipulation of genetic-cell information involves enzymatic, transport and regulatory gene which are the goals of direct modifications and the improvement of cellular activities (Ka et al., 1998). The classical approach of metabolic engineering requires detailed knowledge of enzyme kinetics, the system of work, intermediate pools involved and genetic manipulation (Gregory, 1999).

However, metabolic engineering is usually faced with the challenges of effectively developing and designing the cell metabolism with respect to the metabolism regulation. In order to address this, it is necessary to generate a mathematical model which can efficiently describe the dynamic behavior of the cell in response to the changes in the cultural environment and/or the specific genetic modification (Kadir et al., op cit). In fact, analysis of the sensitivity, genetic optimizing and regulatory processes are the metabolic engineering practice within cells which are done to increase the cellular production of a certain substance.

With a view to studying the dynamics of the metabolic engineering system, there is need to consider how the substrate is converted to Substrate or to a Product and which enzymes should be involved in the conversion process. The conversions in the

metabolic networks consist of a substrate and product and also between them the enzymes which can convert the substrate to product either in an irreversible or reversible way (Kadir et al., 2010). This is described in Figure 1.1. The study of the substrate, enzyme and product conversions are achieved by metabolic computational.

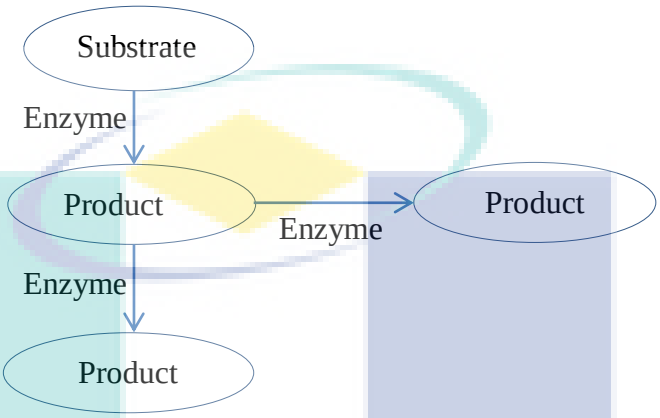


Figure 1.1: Metabolic conversion

1.1.1 Metabolic Computational

Metabolic computational modeling plays a substantial role in the biological system. Every modeling has been constructed using ordinary differential equations (ODEs). The accuracy of the model output prediction would, however, depend on the behavior system physiology, which has a set of parameters such as temperature, reaction rates and kinetic constants. It had been reported that one of the powerful tools for explaining the properties of the dynamic metabolic engineering system as well as to guide experimentation is metabolic network model (Maggioa et al., 2010). Also, it was reported that to build a kinetic metabolic network model requires a large number of kinetic parameters, which has been developed to detect the concentration changes in the metabolites and reactions (Chassagnole et al., 2002). Some of the mathematical models which can describe the dynamic models have been suggested with a view to survey the behavior of the cell. Some used flux balance analysis (Reed and Palsson, 2003), (Radhakrishnanet al., 2002), (Edwards et al., 2001), network component analysis (S. Shuster et al., 2000; Liao et al., 2003), C-metabolic flux analysis (Siddiquee et al.,2004;

Toya et al., 2010), dynamic modelling (Chassagnole et al., 2002; Usuda et al., 2010), metabolic analysis design (Simon and David, 1996), Metabolic control analysis (Diana and Joseph, 2002) and the steady-state of the model (Barbara et al., 1992). In order to simulate the kinetic parameters in a model, there is the need to consider the mathematical equations as simple as possible so that the implementation will become easy.

If the modeling can be effectively simulated, it can be of a great help at in answering some specific questions such as the accuracy of the model outputs. These models are however declared using simulation and represented by specific or some part of metabolic pathways (Chassagnole et al., 2002; Yugi et al., 2005; Ishii et al., 2007; Kremling et al., 2007; Nishio et al., 2008; Kadir et al., op cit). After the model is build, there would be need for sensitivity analysis in order to optimize the model.

1.1.2 Sensitivity analysis

Engineering and science are often studied with the aid of mathematical models designed to simulate the complex physical process (Gangelosi and Parisi, 2001). One of the steps in mathematical model development is the determination of the most effective parameter in the model outputs. A "sensitive analysis" of these parameters is not only definite to model validation, but also it can lead to future research. Sensitivity analysis is often referred to as either local or global. The local analysis addresses sensitivity relative to point estimates of parameter values while global analysis examines sensitivity with regard to the entire parameter distribution. Sensitive analysis can help the researcher to determine which parameter enables the very effectiveness of the model's result (Saltelli, 2000).

The sensitive analysis method can be classified in a variety of ways: statistical, mathematical or graphical. The statistical method involves running a simulation in which an input is assigned some probability distributions, and later the assessment of the effect of variance on the input is done to identify the output distribution. Also, it can allow one input to identify the effects of the interaction among multiple inputs (Griensven et al., 2006). The mathematical method is the sensitivity of a model output

to the range of variation of an input. The method typically involves calculating the output for a few values of an input that represent the possible range of the input (Salehi et al., 2000). The graphical methods give representation of sensitivity in the form of graphs, charts or surface. Generally the graphical method is used to give visual indication of how an output is affected by variation in inputs (Gelderman and Rentez, 2001).

1.1.3 Optimization

Optimization means to find out the best alternative with the most cost effective or highest achievable performance under the given constraints, such as the best result, the best design, among other options. The optimization problem is generally aimed towards minimizing the difference between the model outputs estimated parameters and the respective experimental measurements. It had been stated that in the metabolic engineering model, the kinetic parameter optimization problem of kinetic model can be formulated as an estimation problem (Yukako et al., 2013).

However, recently several researches have been done with some algorithms in order to study the structure and behavior of the cell. Some of the algorithms used are Least Squares Minimization (Rizzi et al., 1997), Simulated Annealing (Chassagnole et al., 2002), a second order polynomial model in RSM (Ismail, 2005), DEPSO algorithm (Rui et al., 2007), a weight least squares objective function (Won et al., 2012), IDE algorithm (Chong et al., 2012), a deterministic outer-approximation algorithm (Miro et al., 2012), a real-coded genetic algorithm (Yukako et al., 2013), formulation of a parameter optimization problem within a control vector parameterization approach (Maggio et al., 2013) and PSEO algorithm (Abdullah et al., 2013).

From the findings of these researchers, it was noted that the problem of designing and validating the metabolic engineering model can be solved through simulation, modeling, analysis and optimization when there are some data available for the pathways which include metabolites, enzymes and co-factors.

1.2 Problem statement

Metabolic engineering allows the direct application of the core subjects of kinetics, transports and thermodynamics to the analysis of the reactions of metabolic networks (Gregory, 1999). On the other hand, metabolic network description provides convenient ways of summarizing and codifying the information gathered from the metabolism of an organism. The most successful scientific tools that can represent the metabolic networks are the mathematical modelling (Wagner, 2012). However, mathematical modeling of metabolism is usually closely associated with changes in compound concentration in terms of rates of biochemical reaction (Gombert and Nielsen, 2000).

In metabolic engineering, one of the biggest challenges is how to design an accurate kinetic model that represents the large-scale number of parameters in the pathways. Also, kinetic constant and initial metabolite constant of the metabolic network of *E. coli* from an estimated measurement value or from vitro are a big challenge, because the kinetic parameters that are usually obtained or estimated from measurements reported by different laboratories using different models and conditions stored in databases are insufficient (Yukako et al., 2013).

During the last years, the design of large-scale metabolic network of *E. coli* (build – develop) has been greatly advanced by a systematic application of modeling, simulation and optimization based on the available data (Jeong et al., 2000). The model that has been working is tested in software programing and point towards genetic modification in the pathway reaction that will lead to predicting new design models (Kadir et al., 2010).

The issues of designing an accurate kinetic metabolic model especially in *E. coli* is solved by many methods such as sensitivity analysis and optimization methods. To this end, the latest methods involved are stated below with their problems and solutions.

1.2.1 The sensitivity methods

Several dynamic models have been proposed in order to study the sensitivity of the large scale kinetic parameters inside the *E. coli*:

The model of glycolysis and pentose phosphate pathways was investigated by stepwise internalization method and applied to 85 kinetic parameters (Chassagnole et al., 2001).

Twelve kinetic parameters were stated as the most affected parameters using global sensitivity analysis of *Sobol Methods and Monte carlo simulation* which was applied to Embden-Meyerhof and pentose phosphate pathways in addition to phosphortransferase system, bringing the number of the kinetics to 85 (Maggio et al., 2010).

The kinetic parameters were investigated by scaling each kinetic parameter individually and then quantify the changes using the Pearson correlation coefficient for V^{max} parameters only, whereby ten kinetic parameters were stated as most sensitive ones in the glycolysis and pentose phosphate pathways (Yukako et al., 2013).

The main idea behind the sensitivity analysis is to identify the model inputs that cause significant changes in the outputs and should therefore be the focus of attention if the robustness is to be increased (perhaps by further research). Therefore, the latest researchers have been working in the *E. coli* bacteria model. They investigated only two pathways which are glycolysis and Pentose Phosphate pathways by applying local or global sensitivity analysis for the kinetic parameters of that pathways either for V_{max} or K or for the both. Therefore, the number of kinetic parameters investigated in the pathways are 85 kinetics only. But in the model under study we have 194 kinetic parameters which are distributed in five pathways in addition to acetate formation. Moreover, the need for local sensitivity analysis is the simplest method to be used in order to achieve our main target.

1.2.2 The optimization methods

Several dynamic models have been proposed in order to study the behavior and to identify the importance of the large scale kinetic parameters inside the *E. coli*:

Many of such models used the central carbon metabolism, which contain PTS system, Glycolysis and Pentose-Phosphate pathways in the central carbon metabolism; they fit the time course of unbalanced metabolite concentration with analytical function by using *Simulated Annealing* for the whole K_m Kinetics only (Chassgnole et al., 2002).

Others used large scale dynamic metabolic which contain Embden Meyerhof-Paranas, Pentose-Phosphate pathways and PTS system with the acetate formation of *E. coli*. Nine parameters were estimated using the optimization technique (*GRAMS*) for the whole kinetics V_{max} and K_m (Maggio et al., 2010).

(Baker et al., 2010) optimized 4 kinetic parameters by applying four algorithms to see which algorithm are good in order to correct the kinetic parameters simulation result to be closer to the experimental data.

The Large kinetic model using Real-Coded Genetic Algorithm (RC-GA) for the optimization also uses the same model formulated by other researchers (Chassagnole et al., 2002) in *E. coli*. The target kinetic parameter is V_{max} whereas ten kinetic parameters have been often identified through the application of sensitivity analysis by increasing each parameter individually in percentage to be optimized (Yukako et al., 2013).

As explained in the previous paragraphs we concluded that, the optimization of large-scale kinetic parameters in complex models becomes difficult due to the model's behavior which requires sensitivity analysis in order to identify the most affected parameters in the model response. Moreover the use of PSO is to correct the kinetic parameters simulation result to be more close to real experimental data of the model under study.

1.3 Objective

The main objective of this research is to optimize the kinetic parameters needed for large scale of the metabolic network of *E. coli*. In order to achieve the main objective, several sub-objectives were considered as listed below:

i. To identify the most sensitive kinetic parameters in the main metabolic pathway of *E. coli* using the local sensitivity analysis technique.

ii. To optimize the kinetic parameters using the PSO algorithm.

iii. To validate the optimization result based on real experimental data.

1.4 Research scope

This study aims at the large-scale kinetic parameters issues of the metabolic network model of *E. coli* formulated by (Kadir et al., 2010) which contain Glycolysis, Pentose Phosphate, TCA cycle, Gluconeogenesis and Glycoxylate pathways, PTS system as well as Acetate Formation.

There are 194 Kinetic parameters, 23 metabolites, 28 enzymatic reactions with 10 co-factors used in this research.

The condition considered in this study was in a continuous culture with steady-state condition in the dilution rate of 0.1 and 0.2.

This study considers only the local sensitivity analysis technique of One-At-A-Time Sensitivity Measures and the minimization of errors between the simulation result and experimental data after applying the sensitivity analysis using PSO Algorithm. The analysis of One-At-A-Time sensitivity measures and Particle Swarm Optimization will be coded in MATLAB and applied to the model under study. The validation will be conducted by comparing three output results which are (Kadir et al., 2010), (Hoque et al. 2005) and our result after we got the result of PSO.

1.5 Thesis organization

Chapter 1 generally describes the introduction (Modelling and Simulation in Computational field, Sensitivity Analysis and Optimization), the problem statement, objectives, and scope of the research. Chapter 2 reviews the metabolic engineering, computational biology, dynamic modeling, Sensitivity analysis and PSO algorithm. Chapter 3 presents the framework of this study, the model description, sensitivity analysis method, optimization algorithm and validation. Chapter 4 elaborates the sensitivity analysis calculation and Particle Swarm Optimization implementation of sensitivity results, discussion of results, and validating the optimization result by comparing results from experiment with (Kadir et al., 2010) and (Hoque et al.,). The conclusions of the present research are summarized and presented in Chapter 5 with suggestions and recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents the detailed description of the concepts, classification and architecture of the metabolic engineering, computational biology, dynamic modeling, and sensitivity analysis and optimization algorithms.

2.2 Metabolic engineering

Metabolic engineering is the direct improvement of product formation or cellular properties through the modification of specific biochemical reactions, and these reactions are connected to each other to shape the metabolic pathways (Stephanopoulos et al., 1998). Metabolic pathway is a sequence of feasible and observable biochemical reaction steps connecting a specified set of input and output metabolite; the word metabolite refer to metabolism which mean is a set of chemical reaction that happen in the cell of living organisms in order to sustain life. Those reactions allow organisms to grow, reproduce, maintain their structure and respond to their environment; the collection of the pathways is called metabolic network (Schilling et al., 1999). Metabolic network is a complete set of metabolic and physical processes that determine the physiological and biochemical properties of a cell (Mathews and Van Hold, 1996).

Metabolic engineering is a abroad field, which contributes to flux measurement, understanding of flux control in vivo, engaging the chemical engineering's in biological research, which allows the direct application of the core subjects of kinetics, transports,

and thermodynamics to analysis of reactions; also contribute in the medical field such as the analysis of the function and general metabolism of tissues and whole organs in vivo; also in industrial production such as the production of new materials, gums and solvent etc (Stephanopouls et al., 1998).

One of the important problems in metabolic engineering is the production of some products from the metabolism of *E. coli*, which the product requires in details gene regulation. It was reported by (Stephanopoales et al., 1998) on how to improve the production of Lysine in corynebacterium-glutamicum and the insertion of new genes. Also, (Dellomonace 2011) reported on how fatty acids and alcohols can be catalytically converted to chemicals and fuels. Moreover the improvement of the DHAP production was produced by *E. coli* (Patnaik and Liao, 1994).

Genetic regulation occurs at genome level, controlling the expression of certain genes. This regulation affects the presence or absence of enzymes in the metabolic engineering, and it also activates or inhibits particular enzymes. In order to study the gene regulation, there is every need to do computational biology, which gives a clear picture about the development of metabolic engineering (Machado et al., 2012).

2.3 Computational biology

In the recent years, the system biology becomes very important for developing the metabolic engineering and the genetic in investigating the components of cellular networks and their interactions, or applying experimental in genome scale, or integrating computational methods with experimental data. A true understanding of genetic and metabolic function and design will crucially depend on mathematical and computational methods for analyzing biochemical systems. To this end, system biology is how to combine biological experiments with computational modeling. The system biology has two major approaches: the first one is dynamic model which offers computational tools used for analyzing, integrating and interpreting biological data and hypotheses such as *E. coli* (chassagnole et al., 2002), saccharomyces cerevisiae (Joseph et al., 1997; Sam et al., 1999) and morphogenesis (Igoshin et al., 2004); the second one is the static model which induces the formulation of new concepts and the existing

application such as theory of dynamic model, the analysis of molecular noise and robustness as well as fragility of the dynamic model.

2.4 Dynamic modeling

Dynamic modelling is a set of equation or rules specifying how the state variables change over time, as a function of the current and past values of the state variables (Kadir et al., 2010). The modeling of the dynamics of biological systems is essentially based on the modeling of a dynamic system of some bio-chemical reactions using deterministic rate laws. This has been proven to be extremely successful in both chemistry and bio-chemistry for many years. These approaches have as their core the law of mass action, which is an empirical law giving a simple relation between reaction rates and molecular component concentration at all points of future time (Provost and Bastin, 2004). Reaction rate can be used to construct mathematical models based on ordinary differential equations (ODEs) of the dynamic set of chemical reactions. ODEs are mathematical equations for un-known functions of one or of several variables that relates the values of the function itself and derivatives (how many variables of un-known function changes over time, the known function could be obtained by integrating the corresponding ODE) (Andrews and Arkin, 2006).

Currently, several methods have been proposed in order to study the dynamic behavior of the metabolic of *E. coli* such as a dynamic system computer analysis program (Wright et al., 1992), a fluxAnalyzer (Kalmatet al., 2003), Petri Nets (Koch et al., 2005), flux balance analysis (Reed and Palsson, 2003, Varma and Palsson, 1997; Kauffman et al., 2003; Edward et al., 2001), network component analysis (Liao et al., 2003); 13C-metabolic flux analysis (Arauzo-Bravo and Shimizu, 2003; Zhao and Shimizu, 2003; Matsuoka and Shimizu; Toya et al., 2010), and even for dynamic modeling (Chassagnole et al., 2002; Usuda, 2010; Kadir et al., 2010).

In order to study the dynamic model, there is a need to employ mathematical modelling, which is used to describe the underlying mechanism of a large number of processes in the natural, physical or social sciences so that mathematical techniques can assist in understanding the system. The phonetic statement is translated into an equation

called dynamic equation of the model knowledge of the initial state of a system and the dynamic equation that describes the forces of change in the system is often sufficient to forecast an observed pattern of the system (Steuer et al., 2006).

The best way to design and develop an accurate dynamic model is through simulating, sensitivity analysis and optimization algorithm methods.

Each dynamic modeling contains mass balance equation, kinetic rate equation, metabolite, enzyme, kinetic enzyme and co-factor. They are all described below:

2.4.1 Pathways

The pathways which are long chains of chemical reactions take place in the normal operation of living system. In the model of (Kadir et al., 2010), there are five pathways involved which are Glycolysis, Pentose Phosphate, TCA cycle, Gluconeogenesis, Glycoxylate pathways, PTS system and acetate formation. Each enzyme described by mass balance equation is the quantity of all species in a solution containing a particular atom; it must be equal to the amount of that atom delivered to the solution. The solution of the equation may be derived from the dynamic equation and the initial state of the system as well as a graph or table of values of the solution may then be compared with the observed pattern of nature; i.e. to what extent the solution of equation matches the pattern is a measure of the validity of the mathematical model. The metabolite concentration rate of the changes in this metabolic network is given by the following equation

$$\frac{\mathrm{d}C_i}{\mathrm{dt}} = \sum_j R_{ij}\mathrm{v} - \mu\mathrm{C_i} \qquad (2.1)$$

Where, C_i is the concentration of metabolite i , R_{ij} is the stoichiometric coefficient of metabolite i in the reaction j , v_j is the rate of the reaction j and μC_i is the growth rate on the dilution effect.

2.4.2 Metabolite and co-metabolites

The metabolite means any substance or product produced by metabolism or by a metabolic process. For example the metabolite of F6P which is called Fructose-6-phosphate. Co-metabolites are organic molecules that are required by certain enzymes to carry out catalysis, they bind to the active site of the enzyme and participate in catalysis but are not considered as substrate of the reaction, for example CoA is a coenzyme, notable for its role in the synthesis and oxidation of fatty acids, and the oxidation of pyruvate in the citric acid cycle (Chassagnole et al., 2002).

2.4.3 Enzyme

Enzymes are natural proteins produced in tiny quantities by all living organisms, functioning as highly selective biochemical catalysts in converting one molecule into another. For example, the enzyme of PTS is called phosphotransferase system, which converts the extra glycolysis to glucose-6-phoshate (G6P).

2.4.4 Kinetics enzyme

Enzyme-kinetics is central to every biological process that ever has been or will be studied and is the basis for a great number of essays that are routinely undertaken in every research laboratory (Irwin, 1993). The important parameters in a kinetic reaction are Vmax, Km and Vmax/Km. Vmax: Maximum velocity of the reaction aka how the enzyme behaves when the substrate is in abundance. Km: The affinity of the enzyme for the substrate the lower the value, the better the affinity. Vmax/Km = 1/2 Vmax aka how the enzyme behaves at low substrate concentrations (Cleland, 1963).

2.4.5 Kinetic rate equation

It means the change in concentration of a reactant in a given period of time; for example, the kinetic rate equation of PTS is described by the equation below:

$$v_{PTS} = \frac{v_{PTS}^{max}[GLc^{ex}]\frac{[PEP]}{PYR}}{\left(K_{a1}+K_{a2}\frac{[PEP]}{[PYR]}+K_{a3}[GLc^{ex}]+[GLc^{ex}]\frac{[PEP]}{[PYR]}\right)\left(1+\frac{[G6P]^{n}G6P}{K_{G6P}}\right)} \quad (2.2)$$

Where PTS is phosphotransferase system, v_{PTS}^{max} is Maximum velocity of the PTS reaction, GLc^{ex} is extracellular glucose, PEP is Phosphoenolpyruvate, PYR is Pyruvate, $G6P$ is Glucose-6-phosphate, K_{a1} is the affinity of the enzyme for the substrate (Chassagnole et al., 2002).

2.5 Sensitivity analysis

The idea of the sensitivity analysis as a technique is to identify how different input variables will affect a model result under a given set of assumption (Fasso and Perri, 2002). In order to optimize large-scale kinetic parameters according to the dynamic model of *E. coli*, there is need to keep the sensitivity analysis at a steady state condition, in which the concept of steady state is a mathematical idealization, which plays an important role in kinetic modeling (Heinrich and Sehuster, 1996).

There are different methods applied to metabolic model such as Metabolic Control Analysis to describe how the sensible properties of the ingredients, enzymes and which metabolic variables respond in a metabolic pathway (Diana et al., 2002).

Other researchers applied Metabolic Control Designs, which address the inverse problem of getting the sensitivity properties of the components of enzymes which are required for the system to show a pre-established pattern of responses (Simon et al., 1996).

Even Sobol' method (Sobol, 1990) applied to the model of *E. coli* (Maggioa et al., 2010).

Recently, Stepwise Internalization method was applied to *E. coli* model (Chassagnole et al., 2002). Monte Carlo simulation with Sobol method and variance based techniques was used to study the glycolysis, phosphotransferase system, and pentose-phosphate pathways. In order to identify the kinetic parameters sensitivity, the

time profile was indicated there and eleven kinetic parameters were affected in the model response. The kinetics including maximum reaction rate, inhibition and half saturation constant have been addressed to formulate a parameter estimation problem (Maggioa et al., 2010); moreover scaling all the kinetics concentration individually from 1% to 100% was done before calculating the changes, and ten kinetic parameters were identified as the most affected in the model response (Yukako et al., 2013).

The latest researchers applied different sensitivity analysis methods to study the sensitivity in the metabolic network of *E. coli* to achieve the optimal goal for glycolysis and PP pathways (Yukako et al., 2013; Mauch et al., 1997; Noacka et al., 2008); Embden-Meyerof-Paranas, PP pathways and PTS system (Maggioa et al., 2010).

In order to optimize large scale kinetic parameters which has become difficult in dynamic modeling, sensitivity analysis has to be employed so that kinetic parameters will be reduced. There is need for the sensitivity analysis method to be applied when using an optimization algorithm for large-scale metabolic model of *E. coli* due to the model complexity. To this end, all the researchers applied different methods in the model of *E. coli*, but they did not apply for the model formulated by (Kadir et al., 2010).

At this juncture, we will attempt a brief description of some sensitivity analysis and optimization methods that were applied to *E. coli* model below.

2.5.1 One at a time sensitivity measures

The core idea behind this model is to conduct a sensitivity analysis for one parameter at a time in percentage, while the others parameters are fixed at their nominal values then quantify the changes over the model response using suitable mathematics formula (Herve et al., 2006).

2.5.2 Variance based sensitivity analysis

It has been reported that, the variance based sensitivity analysis is a form of global sensitivity analysis working within a probabilistic framework. It decomposes

the variance of the output of the model or system into fractions which can be attributed to inputs or sets of inputs. For example, given a model with two inputs and one output, one might find out that 70% of the output variance is caused by the variance in the first input, 20% by the variance in the second, and 10% due to interactions between the two. These percentages are directly interpreted as measures of sensitivity (Tomasz, 2013).

2.6 Optimization

Optimization is not a new concept in biology. It helps to predict new designs which may evolve. In line with this, chemical processes are affected by numerous parameters such as reaction rates and kinetic constant; therefore, the metabolic pathways and regulation analysis which takes place in a cell result as the evolutionary optimization process (Yukako et al., 2013). Optimization methods had significant impact in biological systems as it helps researchers to study, investigate and develop the biological models. Some researchers have applied Differential Evolution Algorithm (DE) to simulate glycolysis pathway in yeast and to estimate the optimal kinetic (Chuii et al., 2012); a Mixed Integer Non-liner Problem (MINLP) is used to calculate which enzyme levels should be modulated to obtain stable optimization of large scale models (Nikolaev, 2010); a decoupling method and minimizing concentration errors is used to estimate the metabolites through the minimizing of slope errors (Gengjie et al., 2011); branch-and-bound principle is used to find the best set of model parameter (Pradeep et al., 2006); a popularization of PSO is used to estimate the parameters from the central carbon metabolism of *E. coli* (Mingshou et al., 2009); degree probability distribution is used to enrich the parameter which needs optimization (Joshua and Franz, 2010); NMR(Nuclear Magnetic Resonance) and GC-MS(Gas Chromatography-Mass Spectroscopy) is used to estimate the flux distribution and to simulate the glycolysis and pentose-phosphate pathways in *E. coli* (Hoque et al., 2011); A Spatial Branch-and-Bond algorithm is used to address the global optimization of metabolic network (Pozo et al., 2010); Particle Swarm Evolutionary Optimization (PSEO) to estimate the parameters in the complex and nonlinear biological models and to reduce the computational times (Abdullah et al., 2013); Improved Differential Evolution Algorithm to (IDE) is used to find a solution for the existence of noisy data and sees to the perplexity of the model (Chong et al., 2012); a nonlinear optimization problem is used to analyze the sensitivity

and the stability properties of kinetic representation of the central carbon metabolism of E. coli at optimal enzyme levels (Francisco et al., 2006); Differential Evolution Particle Swarm Optimization (DEPSO) is used to demonstrate the nonlinear dynamics of gene networks and revealing genes regulatory interactions (Rui et al., 2007); a deterministic outer-approximation reformulates the set of ordinary deferential equation into an equivalent set of algebraic equation (Miro et al., 2012); mass balance and the relevant flux configuration optimizes the global growth of the system and reduces the empirical statistics of flux in *E. coli* (Martelli et al., 2009); a genetic algorithm (GA) optimize mixtures of 13C-labeled glucose and glutamine (Jason et al., 2012). To this end, all the researchers applied or formulated different optimization methods to solve biological systems with different problems. Some of the most popular recent algorithms and their uses in the *E. coli* are stated below.

2.6.1 Differential evolution (DE) algorithm

This is a method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality (Storn and K. V, 1995).

DE optimizes a problem by maintaining a population of candidate solutions and creates new candidate solutions by combining the existing ones according to its simple formulae, but keeping whichever candidate solution has the best score or fitness on the optimization problem at hand. In this way, the optimization problem is treated as a black box that merely provides a measure of quality when given a candidate solution and the gradient is therefore not needed.

A basic variant of the DE algorithm works by having a population of candidate solutions (called agents). These agents are moved around in the search-space by using simple mathematical formulae that combine the positions of existing agents from the population. If the new position of an agent is improved, it is accepted and forms part of the population, otherwise the new position is simply discarded. The process is repeated and by doing so it is hoped, but not guaranteed, that a satisfactory solution will eventually be discovered.

Formally, let $f: R^n \rightarrow R$ be the cost function which must be minimized or fitness function which must be maximized. The function takes a candidate solution as argument in the form of a vector of real numbers and produces a real number as an output which indicates the fitness of the given candidate solution. The gradient of f is not known. The goal is to find a solution m for which $f(m) \leq f(p)$ for all p in the search-space, which would mean m is the global minimum. Maximization can be performed by considering the function $h := -f$ instead.

Let $X \in R^n$ designate a candidate solution (agent) in the population. The basic DE algorithm can then be described as follows:

Algorithm 2.1

1. **Initialize all agents X with random positions in the search-space.**
2. **Until a termination criterion is met (e.g. number of iterations performed, or adequate fitness reached), repeat the following**
3. **For each agent X in the population do:**
4. **Pick three agents a, b, and c from the population at random, they must be distinct from each other as well as from agent X.**
5. **Pick a random index $I \in \{1, \ldots, n\}$ (n being the dimensionality of the problem to be optimized).**
6. **Compute the agent's potentially new position $y = \{y_1, \ldots, y_n\}$ as follows:**
7. **For each i, pick a uniformly distributed number $r_i \in (0, 1)$**
8. **If $r_i < CR$ or $i = R$ then set $y_i = a_i + F * (b_i - c_i)$ otherwise set $y_i = x_i$) (In essence, the new position is the outcome of binary crossover of agent x with intermediate agent $z = a + F * (b - c)$.**
9. **I f $f(Y) < f(X)$ then replace the agent in the population with the improved candidate solution, that is, replace X with Y in the population.**
10. **Pick the agent from the population that has the highest fitness or lowest cost and return it as the best found candidate solution.**

Note that $F \in [0,2]$ is called the differential weight and $CR \in [0,1]$ is called the crossover probability, and both of these parameters are selectable by the practitioner along with the population size≥ 4 .

In line with this, (Chuii et al., 2012) improved the (DE) Differential Evolution Algorithm for the purpose of to developing, executing and estimating the relevant parameter, for the metabolic pathway data in order to simulate glycolysis pathway in

yeast, and they named this new algorithm (IDE) Improved Differential Evolution Algorithm. This algorithm is a crossbred of (DE) and (KF) Kalman Filter. By using the fitness function of IDE algorithm below, estimation of the relevant parameters for metabolic pathway data in order to simulate glycolysis pathway for yeast can be achieved:

$$j = \sum_{i=1}^{N} |f(X, X0, \theta 0) - f(Y, X0, \theta)|^2 \quad (2.3)$$

The fitness function is applied to evaluate the fitness of each individual parameter. X represents the state vector of the measurement system, Y represents the state vector of the simulated system, $\theta 0$ represents a set of original parameters, θ represents a set of estimated parameters $X0$ represents the initial state, $N \equiv$ the ending index, $i \equiv$ the index variable.

2.6.2 Mixed integer non-liner problem (MINLP)

Mixed Integer Nonlinear Programming (MINLP) problems contain nonlinear expressions and integer variables. Mixed integer Nonlinear programming problems are in general more difficult to solve than mixed integer programming problems and nonlinear programming problems (Wong and Tsai 2011).

A general MINLP can be formulated as:

min f(x)
such that g(x) <= 0
l <= x_i <= u_i
x_i in Z for all $i \in$ I

Where f: $R^n \rightarrow$ R and g: $R^n \rightarrow R^m$ are twice continuously differentiable functions, l, u $\in R^n$ determine lower and upper bounds of the variables, and I $\subseteq$ {1,...,n} denotes the set of variables with integral requirements.

In line with this, (Nikolaev, 2010) present the Mixed-Integer Nonlinear Programming (MINLP) formulation to automatically calculate which enzyme levels

should be modulated and which enzyme regulatory structures should be altered in order to achieve the given optimization goal using non-linear kinetic models of the relevant cellular system. He used the glucose uptake through the phosphotransferase system and serine biosynthesis to obtain the stable optimization of large-scale kinetic models of the cellular system.

2.6.3 Nonlinear programming (NLP)

In mathematics, nonlinear programming (NLP) is the process of solving an optimization problem defined by a system of equalities and inequalities, collectively termed constraints, over a set of unknown real variables, along with an objective function to be maximized or minimized, where some of the constraints or the objective function are nonlinear (Kuha and Tucker, 1950).

The problem can be stated simply as:

$max_{x \in X}\, f(x)$ to maximize function product throughput

or

$min_{x \in X}\, f(x)$ to minimize some function such as a cost function

where

$f: R^n \rightarrow R$

$x \in R^n$

s.t. (subject to)

$h_i(x) = 0, i \in I = \{1, \dots, p\}$

$g_i(x) \leq 0, i \in E = \{1, \dots, m\}$

In line with this, (Pradeep et al., 2006) utilizes from the branch-and-bound principle to find the best set of model parameter by creating tight upper and lower for the objective function value of the global solution. The lower bound is used in solving the convex relaxation of the nonconvex NLP problem.

2.6.4 Simulated annealing

Simulated annealing (SA) is a generic probabilistic met-heuristic for the global optimization problem of locating a good approximation to the global optimum of a given function in a large search space. It is often used when the search space is discrete. The name and inspiration come from annealing in metallurgy, a technique involving heating and controlled cooling of a material to increase the size of its crystals and reduce their defects. Both are attributes of the material that depend on its thermodynamic free energy. Heating and cooling the material affects both the temperature and the thermodynamic free energy. While the same amount of cooling brings the same amount of decrease in temperature, it also brings a bigger or smaller decrease in the thermodynamic free energy depending on the rate its occurrence. A slower rate produces a bigger decrease. This notion of slow cooling is implemented in the Simulated Annealing algorithm as a slow decrease in the probability of accepting worse solutions as it explores the solution space. Accepting worse solutions is a fundamental property of met-heuristics because it allows for a more extensive search for the optimal solution (Metropolic et al., 1953).

Algorithm 2.2

1 let $x = x_0$
2 for $k = 0$ through k_{max} (exclusive):
3 T $\leftarrow$ $teperature\left(\frac{k}{k_{max}}\right)$
4 Pick a random neighbour, x_{new} $\leftarrow$ neighbour (x)
5 If $f(E(x), E(x_{new}), T))$ $random\ (0, 1), move\ to\ the\ new\ state$:
6 $x \leftarrow x_{new}$
7 $output: the\ final\ state\ x$

The Algorithm 2.2 above presents the simulated annealing heuristic as described above. It starts from a state x_0 and continues to either a maximum of k_{max} steps or until a state with an energy of e_{min} or less is found. In the process, the call neighbour (*s*) should generate a randomly chosen neighbor of a given state x; the call random (0, 1) should pick and return a value in the range [0, 1), uniformly at random. The annealing schedule is defined by the call temperature (*r*), which should yield the temperature to use, given the fraction *r* of the time budget that has been expended so far.

In line with this, (Chassagnole et al., 2002) applied Simulated Annealing to estimate the kinetics in phosphotransferase system with the reaction of glycolysis and pentose phosphate pathway.

2.6.5 Genetic algorithm

In the computer science field of artificial intelligence, a genetic algorithm (GA) is a search heuristic that mimics the process of natural selection. This heuristic (also sometimes called a metaheuristic) is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover (John, 1960).

In line with this, (Yukako et al., 2013) applied a real-coded genetic algorithm using the objective function of RC-GA to reduce the errors between the actual model values obtained from the simulated result of (Chassagnole et al., 2002) and the estimated values provided by the following equation:

$$F = \sum_{i=1}^{R} \sum_{j=1}^{T} \left| \frac{x_{ij} - y_{ij}}{y_{ij}} \right| \quad (2.4)$$

Where R is the number of metabolite, T is the number of sample point, x_{ij} is the estimated concentration of metabolite i at the jth sampling, and y_{ij} is the true concentration of metabolite i at the jth sampling point.

2.6.6 Control vector parameterization

Control vector parameterization, also known as direct sequential method, is one of the direct optimization methods for solving optimal control problems. The basic idea of direct optimization method is to discretize the control problem, and then apply nonlinear programming (NLP) techniques to the resulting finite-dimensional optimization problem (Saziye et al., 2009).

In line with this, (Maggio et al., 2010) formulated a parameter estimation problem in an equation oriented control vector parameterization environment with a maximum likelihood objective function in order to determine the values of kinetic parameters and variance model parameters by this equation below:

$$\emptyset = \frac{N}{2}1n(2\pi) + \frac{1}{2}min_p \sum_{i=1}^{NM}\sum_{j=1}^{NT}\left[1n\left(\sigma_{ij}^2\right) + \frac{(C_{ij}^M - C_{ij})^2}{\sigma_{ij}^2}\right] \quad (2.5)$$

Where the summation over NM measured state variables (C_{ij}, metabolite concentration) and NT data points for each measured variable; σ_{ij} is the variance of the jth measrment of variable i, which is determined by the measured variable variance model whose elements correspond to variances of the measured variables. N is the total number of measurements. Vector p corresponds to estimated parameters.

2.6.7 PSO algorithm

This method was proposed in 1995 by Kennedy and Eberhart. It is inspired by social behavior and movement dynamics of insects, birds and fishes. The swarm is typically modeled by particles that have a position and a velocity in multidimensional space. It is used to find the best global position and its information of the best neighbor (Russel and James 1995). (Baker et al., 2010) estimated 4 parameters based on time course data by this objective function below:

$$f = \int_0^t (y_{mes}(t) - y_{pre}(\theta, t))^T W(t)(y_{mes}(t) - y_{pre}(\theta, t))dt \quad (2.6)$$

Where f is the cost function, θ is the vector of parameters, $y_{mes}(t)$is the vector computed values of the state at time t, and $\mathrm{W(t)}$ is the weighting matrix.

2.6.8 Related work

In the systems biology, the computational models are usually described by a lack of reliable parameter values; its true in kinetic metabolic models and the optimization of the models are considered to be as a parameter estimation which minimizes the errors

between the simulated model data and the experimental data. Recently, different methods were proposed in order to achieve large scale optimization in *E. coli* by applying Simulated Annealing which helps to consider 85 kinetics in phosphotransferase system with the reaction of glycolysis and pentose phosphate pathway (Chassagnole et al., 2002), or formulating algorithms like A maximum LikeLihood problem to glycolysis pathway, phosphotransferase system and pentose phosphate pathway (Maggio et al., 2013); or (Yukako et al., 2013) investigating the Real-Coded Genetic Algorithm efficiency to estimate the kinetic parameters and sensitivity analysis using the model formulated by (Chassagnole et al., 2002).

Unfortunately, the model under study contain five pathways in addition to two systems that has been built through the simulation, but has not been investigated, though has a large kinetic parameters than the other models (Chassagnole et al., 2002; Maggio et al., 2013; Yukako et al.,2013). Most of the values are not fully correct because of the simulation, which this model needs in order to be fitted closely into real experimental data using suitable optimization algorithm that minimizes the errors between the model experimental and the experimental data. Recently, it has been reported that the Particle Swarm Optimization is very good in estimation than the most popular optimization algorithms such as Evolutionary Computational, Evolutionary Programing and Genetic Algorithm which are often applied in the upper part of (Baker et al., 2010).

Based on (Baker et al., 2010) they stated that PSO are more efficient in correcting the kinetic parameters after compared to others three algorithms, the use of PSO in this study is to correct the kinetic parameters of the model under study.

2.7 Summary

Metabolic engineering component involve analysis and synthesis in the area of traditional fields of engineering. This chapter first introduces the concept of metabolic engineering, computational biology, dynamic modeling, sensitivity analysis and optimization algorithm and the use of objects of kinetic model of *E. coli*.

The methods of one at a time sensitivity measures and variance based sensitivity analysis are described in details, which these methods has various application in many fields, including engineering and economics etc. The statistical analysis of those methods is used to cover all stages of an investigation for the kinetic parameters in this study.

To this, there are many local and global sensitivity analysis methods applied to the kinetic parameters, but still that methods are not applied to many pathways.

In addition to the optimization part explanation, there are some methods described in details including their solution to metabolic models and kinetic parameters.

However, the PSO algorithm was used in addition to Evolutionary Programming, Simulating Annealing and Genetic Algorithm to insure the performance for each algorithm in the kinetic estimation which the kinetics numbers are 4. The best result was found by PSO among the others algorithms. To this, these researches intend to optimize more than 4 kinetics using PSO algorithm.

Moreover, the sensitivity analysis methods and optimization algorithms are an important in the metabolic model of *E. coli* because these methods help to assess the development and exploration of kinetic models.

Furthermore, the overview of sensitivity analysis and optimization algorithm is explained in details, as well as the methods that were used to solve the large scale kinetic parameters.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the methodology used in this research which contains five sections. The first section is about the framework and it described the following in details; the pathways, metabolites, enzyme and kinetic equation, and kinetic parameters. The second section is about the sensitivity analysis method and their application in this research. The third section is about the Particle Swarm Optimization (PSO) method which was applied to correct the kinetic parameters of the model under study. The last section is about validation, which describes how to prove that the sensitivity analysis and optimization algorithm methods are highly sufficient in order to achieve our goals in this research.

3.1.1 The condition used in the sensitivity and optimization methods

Based on the method of one at a time sensitivity measures we define the degree of sensitivity to measure the sensitivity among kinetics, there are 194 kinetics involved, for 0.1 dilution rate each kinetic is increased by 10%, 20%, 40% and 80% then quantify the changes using the Highest Mean for each kinetic parameters on the model response it will be described in the sensitivity analysis method. Moreover, in dilution rate 0.2 each kinetic parameter increased by 10% and 20% then quantifies the changes using the Highest Variance.

The PSO factors used in this study in order to reach our objectives are fitness function, dimension problem, population size, upper and lower values, c1&c2 they

represent the exploitation coefficient, r1 & r2 are random numbers between 0 and 1 with weight constant ω which are explained PSO part. The metabolite concentration of (Hoque et al., 2005) are Fructose 1,6-biophosphate *FDP* (0.67 mM), Phosphoenol-pyruvate *PEP* (1.04 mM), Isocitrate *ICIT* (0.21 mM) and 2-Keto-D-gluconate *2KG* (0.134 mM) are used in PSO algorithm execution in order to fit the optimization result closely with that experimental data picked-up from (Hoque et al., 2005) by test the optimization result in (Kadir et al., 2010) model.

3.2 Framework of the research

In order to achieve large-scale kinetic parameters optimization of metabolic network of *E. coli*, this framework in Figure 3.1 was employed to describe the 4th phase in this chapter for the purpose of solving large scale kinetic parameters.

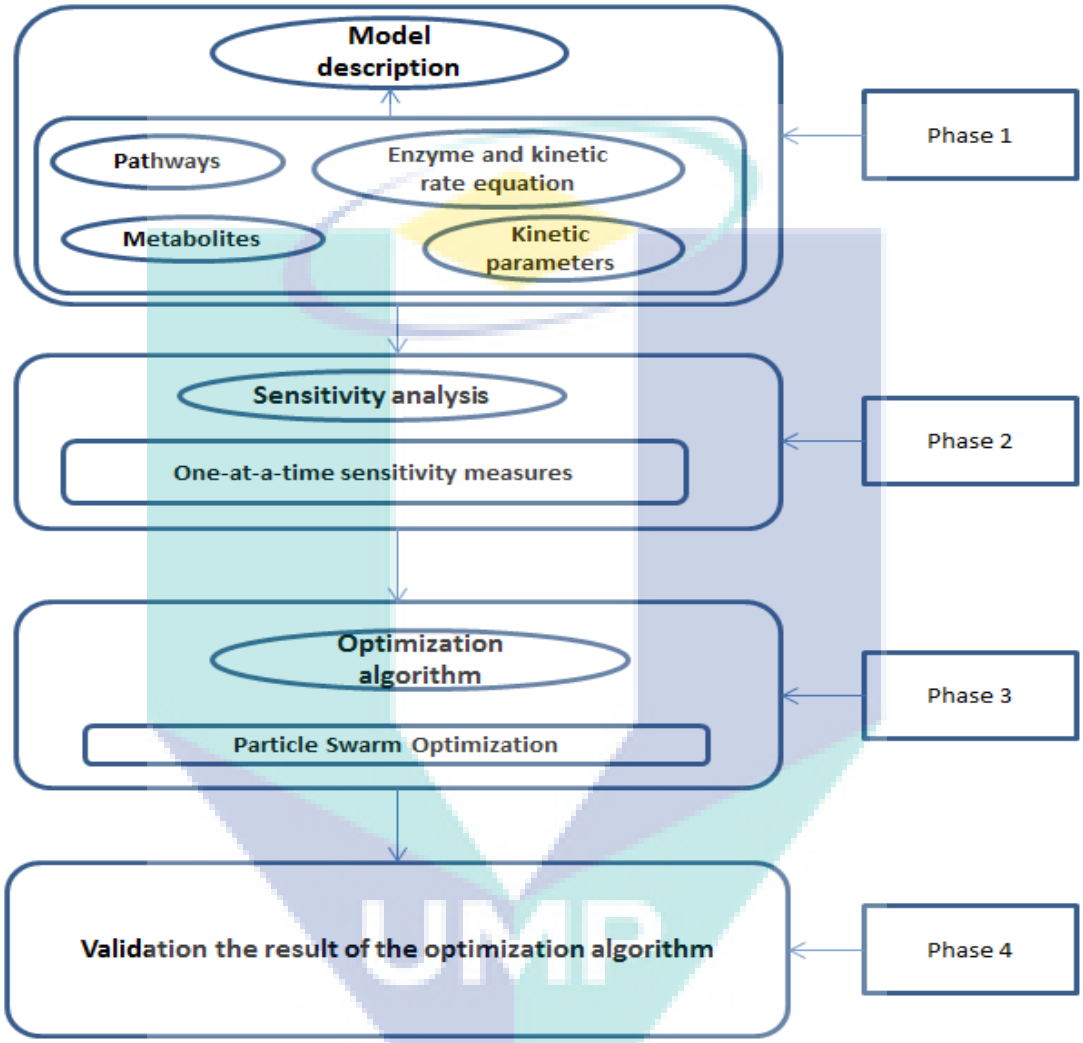
The first phase is describing the model that used in this study which contains pathways, conditions, equations, kinetic values, enzymes and metabolite concentrations.

The second phase is the method of sensitivity analysis technique that we are going to use for the analyization of the model under study, and it contains the implementation of the sensitivity analysis algorithm in dilution rate of 0.1 and 0.2 using One-At-A-Time Sensitivity Measures.

The third phase is the optimization algorithm of PSO which is proposed to optimize the sensitivity analysis result and their implementation to the sensitivity analysis result of 0.1 dilution rates in order to achieve optimization for large-scale kinetic parameters.

The validation is proposed in the fourth phase by replacing the kinetic sensitivity analysis result with optimum values founded by PSO then run the simulation if the output is close to real experimental data we accept the optimization result if not it will be repeated till our result got close to experimental data. The criteria to prove our result is more close to experimental data than the model output result under study by measure

the percentage errors between our result, the model under study result with real experimental data.



Figuer 3.1: Framework of the study

3.3 Model description

The metabolic network model under study was used as a case study in order to achieve a large-scale kinetic parameters optimization of the metabolic network of *E. coli*. Identification of the pathways, metabolites, enzyme and kinetic rate equations, and kinetic parameters involves in this model has to be done so as to realize the goal.

3.3.1 Pathways

A pathway is a long chain of chemical reaction that happens normally in living systems. In each pathway there is a components involved, which are metabolites, enzymes and co- metabolites. In the model of (Kadir et al., 2010) under study, five pathways are involved and described in Figure 3.2 with the abbreviation of metabolites, enzymes and co-factors and how the conversions happened inside the model for the metabolites and the enzymes which described by this symbol → this symbol describe the conversions in one direction, in two direction described by this symbol ↕, the co-factors described by this symbol -------.

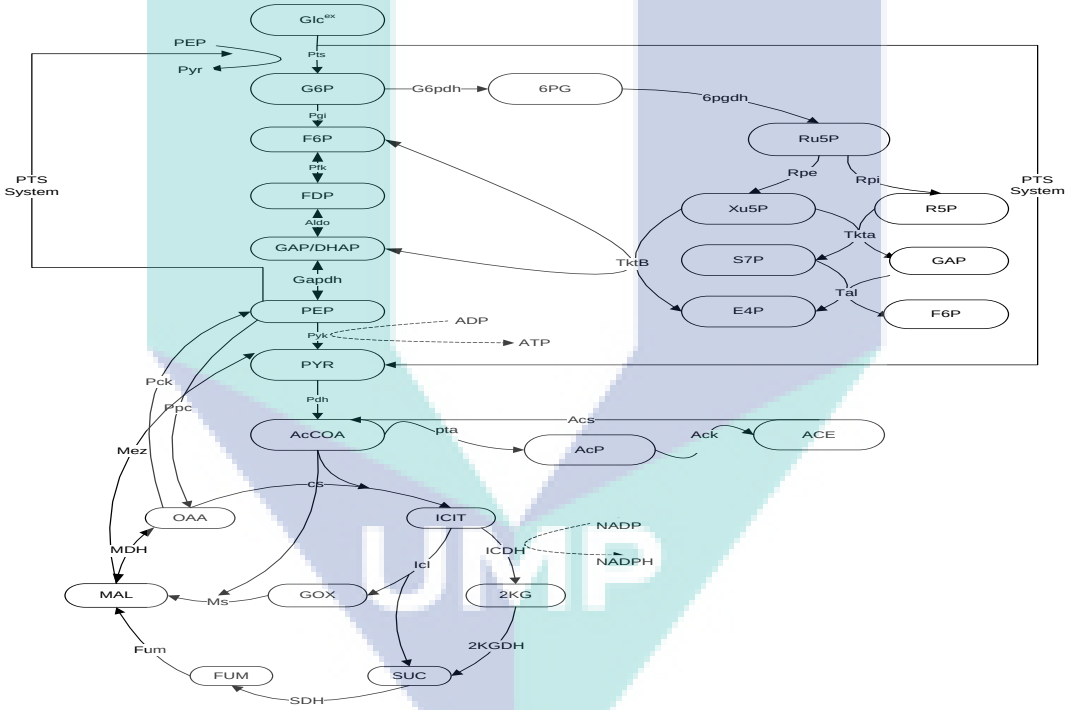


Figure 3.2: Metabolic pathway

The first pathway is Glycolysis has 7 metabolites which started by ($\boldsymbol{Glc^{ex}}$) Glucose, (**G6P)** Glucose 6-phosphate, (**F6P)** Fructose 6-phosphate, (**FDP)** Fructose 1,6-bisphosphate, (**GAP)** Glyceraldehyde 3-phosphate, (**DHAP)** Dihydroxyacetone phosphate, (**PEP)** Phosphoenolpyruvate, and (**PYR)** Pyruvate; and 6 enzymes started by (**Pts)** Phosphotransferase system, (**Pgi)** Phosphoglucose isomerase / Glucosephosphate isomerase, (**Pfk)** Phosphofructokinase-1, Aldo Aldolase, (**GAPDH)** Glyceraldehyde 3-phosphate dehydrogenase, and (**Pyk)** Pyruvate kinase in addition to **PTS**

phosphotransferase system and the co-factor of **ADP** Adenosine diphosphate, **ATP** called Adenosine-5-triphosphte.

The second pathway is Pentose Phosphate pathways which consists of 6 metabolites started by (**6PG)** 6-Phosphogluconolactone, (**Ru5P)** Ribose 5-phosphate, (**Xu5P)** Xylulose 5-phosphate, (**R5P)** Ribulose 5-phosphate, (**S7P)** Sedoheptulose 7-phosphate and (**E4P)** Erythrose 4-phosphate and 7 enzymes started by **(G6pdh)** Glucose-6phosphate dehydrogenase, (**6Pgdh)** 6Phsophogluconate dehydrogenase, (**Rpi)** Ribulose 5phosphate 3-isomerase, (**Rpe)** called Ribulose phosphate 3epimerase, (**Tkta)** TransketolaseI, (**Tktb)** TransketolaseII and (**Tal)** Transaldolase.

The third pathway is TCA cycle which consist of 6 metabolites started by (**ICIT)** Isocitrate, (**2KG)** 2-Keto-Dgluconate, (**SUC)** Succinate, (**FUM)** Fumarate, (**MAL)** Malate and (**OAA)** Oxaloacetate and 6 enzymes started by (**cs)** Citrate synthase, (**ICDH)** Isocitrate dehydrogenase, (**2KGDH)** 2-Keto-D-gluconate Dehydrogenase, (**SDH)** Succinate dehydrogenase, (**Fum)** Fumarase and (**MDH)** Malate dehydrogenase in addition to co-factors of (**NADP/NADPH)** Nicotinamide adenine dinucleotide phosphate).

The fourth pathway is Gluconeogenesis which consists of 4 metabolites started by (**PEP)** Phosphoenolpyruvate, and (**PYR)** Pyruvate, (**MAL)** Malate (**OAA)** called Oxaloacetate; and 3 enzymes started by (**Mez)** Malic enzyme, (**Pck)** Phosphoenolpyruvate carboxykinase and (**Ppc)** PEP carboxylase).

The fifth pathway is Glycoxylate which consist of 4 metabolites started by (**Gox)** Glyxoylate, (**ICIT)** Isocitrate, (**SUC)** Succinate and (**MAL)** called Malate; and 2 enzymes started by (**ICL)** Isocitrate lyase and (**Ms)** Malate synthase. In addition to acetate formation which consists of 3 metabolites started by (**AcCOA)** Acetyl-CoA, (**AcP)** Acetylphosphate and (**ACE)** Acetate; and 3 enzymes started by (**Pdh)** Pyruvate dehydrogenase, (**Acs**) Acetylcoenzyme A synthetase, (**Pta)** Phosphotransacetylase and (**Ack)** Acetate kinase.

3.3.2 Metabolites and Co-Metabolites

Metabolites are organic compounds that are not directly involved in the normal growth, development or reproduction of organisms. There are 23 metabolites involved in this model and they are stated in Table 3.1 with their initial concentration simulation values picked up from (Kadir et al., 2010) in order to fit this result closely to experimental data whether all or some of them.

Table 3.1: Metabolites values

Metabolites	**Metabolites name**	**Values *mM***
GLCex	Extra Glucose	0.12203
G6P	Glucose-6-phosphate	0.12989
F6P	Fructose-6-phosphate	0.021457
FDP	Fructose 1,6-bisphosphate	1.5186
GAPDHAP	Glyceraldehyde 3-phosphate Dihydroxyacetone phosphate	0.31487
PEP	Phosphoenolpyruvate	1.5076
PYR	Pyruvate	2.8279
ACCOA	Acetyl-CoA	1.0021
ICIT	Isocitrate	0.22057
2KG	2-Keto-Dgluconate	5.3784
SUC	Succinate	0.57079
FUM	Fumarate	0.35556
MALATE	Malate	0.14256
OAA	Oxaloacetate	0.029714
GOX	Glyoxylate	0.34542
ACP	Acetylphosphate	2.0035
ACETATE	Acetate	0.0034493
6PG	6-Phosphogluconolactone	0.017854
RU5P	Ribose 5-phosphate	0.021398
R5P	Ribulose 5-phosphate	0.076388

XU5P	Xylulose 5-phosphate	0.026516
S7P	Sedoheptulose 7-phosphate	0.0047299
E4P	Erythrose 4-phosphate	0.027837

Co-Metabolites are organic molecules that required by certain enzymes to carry out catalysis. The Table 3.2 below describes the Co-Metabolites involved in this study and their values with their references which are 10 Co-Metabolites.

Table 3.2: The Co-Metabolites values

Co-Factors	Co- Names	Values *mM*	Reference
NAD	Nicotinamide adenine dinucleotide	1.47	Chassagnole et al., 2002
NADH	Nicotinamide adenine dinucleotide	0.1	Chassagnole et al., 2002
NADP	Nicotinamide adenine dinucleotide phosphate	0.195	Chassagnole et al., 2002
NADPH	Nicotinamide adenine dinucleotide phosphate	0.062	Chassagnole et al., 2002
COA	Coenzyme A	0.001	Chassagnole et al., 2002
ADP	Adenosine diphosphate	0.595	Chassagnole et al., 2002
ATP	Adenosine-5-triphosphte	4.27	Chassagnole et al., 2002
AMP	Dihydroxyacetone phosphate	0.955	Chassagnole et al., 2002

H	Histidine	0.001	Chassagnole et al., 2002
P	Phosphate	10	Hoefnagel et al., 2002

3.3.3 Kinetic rate equations

There are 28 enzymes involved in the model of (Kadir et al., 2010). These enzymes and their values are shown in Table 3.3. The enzymes are responsible for thousands of metabolites processes that sustain life; the enzymes are mathematically described by kinetic rate equations and these equations are described in Table 3.3.

All the kinetic rate equations and dynamic equations (*mass balance*) involved in this study are based on the pathways of (Kadir et al., 2010) which the kinetic rate equation she used are picked up from literature article as she stated in here article. In order to identify the kinetic reaction rate we should explain the rate equation and reaction rate. Which in a chemical the reaction is described by the rate equation that connect the reaction rate with the concentration of reactants and constant parameters through specific equation and the reaction rate means how fast or slow the reactant or product in particular reaction. The kinetic rate equations for that mode under study are stated bellow in Table 3.4.

Table 3.3: Kinetic rate equations

Reaction's	**Kinetic equation**
Cell growth	$\begin{cases} \mu_m \left(1 - \frac{[X]}{X_m}\right)\left(\frac{[GLc^{ex}]}{K_s + [GLc^{ex}]}\right) k_{ATP} v_{ATP}(.), ([GLc^{ex}] > 0) \\ \frac{\mu_{mA}[Ace^{ex}]}{K_{sA} + [Ace^{ex}]}\ k_{ATP} v_{ATP}(.), ([GLc^{ex}] \leq 1\ and [Ace^{ex}] > 0) \end{cases}$
PTS	$\frac{v_{PTS}^{max}[GLc^{ex}]\frac{[PEP]}{PYR}}{\left(K_{a1} + K_{a2}\frac{[PEP]}{[PYR]} + K_{a3}[GLc^{ex}] + [GLc^{ex}]\frac{[PEP]}{[PYR]}\right)\left(1 + \frac{[G6P]^n G6P}{K_{G6P}}\right)}$

PGI	$\frac{v_{PGI}^{max}\left([G6P]-\frac{[F6P]}{Keq}\right)}{K_{G6P}\left(1+\frac{[F6P]}{K_{F6P}\left(1+\frac{[F6P]}{K_{6pginh}^{F6P}}\right)}+\frac{[6PG]}{K_{6pginh}^{G6P}}\right)+G6P}$
PFK	$\frac{v_{PFK}^{max}K_{ATP}[F6P]}{K_{(ATP,ADP)}\left([F6P]+K_z^{F6P}\frac{K_{b\left(ADP,AMP\right)}+\frac{[PEP]}{K_{PEP}}}{K_{a(ADP,AMP)}}\right)\left(1+\frac{L_{pfk}}{\left(1+[F6P]\left(\frac{K_{a(ADP,AMP)}}{K_S^{F6P}\left(K_{b(ADP,AMP)}+\frac{[PEP]}{K_{PEP}}\right)}\right)\right)^{n_{PFK}}}\right)}$
Aldo	$\frac{v_{ALDO}^{max}\left([FDP]-\frac{[DHAP][GAP]}{K_{eq}}\right)}{\left(K_{FDP}+[FDP]+\frac{K_{GAP}[DAHP]}{[K_{eq}V_{bif}]}+\frac{K_{DHAP}[GAP]}{[K_{eq}V_{bif}]}+\frac{[FDP][GAP]}{K_{inh}^{PEP}}+\frac{[DHAP][GAP]}{K_{eq}V_{bif}}\right)}$
GAPDH	$\frac{v_{GAPDH}^{max}\left([GAP]-\frac{[PEP][NADH]}{K_{eq}[NAD]}\right)}{\left(K_{GAP}\left(1+\frac{[PEP]}{K_{PGP}}\right)+[GAP]\right)\left(\frac{K_{NAD}}{NAD}\left(1+\frac{[NADH]}{K_{NADH}}\right)+1\right)}$
PYK	$\frac{v_{PYK}^{max}[PEP]\left(\frac{PEP}{K_{PEP}}+1\right)^{n_{pyk}-1}[ADP]}{K_{PEP}\left(L_{PYK}\left(\frac{1+\frac{[ATP]}{K_{ATP}}}{\frac{[FDP]}{K_{FDP}}+\frac{[AMP]}{K_{AMP}}+1}\right)^{n_{pyk}}+\left(\frac{[PEP]}{K_{PEP}}+1\right)^{n_{pyk}}\right)([ADP]+K_{ADP})}$
Ppc	$\frac{K_1+K_2[AcCOA]+K_3[FDP]+K_4[AcCOA][FDP]}{1+K_5[AcCOA]+K_6[FDP]}\left(\frac{[PEP]}{K_m+[PEP]}\right)$
G6PDH	$\frac{v_{G6PDH}^{max}[G6P][NADP]}{\left([G6P]+K_{g6p}\right)\left(1+\frac{[NADPH]}{K_{ndph}}\right)\left(K_{nadp}\left(1+\frac{[NADPH]}{K_{nadph}}\right)+NADP\right)}$
PGDH	$\frac{v_{PGDH}^{max}[6PG][NADP]}{\left([6PG]+K_{6pg}\right)\left([NADP]+K_{nadp}\left(1+\frac{[NADPH]}{K_{nadph}}\right)\left(1+\frac{[ATP]}{K_{atp}}\right)\right)}$
Rpe	$v_{Rpe}^{max}\left([Ru5P]-\frac{[R5P]}{K_{eq}^{Rpe}}\right)$
Rpi	$v_{Rpi}^{max}\left([Ru5P]-\frac{[R5P]}{K_{eq}^{Rpi}}\right)$
TktA	$v_{TKtA}^{max}\left([R5P][Xu5P]-\frac{[S7P][GAP]}{K_{eq}^{TKtA}}\right)$
TktB	$v_{TKtB}^{max}\left([Xu5P][E4P]-\frac{[F6P][GAP]}{K_{eq}^{TKtB}}\right)$
Tal	$v_{TaL}^{max}\left([GAP][S7P]-\frac{[E4P][F6P]}{K_{eq}^{TKtB}}\right)$
DAHP	$\frac{v_{DAHPS}^{max}[E4P]^{ne4p}[PEP]^{npep}}{\left(K_{E4P}+[E4P]^{ne4p}\right)\left(K_{PEP}+[PEP]^{npep}\right)}$
PcK	$v_{PcK}^{max}\left(\frac{[OAA]\frac{[ATP]}{[ADP]}}{K_m^{OAA}\frac{[ATP]}{[ADP]}+[OAA]\frac{[ATP]}{[ADP]}+\frac{K_i^{ATP}K_m^{OAA}}{K_i^{ADP}}+\frac{K_i^{ATP}K_m^{OAA}}{K_m^{PEP}K_i^{ADP}}[PEP]+\frac{K_i^{ATP}K_m^{OAA}}{K_i^{PEP}K_l^{ATP}}\frac{[ATP]}{[ADP]}[PEP]+\frac{K_i^{ATP}K_m^{OAA}}{K_i^{ADP}K_l^{OAA}}[OAA]}\right)$
PDH	$\frac{v_{PDH}^{max}}{[NAD]}\left(\frac{1}{1+K_i\frac{[NADH]}{[NAD]}}\right)\left(\frac{[PYR]}{K_m^{PYR}}\right)\left(\frac{1}{K_m^{NAD}}\right)\left(\frac{[COA]}{K_m^{COA}}\right)$ $\left(1+\frac{[PYR]}{K_m^{PYR}}\right)\left(\frac{1}{NAD}+\frac{1}{K_m^{NAD}}+\frac{[NADH]}{K_m^{NADH}[NAD]}\right)\left(1+\frac{[COA]}{K_m^{COA}}+\frac{[AcCOA]}{K_m^{AcCOA}}\right)$

Pta	$\frac{v_{Pta}^{max}\left(\frac{1}{K_i^{AcCOA}K_m^P}\right)\left([AcCoA][P]-\frac{[AcP][CoA]}{K_{eq}}\right)}{\left(1+\frac{[AcCoA]}{K_i^{AcCoA}}+\frac{[P]}{K_i^P}+\frac{[ACP]}{K_i^{ACP}}+\frac{[CoA]}{K_i^{CoA}}+\left(\frac{[AcCoA][P]}{K_i^{AcCoA}K_m^P}\right)+\left(\frac{[AcP][CoA]}{K_m^{ACP}K_i^{CoA}}\right)\right)}$
Ack	$\frac{v_{Ack}^{max}\left(\frac{1}{K_m^{ADP}K_m^{ACP}}\right)\left([AcP][ADP]-\frac{[ACE][ATP]}{K_{eq}}\right)}{\left(1+\frac{[Acp]}{K_m^{ACP}}+\frac{[ACE]}{K_m^{ACE}}\right)\left(1+\frac{[ADP]}{K_m^{ADP}}+\frac{[ATP]}{K_m^{ATP}}\right)}$
Acs	$\frac{v_{Acs}^{max}[ACE][NADP]}{(K_m+[ACE])(K_{eq}+[NADP])}$
Cs	$\frac{v_{CS}^{max}[AcCoA][OAA]}{(K_d^{AcCoA}K_m^{OAA}+K_m^{AcCoA}[OAA])+\left([AcCoA]K_m^{OAA}\left(1+\frac{[NADH]}{K_{i1}^{NADH}}\right)\right)+\left([AcCoA][OAA]\left(1+\frac{[NADH]}{K_{i2}^{NDAH}}\right)\right)}$
ICDH	$\frac{[ICDH]\frac{K_f}{K_m^{ICIT}K_d^{NADP}}\left([ICIT]-\frac{[NADH][2KG]}{K_{eq}^{ICDH}[NADP]}\right)}{\left(\begin{array}{c}\frac{1}{[NADP]}+\frac{[ICIT]K_m^{NADP}}{K_m^{ICIT}K_d^{NADP}[NADP]}+\frac{1}{K_d^{NADP}}+\frac{[ICIT]}{K_m^{ICIT}K_d^{NADP}}+\frac{[ICIT]}{K_d^{ICIT}[NADP]}\frac{[NADPH]K_m^{NADP}}{K_m^{ICIT}K_{einh}^{NADPH}}+\\ \frac{[NADPH]K_{eknh}^{2KG}}{K_m^{2KG}K_{enhe}^{NADPH}[NADP]}+\frac{[2KG]K_m^{NADPH}}{K_m^{2KG}K_{enhe}^{NADPH}[NADP]}+\frac{[2KG]}{K_m^{2KG}}\frac{[NADPH]}{K_{enhe}^{NADPH}[NADP]}+\frac{[2KG]K_m^{NADPH}}{K_m^{2KG}K_m^{NADPH}}\frac{[NADPH]}{K_{ekn}^{NADP}[NADP]}\end{array}\right)}$
IcL	$\frac{v_{lcl-f}^{max}\frac{[ICIT]}{K_m^{ICIT}}}{\left(1+\frac{[ICIT]}{K_m^{ICIT}}+\frac{[SUC]}{K_m^{SUC}}+\frac{[PEP]}{K_m^{PEP}}+\frac{[2KG]}{K_m^{2KG}}+\frac{1}{K_l}\right)}$
MS	$\frac{v_{MS}^{max}\frac{[GOX]}{K_m^{GOX}}\frac{[AcCoA]}{K_m^{AcCoA}}-v_{MS}^{max}\frac{[MAL]}{K_m^{MAL}}}{\left(1+\frac{[GOX]}{K_m^{GOX}}+\frac{[MAL]}{K_m^{MAL}}+\left(1+\frac{[AcCoA]}{K_m^{AcCoA}}\right)\right)}$
aKGDH	$\frac{v_{2KGDH}^{max}[aKG][CoA]}{\left(\begin{array}{c}\frac{K_m^{NAD}[aKG][CoA]}{[NAD]}+K_m^{CoA}[aKG]+K_m^{2KG}[CoA]+[aKG][CoA]+\frac{K_m^{2KG}K_z[aKG][SUC][NADH]}{K_1^{2KG}K_1^{SUC}[NAD]}\\ \frac{K_m^{2KG}K_z[SUC][NADH]}{K_1^{SUC}[NAD]}+\frac{K_m^{NAD}[aKG][CoA][NADH]}{K_1^{NADH}[NAD]}+\frac{K_m^{CoA}[aKG][SUC]}{K_1^{SUC}}\end{array}\right)}$
SDH	$\frac{v_{SDH1}v_{SDH2}\left([SUC]-\frac{[FUM]}{K_{eq}}\right)}{K_m^{SUC}v_{SDH2}+v_{SDH2}[SUC]+\frac{v_{SDH1}[FUM]}{K_{eq}}}$
Fum	$\frac{v_{Fum1}v_{Fum2}\left([FUM]-\frac{[MAL]}{k_{Fum\,eq}}\right)}{K_m^{Fum}v_{Fum1}+v_{Fum2}[FUM]+\frac{V_{Fum1}[MAL]}{K_{eq}}}$
Mez	$\frac{v_{Mez}^{max}[MAL][NADP]}{(K_{MAL}+[MAL])(K_{eq}+[NADP])}$
MDH	$\frac{v_{MDH1}v_{MDH2}\left([MAL]-\frac{[OAA]}{K_{eq}}\right)}{\left(\begin{array}{c}\frac{K_1^{NAD}K_m^{MAL}v_{MDH2}}{[NAD]}+K_m^{MAL}v_{MDH2}+\frac{K_m^{NAD}v_{MDH2}[MAL]}{[NAD]}+v_{MDH2}[MAL]+\frac{K_m^{OAA}v_{MDH1}[NADH]}{K_{eq}[NAD]}+\frac{K_m^{NADH}v_{MDH1}[OAA]}{K_{eq}[NAD]}+\\ \frac{v_{MDH1}[NADH][OAA]}{K_{eq}[NAD]}+\frac{v_{MDH1}K_m^{OAA}[NADH]}{K_{eq}K_1^{NAD}}+\frac{v_{MDH2}K_m^{NAD}[MAL][OAA]}{K_1^{OAA}[NAD]}+\frac{v_{MDH2}[MAL][NADH]}{K_1^{NAD}}\\ +\frac{v_{MDH1}[MAL][NADH][OAA]}{K_{eq}K_1^{MAL}[NAD]}+\frac{v_{MDH2}[MAL][OAA]}{K_{II}^{OAA}}+\frac{v_{MDH1}[NADH][OAA]}{K_{II}^{NAD}K_{eq}}+\frac{K_1^{NAD}v_{MDH2}[MAL][NADH][OAA]}{K_{II}^{NAD}K_m^{OAA}K_1^{NADH}}\end{array}\right)}$

Each enzyme described by mass balance equation is the quantity of all species in a solution containing a particular atom; it must be equal to the amount of that atom delivered to the solution. The solution of the equation may be derived from the dynamic equation and the initial state of the system as well as a graph or table of values of the solution may then be compared with the observed pattern of nature; i.e. to what extent

the solution of equation matches the pattern is a measure of the validity of the mathematical model. The metabolite concentration rate of the changes in this metabolic network is given by the following equation

$$\frac{dC_i}{dt} = \sum_j R_{ij} v - \mu C_i \quad (3.1)$$

Where, C_i is the concentration of metabolite i, R_{ij} is the stoichiometric coefficient of metabolite i in the reaction j, v_j is the rate of the reaction j and μC_i is the growth rate on the dilution effect.

The need for the mass balance equation in kinetic rate equation is to analysis the system, which mass balance is an application of conservation of the mass and described in Table 3.4 below.

Table 3.4: Mass Balance

Metabolites	**Mass balance description**
Cell	$\frac{d[X]}{dt} = \mu[X]$
Extra Glucose	$\frac{d[GLC^{ex}]}{dt} = -v_{PTS}[X]$
Glucose-6-phosphate	$\frac{d[G6P]}{dt} = v_{PTS} - v_{PGI} - v_{G6PDH} - \mu[G6P]$
Fructose 6-phospahte	$\frac{d[F6P]}{dt} = v_{PGI} - v_{PFK} + v_{TKTB} + v_{TAL} - \mu[F6P]$
Fructose 1,6-Phosphate	$\frac{d[FDP]}{dt} = v_{PFK} - v_{ALDO} - \mu[FDP]$
Glyceraldehyde 3-phosphate	$\frac{d[GAP]}{dt} = 2v_{ALDO} - v_{GAPDH} + v_{TKTA} + v_{TKTB} - v_{TAL} - \mu[GAP]$
Phosphoenol-pyruvate	$\frac{d[PEP]}{dt} = v_{GAPDH} + v_{PCK} - v_{PTS} - v_{PYK} - v_{PPC} - \mu[PEP]$
Pyruvate	$\frac{d[PYR]}{dt} = v_{PYK} + v_{PTS} + v_{MEZ} - v_{PDH} - \mu[PYR]$
Acetyl-CoA	$\frac{d[AcCoA]}{dt} = v_{PDH} + v_{ACS} + v_{CS} - v_{PTA} - \mu[AcCoA]$
Isocitrate	$\frac{d[ICIT]}{dt} = v_{CS} - v_{ICDH} - v_{ICL} - \mu[ICIT]$
2-Keto-D-gluconate	$\frac{d[2KG]}{dt} = v_{ICDH} - v_{2KGDH} - \mu[2KG]$
Succinate	$\frac{d[SUC]}{dt} = v_{2KGDH} + v_{ICL} - v_{SDH} - \mu[SUC]$

Fumrate	$\frac{d[FUM]}{dt} = v_{SDH} - v_{FUM} - \mu[FUM]$
Malate	$\frac{d[MAL]}{dt} = v_{FUM} + v_{MS} - v_{MDH} - v_{MEZ} - \mu[MAL]$
Oxaloacetate	$\frac{d[OAA]}{dt} = v_{MDH} + v_{PPC} - v_{CS} - v_{PCK} - \mu[OAA]$
Glyoxylate	$\frac{d[GOX]}{dt} = v_{ICL} - v_{MS} - \mu[GOX]$
Acetyl phosphate	$\frac{d[ACP]}{dt} = v_{PTA} - v_{ACK} - \mu[ACP]$
Acetate	$\frac{d[ACE^{ex}]}{dt} = (v_{ACK} - v_{ACS})[X]$
6-Phosphogluconolactone	$\frac{d[6PG]}{dt} = v_{G6PDH} - v_{6PGDH} - \mu[6PG]$
Ribose 5-phosphate	$\frac{d[Ru5P]}{dt} = v_{6PGDH} - v_{RPE} - v_{RPI} - \mu[Ru5P]$
Ribulose 5-phosphoenolpyruvate	$\frac{d[R5P]}{dt} = v_{RPI} - v_{TKTA} - \mu[R5P]$
Xylulose 5-phsophate	$\frac{d[Xu5P]}{dt} = v_{RPE} - v_{TKTA} - v_{TKTB} - \mu[Xu5P]$
Sedoheptulose 7-phosphate	$\frac{d[S7P]}{dt} = v_{TKTA} - v_{TAL} - \mu[S7P]$
Erythrose 4-phsophate	$\frac{d[E4P]}{dt} = v_{TAL} - v_{TKTB} - \mu[E4P]$

Where in Table 3.4 the [.] denotes the concentration, μ is the specific growth rate, v_i stands for the intracellular fluxes, and the superscript $^{(ex)}$ means extra cellular. In the end of each equation there is term called μ and it denotes the dilution effect because of the increases in cell volume which happens as the cell grows (Chassagnole et al., 2002). In Figure 3.2 GAP and DHAP is lumped together and are considered to be in equilibrium for simplicity purposes (Kadir et al., 2010).

3.3.4 Kinetic Parameters

Kinetic parameter or (*kinetic enzyme*) is the study of the chemical reactions that are catalyzed by enzymes; which the values that used in this study are the initial values of the kinetic parameters measured by Milli-molar. These proteins are commonly enzymes that assist in biochemical transformation. There are 194 kinetic parameters involved in this model as stated in Table 3.5 below with their reference.

Table 3.5: The kinetic parameters used in this study

Enzyme activities	**Kinetic Parameters with values *Mm***	**Reference**
Miu	miu_Max=0.6 , K_Xs= 0.1, K_ACETs= 0.05	Kadir et al., 2010
Pts	V_PTSmax= 25.739, K_PTSa1=1, K_PTSa2=0.01, K_PTSa3=1, n_PTSg6p=4, K_PTSg6p =0.5;	Kadir et al., 2010
Pgi	V_PGImax=26.3711, K_PGIg6p=2.46, K_PGIf6p=0.2, K_PGIeq=0.43, K_PGIg6p_6pginh=0.2, K_PGIf6p_6pgi=0.2;	Kadir et al., 2010
Pfk	V_PFKmax=24.613, K_PFKatp_s=0.16, K_PFKadp_a=239, K_PFKadp_b=0.25, K_PFKadp_c=0.36, K_PFKamp_a=8.74 K_PFKamp_b=0.01, n_PFK=4, L_PFK=4000000, K_PFKf6p_s =0.14, K_PFKpep=3.26;	Kadir et al., 2010
Aldo	V_ALDOmax=2.8337, K_ALDOfdp=0.133 K_ALDOgap=0.088, K_ALDOdhap=0.088, K_ALDOgapinh=0.6, K_ALDOeq=0.14, V_ALDOblf=2;	Kadir et al., 2010
Gapdh	V_GAPDHmax=121.29, K_GAPDHgap=0.15, K_GAPDHpgp0.1, K_GAPDHnad=0.45, K_GAPDHnadh=0.02, K_GAPDHeq=0.63;	Kadir et al., 2010
Pyk	V_PYKmax=1.085,Km_PYKpep=0.31, Km_PYKfdp=0.19, Km_PYKamp=0.2, Km_PYKatp=22.5, Km_PYKadp=0.26, L_PYK=1000, n_PYK=4, n_PK=3	Kadir et al., 2010
PDH	V_PDHmax2= 27171, Km_PDHpyr2=1, Km_PDHnad2=0.4, Km_PDHcoa2 0.014, Km_PDHaccoa2=0.008, Km_PDHnadh2=0.1, Ki_PDH2 = 46.4;	Kadir et al., 2010
G6PDH	V_G6PDHmax=0.97922*50, K_G6PDHnadp=14.4, K_G6PDHnadphnadpinh=0.01,K_G6PDHnadphg6pinh=0.18,	Kadir et al., 2010

	K_G6PDHg6p=0.07;	
6PGDH	V_6PGDHmax=1.81, K_6PGDH6pg=0.1, K_6PGDHnadp=0.028, K_6PGDHnadphinh=0.01, K_6PGDHatpinh=3.0;	Kadir et al., 2010
RPE	V_RPEmax=18.486, K_RPEeq=1.4;	Kadir et al., 2010
RPI	V_RPImax=13.319, K_RPIeq=4.0;	Kadir et al., 2010
TKTA	V_TKTAmax=29.348, K_TKTAeq=1.2;	Kadir et al., 2010
TKTB	V_TKTBmax=316.22, K_TKTBeq=10;	Kadir et al., 2010
TAL	V_TALmax =24.499, K_TALeq=1.05;	Kadir et al., 2010
CS	CSmax=34.7244*2, Km_CSaccoa=0.18, Km_CSoaa=0.04, Kd_CSaccoa=0.1, Kd1_CSh=1e-5, Kd2_CSh=2e-4, Ki_CSatp=0.58, Ki1_CS2kg=0.015, Ki2_CS2kg=0.256, Ki1_CSnadh=3.3e-4, Ki2_CSnadh=8.4e-3, K_cat0=1;	Kadir et al., 2010
ICDH	ICDH=1.8785*13, Keq_ICDH=1000, Kf_ICDH=4830*60, Km_ICDH2kg=0.038, Km_ICDHco2=2.2, Km_ICDHicit= 0.011, Kd_ICDHnadp=0.006, Km_ICDHnadp=0.017, Kd_ICDHicit= 0.003, Kd_ICDHnadph=0.00014, Keinh_ICDHnadph=7e-3, Keknh_ICDH2kg=5.5, Kd_ICDHco2=1.6, Keke_ICDHco2=1.6, Kekn_ICDHnadp=1.6e-4, Km_ICDHnadph=0.0036, Kenhe_ICDHnadph=0.028;	Kadir et al., 2010
2KGDH	V_2KGDHmax=149.74/4, K_2KGDH2kg=1, KI_2KGDH2kg=0.75, K_2KGDHcoa=0.002, K_2KGDHnad=0.07, K_2KGDHsuc=1, K_2KGDHnadh=0.018, K_2KGDHz=1.5;	Kadir et al., 2010
SDH	Km_SDHsuc=0.1, V_SDH1=1.1334,	Kadir et al.,

	V_SDH2=1.1334, Keq_SDH=10;	2010
FUM	Km_FUM=0.1, V_FUM1=1.1334, V_FUM2=1.1334, Keq_FUM=10;	Kadir et al., 2010
MDH	V_MDH1=25.874, V_MDH2=25.874, Keq_MDH=1, KI_MDHnad=0.31, KI_MDHnadh=0.04, KI_MDHmal=3.30, KI_MDHoaa=0.27, Km_MDHnad=0.10, Km_MDHnadh=0.04, Km_MDHmal=1.33, Km_MDHoaa=0.27, KII_MDHnad=0.31, KII_MDHoaa=0.17;	Kadir et al., 2010
PTA	V_PTAmax=0.83902*15, Ki_PTAaccoa=0.2, Km_PTAp=2.6, Ki_PTAp=2.6, Ki_PTAacp=0.2, Ki_PTAcoa=0.029, Km_PTAacp=0.7, Keq_PTA=0.0281;	Kadir et al., 2010
ACK	V_ACKmax=191.02*12, Km_ACKacp=0.16, Km_ACKadp=0.5, Km_ACKacet=7, Km_ACKatp=0.07, Keq_ACK= 174.2;	Kadir et al., 2010
ACS	V_ACSmax=0.089971*150, Km2_ACS=0.07, Keq_ACS=0.15;	Kadir et al., 2010
MEZ	V_MEZ=0.1058, Km_MEZmal=0.37, Keq_MEZ=0.10;	Kadir et al., 2010
PCK	V_PCKmax=4.5116, Km_PCKatp=0.06, KI_PCKatp=0.04, Ki_PCKatp=0.04, Km_PCKoaa=0.67, Ki_PCKpep=0.06, KI_PCKoaa=0.45; Km_PCKpep=0.07, Ki_PCKadp=0.04;	Kadir et al., 2010
PPC	V_PPCmax2=3.7719/20,Km_PEP=0.3231, k1=0.03176, k2=1.2878, k3=0.05425, k4=0.8139, k5=0.0939, k6=0.2693;	Kadir et al., 2010
ICL	r_ICLmax_f=3.8315, K_ICLicit=0.0104, K_ICLsuc =1.19, K_ICLpep=0.91, K_ICL2kg=1.35;	Kadir et al., 2010
MS	Vf_MSmax=3.6869, Vf_MS=3.6869, Vr_MS =3.6869/100,	Kadir et al., 2010

```
Km_MSgox=2,  Km_MSaccoa=0.01,
Km_MSmal=1,  Km_MScoa=0.1;
```

3.4 Sensitivity analysis technique

In biology systems, the stability analysis shows how a biochemical system responds to small a perturbation that alters in steady-state condition, which is an important part in biology system development. The need for the sensitivity in the system biology mostly is to reduce the huge number of the parameters by choosing the most sensitive parameters when we need to be identified. Sensitivity analysis methods have been widely applied to study the biological system and it can provide valuable insights on how robust the biological responses are with respect to the changes of biological parameters and which model inputs are the key factors that affect the model outputs; moreover sensitivity analysis is valuable for guiding experimental analysis, model reduction and parameter estimation. Sensitivity analysis has two approaches and they are; local methods which study the impact of small perturbation on the model output and the global methods which studies how the model output are affected by large variations of the model outputs.

In the Algorithm 3.1 below, represents the simulation of the model under study with four major steps start by 194 parameters (*input*), 53 model response (*output*), 28 reaction rate in addition to ODE to be solved steps 1 to 4. The fifth step is the validation.

Algorithm 3.1: Model Simulation

1. **Input :** $x \in R^n$ **with** $n = 194$,
2. **Output :** $y \in R^m$ **with** $m = 53$,
3. **Set the** x **values in their equations 1 to 28,**
4. **Solve ODE,**
5. **The validation is set to** y.

Where x is input of the model, n is the number of inputs, y is output of the model, R is the range of inputs and outputs, and m is the number of outputs.

In this study therefore, the local method of one-at-a-time sensitivity measures was used to study the sensitivity analysis for large-scale of kinetic parameters of Algorithm 3.1. It is conceptually the simplest method to sensitivity analysis. The repetition of every one parameter at a time while holding the others, fixed is a sensitivity ranking that can be obtained quickly by increasing each parameter by a given percentage while leaving all others constant, thereby quantifying the change in model output. This method is applied to 194 kinetic parameters represented in the model under study. In order to achieve the sensitivity, the Algorithm 3.2 below should be followed and the equations used in these methods are stated below should be used.

There are four major steps and they are stated in Algorithm 3.2. In order to apply the sensitivity analysis for large scale kinetic parameters using one-at-a-time sensitivity measures, the first step identify how many inputs to be investigated by the sensitivity method; the second step in each individual kinetics should be increased by percentages of 10%, 20%, 40% and 80%. The third step find the differences between the actual model response outputs represented by y_0 and the simulation response after $k \pm$ (means k increased or decreased) represented by y. The fourth steps are how to quantify the changes over the outputs which proposed the highest variance and the highest Mean.

To identify the changes in the model output by increasing the whole kinetics into percentages, follow these equations below:

$$S = K + t\,K \tag{3.2}$$

Where S is the new simulation result, K is the kinetic parameter and t is the percentage of kinetic K.

For quantifying the changes this two equations below should use:

$$V = K_m - S \tag{3.3}$$

Where V is the differences, K_m is the actual model kinetic values and S is the new simulation result.

$$SA = \frac{\sum y}{\sum y_0^m} * 100 \qquad (3.4)$$

Where SA is the Sensitivity Analysis over the model output, y is the summation percentage over the all metabolites after increased or decreased by the kinetics and y_0^m is the summation number of the model response. Bellow the Algorithm 3.2 used in this study as described in the previous paragraphs.

Algorithm 3.2: Sensitivity Analysis

1. **Input :** x, j
2. $x_j \in [x_{j,min}, x_{j,max}]$,
3. $x_j \mid j = 1, \dots, 194$,
4. *Input*
5. $k = x_j$,
6. $x_j = 0.1$,
7. *For*
8. $i = 1:4$
9. *Use algorithm 3.1 and to be simulated with the value of* k.
10. *If* $k : \pm$,
11. *Find* $|y_0 - y|$
12. *Then*
13. $S = K + tK$,
14. $t = t * 2$,
15. **End**
16. **Quantify the changes using the highest variance in step 11.**
17. **Quantify the changes using the highest Mean using these equation** $SA = \frac{\sum y}{\sum y_0^m} * 100$.

The Algorithm 3.2 is translate to 9 major steps below:

In the first step each parameter has minimum and maximum values in the model under study represented in step 1 and 2 where x are the paremeters and j is the parameter position start from 1 to 194.

The secound step set $k = x_j$ where k is the kinetic parameters and x_j is the kinetic parameter to jth position in step 5.

The third step set the percentage changes values to t step 6.

The fourth step the range percentage iteration set to i step 8.

The fifth step the Algorithm 3.1 should used to be investigated step 9.

The sixth step apply the sensitivity S by using this equation $S = K + t\,K$ step 10.

The seventh step if the the value of the sensitiviy increased or decreased $S \pm$ find the defferences between the new sensitivity result and the orignal result of the model output using these equation $y_0 - y$, where y_0 is the actul model outputs result and y is the new sensitivity result steps 11 and 12.

The eighth step set the new iteration t by these equation $t = t * 2$ step 14.

The ninth step is how to quantify the changes using steps 16 and 17.

3.5 Optimization algorithm

Mathematical optimization plays an important role in the development of system biology. The PSO algorithm is proposed to achieve large-scale kinetic parameters optimization in the metabolic network model of *E. coli* which was formulated by (Kadir et al., 2010) to minimize the model errors with experimental data. PSO was introduced as a heuristic method (Eberhart and Kennedy, 1995) inspired by the food-searching behaviors of fish and their activities or a flock of birds in D-dimensional search space.

In order to apply the PSO algorithm for our problem of this study, we first initialize the PSO particles; the particles are number of birds, the steps of the birds, problem dimension and inertia weight.

Second, initialize the acceleration coefficients of PSO c_1, c_2, $r_1$and r_2.

Third, initialize the swarm velocity and position by choosing the range of searching for each dimension problem.

Fourth, initialize the kinetic rate equation with their values of metabolites and kinetic parameters that are stated in the sensitivity analysis result. Then initialize the kinetic parameters that are stated as the most sensitive parameters to be unknown and should be found by the fitness function during PSO execution and replace the metabolites values with the experimental data.

Fifth, the fitness function used to optimize the large-scale metabolic network of *E. Coil* system model in this study and to reduce the errors between the model response and the real experimental data transfer as follows:

$$f = \left|(y_{0,1} - y_1) + (y_{0,2} - y_2) + \ldots + (y_{0,i} - y_i)\right| \quad (3.5)$$

Where $y_{0,i}$ is the actual model reaction rate resulting from $0, i$ kinetics and y_i is a simulation reaction result for i kinetics.

Sixth, calculate the particle velocity by this equation below:

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (p_i(t) - x_i(t)) + c_2 r_2 (G(t) - x(t)) \quad (3.6)$$

Where; $p(t)$ is the best position already found by particle i until time t and $G(t)$ is the best position already found by a neighbor of particle i until time t, ω is an inertia weight parameter to explore the search space. c_1, c_2 are acceleration coefficients toward p and G respectively, while p is the current local best particle and G is the current global best particle at iteration t; and r_1, r_2 are random number between 0 and 1.

Seventh, update the particle position by this equation below:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (3.7)$$

In each iteration, the particles will use eq (3.6) & (3.7) to update their position (x_i) and velocity (v_i), but the algorithm used in this work will be described in Figure 3.4. Normally, the estimation of the unknown parameters techniques is based on the difference between the simulated model and actual system dynamic behavior (Yukako et al., 2013).

Eighth, if the fitness function values is better than the best fitness values of the best position already found by particle i until time t described in equation 3.6 by ($p_i(t)$) set the current values as the new best position already found by a neighbor until time t described in equation 3.6 by ($G(t)$).

Ninth, repeat the previous steps until the number of iteration or the exit criteria is met, or detecting high optimum result solution for study problem.

During the PSO execution, the maximum number of generation is set as 100 (bird-steps), and the dimension problem will be identified based on the sensitivity analysis result, while the population size (iterations) was repeated for 100 times. The linear inertia weight ω is 0.9, PSO parameter c_1=1.5 and c_2=0.8 with lower and upper values 1± for each kinetics. The pseudo code explaining the relation between the optimization theory and their equations used in this study is shown below in Algorithm 3.3.

Algorithm 3.3 PSO

1. Initialize particles of PSO (number of birds, maximum of bird steps, problem dimension, and ω).

2. Initialize the parameters of r_1, r_2, c_1, c_2;

3. Initialize the swam velocities and position;

4. For each particle

5. Iter i=1: bird_step: i++;

6. Initialize the kinetics and their reaction rates

7. Calculate the fitness of each reaction kinetic parameter involved by this equation $\mathbf{f} = \left|(y_{0,1} - y_1) + (y_{0,2} - y_2) + \ldots + (y_{0,i} - y_i)\right|$

8. Find in the neighborhood the particles with best fitness.

9. Calculate the particle velocity using this equation

$$v(\mathrm{t}+1) = \omega v(\mathrm{t}) + c_1 r_1\big(p(t) - x(t)\big) + c_2 r_2\big(G(t) - x(t)\big),$$

10. Update the particle position by this equation

$x(\mathrm{t}+1) = x(t) + v(\mathrm{t}+1)$ **position (x) and velocity (v).**

11. If the fitness values is better than the best fitness values of (p_i best) set the current values as the new (G_i best).

12. Repeated until a stopping iteration is met, or discovering high quality solution.

13. End.

A simple explanation of the PSO's operation above in Algorithm 3.3 is as follows. Each particle represents a possible solution to the optimization. During each iteration in PSO, the accelerating direction of one particle was determined by its own best found solution so far and the global best position discovered so far by any of the particles in the swarm. This means that if a particle discovers a promising new solution, all the other particles will move closer to it, exploring the region more thoroughly in the process (Yin et al., 2006).

Let H denote the number of birds (Swarm) and n the dimension of the problems. Each individual ($1 \leq i \leq H$) has the attributes: A current position in the search space $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$, a current velocity $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$, and a personal best ($p_i\ Best$) position (the position giving the best fitness value experience by particle) $p_i = (p_{i1}, p_{i2}, \ldots, p_{in}$). At each iteration, each particle in the swarm updates its velocity according to (Kennedy and Eberhart, 1995), assuming that the function fitness is to be minimized, and that r_1, and r_2 are two random numbers uniformly distributed in the interval (0 to 1).

The parameter ω in equation 3.5 is called the inertia weight that is typically set up to vary linearly from 0.9 to 0.4 during the course of search process. The inclusion of inertia weight leads to faster convergence of the PSO algorithm.

The acceleration coefficients c_1 and c_2 can be used to control the rate in which a particle will move in a single iteration and thus, may exert a great influence on the convergence speed of PSO. Typically, these are both set to a value of 2.0, although assigning different values to c_1 and c_2 sometimes leads to better performance (Salamn et al., 2002).

3.6 Validation

The conditions considered for this model are in a continuous culture (*it was used in a particular phase of the cell growth to grow microorganisms or cell continually*) but at a steady-state condition (*is a situation in which all state variables are constant in spite of ongoing process that strive to change them*) with dilution rates (*it's the process of reducing the concentration of a solute in solution, usually simply mixing with more solvent*) of 0.1 and 0.2 dilution rates.

The proposed methods, One-At-A-Time Sensitivity Measures and PSO algorithm aims to minimize the errors between the sensitivity simulation and original results as well as to achieve a large-scale kinetic parameter optimization in the metabolic network model of *E. coli*. The validation Algorithm steps 3.4 were proposed to enhance the strength of sensitivity analysis and optimization algorithm methods for large-scale kinetic parameters optimization of (Kadir et al., 2010) model which is described as follows:

Algorithm 3.4 Validation

1. **Initialize the pathways involved in the model of *E. coli*.**
2. **Initialize the kinetic parameters.**
3. **Apply kinetic rate computation.**
4. **Formulate the ODE of the metabolites for the simulation.**
5. **Solve the ODE.**
6. **Apply the One At A Time Sensitivity Measure analysis technique to step 2.**
7. **Optimize the sensitivity analysis of step 6 using PSO optimization algorithm.**
8. **Cross-check with experimental data and output of ODE.**
9. **If the optimization result minimizes the errors in step 8 then stop, if not repeat steps 6, 7, 8 till 9.**

The first five steps it was introduced previously. To this, the kinetic parameters that involved in step 3 it's our main problem need to be solved. The explanation steps stated as follow, the first step initializes the pathways that were used in model of *E. coli*.

The second step, initialize the kinetic parameters and the kinetic rate equation that involved in the model under study.

The third step applies the kinetic rate equation computation in addition to the pathways rules.

The fourth step formulates the ODE in order to simulate the pathways and the fifth step solves the ODE.

The sixth step applies the sensitivity analysis technique to the kinetic parameters that initialized in step 2.

The seventh step applies the optimization algorithm to the sensitivity analysis result of step 6 and replaces the kinetic sensitivity result with the optimized kinetic parameters in step 7.

The eighth step compares the ODE result of PSO algorithm with experimental data. The final step, if the optimization result after we executed in the model under study and the model responses have moved closely to experimental data stop, if not repeat the steps of 7, 8 till 9.

Finally to ensure that sensitivity analysis and PSO algorithm methods achieved the error minimization, accounts the errors minimization percentage for the model response optimization result under study and (Kadir et al., 2010) model response with experimental data using this equation below:

$$The\ error\ minimization\ percentage\ = \frac{\sum D_m - \sum E_m}{\sum E_m} * 100 \quad (3.8)$$

Where D_m are the metabolites model summations under study and E_m are the summation of the experimental data metabolites. Then quantify the errors minimization by account the variance between our metabolites model under study and (Kadir et al., 2010) metabolites.

3.7 Summary

In conclusion, Chapter 3 explained the methodology of large-scale kinetic parameter optimization of the metabolic network of *E. coli*. Therefore, four phases are included. They are model description, sensitivity analysis, optimization and validation.

Each phase has its own objectives and interests. The significant phase in order to implement the sensitivity analysis is a model descriptions phase, which consists of pathways, metabolites, values of enzymes, reaction rates, dynamic equation and kinetics values that were used in this study. The very influential sensitivity analysis phase to large-scale kinetics is the one-at-a-time sensitivity measures technique to achieve this

phase. The most important phase is the Optimization phase, which has with the view to optimize the kinetic parameter's in PSO algorithm. In addition, the implementation to reach the kinetics' estimation is combined with experimental data to achieve this phase. To prove that, this PSO algorithm is good enough in order to fix the model inputs, it proposed the model validation.

The result obtained from the experiments of sensitivity analysis and optimization phases are discussed in the next chapter with the validation.

CHAPTER 4

RESULT

4.1 Introduction

In order to evaluate the performance and the effectiveness of large-scale kinetic parameters in a metabolic network of *E. coli*, this thesis applied a set of tests carried out by 194 kinetic parameters. This research assumes that there are three parts that are to be executed by local sensitivity analysis technique, optimization algorithm and validation.

The performance of One-At-A-Time Sensitivity Analysis Measure in 0.1 and 0.2 dilution rate respectively and PSO algorithm with the large-scale kinetic parameters is used to identify the kinetic sensitivity and to optimize the kinetic parameters sensitivity result. The proposed methods with all the kinetic values are executed to achieve the sensitivity analysis goal; the Particle Swarm Optimization algorithm is applied for seven kinetic parameters which appear in 0.1 dilution rate. The validation part is explained in details.

4.2 Experimental results and analysis for sensitivity analysis and optimization

The experiment is implemented using Matlab platform. The algorithm methods of the experiment include using the one-at-a-time sensitivity measures and PSO algorithm. The methods and the metabolic network of *E. coli* introduced in chapter 3, and the kinetics on the metabolic network of *E. coli* was formulated by (Kadir et al., 2010).

4.3 Sensitivity analysis result

The kinetic parameters are tested using One-At-A-Time Sensitivity Measures quantified by the Mean in 0.1 and the highest variance in 0.2 dilution rate, 29 algebraic equations for kinetic expression and co-metabolites concentration and twenty nine differential equations, it was targeted to perform sensitivity analysis on the large-scale dynamic metabolic network under steady-state condition of *E. coli* for this model formulated by (Kadir et al., 2010) where is kinetic are 194.

4.3.1 Dilution rate 0.1 result

In 0.1 dilutions rate all the kinetics individually increased by 10%, 20% and 40% then the result of each kinetic parameters are quantified using the mean. The highest mean obtained being 40% shows that there are seven kinetic parameters that have the highest impact on the dynamic model output of *E. coli* as formulated by (Kadir et al., 2010). The sensitivity percentage and the analysis explanation are shown below in a Table 4.1.

Table 4.1: Sensitivity Percentage

Metabolites and Fluxes	Original values	V_PYKmax	n_PK	ICDH	Kf_ICDH	Kd_ICDHnadp	Km_ICDHnadp	V_ICLmax
Cell Con	1.5783	-11.98%	-27.87%	-3%	-3.00%	-4.59%	4.12%	3.99%
GLCex	0.022105	-45.95%	-263.06%	-10.98%	-10.98%	-17.82%	9.51%	12.16%
G6P	0.20345	12.24%	25.45%	3.35%	3.35%	5.05%	-4.92%	-4.77%
F6P	0.21311	9.99%	21.27%	2.71%	2.71%	4.09%	-3.89%	-3.81%
FDP	1.4621	60.42%	90.18%	13.44%	13.44%	20.09%	-17.84%	-20.37%
GAPDHAP	0.31094	34.37%	75.73%	4.63%	4.63%	7.28%	-4.3%	-5.76%
PEP	1.4914	33.71%	75.80%	3.9%	3.9%	6.21%	-2.92%	-4.67%
PYR	2.8117	-4.42%	9.74%	-8.92	-8.92	-13.96	10.19%	10.96%
AcCOA	1.0018	-0.43%	5.77%	-6.69%	-6.69	-10.46%	7.81%	8.26%
ICIT	0.21101	63.16%	90.59%	87.88%	87.88%	93.11%	-1155.72%	-3.95%
2KG	5.3724	9.47%	36.54%	8.99%	8.99%	15.77%	28.99%	-14.74%
SUC	0.57217	-4.57%	-4.87%	15.68%	15.68%	22.94%	-18.63%	-28.49%
FUM	0.35609	-2.74%	-2.16%	11.69%	11.69%	17.53%	-11.9%	-18.69%
MAL	0.14263	-0.27%	4.25%	11.94%	11.94%	18.37%	-8.07%	-17.49%
OAA	0.029637	8.35%	29.87%	11.99%	11.99%	18.98%	5.55%	-18.53%
GOX	0.34577	-7.06%	-11.61%	23.07%	23.07%	34.97%	-27.13%	-33.55%
AcP	2.0199	-30.76%	-127.28%	-10.24%	-10.24%	-16.42%	9.72%	11.6%

Metabolites and Fluxes	Original values	V_PYKmax	n_PK	ICDH	Kf_ICDH	Kd_ICDHnadp	Km_ICDHnadp	V_ICLmax
Cell Con	1.5783	-11.98%	-27.87%	-3%	-3.00%	-4.59%	4.12%	3.99%
ACE	0.000209	-55.5%	-408.78%	-12.96%	-12.96%	-21.18%	11.07%	14.04%
6PG	0.017832	-0.17%	-0.87%	-1.41%	-1.41%	-2.22%	0.73%	1.84%
Ru5P	0.02134	4.38%	8.52%	0.33%	0.33%	0.47%	-0.79%	-0.5%
R5P	0.07617	4.75%	9.36%	0.46%	0.46%	0.68%	-0.92%	-0.69%
Xu5P	0.026436	5.11%	9.98%	0.57%	0.57%	0.84%	-1.02%	-0.82%
S7P	0.004747	-28.5%	-185.29%	-2.3%	-2.3%	-3.89%	0.68%	2.13%
E4P	0.027433	33.67%	70.4%	5.88%	5.88%	9.07%	-6.41%	-7.8%
Miu	0.099617	-0.16%	-0.33%	-0.1%	-0.1%	-0.16%	-0.07%	0.1%
Pts	1.4003	10.93%	22.93%	2.99%	2.99%	4.48%	-4.36%	-4.22%
Pgi	1.3	11.58%	24.35%	3.23%	3.23%	4.87%	-4.65%	-4.57%
Pfk	1.3402	11.18%	23.54%	3.08%	3.08%	4.64%	-4.47%	-4.36%
Aldo	0.52536	80.38%	141.15%	14.52%	14.52%	22.49%	-17.7%	-18.48%
Gapdh	2.3756	4.65%	14.31%	1.72%	1.72%	2.61%	-2.78%	-2.29%
Pyk	0.62509	-28.6%	-56.26%	0.7%	0.7%	1.13%	-0.49%	-0.78%
Pdh.	1.766	-0.7%	-3.15%	4.16%	4.16%	6.39%	-5.27%	-5.52%
Cs	1.4682	3.47%	13.52%	6.91%	6.91%	10.72%	-8.20%	-8.85%
ICDH	0.93296	7.42%	25.01%	-2.01%	-2.01%	-2.1%	24.63%	0.53%
2KGDH	0.40201	5.4%	10.6%	-16.47%	-16.47%	-25.65%	19.16%	20.34%
Icl	0.51436	-6.1%	-10.45%	19.79%	19.79%	30.64%	-21.98%	-26.05%
Ms	0.47975	-6.05%	-10.37%	19.55%	19.55%	30.32%	-21.62%	-25.51%
SDH	0.85922	-0.83%	-0.96%	3.09%	3.09%	4.81%	-2.95%	-4.17%
Fum	0.8237	-0.74%	-0.91%	2.72%	2.72%	4.26%	-2.57%	-3.54%
MDH	1.2698	-2.76%	-4.6%	8.88%	8.88%	13.8%	-9.66%	-11.55%
Pita	0.2504	-35.12%	-161.62%	-10.85%	-10.85%	-17.51%	9.97%	12.29%
Ask	0.052391	-48.51%	-278.56%	-11.74%	-11.74%	-19.05%	10.23%	13%
Aces	0.15652	-48.54%	-278.48%	-11.74%	-11.74%	-19.06%	10.24%	13.01%
Pck	0.068774	-35.45%	-157.26%	8.55%	8.55%	13.82%	8.11%	-13.41%
Ppc	0.2702	22.89%	55.39%	-1.88%	-1.88%	-2.83%	2.96%	2.63%
Mez	0.019458	-0.2%	3.1%	8.91%	8.91%	13.97%	-5.7%	-12.04%
G6pgdh	0.079927	-0.15%	-0.74%	-1.2%	-1.2%	-1.88%	0.62%	1.58%
6pgdh	0.078143	-0.14%	-0.74%	-1.19%	-1.19%	-1.87%	0.62%	1.57%
Rpe	0.04516	-1.3%	-2.71%	-1.52%	-1.52%	-2.38%	0.95%	2.01%
Rpi	0.030597	1.26%	1.56%	-0.81%	-0.81%	-1.28%	0.23%	1.05%
Tkta	0.022996	0.14%	-0.97%	-1.2%	-1.2%	-1.89%	0.63%	1.6%
TktB	0.019783	-3.81%	-6.41%	-2.15%	-2.15%	-3.33%	1.65%	2.88%
TaL	0.022522	0.74%	2.92%	-1.18%	-1.18%	-1.85%	0.63%	1.59%
Mean	-	16.06%	54.87%	8.22%	8.22%	11.99%	8.85%	29.36%

We found that seven kinetic parameters were affected in the model output, in which the kinetics V_PYKmax, n_PK, ICDH, Kf_ICDH, Kd_ICDHnadp, Km_ICDHnadp and V_ICLmax represent the reaction rate of V_{pyk}, V_{icdh} and V_{icl} . The

concentration of the metabolite which are substrates and products of the reaction rate C_{PEP}, C_{PYR}, C_{ICIT}, C_{2KG}, C_{GOX}, and C_{SUC} are represented in Table (4.1) above.

The deviations in V_PYKmax result shows that, the Mean is 16.06% which the metabolites of FDP and ICIT are highly increased and ACE is highly decreased. The enzyme of ALDO increased a lot due to a high decrease in GLcex, which in turn is regulated by its effectors ATP, ADP and PEP. The deviations in n_PK shows that, the Mean is 54.87% which suggests that the metabolites of FDP, GAPDHAP, PEP, ICIT and E4P were highly increased while ACP, ACE and S7P were highly decreased; the enzymes of Aldo increased a lot while the Ack and the Pck decreased a lot due to a high increase in GLCex which in turn is regulated by the same V_PYKmax effectors. The deviations in ICDH, Kf_ICDH, Kd_ICDHnadp and Km_ICDHnadp gave rise to the Means, 8.22%, 8.22%, 11.9% and 8.85% respectively. The results has shown that, a high increase in the metabolite of ICIT, and the deviation in the result of Km_ICDHnadp causes a high decrease in ICIT also; which in turn is regulated by its effectors NADP, NADPH and 2KG. Moreover, the kinetics of ICDH and Kf_ICDH has the same results. These results may lead to a high increase in ICIT metabolites. The deviation in V_ICLmax Mean of 29.36% in the result shows that, the metabolites of SUC and GOX decreased. This deviation is regulated by its effector ICIT. In addition, the changes caused by four kinetic parameters in the reaction rate of V_{icdh} are important due to the changes in ICIT and 2KG during execution time. The rest of the kinetics changes and their quantifications are shown in appendix (A).

4.3.2 Dilution rate 0.2 result

At 0.2 dilutions rate all the kinetics individually increased by 10% and 20%. However, the formal analysis showed that at 10% increase, no affection was noticed while at 20% increase eight kinetic parameters affected the response of the model as quantified by the highest variance. The kinetics which are *V_ALDOmax, n_PK, Ki_PDH, ICDH, Kf_ICDH, V_SDH, V_FUM* and *V_ICL* represent the reaction rates of ($V_{aldo}, V_{pyk}, V_{pdh}, V_{icdh}, V_{icl}, V_{sdh}\ and\ V_{fum}$) with concentration of the metabolites which are the substrate and products of the reaction rates of ($C_{FDP}, C_{GAPDHAP}, C_{PEP}, C_{PYR}, C_{AcCOA}, C_{ICIT}, C_{2KG}, C_{SUC}, C_{GOX}, C_{FUM}\ and\ C_{MAL}$). The

kinetic also quantify the changes by using the highest variance, while the interaction of *V_ALDOmax* in the model response caused high changes in glycolysis pathway for *FDP*, *GAPDHAP*, *PEP* and *PYR* metabolites and *pts*, *pgi*, *pfk*, *aldo*, *gapdh* and *pyk* enzymes; also in Acetate formation there is increase in ACP metabolite and *pta*, *ack* and *acs* is described in Figure 4.1 & 4.2.

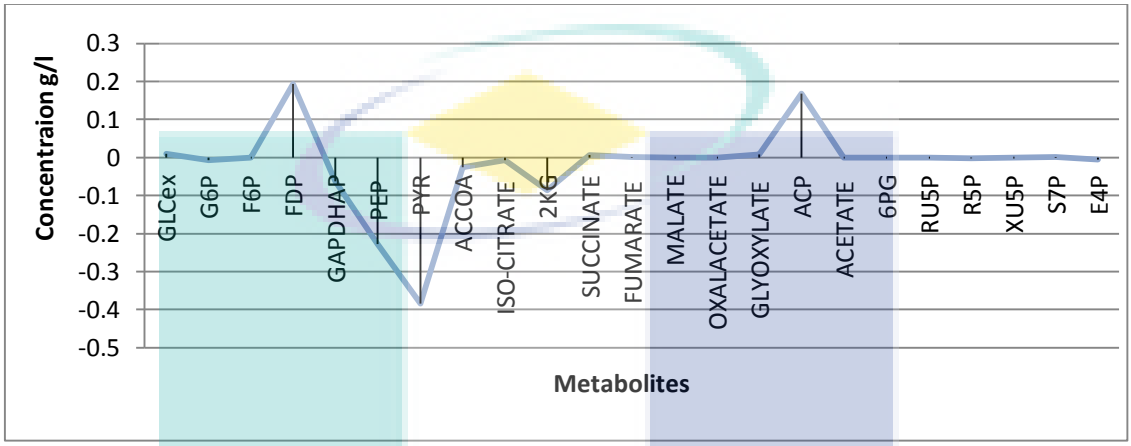


Figure 4.1: Metabolites affection by V_ALDOmax

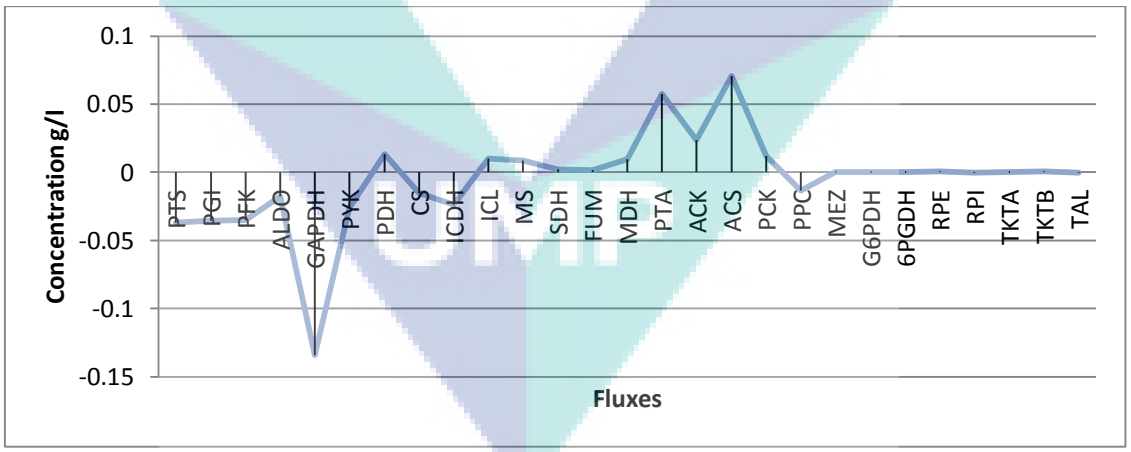


Figure 4.2: Fluxes affection by V_ALDOmax

The interaction of *n_PK* in the model response cause high changes in glycolysis, *TCA* cycle pathways in addition to Acetate formation for *FDP*, *PEP*, *PYR*, *2KG* and *ACP* metabolites, and *pyk*, *pta*, *ack* and aces enzymes. The interaction of *Kid_PDH* in the model response cause changes in glycolysis and TCA cycle pathways

in addition to Acetate formation for *PEP*, *PYR*, *AcCOA*, *2KG*, *SUC*, *ACE* and *ACP* metabolites, and *gapdh*, *pdh* enzymes as described in Figure 4.3 & 4.4.

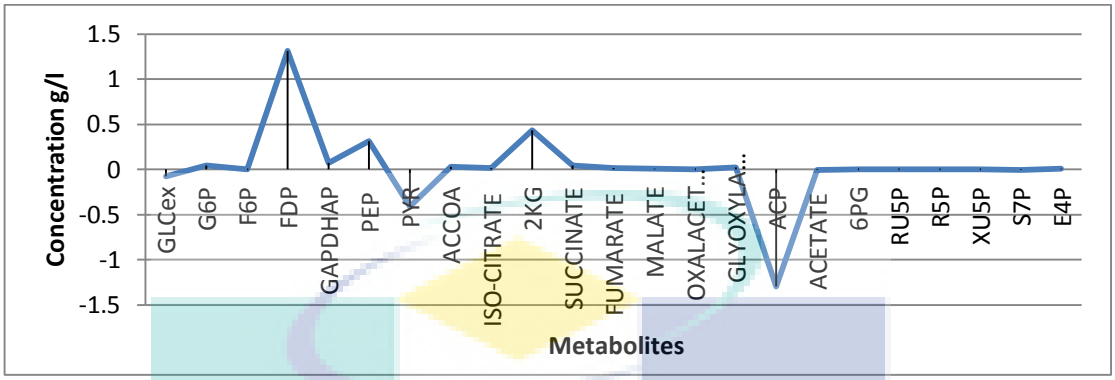


Figure 4.3: Metabolites affection by n_PK

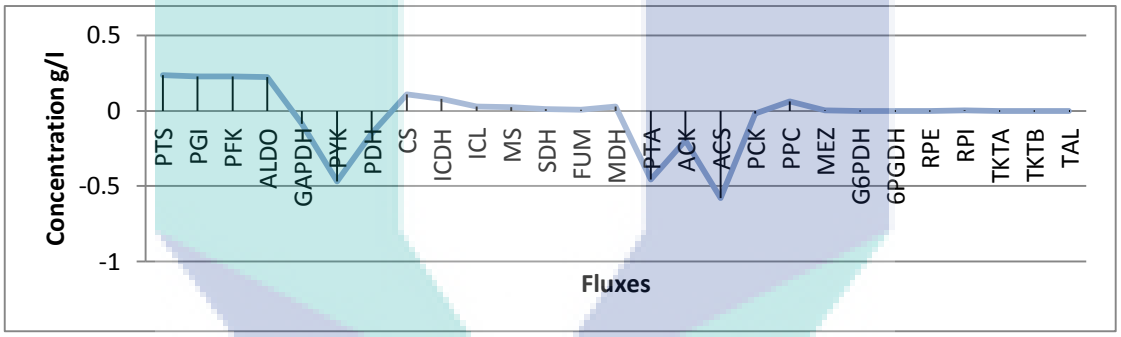


Figure 4.4: Fluxes affection by n_PK

The interaction of *ICDH* in the model response, cause high changes in glycolysis and *TCA* cycle pathways in addition to Acetate formation for these *FDP*, *PYR*, *AcCOA*, *ACE*, *ACE*, *ICIT*, *2KG*, *SUC*, *FUM* and *MAL* metabolites; and for these *pdh*, *pta*, *acs*, *sdh*, *fum*, *mdh*, *icl* and *ms* enzymes as described in Figure 4.5 & 4.6.

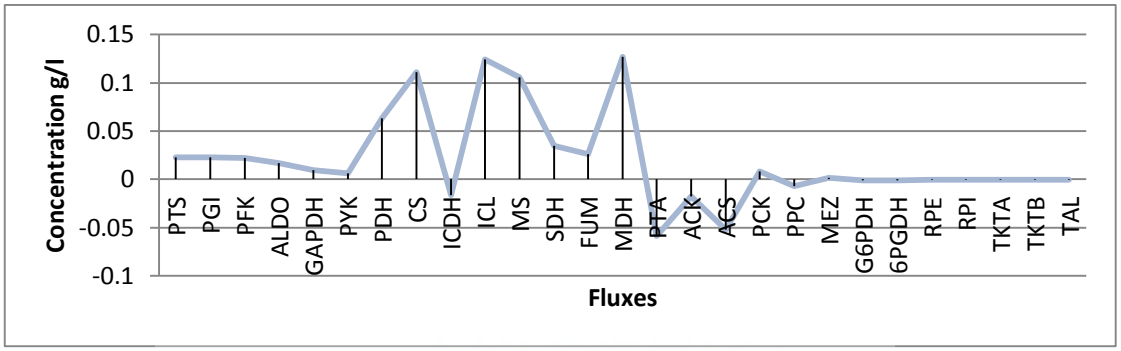


Figure 4.5: Metabolites affection by ICDH

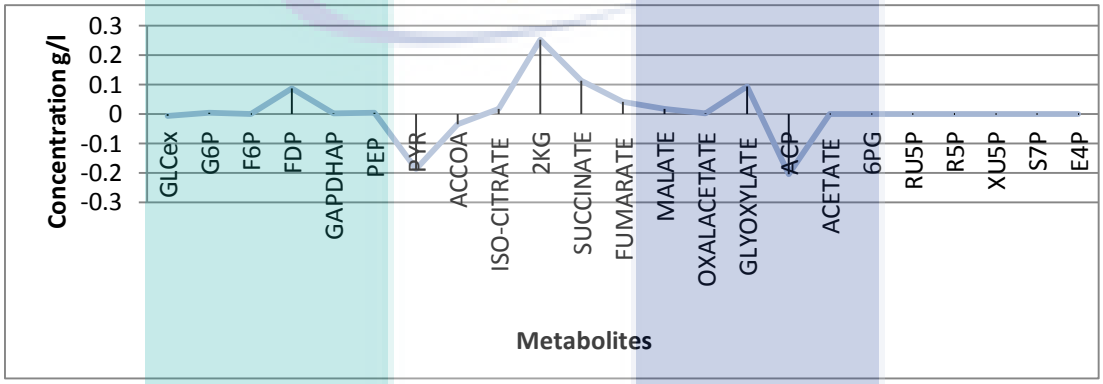


Figure 4.6: Fluxes affection by ICDH

The interaction of *Kf_ICDH* in the model response, caused high changes in glycolysis and *TCA* cycle pathways in addition to Acetate formation for these *FDP, PYR, AcCOA, ACE, ACE, ICIT, 2KG, SUC, FUM* and *MAL* metabolites as well as for these *pdh, pta, acs, sdh, fum, mdh, icl* and *ms* enzymes. The interaction of *V_SDH* in the model response, cause high changes in glycolysis and *TCA* cycle pathways in addition to Acetate formation for these *FDP, PYR, ACP, ICIT, 2KG, SUC, FUM, MAL* and *GOX* metabolites; and for these *pdh, cs, icl, sdh, fum, mdh, pta* and *acs* enzymes. Furthermore, the interaction of V_FUM in the model response cause high changes in glycolysis and *TCA* cycle pathways in addition to Acetate formation for these FDP, PYR, 2KG, FUM and *ACP* metabolites; and in these *pdh, cs, icdh, icl, ms, sdh, fum, mdh, pta* and *acs*. The interaction of *V_ICLmax* in the model, cause high changes in glycolysis and TCA cycle pathways in addition to Acetate formation for these *FDP,*

PYR, *2KG*, *SUC*, *FUM*, *AcCOA* and *ACP* metabolites; and for these *cs*, *icl*, *ms* and *mdh* enzymes.

Overall, there are three pathways that were affected by the increase in the 20% concentration in the glycolysis, TCA cycle pathways and Acetate formation. Moreover, the effect was more in these FDP, PYR, 2KG and ACP metabolites.

The rest of the kinetics changes and quantification are shown in appendix (B).

4.4 Kinetic Parameters identification result for 0.1 and 0.2 dilution rates

In this study the sensitivity analysis results of 0.1 and 0.2dilution rate were used for optimization due to the available data in the articles.

There are seven kinetic parameters that affected the response of (Kadir et al., 2010) model in 0.1 dilution rate. These kinetic parameters were involved in these three reaction rates of V_{pyk}, V_{icdh} and V_{icl} which contain these F6P, SUC, PEP, ICIT and 2KG metabolites, by applying PSO algorithm to these three reaction rates to fit closely with the experimental date that was taken from (Hoque et al., 2005). In the PSO algorithm that 3 reaction rate are initialized with objective function described in Ch 3, the fourth metabolites of F6P, PEP, ICIT and 2KG are replaced by the old metabolites of that three reaction rates under study and left the others kinetic parameters involved as their original values except the seven kinetic parameters identified in the sensitivity analysis result will be unknown and should be found by the PSO algorithm. The lower and upper values are determined based on (Kadir et al., 2010) publication for the purpose of achieving the best optimum values. The data is shown in Table (4.2). In order to prove that these algorithms fit the model simulation result closely to the experimental data are shown in validation part.

Table 4.2: Kinetic Parameters identification for 0.1 dilution

Kinetics	Nominal value	Optimal values
V_PYKmax	1.085	0.921
n_PK	3	3.32
ICDH	24.421	24.62
Kf_ICDH	289800	2829800
Kd_ICDHnadp	0.006	0.012
Km_ICDHnadp	0.017	0.013
V_ICLmax	3.8315	3.942

There are eight kinetic parameters that affected the response of (Kadir et al., 2010) model in 0.2 dilution rate. These kinetic parameters were involved in these seven reaction rates of $V_{aldo}, V_{pyk}, V_{pdh}, V_{icdh}, V_{icl}, V_{sdh}$ and V_{fum} which contain these $FDP, GAPDHAP, PEP, PYR, AcCOA, ICIT, 2KG, SUC, GOX, FUM,$ and MAL metabolites, by applying PSO algorithm to these seven reaction rates to fit closely with the experimental date that was taken from (Ishii et al., 2007). In the PSO algorithm that 7 reaction rate are initialized with objective function described in Ch 3, the metabolites of $Glc, G6P, F6P, FDP, PYR, 2KG, SUC, FUM, MAL, ACET, RU5P, R5P,$ and $S7P$ are replaced by the old metabolites of that model under study and left the others kinetic parameters involved as their original values except the eight kinetic parameters identified in the sensitivity analysis result will be unknown and should be found by the PSO algorithm. The lower and upper values are determined based on (Kadir et al., 2010) publication for the purpose of achieving the best optimum values. The data is shown in Table (4.3). In order to prove that these algorithms fit the model simulation result closely to the experimental data are shown in validation part.

Table 4.3: Kinetic parameters identification for 0.2 dilution

Kinetics	Nominal value	optimized values
V_ALDOmax	2.8337	2.5
n_PK	3	4.787
Ki_PDH	46.4	46.6082
ICDH	24.4205	24.6483
Kf_ICDH	289800	289800
V_SDH1	1.1334	1.498
V_FUM1	1.1334	1.594
V_ICLmax	3.8315	3.142

4.5 Validation and error minimization

The optimization problem in biological systems sometimes refer to an estimation problem, and the metabolite's estimation values from experimental data is a big challenge due to the model's integrity and complexity. To this end, the kinetic parameter's sensitivity result of 0.1 dilution rate values was used to be optimized and then tested in (Kadir et al., 2010) model to validate how the sensitivity analysis and PSO algorithm aid to fit the (Kadir et al., 2010) model metabolite's closely with real experimental data. The validation was performed by replacing the nominal sensitivity kinetic parameters values with kinetic parameters optimization result and leaves the kinetic of others with their actual values so that they can fit these FDP, PEP. ICIT and 2KG metabolites of (Kadir et al., 2010) to the experimental data of (Hoque et al., 2005). The validation result shows that the PSO algorithm can achieve the optimization. It was also discovered that 3 metabolites got close to the real experimental data which are *FDP, PEP* and *2KG*; while *ICIT* metabolite got a little bit far. This is due to the model's complexity. The analysis shows the real metabolite data which (Hoque et al., 2005) provided. Furthermore, because of the 3 reaction rates that contains 5 metabolites, 4 of the metabolites are optimized and the other metabolite was not found in the literature

review. The errors between (Kadir et al., 2010) and (Hoque et al., 2005) is 131% and the error minimization was achieved by reduce 12% from (Kadir et al., 2010) for our validation by account the errors percentage for (Kadir et al., 2010) with (Hoque et al., 2005) and our optimization result with (Hoque et al., 2005) and account the variance between them after we account the error minimization percentage. The metabolite of the model under study (Kadir et al., 2010) and our optimization result which was done to prove the strength of the proposed algorithm and the errors are minimized in Figure 4.7 below:

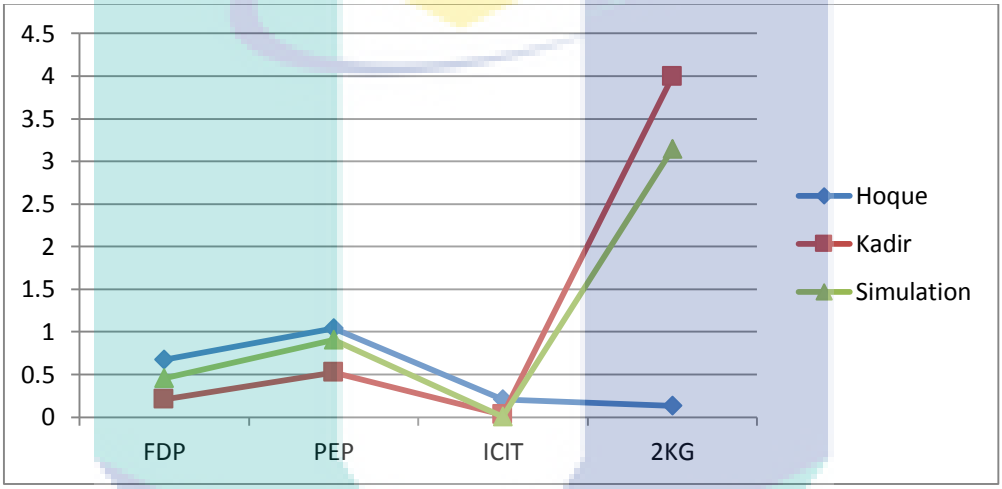


Figure 4.4: Validation of 0.1 dilution

The kinetic parameter's sensitivity result of 0.2 dilution rate values was used to be optimized and then tested in (Kadir et al., 2010) model to validate how the sensitivity analysis and PSO algorithm aid to fit the (Kadir et al., 2010) model metabolite's closely with real experimental data. The validation was performed by replacing the nominal sensitivity kinetic parameters values with kinetic parameters optimization result and leaves the kinetic of others with their actual values so that they can fit these $GLc_{ex}, G6P, F6P, FDP, PYK, 2KG, SUC, FUM, MAL, ACET, Ru5P, R5P\ and\ S7P$ metabolites of (Kadir et al., 2010) to the experimental data of (Ishii et al., 2007). The validation result shows that the PSO algorithm can achieve the identification. It was also discovered that 10 metabolites got close to the real experimental data which are $G6P, F6P, FDP, PYK, 2KG, SUC, FUM, MAL, R5P\ and\ S7P$; while $GLc_{ex}, ACET\ and\ Ru5P$ metabolite got a little bit far. This is due to the model's

complexity. The analysis shows the real metabolite data which (Ishii et al., 2007) provided. Furthermore, because of the 7 reaction rates that contains 13 metabolites, 10 of the metabolites are identified and the other metabolite was not identified due to the complexity of the model. The errors between (Kadir et al., 2010) and (Ishii et al., 2007) is 938% and the error minimization was achieved by reduce 294% from (Kadir et al., 2010) for our validation by account the errors percentage for (Kadir et al., 2010) with (Ishii et al., 2007) and our identification result with (Ishii et al., 2007) and account the variance between them after we account the error minimization percentage. The metabolite of the model under study (Kadir et al., 2010) and our identification result which was done to prove the strength of the proposed algorithm and the errors are minimized in Figure 4.8 below:

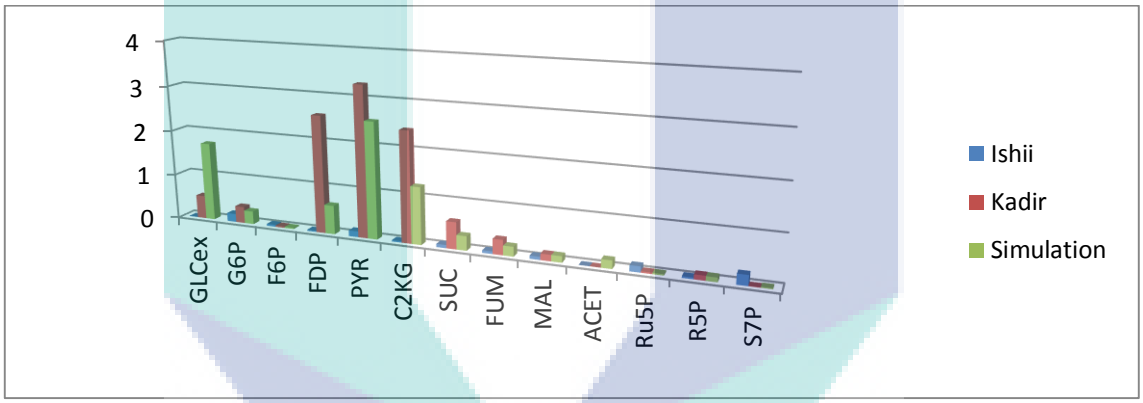


Figure 4.5: Validation of 0.2 dilution

The error minimization it shown in Table 4.9 for 0.1 and 0.2 dilution rates, which accounted by the percentage changes of our simulation result metabolite outputs summation with the summation of real experimental data then compared mathematically to the model result metabolite outputs summation under study; if the errors are minimized means we achieve our goals if not we repeat the PSO algorithm till the errors minimized. Furthermore, the errors was minimized by 11% in 0.1 dilution rate and 294% in 0.2 dilution rate, which described in Table (4.7) below.

Table 4.7: Kinetic parameters percentage minimization errors

Kadir 0.1 dilution	**131%**
Simulation	**120%**
Errors Minimization	**11%**
Kadir 0.2 dilution	**938%**
Simulation	**644%**
Errors Minimization	**294%**

4.6 Summary

The kinetic parameters in the metabolic network of *E. coli* is important, since small perturbation causes high changes either as a decrease or as an increase in the model outputs. However, the giving of assessment in order to develop or discover a new model has not yet evolves in biological system.

Therefore, the model under study, is presented in Ch3 and was investigated to study the sensitivity of over 53 outputs via one-at-a-time sensitivity measures. The time profile indicates that there are seven and eight kinetics in 0.1 and 0.2 dilution rates respectively and they are sensitive to the model response, where PSO algorithm is used to optimize the sensitivity result of 0.1 and 0.2 dilution rates with lower and upper bounds for each kinetic.

The validation of the identification result using the same model was performed and the sensitivity of replacing the original kinetic values by the kinetic optimization result was also done with good result according to the errors minimization that was proved in the validation with error minimization part.

In order to evaluate the optimization algorithm prediction ability deeply, we need to make some new predictions and compare it with new experimental data that are different from the data used for identification.

Finally, we found that the sensitivity analysis and PSO algorithm methods can achieve the large-scale kinetic parameters identification, it was proved in table (4.7) by clarify the errors percentage minimization.

CHAPTER 5

CONCLUSION

5.1 Introduction

The previous chapters have provided a review Metabolic Network, Local Sensitivity Analysis and Particle Swarm Optimization for metabolic network of *E. coli* to analyze and identify large-scale kinetic parameters in a dynamic biological metabolic system. In addition, PSO was used to minimize the errors that appear between the simulation and actual model output values. In this final chapter, the primary results of this investigation are summarized, and several key areas are listed as possible avenues of further research in this field.

This research aims at large-scale kinetic parameters issues of the metabolic network based on sensitivity analysis and optimization techniques. It even proposes a local sensitivity analysis and optimization algorithm for large-scale kinetic parameters in metabolic networks. This can be realized by introducing One-At-A-Time Sensitivity Measures in order to reduce the kinetic parameters and PSO algorithm to identify large-scale kinetic parameters of the metabolic network of *E. coli*. These techniques were employed to analyze and optimize the kinetics. In addition to estimating the kinetics, the result arising from the sensitivity analysis of 0.1 and 0.2 dilution rates as well as optimization algorithm implementation is the primary objective of this research. Based on the large-scale kinetic parameters, One-At-A-Time Sensitivity Measures and Particle Swarm Optimization Algorithm is proposed. The experiments show that, it could effectively analyze and identify the kinetics. The identified kinetics was found when

using the PSO as compared to the actual model result. The algorithm shows good result of kinetic identification.

This thesis proposes One-At-A-Time Sensitivity Measure and PSO algorithm for large-scale kinetic parameters in the metabolic network of *E. coli* as formulated by (Kadir et al., 2010) using the MATLAB simulation tool. The simulation results of the kinetic parameters showed a seven kinetics and the heights Mean percentage overall of 194 kinetics with 0.1 dilution rate; and eight kinetic parameters are the highest overall variance with 194 kinetics and dilution rate of 0.2 in addition to parameter identification of five metabolites which shows that the validation part moved closer to the experimental data.

In the experiments and simulation, the local sensitivity analysis and optimization algorithm were used. The research focuses on sensitivity analysis and optimizing the large-scale kinetic parameters. Then, we increased each kinetic parameter in percentage and quantified them by the Mean and the highest variance in 0.1 and 0.2 dilution rate respectively. The sensitivity result used in the PSO algorithm in order to be identified, the upper and lower values in the algorithm was used to adjust the searching in an allowable range with PSO parameters. Finally, our proposed methods deal with large-scale kinetic parameters identification. The result of the PSO algorithm was performed in the same model under study to prove that this algorithm is very good in the identification of large scale kinetic parameters.

This research provides the following novel contributions:

The first major contribution of the research has applied local sensitivity analysis technique named One-At-A-Time Sensitivity Measures applied to five pathways and quantifies the changes using the highest variance and the highest Mean.

The second major contribution is the minimizing of the errors that appear often between the experimental data and the actual model results by applying the particle swarm optimization algorithm.

The third important contribution is about the capability of our proposed methods in order to achieve large-scale kinetic parameters identification through the errors percentage changes which has done in the result validation part.

5.2 Future work

In order to investigate the influence of large-scale kinetic parameters in the metabolic network, the present methods can be applied. The sensitivity method is a local sensitivity analysis which lacks the ability to search out the sensitivity for all the kinetics simultaneously. We believe that, this method can be blamed as long as there are global sensitivity analysis methods. We also believe that, the present study give ideas on how to apply global sensitivity analysis method implement another optimization algorithm and propose new sensitivity analysis or new algorithm in term of large-scale kinetics in the future.

REFERENCE

A Berry, TC Dodge, M Pepsin and W Weyler. "application of metabolic engineering to improve both the production and use of biotch indigo." *Journal of Industrial Microbiology & Biotechnology* 28 (2002): 127 – 133.

A, Salman. "particle swarm optimization for task assignment problem ." *microprocessors and microsystems* (2002): 363-371.

Afnizanfaizal Abdullah, Safaai Deris, Mohd Saberi Mohamad and Siti Zaiton Mohd Hashim. "A New Particle Swarm Evolutionary Optimization for Parameter Estimation of Biological Models." *International Journal of Computer Information Systems and Industrial Management Applications* (2013): 571-580.

Al Zaid Siddiquee K, Arauzo-Bravo MJ, Shimizu K. "Metabolic flux analysis for a ppc mutant Escherichia coli based on 13C-labelling experiments together with enzyme activity assays and intracellular metabolite measurements." *Biotechnol* (2004): 407-417.

Alexander Kerna, Emma Tilleyb, Iain S. Hunterb, Matic Legisac and Anton Glieder. "Engineering primary metabolic pathways of industrial micro-organisms." *Journal of Biotechnology* 129 (2007)): 6–29.

Andreas Drager, Marcel Kronfeld, Michael J Ziller, Jochen Supper, Hannes Planatscher, Jørgen B Magnus, Marco Oldiges, Oliver Kohlbacher and Andreas Zell. "Modeling metabolic networks in C. glutamicum: a comparison of rate laws in combination with various parameter optimization strategies." *BMC Systems Biology* 3.5 (2009): 1-24.

Andreas Kremling, Katja Bettenbrock and Ernst Dieter Gilles. "Analysis of global control of Escherichia coli carbohydrate uptake." *BMC Systems Biology* (2007): 1-16.

Anton Miro, Carlos Pozo, Gonzalo Guillen-Gosalbez,JoseAEgea and Laureano Jimenez. "Deterministic global optimization algorithm based on outer approximation for the parameter estimation of nonlinear dynamic biological systems." (n.d.).

Anton Miro, CarlosPozo, Gonzalo Guill en-Gosalbez, JoseAEgea and Laureano Jimenez. "Deterministic global optimization algorithm based on outer approximation for the parameter estimation of nonlinear dynamic biological systems." *BMC Bioinformatics* (2012): 1-12.

Arkin, Steven S. Andrews and Adam P. "Simulating cell biology." *Magazine* 16 (2006): 1-5.

Asenjo AJ, Ramirez P, Rapaport I, Aracena J, Goles E and Andrews BA. "A discrete mathematical model applied to genetic regulation and metabolic networks." *Journal of Microbiology and Biotechnology* 17.3 (2007): 496-510.

Badowski, Tomasz. "Variance-based sensitivity analysis and orthogonal approximations for stochastic models." *Variance-based sensitivity analysis and orthogonal approximations for stochastic models*. 2013. 1-99.

Bastin, A. Provost and G. "Dynamic metabolic modelling under the balanced growth condition." *Journal of Process Control* (2004): 717–728.

Bournholdt, Stefan. "Less Is More in Modeling Large Genetic Networks." *Scinece Journal* (2005): 449-451.

Boyaci, Ismail Hakki. "A new approach for determination of enzyme kinetic constants using response surface methodology." *Biochemical Engineering Journal* (2005): 55–62.

C. Pozo, G. Guillen-Gosalbez, A. Sorribas, and L. Jimenez. "A Spatial Branch-and-Bound Framework for the Global Optimization of Kinetic Models of Metabolic Networks." *American Chemical Society* 50 .9 (2010): 5225–5238.

C.K. Chong, M. Mohd Saberi, D. Safaai, S. Shahir, W.C. Yee and E.C. Lian. "Aspartate Biosynthesis Pathway Simulation Using an Improved Differential Evolution Algorithm through Parameter Estimation." Terengganu, Malaysia: UMT, 2012.

Carlotta Martellia, Andrea De Martinob, Enzo Marinaric, Matteo Marsilid, and Isaac Perez Castillo. "Identifying essential genes in Escherichia coli from a metabolic optimization principle." *PNAS* 106 .8 (2009): 2607–2611.

Chaouiya, Claudine. "Petri net modelling of biological networks." *Briefings in Bioinformatics* (2007): 210-219.

Christophe Chassagnole, Naruemol Noisommit-Rizzi, Joachim W. Schmid, Klaus Mauch and Matthias Reuss. "Dynamic modeling of central metabolism of Escherichia coli ." *Biotechnology and Bioengineering* (2002): 53-73.

Christophe H. Schilling, Stean Schuster, Bernhard O. Palsson, and Reinhart Heinrich. "Metabolic pathway analysis: Basic concepts and scientific applications in the post genomic era." *American chemical society and american institute of chemical engineers* (1999): 296-303.

Chuii Khim Chong, Mohd Saberi Mohamad, Safaai Deris, Shahir Shamsir, Afnizanfaizal Abdullah, Yee Wen Choon, Lian En Chai, Sigeru Omatu. "Using an Improved Differential Evolution Algorithm For Parameter Estimation to Simulate Glycolysis Pathway." *Advances in Intelligent and Soft Computing* 151 (2012): 709-716.

Cleland, W.W. "The kinetics of enzyme-catalyzed reactions with two or more substrates or products: III. Prediction of initial velocity and inhibition patterns by inspection." *Biochimica et Biophysica Acta (BBA) - Specialized Section on Enzymological Subjects* (1963): 188–196.

Dellomonaco, Clementina. "Engineered Reversal of the beta oxidation cycle for the Synthesis of Fuels and Chemicals." *Nature* (2011): 355-359.

Edda Klipp, wolfarm liebermeister, christoph wierling, axel kowald, hans lehrach, and ralf herwig. *system biology*. John Wiley & Sons, 2009.

Edwards JS, Ibarra RU, Palsson BO. " In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. ." *Nat Biotechnol* (2001): 125–130.

Fell, Simon Thomas and David A. "Design of Metabolic Control for Large Flux Changes." *J. theor. Biol* (1996): 285–298.

Fonseca, E.O. Voit G. Goel I.-C. Chou L.L. "Estimation of metabolic pathway systems from different data sources." *System Biology* 3.6 (2009): 513–522.

Francisco G. Vital-Lopez, Costas D. Maranas, and Antonios Armaou. "Bifurcation analysis of the metabolism of E. coli at optimal enzyme levels." Minneapolis, MN: IEEE, 2006.

Geldermann, J., and O. Rentz. "Integrated Technique Assessment with Imprecise Information as a Support for the Identification of Best Available Techniques (BAT)." *OR Spektrum*, (2001): 137-157.

Gengjie Jia, Gregory N. Stephanopoulos and Rudiyanto Gunawan. "Parameter estimation of kinetic models from metabolic profiles: two-phase dynamic decoupling method." *BIOINFORMATICS* 27 .14 (2011): 1964–1970.

gregory N. stephanopoulos, aristos A. aristidou and jens nielsen. *metabolic engineering* . San Diego : elsevier science , 1998.

Griensven, A. V., et al. "A global sensitivity analysis tool for parameters of multi-variable catchment models." *Journal of Hydrology* 324.1-4 (2006): 10–23.

H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabasi. "The large-scale organization of metabolic network." *Nature* 407 (2000): 651-654.

Harry Ako, Lisa A. Shimeld, Christopher K. Mathews, K. E. Van Holde. *Study Guide for Biochemistry*. Pennsylvania State University: Benjamin/Cummings Publishing Company, 1996, 1996.

Heijnen, Diana Visser and Joseph J. "The Mathematics of Metabolic Control Analysis Revisited." *Metabolic Engineering* (2002): 114–123.

Heijnen, Diana Visser and Sef J. "The mathematics of metabolic control analysis revisited." *Metabolic engineering* (2002): 114-123.

Herve Monod, Cedric Nard and David Makowski. "Uncertainty and Sensitivity analysis for crop models." *Elsevier* (2006): 55-100.

Hoefnagel MHN, Starrenburg MJC, Martens DE, Hugenholtz J, Kleerbezem M, Van Swam II, Bongers R, Westerhoff HV and Soep JL. "Metabolic enginnering of lactic acid bacteria, the combined approach: kinetic modeling, metabolic control and experimental analysis ." *Microbiology* (2002): 1003-1013.

Hover, Joshua A. Taylor and Franz S. "Numerical optimization of generative network parameters." Ottawa, ON, Canada: Association Computing Machinery , 2010.

Ina Koch, Björn H. Junker and Monika Heiner. "Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber." *Bioinformatics* (2005): 1219-1226.

Ishii N1, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y, Nish. "Multiple high-throughput analyses monitor the response of E. coli to perturbations." *Science* (2007): 593-597.

J. Di Maggioa, J.C. Diaz Riccib and M.S. Diaza. "Global sensitivity analysis in dynamic metabolic networks." *Computers and Chemical Engineering* (2010): 770-781.

James C. Liao, Riccardo Boscolo, Young-Lyeol Yang, Linh My Tran, Chiara Sabatti, and Vwani P. Roychowdhury. "Network component analysis: Reconstruction of regulatory signals in biological systems." *PNAS* (2003): 15522–15527.

Jason L. Walther, Christian M. Metallo, Jie Zhang, Gregory Stephanopoulos. "Optimization of 13C isotopic tracer for metabolic flux analysis in mammalian cell." *Metabolic Engineering* 14 (2012): 162-171.

Jems C. Liao, Riccardo Boscolo, Young-Lyeol Yang, Linh My Tran, Chiara Sabatti, and Vwani P. Roychowdhury. "Network component analysis: Reconstruction of regulatory signals in biological systems." *PANS* (2003): 15522–15527.

Jeremy S. Edwards, Rafael U. Ibarra, and Bernhard O. Palsson. "In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data." *NATURE BIOTECHNOLOGY* 19 (2001): 125-130.

Joseph L. DeRisi, Vishwanath R. Iyer and Patrick O. Brown. "Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale." *SCIENCE* 278 (1997): 680-686.

K. Al Zaid Siddiquee, M. J. Arauzo-Bravo, and K. Shimizu. "Metabolic flux analysis of pykF gene knockout Escherichia coli based on 13C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations." *Appl Microbiol Biotechnol* (2004): 407–417.

Katsuyuki Yugi, Yoichi Nakayama, Ayako Kinoshita and Masaru Tomita. "Hybrid dynamic/static method for large-scale simulation of metabolism." *Theoretical Biology and Medical Modelling* (2005): 1-11.

Kauffman KJ, Prakash P, Edwards JS. "Advances in flux balance analysis." *Biotechnol* (2003): 491–496.

Ka-Yiu San, George N. Bennett, and Yea-Tyng Yang. "Genetic and metabolic engineering." *Elctronic Journal of Biotechnology* 1 (1998): 1-8.

Klamt S, Stelling J, Ginkel M, Gilles ED. "FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps." *Bioinformatics* (2003): 261-269.

Klaus Mauch, S. Arnnold and M. Reuss. "Dynamic sensitivity analysis for metabolic system." *Chemical Engineering Science* 52 (1997): 2589-2598.

Kleijnen, Jack P.C. "Sensitivity analysis and related analyses: A review of some statistical techniques." *Journal of Statistical Computation and Simulation* (1997): 111-142.

Koch I, Junker BH, Heiner M. "Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber." *Bioinformatics* (2005): 1219-1226.

Liao. C James, Riccardo Boscolo, Young-Lyeol Yang, Linh My Tran, Chiara Sabatt, and Vwani P. Roychowdhur. "Network component analysis: Reconstruction of regulatory signals in biological systems." *PNAS* (2003): 15522-15527.

Linh M. Tran, Matthew L. Rizk, and James C. Liao. "Ensemble Modeling of Metabolic Networks." *Biophysical Journal* (2008): 5606–5617.

M. Antoniotti, F.C. Park, A. Policriti, N. Ugel and B. Mishra. "Foundations of a query and simulation system for the modeling of biochemical and biological processes." *Bioiformatic and Computational biology Journals* (2003): 116-117.

Manfred Rizzi, Michael Baltes, Uwe Theobald and Matthias Reuss. "In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model." *Biotechnology and Bioengineering* (1997): 592–608.

Maria, Anu. "INTRODUCTION TO MODELING AND SIMULATION." *Winter Simulation*. China : Arena , 1997. 7-13.

Mauch, K., Arnold, S. and Reuss, M. "Dynamic Sensitivity Analysis for Metabolic Systems." *Chemical Engineering Science* (1997): 2589-2598.

Md. Aminul Hoque, Kotb Attia, Omar Alattas and Amir Feisal Merican. "Metabolic flux distribution and mathematical models for dynamic simulation of carbon metabolism in Escherichaia coli ." *African Journal of Biotechnology* 10.12 (2011): 2340-2352.

Merryn H. Tawhai, Alys R. Clark, and Kelly S. Burrowes. "Computational models of the pulmonary circulation: Insights and the move towards clinically directed studies." *Pulm Circ* 1.2 (2011): 224–238.

Mingshou Liu, Dongil Shin and Hwan Il Kang. "Parameter Estimation in Dynamic Biochemical Systems Based on Adaptive Particle Swarm." Macau: IEEE, 2009.

Mirny, Zeba Wunderlich and Leonid A. "Using the Topology of Metabolic Networks to Predict Viability of Mutant Strains." *Biophysical Journal* 91 (2006): 2304–2311.

Monika Maciag, Qariusz Nowicki, Laurent Janniere, Agnieszka Szalewska-Palaz and Grzegorz Wegrzyn. "Genetic responce to metabolic fluctuations: correlation between central carbon metabolism and DNA replication in Escherichia coli." *Microbial Cell Factories* (2011): 1-11.

N.Oltvai, Albert-László Barabási and Zoltán. "Network Biology: Understanding The Cell'S Functional Organization." *Nature Reviews Genetics* (2004): 101-113.

Nielsen, Andreas Karoly Gombert and Jens. "Mathematical modeling of metabolism." *Biotechnology* 11 (2000): 180-186.

Nikolaev, Evgeni V. "The elucidation of metabolic pathways and their improvements using stable optimization of large scale kinetic models of cellular systems." *Metabolic Engineering* 12.1 (2010): 26-38.

Noack, S., Wahl, A., Haunschild, M., Qeli, E., Freisleben, B., Wiechert, W. "Visualizing regulatory interdependencies and parameter sensitivities in biochemical network models." *Mathematics and Computers in Simulation* (2008): 991–998.

Nobuyoshi Ishii, Kenji Nakahigashi, Tomoya Baba, Martin Robert, Tomoyoshi Soga, Akio Kanai, Takashi Hirasawa, Miki Naba, Kenta Hirai, Aminul Hoque, Pei Yee Ho, Yuji Kakazu, Kaori Sugawar, Saori Igarashi, Satoshi Harada, Takeshi Masuda, Naoyuki

Sugiyama,. "Multiple High-Throughput Analyses Monitor the Response of E. coli to Perturbations." *Science* (2007): 593-597.

Ø.Palsson, Neema Jamshidiand Bernhard. "Metabolic Network Dynamics: Properties and Principles." *Understanding the Dynamics of Biologica lSystems* (2011): 19-37.

Oleg A. Igoshin, Albert Goldbeter, Dale Kaiser, and and George Oster. "A biochemical oscillator explains several aspects of Myxococcus xanthus behavior during development." *PNAS* (n.d.): 15760–15765 .

Palsson, Jennifer L. Reed and Bernhard Q. "Genome-scale in silico models of E. coli have multiple equivalent phenntypic states: assessment of correlated reaction subset that comprise network states." *Genome Research* (2004): 1797-1805.

Palsson, Jeremy S. Edwards and Bernhard O. "How will bioinformatic influence metabolic engineering." *Biotechnology and Bioinformatic* 58 (1998): 162-169.

Patnaik, R. and Liao, J. "Engineering of Escherichia coli central metabolism for aromatic metabolite production with near theoretical yield." *Appl. Environ. Microbiol* (1994): 3903-3908.

Perri, Alessandro Fasso and Pier Francesco. "Sensitivity analysis." *Encyclopedia of Environmetrics* 4 (2002): 1968-1982.

Pradeep K Polisetty, Eberhard O Voit and Edward P Gatzke. "Identification of metabolic system parameters using global optimization methods." *Theoretical Biology and Medical Modelling* (2006): 1-15.

R, Kennedy J and Eberhart. "Particle Swarm Optimization ." *IEEE International Conference on Nerual Network* . 1995. 942-1948.

Radhakrishnan Mahadevan, Jeremy S. Edwards, and Francis J. Doyle. "Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli." *Biophysical Journal Volume* 83 (2002): 1331–1340.

Radhakrishnan Mahadevan, Jeremy S. Edwards, and Francis J. Doyle,. "Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli." *Biophysical Journal* 83 (2002): 1331–1340.

Ralf Steuer, Thilo Gross, Joachim Selbig, and Bernd Blasius. "Structural kinetic modeling of metabolic networks." *PNAS* 103 (2006): 11868–11873.

Rocha, Sara Correia and Miguel. "In silico strain optimization by adding reactions to metbolic models." *Journal of Integrative Bioinformatics* (2012): 1-13.

Rui Xua, Ganesh K. Venayagamoorthyb, and Donald C. Wunsch. "Modeling of gene regulatory networks with hybrid differential evolution and particle swarm optimization." *Neural Networks* (2007): 917–927.

S. Noacka, A.Wahla, M. Haunschildb, E. Qelic, B. Freislebenc and W. Wiechertb. "Visualizing regulatory interdependencies and parameter sensitivities in biochemical network models." *Mathematics and Computers in Simulation* (2008): 991–998.

Salehi, F., S.O. Prasher, S. Amin, A. Madani, S.J. Jebelli, H.S. Ramaswamy, and C. T. Drury. "Prediction of Annual Nitrate-N Losses in Drain Outflows with Artificial Neural ." *Transactions of the ASAE* (2000): 1137-1143.

Saltelli, Andrea. "Sensitivity Analysis For Importance Assessment." *Risk Analysis* 22.3 (2002): 579–590.

Sam Vaseghi, Anja Baumeister, Manfred Rizzi, and Matthias Reuss. "In Vivo Dynamics of the Pentose Phosphate Pathway in Saccharomyces cerevisiae." *Metabolic Engineering* (1999): 128-140.

Sauer, Eliane Fischer and Uwe. "Metabolic flux profiling of Escherichia coli mutants in central carbon metabolism using GC-MS." *European Journal of Biochemistry* 270.5 (2003): 880–891.

Schilling, C.H., Edwards, J.S., and Palsson, B.O. "Towards metabolic phenomics: Analysis of genomic data using flux balances." *Biotechnol Prog* (1999): 288–295.

Segal, Irwin H. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. usa: willy, 1993.

Segel, Irwin H. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. California : Wiley Classic library Edtion , 1993.

Stefan Schuster, David A. Fell, and Thomas Dandekar. "A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks." *Nature Biotechnology* 18.3 (2000): 326-332.

Steffen Klamt, Jorg Stelling, Martin Ginkel and Ernst Dieter Gilles. "FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps." *BIOINFORMATICS* (2003): 261–269.

Stelling, Steffen Klamt and Jorg. "Combinatorial complexity of pathway analysis in metabolic networks." *Molecular Biology Reports* 29.1-2 (2002): 233-236.

Stephanopoulos, Gregory. "Metabolic fluxes and metabolic engineering." *Metabolic Engineering* (1999): 1-11.

thomas , simon and david a fellf. "Design of Metabolic Control for Large Flux Changes." *Theoretical Biology* (1996): 285–298.

Tuty Asmawaty Abdul Kadir, Ahmad A Mannan, Andrzej M Kierzek, Johnjoe McFadden and Kazuyuki Shimizu. "Modeling and simulation of the main metabolism in Eschercichia coli and its several single-gene knockout mutants with experimental verification." *Microbial Cell Factories* (2010): 1-21.

Varma A, Palsson BO. "Metabolic capabilities of Escherichia coli: I. Synthesis of biosynthetic precursors and cofactors." *Theoretical Biology* (1993): 477–502.

Wagner, Andreas. "Metabolic network and their evaluation." *Evolutionary Systems Biology* (2012): 29-52.

Wagner, D. A. Fell and A. "Structural properties of metabolic networks." *Nature Biotechnology* 18 (20): 1121-1122.

Wagner, David A. Fell and A ndreas. "The small world of metabolism." *NATURE BIOTECHNOLOGY* 18 (2000): 1121-1122.

Wright, Kathy R. Albe and Barbara E. "Systems Analysis of the Tricarboxylic Acid Cycle in Dictyostelium discoideum." *The Journal of Biological Chemistry* 267 (1992): 3106-3114.

Yin P. Y, Yu S. S, and Wang Y. T. "A hybrid particle swarm optimization algorithm for optimal task assignment in distributed system ." *computer standards and interfaces* (2006): 441-450.

Yoshihiro Toya, Nobuyoshi Ishii, Kenji Nakahigashi, Takashi Hirasawa, Tomoyoshi Soga, Masaru Tomita, and Kazuyuki Shimizu. "13C-metabolic flux analysis for batch culture of Escherichia coli and its pyk and pgi gene knockout mutants based on mass isotopomer distribution of intracellular metabolites." *Biotechnology Progress* (2010): 975–992.

Yoshihiro Usudaa, Yosuke Nishioa, Shintaro Iwatania, Stephen J. Van Dienb, Akira Imaizumib, Kazutaka Shimbob, Naoko Kageyamab, Daigo Iwahatab, Hiroshi Miyanob and Kazuhiko Matsuia. "Dynamic modeling of Escherichia coli metabolic and regulatory systems for amino-acid production." *Journal of Biotechnology* (2010): 17–30.

Yousuke Nishio, Yoshihiro Usuda, Kazuhiko Matsui and Hiroyuki Kurata. "Computer-aided rational design of the phosphotransferase system for enhanced glucose uptake in Escherichia coli." *Molecular Systems Biology* 4.1 (2008): 1-12.

Yukako Tohsato, Kunihiko Ikuta, Akitaka Shionoya, Yusaku Mazaki, and Masahiro Ito. "Parameter optimization and sensitivity analysis for large kinetic models using a real-coded genetic algorithm." *Gene* (2013): 84–90.

RESEARCH PUBLICATION

Mohammed Adam Kunna, Tuty Asmawaty Abdul Kadir, Aqeel S. Jaber, Rofilde Hasudungan. 2014. Large-scale kinetic parameters in metabolic network of *Escherichia coli* using local sensitivity analysis. *International Journal of Scientific & Engineering Research*, 5(10): 1299-1303.

Mohammed Adam Kunna , Tuty Asmawaty Abdul Kadir , Aqeel S. Jaber , Julius B. Odili . 2014. Large-scale kinetic Parameter Identification of Metabolic Network Model of *E. Coli* Using PSO. *Advances in Bioscience and Biotechnology,* accepted.

Mohammed Adam Kunna, Tuty Asmawaty Abdul Kadir, Alrasheed I. S, Aqeel S. Jaber. 2014. Large-Scale of Metabolic Network of *Escherichia Coli* using MATLAB. *International Journal of Computer Applications,* accepted.

Mohammed Adam Kunna, Tuty Asmawaty Abdul Kadir, Alrasheed I. S, Aqeel S. Jaber. 2014. Large-Scale of Metabolic Network of *Escherichia Coli* using MATLAB. *Majan College International Conference 2014*.

Mohammed Adam Kunna, Tuty Asmawaty Abdul Kadir. 2013. Sensitivity Analysis in Large-Scale of Metabolic Network of *E. Coli*. *International Conference on Advanced Computer Science Applications and Technologies*. 346-35.

Julius Beneoluchi Odili, Mohd Nizam Mohmad Kahar, Adam Kunna Mohammed and Anwar Shahid Safi. 2015. A Comparative study of African Buffalo Optimization and Randomized Insertion Algorithm for Asymmetric Travelling Salesman's Problem. *Congress on evolutionary computation (CEC2015)*.

APPENDIX A

The Rest of 40% Changes in Dilution Rate 0.1

miu_Max affection	original	changes	simulation	percentage
CELL concentration	1.5783	-0.095	1.6733	-6.02%
GLCex	0.022105	0.00406	0.018045	18.37%
G6P	0.20354	0.01275	0.19079	6.26%
F6P	0.021311	0.001071	0.02024	5.03%
FDP	1.4621	0.2558	1.2063	17.50%
GAPDHAP	0.31094	0.01362	0.29732	4.38%
PEP	1.4914	0.035	1.4564	2.35%
PYR	2.8117	0.3809	2.4308	13.55%
ACCOA	1.0018	0.0154	0.9864	1.54%
IsoCitrate	0.21101	0.04274	0.16827	20.25%
2KG	5.3724	0.1086	5.2638	2.02%
SUCCINATE	0.57217	-0.00141	0.57358	-0.25%
FUMARATE	0.35609	-0.00032	0.35641	-0.09%
MALATE	0.14263	0.0005	0.14213	0.35%
OAA	0.029637	0.000576	0.029061	1.94%
GOX	0.34577	-0.00186	0.34763	-0.54%
ACP	2.0199	0.2947	1.7252	14.59%
ACETATE	0.000209	4.33E-05	0.00016523	20.77%
6PG	0.017832	-1.9E-05	0.017851	-0.11%

Ru5P	0.02134	0.000252	0.021088	1.18%
R5P	0.07617	0.000982	0.075188	1.29%
Xu5P	0.026436	0.000366	0.02607	1.38%
S7P	0.004747	-1E-07	0.0047472	0.00%
E4P	0.027433	0.001926	0.025507	7.02%
miu	0.099617	-5.4E-05	0.099671	-0.05%
PTS	1.4003	0.0782	1.3221	5.58%
PGI	1.3	0.077	1.223	5.92%
PFK	1.3402	0.0767	1.2635	5.72%
ALDO	0.52536	0.11904	0.40632	22.66%
GAPDH	2.3756	0.1003	2.2753	4.22%
PYK	0.62509	0.0026	0.62249	0.42%
PDH	1.766	0.0434	1.7226	2.46%
CS	1.4682	0.0131	1.4551	0.89%
ICDH	0.93296	0.01144	0.92152	1.23%
2KGDH	0.40201	0.00193	0.40008	0.48%
ICL	0.51436	-0.0025	0.51686	-0.49%
MS	0.47975	-0.00235	0.4821	-0.49%
SDH	0.85922	-0.00045	0.85967	-0.05%
FUM	0.8237	-0.00041	0.82411	-0.05%
MDH	1.2698	-0.0028	1.2726	-0.22%
PTA	0.2504	0.03894	0.21146	15.55%
ACK	0.052391	0.010152	0.042239	19.38%
ACS	0.15652	0.03029	0.12623	19.35%
PCK	0.068774	-0.00022	0.06899	-0.31%
PPC	0.2702	0.01583	0.25437	5.86%
MEZ	0.019458	4.9E-05	0.019409	0.25%
G6PDH	0.079927	-7.3E-05	0.08	-0.09%
6PGDH	0.078143	-7.1E-05	0.078214	-0.09%
RPE	0.045416	-0.00018	0.045595	-0.39%
RPI	0.030597	8.4E-05	0.030513	0.27%
TKTA	0.022996	-1.1E-05	0.023007	-0.05%

TKTB	0.019783	-0.0002	0.019985	-1.02%
TAL	0.022522	-1.1E-05	0.022533	-0.05%
The total affection				113.61%

K_ALDOeq affection	original	changes	simulation	percentage
CELL concentration	1.5783	0.0074	1.5709	0.47%
GLCex	0.022105	0.000397	0.021708	1.80%
G6P	0.20354	-0.00111	0.20465	-0.55%
F6P	0.021311	-9.6E-05	0.021407	-0.45%
FDP	1.4621	0.1136	1.3485	7.77%
GAPDHAP	0.31094	-0.02567	0.33661	-8.26%
PEP	1.4914	-0.1151	1.6065	-7.72%
PYR	2.8117	-0.1529	2.9646	-5.44%
ACCOA	1.0018	-0.0143	1.0161	-1.43%
IsoCitrate	0.21101	-0.01919	0.2302	-9.09%
2KG	5.3724	-0.0002	5.3726	0.00%
SUCCINATE	0.57217	0.00631	0.56586	1.10%
FUMARATE	0.35609	0.0027	0.35339	0.76%
MALATE	0.14263	0.00086	0.14177	0.60%
OAA	0.029637	2.5E-05	0.029612	0.08%
GOX	0.34577	0.00538	0.34039	1.56%
ACP	2.0199	0.0165	2.0034	0.82%
ACETATE	0.000209	4.23E-06	0.00020431	2.03%
6PG	0.017832	-1.2E-05	0.017844	-0.07%
Ru5P	0.02134	-0.00018	0.021516	-0.82%
R5P	0.07617	-0.00068	0.076846	-0.89%
Xu5P	0.026436	-0.00025	0.026689	-0.96%
S7P	0.004747	0.000233	0.0045142	4.91%

E4P	0.027433	-0.0019	0.029328	-6.91%
miu	0.099617	-1E-06	0.099618	0.00%
PTS	1.4003	-0.0066	1.4069	-0.47%
PGI	1.3	-0.0065	1.3065	-0.50%
PFK	1.3402	-0.0064	1.3466	-0.48%
ALDO	0.52536	-0.12239	0.64775	-23.30%
GAPDH	2.3756	-0.0327	2.4083	-1.38%
PYK	0.62509	-0.0078	0.63289	-1.25%
PDH	1.766	0.0008	1.7652	0.05%
CS	1.4682	0.0005	1.4677	0.03%
ICDH	0.93296	-0.00421	0.93717	-0.45%
2KGDH	0.40201	-0.00433	0.40634	-1.08%
ICL	0.51436	0.00666	0.5077	1.29%
MS	0.47975	0.00612	0.47363	1.28%
SDH	0.85922	0.0017	0.85752	0.20%
FUM	0.8237	0.00144	0.82226	0.17%
MDH	1.2698	0.0074	1.2624	0.58%
PTA	0.2504	0.00261	0.24779	1.04%
ACK	0.052391	0.000972	0.051419	1.86%
ACS	0.15652	0.00291	0.15361	1.86%
PCK	0.068774	0.004794	0.06398	6.97%
PPC	0.2702	-0.00202	0.27222	-0.75%
MEZ	0.019458	8.4E-05	0.019374	0.43%
G6PDH	0.079927	-4.3E-05	0.07997	-0.05%
6PGDH	0.078143	-4.3E-05	0.078186	-0.06%
RPE	0.045416	7.5E-05	0.045341	0.17%
RPI	0.030597	-1E-04	0.030697	-0.33%
TKTA	0.022996	-3.3E-05	0.023029	-0.14%
TKTB	0.019783	0.000134	0.019649	0.68%
TAL	0.022522	-5.6E-05	0.022578	-0.25%
The total affection				111.55%

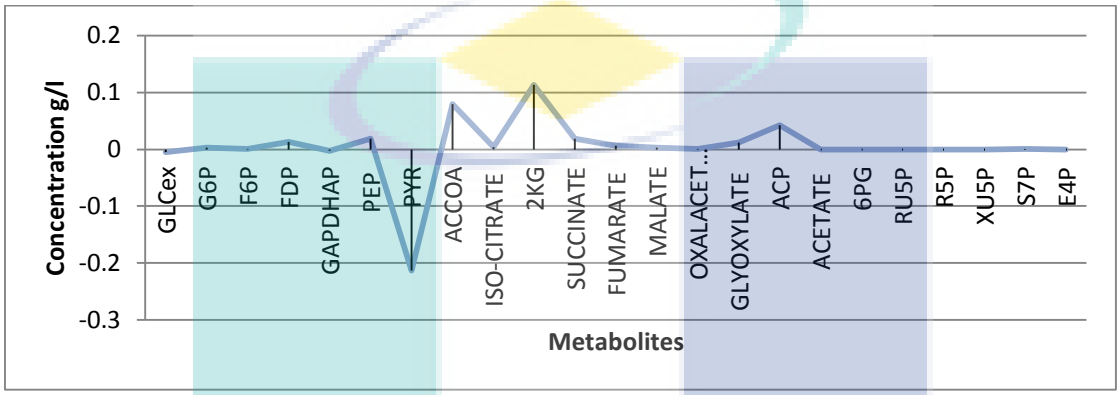
Km_PYKadp affection	original	changes	simulation	percentage
CELL concentration	1.5783	0.0539	1.5244	3.42%
GLCex	0.022105	0.001761	0.020344	7.97%
G6P	0.20354	-0.00825	0.21179	-4.05%
F6P	0.021311	-0.00069	0.021996	-3.21%
FDP	1.4621	-0.3616	1.8237	-24.73%
GAPDHAP	0.31094	-0.02742	0.33836	-8.82%
PEP	1.4914	-0.1272	1.6186	-8.53%
PYR	2.8117	0.0854	2.7263	3.04%
ACCOA	1.0018	0.00735	0.99445	0.73%
IsoCitrate	0.21101	-0.073	0.28401	-34.60%
2KG	5.3724	-0.0842	5.4566	-1.57%
SUCCINATE	0.57217	0.00659	0.56558	1.15%
FUMARATE	0.35609	0.00258	0.35351	0.72%
MALATE	0.14263	0.00026	0.14237	0.18%
OAA	0.029637	-0.0005	0.030141	-1.70%
GOX	0.34577	0.00598	0.33979	1.73%
ACP	2.0199	0.1237	1.8962	6.12%
ACETATE	0.000209	1.9E-05	0.00018952	9.12%
6PG	0.017832	-6E-06	0.017838	-0.03%
Ru5P	0.02134	-0.00029	0.021632	-1.37%
R5P	0.07617	-0.00113	0.077299	-1.48%

Xu5P	0.026436	-0.00042	0.026857	-1.59%
S7P	0.004747	0.000167	0.0045805	3.51%
E4P	0.027433	-0.00258	0.030017	-9.42%
miu	0.099617	-1.1E-05	0.099628	-0.01%
PTS	1.4003	-0.05	1.4503	-3.57%
PGI	1.3	-0.0492	1.3492	-3.78%
PFK	1.3402	-0.049	1.3892	-3.66%
ALDO	0.52536	-0.09897	0.62433	-18.84%
GAPDH	2.3756	-0.0229	2.3985	-0.96%
PYK	0.62509	0.06013	0.56496	9.62%
PDH	1.766	0.0015	1.7645	0.08%
CS	1.4682	-0.0117	1.4799	-0.80%
ICDH	0.93296	-0.01251	0.94547	-1.34%
2KGDH	0.40201	-0.00527	0.40728	-1.31%
ICL	0.51436	0.00775	0.50661	1.51%
MS	0.47975	0.00717	0.47258	1.49%
SDH	0.85922	0.00184	0.85738	0.21%
FUM	0.8237	0.00161	0.82209	0.20%
MDH	1.2698	0.0088	1.261	0.69%
PTA	0.2504	0.01679	0.23361	6.71%
ACK	0.052391	0.004404	0.047987	8.41%
ACS	0.15652	0.01317	0.14335	8.41%
PCK	0.068774	0.004121	0.064653	5.99%
PPC	0.2702	-0.01629	0.28649	-6.03%
MEZ	0.019458	2.6E-05	0.019432	0.13%
G6PDH	0.079927	-2.2E-05	0.079949	-0.03%
6PGDH	0.078143	-2.3E-05	0.078166	-0.03%
RPE	0.045416	0.000145	0.045271	0.32%
RPI	0.030597	-0.00014	0.030736	-0.45%
TKTA	0.022996	-2.6E-05	0.023022	-0.11%
TKTB	0.019783	0.000214	0.019569	1.08%
TAL	0.022522	-4.3E-05	0.022565	-0.19%

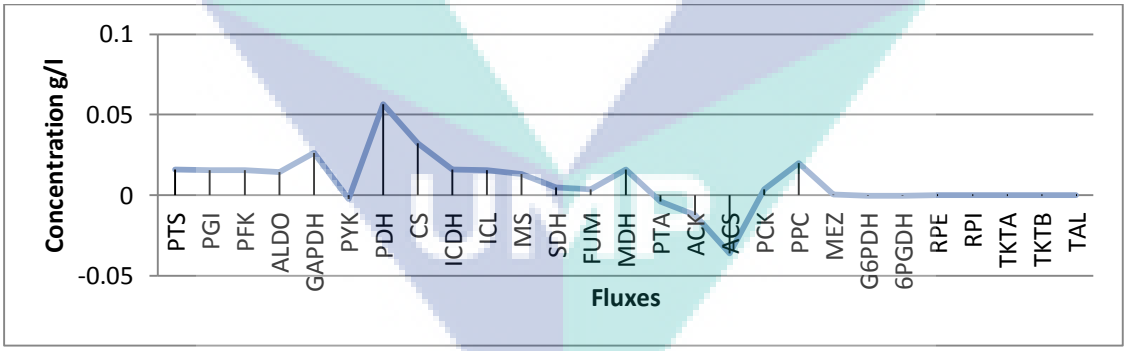
The total affection	224.78%

APPENDIX B

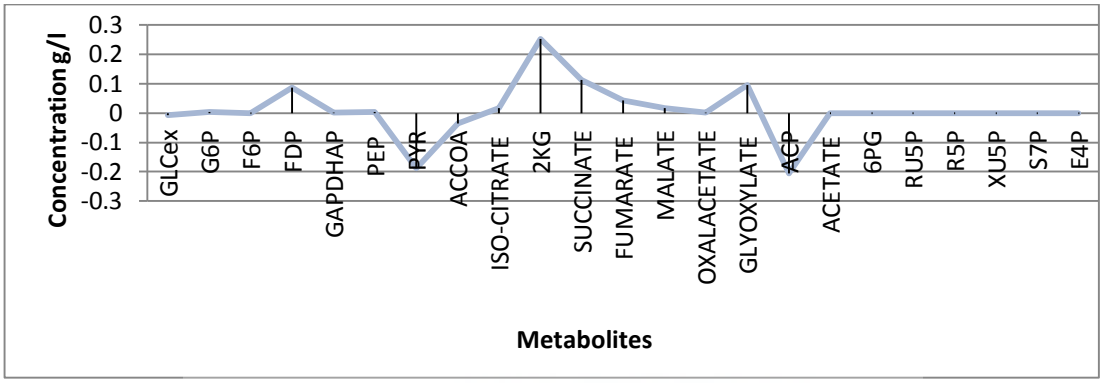
THE REST OF 20% CHANGES IN DILUTION RATE 0.2



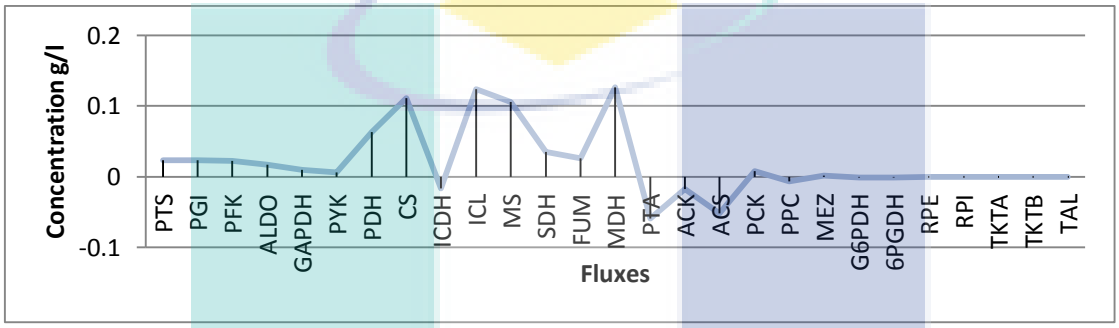
The kinetic interaction of *Ki_PDH* in metabolites



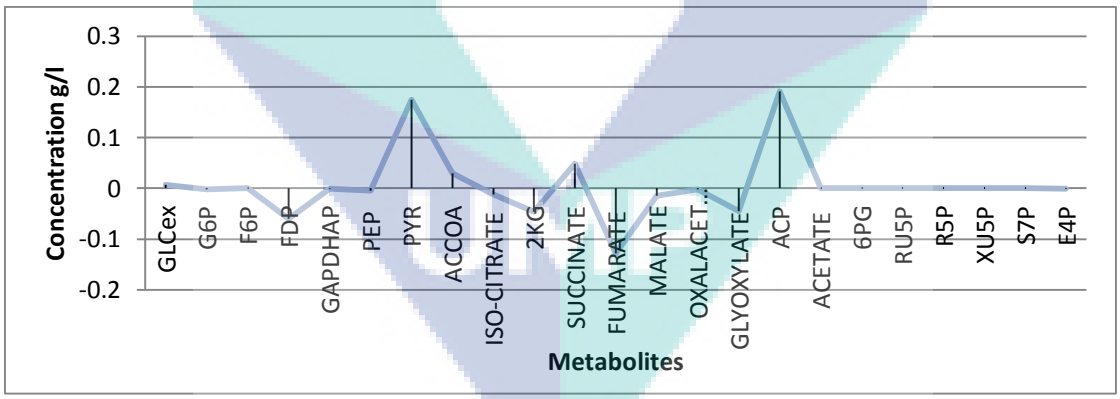
The kinetic interaction of *Ki_PDH* in fluxes



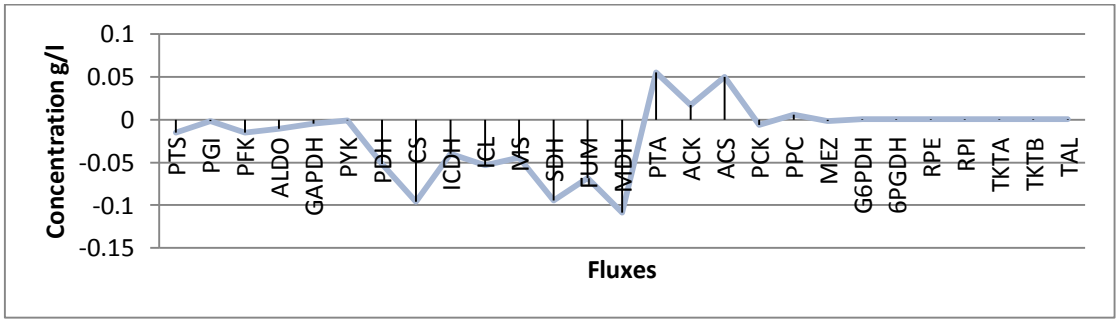
The kinetic interaction of *Kf_ICDH* in metabolites



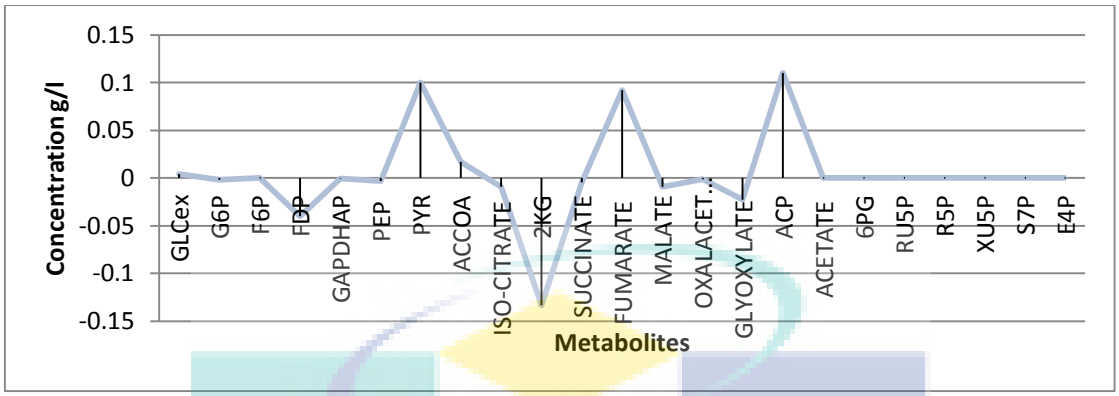
The kinetic interaction of *Kf_ICDH* in fluxes



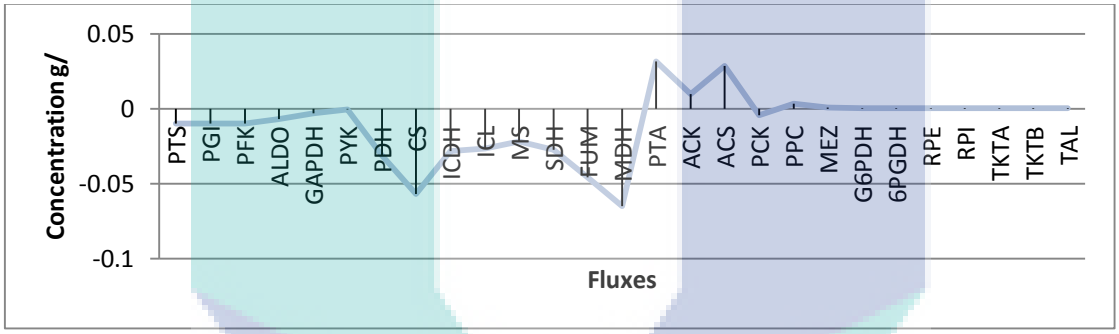
The kinetic interaction of *V_SDH in* metabolites



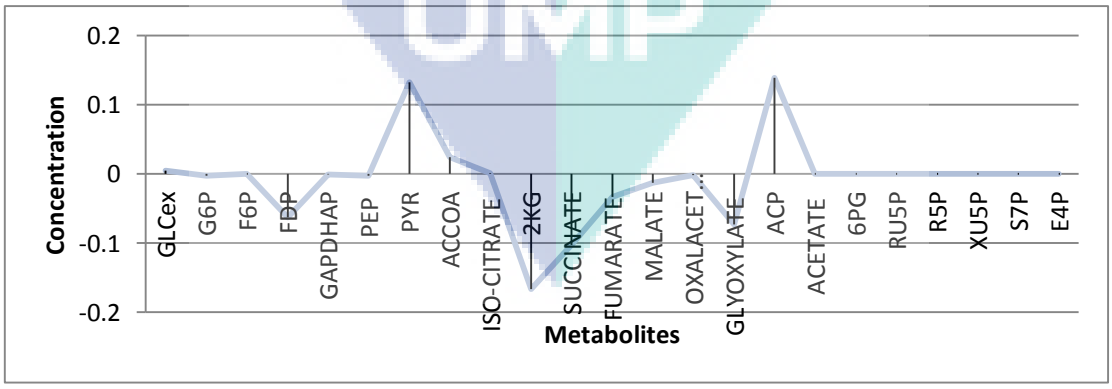
The kinetic interaction of *V_SDH* in fluxes



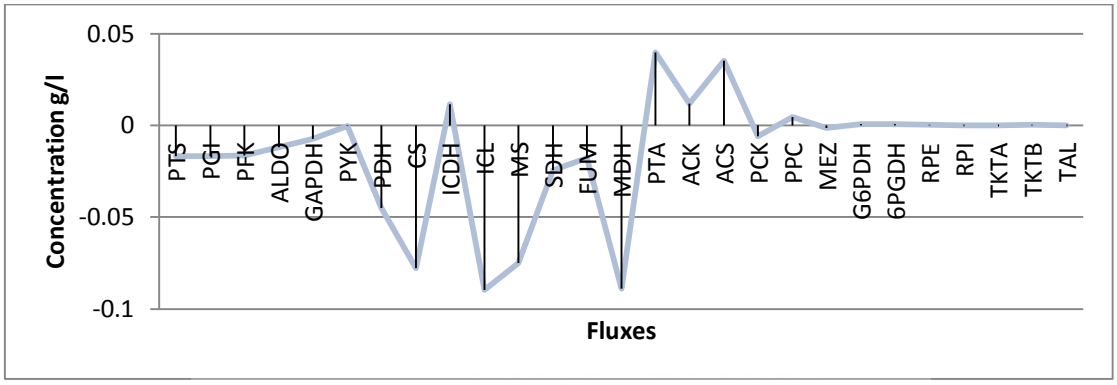
The kinetic interaction of *V_FUM* in metabolites



The kinetic interaction of *V_FUM* in fluxes



The kinetic interaction of *V_ICL* in metabolites



The kinetic interaction of *V_ICL* in fluxes