MODELLING AND SIMULATION OF LARGE SCALE STREPTOKINASE PRODUCTION USING *E. coli* AS A HOST

Ph.D. THESIS

by

PAVAN KUMAR

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247 667 (INDIA)
JANUARY, 2014

MODELLING AND SIMULATION OF LARGE SCALE STREPTOKINASE PRODUCTION USING *E. coli* AS A HOST

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

BIOTECHNOLOGY

by

PAVAN KUMAR

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247 667 (INDIA)

JANUARY, 2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **"MODELLING AND SIMULATION OF LARGE SCALE STREPTOKINASE PRODUCTION USING *E. coli* AS A HOST"** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from January, 2008 to January, 2014 under the supervision of Dr. Sanjoy Ghosh, Associate Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute/University.

(Pavan Kumar)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date: (Sanjoy Ghosh)
Supervisor

The Ph.D. Viva-Voce examination of **Mr. Pavan Kumar**, Research Scholar, has been held on ……

Signature of Supervisor Chairman, SRC External Examiner

Head of the Department/ Chairman, ODC

ACKNOWLEDGEMENT

I express my gratitude towards my supervisor Dr. Sanjoy Ghosh, Associate Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee, for his supervision and interest in my work.

I express my sincere thanks to Prof. R. Prasad, Head, Department of Biotechnology, for his kind attitude and understanding during the span of my research work.

I am equally thankful to my student research committee members, Dr. Partha Roy (DRC, Chairman) and Dr. Bijan Choudhury and Dr. C. B. Majumdar for their constructive comments during the various stages of my work. I express my sincere thanks to faculty members of the institute for their encouraging words, which contributed directly or indirectly in a significant way towards the completion of this thesis.

It is a pleasure to acknowledge the support extended by fellow research scholars in general Mr. Pankaj Pratap Singh, Mr. Naresh kumar, Mr. Kamlesh kumar, Mr. Vivekanand, Mr. Chinmoy, Mr. Sheetala Prasad, Mr. Raghu and Mr. Vijay in particular. I also have been lucky to have junior colleagues in the lab and friends in the institute who never hesitated to render their help whenever required.

I am extremely thankful to Dr. Gopal Raju, Dr. Mohd. Arif, Dr. Atul Tripathi, Dr. Manoj Tripathi, Mr. Rishi Raj, Mr. Puthooran, Mr. T.N. Mishra and Mr. Krishna Yadav for having spent their valuable time in constructive and countless discussions. I would like to acknowledge the co-operation and help rendered by Mr. Sandeep and Mr. Anil the staff of Biotechnology department, IIT Roorkee. I am equally thankful to the other departmental staff and institute staff for their support.

I cannot close these prefatory remarks without expressing my deep sense of gratitude and respect to my parents, Sri. Ram Lal Srivastava and Smt. Janki Devi for their blessings and endeavor to keep my moral high throughout the period of my work. I have no words to express the appreciation for my jijaji Sri. Ramesh Chandra Srivastava and sister Smt. Kamal Srivastava, and my sisters Smt. Sudha and Shashi Srivastava who inspired me to do research, who stood by me at every moment, kept encouraging me in low moments.

My deepest gratitude and sincere thanks to my elder brother Dr. Durgesh kumar and my bhabhiji Dr. Pankhuri Johari to help me in research and keep myself free from almost all the liabilities of home issues during my research work. Words can hardly explain the co-operation, immeasurable support and patience of nephews and nieces, Rashmi, Amit, Roopam, and to baboo Meenakshi Jyoti and Vinayak Shran who boosted me up, made me smile and encouraged me to do hard work. In addition, I want to emphasize the immense support of nephews and nieces, Shubham, Shreya, Naveen, Nitin, Sebum and Ravi. I am also thankful to all other family members at Gorakhpur, gaon-Charu, Gonda, Lucknow and Allahabad involved at various stages during my research work.

I want to express my sincere thanks to all those who directly or indirectly helped me at various stages of this work.

I pray to mother goddess for her blessings and to provide me mental and physical strength for successfully completing the entire work.

(Pavan Kumar)

ABSTRACT

The enzyme streptokinase is secreted by β-hemolytic *streptococcus* sp., which is often used to treat acute myocardial infarction and pulmonary embolism, being a vital drug it is required to be produced in recombinant form for thrombolytic therapy. This enzyme can be produced by using recombinant *E. coli* cells. The unstructured model factors were found to apparently influence the existence of active cells in the bioreactor environment. Our endeavour was to make out obligatory constraints that deal with the plasmid instability with an approach to apply composite model, which explained the relevant part of dynamics in bioreactor operation.

Computational models were developed utilizing structured and unstructured approaches. A range of dilution rates were selected starting from 0.1 to 0.65. On simulation of the process, the patterns obtained noticeably depicts the role of relevant parameters governing bioprocess system, particularly metabolite concentration and dilution rate, to present segregational instability and competitive dynamics in cell population. A set of parameters, including plasmid bearing cell population, plasmid lacking cell population, substrate concentration, metabolite concentration and probability of plasmid loss were taken into account. The idea was to measure the instability of plasmid, which could be directly derived from the growth of plasmid lacking cell population. This strategy ensures high flexibility in bioprocess modelling framework since it has a number of adjustable parameters.

Other bioprocess models were assumed to reveal the significance of dilutions and antibiotic concentration regulation during continuous culture. The structural machinery of a cell itself could assume to be an entire structured system that presented the functional role of various sub-cellular entities. The rate of failure of any cellular entity was found to be governed by prime metabolic events and partitioning phenomenon. Plasmid copy number dynamics trend was observed to evaluate the effect of metabolite concentration in time dependent manner. The copy number was estimated particularly after 2-6 hours of induction to understand its variability.

Firstly, the production media was statistically assessed using Plackett Burman design and later central composite design was used to estimate the interacting media components and culture condition factors. The four selected media components were put for CCD analysis and optimization. The production of streptokinase with optimized

medium and culture conditions was found up to 40% higher in magnitude in comparison to usual based conditions. An effective numerical system had been further considered using neural network and statistical method together where the prior one served as a potent tool for identifying and optimizing the output parameter. The statistical and neural network approaches were compared in predicting the output of different set of optimization systems; the later had revealed results that are more accurate. The different inputs of population dynamics simulation had been taken to neural network and prediction accuracy with high value of r^2 0.98 was achieved in estimation product formation. The production of highly valuable recombinant enzymes is being done using fermentation technology and in similar way the computational bioprocess methodologies can be used for its large scale production.

CONTENTS

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
ANOVA	Analysis of Variance
CCD	Central Composite Design
DCW	Dry Cell weight
DO	Dissolved Oxygen
FFNN	Feed Forward Neural Network
GA	Genetic Algorithm
GUI	Graphics User Interface
HNN	Hybrid Neural Network
IPTG	Isopropyl β-D-1-thiogalactopyranoside
$\mathbf{M_{CM}}$	Media Components
$\mathbf{M_{CD}}$	Media Culture Condition
$\mathbf{M_{IP}}$	Model inputs
OD	Optical Density
PB	Plackett Burman
RSM	Response Surface Methodology
SKC	Streptokinase
skc	streptokinase gene

GLOSSARY

	Concentration of two cell populations (g/l)
X_1, X_2	X_1- Concentration of plasmid bearing cells, X_2- Concentration of plasmid lacking cells
Y	Yield i.e., ratio of gram of cells formed and gram of substrate consumed
Y_M	Yield in terms of metabolite
M_{th}	Threshold concentration of metabolite that can effect population growth (g/l)
m_1	Metabolite toxicity coefficient for plasmid bearing cells
m_2	Metabolite toxicity coefficient for plasmid free cells
r_1	Selection stress coefficient
M	Metabolite concentration (g/l)
m_f	Collective probability factor
m_p	Probability factor associated to metabolite toxicity
r	Probability of plasmid retention due to selective stress
z	Probability of plasmid loss due to other structural or rare factors
p	Probability of plasmid loss
K_1	Monod coefficient for plasmid bearing cells (g/l)
K_2	Monod coefficient for plasmid lacking cells(g/l)
k_d	Decay constant for plasmid bearing cells (1/h)
k_p	Decay constant for product (1/h)
D	Dilution rate (1/h)
q	Probability of plasmid loss

S_0	Initial substrate concentration (g/l)
μ_1	Specific growth rate of plasmid bearing cells (1/h)
μ_2	Specific growth rate of plasmid free cells (1/h)
μ_{1max}	Maximum specific growth rate of plasmid bearing cells (1/h)
μ_{2max}	Maximum specific growth rate of plasmid free cells (1/h)
α	Growth rate of prey or population
β	Rate at which predators destroy prey
γ	Death rate of predators
δ	Rate at which predators increase by consuming the prey
N_1	Population of plasmid bearing cells
N_2	Population of plasmid free cells
τ_{pl}	Probability of losing one plasmid copy in division
ϕ	Intracellular parameter
$\mathbf{y}$	Extracellular state vector
x	Intracellular state vector
C_{a0} **to** C_d	Intra-cellular components (a-d) in a topological model
R_S	Existing system reliability
ξ	Failure rate of cellular subcomponent
fsu	Functional subunit
cfu	Colony forming unit
$€$	Error function
η	Learning rate
θ_p	Actual data with index *p*
ϕ_p	Predicted data with index *p*
N_X	Normalized concentration of data variable *x*

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

Streptokinase was the first thrombolytic drug, to be described and introduced for treatment of acute myocardial infarction and pulmonary embolism for more than forty years ago. It is now the leading fibrinolytic agent in the treatment of thrombo-embolic conditions and it is included in the World Health Organization Model List of Essential Medicines. The enzyme streptokinase is secreted by β-hemolytic *Streptococcus equisimilis*, a gram-positive bacterium. Being a vital drug of high demand it is required to be produced in recombinant form for thrombolytic therapy. Biochemically, streptokinase is regarded as an extracellular 47kDa glycoprotein consists of 415 amino acid residues. It interacts with plasminogen to form a stoichiometric complex, which activates plasminogen to plasmin, the later being the active form can degrade the fibrin matrix of blood clots (Wong *et al.*, 1994). Streptokinase has been widely used as thrombolytic agent since long. Its ability to induce reperfusion of the occluded coronary arteries and to reduce mortality has been firmly established.

Workers have over-expressed the streptokinase in recombinant *E. coli* (Thangadurai *et al.*, 2008). It is encoded by usually occuring *skc* gene of *Streptococcus* sp. in native form. Earlier the constructs were transformed into BL21 (DE3) and transformants grown in LB medium till 0.6 OD at absorbance wavelength of 600 nm. Then the cultures were induced with 1 mM IPTG at 37°C for 3h. The expression profile of the streptokinase samples were analyzed by resolving in 10 percent SDS and staining in Coomassie Blue G-250. Though *skc* gene is known to have many rare codons in its composition, overexpression was achieved instead of having some negative effect. The analysis of relative codon frequency of *skc* gene in *E.coli* reveals the presence of few position specific rare codons that affect the heterologous protein expression significantly. The specific growth rate decreased sharply upon induction of recombinant protein expression i.e. streptokinase, thus various feed profiles employed in post induction phase, with varying feed rates (Ramalingam *et al.*, 2007).

The *S. equisimilis* streptokinase gene expressed in *E. coli* has led to a ten-fold greater streptokinase titer than magnitude obtained in culturing of any other group C bacterium. Work had been reported on localizing the core promoter region of the

streptokinase gene as *skc* (Malke *et al.*, 2000). Usually recombinant streptokinase was produced from *E. coli* culture using the LB medium at 37°C from the expression of *skc* gene (Lee *et al.*, 1997). The plasmid used had previously imparted ampicillin resistance marker that provided the bacterium resistance to ampicillin in turn rendered the selection pressure for plasmid retention. The production of streptokinase was induced by adding 1.0mM IPTG to the medium. This enhanced the productivity of the recombinant protein and enabled its secretion into the extracellular medium. A radial caseinolysic method with the agarose gel containing both casein and plasminogen is commonly used (Saksela, 1981).

High cell density could be obtained with high cell density cultivation of recombinant cell system by using a proper feed strategy. The T7 RNA polymerase was induced with IPTG when biomass was 1.2mg DCW per ml and pulse feeding of concentrated substrate was done with different time intervals (Yazdani and Mukherjee, 1998). Likewise, an inducer concentration of 0.1mM IPTG was used while recombinant streptokinase accumulated to about 20% of the total soluble protein in the cell (Balagurunathan *et al.*, 2008). Different dilution rate was used in continuous culture for expression of streptokinase protein and found that a high dilution rate of 0.3 h^{-1} before and after induction helped in increasing the product concentration up to a level of 421 plasmin units per ml (Yazdani and Mukherjee, 2002). Changes in dilution rate during continuous culture would lead to an attainment of higher level of plasmid stability (Patnaik, 1995).

Streptokinase is a non-growth associated product since its production and cell growth are not linearly proportional. The culture of *E. coli* cells was routinely grown and maintained in LB medium and production medium used for batch fermentation contained in a medium contains ampicillin (100μg/ml). In performing streptokinase assay to test its activity in culture supernatant fluids, comparison of experimental samples with dilutions of standard purified streptokinase solution using casein/plasminogen plate technique was performed.

While conducting fermentation process in the bioreactor, recombinant cells are found to lose their plasmid. After some time as the fermentation proceeds two types of cell populations are to be developed. Since substrate is a growth-limiting nutrient factor so there starts a competition between two populations. The organisms carrying the plasmid are likely to be weaker competitors than one without because of the added load on its extra metabolic machinery (Lu and Hadeler, 1998).

The Lotka and Volterra model had significant relevance in bioprocess since chemostat used adopts the behaviour of theoretical ecology. The organisms carrying the plasmid used to compete with plasmid lacking cells for their existence in culture (Lu and Hadeler, 1998). There could be a number of factors that regulates the dynamics of plasmid carrying cells within the reactor. One such major factor was probability of plasmid loss due to segregation during cell division that could be described by segregative instability coefficient (Syamsu *et al.*, 1992).

The behaviour of the systems like bioreactor was found to be solely dynamic. Population dynamics model for plasmid bearing and plasmid lacking cells in bioreactor were being made more robust to develop an *insilico* dynamical system, which had the characteristics of a chemostat that several workers used to employ for streptokinase production. The representation of the dynamical system through modelling had its relevance in predicting the behaviour of the system on disturbances made to its initial conditions. The most significant consideration of the present day model is the challenge regarding instability of plasmid or plasmid loss with respect to time, which incorporates the effect generated from different means. Various other time invariants and intrinsic constraints together were now taken into account to study such responsible factors. A sophisticated model ensures a higher degree of flexibility since it has a number of adjustable parameters. Still efforts regarding the development of operational approaches imparting certain non-ideal bioreactor conditions are much preferable in this regard.

So, on the coexistence of two species which occurs in handling the recombinant cells population, an interaction occurs which may not be simply evaluated as of competitive type. The formation of recombinant product streptokinase is directly related to the population of plasmid bearing cells which often having decreasing trend with respect to process duration due to plasmid instability. Mortality is considered to lower the extent of inter specific competition and thereby promote the coexistence of competing species (Abrams, 2001). Preferably our endeavour is to bring up a developed model that has close resemblance to a natural chemostat in behaviour retaining the involved noticeable factors that seems to play some relevant role in dynamics.

Streptokinase production in bioreactor is associated to the development of bioprocess models pertaining to various aspects of the fermentation system. It is an established fact that plasmid lacking cells are found to emerge from the initial pool of recombinant cell population. This phenomenon leads to an undesired loss in yield of the

product. Primary metabolites, particularly acetic acid etc are formed as by-product of metabolism, has its influential role in regulating the competitive and inhibitory dynamics. Metabolites were found to play a crucial role since their threshold amount was assumed to promote the formation of plasmid lacking cells. On the other hand, probability of plasmid loss had shown to follow a variable trend in the model dynamics. In this study a process model with coupled differential equations is required to reveal the interaction among various parameters. Numerical simulation of delay and periodic model of partially recycled nutrient after extinction of cells due to decomposition during fermentation (Yuan and Zhang, 2012).

Our effort is to make out obligatory constraints that deals with the plasmid instability i.e., time dependent loss of plasmid. A brief review of probability of plasmid loss has been presented which has been shown to follow a variable trend in the model dynamics. The bioreactor is often used to gather information in form of data, regarding microorganism activity and bioprocess phenomenon in order to generate a mathematical model utilizing a set of culture parameters. Preferably, the endeavour was to develop a model that had close resemblance to a natural dynamic system that involved noticeable factors.

In producing the recombinant enzyme using bioprocess technique the forth most problem was the cease of production magnitude with progress in process duration of fermentation. Since the formation of product (streptokinase) would be directly proportional to the number of plasmid bearing recombinant cells in the media and also the average plasmid copy number present in individual cell at the point of time, so our emphasis was to consider plasmid copy number into our account and to study its variations. A stochastic model for plasmid copy number is studied (Seneta and Tavare, 1983). In continuous culture the increase in dilution rate to certain point seemed to induce a rapid decrease in plasmid copy numbers (Reinikainen and Virkajarvi, 1989). A method of rapid quantification of plasmid copy number is regarded to be an important process variable that can be used in the process control dynamics studies (Schmit *et at.*, 1996). Hence, subsequent determination of the copy number was very much essential and a kinetic model was required to analyse the decrease in plasmid copy number. In order to maintain a high copy number and high fraction of recombinant cells in the culture medium, strategy would be to optimize key parameters in production process of streptokinase. The magnitude of product formation was regulated through sustenance of recombinant cells fraction.

Specifically, the research work deals with the task pertaining to strength of plasmid bearing and plasmid lacking cell population, plasmid copy number and their dynamics. It also deals with media and culture condition optimization and its analysis. Several modelling strategies in bioprocess were found to play a relevant role in the production of valuable recombinant products. Modelling of plasmid stability is presently dealt in mechanistic way to support structured and unstructured models.

The unstructured model factors were found to be apparently influence the existence of active cells in the bioreactor environment. On the other side, the structural features of a cell itself can be assumed to be an entire structured system that presents the functional role of various sub-cellular entities. An approach was made to apply composite model, which explains the relevant part of dynamics in bioreactor operation. Genetic Algorithm simultaneously evaluates many points in the given parameter space and it often used to converge towards the global solution; hence it is helpful in identifying bioprocess parameters in a non-linear system (Ranganath *et at.*, 1999).

In the substrate inhibition kinetics of *Saccharomyces cerevisiae* in fed-batch cultures operated under the condition of constant glucose and maltose concentration level. The observed sugar inhibition effect in glucostat cultures was taken into account in modelling the growth kinetics (Papagianni *et al.*, 2007). In glucose grown culture there would be an effect of lactic acid accumulation which results in the lowering of pH, the acid tolerance response results into a varying amount of recombinant protein i.e., streptokinase produced (Sriraman and Jayaraman, 2006). The probability of plasmid loss in selective medium with difference in specific growth rate of recombinant and plasmid free cells (Zabriskie and Arcuri, 1986) were found to be affected by factors like genetic make-up of the host cells and the reactor operating parameters such as temperature, pH and growth medium composition. The selective pressure of selection medium is less effective due to the leakage of gene product (Sardonini and Dibiasio, 1987) which is responsible for selective mechanism. The effect of substrate feed concentration and forcing via dilution rate was examined on the performance of the bioreactor (Ali et al., 2012).

Computational models were developed utilizing structured and unstructured approaches. On simulation of the process, the patterns obtained noticeably depicted the role of relevant parameters governing bioprocess system, particularly metabolite concentration and dilution rate, to present segregational instability and competitive dynamics in cell population. A numerical system had been further designed and considered

using neural network and statistical method together where the goal was to optimize the input parameter. Now an improvement in the technology is required to increase the magnitude of enzyme production from the developed recombinant strain. A kinetic structured model for cell cultivation in reactor was assumed on basis of structuring biomass in two main groups including dividing and non-dividing cells (Klykov and Kurakov, 2012). Plasmid dynamics was studied in respect to generation time and half-elimination time of plasmids from recombinant cell population (Ganusov and Brilkov, 2002).

Secretion rates of microbial products can be affected by cell-cycle position of individual cells and thus on the degree of heterogeneity in terms of intracellular constituents (Henson, 2003). Growth conditions have profound effect over plasmid stability as well as specific growth rates of two cell populations (Stephanopoulos and Lapidus, 1988). On the other hand formation and release of by-products may have its profound impact over the functional subunits involve in expression and translation pathway, it results into decline in the reliability of functional metabolic subunits and hence ceases the production.

The computational intelligence techniques including regulatory control are of major help in artificial framing and dynamics associated to bioprocess (Wang *et al.*, 2010)**.** Cellular intelligence is now-a-days a most excellent tool to amplify the information taken from cellular level activities in order to utilize it for higher level computational applications in prospects of getting automated (Nicoletti *et al.*, 2009). A highly constituted model framework presents designing of composite architecture that optimally blends cellular and artificial intelligence as well as mechanistic models (Patnaik, 2009). The signal oriented modelling is remarkably an approach for utilizing the intra and inter cellular level signals to emphasize over the fundamentals of structured model. The use of regulatory control enable the models based on cells to utilize information gained from experience and thereby respond intelligently to external environmental stimulating factor.

The rationale of structural instability is described to be manifold (Summers *et al.*, 1993). Insufficient repair mechanisms are further reason for mutations, which results into failure of expression. It is a fact that generally quite complex mechanistic models are required to adequately describing the metabolic dynamics of multi-cellular systems especially under non-ideal conditions. It is evident that cells have internal regulatory control to govern all biochemical pathways in a legitimate manner. Hence, it coordinates

and directs the adaptive machinery to cope up with external or extra-cellular variations maintaining the supportive mechanism that operates to serve simultaneously. This instrumental work can only be made feasible employing stochastic approaches, which easily support to constitute the variable pathway types that can be evaluated using simulation method. The identification of a key functional unit within the network of pathways is a vital task to complete for providing a minimal framework to our model. An optimal and reliable representation of the overall model using integrated approach for bioprocess design (Rouf, 1999) is a mandatory task in this direction.

It was observed that after 60h of continuous cultivation in LB medium without ampicillin the stability of plasmid ceases to zero i.e., the entire culture contained only plasmid free cells. On using ampicillin concentration of $100mgl^{-1}$ the stability was found to drop upto 60% and did not significantly cease further (Friehs, 2004). All the model parameters of plasmid replication control can be obtained independently and no adjustable parameters are needed (Ataai and Shuler, 1987). The accuracy of simulations allows description of the complex interaction between substrate, microorganism and metabolite, acetate, which is a product and also a carbon source for the microbes (Reinikainen and Virkajarvi, 1989).

In microbial culture, assay with different inoculums ages and different concentrations were performed, the biomass obtained was analysed in terms of their chemical composition (Pelizer *et al.*, 2003). The *Pischia pastoris* clone producing streptokinase was optimized and the effects of carbon and nitrogen sources were observed. Response surface methodology has been widely used to evaluate and understand the interactions between different process parameters. Two level Plackett and Burman design was used for the screening of carbon sources. The present study was aimed at screening of the important carbon and organic nitrogen sources with respect to their main effects and not the interaction effects between various medium constituents and hence such design was used for the screening of different carbon sources (Vellanki *et al.*, 2009). A response surface methodology together with central composite rotator design was also employed to optimize the fermentation medium for Nattokinase production by *Bacillus subtilis*. (Deepak *et al.*, 2008) Streptokinase activity of culture supernatant fluids was estimated by comparison with dilutions of standard purified streptokinase solution using casein-plasminogen plate technique (Malke and Ferretti, 1984).

Media optimization for the strain of *E. coli* BL21, which was taken as a host organism with plasmid vector pRSET-B had not been done yet. From the literature survey, it was clear that the parameters pertaining to culture conditions were not yet optimized for the same. Consequently, the optimization of various media constituents and condition parameters were to be done. In order to achieve an optimal level of production primarily the effort was made to optimize the production media components and culture conditions statistically. Statistical methods were helpful in minimizing the number of experiments required to test the combination of factors for best output. In statistical technique screening design should be carried out to determine which of the several experimental variables and their interactions present more significant effects (Bezerra *et al.*, 2008). ANOVA was constructed for the second order response surface model; the significance of each coefficient was determined by Student's *t*-test and *p*-value to identify the corresponding significant factors. A mathematical method was used to solve the regression equation (Vellanki *et al.*, 2009).

So, statistical techniques viz. PB and CCD were employed to screen the parameters and to optimize those in successive experiments. A mathematical method had been used to compute the regression and correlation among output data. Optimization would be the first major step before carrying the bioreactor operations. In next step, streptokinase production had to be optimized using appropriate conditions under process operation. Finally, fermentation would be done employing the microorganism based on maximum yield of streptokinase providing most appropriate conditions for growth of cells and plasmid retention criteria.

Building ecological model utilizes many methods, ranging from numerical, mathematical and statistical methods to techniques originating from artificial intelligence, like neural networks. A multilayer feed forward neural network with multi-layer perceptron, is very popular and it is more commonly used than any other neural network types (Lek and Guegan, 1999). Feed-forward ANNs, modified to include dynamic characteristics and form the basis of estimator models. ANNs, serve as a powerful tool for non-linear modelling and process control (Glassey *et al.*, 1994). Use of about more than hundred iterations was enough for successful convergence of the error under the conditions employed. Obtaining the observed data and simulated results were close, suggests the superior capability of neural network modelling for the dynamic behaviour of the system (Horiuchi *et al.*, 2001). The ANN model showed a better correlation with the experimental

values than the regression model (Haider *et al.*, 2001). Various training procedures proposed for the hybrid modelling approach is generally based on gradient optimization method, so performing repeatedly the exploration analysis with different start vectors of the ANN weights should be taken and the network led to minimal “test error” is to be adopted (Simutis and Lubbert, 1997) (Simutis and Lubbert, 1997).

The Neural Network based simulation of the process was done using Matlab 2008Ra, the predicted values obtained from neural network had been taken into account together with the statistically predicted output. Neural network modelling was applied for control of fed-batch like process and to evaluate the bioprocess performance under non-ideal conditions (Patnaik, 2004). The previously developed model for fed-batch culture in respect to streptokinase production has remarkably depicted the sole dynamic behaviour of the bioprocess system (Patnaik, 1995) (Patnaik, 2002). Online estimators for biomass and recombinant protein concentration were constructed using information available online by the application of Neural Network (Glassey *et al.*, 1994). The ANN has been proved to be a useful tool for model building; there was a striving need to improve bioprocess operability dealing with the large scale industrial fermentation systems (Massimo *et al.*, 1991).

Significant attempts were made to configure a composite model to represent the overall dynamics in a well-defined algorithm that depicts the behaviour of the microbial population in the entire bioreactor operational environment. The estimators of instability were utilized to estimate some of the aspects of fermentation process, which may lead to an improved supervision. The simulation of the process had been done using numerical method to evaluate and predict the model behaviour using neural network approach.

Pilot-scale bioreactors differ from small laboratory-scale reactors in terms of a greater occurrence of noise and incomplete mixing of the broth. Conventional control tries to induce good mixing and to filter out the noise as completely as possible. As such an ideal operation is difficult to achieve, recent work has tried to exploit the non-ideal features to improve the performance using computational heuristics approaches. The bioreactor is often used to generate online data of a set of parameters from the microbial culture to constitute a mathematical model. Other bioprocess models had been incorporated to the work suggesting the means to exploit other domains of biochemical technology for developing a more worthy approach. The structural features are assumed to model the productive span of cells in the bioreactor system.

In this research the work focus would be on the modelling of production of recombinant streptokinase from the bioreactor. Another approach would be the use of heuristic approaches that might be a most desirable and ultimate step in the direction of improving the model output. A novel multidimensional approach was used to improve the robustness of the problem of data management, and was supported by computational and experiment framework. Moreover the idea was to exploit the features concerning to such dynamics, particularly plasmid vector stability that indirectly governs the product formation.

CHAPTER-2

POPULATION DYNAMICS MODEL

2.1. An Overview

The production of streptokinase in bioreactor can be well depicted as cell population dynamics. It is an established fact that two types of cell populations are found to emerge from the initial pool of recombinant cell population. This phenomenon leads to an undesirable loss in yield of recombinant product. Primary metabolites are formed as by-product of metabolism, has its influential role in regulating the competitive dynamics. On the other hand, probability of plasmid loss has shown to follow a variable trend in the model dynamics. In this study a process model with coupled differential equations is designed to reveal the interaction among various system parameters.

The population dynamics model in context of instability of plasmid during fermentation were taken elaborately. In present configured model the coefficients were either taken from the earlier known classical models or may have been derived using multiple linear regression technique taking standard validated values from simulation in respect to experimental results. The five parameters, including a) plasmid bearing cell population, b) plasmid lacking cell population, c) substrate concentration, d) metabolite concentration and e) probability of plasmid loss were taken into account. The idea was to evaluate the instability of plasmid, which can be directly derived from the growth of plasmid lacking cell population. Mortality was another indispensable factor, which was found to lower the extent of inter-specific competition and thereby promote the coexistence of competing species (Abrams, 2001).

The toxicity of metabolite equally harm both the population simultaneously but since plasmid bearing cells are liable to lose their plasmid in response, so plasmid free cell population was strengthened during the time. So, the plasmid free cell population would have an increasing trend despite of decreasing due to this event. The probability of plasmid loss was not constant throughout the fermentation process due to the formation of metabolites, which showed its presumed toxic effect after certain threshold concentration, so it was taken as variable parameter in the model with respect to time. A criterion of threshold policy could be implemented to evaluate collective probability factor which influenced the time dependent variation of probability. Also mortality or formation of

dormant cells was possible which was due to toxification developed on account of toxic metabolite or by-product formation, like acetic acid formation in case of fermentation process carried for streptokinase production.

The model was simulated at different dilution rates and different initial substrate concentration and thereby differences in the model behaviour had been observed on simulation. There was a remarkable decline in the percentage of plasmid bearing cells above a certain level of dilution rate. System had shown its prevalent sensitivity to change in dilution rate. The probability of plasmid loss had shown a gradual trend at high dilution rates. The numerical simulation of the model equations showed plasmid loss tends to occur after a certain limit of metabolite level, which shows that the increasing concentration of metabolite tends to support the population of plasmid lacking cells that witnessed the trend in loss of plasmid from recombinant cells. Simulation was done using most of the standard values from existing models and assumed constraints required to explicate this dynamical system. Since the production of streptokinase was directly depending upon the dynamics of plasmid bearing cells, so to enhance the production it was inevitable to reveal the kinetics operating factors and to evaluate the instability of the plasmid.

2.2. Model Development Parameters

2.2.1. Segregational Instability

Segregational plasmid stability would have significant influence on the production of heterologous proteins since it ensures the degree of sustenance of plasmid. Reportedly stabilization machineries found on natural plasmids which somewhat control the partitioning of plasmids on to daughter cells (Friehs, 2004). The partitioning phenomenon depends on mechanisms of segregational plasmid stability, which is governed by several factors including, plasmid size and form, high copy number and plasmid distribution, difference in specific growth rate by cell internal factors, post segregational killing of plasmid free cells etc. Possibly our effort was to bring up a model that incorporates noticeable factors that seems to play some relevant role in dynamics of the vector.

2.2.2. Dilution Rate

In a continuous culture, the dynamics of the streptokinase was observed at different dilution rates. Simulation was performed for a range of dilutions. Thus attempt had been taken to investigate the bioprocess potential keeping a dynamic view into account.

2.2.3. Substrate Concentration

After some time as the fermentation proceeds, two types of cell populations were obtained. Since substrate was a growth-limiting factor, there started a competition between two populations. The organisms carrying the plasmid or plasmid bearing cells were likely to be weaker competitors than one without because of extra load of vector machinery; on the other hand, plasmid carrying cells had an advantage of selection stress due to antibiotic in the medium. The production media comprising of carbon, nitrogen and salt concentration was used and its characteristic consumption with either type of population during bioprocess was observed simultaneously.

2.2.4. Metabolite Concentration

Various indispensable factors served as significant factors in influencing the model. A threshold criterion was implemented for toxicity effect generated from primary metabolite for the growth of cells. Their effective concentration could bring mortality or sluggishness in growth, moreover it used to facilitate the cells to carry out and enhance the state of plasmid instability.

2.3. Previous Population Dynamics Models

The earliest models for population dynamics was proposed earlier by Lotka and Volterra (Bailey and Ollis, 1986) which interpreted the population interaction for a prey-predator like system. The folowing relation gives the dynamics,

$$\frac{dA}{dt} = \alpha A - \beta AB \quad (1)$$

$$\frac{dB}{dt} = -\gamma B + \delta AB \quad (2)$$

In the above expression, α growth rate of prey or population A, β is the rate at which predators or population B, destroys prey, γ is the death rate of population B and δ is the rate at which predators increase by consuming the prey. The above model Equations (1) and (2) has good relevance in bioprocess since chemostat adopts the behaviour of theoretical ecology. A microbial ecosystem was considered as a functional entity characterized by certain macroscopic measurements such as the total quantity of biomass or the total number of cells in the medium. It is possible to work with rapid growth of species in well controlled environments, such as ‘‘chemostat” (Harmand et al., 2008). During the culture and production of micro-organisms, the control of the bioprocess sometimes depends on the micro-organism concentration or the biomass density and

conditions provided during that time period, for instance, in case some aerobic microbial community the Dissolved Oxygen (DO) is a key factor to their growth (Guo and Chen, 2009).

The expression profile of the streptokinase samples were analysed by workers (Thangadurai et al., 2008) and the specific growth rate noticeably found to decrease upon induction for recombinant protein expression (Ramalingam et al., 2007). While performing fermentation process in the bioreactor since recombinant cells are found to lose their plasmid, after some time as the fermentation proceeds two types of cell populations are developed. Since substrate is a growth-limiting factor so there starts a competition between two populations. The organisms carrying the plasmid or plasmid bearing cells are likely to be weaker competitors than one without because of the added load on its metabolic machinery (Lu and Hadeler, 1998).

There could be a number of factors, which regulate the dynamics of plasmid carrying cells within the reactor. One such major factor was probability of plasmid loss due to segregation during cell division that could be described by segregative instability coefficient (Syamsu et al., 1992). On the other hand, the likelihood of segregation for a plasmid bearing cell is independent of the frequency of plasmid-free cells generation in the population (Lenski and Bouma, 1987) while the selection intensity against plasmid carriage is influenced by occurrence of plasmid lacking cells in population. In cell division process occasionally a daughter cell results that doesn't contain the plasmid and can no longer produce the desired product (Stephens et al., 1992). Plasmid free cells being faster in adapting to environmental changes, it is indeed possible to give a competitive edge to a plasmid containing population through cycling of dilution rate. But the periodic variation of dilution rate noticeably proved to have insignificant role in improving the performance of biomass production in a continuous process (Stephens et al., 1992).

The probability of plasmid loss in selective medium and difference in specific growth rate between recombinant and plasmid free cells (Zabriskie and Arcuri, 1986) are affected by factors like genetic make-up of the host cells and the reactor operating parameters such as temperature, pH and growth medium composition. Moreover, the selective pressure of selection medium is less effective due to the leakage of gene product (Sardonini and Dibiasio, 1987), which is responsible for selective mechanism. The

continuous system of production often serves as a better choice even in the packed column reactor for production of antibiotics (Banerjee and Debnath, 2007).

A chemostat model of competition could be established between plasmid bearing and plasmid free organism for a single nutrient where plasmid-bearing organism can produce toxins against plasmid free organism at same cost to its reproductive abilities (Hsu et al., 2000). Lenski and Hattingh (Lenski and Hattingh, 1986) studied the effect of an inhibitor on two populations. They considered the degree of inhibition in presence of inhibitor or toxicant on growth rate. It was studied that estimation of unmeasurable biological variables was important in fermentation process, directly influencing the optimal control performance of the fermentation system as well as quality and yield of the targeted recombinant product. Application of some novel strategy for state estimation of fed-batch fermentation was suggested (Wang et al., 2010).

The input substrate concentration and dilution rate serve as operating parameters and these are to be controlled by the experimenter. The study for the cases where nutrient supplied at constant rate and time dependent manner were performed earlier and a delay in the growth response of organism to nutrient uptake was obtained. Varying feed profiles were employed in the post-induction phase of recombinant streptokinase protein expression, including constant feed rates, linearly increasing feed rate and exponentially varying feed rates (Ramalingama et al., 2010) to evaluate the requirement of variable feed strategy. Modulation of an input such as a substrate or a nutrient concentration or the cell environment such as the pH can enhance the rates of biochemical reactions that were occurring (Silveston et al., 2008). A twofold increase in the concentration of plasmid bearing cells using square wave modulation of the dilution rate had been found. The use of altering dilution rate for the same environment setting could revert competitive exclusion to species coexistence (Costa et al., 2006).

According to model proposed by Imanaka and Aiba (Imanaka and Aiba, 1981) for a continuous culture the two types of population of cells, plasmid bearing and non bearing cells having interaction given by the following expression, Eqs. (3) and (4):

$$\frac{dN_1}{dt} = \mu_1 N_1 - DN_1 - p\mu_1 N_1 \tag{3}$$

$$\frac{dN_2}{dt} = p\mu_1 N_1 + \mu_2 N_2 - DN_2 \tag{4}$$

Where, N_1 population of plasmid bearing cells and N_2 population of plasmid free cells. The above model Equations (3) and (4) considered, is purely of exploitative competition since no toxin is assumed to be formed. Another model was given by Stephanopaulis and Lapidus (Stephanopoulos and Lapidus, 1988) which is very close to the model given earlier researchers (Imanaka and Aiba, 1981), it includes three parameter variables plasmid bearing and plasmid lacking cell populations and substrate concentration which is limiting, expression shown in set of mass balance Eqs. (5) - (7):

$$\frac{dX_1}{dt} = \mu_1 X_1(1-q) - DX_1 \quad (5)$$

$$\frac{dX_2}{dt} = \mu_2 X_2 + q\mu_1 X_1 - DX_2 \quad (6)$$

$$\frac{dS}{dt} = D(S_0 - S) - \frac{1}{Y}(\mu_1 X_1 - \mu_2 X_2) \quad (7)$$

$S_0 \geq 0, X_i(0) > 0, i = 1, 2.$

2.4. New Model Development

2.4.1. Model Assumptions and framework

In case selective media is considered, a relevant factor selection stress coefficient can be taken into account, which is favourable for plasmid bearing cells (Sardonini and Dibiasio, 1987) while it does not favour the plasmid free cells. Since it favours the population of plasmid bearing cells so this parameter would be considered to resist the phenomenon of plasmid loss. Therefore, probability of plasmid loss in selective medium is smaller than that in non-selective medium.

During the process of fermentation metabolites formation occur (Lee and Papoutsakis, 1999) as by-products of metabolism which are toxic (Stuebner et al., 1991) and inhibit the growth of both types of cells to different extent. The plasmid bearing cells are likely to lose their plasmid because of the permeability of such metabolites into the cells from the environment while other type that is free of plasmid do not have that much extent of harm. In other words this can be said that toxicity of metabolite equally harm both the population simultaneously but since plasmid bearing cells are liable to lose their plasmid in response, so plasmid free cell population is strengthening during the same time.

The probability of plasmid loss is not constant throughout the fermentation process due to the formation of metabolites, which shows its presumed toxic effect after certain threshold concentration, so it needs to be taken as variable parameter in the model

with respect to time. A criterion of threshold policy can be implemented to evaluate collective probability factor which influence the time dependent variation in probability. Also mortality or formation of dormant cells is possible which is due to toxification developed on account of toxic metabolite or by-product formation, like acetic acid formation in case of fermentation process carried for streptokinase production. Therefore, considering the above assumptions we may write the dynamic model as:

$$\frac{dX_1}{dt} = \mu_1 X_1(1-q) - DX_1 - m_1\mu_1 X_1 + r_1\mu_1 X_1 \quad (8)$$

$$\frac{dX_2}{dt} = \mu_2 X_2 + q\mu_1 X_1 - DX_2 + m_1\mu_1 X_1 - m_2\mu_2 X_2 - r_1\mu_1 X_1 \quad (9)$$

$$\frac{dS}{dt} = D(S_0 - S) - \frac{1}{Y}(\mu_1 X_1 + \mu_2 X_2) \quad (10)$$

$$\frac{dq}{dt} = \left(1 - e^{-mf\left(\frac{X_2}{X_1+X_2}\right)}\right)\mu_2 \quad (11)$$

$$\frac{dM}{dt} = Y_M(X_1\mu_1 + X_2\mu_2) - DM \quad (12)$$

$$\frac{dP}{dt} = Y_p(\mu - qk_d)X_1 - P(k_p - D) \quad (13)$$

$$\mu_1 = \mu_{1\max} S/(K_1 + S) \quad (14)$$

$$\mu_2 = \mu_{2\max} S/(K_2 + S) \quad (15)$$

$$f_1 = m_p - r + z \quad (16)$$

$$f_2 = z - r \quad (17)$$

where, $m_f = \begin{cases} f_1 \text{ if } M > M_{th} \\ f_2 \text{ if } M < M_{th} \end{cases}$

also, $(0 < f_1, f_2 < 1)$

Here, $m_1 > m_2$

Since metabolic toxicity has an influence over plasmid bearing cell population to a larger amount.

m_1 & $m_2 = 0$, if $M < M_{th}$

M_{th} was required to be evaluated for different recombinant strain of micro-organisms and media composition under varying set of operational conditions and it depends upon experimental setup with presumed parameters for a bioprocess. Likewise, values of constants m_1, m_2, m_p and r_1 were assessed for a defined set of conditions.

The dynamics related to plasmid bearing and lacking cell populations together with effects impart due to metabolite toxicity and selective stress has been shown in mass balance Eqs. (8) and (9). The numerical value of q, m_1 and r_1 together at any instance

cannot be more than unity. Material balance for substrate concentration is expressed in Eq. (10) where the substrate consumption rates were considered proportional to the rates of formation of the two types of cells through their yield factors. The kinetics shown in Equations (11) and (12) ensures the role of varying probability of plasmid loss in the model and variation in metabolite concentration respectively. The Eq. (13) is mass balance for product i.e., streptokinase formation, Eqs (14) and (15) are Monod expressions to compute specific growth rates. For the collective probability factor m_f two variants are used as f_1 and f_2, in expression (16) and (17) that has got variable net probability tested with the level of threshold metabolite concentration to assume its value.

Since initially plasmid lacking cell population and metabolite concentration is absent in the medium, both were taken as zero. In the very start of the process all cells present are plasmid bearing, so the probability of plasmid loss was also to be taken as zero. The magnitude of recombinant cell population and substrate had a pivotal role in governing the dynamics.

2.4.2. Genetic algorithm (GA) approach for optimizing of model constants

2.4.2.1. Algorithm topology

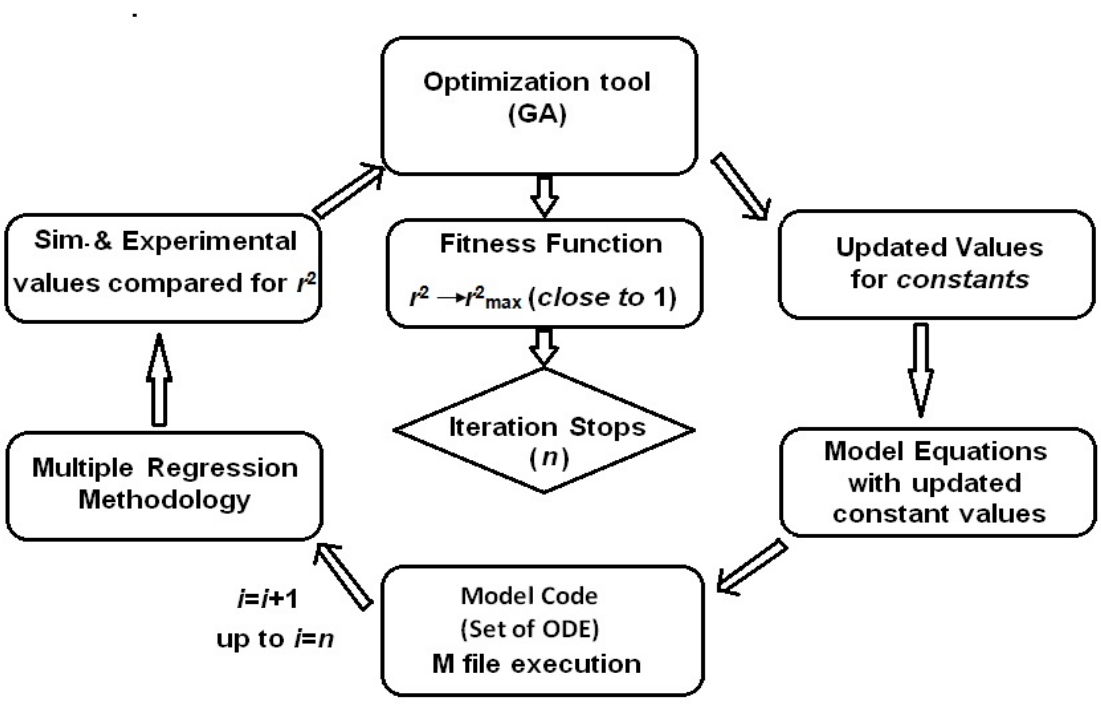


Figure 2.1. Algorithm of model development.

2.4.2.2. Modified Genetic Algorithm Paradigm

Genetic Algorithm (GA) simultaneously evaluates many points in the given parameter space it often used to converge towards the global solution, **Figure 2.2**); hence it was helpful in identifying bioprocess parameters in a non-linear system [18].

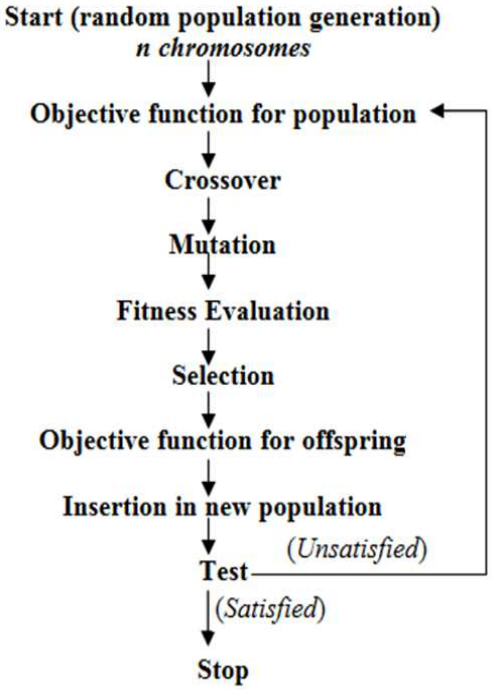
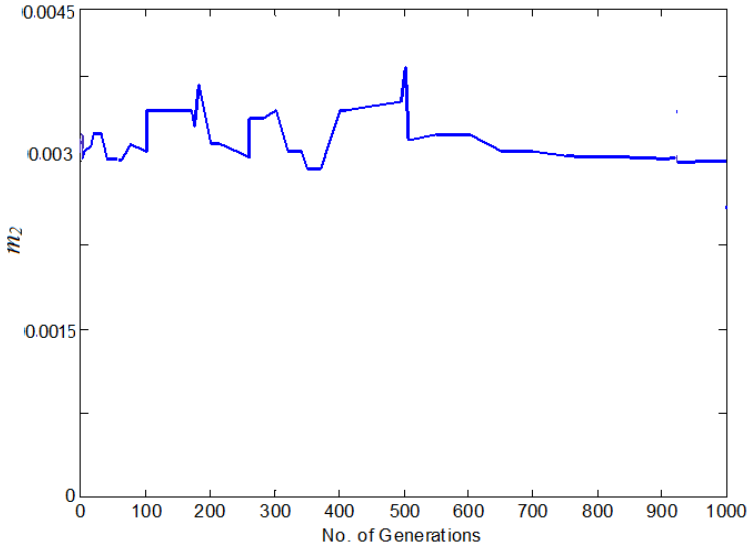
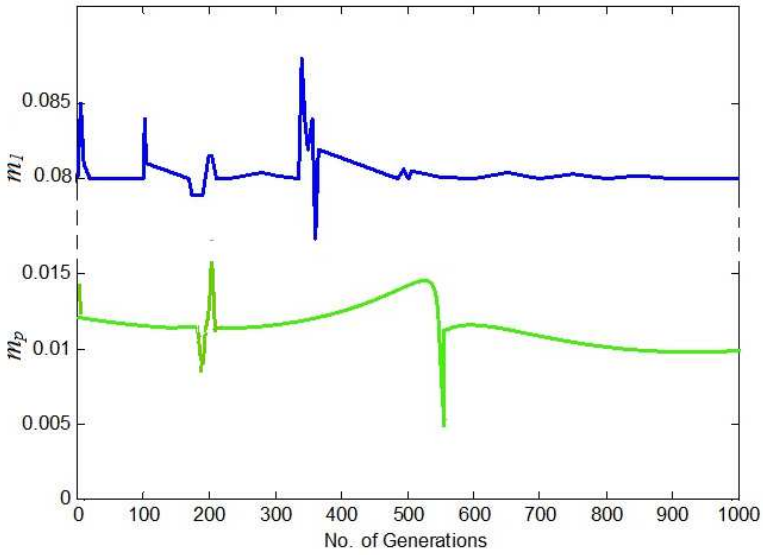


Figure 2.2. Global optimization paradigm for genetic operation with condition testing.

2.4.2.3. Parameter Identification and Optimization



2.3 (a)



2.3 (b)

Figure 2.3. a and b. Profile for process coefficients m_1, m_2 and m_p, in continuous culture with number of generations.

Table 2.1. Numerical values of Genetic algorithm against Genetic operators with adjusted parameters.

Genetic Operator and Type	**Genetic Parameter**	**Numerical value of G.A.**
Encoding-*binary*	*Gap value in Generation*	0.90
Crossover-*multipoint*	*Prob. of Cross-over*	0.05
Mutation-*inversion/duplication*	*Prob. of Mutation*	0.01
Selection-*Roulette wheel method*	*Precision of Cycle*	30
Fitness-*linear type*	*Screened individuals*	200

Table 2.2. Results of continuous culture, showing GA estimated converged values and the actual values.

GA estimated values for	m_1	m_2	m_p
parameters	0.0811	0.0032	0.0121

The plots in **Figure 2.3 (A and B)**, of continuous culture parameter identification was generated using adjusted genetic-operator parameters in **Table 2.1**, and GA estimated converged values of three model variables as depicted in **Table 2.2**.

Finally, a simulated plot (**Figure 2.4)** for estimating the key factor, threshold metabolite concentration, which may influence the recombinant state dynamics, was obtained experimentally.

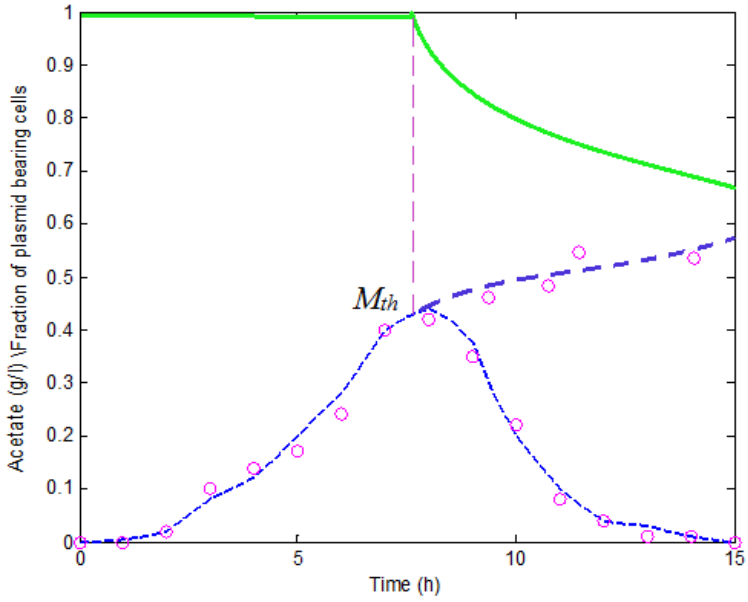


Figure 2.4. Showing simulated (dotted line) and experimental (pink circle) plots of metabolite level variation (with bifurcation for batch (lower thin blue line) and continuous(upper thick blue culture) and fraction of plasmid bearing cells (green line) to obtain M_{th} using GA optimized model variables

2.4.3. List of model parameters

Table 2.3. Initial values of model variables

Variable	Unit	Initial Value
X1	g l^{-1}	0.70
X2	g l^{-1}	0.00
S	g l^{-1}	70.0
M	g l^{-1}	0.00
q	-	0,00

Table 2.4. Values of parameters assumed in the model

Variable	Unit	Initial Value
S_0	g l^{-1}	70.00
μ_{1max}	h^{-1}	0.74
μ_{2max}	h^{-1}	0.80
Y	—	2.00
Y_M	—	4.80
Y_P	—	0.44
k_d	h^{-1}	0.02
k_p	h^{-1}	0.0005
K_1	g l^{-1}	20.00
K_2	g l^{-1}	10.00
m_1	—	0.081
m_2	—	0.003
m_p	—	0.012
r	—	0.01
r_1	—	0.02
z	—	0.001

2.5. Results and Analysis

2.5.1. Model simulation

The simulation of the above model was done using Matlab 2010Ra. The initial values (Patnaik, 1995) taken for different parameters at time zero has been shown in **table 2.3**. The simulation was done using most of the standard values taken from a previous model data meant for streptokinase (Patnaik, 2002) together with various other model constraints with smaller magnitude assumed for different set parameters which have been given in **table 2.4**., taken on the basis of their apparent role in this dynamical system. The probability of plasmid loss was zero at time t_0 since at the beginning of process all recombinants cells had plasmid machinery. Different dilution rates were considered which proved to be the most relevant factor for continuous operation. Dilution rate was started at a

very low value and increased to high values to evaluate the sensitivity of response at different levels.

Specific growth rates taken for plasmid bearing and plasmid lacking cells in the model were different since specific growth rate of plasmid lacking cells would be noticeably higher. Yield factor which was regarded to be the ratio of gram of cells formed and gram of substrate consumed was taken together with another yield term associated to the formation of metabolite. Other constants governing the effect of metabolite concentration were the two toxicity coefficients m_1 and m_2. While various model related probability factors were effectively incorporated pertaining to plasmid loss due to metabolite toxicity and plasmid retention for selective stress, some other rare factors were also considered.

The approximate value of exponential term was taken e=2.71828, as the base of natural logarithm. The value of threshold concentration of metabolite taken here after estimation, m_{th} = 0.50, which had its pronounced key role in leading the major aspects of the process. Additionally another factor, selection stress coefficient was incorporated to strengthen the model in respect to the selective operation carried over by traces of antibiotics present in the medium. The stress causing substances present in the environment which often found to inhibit the formation of cells devoid of any plasmid. The simulation was done at low and a fairly high different dilution rate *viz*. D_1 and D_2, as being shown in subsequent plots. In figures all three variables with different initial value (**table 2.3**) were taken together to generate the dynamics. The cases where dilution going outside the limit i.e., $0.23 \leq D_i \leq 0.65$, i=1, 2,...; had not been studied because of insignificant changes beyond the range for this context. The model plot for three vital variable parameters were obtained on simulation. The uppermost prominant dashed line showing X_1 and the just lower solid line is representing X_2, the third dash-dot line is showing product (streptokinase) concentration. Plotting all the three variable parameters in **Figs. 2.5.1. to 2.5.5**, together consequently justifies the correlation among different component variables simultanaeously.

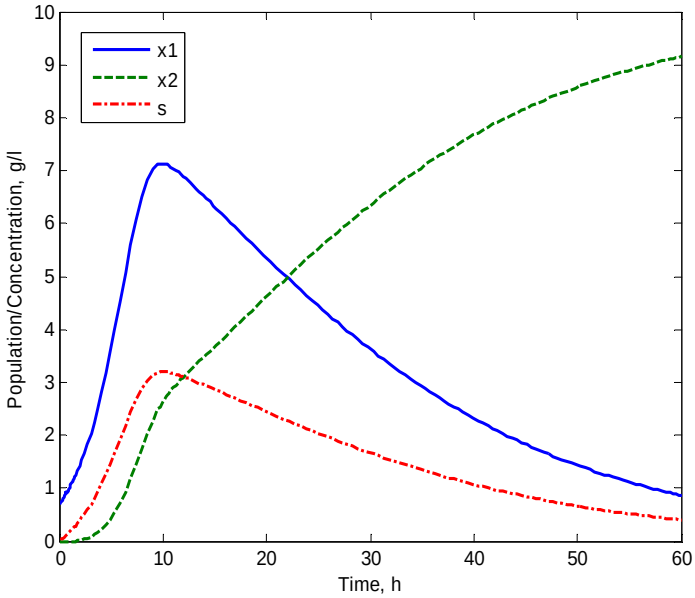


Figure 2.5 (a). Dynamics shown in plot at low dilution rate, D_1= 0.10

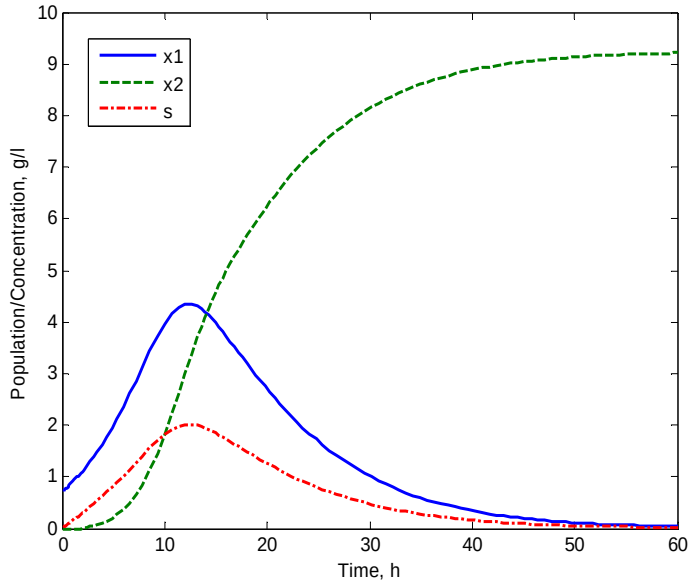


Figure 2.5 (b). Dynamics shown in plot at dilution rate, D_2= 0.23

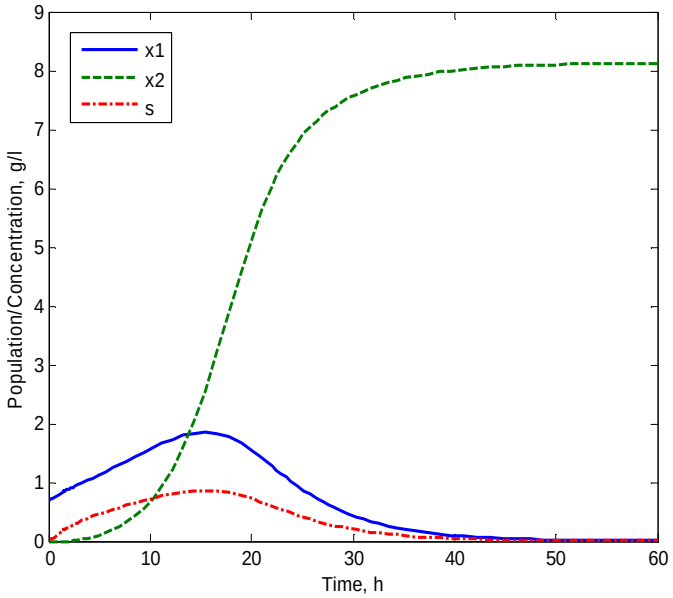


Figure 2.5 (c). Plot shownatdilution rate D_3= 0.35

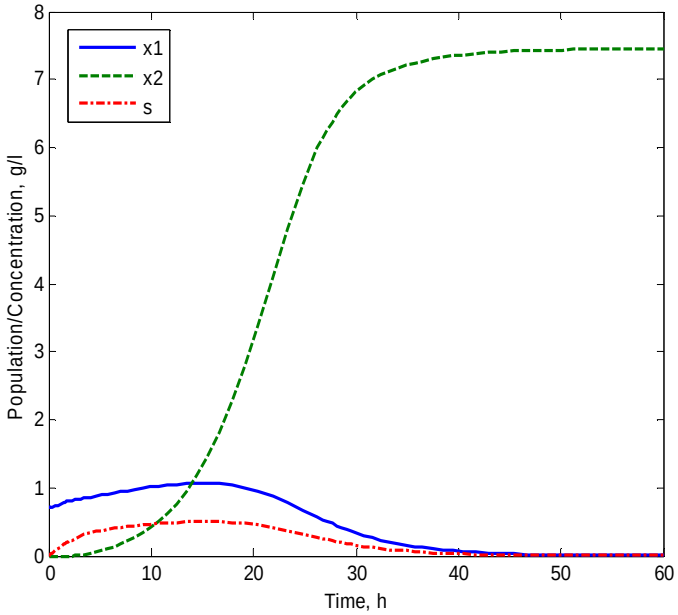


Figure 2.5 (d). At dilution rate, D_4= 0.40

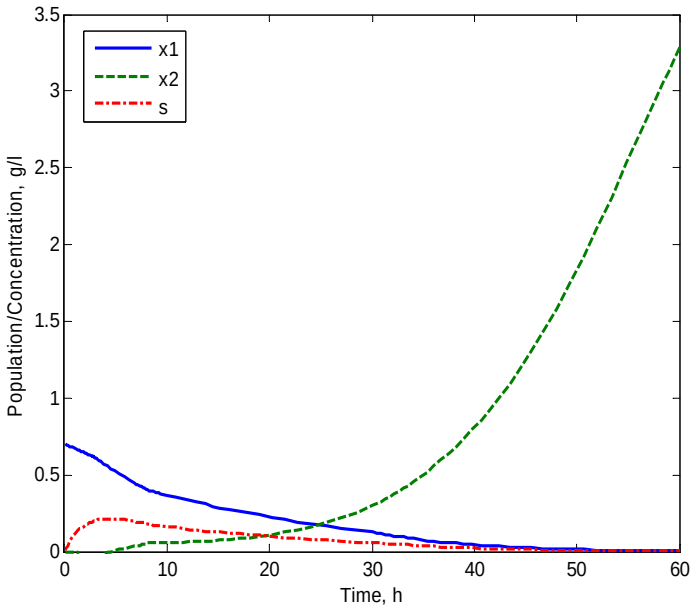


Figure 2.5 (e). Dynamics shown in plot at dilution rate, D_5= 0.52

It is quite reasonable to investigate that low dilution rate allows the fermentation to last longer generates higher magnitude of concentration of plasmid bearing cells and favour the consumption of substrate. The numerical simulation of the model equations shows plasmid loss tends to occur after a certain limit of metabolite level, which depicts that the increasing concentration of metabolite tends to support the population of plasmid lacking cells that witnesses the trend in loss of plasmid from recombinant cells. Using the simulation dynamics it becomes easier to resolve the plot in order to understand the trend followed by the two population of cells. The probability of plasmid loss has its dependence over metabolite concentration and the two population of cells. The varying probability has its impact over the dynamics of the model thereby involving almost all the parameters that are being substantially affected due to its variability. It is very interesting to note that only a threshold amount of metabolite concentration is responsible for starting the variable dynamics of plasmid lacking and plasmid bearing cells. The model simulation was performed using most of the standard values from existing models and assumed constraints required to explicate this dynamical system.

Dynamics observed in initial plots, **Figure 2.5 (a)**, clearly depicts the bifurcation in population of pure recombinant cells and growth in plasmid lacking cell population. Noticeable decline in X_1 with subsequent elevation in X_2 level is evident on moving from

Figs. 2.5 (a) to 2.5 (e). After a certain dilution rate the downward steepness of the the curve elucidate the event of ongoing increase in plasmid lacking cells, indicates that there is a gradual decline in the level of recombinant cells on enhancing the dilution, shown in **Figure 2.5 (c)** Since the production of streptokinase is directly depending upon the dynamics of palsmid bearing cells so to enhance the production it is inevitable to reveal the kinetics operating factors and to evaluate the instability of the plasmid. The rise in probability can also be well marked in the plots in respect to the remarkable hike in metabolite formation. These two events can be supposed to deduce the trend followed by growth of plasmid lacking cell population.

The influence of dilution rate *D* on concentration of plasmid bearing and plasmid lacking cells with respect to time duration of continuous process is represented in two plots, **Figs. 2.6 and 2.7**. The plots in respect to dilution rate have their importance in deciphering the behavior of biased inter-population cell dynamics in this case. The results emphasizes that delayed plasmid loss occurs at lower dilution rates. Numerical simulation of the continuous fermentation process could be rather helpful to enhance the performance adjusting dilution rate to achieve product in amplified amount.

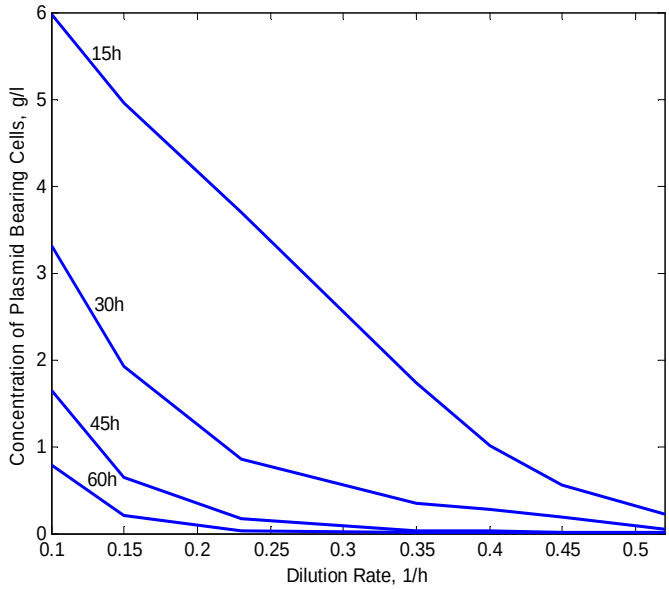


Figure 2.6. Effect of dilution rate on plasmid bearing cells with respect to time duration of fermentation process

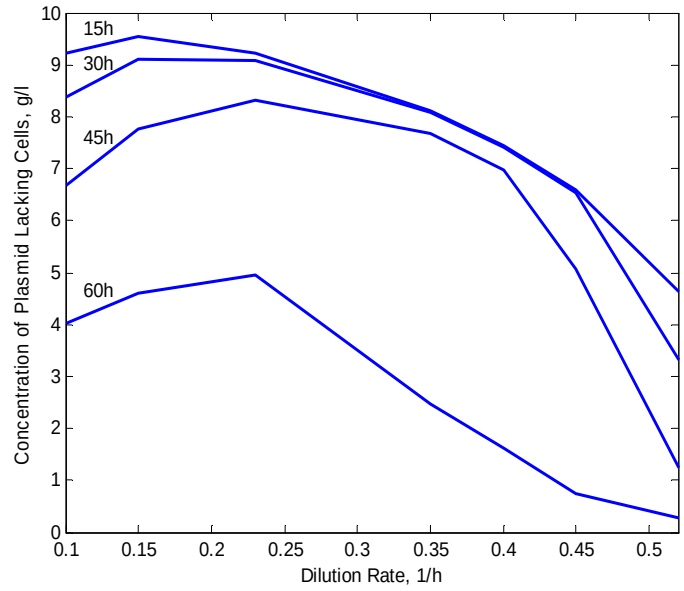


Figure 2.7. Effect of dilution rate on plasmid lacking cells against time duration

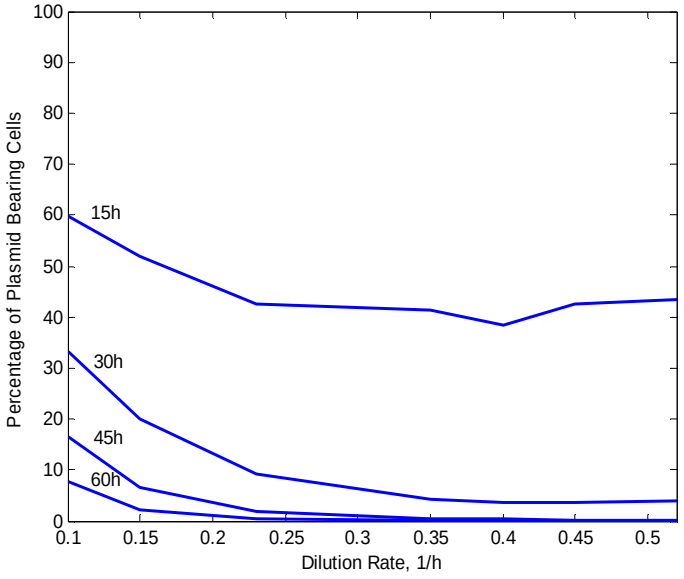


Figure 2.8. Effect of dilution rate on percentage of plasmid bearing cells with respect to time duration

The observed trends were matched by the concentration profiles at different time duration of the fermentation process. It could be noted that better dynamics was obtained at late hours. Two cases were studied for plasmid bearing and plasmid lacking cells. The concentration profile is plotted at duration in hours: 15, 30, 45 and 60, for X_1 and X_2 cell concentrations, **Figures 2.6** and **2.7** respectively. There is a remarkable increase in abruptness of the slope on progressing ahead in respect to process duration, from 15h –

60h. The **Figure 2.6**, showing nearly smooth variation with increasing duration on the other hand irregular variability is there with respect to concentration (**Figure 2.7)**. An interesting feature reflected by plotting concentration against dilution rate is assessment of time in governing the interaction between cell populations (X_1 and X_2) and dilution rate regardless of how intense the phenomena associated to other variables are progressing. **Figure 2.8** shows the effect of dilution rate on percentage of plasmid bearing cells with respect to time duration. An interesting feature of changing probability of plasmid loss with respect to metabolite concentration at different dilution rates, is depicted in **Figure 2.9**. There is a very high probability of plasmid loss even at low metabolite concentration on moving towards a high dilution rate. Particularly the probability has shown a rapid hike on increasing dilution rate from D=0.45 to D=0.52. The percentage of plasmid bearing cells with respect to time in **Figure 2.11**, showing a gradual decreasing trend on increasing the dilution rate to some extent, D=0.35, the declining trend is not found to exist on further dilution rates.

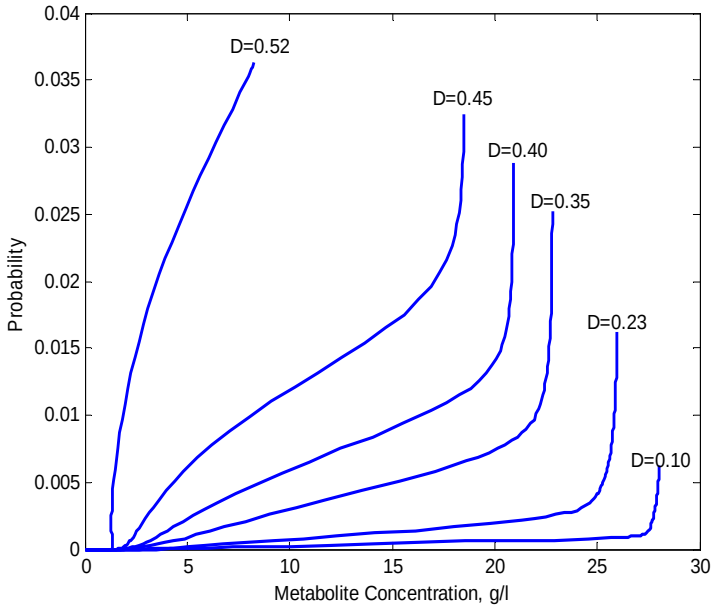


Figure 2.9. Changing probability of plasmid loss with respect to metabolite concentration at different dilution rates

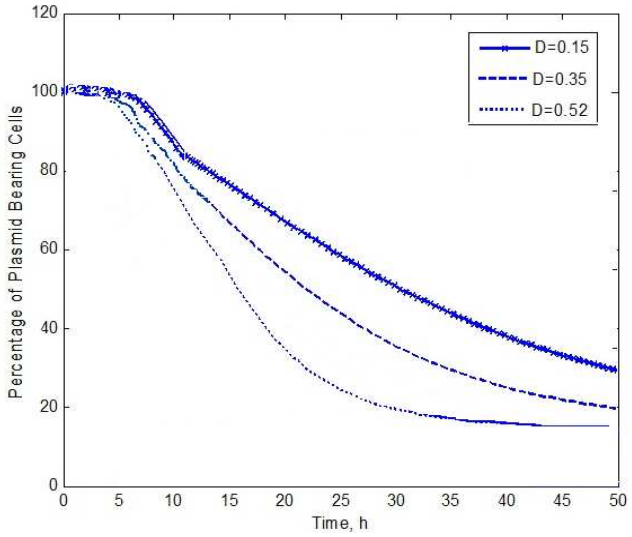


Figure 2.10. Percentage of plasmid bearing cells with respect to time

2.5.2. Model Validation

The model dynamics has beed validated for specific dilution rates. Experimental results (Yazdani and Mukherjee, 2002) were used to compare with the simulated numerical values to compute the statistical significance of the closeness of two data using multiple regression analysis via mutivariate optimization tool. Matlab contains optimization tool commands, *nlin* function, that was used to obtain results in this regard. The **Figure 2.10.** shows together the experimental and simulated data plot obtained at a particular dilution and substrate concentration in continuous culture.

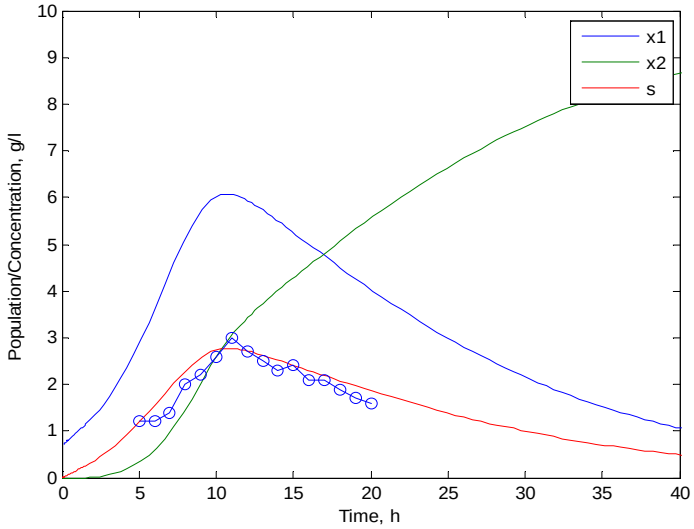


Figure 2.11. Showing diversion of X_1 and X_2 cells growth dynamics at dilution rate, D=0.15 and substrate concentration in g/l, experimental findings for product streptokinase in mg/l after induction (5h to 20 h) has been represented by open circles

The model validation was done in respect to declining trend of plasmid bearing cells for which simulation plots obtained at different dilution rates were correlated to experimental results obtained at respective dilutution. Similar data characteristic were observed (Kim et al., 1998) in study of unstable bacterial population strain in chemostat. The obtained correlated trends are depicted in the **Figure 2.11.** below.

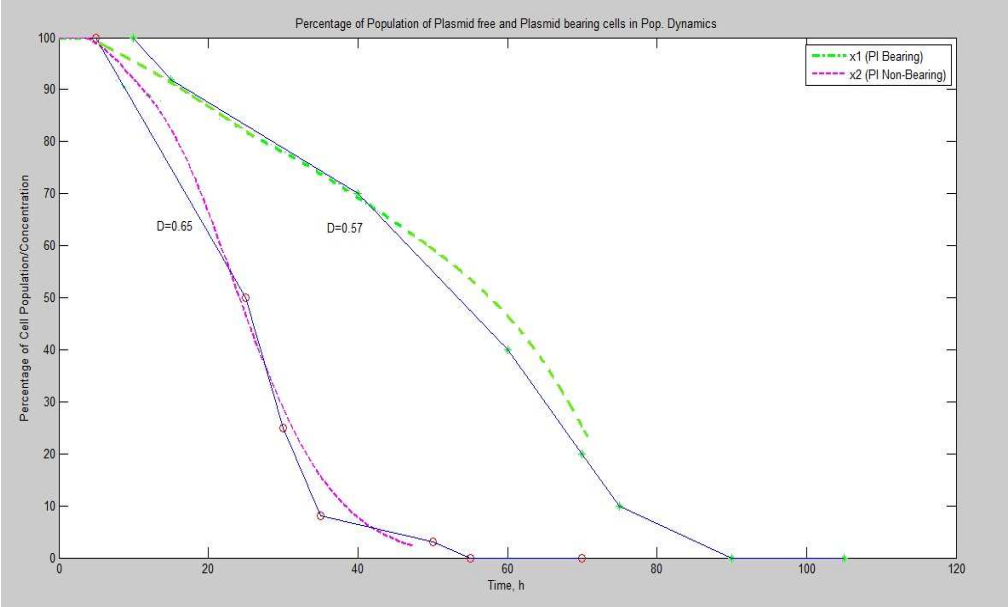


Figure 2.11. Simulated and experimental data plot has been shown, dynamics depicting the percentage decline of plasmid bearing cells at respective dilution rates of D=0.65 and D=057.

2.6. Discussion

The behaviour of the systems like bioreactor is found to be solely dynamic. The effort was to organize the information obtained from fermentation regarding set of cultivation parameters. The key factors noticeably found to play the key role would be taken together into account to resolve the simultaneous variation in system dynamics (Friehs and Schügerl, 1990).

Population dynamics model for plasmid bearing and plasmid lacking cells in bioreactor had been made more robust to develop an *insilico* dynamical system which had having the characteristics of a chemostat that we used to employ for streptokinase production. The dynamical system representation through modelling has its relevance in

predicting the behaviour of the system on disturbances or changes made in initial conditions.

All the considered five parameters, including plasmid bearing cell population, plasmid lacking cell population, substrate concentration, metabolite concentration and probability of plasmid loss were taken into account. The idea was to measure the instability of plasmid, which can be directly derived from the growth of plasmid lacking cell population. The model had been simulated at different dilution rates and different initial substrate concentration and thereby differences in the model behaviour were observed on simulation. There was a remarkable decline in the percentage of plasmid bearing cells above a certain level of dilution rate. System had shown its prevalent sensitivity to change in dilution rate. The probability of plasmid loss had shown to follow a gradual trend at high dilution rates. In the proposed model new factors were taken into account like selection stress coefficient and metabolite toxicity coefficient that had resolved the simultaneous variation in other parameters and their interaction criteria. Selection stress coefficient showed to resist the process of plasmid loss up to some extent. Increasing concentration of metabolites had shown to inhibit the growth of two cell populations after attaining a certain threshold concentration, it could lead to the fact that plasmid lacking cells population had adopted an increasing trend while plasmid bearing cell population was found to follow a decreasing trend. The most significant consideration of the model was the changing probability of plasmid loss with respect to time which incorporated the effect generated from toxicity developed by the metabolites. Various other time invariants and intrinsic constraints were together taken into account to plot their collective effect with probability factor. The model ensures a higher degree of flexibility since it has a number of adjustable parameters. Still efforts are required for including influence of genetic factors and use of heuristic approaches to impart certain non-ideal conditions to bioprocess phenomenon.

CHAPTER-3

DEVELOPMENT OF MODEL SIMUTALION TOOL AND ANTIBIOTIC REGULATION STRATEGY

3.1. An Overview

In this chapter, the problem of stating adjustable overall bioreactor dynamics has been studied. Since the formation of product streptokinase was found to be directly proportional to the number of recombinant cells in the media so our emphasis here was to regulate the antibiotic concentration since its effect ceases after few hours of the start of fermentation process in continuous culture. Using culture medium without ampicillin the plasmid stability decreased to 0% after 60hours of cultivation while it dropped to only 60% on using 100 mgl^{-1} of antibiotic in the medium (Kim *et al.*, 1998). Sometime use of combination of two antibiotics provided a better result (Friehs and Schügerl, 1990).

Since the effect of antibiotic in culture medium did not control the plasmid lacking cells with same strength for a longer span, a model with variable antibiotic concentration was proposed to show the shifting of $T_{1/2}$ signifying the generation time (Ganusov et al., 2000) of plasmid bearing cells. A strategy to regulate antibiotic concentration during bioprocess was adapted. Plasmid bearing and non-bearing cell growth dynamics in respect to generation time was observed, utilizing time of half-elimination of plasmids from the recombinant cell population. If it would be possible to increase $T_{1/2}$ we may compute for better production magnitude. There seems to be a probable solution and it could be possible via applying this variable antibiotic concentration strategy. Since addition of antibiotic ampicillin improves the health of plasmid bearing cells its variation after initial addition of an amount may further regulate the competitive dynamics and relative growth of two kinds of cell populations. Hence, adding the antibiotic to the bioreactor in a fashion reduced the probability of losing at least one plasmid copy per cell during cell division. Implementing the above theme, we incorporated the effect of antibiotic concentration in the prior developed basic model equations. The effect of antibiotic on both the populations and other model parameters can be observed performing the simulation in Matlab. It noticeably represented the delay in percentage decline trend for plasmid bearing cells in reactor with the progress of fermentation.

A Graphics User Interface (GUI) was constituted to predict the model behaviour at different dilution rate and other adjusted parameters. Development of bioprocess model simulation tool had been done and this bioprocess simulator was helpful in simulating different models dealing with various bioprocess operations (Sevella and Bertalan, 2000). The general mathematical modelling tool was developed on the basis of Matlab inbuilt tool box. The programming system with GUI having the functions to solve differential equations based system had been configured with ease to use graphical user interface. It became easy to adjust different parameters of the dynamical system and observe the behaviour in response to the assumed fermentation systems via *insilico* environment.

In the way to exploit the relevance of other bioprocess models, fed-batch had been particularly considered and simulation of the process model was performed on the basis of conceptual algorithm. In the model for batch continuous culture the additional criteria of intermittent addition of substrate or production media were summed up to develop a model for fed batch culture. It was well noticeable from the validated simulation plot that higher yield in this process was possible in a sustained fashion for longer duration. Several bioprocess approaches include temperature up-shift approaches, feeding strategies, timing of induction, and implementation of two-stage culture mode for stabilization of plasmid had been implemented (Razali *et al.*, 2007). Moreover, a structured model framework to evaluate production in bioprocess had been taken into account. Mechanistic model for reliability based assessment of streptokinase production from a bacterial cell using stochastic Monte-Carlo based principles had been configured to make another application of structural modelling strategies. Our product enzyme is formed from the cell as the result of interaction among various intra-cellular factors within the cell and the composite effect of its environment. Since a highly developed model presents designing of composite architecture that optimally blended cellular intelligence, artificial intelligence and mechanistic models. Hence several interacting sub-components pattern together have their influential role in enzyme production kinetics.

The concept of reliability has its relevance in assessing the durability of any functional mechanistic unit within a defined time frame. The idea was to support the dynamic pathway associated networks in view of conducting possible estimation of the enzyme production, taking streptokinase into consideration as an instance. Here time dependent interaction of intra-cellular subcomponents was plotted out, few prominent

parameters among the ample of parameters in natural cellular system had been given a desirable weight age in developing the model.

3.2. Bioprocess Simulation Tool

3.2.1. Design of simulation platform

The GUI was created to perferm *insilico* simulation of the bioprocess. Few functions which were configured in Matlab were together simulated to perform the simulation task. The snap shot of the simulation platform is visualized in **Figure 3.1**.

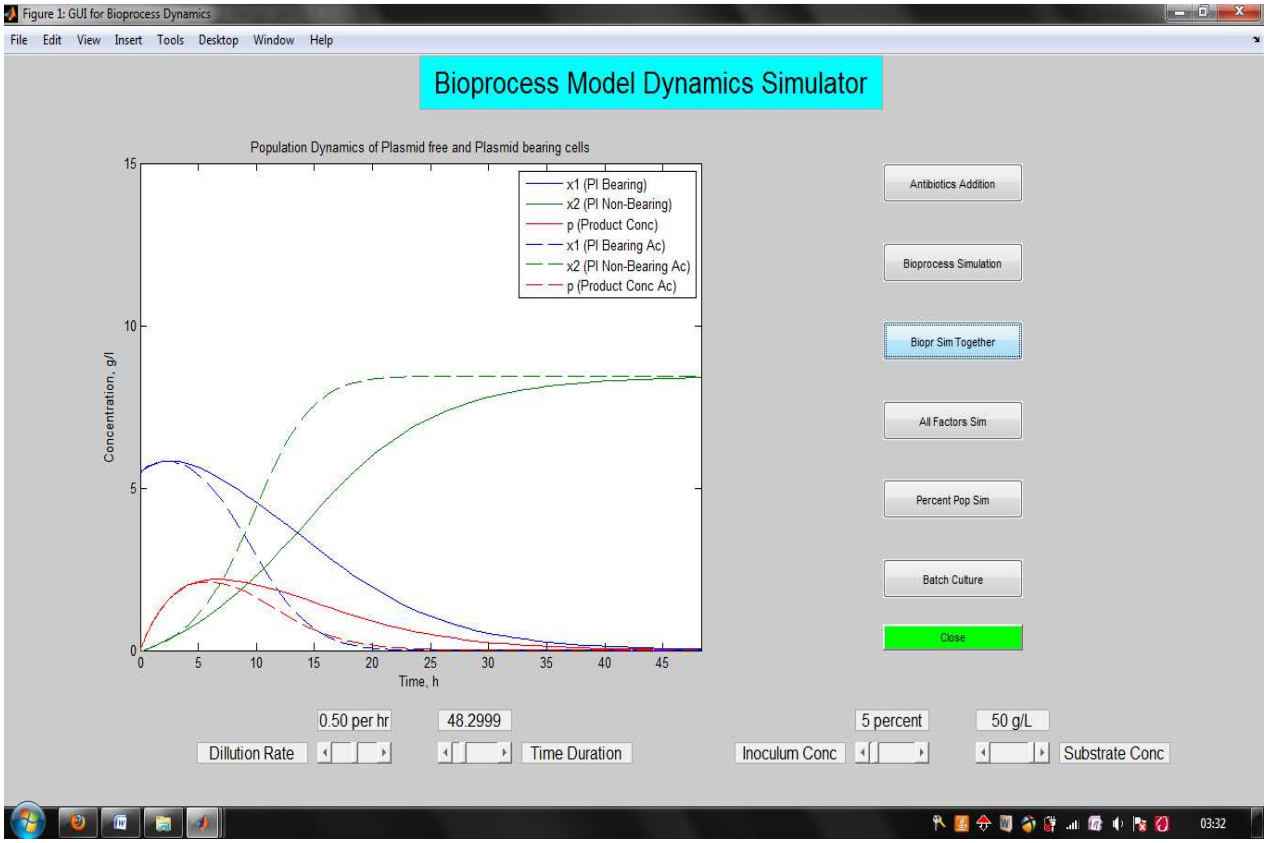


Figure 3.1. Bioprocess Simulator

A general mathematical modelling tool was developed on the basis of Matlab software. The program system with GUI provides chance to solve differential equation based system with the help of easy to use graphical surface (Sevella and Bertalan, 2000). The Matlab code of creating GUI is given in *Appendix A1*.

The model simulator can present the various stages of the model dynamics. In order to compute the concerned numerical domain related to work, process simulation has been done using Matlab R2010a, it is thus possible to numerically evaluate the role of each and every parameter using initial values (Patnaik, 1995) of different parameters. Using

standard values, the simulation can be performed via GUI platform following the previous model input data intended for streptokinase production (Patnaik, 2002) in culcating various other constraints with little magnitude considered for a different set of model parameters (Kumar and Ghosh, 2010), incorporated in our earlier work.

The dilution rates of discrete magnitude are applicable which provide a most significant factor for continuous process. In the beginning, the dilution rate was set at very low value and subsequently increased to a high magnitude to assess the level of response at various extents. Artificial intelligence has been invoked extensively to model macroscopic microbial behaviour under the influences of noise and certain spatial variations in fermenter [30]. Dilution rate regulates the persistence of cells in the production broth for the definite duration.

3.2.2. Topology of the Simulator

The working topology of the bioprocess simulator for culture and production of streptokinase is described in the **Figure 3.2.** It comprised of Ordinary differential equation (ODE) functions which were linked to callback buttons. The GUI terminal was configured to provide a sophidticated platform to conduct process operation. The adjustment of various parameters pertaining to bioprocess parameters could be well feasible via parmeter selection buttons in slider window form. The user can adjust the parameters including, dilution rate, time duration, innoculum and substrate concentration before visualizing the results using buttons on the graphics platform.

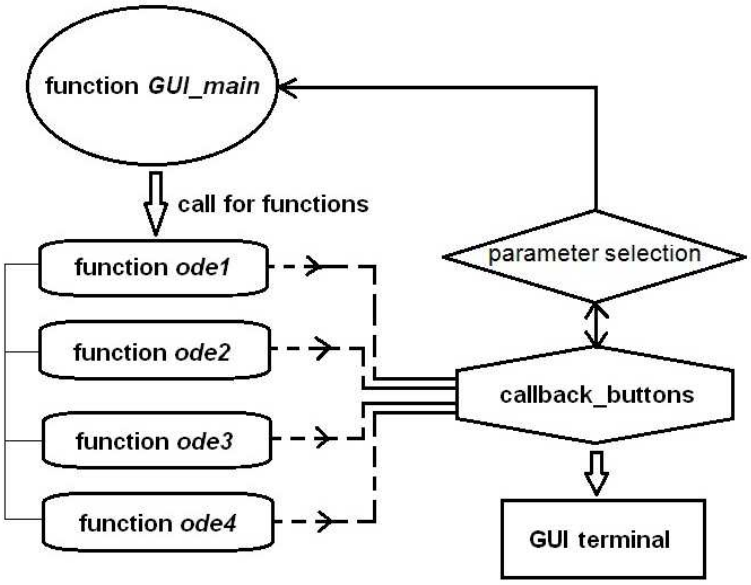


Figure 3.2. The topology of the GUI scheme in working

3.2.3. Advantages of using simulator in evaluating fermentation process

Bioprocess simulator is helpful in simulating different models dealing with various bioprocess operations. It becomes easy to adjust different parameters of the dynamical system and observe the behavior in response to the assumed bioreaction systems. The batch as well as continuous cultivation can be visualized adjusting dilution rate from zero to some decimal value and other parameters too.

It is evident from numerical simulation of the model that plasmid loss tends to occur after a certain value of metabolite level, it certainly depicts that the elevation in concentration of metabolite have a propensity to support the plasmid lacking cells population. The data obtained experimentally is of great significance in understanding and validating the computational results.

In case of continuous operation, varying dilution rate *D* with slider have shown noticiable influence on recombinant cells concentration with respect to process duration. The dynamics pertaining to *D* have their relevance in interpreting the behavior of a blended inter-population dynamics at an instance. Outcome emphasizes over the delayed plasmid loss that often occurs at lower dilutions. Since continuous cultivation has number of adjustable parameters so their inter-dependent role in dynamics behaviour is appreciable. The numerical simulation of continuous process using a graphics platform would be rather much helpful to improve the performance regulating dilution rate in order to achieve amplified amount of product.

On utilizing the graphics facility of different parameter adjustment, it was clear from the simulation that rapid variation particularly hike in metabolite concentration after a certain concentration caused profound increase in probability of plasmid loss. Using the means of simulation it was evident that even at a very low metabolite concentration level there found a high probability of plasmid loss on shifting towards high dilution rates.

3.3. Relevance of other Bioprocess Models

3.3.1. Fedbatch

In the model for continuous culture (in **chapter 2**, **Eq.8 - Eq.17**) the following condition were are added after Eq. 17, to develop a model for fed batch culture.

Case I: $X > X_{max}$

$D = D_F$

Case II: $X < (1 - S_f)\, X_{max}$

$D = D_0$

Here, X_{max}=3.2g/l, S_f=30g/l , D_0=0 and D_F=0.20;

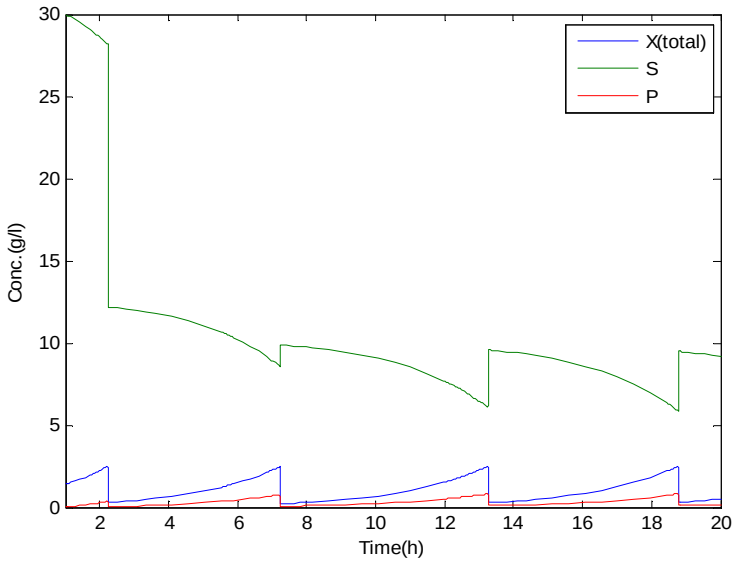


Figure 3.3. Fed-Batch Simulation Pattern

Plotting all the three variable parameters together, subsequently justifies the correlation among various component variables in a fed batch culture simultaneously, in which substrate is being added at time intervals. The similar effort was made by researchers (Sevella and Bertalan, 2000) earlier, but in our case the conditions applied with reference to production of streptokinase and the comparable results were obtained by experiments (Yazdani and Mukherjee, 1998). Implementing the defined criterion the simulated model obtained is shown in the plot above, **Figure 3.3.** Progressive induction by means of continuous IPTG dosage in *E. coli* fed-batch cultures yield a higher specific levels of recombinant protein (Pinsach *et al.*, 2008). It is noticeable that sustained yield is possible in a continued fashion for longer duration.

3.3.2. Overall Bioreactor System model

The model equations were developed by incorporating different model parameters (Sevella and Bertalan, 2000). The initial conditions taken in the model at t_0 are:
$X(0)$=1.40; $S(0)$=30.0; $P(0)$=0.00; $C(0)$=5.20; $V(0)$=2.00;
Various other model constraints used, k_o=0.0002; K_La=150; α=0.64; C*=0.005; Y=0.5; and μ_{max}=0.2;
The simulation result is shown, in the **Figure 3.4**;

In overall bioreactor model, for representing the characteristic dynamics of different parameters together, our culture system is described by the given set of differential equations, **Eq.18-21**.

$$\frac{dX_{total}}{dt} = \mu_{\max} \frac{S}{k_s+S} \cdot \frac{C}{k_o+C} . X_{total} \tag{18}$$

$$\frac{dS}{dt} = -\frac{1}{Y} \mu_{\max} \frac{S}{k_s+S} \cdot \frac{C}{k_o+C} . X_{total} \tag{19}$$

$$\frac{dC}{dt} = -\frac{1}{Y_o} \mu_{\max} \frac{S}{k_s+S} \cdot \frac{C}{k_o+C} . X_{total} - K_{La}(C^* - C) \tag{20}$$

$$\frac{dP}{dt} = \alpha . \frac{dX_{total}}{dt} \tag{21}$$

The overall bioreactor model is described in fed batch phase like fermentation system, it mimics the parametr dynamics obtained from bioreactor system.

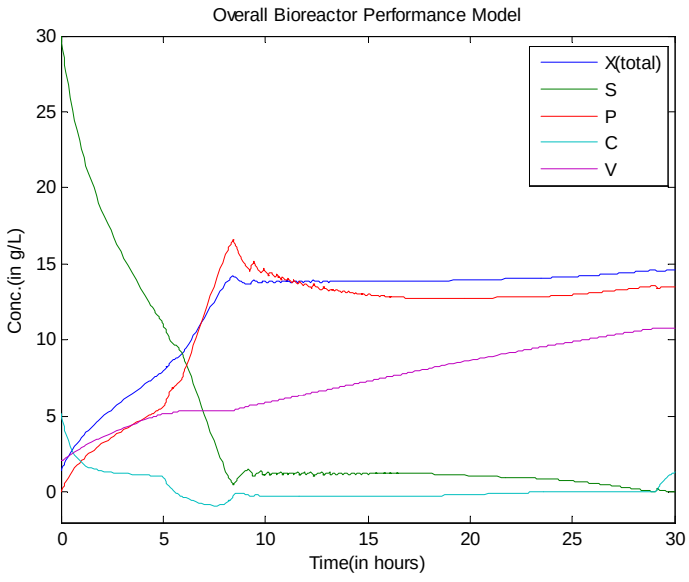


Figure 3.4. Overall Bioreactor Performance

3.4. Strategy to Regulate a Variable Antibiotic Concentration during Bioprocess

Plasmid bearing and non-bearing cells growth dynamics in respect to generation time utilizing time of half-elimination of plasmids from the recombinant cell population had been studied (Lu and Hadeler, 1998), (Ganusov *et al.*, 2000). It was noticeable, if it would be possible to increase $T_{1/2,}$ it might be possible to compute better

production magnitude, **Figure 3.5.** There seems to be a probable solution and it could be possible via applying regulation of variable antibiotic concentration. Since addition of antibiotic ampicilin improves the health of plasmid bearing cells its additional variation after initial addition of its requisite amount may further regulate the competitive dynamics and relative growth of two kinds of cell populations.

If, at an instance *t*, number of plasmid bearing cell ≠ 0 then,

In case, $dX- = 1$, $dX+/dX- = dX+$, $dX+ = dX+_{max}$

This is the maximum value for plasmid bearing cells in the bioreactor,

Therefore,

$\frac{dX^+}{dX^-}$, always less than dX^+_{max}

Hence, adding the antibiotic to the bioreactor reduces the probability of losing one plasmid copy per cell during cell division which is inversely proportional to generation time.

Therefore $T_{optimal} = g/\tau_{pl}$

Where $T_{optimal}$ is the optimal time of continuing the bioreactor, g is the mean generation time,

$g = In2/D$

and τ_{pl} is the probability of losing one plasmid copy in division.

Ganusov *et al* (Ganusov et al., 2000) stated that the time of half-elimination of plasmid carrying cells from the entire population in the bioreactor is half of the time of generation,

i.e. $T_{1/2} = g/\tau$

This proves that the $T_{1/2} < T_{optimal}$

Where $T_{optimal} = T_{1/2} + \alpha$

has been discussed (Lu and Hadeler, 1998) in mathematical model for plasmid bearing and non-bearing population. The simulation of two kinds of plots can be well visualized in **Figure 3.6** and **3.7**, noticeably showing the shifting of $T_{1/2}$ time.

Assuming the relative dynamics of two population,

$$alpha = (dX_1/dt)/(dX_2/dt) = dX_1/dX_2 \quad (22)$$

$$\begin{cases} \text{if } alpha <= 1, & a_c = f(a_f) \\ a_c = f(b_f), & \text{otherwise} \end{cases}$$

In above expression, $a_c = f(a_f)$ and $a_c = f(b_f)$; where,

$$a_f = (\mu_2/\mu_1)*t_c \quad (23)$$

$b_f = (\mu_1/\mu_2)*t_c$ here, t_c - the antibiotic toxicity coefficient (24)

After addition of a necessary quantity of antibiotics, the rest of other quantity would be regulated in variable fashion following the relatively varying population trend of two cell types.

Implementing the above theme we need to incorporate the affect of antibiotic concentration in basic model equation (in **chapter 2**, **Eq. 11**);

$$dq/dt = (1 - e \wedge (-mf\,(X_2/(X_1+X_2))*a_c)) *\mu_2 \qquad (25)$$

The effect of antibiotic on X_1, X_2 population and other model parameters can be observed performing the simulation in Matlab. It noticeably represents the delay in percentage decline trend for plasmid bearing cells in reactor with the progress of fermentation**,** also been found experimentally. The simple generation time curve is designed to present the possibility of improving $T_{1/2}$ by some finite magnitude, **Figure 3.5**.

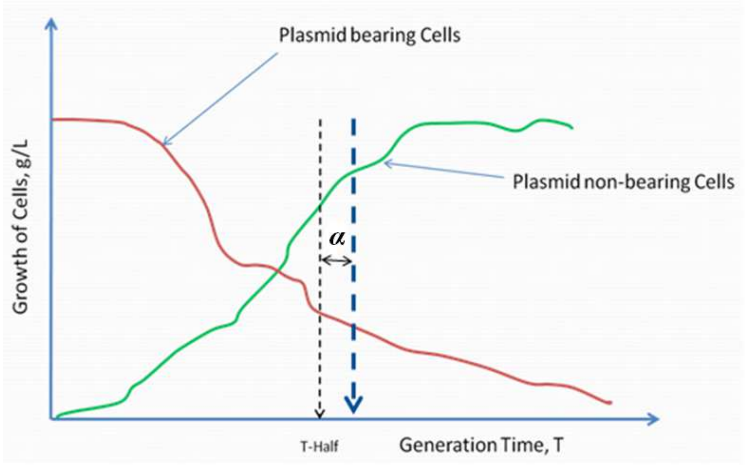


Figure 3.5. Diagram showing $T_{1/2}$ of generation time (Ganusov et al., 2000) in respect to plasmid-bearing (X_1) and non-bearing cells (X_2) cell population and achievable increase in $T_{1/2}$ by additional α magnitude

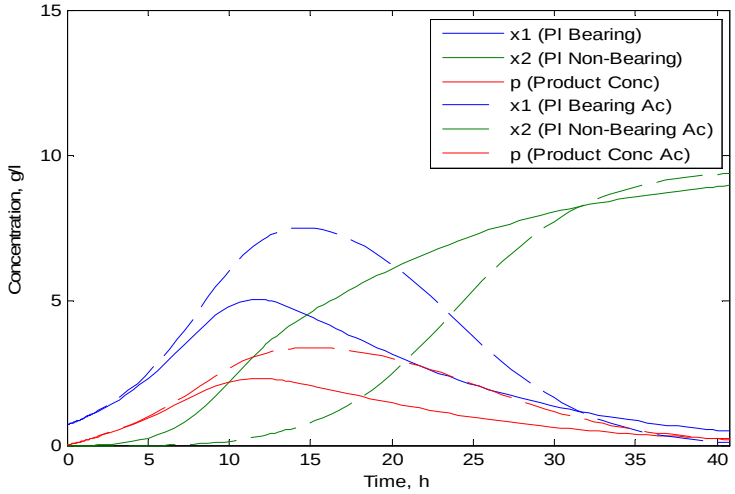


Figure 3.6. Profile of X_1, X_2 & P with fixed antibiotic (solid lines) and antibiotic concentration (dashed lines).

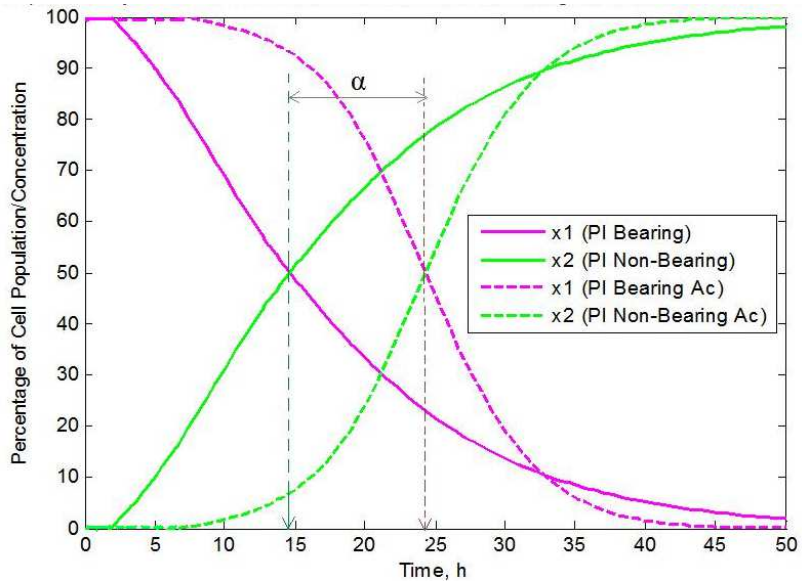


Figure 3.7. Population dynamics of % plasmid free and plasmid bearing cells at dillution rate 0.20 per h^{-1}

In the figure solid and dotted purple line shows percentage decline in plasmid bearing cell population while dotted line depicts the trend of population with administered ampicillin concentration at D=0.20, likewise the trend is presented for plasmid lacking cells in green line.

The population of plasmid bearing and lacking cells system is shown, **Figure 3.6**, with and without antibiotic regulation strategy together with variation in product concentration. The simulated plot, **Figure 3.7**, illustrated $T_{1/2}$ of generation time with respect to both type of cell populations with achievable shifting in $T_{1/2}$ by α magnitude on using the stated strategy.

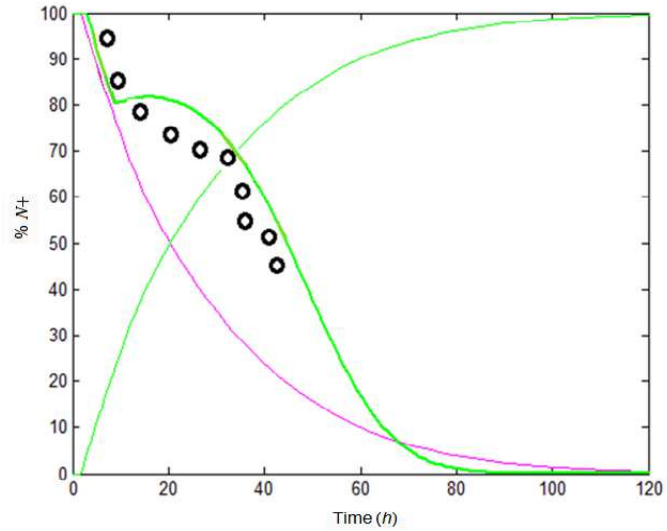


Figure 3.8. Plot showing decline in percentage of N+ at only fixed and variable antibiotic concentration at D=0.14

Using variable antibiotic concentration **Figure 3.8**, it was clearly observed that there found to be a delay in percentage decline of recombinant cells during fermentation process on applying the strategy of variable ampicillin concentration with devised algorithm.

3.5. Structured Pathway Model Approach

3.5.1. General Framework

The models have sufficient details to get compatible with new experimental techniques and together with experimental and modelling work in fermentation technology (GanNielsen, 1992). Focus is required on the structured models which describe microbial kinetics by means of selected cell components rather than by undifferentiated mass. To understand the plasmid-bacteria interaction, spatially structured model of plasmid dynamics have been often studied using simulation of dynamical system (Krone et al., 2007).

Production of an enzyme or product within a cell is associated to various pathways. So, several components and sub-components together have their role in an enzyme production dynamics, depicted in **Figure 3.9.** Several subcomponents or intracellular state vector parameters (Palsson and Joshi, 1987) involve in the production dynamics used to contribute their some vital role and hence control the product formation. Although there is a decline in the product formations in respect to the considered time frame in general. We can evaluate the performance of a unit carrying out production process considering the associated prime pathways with their subcomponents and ultimate component which is giving the product after the entire interaction of all subunits substantially in a defined manner. The sequential interaction of subunits, their failure and repair in the considered time span is quite interesting. A probabilistic framework can show a better implication of the approach to design a model in order to evaluate the performance of such dynamic units. The feasibility of this approach is to be made on the basis of reliability assessment of the production unit in a mechanistic way. The sole network could be expressed in the form of a unitary system which is responsible for the production of a particular product which lies on the retaining of its productive state with varying level of product formation which is very likely to be time dependent. The release of metabolites etc may have its impact to negatively affect the subunits involve in pathways, results into cease in production and hence decline in the reliability of such units taken into account.

It is also a remarkable thing that production can be effected by interference of large number of intra-cellular factors that may somehow interact to alter the production

level in the process. It may be possible that failure of a few sub-components can have insignificant impact to decline the entire production on the other hand there may be other set of components whose dormancy may result into termination of productive strength. A combinatorial approach is required to device the 'on' and 'off' state of such subunits in various feasible ways. The study of the performance of even a single cell under the presumed assumption is somewhat sufficient to model such process in respect to the reliability criteria.

The major task is to define chances of failure of individual subunits and topology of possible paths, to fix their assumptions for the purpose of large number of trials over a population model. The study pertaining to presumably large number of particular pathways with experimental substantiation is required to impose the hypothesis for the existing system. The random trial dependent Monte-Carlo approach considering the utilization of an ample of database have an effective potential to support the model functionality in lieu of the natural process, in a putative manner. The noise prone behaviour of the natural system dynamics will still be a challenging task in the direction of updating the neo configured artificial model. Any production system unit in cell is highly complex in its topological pathways and fairly non-linear in their interaction paradigm. So representation their sole interaction require indispensable computational efficacy to plot their multivariate routine.

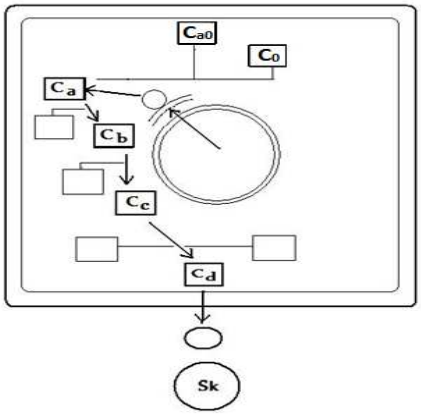


Figure 3.9. A model framework showing interacting process subunits

State vector of each cell can be represented by, **Eq. 26**,

$$dx/dt = f(\boldsymbol{\phi}, y, x) \quad (26)$$

$\boldsymbol{\phi}$ -Intracellular parameter

y -Extracellular state vector

x -Intracellular state vector

Interaction of intra-cellular parameter ($\boldsymbol{C_{a0}}$ to $\boldsymbol{C_d}$) has been shown in a topological model.

Model frame and its dynamics

The concept of reliability has its relevance in assessing the durability of the functionality of any mechanistic unit for a defined time frame. The overall idea is to support the associated pathway networks in view of conducting possible estimation of the enzyme streptokinase production taking into consideration as an instance. The time dependent interaction of few prominent parameters among the ample of parameters in natural cellular system has been given a desirable weight age in the developing model. The general theme was to make the model robust in terms of its configured subunits and dynamic pathways to ensure an optimal outcome in numerical terms.

3.5.2. Major assumptions in governing topology

The first assumption envisages over the fact that any production unit has initially maximum reliability and it decreases on the basis of interaction of several subunits in a specific topological fashion with variable dynamics. The varying consequence of transformation in topology can be taken into account to depict the observed level of productivity in any terms. Basic formula to estimate any existing system reliability R_S (having failure rate ξ) at a time t is given by the general formula;

$R_S = e^{-\xi * t}$

Three possibilities are assumed in this model for cellular system:

- The running original part with general decline in reliability (Frequently)
- Repair of the partially failed component at any time instant (Sometimes) (to resume some degree of reliability)
- Replacement of the completely failed component at the instant (Rarely) (to resume a high (initial) level reliability for that component again)

Overall the process is Stochastic

Basic Assumptions

In General, reliability is just the reverse of failure.

- Four pathways are considered in the model
- Failure of the support or completion of any two or more pathways at a time may lead to failure in the formation of streptokinase
- In all sorts of possible combinations failure of the assumed four pathways are possible
- Each component in the pathway has its specific lambda value

(*lbd*: failure rate or probability of failure with respect to time)

- The failure of the pathways has been made to be governed by Monte -Carlo method

The possible intracellular mechanistic framework is represented in the pathway topology, **Figure 3.10**. It was found that sometimes growth conditions have to change in order to increase structural stability (Samanta et al., 1998). The UV repair system and SOS repair pathway show their role for better structural plasmid stability (Greener, 1996). Plasmid replication mechanism Was found to have key influence on segregational plasmid stability (Sharma, 1993), (Bingle and Thomas , 2001). The sequences of ribonucleic acid RNAI and RNAII generally become the targets for the mutations (Moser and Campbell, 1983), (Lin-Chao et al., 1992), due to which their involvement in regulation and replication is effected. In reverse cases of mutation or deletion the condition leads to production of low plasmid copy and even impairment of replication (Phillips, 1998). So the biological reasons for the structural instabilities of vector are manifolds (Summers et al., 1993). Recent advances have been made in cell population modelling which allows the effects of cell heterogeneity in culture dynamics using intra cellular state and metabolite production (Henson, 2003).

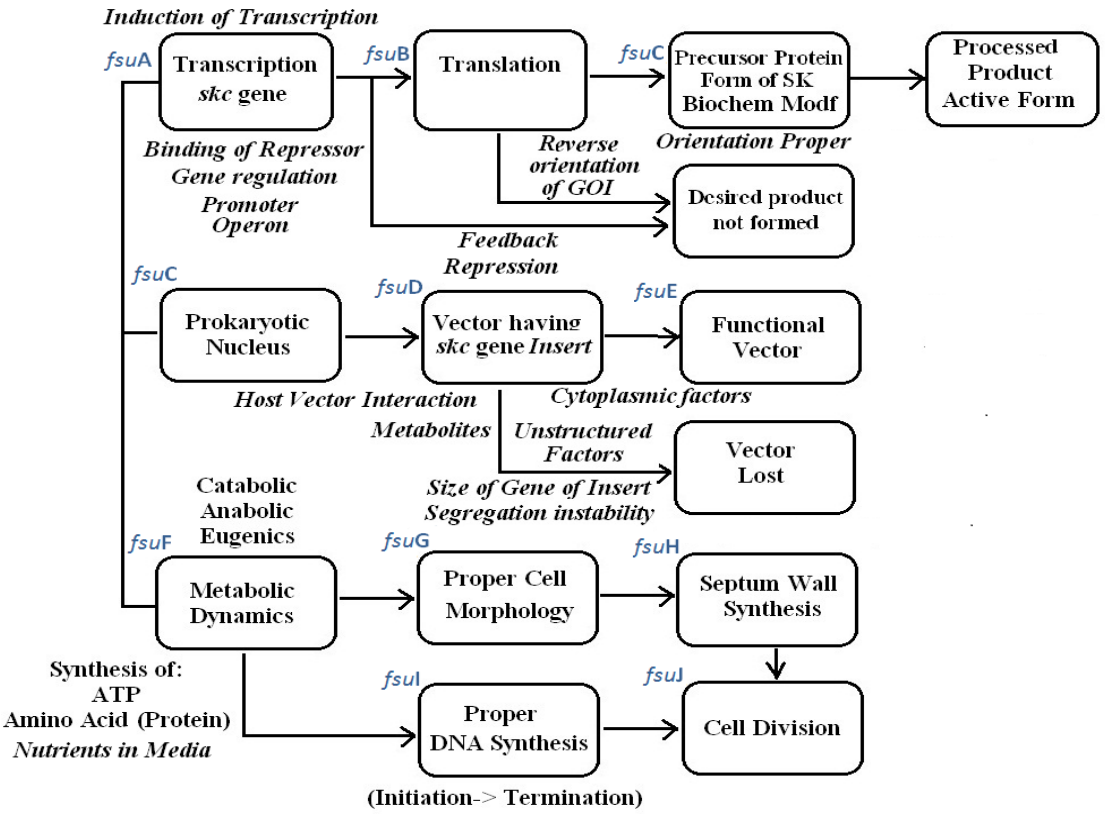


Figure 3.10. Pathways showing assumed interactions among components to facilitate the formation of resulting product (fsu=>functional subunit)

A chemically structured growth model of *E. coli* had been formulated and essential dynamic part of the model was considered (Palsson and Joshi, 1987). Abstract mathematical quantities lead to a particular interpretation when linear analysis is applied to chemical kinetics (Palsson, 1984). The secretory production of complex proteins particularly peri-plasmic proteins like proteases and chaperones can be manipulated to enhance the yield of secreted recombinant proteins (Choi and Lee, 2004). The segregation and selection both had been demonstrated to evaluate vector constructs that could be stabilized by partitioning factors (Ray and Skurray, 1984). Partitioning function is often the breakage function and denotes the probability of cell division of a cell with intracellular mass content (Fredrickson et al., 1970). The stoichiometric study of glucose and acetate metabolism of elementary compounds was found significant in respect to conversion kinetics framework pathway (Guardia and Calvo, 2001).

A few attempts had been made to construct a general single cell model including most of the intra-cellular pathways (Heinmets, 1969). According to Shuler and Domach (Shuler and Domach, 1982) the advantages of single cell-models were to account explicitly the cell in various respects. Particularly in respect to its geometry for potential effects on protein/nutrient transport, spatial arrangements of intracellular components, and temporal events during cell cycle, biochemical pathway and metabolic control models. Four pathways were configured showing interactions among various components of the recombinant cell during its active state, **Figure 3.10**.

3.5.3. Structured Pathways

Apparent Model Topology

The general topology of the structured pathways had been constituted using the biological interactive pathway model. The subcomponents are derived from their relevance in regulating metabolism. The topology of the model with interacting subcomponents has been illustrated below, **Figure 3.11.**

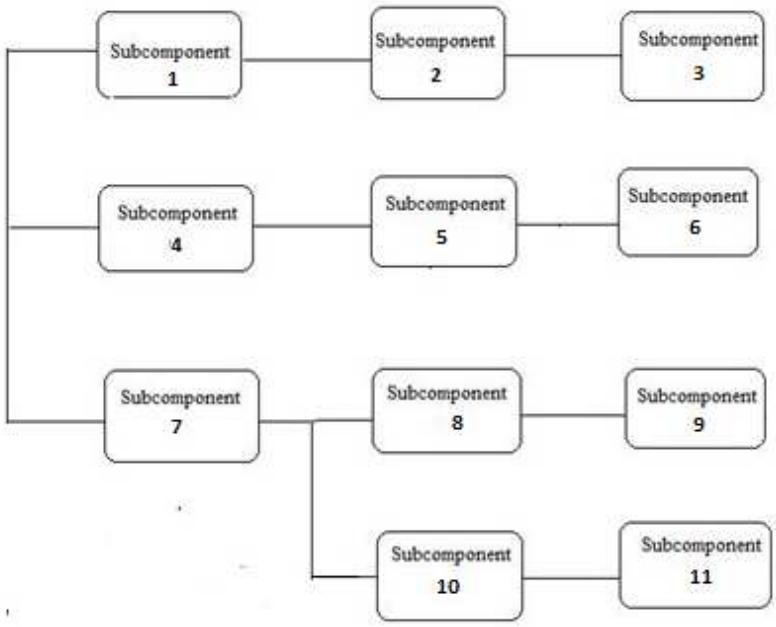


Figure 3.11. The assumed topology of heuristic based model with 11 subcomponents and 4 prime pathways

Combination of chances constitutes five different cases (considering six pathways: *a*, *b*, *c*, *d*), thereby four Pathways (each may having several units) related to a Product Formation i.e., their interaction together assumed to result into formation of the product;

$R_S(t) = r_s a(t) + r_s b(t) + r_s c(t) + r_s d(t)$ (27)

According to topology,

$R_S(t) = (r_s1.r_s2.r_s3) + (r_s4.r_s5.r_s6) + (r_s7.[(r_s8.r_s9) + (r_s10.r_s11)])$ (28)

Keeping in consideration, the existing pathway depicted, Figure **3.10**, the chances of failure at a junction was kept declining from $\xi 1$ to $\xi 11$.

Using the basic form, it would be computed as;

$$R_S(t) = e^{(-(\xi 1 + \xi 2 + \xi 3) * t)} + e^{(-(\xi 4 + \xi 5 + \xi 6) * t)} + e^{(-(\xi 7) * t)} .(e^{(-(\xi 8 + \xi 9) * t)} + e^{(-(\xi 10 + \xi 11) * t)}) \quad (29)$$

Model Assumptions in Detail:

Combination of chances constitutes five different cases (considering four pathways: P_a, P_b, P_c and P_d can be mentioned as 1, 2, 3 and 4 respectively):

1st case None fails: (1) Failure of none (as usually expected)

Failing pathways: [0]

2nd case One fails: (4) P_a or P_b or P_c or P_d

Failing pathways combination: [1 2 3 4]

3rd case Two fail: (6) P_aP_b, P_aP_c, P_aP_d, P_bP_c, P_bP_d, P_cP_d

Failing pathways combination: [12 13 14 23 24 34]

4th case Three fail: (4) $P_aP_bP_c$ $P_bP_cP_d$ $P_cP_dP_a$ $P_aP_bP_d$

Failing pathways combination: [123 234 341 124]

5th case Four fail: (1) only one possibility is there of getting all in failed state

Failure of all (very less occurrence)

Failing pathways combination: [1234]

The expressions, **Eq. 27-29**, were simulated incorporating the stated model assumptions. Individually each pathway has its reliability for production and as a system it showed a characteristic trend. The structured system dynamics simulation result is shown in **Figure 3.12**, on basis of Monte Carlo simulation method which is reflecting the declining stability of recombinant cells in cultivation.

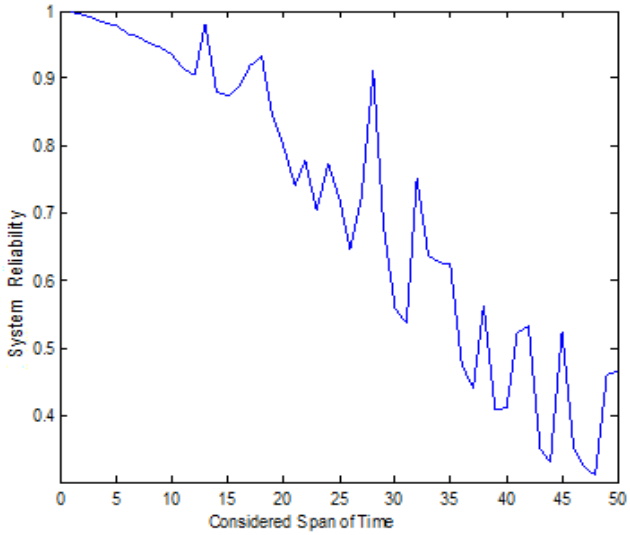


Figure 3.12. Reliability variation trend in stochastic manner shown for a cellular system

3.5.4. Supportive algorithms and property of existing regulatory networks

Stochastic assumptions are made to generate random decision to regulate the variable pathways by the predefined optimal criteria. Time dependent variables play

crucial role in governing the multidimensional dynamics thus preparing the model for showing close proximity to the real-time processes.

Incorporating regulatory control, which is a necessary characteristic to mimic the regulatory pathways showing their property of being autonomous in performing repair and other adaptability steps. Cellular intelligence is now-a-days a most excellent tool to amplify the information from cellular level activities in order to utilize it for higher level computational applications.

Incorporating regulatory control enable the models based on cells to utilize information gained from external stimulus or conditions, so doing that approach makes feasible for cybernetics modelling to overcome the rigidity and limitation of mechanistic modelling. Even quite complex models are required to describe adequately the metabolic dynamics of multicellular systems, especially under non-ideal conditions.

3.5.5. Applicability in estimating enzyme production

Present day logical models have their passive role in bioprocess modelling since untouched dimensions cannot be resolve unless applying hybrid approaches. Now recently cybernetics modelling has come forward to overcome the rigidity and shortcomings of mechanistic model which is too instrumental and conceptual.

It is true that cells have internal regulatory control to govern all biochemical pathways in a legitimate manner. Hence it coordinates and directs the adaptive machinery to cope up with external or extra-cellular variations maintaining the supportive mechanism that operates to serve simultaneously. It seems to be a good model for representing the complex biological processes and their associated vital events in terms of mechanical dynamics.

3.6. Discussion

The bioprocess simulation tool was constituted to visualize the predictive stages of the bioprocess operation. The antibiotic regulation strategy could be helpful in utilizing selective pressure resulting into the strengthening of recombinant cells population. The other bioprocess models can comparatively be taken into account to obtain the relative model dynamics in the different process operations.

Optimal allocation of metabolic pathways for producing the desired protein is a quite tough deed. This work can only be made feasible employing computational and stochastic approaches which easily support to constitute the variable pathways.

Identification of functional units within the network of pathways is a vital task to complete for giving a minimal framework. On the other hand optimal and reliable representation of the overall model is a mandatory task in this direction. Classical interpretation has always been appreciated to support and realize the emergent role of any component. Indeed constituting a framework revealing some link between intra-cellular and extra-cellular processes is reasonably of high utility to generate a robust model with an ample of indispensable dimensions.

Predictive models are now found to be of great help to forecast the model performance resolving its dynamics with several feature anticipating the outcome of different time dependent parameters and their interactions at different instances. Thus this sort of approach is of great help in understanding the variable facets of the model dynamics and their extent visualizing the apparent aspects arriving in the way.

Although sub-cellular topology based reliability model can be a supplement to our structured model system but it is considered to be novel approach pertaining to assessment of a cellular function in bioprocess system. It facilitates to design and identify the pathway failures during the operation which thereby quite helpful to improve the production system, since the population of cells would have been experiencing the same sort of complication. The structured fame work would be of worth in understanding the specific effect of external factor that has most significant influence in regulating and mal-functioning the intra-cellular dynamical system.

CHAPTER-4

PLASMID COPY NUMBER DYNAMICS

4.1. An Overview

This chapter of thesis deals with an important problem pertains to the regulation of plasmid copy number that has role in managing the production of recombinant enzyme streptokinase. Experiments using production medium was performed to achieve high production of biomass and maximum streptokinase activity. The basic work was primarily associated to estimate the plasmid loss and to evaluate the changes in copy number, if any, during the progress of the culture for which different factors were evaluated using the known experimental conditions.

It was evident from the outcome of data chart that plasmid copy number was apparently changing with time and depended on the changing environment in the chemostat system. The approximate estimation of the plasmid bearing and lacking cell population together with acetate, a primary metabolite concentration was required that had great importance for understanding the overall dynamics. It was found that a rapid hike in metabolite concentration at later stages of batch culture might lead to the abrupt decline of plasmid copy number, which was increasing previously. After performing several experiments in this direction, a plenty of facts regarding competitive trend of parameter dynamics had come forward to supplement our existing knowledge. Plasmid copy number dynamics was studied to evaluate the instability criteria, thus most significant consideration of the present work was to handle the challenge of plasmid instability with respect to process duration.

4.1.1. Earlier work

According to the known biochemical mechanism streptokinase binds to plasminogen forming an "activator complex" that converts plasminogen into the proteolytic enzyme plasmin. Streptokinase assays rely on its ability to activate plasminogen to plasmin. Plasmin then hydrolyses an indicator substrate and the extent of hydrolysis over a given period is related back to the concentration of streptokinase. The indicator substrates for plasmin include the fibrin clot, casein and other proteins including various synthetic esters.

The digestion of casein for estimating streptokinase had been established too early. The fibrin plate method originally introduced for determining proteolytic activity in blood has been widely used for measuring fibrinolysis. A radial caseinolytic method with the agarose gel containing both casein and plasminogen is commonly used (Saksela, 1981). In glucose grown culture there may be an effect of lactic acid accumulation which results in the lowering of pH, the acid tolerance response may result into a varying amount of recombinant streptokinase produced (Sriraman and Jayaraman, 2006).

High cell density can be obtained with high cell density cultivation of recombinant cell system by using a proper feed strategy. The T7 RNA polymerase was induced with IPTG to produce biomass and pulse feeding of concentrated substrate was done with different time intervals (Yazdani and Mukherjee, 1998). With an inducer concentration of 0.1mM IPTG, recombinant streptokinase accumulated was about 20% of the total soluble protein in the cell (Balagurunathan et al., 2008). Using microbial culture, assay for four different inoculum ages and four different inoculums concentrations were performed. The biomass obtained was analysed in terms of their chemical composition (Pelizer et al., 2003).

Batch cultures of *streptococci* had been characterized of having a rather extended lag phase followed by a relatively short period of exponential growth period. Comparison with continuous culture had revealed a lower productivity of the equivalent batch fermentation. In contrast, the production of streptokinase was maximized when glucose was in excess and the other nutrients were present in limiting amounts. There might be a number of factors that used to regulate the dynamics of plasmid carrying cells within the reactor. One such major factor was probability of plasmid loss due to segregation during cell division that could be described by segregative instability coefficient (Syamsu et al., 1992).

On the other hand, the likelihood of segregation for a plasmid bearing cell is independent of the frequency of plasmid-free cells generation in the population [66] while the selection intensity against plasmid carriage is influenced by occurrence of plasmid lacking cells in population. Occurrence of point mutation in the gene for RNA I might leads to high copy type due to change in replication control of vector (Boros et al., 1984). Several other mutations are known to enhance the copy number (Müller et al., 1995). The interaction between two RNA's, tRNA and RNA I consequently diminishes the inhibitory effect of the later one and this leads to the phenomenal increase in the copy up to the time

when metabolic capacity of the host cell is exhausted (Kramer et al., 1996), (Cserjan-Puschmann et al., 1999). In cell division process occasionally a daughter cell results that doesn't contain the plasmid and can no longer produce the desired product (Stephens et al., 1992). Modulation of an input such as nutrient concentration or the cell environment such as the pH could enhance the rate of biochemical reactions that were occurring (Silveston et al., 2008). Excessive plasmid replication is observed after induction, it greatly contributes to the metabolic overload of the cell (Grabherr et al., 2002). In genetic study the fraction of repressor-free operators affect multi-copy plasmids containing the *lac* promoter-operator to increase plasmid copy number (Lee et al., 1984d). Also the compartmental model used to describe the variation in intra cellular RNA and copy number with varying specific growth rate (Nielsen et al., 2008). Moreover, the decrease of specific growth rate with increasing plasmid content was observed in mechanistic way due to enhancing metabolic burden. The vector copy number with binding affinity of tRNA and RNA I in relaxed strain were assumed that were starved for some amino acids (Wang et al., 2002). Chemostat cultivation conditions were found to work for stable genetic inheritance of vectors in the hosts (Noack et al., 1981). The consequence of theta and rolling circle modes of replication on maintenance of recombinant bacterial strain with their plasmid copy were studied (Kiewiet et al., 1981).

4.1.2. Relevance of Evaluating Plasmid Copy Number

Study of plasmid copy number dynamics has its relevance in pointing out the span of bioprocess in which the copy number is high. The specific growth rate of recombinant cells and average plasmid copy per cell can be of much significance to harvest high yield from the fermentation process. There are several factors that were found affecting plasmid copy number during fermentation process. The structured as well as unstructured factors are collectively supposed to influence the copy number dynamics.

4.2. Materials and Methods

4.2.1. Materials

The overall general requirement for conducting the experiment is given below,

Recombinant strain and vector: *E. coli* BL21 as host with vector pRSET-B was used.

4.2.1.1. General Requirements

Glassware requirement: Micropipettes, Conical Flasks of 100, 250, 1000 ml and 3000ml were required. Beaker 250ml, Test-tube and disposable Petri-Plates were also used.

Miscellaneous requirement: Aluminium foil, sterilized cotton, inoculation needle, lamp, tissue paper, permanent marker, strider, tooth-pick, paraffin paper, vacuum pump assembly, filter 0.3µ, syringe etc were needed.

***Growth Media*:** LB media composition (Tryptone, Yeast Extract and NaCl)

***Production Media*:** Major and trace salt components

Major Components: K_2HPO_4, KH_2PO_4, NH_4Cl, NaCl, $MgSO_4$ and $CaCl_2$

Trace Components: $Al_2(SO_4)_3.7H_2O$, H_3BO_3, $CuSO_4.H_2O$, $MnCl_3.4H_2O$, $NiCl_2.6H2O$, $Na_2MoO_4.2H_2O$ and $ZnSO_4.7H_2O$ etc

***Antibiotics*:** Ampicillin

***Inducer*:** IPTG for induction

ddH_2O was used for media preparation

4.2.1.2. Requirement for Streptokinase Assay (Caesinolytic assay)

- Plasminogen
- Pure Streptokinase
- Tris-HCL Buffer
- Caesin milk powder (Analytical grade)
- Sodium Azide
- Agarose
- Chromozym PL

4.2.1.3. Requirement for Plasmid Isolation

Zyppy Plasmid Miniprep Kit was required.

It contained Lysis Buffer, Neutralization Buffer, Endo-wash Buffer, Wash Buffer, Elution Buffer, RNAse, IIN Columns and Collection Tubes.

4.2.1.4. Equipments Required

Bioreactor system (New Brunswick Scientific with *Bio-command* software) and accessories, Autoclave, Laminar flow, Refrigerator, Centrifuge, Shaker, UV-

spectrophotometer, SDS-PAGE assembly, Agarose Gel Electrophoresis Unit, Compound Microscope, Lyphilizer etc.

4.2.1.5. Organism Strain and Vector Used

Bacterial recombinant strain used in our experimental work was *E. coli* BL21 having shuttle vector pRSET-B. The organism was preserved in glycerol stock at -80^{0}C. For working, time to time maintenance media was prepared using pure colony. The vector map of pRSET-B showing important features is depicted in **Figure 4.1**. The detail about the vector properties was described in **Table 4.1**.

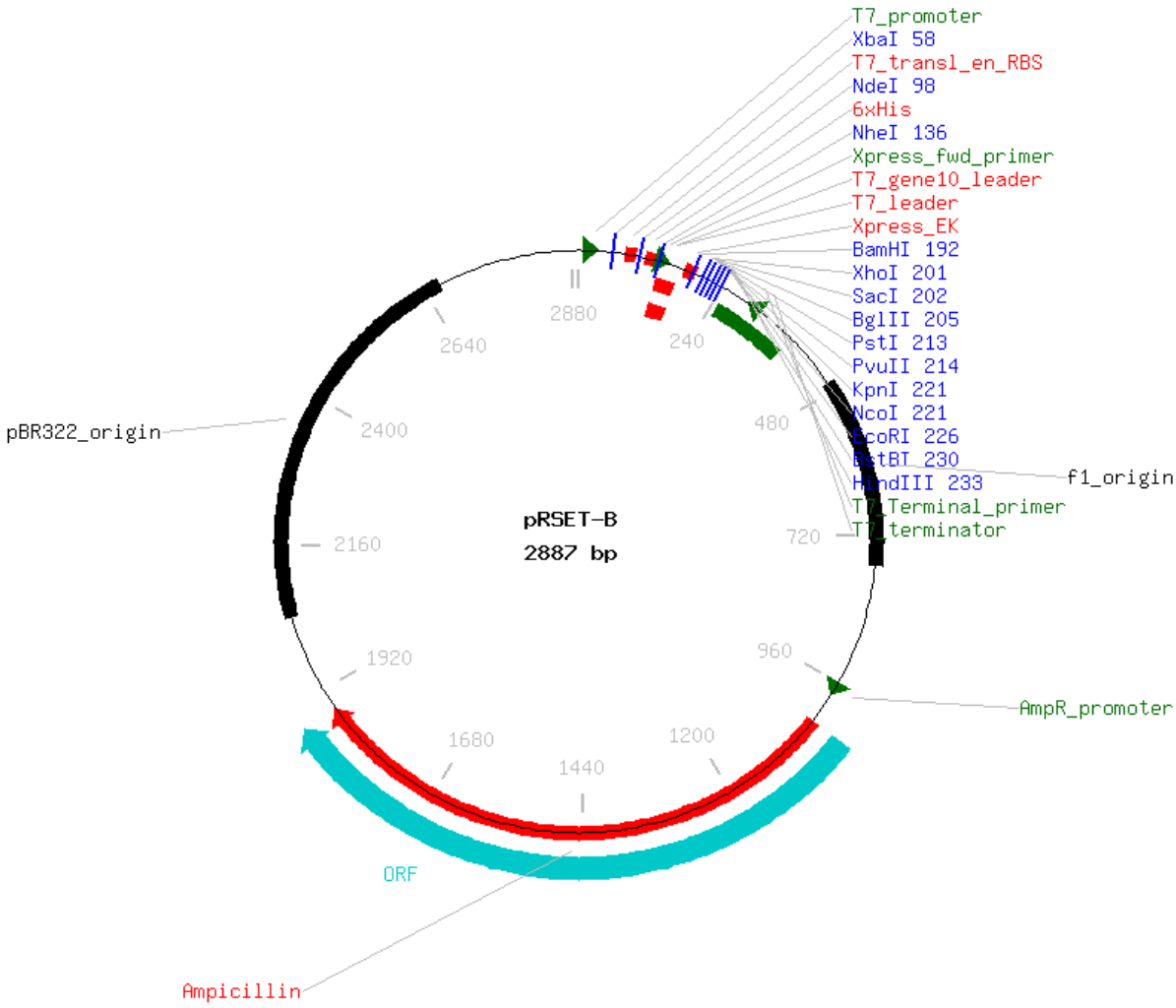


Figure 4.1. *pRSET-B* Vector Map

Table 4.1. Important features of pRSET-B vector

Features	Advantage
T 7 promoter	Provides tight, dose-dependent regulation of heterologous gene expression. Provides a binding site for most T7 promoter primers for sequencing into the insert.
Multiple cloning site	Allows insertion of our gene of interest and facilitates cloning in frame with N-terminal epitope tag
Ampicillin resistance gene (β-lactamase)	Allows selection of the plasmid carrying *E. coli* cells.
pUC origin	Meant for high copy replication and growth in *E. coli*.

4.2.1.6. Various Media Used

A. Growth Media

E. coli cells were routinely grown and maintained in LB medium (Tryptone, 10g/l; yeast extract, 5g/l; NaCl, 10g/l) containing 100μg Ampicillin per ml at 37 °C. The pH of the medium was always adjusted to 7.0 with 0.1N NaOH and 0.1N HCl solution. When required, the medium was solidified by adding 15g/l agar before autoclaving to prepare the maintenance media.

B. Production Media

The production media constituents (Yazdani and Mukherjee, 2002) taken for producing the recombinant enzyme in shake flask and bioreactor batch culture is described below under the heads of Major and Micro-elements, **Table 4.2(a)** and **4.2(b)**;

Table 4.2(a). Major-elements

Constituents	Amount (g/l)
K_2HPO_4	6.0
K_2HPO_4	3.0
NaCl	0.5
NH_4Cl	1.0
Yeast Extract	5.0
Glucose	5.0
$MgSO_4$	2.0
$CaCl_2$	1.0

Table 4.2(b). Micro-elements

Constituents	Amount (mg/l)
$FeSO_4$	100.0
$Al_2(SO)_4.7H_2O$	10.0
$CuSO_4.H_2O$	2.0
H_3BO_3	1.0
$MnCl_3.4H_2O$	20.0
$NiCl_2.6H_2O$	1.0
$Na_2MoO_4.2H_2O$	50.0
$ZnSO_4.7H_2O$	5.0

The pH of the solution was adjusted to 7.0 with 0.1N NaOH and 0.1N HCl and then autoclaved at 121^0C and 15 *lb* for 20 minutes as usual.

4.2.2. Methods

The streptokinase production could be efficiently done understanding the behaviour of the population and plasmid copy dynamics. Logical models were now being developed after identifying the missing indispensable factors in existing models. In order to impart the effect of relevant parameters and supplement the existing knowledge

experimental support was required. The improvement in the bioreactor performance was possible using experimental methodology.

Now in order to make effort to reveal the characteristics of a natural non-ideal bioreactor system and making the ongoing models closer to our experimental setup, the endeavour would be to evaluate and adjust the numerical values of the parameters to get maximum response. The constituent of the media and culture conditions were being given with reference to the earlier works in this field. The sole objective of understanding the system behaviour was to improve the performance of the existing bioreactor related models.

The Streptokinase producing cloned strain of *E.coli* BL21 had been taken with plasmid vector pRSET-B, it was obtained from Dr. V. Murugan's Lab, Department of Biotechnology, Anna University, Chennai. This vector was carrying ampicillin marker hence this antibiotic was used in preparing growth media for promoting the growth of selective microbial strain. Having *lac* promoter the induction of recombinant protein expression was effected by addition of the non-metabolically *lac* inducer IPTG, it was to be added when OD reached just above to a level, generally 0.6 to 0.8. Production media which contained glucose, major and trace metal salts, and yeast extract, as mentioned in **Table 2 (a)** and **(b)** that was found suitable to get the desired level of product. The response in terms of maximum biomass formation and streptokinase activity was considered to be relevant in this direction.

The culture condition could be assumed to be responsible for maximum productivity since other constraints were kept constant. The goal was to understand the plasmid bearing and lacking cell dynamics at the given culture conditions like pH and agitation that can be done at shake flash level or in a bioreactor. The vital steps in the methodology constituted of the following techniques;

- ***Bacterial Culture to Estimate Cell Population***
- ***IPTG Induction***
- ***Streptokinase Assay for Estimation of Streptokinase***
- ***Plasmid Isolation and Quantification***
- ***Population Screening***

4.2.2.1. Bacterial Culture to Estimate Cell Population

E. coli Growth Curve was made using the cell concentration readings at 600nm wavelength using UV Spectrophotometer.

- Since ampicillin containing plates were to be made (having final conc. 100µg/ml in the plate) so for this purpose stock solution was prepared in 10ml ddH_2O adding 1g of pure ampicillin powder in sterilized environment and kept as stock at -20^0C.
- A starter culture was prepared by inoculating a single colony from a freshly streaked selective plate into 2–10 ml LB medium containing the ampicillin antibiotic. It was allowed to grow at 37°C for about 8 hours (having logarithmic growth phase) with shaking (at 200 rpm) in incubator-shaker.
- Diluted the starter culture, 1/500 to 1/1000 into a larger volume of selective LB medium,
- The culture was grown at 37°C at 200 rpm for 12–16 hours.
- The bacterial culture was now harvested 12–16 hours after inoculation. It was harvested by centrifugation at 6000 X g for 15 min at 4°C. All traces of supernatant were removed by inverting the open centrifuge tube until the entire medium would be drained. The cells were now ready for the lysis procedure, as indicated in the appropriate plasmid purification protocol.
- The growth curve of an *E. coli* culture could be divided into distinct phases.
- The first, lag phase occurred directly after dilution of the starter culture into fresh medium. During this phase, cell division was slow as the bacteria adapt to the fresh medium.
- The bacteria then started to divide more rapidly and the culture entered logarithmic (log) phase (4–5 hours after dilution), during which the number of cells increased exponentially.
- As the available nutrients in the medium were used up and released metabolites inhibited bacterial growth, the culture became saturated and entered stationary phase (in about 16 hours after dilution), during which cell density used to remain constant.
- Eventually the culture entered into declining phase cells started to lyse, the number of viable bacteria turned down, and DNA became partly degraded.

- The growth curve of a bacterial culture had been monitored photo-metrically by reading the optical density at 600 nm.

4.2.2.2. IPTG Induction

A stock solution of IPTG was prepared in ddH_2O of strength 100mM (2.38g/100ml). It was then filtered with syringe and sterilized. The starting point to perform the IPTG induction was to obtain an inoculated freshly prepared culture. The desired OD_{600} was measured in the range of 0.5-0.8 at which IPTG induction had been added. One more flask had to be taken just as control in which IPTG was added. The starting reading should be same for both the cultures.

The method of IPTG induction suggested by researchers (Thangadurai et al., 2008) was slightly modified to obtain better expression in present case. An IPTG concentration of 0.8mM was used instead of 1.0mM IPTG to get better expression of *skc* gene.

4.2.2.3. Streptokinase Assay for Estimation of Streptokinase

Two types of assays are mention in literature one is Caesinolytic assay and the other one is enzymatic assay. A simple, easy, sensitive and reproducible method, Caesinolytic assay (Saksela, 1981) was developed for detection of plasminogen activator in cell growth media and cell extracts. Samples to be assayed were applied into wells cut in opaque casein containing agarose gels with and without plasminogen. Diameters/area of the clear circular zone (plaque) resulting from proteolysis around the wells were proportional to the amount of applied plasminogen activator.

Firstly, agarose was made to final concentration of 1% in 0.1M Tris-HCl buffer, pH 8.0 [38]. Now purified plasminogen was added just before preparation of the plates to give a concentration of 2µg protein per millilitre into agarose melted and kept at 45^0C. Non-fat dry milk served as the source of casein and it was stored frozen in 15% (w/v) stock solution in the tris-HCl buffer. The final concentration of milk powder in the test plate was made to be 0.6%. Lastly sodium azide was added to a final concentration of 0.1% in order to prevent the microbial growth. On the other hand, no plasminogen was added into the control plates. Plasminogen was obtained from MP biomedical (25U, 1mg vial). It was then diluted in 0.05M Tris HCl buffer to 0.05mg/ml was kept as stock.

While preparing the plates, the mixture was rapidly poured into petri-plate and a 2mm thick gel was made. After cooling a suitable, 3mm diameter size wells were punched into the gel. Later 10µl of samples were filled in the well. One such plate was suitable for the analysis of 3-4 samples. During incubation, the gels were placed at 37^0C and the results were recorded after 48 h of incubation.

The plaques size (i.e., diameter) was measured using finely graduated scale, to take the accurate size dimension diameter was measured from two sides to get the average size of plaque zone and its area was considered in final reading. Pure Streptokinase (from MP Biomedicals, 10kU, 2mg in vial) with known concentration had been used to draw a standard curve. For this purpose pure streptokinase was diluted in Tris-HCl buffer (0.15M) and different dilution viz. 25U, 25U, 100U, 250U, 500U, 1000U were done. Now, since the plaque size for different concentrations was known so using the regression approach the unknown sample concentrations were calculated.

4.2.2.4. Plasmid Isolation and Quantification

Plasmid isolation was done for samples obtained at various stages of the culture to know the plasmid content at different stage of the culture process. Firstly the plasmids samples were isolated (using Prolab protocol) then diluting it to 1:1000 times with ddH_2O. It was then, taken for measuring the OD through UV Spectrophotometer at 260nm wavelength in order to quantify the plasmid concentration.

A. Principle Involved in Plasmid Isolation

Exposure of bacterial suspension to strongly anionic detergents at high pH opens the cell wall, denatures chromosomal DNA and proteins, and releases plasmid DNA into the supernatant. Although the alkaline solution disrupts the base pairing, the strands of closed circular plasmid DNA are unable to separate from each other because they are topologically inter-wined as long as the intensity and duration of exposure to -OH ions is returned to neutral. During lysis, denatured chromosomal DNA becomes enmeshed in large complexes that are coated with dodecyl sulphate. These complexes are efficiently precipitated from solution when Na^+ ions are replaced by K^+ ions. After denatured material has been removed by centrifugation, native plasmid DNA can be recovered from the supernatant**.**

B. Procedure for Plasmid Isolation

In our experiment, the plasmid isolation step had been conducted using Zyppy Plasmid Miniprep Kit. The sub-steps followed are shortly mentioned as under;

- 100µl volume of 7X Lysis Buffer was added to 600µl of *E. coli* culture in a 1.5ml microcentrifuge tube was then mixed by inverting the tube 4-6 times.
- 350µl of cold Neutralization Buffer was added and mixed thoroughly.
- Centrifuged at 11000-16000X g for two minutes.
- The supernatant was then transferred into Zymo-spin IIN column provided with the kit.
- The column was placed into the collection tube and centrifugation was done for 15 seconds. The flow was discarded through it and placed the column back into the same collection tube.
- Endo-Wash buffer of 200µl volume was added to the column. It was now centrifuged for 15 seconds.
- Again added 400µl of Zyppy Wash buffer to the column and then centrifuged for 30 seconds.
- The column was then, transferred into a clean 1.5ml centrifuge tube then added 30µl of Zyppy Elution buffer directly to the column matrix and it was let to stand for one minute at room temperature. Then it was centrifuged for 15 seconds to elute the DNA.

C. Plasmid Quantification

Plasmid quantification was done using the isolated plasmid sample from 2.5ml of fresh *E. coli* culture sample. So to get better yield replica of the samples were taken and isolation had been done using Zyppy Plasmid Isolation kit. After isolating the plasmid the yield was mixed with ddH_2O and OD was taken at 260nm using UV spectrophotometer. Likewise OD had been taken for each sample from different stages of the culture process.

4.2.2.5. Population Screening

Population screening was done using cell culture at different stages. A small volume of culture i.e., 0.1ml at regular time gap of 2 h had been collected and spread on separate non-ampicillin plates after dilution of 10^{-4} times, then on the basis of *cfu* (colony

forming unit), individual colonies were selected and transferred over grid marked plate which contained ampicillin. So likewise the process was performed for different collected samples separately. Using this method non-amp and amp resistant colonies could be counted and the percentage of cells which were plasmid bearing was sorted out. This was helpful in estimating the population and differentiates the recombinant population at different stages of the culture.

4.3. Results

Assay of streptokinase enzyme was done using the method described in steps of methodology, reading of shake flask and bioreactor (batch process) has been shown separately.

4.3.1. Standard Curve using plaque size of pure Streptokinase

Table 4.3. Values for Standard Curve, as plaque size for pure streptokinase

Serial No.	Pure Streptokinase Sample	Plaque Size (diameter in mm)
1.	25U	5.0
2.	50 U	7.0
3.	100U	12.0
4.	250U	18.0
5.	500U	27.0
6.	1000U	33.0

4.3.2. Streptokinase assay and finding of unknown concentration using standard

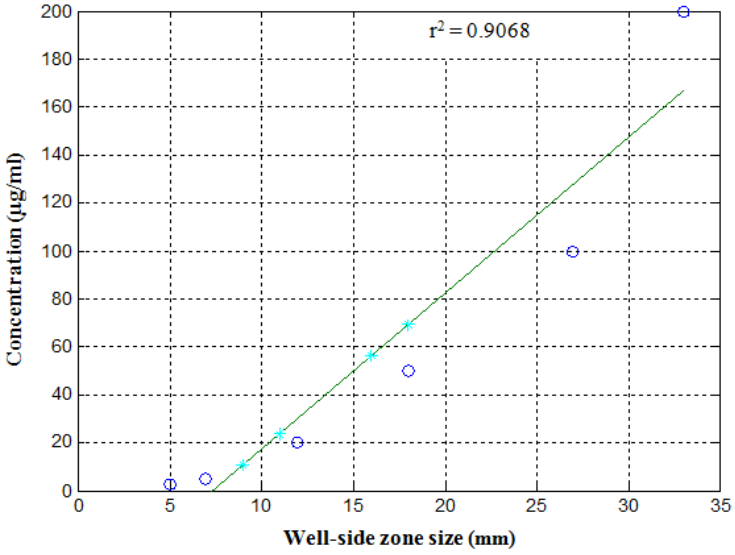


Figure 4.2. Standard curve for streptokinase assay

Pure streptokinase had been diluted using Tris-HCl buffer to draw a standard curve. The corresponding concentration for units of streptokinase, **Table 4.3**, from serial no. 1 to serial no. 6 would be 2.5µg/ml, 5µg /ml, 20µg/ml, 50µg/ml, 100µg/ml and 200µg/ml respectively. The standard plot was made with the help of pure streptokinase sample, **Figure 4.2**.

4.3.3. Growth of E. coli Cells and Streptokinase Production (in Shake Flask)

Table 4.4(a). Results of shake flask study for streptokinase production

S. No.	Time after induction (hours)	Concentration of streptokinase (µg/ml)
1.	2	10.9
2.	4	56.41
3.	6	69.42
4.	8	23.91

Table 4.4(b). Results of batch bioreactor studies for streptokinase production

S. No.	Time after induction (hours)	Concentration of streptokinase (µg/ml)
1.	2	4.4
2.	4	43.41
3.	6	36.91
4.	8	85.67
5.	10	10.90
6	12	1.15

Readings taken to plot out overall cell concentration in terms of OD. Weight of a single *E. coli* is about 7×10^{-13}g. On this basis the weight of *E. coli* to the respective cell OD was calculated, **Table 4.5(a)**.

Table 4.5(a). Cell growth reading from Shake Flask experiment

S.No.	Time (hours)	Absorbance (600nm)	Number of *E. coli* cells (per ml) x 10^{12}
1.	0	0.0175	0.0250
2.	1	0.0391	0.0558
3.	2	0.0903	0.1290
4.	3	0.2459	0.3510
5.	4	0.4350	0.6210
6.	5	0.6590	0.9410
7.	6	0.9896	1.4130
8.	7	1.1687	1.6690
9.	8	1.2228	1.7460
10.	9	1.2243	1.7490
11.	10	1.2152	1.7360
12.	11	1.1806	1.6860

Curve was plotted, **Figure 4.3**, showing cell population growth and streptokinase activity on the same scale to understand the dynamics.

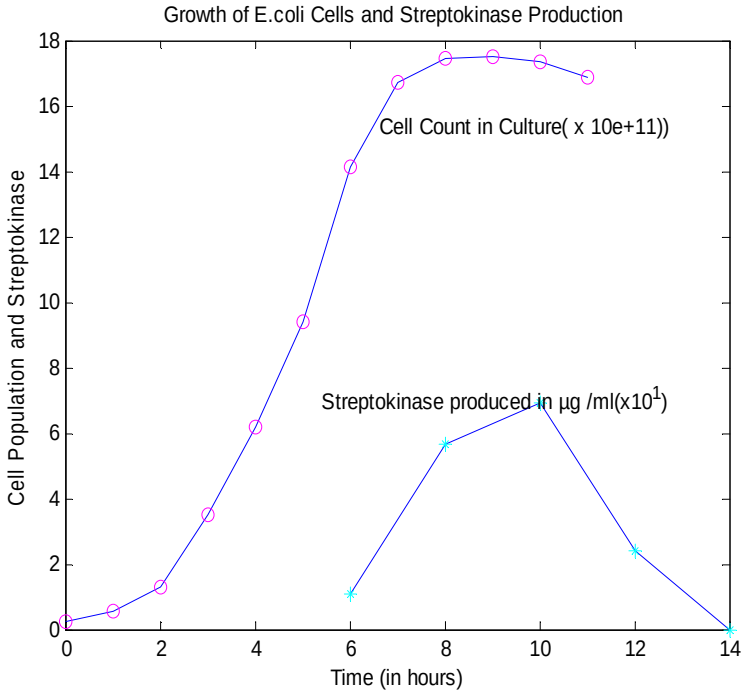


Figure 4.3. Cell growth and streptokinase production profile in shake flask

4.3.4. Growth of E. coli Cells and Streptokinase Production (in Bioreactor)

Table 4.5(b). Cell growth reading from Bioreactor

S. No.	Time (hours)	Absorbance (600nm)	Number of *E. coli* cells (per ml) x 10^{12}
1.	0	0.0078	0.0114
2.	1	0.0478	0.0683
3.	2	0.4123	0.5890
4.	3	0.4354	0.6220
5.	4	0.6086	0.6694
6.	5	0.7743	1.1061
7.	6	0.8115	1.1593
8.	7	1.0070	1.4386
9.	8	1.1120	1.5886
10.	9	1.2936	1.8480
11.	10	1.3230	1.8900
12.	11	1.4234	2.0334
13	12	1.4519	2.7410
14	13	1.6234	2.3191
15	14	1.7615	2.5164

The experiment was carried out in bioreactor (batch mode), the obtained values were given in Table 4.5 and plotted (Figure 4.4).

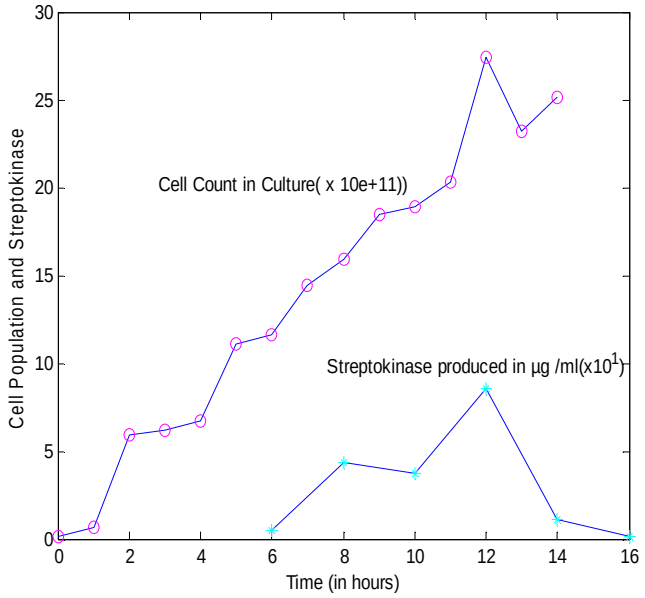


Figure 4.4. Cell growth and streptokinase profile in batch bioreactor studies

4.3.5. Vector Quantification

Plasmid concentration was measured in terms of OD as shown in the following **Tables 6(a)** and **(b)**. Firstly plasmid was isolated from 1ml of culture media then it was diluted in 1ml ddH_2O and absorbance was measured. Plasmid concentration was derived from absorbance directly multiplying it by 50µg/ml, which had been the known plasmid concentration at OD 1.0 and $A_{260.}$

Since each pRSET-B vector consists of 2887basepairs (additionally 2.5kb gene insert), so each base pair is having known mass of 660daltons. The calculation was done using (Kramer et al., 1996) established expression, hence gram weight of each plasmid copy would be,

$\{(2887+2500) \times 660\}/(6.023 \times 10^{23})$

$= 5.9039 \times 10^{-18}$g

Table 4.6(a). Plasmid quantification in Shake Flask experiment

S. No.	Time (hours)	Absorbance (260nm)	Plasmid Concentration (µg/ml)	Number of Plasmid Copy per ml reaction volume x 10^{13}
1.	0	0.0012	0.060	0.0896
2.	2	0.0117	0.585	0.1849
3.	4	0.0158	0.790	0.2497
4.	6	0.0232	1.160	0.3666
5.	8	0.0336	1.680	0.5310
6.	10	0.1100	5.500	1.7385
7.	12	0.0068	0.340	0.1074
8.	14	0.0128	0.640	0.2023
9.	16	0.0220	1.100	0.3477
10.	18	0.0214	1.070	0.3382

So the plasmid quantification in case of shake flask and bioreactor was done employing the expression and finally the plasmid copy per unit volume was obtained in each case, **Table** 4.**6 (a)** and **(b)**.

Table 4.6(b). Plasmid quantification in Bioreactor medium

S. No.	Time (hours)	Absorbance (260nm)	Plasmid Concentration (µg/ml)	Number of Plasmid Copy per ml volume x 10^{13}
1.	0	0.0019	0.095	0.0300
2.	2	0.1149	5.745	1.8160
3.	4	0.1513	7.565	2.3913
4.	6	0.2034	10.170	3.2147
5.	8	0.1409	7.045	2.2269
6.	10	0.1229	6.145	1.9424
7.	12	0.1014	5.070	1.6026
8.	14	0.1045	5.225	1.6516
9.	16	0.09080	4.540	1.4351

4.3.6. Population Screening

The screening of population of recombinant and plasmid lacking cells was carried out on colony basis as mentioned in *sub-section 4.2.2.5* of this chapter. The experiments were conducted triplicate for batch accuracy; the results obtained were given in **Table 4.7 (a)** and **(b)**.

Table 4.7(a). Recombinant cells screening, in Shake Flask medium

S. No.	Time (h)	Number of Plasmid Bearing Cells Colonies (on *cfu* basis)	Number of Plasmid Lacking Cells Colonies (on *cfu* basis)
1.	0	31	2
2.	2	38	4
3.	4	30	4
4.	6	42	2
5.	8	33	4
6.	10	36	6
7.	12	39	12

Table 4.7(b). Recombinant cells screening in medium from Bioreactor

S. No.	Time (h)	Number of Cell Colonies (on *cfu* basis)	Number of Plasmid Lacking Cells Colonies (on *cfu* basis)
1.	0	43	1
2.	2	36	4
3.	4	40	4
4.	6	43	3
5.	8	40	5
6.	10	41	6
7.	12	44	9
8.	14	42	20

4.3.7. Average Copy Number per Cell

To find the Average Copy Number

Plasmid bearing fraction of cells =

(Total colonies of cells - Plasmid lacking cell colonies)/ Total colonies of cells

Copy Number (Average) at an instance =

Total Plasmid copy / (Plasmid bearing fraction x Total cells presents)

Table 4.8(a). The Average Copy Number (in Shake Flask Experiment)

S. No.	Time (hours)	Plasmid Bearing *E. coli* cells (per ml)x 10^{10}	Plasmid copy present (per ml) x 10^{13}	Plasmid Average Copy number per cell
1.	0	2.3387	0.0896	16.0 (16)
2.	2	11.5421	0.1849	38.3 (38)
3.	4	53.8200	0.2497	06.6 (7)
4.	6	134.5714	0.3666	05.7 (6)
5.	8	153.4364	0.5310	05.5 (5)
6.	10	144.6667	1.7385	12.0 (12)
7.	12	116.7213	0.1074	00.92 (1)

The plasmid average copy number per cell was finally obtained in each case, **Table 4.8 (a)** and **(b)**, at different time instances after induction in fermentation process.

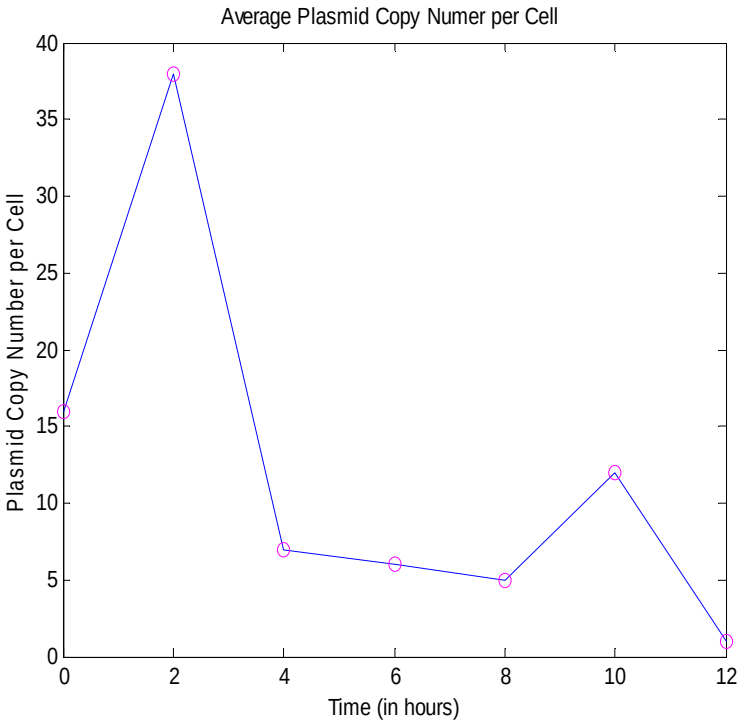


Figure 4.5. Characteristic curve of average plasmid copy number for shake flask study

Together with the count of plasmid bearing cells and plasmid copy present in unit volume of culture, average copy number was computed and illustrated in the plot, **Figure 4.5**.

Table 4.8(b). The Average Copy Number (in Bioreactor Batch Process)

S.No.	Time (hours)	Plasmid Bearing *E. coli* cells (per ml)x 10^{10}	Plasmid copy present (per ml) x 10^{13}	Plasmid Average Copy number per cell
1.	0	1.1135	0.0300	22.9 (23)
2.	2	52.3556	1.8160	34.7 (35)
3.	4	60.2460	2.3913	39.7 (40)
4.	6	107.8419	3.2147	29.8 (30)
5.	8	139.0025	2.2269	16.0 (16)
6.	10	161.3415	1.9424	12.0 (12)
7.	12	218.0341	1.6026	07.0 (7)
8.	14	131.8114	1.6516	12.0 (12)

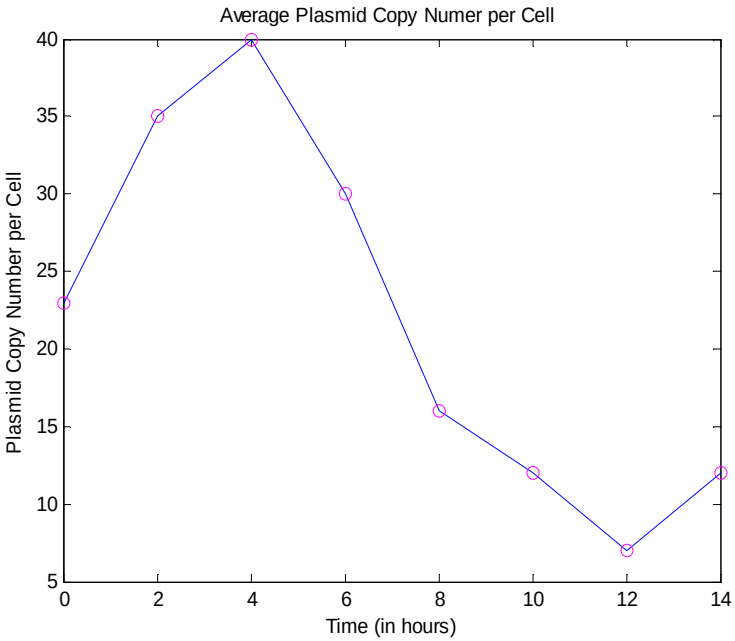


Figure 4.6. Characteristic curve of average plasmid copy number (in Bioreactor)

The curve for average copy number per cell, **Figure 4.6**, was obtained on the basis of plasmid bearing cells count and plasmid copy present in unit volume of bioreactor medium. Experiments were carried out to see the effect of metabolite concentration (in terms of acetate formation) in shake flask (Guardia and Calvo, 2001) and the data is shown in **Figure 4.7.**

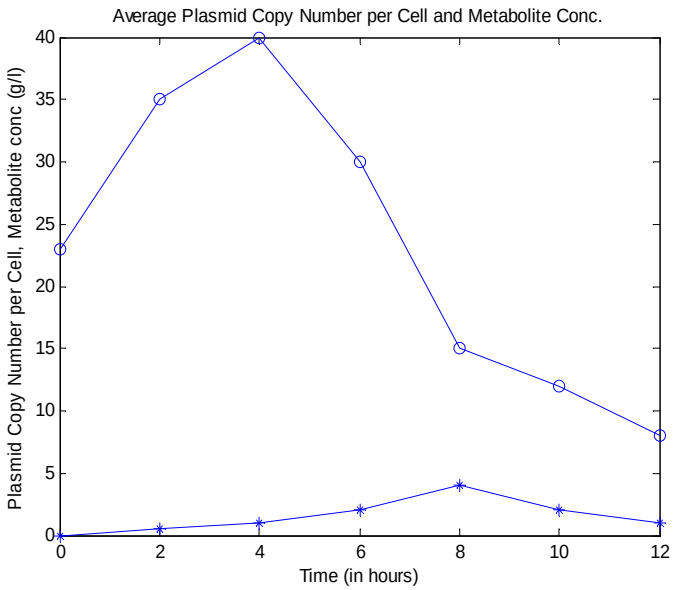


Figure 4.7. Effect of metabolite concentration (asterisk) on average plasmid copy number per cell

4.4. Discussion

Inference could be drawn from the experiments conducted using samples from different time steps viz. overall cell count through OD and DCW, plasmid density, streptokinase expression, plasmid bearing and lacking cells count through population screening. It was observed that plasmid bearing cells and copy number population were going on declining after certain time span with the progress of recombinant culture process. The comparison of shake flask and bioreactor results had been done to evaluate the change in plasmid copy dynamics more firmly.

The dynamics pertaining to plasmid average copy number showed that copy number showing a decreasing trend after 4-6 hours and it became negligible in magnitude after 12 hours of culture time. There found a little hike in average copy number in late hours of batch cultivation due to inertial reasons which work on decline of metabolite concentration.

Experiments using production medium was performed to achieve high production of biomass and to get maximum product i.e., activity. The basic work was primarily associated to estimate the plasmid loss and to evaluate the changes in copy number. From the outcome of chart it is evident that plasmid copy number was apparently changing with time and it depends upon the changing environment in the chemostat system. The approximate estimation of plasmid copy number was required which would be of greater importance for understanding the instability dynamics. After performing several experiments in this direction a plenty of facts regarding competitive trend of dynamics had been emerged. Also role of parameters like primary metabolite were established in a simple way for its contribution in plasmid copy number decline with culture process. It was evident from the experiments that bioreactor supports for a better sustenance of cell biomass and copy number and to a higher magnitude in comparison to the shake flask.

Since the formation of product streptokinase would be directly proportional to the number of recombinant cells and vector copies per cell in the media so our emphasis here was to consider plasmid copy number into our account to study the variations. Moreover the idea was to exploit the dynamical features concerned to structural reasons particularly genetic basis that indirectly govern the copy number sustenance in production medium.

CHAPTER-5

MEDIA AND CULTURE CONDITION OPTIMIZATION

5.1. An Overview

This chapter addresses another interesting area of research in bioprocess system involving efficient optimization of culture conditions together with media. It depicts the use of statistical approach to improve the product formation via test of model and prediction methodology. After performing the task of optimization of production media components the next aim had been to optimize the culture-condition governing parameters which had been done subsequently and success was achieved to an extent.

Streptokinase assay was performed to test its activity in culture supernatant fluid which was done by comparison with dilutions of standard purified streptokinase solution, using casein/plasminogen plate technique. Plackett Burman (PB) design was prepared to screen out media components using experimental results. Central Composite Design (CCD) matrix application was carried out to determine which of the several experimental variables were significant and their interactions present more significant effects. Design Expert Software, DX5 and Matlab installed system were used for conducting the optimization work. ANOVA (Analysis of variance) was done for analysing the second order response surface model; the significance of each coefficient was determined by Student's *t*-test and *p*-value, to identify the corresponding significant factors.

Screening of media components was firstly conducted using PB, since a number of media components were used to constitute production medium to carryout the fermentation. The next step was performed as CCD analysis to estimate the interaction among the relevant factors and evaluate the response. The vital production media components had been considered to optimize for streptokinase production. After performing the task of optimization of production media components the next aim was to optimize the culture-condition governing parameters. Statistically the model was tested for being significant and lack of fit was evaluated simultaneously for finding lack of its significance.

Statistical technique for screening and design experiments would be carried out to determine which of the several experimental variables and their interactions present more significant effects (Yazdani and Mukherjee, 1998). ANOVA was constructed for the

second order response surface model; the significance of each coefficient was determined by Student's *t*-test and *p*-value (Vellanki et al., 2009), to identify the corresponding significant coefficient (Balagurunathan et al., 2008). The statistical technique of optimization was also compared with neural network approach. Box and Behnken (Gilmour, 2006), (Bruns, 2006) studied how to select points from three-level factorial design for optimization using linear polynomial function (Theophilus and Ferreira, 2006). It might help in evaluation of first and second order coefficients of the statistical model (Boxa and Behnkena, 1960). ANOVA often used to compare variations due to process treatment with the randomly generated errors due to measurement of the responses (Vieira, 1989). The screening and experiment design should be carried out to determine the several experimental factors and their interaction that would be more promising to present the more significant effects (Bezerra et al., 2008). The industrial fermentation process directly depends upon factors like specific growth rate, biomass yield, final product quality etc, which influence the production cost (Monaghan et al., 1999), (Stabury et al., 1995). The cellulase activity of *Bacillus* strain was enhanced by optimizing the medium composition by statistical methods, Placket-Burman design and CCD (Deka et al., 2011). Structural characteristics of different combinations of bioprocess variables viz. time, temperature, and acid concentration were used for enhancing bioethanol production (Dhabhai et al., 2013). Central composite experimental design maximizes the acquired information and reduces the number of individual experiments required for any inference (Deepak et al., 2008). The fermentation factors optimization with respect to cultures condition was done in one-at-a-time strategy to enhance the D-amino acid oxidase yield (Gupta et al., 2012). To find the optimal fermentation media different monosaccharide and disaccharides were used as carbon sources such as glucose, fructose etc. RSM is a well known technique usually applied in optimization of medium constituents and other critical condition parameters responsible for production of enzymes (Xiong et al., 2004). RSM used by several researchers to find the requirement of media components malt and yeast extract for lipase production (Kumar and Gupta, 2009). Identification of important factors or key determinants with screening experiment is followed by application of complex response surface design to work for further optimization (Chauhan and Gupta, 2004). The statistical experiment design provides a global means to workers of different field particularly academia, industry and engineering for designing and analysing the experiments (Montogomery, 2001).

5.1.1. Statistical Bioprocess Modelling

It is clear from the literature that the parameters pertaining to culture conditions are not yet optimized for this recombinant strain. Consequently the screening of various production media constituents and optimization of culture condition parameters are to be assumed for maximum streptokinase/plasmin activity. After screening out vital media components for *E. coli*, on the basis of their statistical significance CCD would be the next aim to find out the most significant interaction of culture conditions with respect to maximum streptokinase activity.

Optimization methodology widely used in bioprocess often applying CCD. The Response Surface Methodology (RSM) was originally developed in perspective to fitness of the mathematical model. The CCD was initially configured by Box and Wilson (Gilmour, 2006), (Bruns et al., 2006). Optimization approach was developed for two factors (variables) and three factors with axial values α=1.41 and 1.68 respectively. RSM comprises of a cluster of mathematical and statistical techniques which used to depend on fitness of empirical model with the experimental model in contrast to experimental data design. Linear or square polynomial functions are usually applicable to utilize modelling setup using experimental condition upto its optimization goal (Theophilus and Ferreira, 2006). Two and three variable optimization has been shown diagramatically, **Figure 5.1 (a)** and **(b)**.

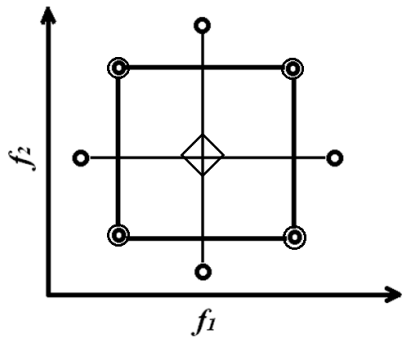


Figure 5.1 (a). Two variable optimization (α=1.41) using CCD, with points of factorial design (double circled dots), axial (single circled dots) and central point at intersection (f_1 and f_2 are the factors)

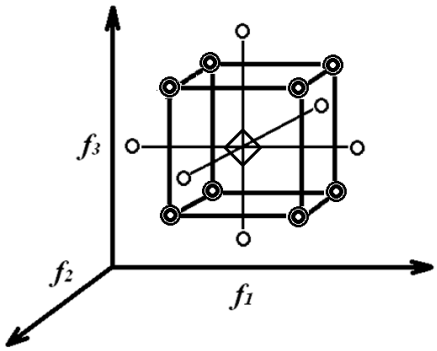


Figure 5.1 (b). Three variable optimization (α=1.68) using CCD, with points of factorial design (double circled dots), axial (single circled dots) and central point at intersection of axes (f_1, f_2 and f_3 are the factors)

The RSM application was carried out with few steps;

- Screening the vital variables
- Selection of suitable experimental design
- Simple codification of factors with levels
- Statistical analysis of experimental data
- Model fitness evaluation
- Solving Quadratic expression to find optimal parameters

In order to apply the linear function of RSM model it is essential that responses should be well fitted to the following expression;

$$y = \beta_0 \sum_{k=0}^{n} \beta_k x_k + \phi \quad (30)$$

Here, *n*-number of factor variables, *k* is the index, β_0 a constant, β_k coefficients of linear experimental parameters, x_k denotes the factors, ϕ the residual term associated to the considered experiments.

5.1.2. Culture conditions optimization

The indispensable task of culture condition optimization had been assumed for maximum productivity using Design Expert 8.0 software. Numerical based simulation of the process was performed in Matlab 2010Ra. In optimization process the contour plots were

obtained that interprets interactions among the pair of variables to achieve best level plasmin activity.

Inoculum Percentage Optimization

Additionally, inoculum concentration and age for best Plasmin Activity had been manually optimized. Several other parameters that were previously employed in similar work (Yazdani and Mukherjee, 2002) incorporated in our process as usual. The inoculum percentage and age optimization was done individually to achieve best level of recombinant streptokinase production, shown in **Figure 5.2 (a)** and **(b)**. Prior proceeding to the composite optimization of several parameters of cultivation it would be necessary to independently evaluate the individual parameters in terms of product formation.

Thus inoculum concentration of 2.75 g/ml and inoculum preparation of 6 hours of age was found to be most appropriate for our purpose to promote streptokinase production.

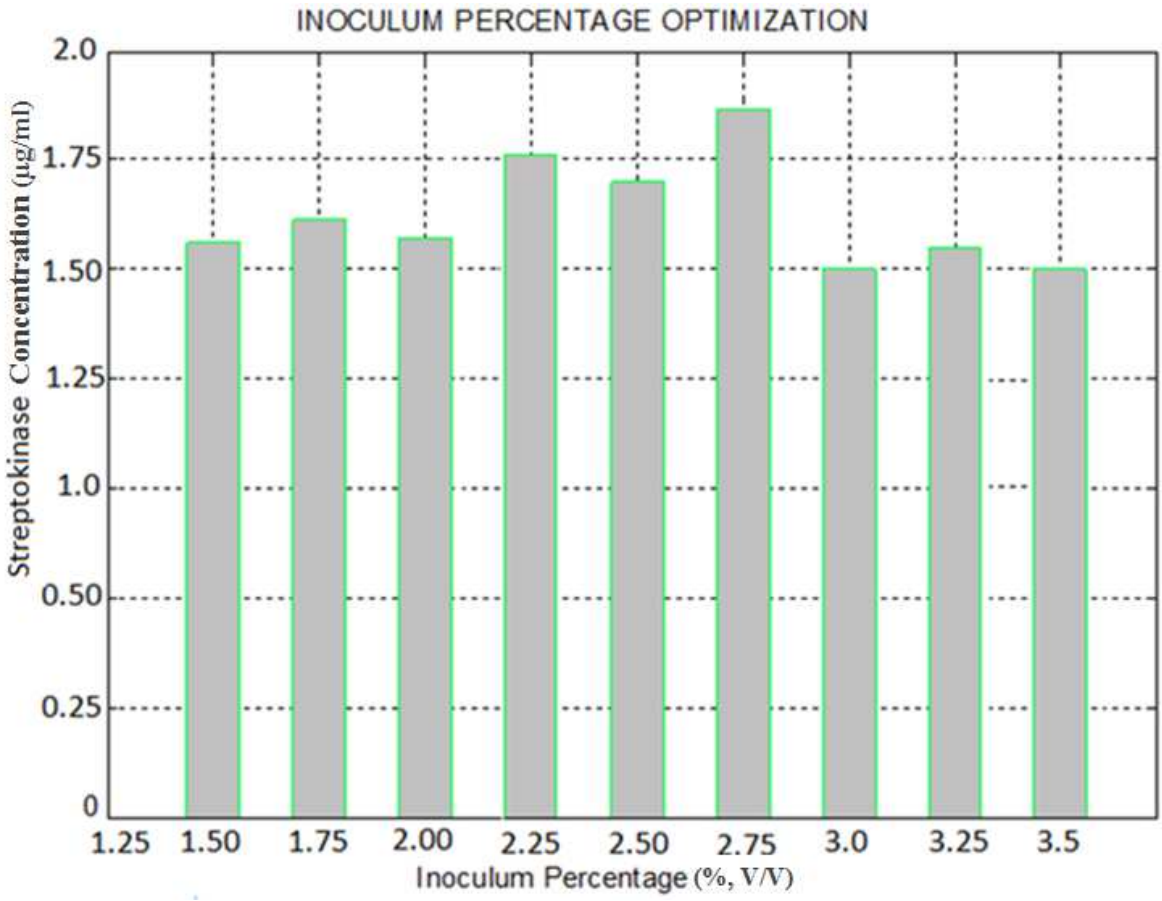


Figure 5.2(a). Inoculum percentage optimization

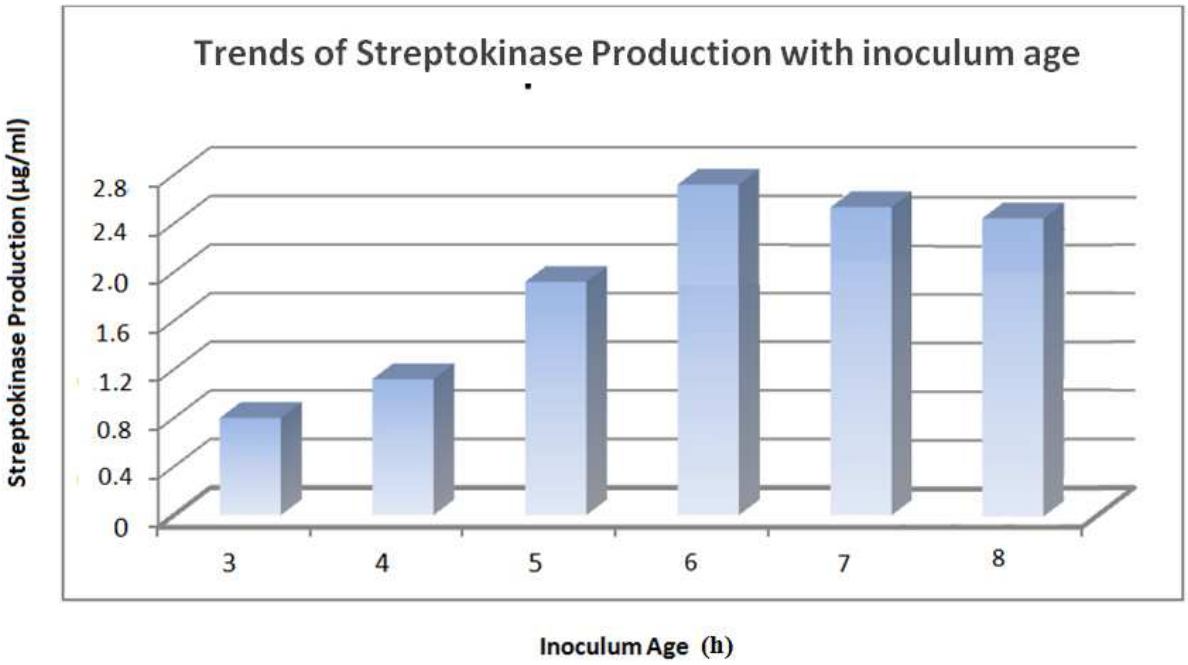


Figure 5.2(b). Inoculum age optimization

5.2. Screening of Production Media Constituents for Streptokinase production:

5.2.1. Plackett Burman Method (PB)

Numerous factors may influence the response of the system and it is generally not possible to evaluate and identify the significance of their role in the system individualy. It would be more convinient and feasible to select the factors that having prime role. Screening design was thus applied to findout which of the experimental factors, those having vital interaction that provide additional significant effects. The full factorial level design was more efficient and found to be applicable to bioprocess.

5.2.2. Screening Media Components using PB Methodology

5.2.2.1. Level of Factors

The PB design with level and range of media components shown in **Table 5.1**, while layout is given in **Table 5.2.**

Table 5.1. Levels and range of various factors used in Plackett-Burman

Factors	Medium Constituents	Levels Used (g/l) (-1)	(+1)
A	Glucose	3.0	7.0
B	Yeast Extract	3.0	7.0
C	NH_4Cl	0.7	1.30
D	Phosphate	5.0	11.0
E	NaCl	0.3	0.7
F	$MgSO_4$	0.3	0.7
G	$FeSO_4$	0.07	0.13

5.2.2.2. PB Layout

The seven media components were together put into consideration to conduct statistical screening, **Table 5.2**.

Table 5.2. Plackett-Burman design layout

Run	A:Glucose g/l	B:Yeast Extract g/l	C:NH_4Cl g/l	D:Phosphate g/l	E:NaCl g/l	F:$MgSO_4.7H_2O$ g/l	G:$FeSO_4$ g/l	SK Production µg/ml
1	3.0	3.0	1.3	11.0	0.3	0.3	0.13	1.624
2	3.0	3.0	0.7	11.0	0.7	0.7	0.07	1.596
3	3.0	7.0	0.7	5.0	0.7	0.3	0.13	1.554
4	7.0	7.0	1.3	11.0	0.7	0.7	0.13	1.918
5	7.0	3.0	0.7	5.0	0.3	0.7	0.13	1.386
6	7.0	3.0	1.3	5.0	0.7	0.3	0.07	1.344
7	7.0	7.0	0.7	11.0	0.3	0.3	0.07	1.484
8	3.0	7.0	1.3	5.0	0.3	0.7	0.07	1.862

5.2.2.3. Numerical Solution for Screening the Vital Components

Effect of each factor variable was computed from the design to find out their respective *t* and *p* values. The **Eq. 31** and **32** are representing expression to calculate effect and *t*-value subsequently.

Glucose,……..$FeSO_4$

X_1, X_2,…….....X_7

As X_i to X_n being encoded to show factors from A to G

Effect(X_i) = $\sum(y(+)_i - y(-)_i) / (N/2)$

(31)

N -> is the number of run (experiments)

E(X_1)=(6.1320-6.6360)/4 = -0.1260

E(X_2)=(6.8180-5.9500)/4= 0.2170

E(X_3)=(6.7480-6.0200)/4= 0.1820

E(X_4)=(6.6220-6.1460)/4= 0.1190

E(X_5)=(6.4120-6.3560)/4= 0.0140

E(X_6)=(6.7620-6.0060)/4=- 0.1890

E(X_7)=(6.4820-6.2860)/4= 0.0490

Degree of Freedom = Number of dummy variables found were two (here, E(X_5) and E(X_7))

$V_{eff} = \sum E_d^2/N_d = 1.3X10^{-3}$

S.E. = $\sqrt{V_{eff}}$ = 0.0360

t(Xi) = E(Xi)/S.E.

(32)

t(X_1) to t(X_7) = -3.50, 6.0278, 5.0556, 3.3056, 0.3889, 5.2500 and 1.3611 respectively

From the table, where p corresponds to t (http://graphpad.com/quickcalcs/PValue1.cfm) is given, thus p value was obtained likewise.

5.2.2.4. Statistical Analysis for Screening

Table 5.3. Statistical analysis from the results of Plackett-Burman design for screening of media components

Factors	Medium Components	Effect, $E(x_i)$	*t-value*	*p*-value
A	Glucose	-0.1260	-3.50	0.0228
B	Yeast Extract	0.2170	6.0278	0.0264
C	NH_4Cl	0.1820	5.0556	0.0806
D	Phosphate	0.1190	3.3056	0.0370
E	NaCl	0.0140	0.3889	0.7348
F	$MgSO_4$	0.1890	5.2500	0.0344
G	$FeSO_4$	0.0490	1.3611	0.3066

So from the *p*-values obtained in **Table 5.3**, it is evident that factors A, B, D and F are found significant. Since *p*-value is regarded to be the probability of rejecting any hypothesis/assumption thus the lesser value (usually <0.05) favours the acceptance of the proposed model.

5.3. Optimization of Production Media Components for Streptokinase production: Using Central Composite Design (CCD)

The level of selected factors are depicted in **Table 5.4**, while CCD layout is given in **Table 5.5**.

5.3.1. Level of factors for CCD

Table 5.4. The level and extent of factors taken for CCD

Factors	Medium Constituents	Levels (g/l)				
		-α	**-1**	**0**	**+1**	**+ α**
A	Glucose	1.0	3.0	5.0	7.0	9.0
B	Yeast Extract	1.0	3.0	5.0	7.0	9.0
C	Phosphate	2.0	5.0	8.0	11.0	14.0
D	$MgSO_4$	0.1	0.3	0.5	0.7	0.9

Note: α=2 ($2^{k/4}$, where k is the no. of factors, k=4)

5.3.2. Optimization of Media Components using CCD

Four production media components with their response as streptokinase prodution and model predicted values for CCD has been shown in the **Table 5.5**.

Table 5.5. CCD for media component optimization

Run	Factor 1	Factor 2	Factor 3	Factor 4	Response 1	Predicted Values	ANN Predicted
	A:Glucose	B:Yeast Extract	C:Phosphate	D:$MgSO_4$	SK Production	SK Production	SK Production
	g/l	g/l	g/l	g/l	µ g/ml	µ g/ml	µ g/ml
1	5	9	8	0.5	2.014	2.1809	2.0535
2	7	7	11	0.3	1.547	1.9017	1.7223
3	5	5	8	0.5	2.184	1.8787	1.9188
4	5	5	8	0.5	1.554	1.8884	1.5188
5	9	5	8	0.5	1.619	2.091	1.4806
6	3	3	11	0.7	2.34	1.9894	2.1991
7	5	5	14	0.5	2.019	1.9449	1.9572
8	3	7	5	0.3	2.461	2.1317	2.5027
9	3	3	5	0.3	1.756	2.5622	1.4964
10	5	5	8	0.9	2.428	2.0639	2.4288
11	7	7	11	0.7	1.544	2.0094	1.6262
12	5	5	2	0.5	2.381	1.9997	2.4717
13	7	3	11	0.3	1.507	2.1774	1.7060
14	7	3	5	0.3	1.506	1.8562	1.4855
15	5	5	8	0.5	2.403	1.7802	2.3188
16	5	5	8	0.5	1.708	1.7479	1.5188
17	5	5	8	0.5	1.573	2.1782	1.5188
18	3	3	5	0.7	1.612	1.8667	1.4066
19	5	1	8	0.5	1.576	2.0287	1.3889
20	5	5	8	0.5	2.291	1.6182	1.9188
21	7	7	5	0.3	2.333	2.2092	2.4378
22	3	7	11	0.3	2.299	2.0677	1.9606
23	3	7	5	0.7	2.31	2.0177	1.8639
24	7	3	5	0.7	1.548	2.0152	1.5152

25	3	7	11	0.7	1.561	1.5403	1.6467
26	1	5	8	0.5	2.279	2.5043	1.5866
27	5	5	8	0.1	2.363	2.1403	1.8560
28	7	7	5	0.7	1.407	1.5401	1.4162
29	3	3	11	0.3	1.256	1.4503	1.2990
30	7	3	11	0.7	1.659	1.8403	1.6552

5.3.3. ANOVA for production medium

Analysis of variance is used to evaluate the significance of the model, **Table 5.6**.

Table 5.6. ANOVA for production medium constituents

Source	Sum of Squares	df	Mean Square Value	F Prob > F	p-value	
Mode	0.38	14	0.027	2.59	0.0388	significant
A-Glucose	*9.963E-003*	*1*	*9.963E-003*	*0.95*	*0.3449*	
B-Yeast Extract	*0.086*	*1*	*0.086*	*8.22*	*0.0118*	
C-Phosphate	*0.029*	*1*	*0.029*	*2.73*	*0.1190*	
D-MgSO4	*0.015*	*1*	*0.015*	*1.42*	*0.2523*	
AB	*0.10*	*1*	*0.10*	*9.61*	*0.0073*	
AC	*0.060*	*1*	*0.060*	*5.72*	*0.0303*	
AD	*1.351E-003*	*1*	*1.351E-003*	*0.13*	*0.7245*	
BC	*0.053*	*1*	*0.053*	*5.04*	*0.0403*	
BD	*2.328E-003*	*1*	*2.328E-003*	*0.22*	*0.6441*	
CD	*9.361E-003*	*1*	*9.361E-003*	*0.89*	*0.3594*	
A^2	*1.534E-004*	*1*	*1.534E-004*	*0.015*	*0.9053*	
B^2	*2.680E-003*	*1*	*2.680E-003*	*0.26*	*0.6203*	
C^2	*9.525E-003*	*1*	*9.525E-003*	*0.91*	*0.3554*	
D^2	*3.713E-003*	*1*	*3.713E-003*	*0.35*	*0.5604*	
Residual	*0.16*	*15*	*0.010*			
Lack of Fit	*0.12*	*10*	*0.012*	*1.45*	*0.3568*	*not significant*
Pure Error	*0.040*	*5*	*8.049E-003*			
Cor Total	*0.54*	*29*				

5.3.4. Inferences from ANOVA

The Model F-value of 2.59 implies the model is significant. There is only a 3.88% chance that a "Model F-Value" this large could occur due to noise. R-Squared is 0.7201. Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B, AB, AC and BC are significant model terms. The "Lack of Fit F-value" of 1.45 implies the Lack of Fit is not significant relative to the pure error. Non-significant lack of fit value is good as it emphasizes is the model to unfit.

5.3.5. Numerical Analysis of Optimization Solving Quadratic Equation

Model Quadratic Equation obtained, given by **Eq. 33**;

(Abbreviations used here, G=> Glucose, Y=> Yeast Extract, P=> Phosphates, and M => Magnesium Sulphate)

$$\begin{aligned} SK\ (y) = {} & + 1.22772 \\ & + 0.033318G + 0.062130Y - 0.034054P + 0.45307M \\ & -0.019828GY + 0.010198GP - 0.022969GM \\ & + 9.57292X10^{-3}YP + 0.030156YM - 0.040312PM \\ & +5.91146X10^{-4}G^2 - 2.47135X10^{-3}Y^2 \\ & -2.07060X10^{-3}P^2 - 0.29089M^2 \end{aligned} \quad (33)$$

A simplified form:

$$\begin{aligned} y = {} & b0 + b1*A + b2*A + b3*C + b4*D \\ & +b11*A^2 + b22*B^2 + b33*C^2 + b44*D^2 \\ & +b12*AB + b23*BC + b34*CD + b41*DA \\ & + b13*AC + b24*BD \end{aligned}$$

Taking partial derivatives of y with respect to each of the four parameters and considering it in matrix form;

$$2*b11*x1 + b12*x2 + b13*x3 + b41*x4 = -b1$$

$$b12*x1 + 2*\ b22*x2 + b23*x3 + b24*x4 = -b2$$

$$b13*x1 + b23*x2 + 2*b33*x3 + b34*x4 = -b3$$

$$b41*x1 + b24*x2 + b34*x3 + 2*b44*x4 = -b4$$

In general, the above set of equation could be solved using matrix method;

A = [2*b11 b12 b13 b14
2*b22 b23 b24
2*b33 b34
2*b44]

B=[- b1 - b2 - b3 - b4];

Now the four simultaneous equations solved for, $\frac{\partial y}{\partial P} = 0$, $\frac{\partial y}{\partial A} = 0$, $\frac{\partial y}{\partial T} = 0$, $\frac{\partial y}{\partial I} = 0$

Since, y = inv (A) * B;

(34)

So using **Eq.34**, the values of matrix coefficients variables can be obtained as,

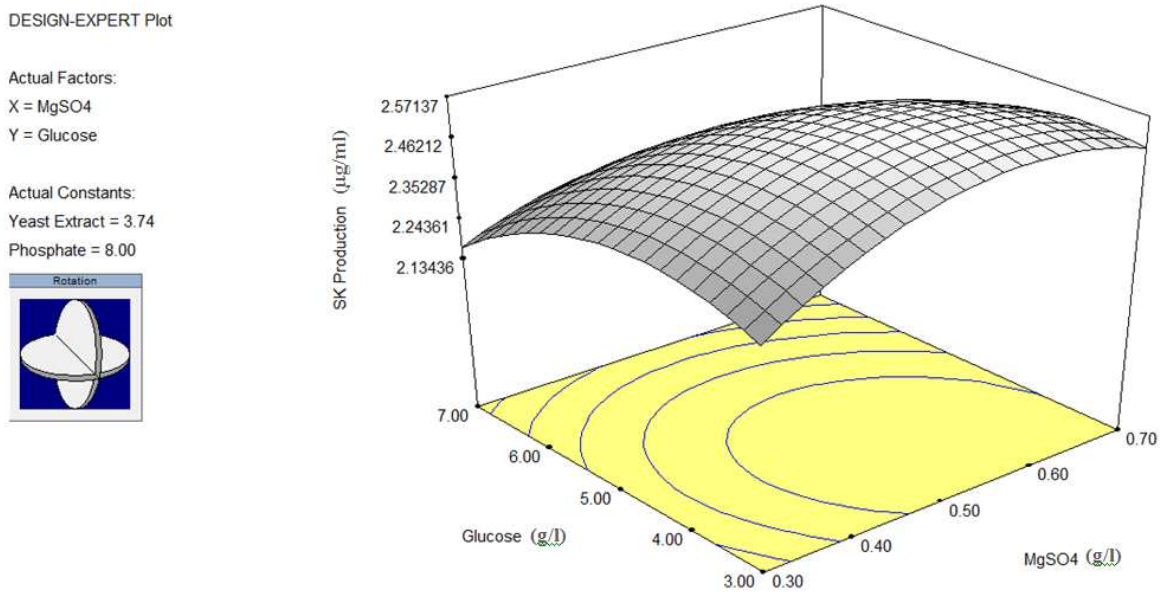
G=4.3695; Y=2.1899; P=1.8210; M=0.5936;

The production magnitude obtained as, y=1.8720

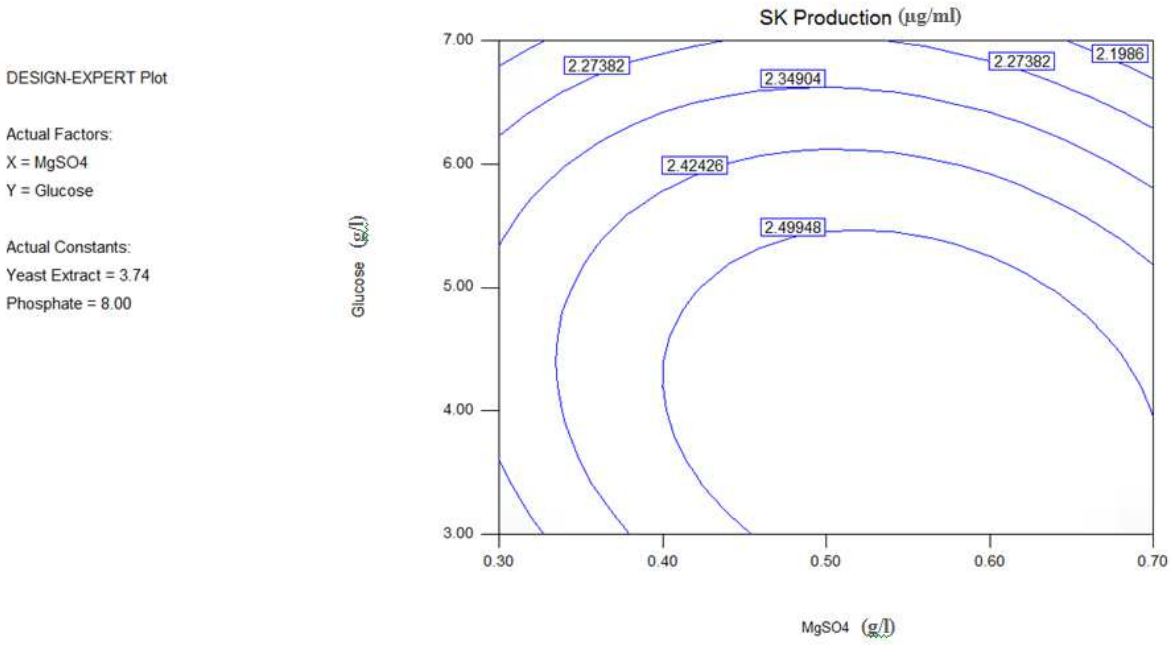
The set of quadratic equations obtained was done in Matlab Editor. The term of correlation r^2 in the case was found to be close to 1.0. The performance efficiency of CCD with respect to error minimization was found to be appreciable. The correlation r^2 computed to be 0.7201 for the statistical model which could also be evaluated using other simulation means. The statistical and ANN predicted output values of **Table 5** has been further analysed in **section 6.3.1, Chapter 6**.

5.3.6. 3-D and Contour Plots

Figures 5.3-5.8, shown with response surface 3-D and contour plots and to represent interaction of different pair of production media components and their response in terms of streptokinase production.

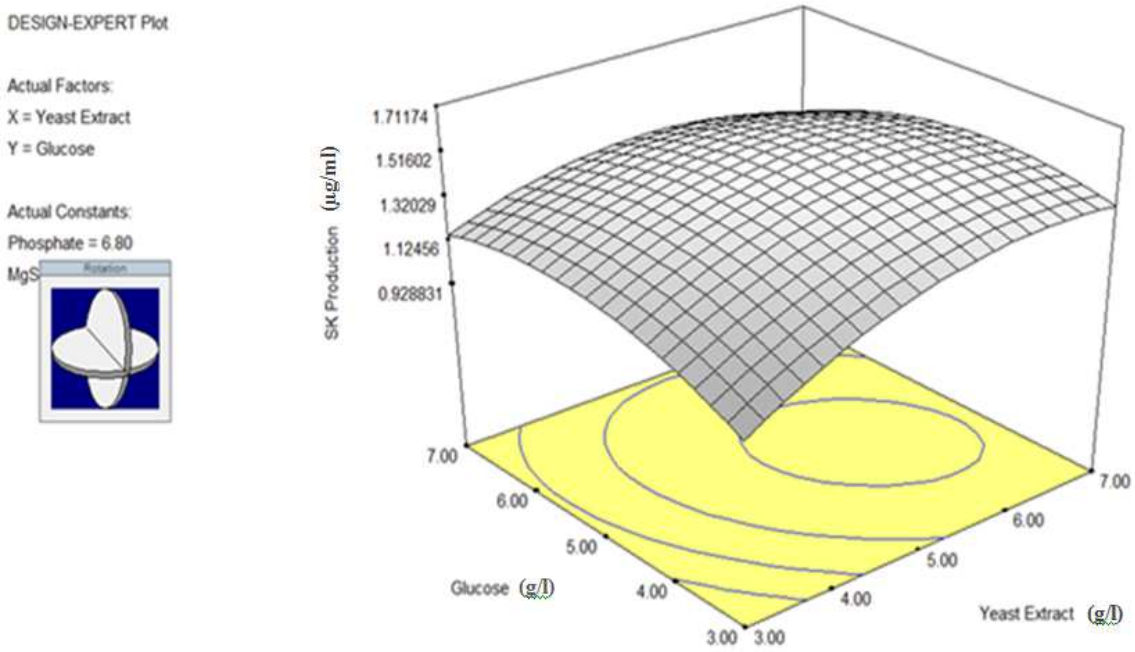


5.3(a)

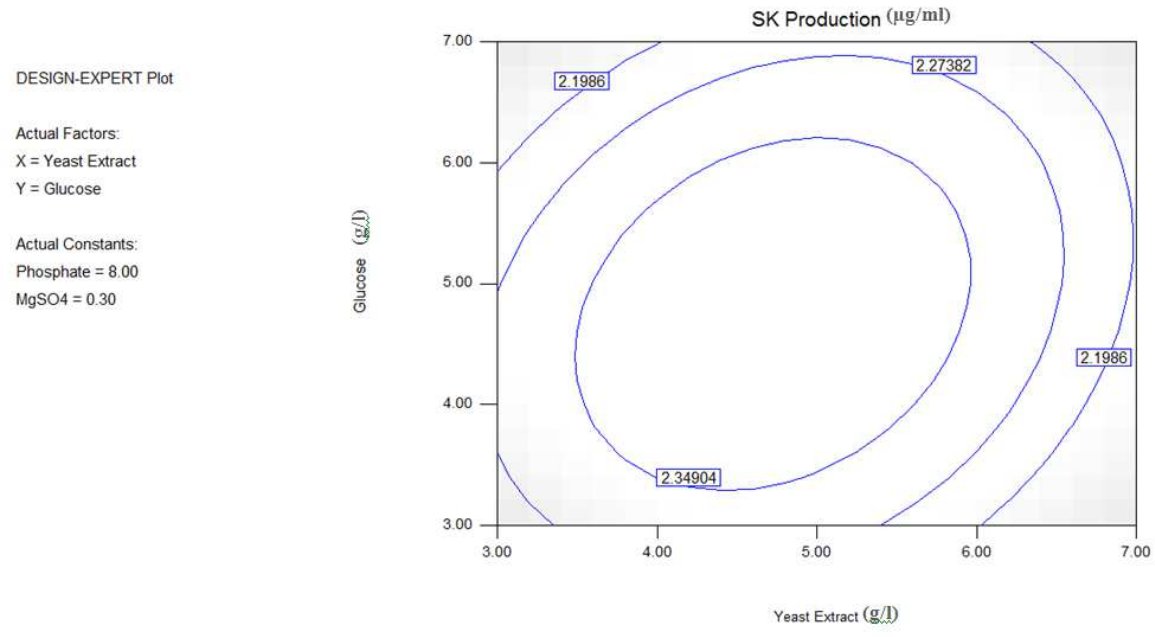


5.3(b)

Figure 5.3(a) and (b). Response surface and contour curve of streptokinase production (μg/ml) as a function of glucose and $MgSO_4$ in medium with yeast extract and phosphate, 3.74 and 8.00 g/l respectively

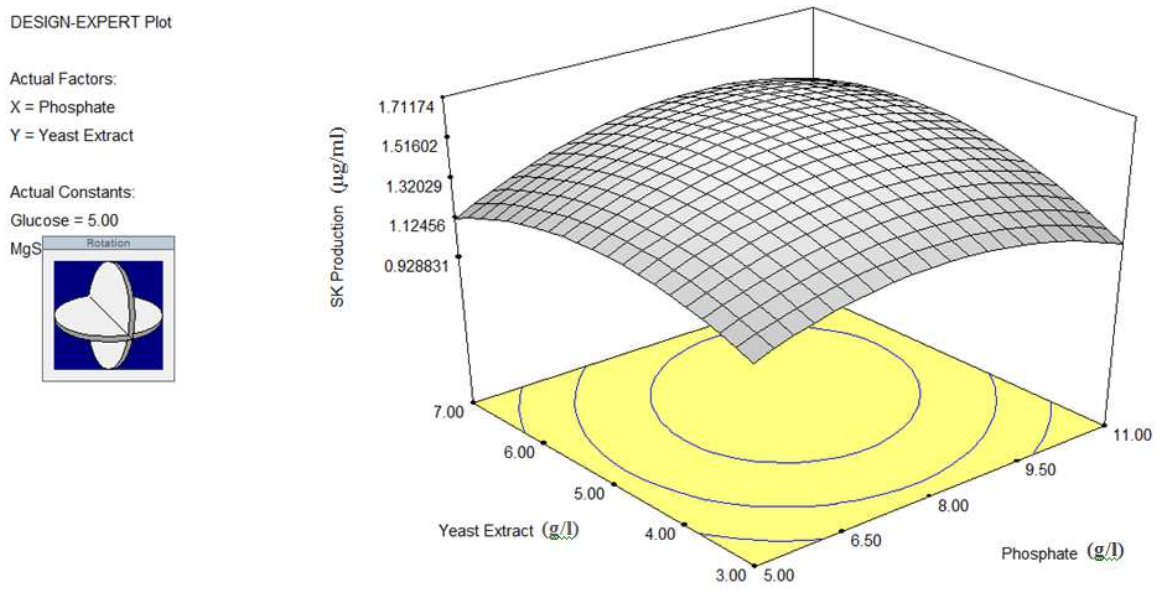


5.4(a).

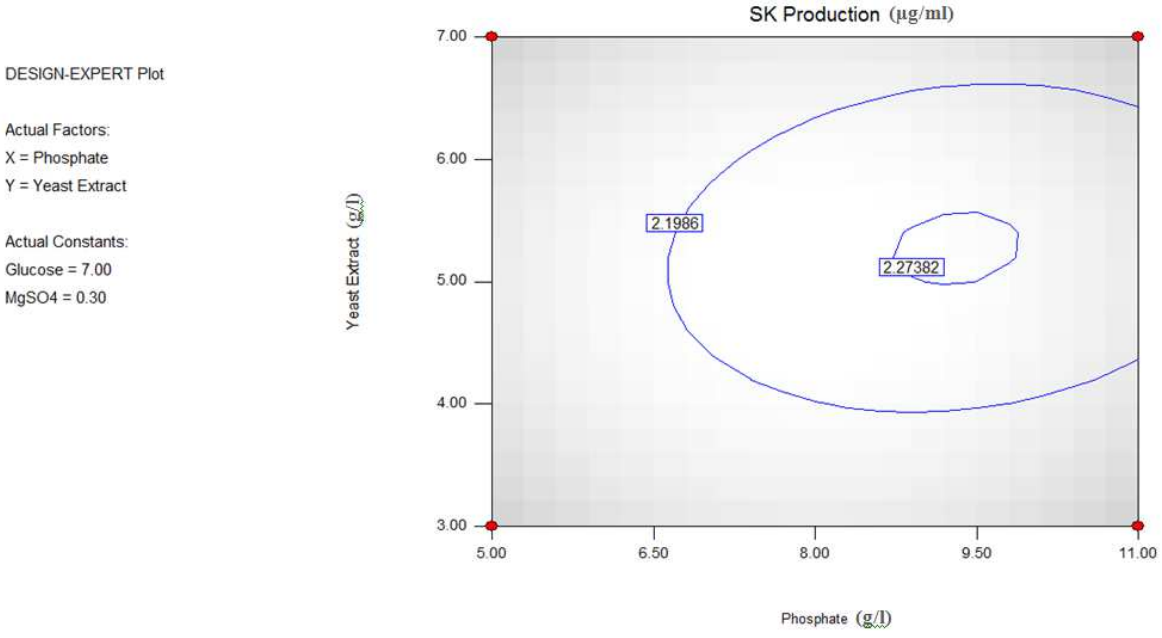


5.4(b)

Figure 5.4(a) and (b). Response surface and contour plot for streptokinase production (μg/ml) as a function of glucose and yeast extract in medium with varying conc. of phosphate and $MgSO_4$ in g/l

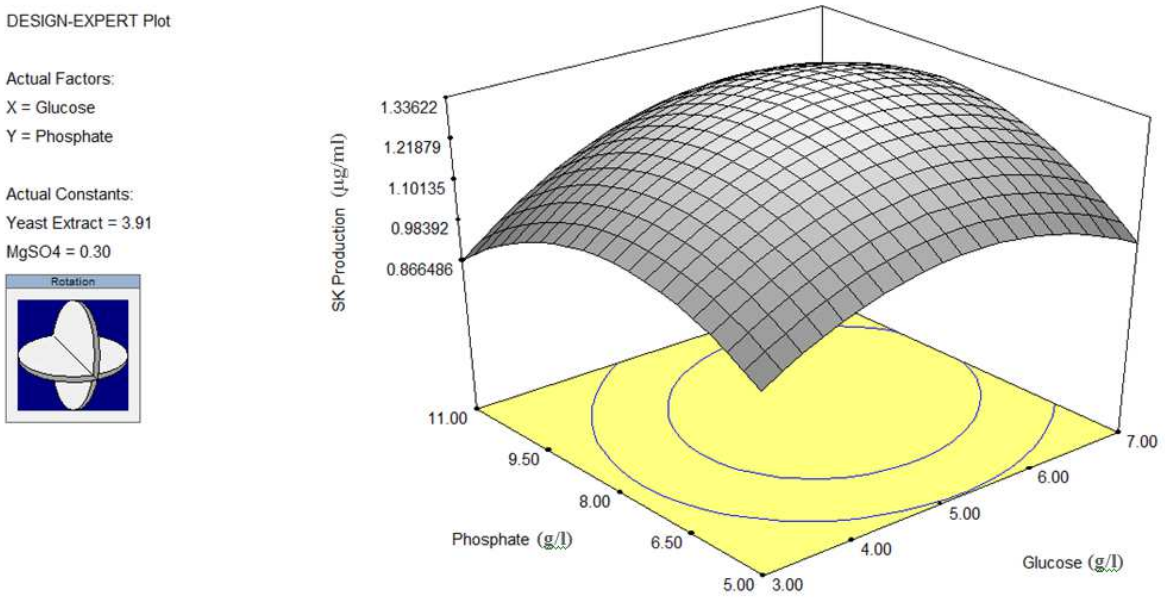


5.5(a).

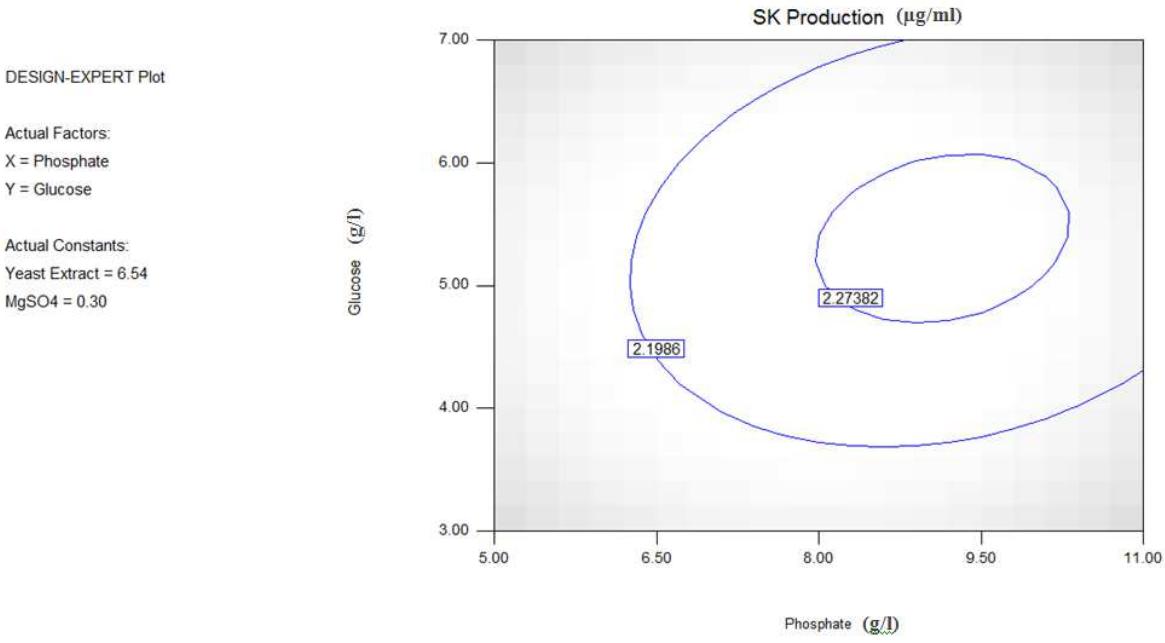


5.5(b)

Figure 5.5(a) and (b). Response surface and contour plot for streptokinase production (μg/ml) as a function of yeast extract and phosphate in medium with varying concentration of glucose and $MgSO_4$ in g/l

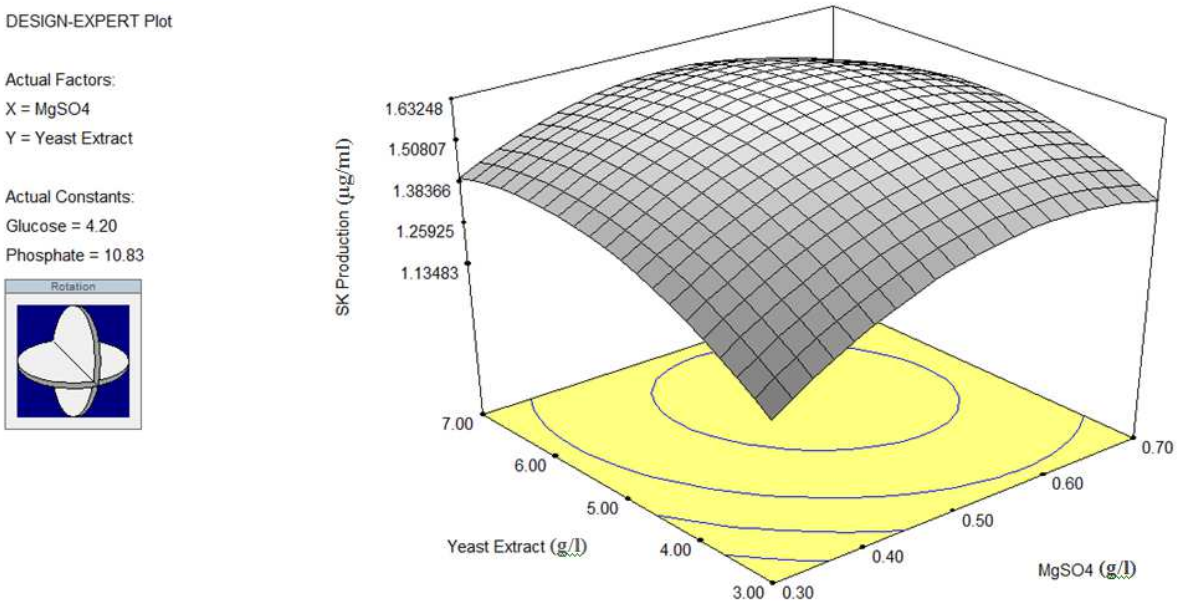


5.6(a)

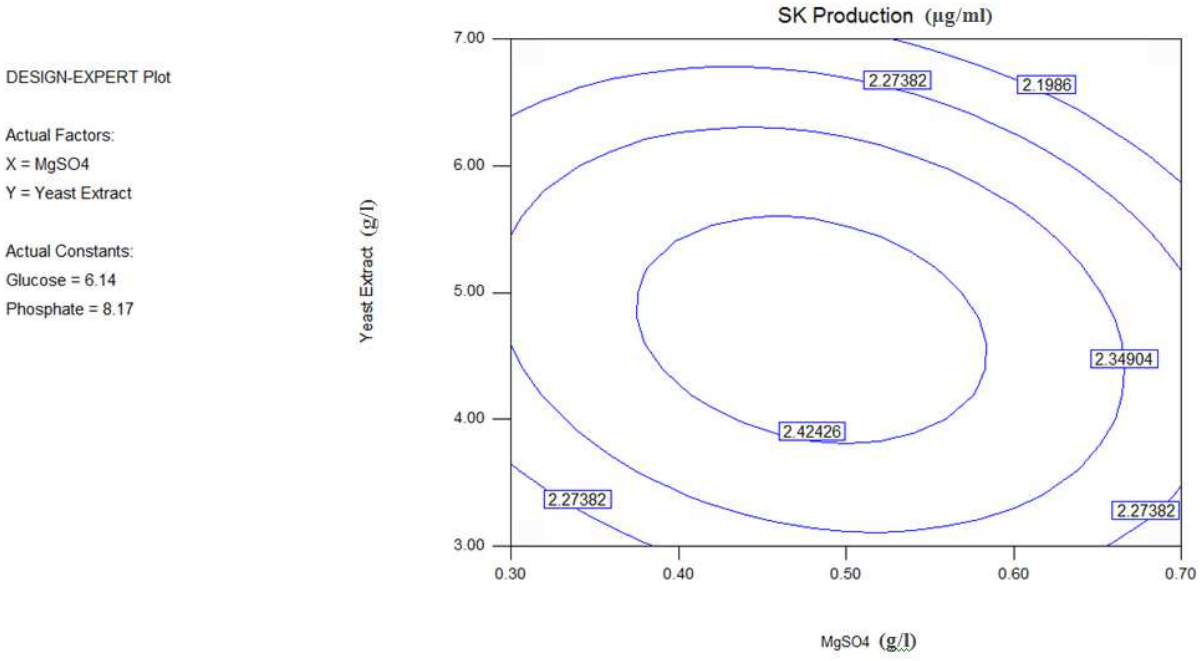


5.6(b)

Figure 5.6(a) and (b). Response surface and contour plot for streptokinase production (μg/ml) as a function of phosphate and glucose in the medium with yeast extract and $MgSO_4$

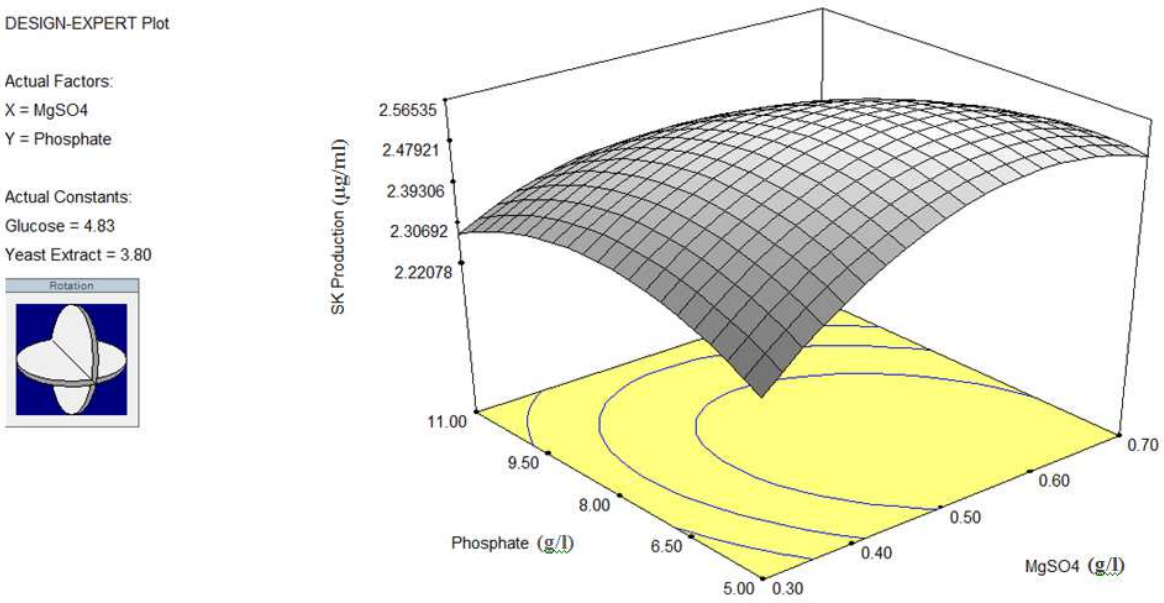


5.7(a)

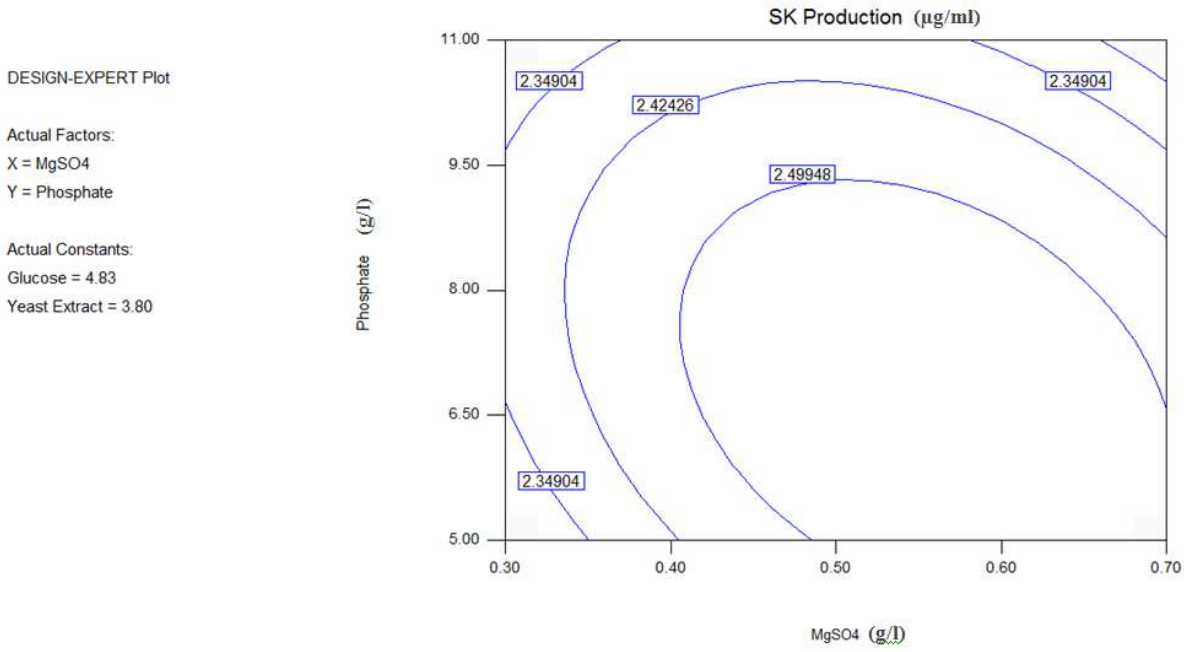


5.7(b)

Figure 5.7(a) and (b). 3-D response surface and 2-D contour plot showing streptokinase production (μg/ml) as a function of yeast extract and $MgSO_4$ in the medium with glucose and phosphate



5.8(a).



5.8(b).

Figure 5.8(a) and (b). 3-D response surface and 2-D contour plot for streptokinase production (μg/ml) as a function of phosphate and glucose in the medium with yeast extract and $MgSO_4$ in concentration 4.83 and 3.80 (g/l) respectively

The optimal level of each medium component for maximum response in terms of streptokinase production was evaluated via response surfaces. It would be helpful in

determining the best magnitude combination of factors to achieve optimal production. So, six pairs of 3-D and contour plots were obtained using Design Expert software, **Figures. 3-8**, a slightly increased concentration of MgSO4 and yeast extract indicated to favours the production of recombinant enzyme streptokinase.

5.4. Optimization of Culture Conditions for Streptokinase production

5.4.1. Level of Factors for CCD

The level of cultivation factors were taken, **Table 5.7** according to CCD method.

Table 5.7. The range of various factors taken for CCD

Factors	Medium Constituents	Levels (g/l)				
		-α	**-1**	**0**	**+1**	**+ α**
A	pH	6.00	6.50	7.00	7.50	8.00
B	Agitation (rpm)	100	150	200	250	300
C	Temperature (°C)	31.0	34.0	37.0	40.0	43.0
D	Inoculum Conc. (%,V/V)	1.25	2.00	2.75	3.50	4.25

Note: α=2 ($2^{k/4}$, where k is the no. of factors, k=4)

5.4.2. Optimization of Culture Conditions Using CCD

The layout of CCD for four culture condition parameters with response and predicted model values are given below in **Table 5.8.**

Table 5.8. CCD for culture conditions optimization

Run	Factor 1	Factor 2	Factor 3	Factor 4	Response 1	Predicted Values	ANN Predicted
	A: pH	**B:Agitation**	**C: Temp.**	**D: Inoculum conc.**	**SK Production**	**SK Production**	**SK Production**
		rpm	(°C)	**(%, V/V)**	**µ g/ml**	**µ g/ml**	**µ g/ml**
1	7.5	250	40	2	2.779	2.0265	2.1663
2	6.5	250	40	3.5	2.072	1.7596	2.2401
3	6.5	250	34	3.5	2.142	2.8586	2.072
4	7.0	200	37	2.75	2.723	2.4430	2.5209

5	6.5	150	40	2.0	1.953	1.4842	1.8403
6	7.0	200	43	2.75	2.002	1.9386	2.4046
7	7.0	200	31	2.75	1.946	2.5211	1.7756
8	7.5	150	34	3.5	2.912	2.0267	2.7029
9	7.5	250	34	2.0	1.778	2.1702	1.6931
10	7.5	150	40	2.0	2.303	2.1151	2.3497
11	6.5	250	40	2.0	2.079	2.0346	2.0773
12	7.0	200	37	2.75	1.435	1.8307	1.7209
13	6.5	150	34	2.0	2.163	2.1023	2.1417
14	7.5	150	40	3.5	2.086	1.9684	2.0065
15	7.0	200	37	4.25	1.365	2.1714	1.7791
16	6.5	150	40	3.5	2.156	1.8888	2.0675
17	7.5	250	34	3.5	2.107	2.2560	2.163
18	7.0	300	37	2.75	1.239	1.7065	2.107
19	6.0	200	37	2.75	2.053	1.8185	2.303
20	7.5	250	40	3.5	1.907	2.5710	2.0201
21	7.5	150	34	2.0	1.953	2.0168	1.7783
22	6.5	150	34	3.5	1.778	1.5327	2.1607
23	7.0	200	37	2.75	2.023	2.1813	1.9209
24	7.0	100	37	2.75	2.121	2.1782	1.9461
25	8.0	200	37	2.75	1.351	1.6241	1.3185
26	6.5	250	34	2.0	2.24	1.9441	2.5123
27	7.0	200	37	2.75	1.323	1.5141	1.9209
28	7.0	200	37	2.75	2.128	1.9241	1.9289
29	7.0	200	37	2.75	2.107	1.9141	1.9209
30	7.0	200	37	1.25	2.072	1.9141	1.9495

5.4.3. ANOVA for culture conditions

The result of CCD analysis is presented in ANOVA framework, **Table 5.9**.

Table 5.9. ANOVA for culture condition factors

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	3.24	14	0.23	2.63	0.0368	significant
A-pH	*0.45*	*1*	*0.45*	*5.15*	*0.0385*	
B-Agitation	*0.85*	*1*	*0.85*	*9.65*	*0.0072*	
C-Temperature	*0.35*	*1*	*0.35*	*4.00*	*0.0640*	
D-Inoculum conc (%)	*5.104E-005*	*1*	*5.104E-005*	*5.801E-004*	*0.9811*	
AB0	*022*	*1*	*0.022*	*0.25*	*0.6233*	
AC	*6.202E-003*	*1*	*6.202E-003*	*0.070*	*0.7942*	
AD	*0.045*	*1*	*0.045*	*0.51*	*0.4863*	
BC	*0.042*	*1*	*0.042*	*0.48*	*0.5006*	
BD	*0.94*	*1*	*0.94*	*10.64*	*0.0052*	
CD	*0.22*	*1*	*0.22*	*2.56*	*0.1307*	
A^2	*7.724E-003*	*1*	*7.724E-003*	*0.088*	*0.7711*	
B^2	*0.14*	*1*	*0.14*	*1.53*	*0.2345*	
C^2	*0.033*	*1*	*0.033*	*0.38*	*0.5476*	
D^2	*0.13*	*1*	*0.13*	*1.42*	*0.2516*	
Residual	1.32	15	0.088			
Lack of Fit	*0.93*	*10*	*0.093*	*1.17*	*0.4569*	*(not-Significant)*
Pure Error	*0.39*	*5*	*0.079*			
Cor Total	4.56	29				

5.4.4. Inference of ANOVA

The Model F-value of 2.63 implies the model was significant. There was only a 3.68% chance that a "Model F-Value" this large could occur due to noise. R-Squared was 0.7102. Values of "Prob > F" less than 0.0500 that used to indicate model terms were significant. In this case A, B, BD are significant model terms. The "Lack of Fit F-value" of 1.17 implies the Lack of Fit was not significant relative to the pure error. Non-significant lack of fit assumed to be good since it was desirable for the model to fit.

5.4.5. Numerical Analysis of Optimization using Quadratic Expression

Model Quadratic Equation obtained as **Eq. 35**, given by,

(Abbreviations used here, P=> pH, A=> Agitation, T=> Temperature, I=> Inoculum Conc. (%))

$$\text{SK Production}(y) = 6.16592 - 0.81958P + (8.06583X10^{-3}A) + 0.12486T - 2.30589I - (1.48750\,X10^{-3}PA) - 0.013125PT + 0.14117PI + (3.41250\,X10^{-4}AT) - (6.45167\,X10^{-3}AI) + 0.052694TI + 0.067125P^2 + (2.80625\,X10^{-5}A^2) - (3.87153\,X10^{-3}T^2) + 0.12006I^2 \quad (35)$$

$\delta y/\delta P = -0.81958 + 2X0.067125P - 1.48750X10^{-3}A - 0.013125T + 0.14117I = 0$

$\delta y/\delta A = 8.06583 - 1.48750X10^{-3}P + 2X2.80625X10^{-5}A + 3.41250TX10^{-4} - 6.45167X10^{-3}I = 0$

$\delta y/\delta T = 0.12486 - 0.013125P + 3.41250X10^{-4}A - 2(-3.87153X10^{-}3)T + 0.52694I = 0$

$\delta y/\delta I = -2.30589 + 0.14117P - 6.45167X10^{-3}A + 0.052694T + 2X0.12006I = 0$

In general, the above set of equation could be solved in matrix method form. Now the four simultaneous equations solved for, $\frac{\partial y}{\partial P} = 0$, $\frac{\partial y}{\partial A} = 0$, $\frac{\partial y}{\partial T} = 0$, $\frac{\partial y}{\partial I} = 0$

Since, y = inv (A) * B, using the values of matrix coefficients, parameter variables can be obtained as,

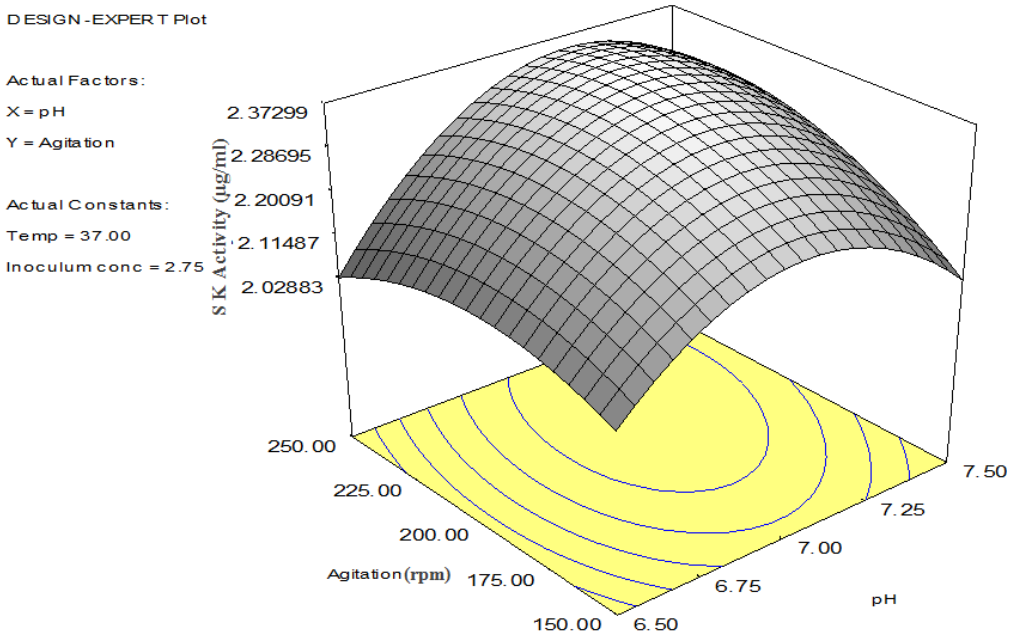
P=8.1351 (pH value), A=283.0167 (rpm), T=37.5653 (°C) and I=3.5930 (%, V/V).

Finally the production magnitude computed as, y=2.8665 (µg/ml)

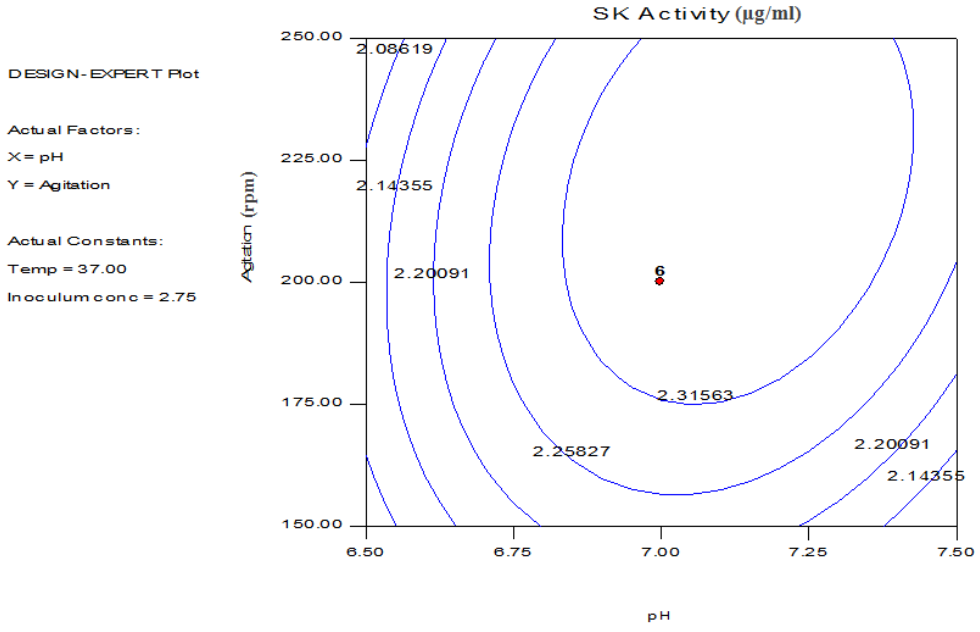
The numerical solution of the quadratic equation was done in Matlab Editor. The correlation r^2 in the later case was found to be close to 1.0. The performance efficiency of CCD in respect to error minimization was found to be appreciable. Plot shows the Actual *vs* Predicted obtained from statistical based methods. The correlation r^2 for the plot, computed to be 0.7102 for the statistical model. Further, the statistical and ANN predicted output obtained in **Table 8,** has been compared and analysed in **section 6.3.2, Chapter 6**.

5.4.6. Optimization 3-D and Contour Plots

Figures 5.9-5.14, Showing 3-D surface and contour plots to represent interaction of different pair of culture condition parameters and their response in terms in streptokinase production

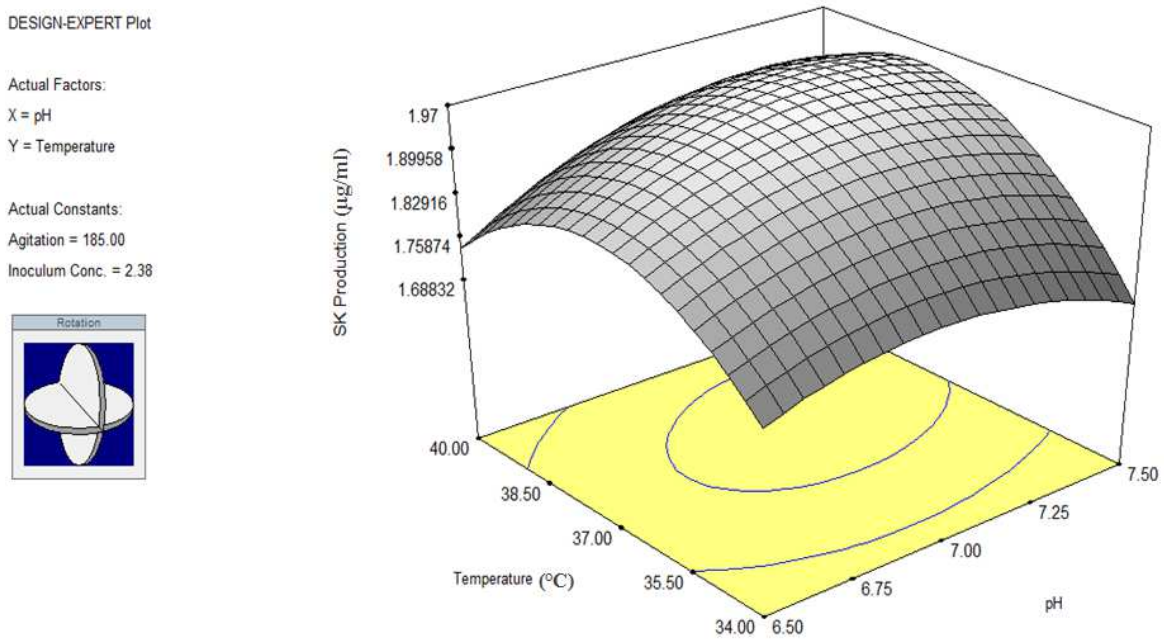


5.9 (a)

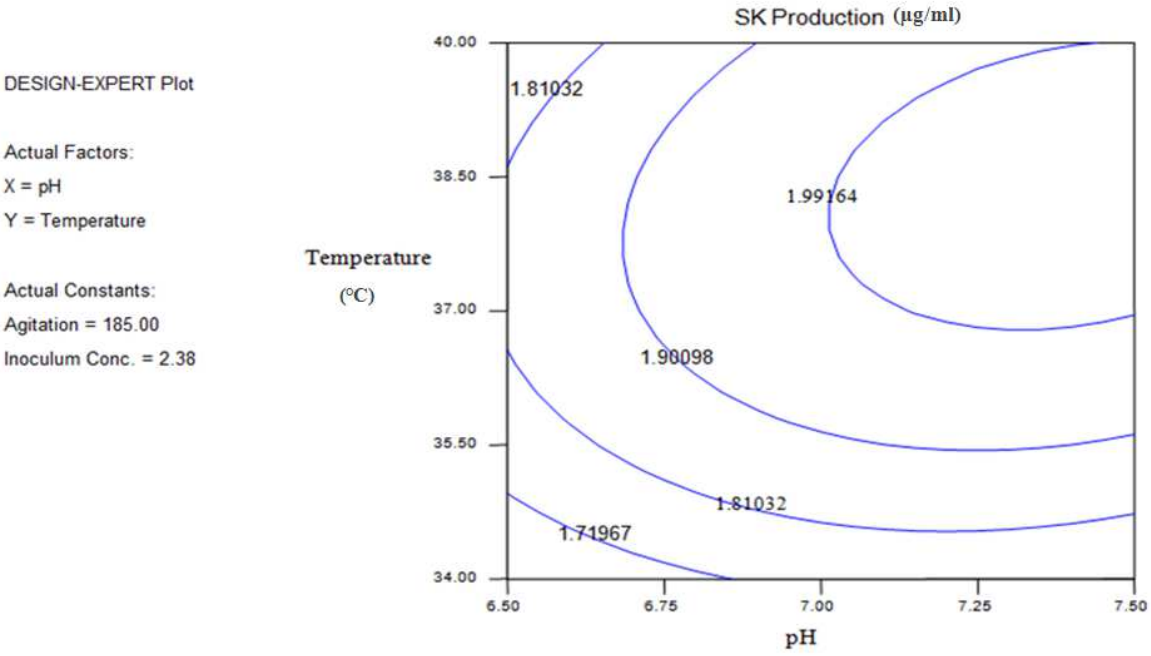


5.9(b)

Figure 5.9(a) and (b). 3-D surface and contour plot showing showing the interaction and effect of agitation and pH on streptokinase production (μg/ml). In plot, actual constraints are temperature 37°C, inoculum concentration 2.75% (V/V)

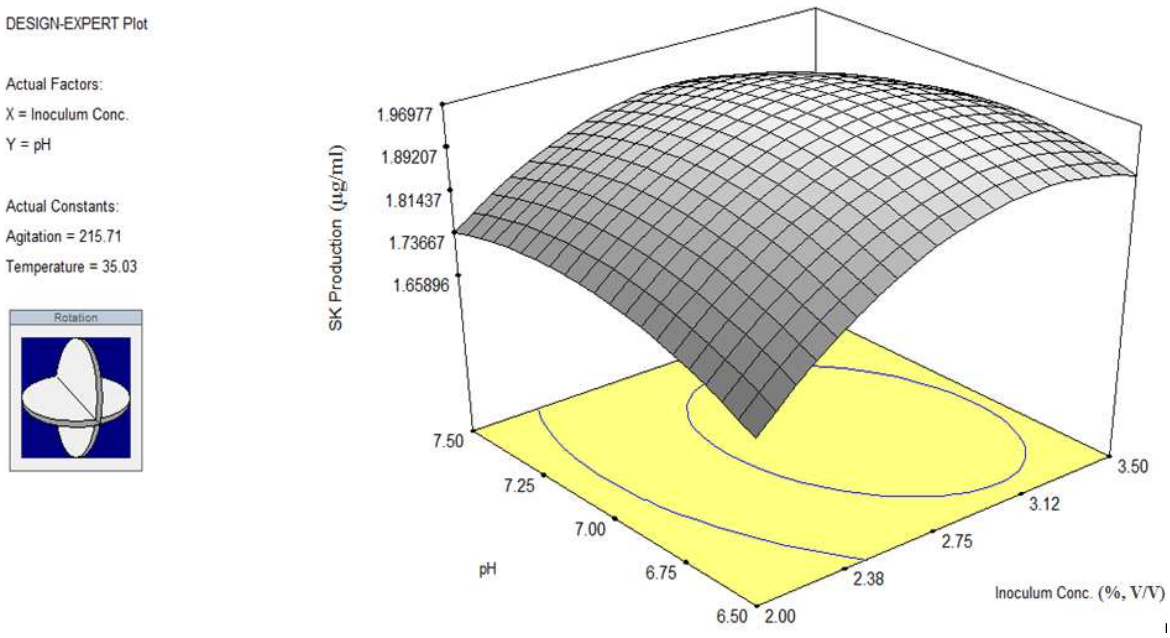


5.10(a)

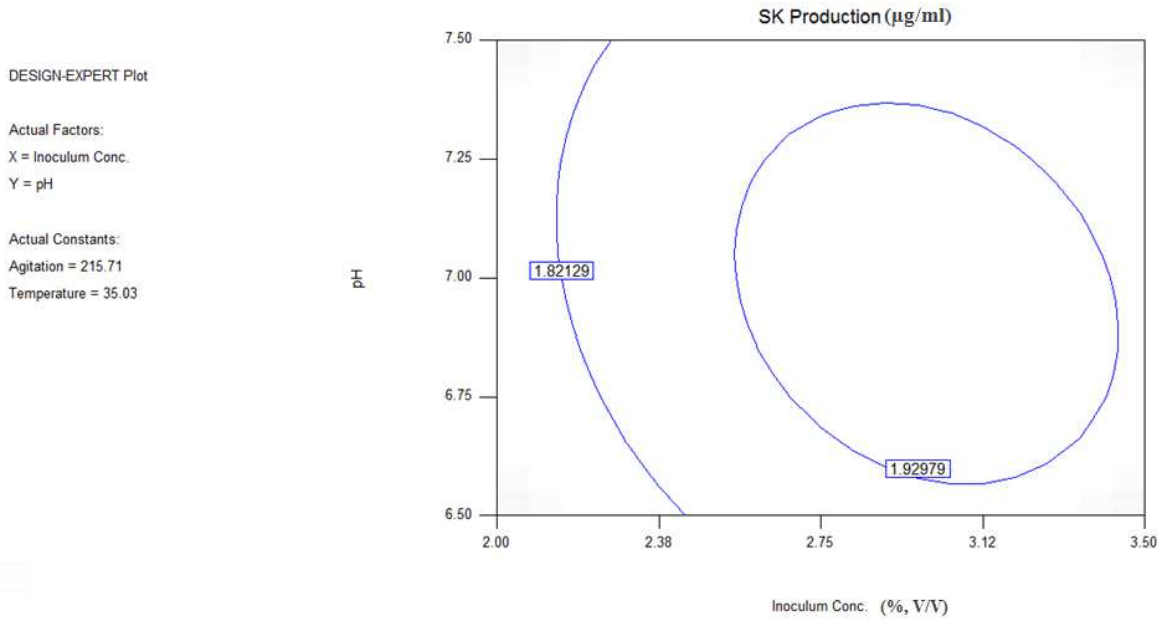


5.10(b)

Figure 5.10 (a) and (b). 3-D surface plot and contour diagram depicting the interaction effect of temperature and pH on streptokinase production, given actual constraints are agitation and inoculum concentration as 185rpm and 2.38g/l respectively

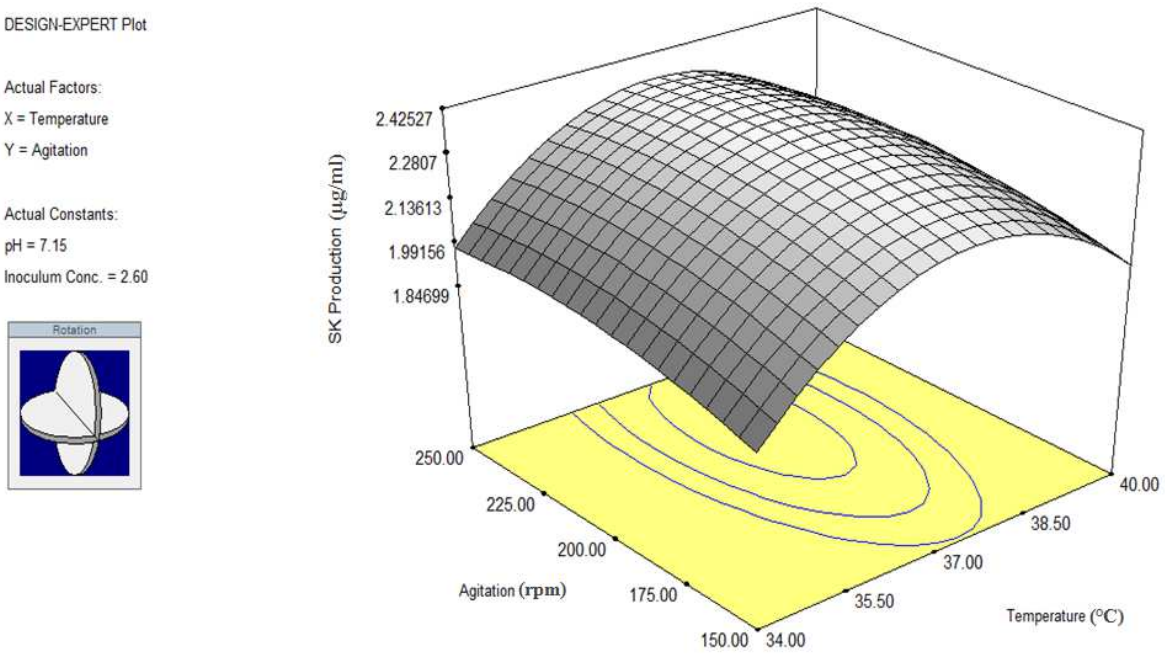


5.11 (a)

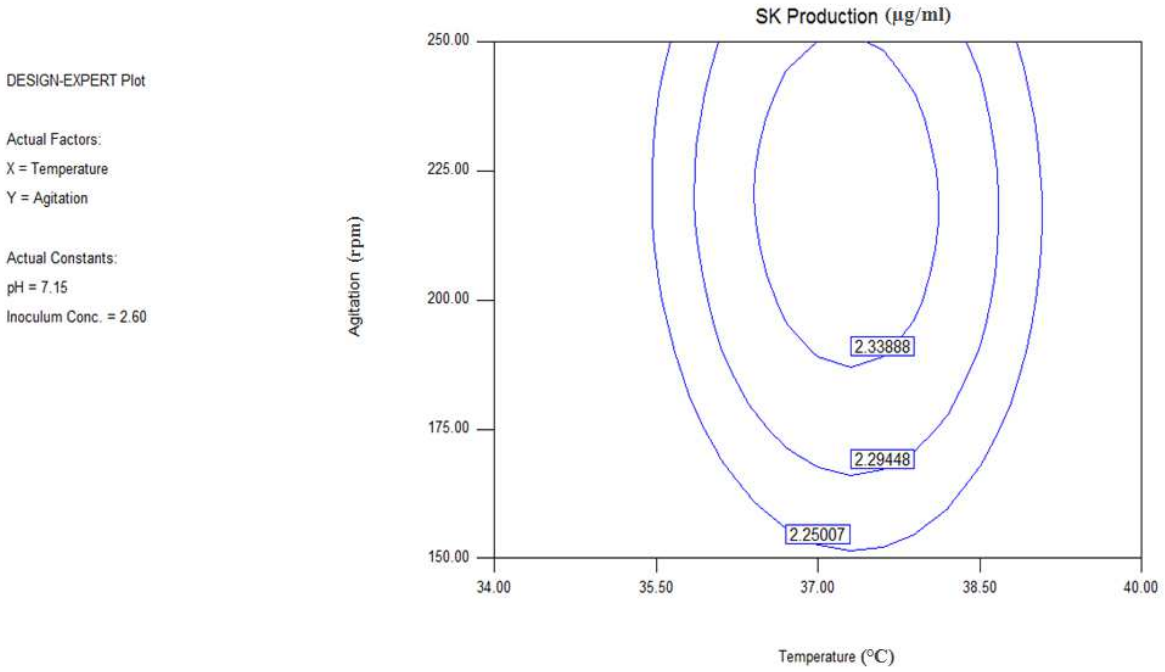


5.11(b)

Figure 5.11 (a) and (b). 3-D surface plot and contours showing the effect of inoculum concentration and pH on streptokinase production at temperature of 36.03^{0}C and agitation 215.71 rpm

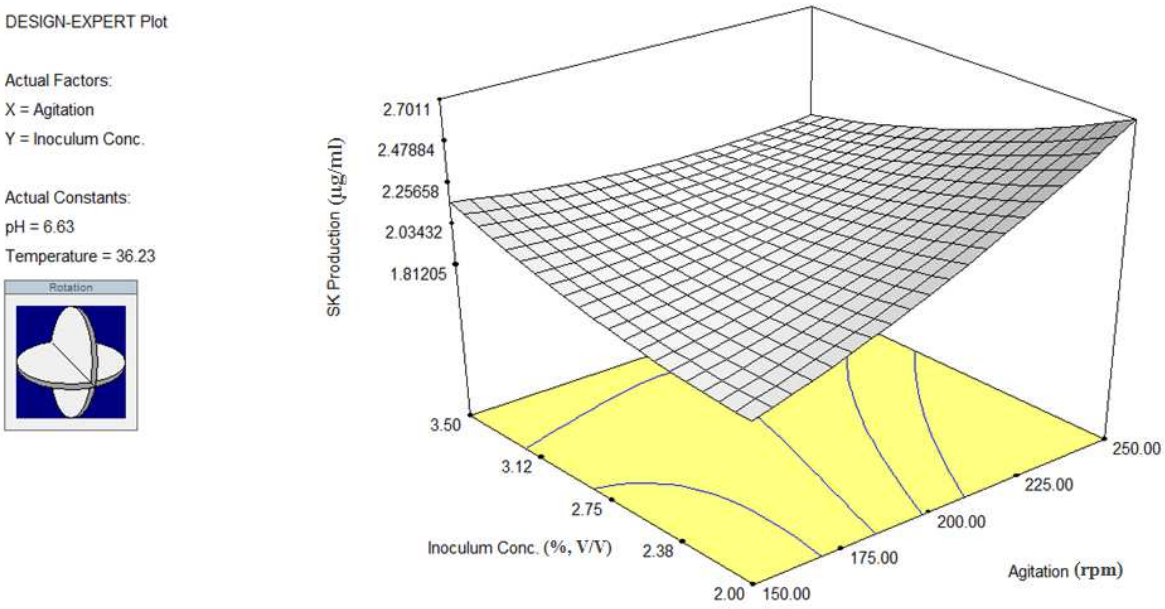


5.12(a)

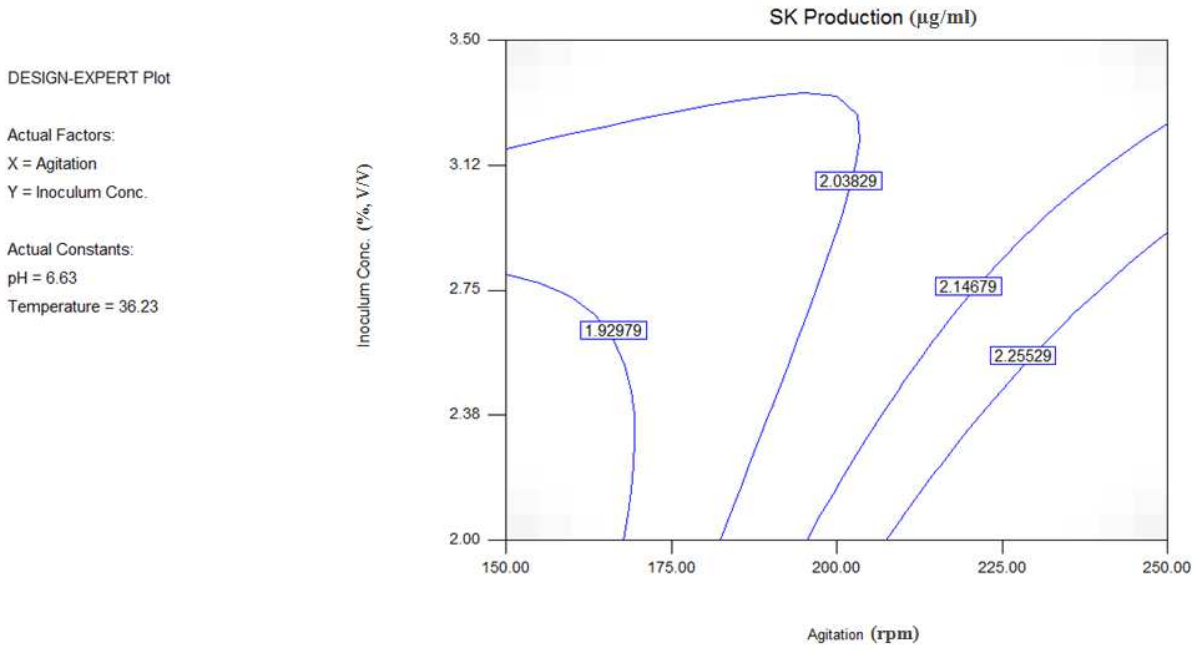


5.12(b)

Figure 5.12 (a) and (b). Surface and contour plots showing interaction effect of agitation vs temperature to evaluate streptokinase production with constrains pH and Inoculcum conc., 7.15 and 2.60g/l respectively

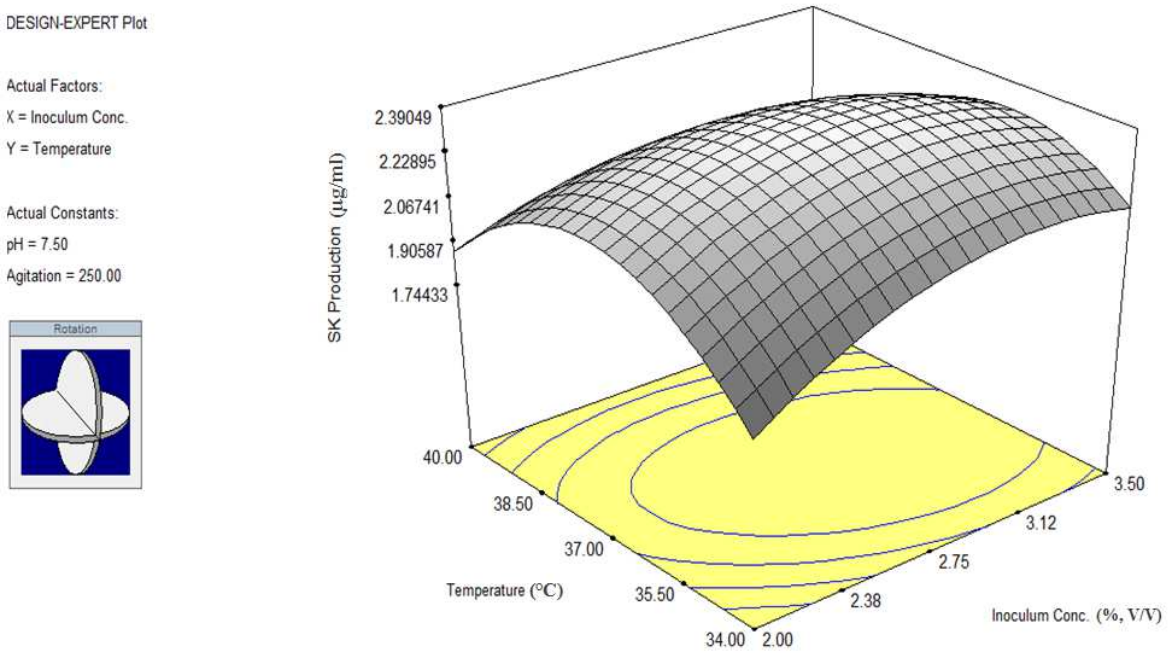


5.13(a)

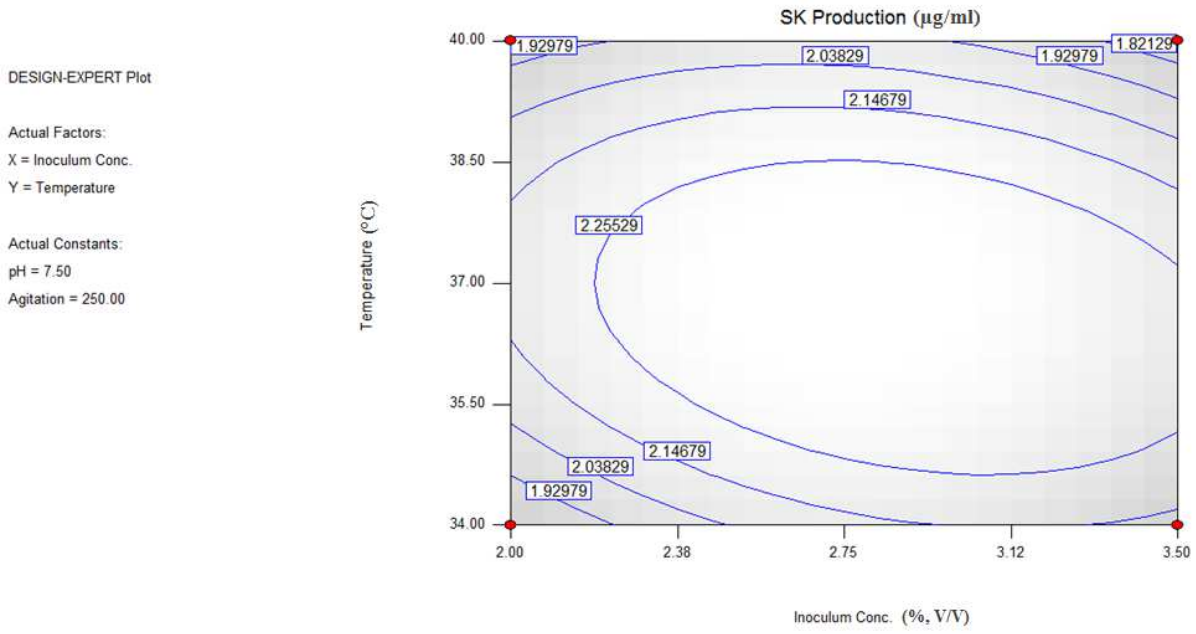


5.13(b)

Figure 5.13 (a) and (b). Response surface 3-D and contour plot between inoculum conc. vs agitation is shown in 3-D surface and contour plot to evaluate their effect in terms of streptokinase production, the other constrains maintained are pH and temperature as 6.63 and 36.23 respectively.



5.14(a)



5.14(b)

Figure 5.14 (a) and (b) 3-D surface and contour plot signifying the interaction and effect of temperature and inoculum concentration on streptokinase production (μg/ml), other contraints of culture conditions are pH 7.50 and agitation 250^{0}C respectively

The 3-D response surface and contour curves were employed to determine the optimal magnitude of each factor for maximum response in terms of streptokinase production. So, six pairs of 3-D and contour plots **Figures. 5.9-5.14**, were obtained on analysing the parameter values using design expert. Increasing agitation, pH and inoculum concentration indicated to favour streptokinase production.

5.5. Discussion

In the dynamical system, it was observed that sustenance of plasmid bearing cells depends upon unstructured conditions provided during the cultivation. The production media components and culture condition factors are the two basic prime requirements which were used to support the cell growth and product formation. The four production media components were screened out by Plackett Burman methodology; these factors viz., Glucose, Yeast Extract, Phosphate and $MgSO_4$ had shown their relevance in recombinant enzyme production having significant role in growth. It would be due to the reason that glucose and yeast extract are prime carbon and nitrogen sources for proper cell growth, phosphate and $MgSO_4$ used to support ATP formation, protein synthesis at intracellular level. Subsequently, the next aim had been to optimize the culture governing parameter which was done subsequently with success. The validation of the model outcome from CCD was done afterward and about 40% enhancement in product was obtained.

The task of optimization of culture conditions was achieved to a good extent; increase in yield had been acquired employing this technique. Statistically both, production media and culture condition models were found to be significant and lack of fit was not significant which was desirable. Enhancing the level of $MgSO_4$ and yeast extract showed to favor the enzyme production. In the surface topology observation, shifting of pH to some higher extent and increased level of inoculum concentration and agitation had been proved to be beneficial for optimal streptokinase production as evident from CCD analysis.

Therefore, CCD analysis was performed to evaluate the interaction and relevance of variation of factors in terms of response that was evident from ANOVA analysis. Due to presence of natural noise in system the observation with very adequate correlation could be sometimes harder to achieve. Also, it might be due to the reason that accuracy in experimental response measurement was not always adequate. Efficient modeling may ensure an enhanced

production of streptokinase on large scale. The optimization task usually marks an increase of one or two fold in magnitude of response while in our case it was up to a certain extent, still it would be quite worthy since streptokinase is a known highly valuable product.

CHAPTER-6

NEURAL NETWORK APPLICATIONS TO BIOPROCESS MODEL

6.1. An Overview

In this chapter, a multilayer feed forward artificial neural networks (ANNs) having a multi-layer perceptron is employed. It is more commonly used than any other neural network types. The building of ecological model that utilizes a lot of methods, ranging from numerical, mathematical and statistical methods to techniques originating from artificial intelligence, like neural networks. Artificial intelligence knowledge is required to be implemented to bioprocess field in recent times to achieve a high degree of accuracy in prediction and process optimization. In our work, Matlab 2010Ra *nn* toolbox was used together with neural network algorithm based programming in Matlab editor and simulation of the model.

A feed-forward neural network was modified to include dynamic characteristics, form the basis of estimator models. Training, testing and validation of data were repeatedly done. A multilayer perceptron is usually designed to process the weighted input (Lek and Guegan, 1999). Sigmoidal function was used as an activation function using which the output was obtained. The given inputs to the model were taken as medium components, culture condition parameters and also as population size of plasmid bearing and lacking cells and metabolite concentration etc, while the output was obtained in terms of streptokinase production. The idea was to iterate through the training set and adjust the weights to minimize the gradient of the error. ANNs, serve as a powerful tool for non-linear modelling and process control (Glassey et al., 1994). The ANN model showed a better correlation with the experimental values than the regression model (Haider et al., 2008), it suggests the superior capability of neural network modelling for the dynamic behaviour of the system. Artificial neural network proved to be a useful model building tool for improving bioprocess operability dealing for industrial level fermentation systems (Massimo et al., 1991).

6.1.1. Development of modelling strategies for bioprocess using ANN

Neural network application are widely used in prediction for model parameters and supervision of bioprocess. Unstructured model factors are found to apparently influence the existence of active cells in the bioreactor environment. The simulation of the process has been done using Matlab 7.2, the patterns obtained noticeably depicts the role of few relevant governing parameter including, metabolite concentration, dilution rate, plasmid copy number etc, to find out about segregational instability and competitive dynamics in cell population. The most potent technique has been further applied using collectively Neural Network and Logical Models together where the later one serves as tool of optimizing the process parameters. This is often reffered as hybrid approach. The production of many highly profitable compounds like recombinant enzymes is being done using fermentation technology

Several modelling strategies in bioprocess are found to play a vital role in the production of high value recombinant products. Modelling of plasmid stability is presently dealt in mechanistic with neural network applications. Our effort is to configure a composite hybrid model which represents the overall dynamics in a well defined algorithm that depicts the behaviour of the microbial population in the bioreactor environment.

ANNs function serves as 'black box' representation of the process, their performance depends solely on their training, i.e., selection of data and training methodology (Patnaik, 2003). Also over-learning or extra number of neuron in a particular layer may decrease the efficiency rather than increasing it (Zhang et al., 1994). It would be sometimes difficult to make ANN applicable to real conditions, so by combining mathematical model of bioprocess to neural network helpful in developing a new approach inferred as Hybrid Neural Network (HNN). In HNN, mathematical models work together with neural network framework having interaction among their operational routines (Schubert et al., 1994), (Van Can et al., 1997). ANN can be modified efficiently in order to model dynamic systems that can be utilized in process applications (Willis et al., 1992). The fed batch fermentation by a recombinant *E. coli* strain was studied in reference to carryout on-line optimization of PID control using ANN (Patnaik, 1999). The bioreactor controller would be adjusted on the basis of data of previous interval. System of neural networks was taken to update continually during successive time intervals (Patnaik, 2002).

The continuous fermentation are more susceptible than batch fermentation. The incorporation of prior knowledge serve as an practical approach which improves the

prediction ability of considered neural network. Use of ANN in modelling of yeast biomass and yield of β glucan was done successfully adjusting nurons in the hidden layer (Desai et al., 2005). The combination of neural network together with the emperical expression providea an efficient representation of estimated parameter values (García-Ochoa and Castro, 2001). It is quite noticeable that extraction of useful information from the complex fuzzy data can be well handled using neural network technique (Hoskins and Himmelblau, 1988), (Chitra, 1993) & (Qian and Tessier, 1995). ANN technique has been utilized for several purposes in biotechnology particularly for bioprocess in process control, prediction of factor variables, optimization and modelling of bioprocess(Ungar et al., 1995), (Thibault et al., 1990), (Di Massimo et al., 1992), (Thomposon and Kramer, 1994), (Kurtanjek, 1994) & (Tholudur, 1996). The application of prior knowledge has been examined as the means of developing and enhancing ANN prediction and estimation in case of noisy data of the bioprocess (Thomposon and Kramer, 1994). The feed forward neural network has been more commonly used in different bioprocess disciplines (Yet-Poleo et al., 1996), while two layered neural network is also sufficient for such applications in fermentation technology (Reuss, 1995). On experimental analysis and validation the output values of hybrid neural network are found to be very close to the numerical values obtained using other means by the non-linear regression analysis of process data (Costa et al., 1982).

6.1.2. Modelling Strategies and Use of Neural Network

There are variety of knowledge based approaches among those ANN has been extensively used in modelling the various biological phenomenona, it is because of its superior capability to configure non-linear na multivariate complicated bioprocess systems (Linko and Zhu, 1991), (Karim and Rivera, 1992), (Tomida, 1999) & (Shimizu et al., 1997). The ANN is often regarded as a data driven model, framed by learning procedure that utilizes back-propagation algorithm and input-output data maping (Rumelhart et al., 1986). The intelligence based has been further applied using collectively Neural Network and Logical Models together where the later one serves as tool of optimizing the process parameters. Supervision of bioprocess can be conducted on the basis of known developed algorithms.

In earlier attempts several deterministic mathematical models describing microbial population were evaluated but due to complex system behavior, the

mathematical and transient response approach were not found to be sufficient (Molleta et al., 1984), (Pavlou, 1987), (Mata-Alvarez, 1987) & (Lessard and Beck, 1993). Since several interacting sub-components within a cell, together have their influential role in enzyme production kinetics. Utilizing the theme we must have to assume the intracellular level dynamics and its sustainability for a certain span. So intra-cellular intelligence with various subcomponents involve in the production dynamics often used to contribute their indispensable role in product formation (Patnaik, 2009). So, it would be interesting to ensure the applicability of stochastic process in running the working state of system at any instance of time. Such entire network can be expressed in the form of a unitary system which is working for the production of a particular product that depends on the retention of its productive state. An artificial neural network (ANN) was developed to estimate the growth of microorganisms during a fermentation process. The two input variables were supervised on-line from a series of batch cultivations, which was used to train the ANN to estimate biomass (Hur and Chung, 2006).

The formation of secretory microbial products can be affected by cell-cycle position of individual cells and thus on the degree of heterogeneity in the system. The intelligence based computing methods including regulatory control are of much help in artificail farming associated to bioprocess dynamics (Wang et al., 2010). A nonlinear statistical approach meant for data analysis and an auto-associative neural network, was taken to fault diagnosis in the optimal production process of a recombinant yeast (Shimizu, 1997). Since production media and growth conditions have profound effect over plasmid stability as well as specific growth rates of two kind of cells, thus improvement of such systems via optimization using neural network is desirable. The pioneer applications of the knowledge-based approach to bioprocess operations were already reported. The expert systems and use of genetic algorithms have been extensively considered and used in control of various bioprocess designs (Shioya et al., 1999). The modelling of the dynamics of chemical processes with non-linear characteristics and multiple input-output can be well suited to such framework (Hotta, 1975).

6.1.3. Neural Network based Bioprocess Supervision

Artificial intelligence knowledge is required to be implemented to bioprocess field in recent times to achieve a high degree of accuracy in prediction and process optimization. ANNs are used for engineering purposes to process information and control

automated systems. By adjusting the weights of an artificial neuron we can obtain the desired output for specific inputs. Intelligent descriptions of microbial processes in biorectors have some decisive avantages over mechanistic model under realistic conditions and has wide range of flexibility (Patnaik, 2009). A generaal topology of the proposed method using ANN and logical model input is depicted below, **Figure 6.1.**

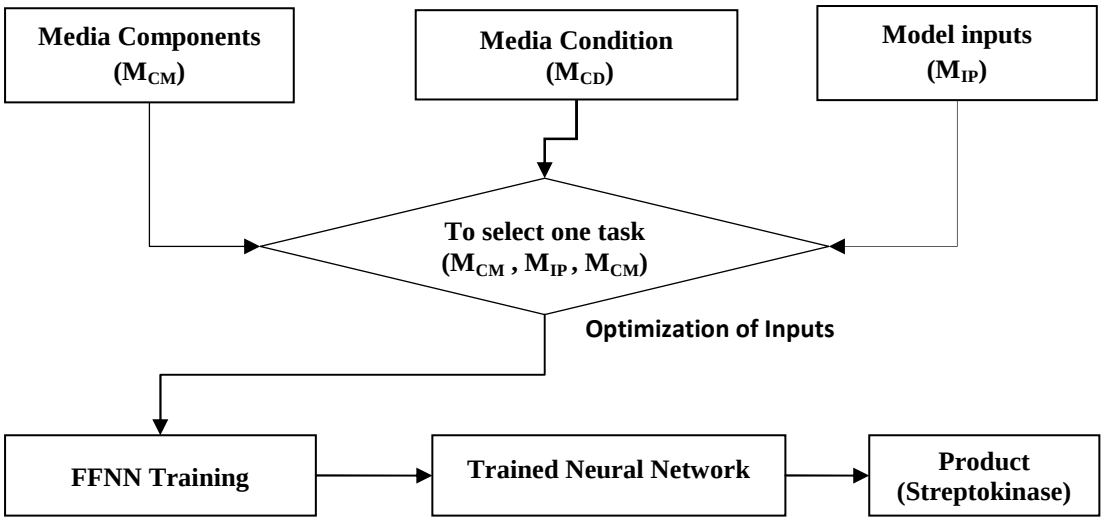


Figure 6.1. Flow diagram of proposed methodology

Recently artificial intelligence working paradigm has been implemented to bioprocess field in order to achieve a high degree of accuracy in process optimization and prediction. We may obtain the desired output for specific data inputs merely by adjusting the strengths of artificially simulated neuron in connecting layer.

6.2. Neural Network Working Plan and its Architechture

6.2.1. Feed Forward Neural Network Architechture

Artificial intelligence knowledge is required to be implemented to bioprocess field in recent times to achieve a high degree of accuracy in prediction and process optimization. Artificial Neural Networks (ANNs) are used for engineering purposes to process information and control automated systems, the scheme is shown in **Figure 6.2**.

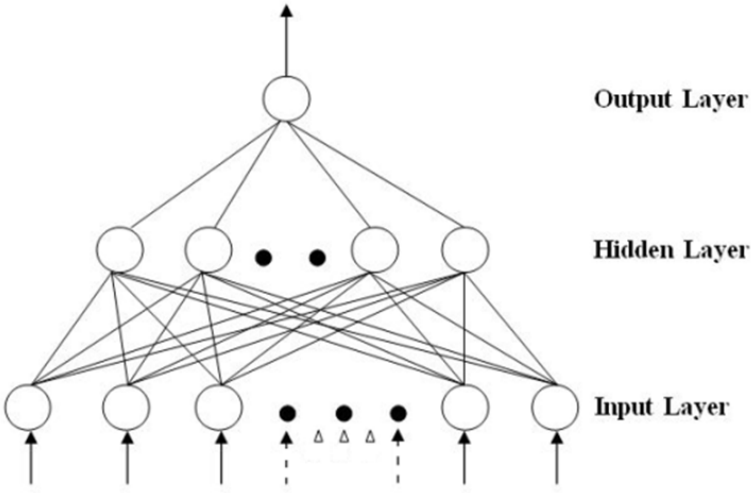


Figure 6.2. A general plan Neural Network with feedforward type architechture, showing input, hidden and output layer

A wide range of applications can be seen using different inputs from the bioprocess system to evaluate the output as production magnitude. Neural network has defined mechanism of adjusting the weights through learning or training. In our case the data is first normalized using upper and lower bound of the numerical value range and then the normalized data is used to serve as input for the neural network model.

6.2.2. Devised Algorithm for ANN Implementation

6.2.2.1. Error Minimization Approach

A multilayer perceptron is designed to process the weighted input. In our case sigmoidal function was used as an activation function in hidden layer using which the output was obtained. The taken input was given as different magnitude of culture conditions while the output is taken in terms of streptokinase production with time duration. To ensure optimal memory for the network ten hidden layers are taken after optimizing the number of layers to ensure optimal memory for the neurons. Training, testing and validation of data is repeatedly done. Allowing about few hundred of iterations is enough for efficient convergence of the error under the employed conditions. On obtaining the experimental data close to the simulated results, suggests the superior capability of neural network modelling for dynamic behaviour of the system (Horiuchi et al., 2001). The functional theme of ANN is clearly configured in famework, shown in **Figure 6.3 (a) and (b)**.

- Error Function € depends on the assumed weight values
- Gradient Descent: To minimize error by shifting weights along the decreasing slope of error
- The idea is just to iterate through the training set and adjusting the weights to minimize the gradient of the error

The process of gradient descent is made to minimize the error via iteration of training set and regulate the weights to by adjusting it along the descending slope of the computed error, **Eqs. 36-38**

$$€ = \sum_{X_p \epsilon\, E} (\theta_p - \phi_p)^2$$

(36)

The gradient of €:

$$\frac{\partial €}{\partial \omega} = \left[\frac{\partial €}{\partial \omega_1}, \dots, \frac{\partial €}{\partial \omega_p}, \dots, \frac{\partial €}{\partial \omega_{n+1}}\right]$$

$$O_q = f\left(\sum\nolimits_r w_{rq} u_r + w_{0q}\right)$$

$$= f\left(\sum\nolimits_r w_{rq}\left(f\left(\sum\nolimits_p w_{pr} x_p + w_{0r}\right)\right) + w_{0q}\right) \quad (37)$$

Where,

$$u_r = f\left(\sum\nolimits_p w_{pr} x_p + w_{0r}\right) \quad (38)$$

6.2.2.2. Back-Propagation and Learning

Data has been allowed to repeatedly pass through training, testing and validation subsequently. The input for the ANN model is the culture conditions while the output is streptokinase production obtained at certain duration of time. Inbuilt tool box in Matlab is employed to simulate the neural network and to make out the output; the statistical significance of the output is evaluated. The error minimization criterion in prediction is well achieved with high degree of accuracy. Back-propagation helps to compute the learning rate η and adjusting the weight. Usually, back-propagation algorithm (**Eq. 39**) is used:

$$\Delta_i w_{pq}(\delta) = -\eta \frac{\partial E^i}{\partial w_{pq}} + \mu \Delta_i w_{pq}(\delta - 1)$$

(39)

6.2.3. Working of Neural Network Framework

6.2.3.1. Data Normaliztion and Processing

From Neural Network Architecture, output is given by **Eqs. 40-41**,

$u_o = f(x_p, \omega_p)$

(40)

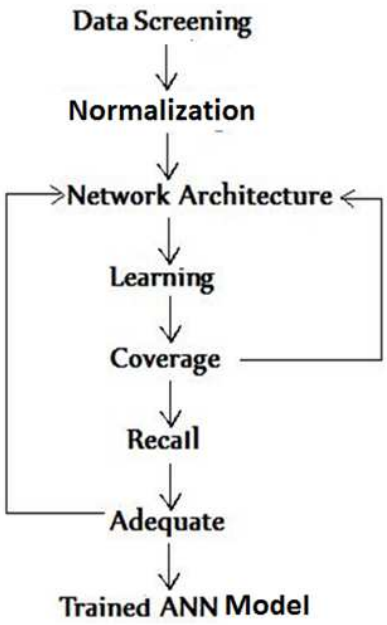
Simply, $u_o = \sum x_p \omega_p$

where, x_p – are inputs, u_o is output vector and ω_p - are the weights

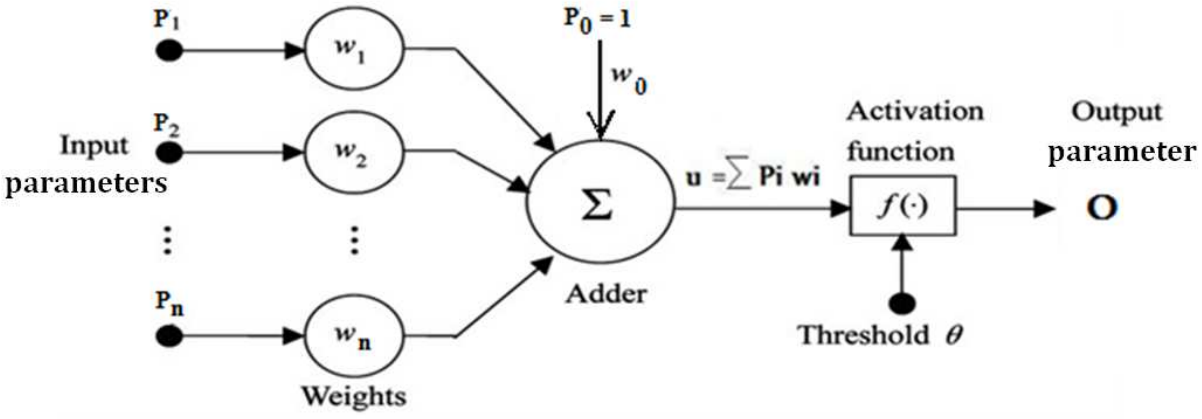
Activation function, $a_f = 1 / (1 + e^{-up})$ (41)

By adjusting the weights of an artificial neuron we can obtain the desired output for specific inputs. Intelligent descriptions of microbial processes in biorectors have some decisive avantages over mechanistic model under realistic conditions and has wide range of flexibility (Patnaik, 2009). A general plan and working of Neural Network is framed out in **Figure 6.3 (a)**. ANN weights should be taken iinitially with different start vectors and the network led to minimal 'test error' which provides adaptibility for its dynamic behavior.

The ANN model is first trained and then testing and validation is done using the part of the same dataset. For our purpose *nn* toolbox and code written in editor is used, available in Matlab 7.0 platform.



3(a)



3(b)

Figure 6.3 (a) and (b). Overall Working Plan of ANN parameters

6.2.3.2. Neural Network Architechture

Neural network has defined mechanism of adjusting the weights through learning or training. In our case the data is first normalized using upper and lower bound of the numerical value range and then the normalized data is used to serve as input for the neural network model.

Training, testing and validation of data is repeatedly done. A multilayer perceptron is designed to process the weighted input. Sigmoidal function is used as an activation function using which the output is obtained. The given input is population size of plasmid bearing and lacking cells and metabolite concentration while the output is taken in terms of streptokinase production with time duration.

The Idea is to iterate through the training set and adjust the weights to minimize the gradient of the error. The prime steps are summerized in the stepwise logical representaion as algorithm for Feed-forward ANN which had been utilized in our task, **Table 6.1**.

Table 6.1. FFNN Algorithm

Step 1: *Select the number of hidden layer neurons (h)*

Step 2: *Initialize the number of iterations with a desirable learning rate (η).*

Initialize the weights with small random values.

Initialize a threshold value (θ) as additional weight i.e. equal to 1.

Step 3: *For every Training set parameters, (Inputs (Media components, Media condition and set of media input, Output O_p) repeat steps 4-7.*

Step 4: *current input parameters to input layer nodes and the output to the output layer node;*

Step 5: *Calculate the input values to hidden layer nodes:*

$m_i = f(p_n, \omega_n)$

Simply, $m_i = \sum p_n \ \omega_n$

where, x_p – are inputs, u_o is output vector and ω_p - are the weights

Calculate the output value from hidden layer nodes

(Activation function):

$X_h = A_f = 1 / (1 + e^{-mi})$

Calculate the input values to the output nodes:

$O_i = f(X_h, \omega_{nk})$

Simply, $O_i = \sum X_h \ \omega_{nk}$

where, X_h – are inputs, O_i is output vector, ω_{nk} - are the weights and k=1 for single output node.

Calculate the corresponding output value:

$O = 1 / (1 + e^{-oi})$

Step 6: *Calculate the error term for the output node:*

The gradient of €:

$$€ = \sum_{X_p \epsilon\, E} (\theta_p - \phi_p)^2$$

$$\frac{\partial €}{\partial \omega} = \left[\frac{\partial €}{\partial \omega_1}, \ldots, \frac{\partial €}{\partial \omega_p}, \ldots, \frac{\partial €}{\partial \omega_{n+1}}\right]$$

$$€_k = (O_p - O) f'(O_i)$$

Calculate the error term for the hidden node:

$$€_h = (\mathrm{A}f)' \sum_k €_k\, \omega_{nk}$$

Step 7: *update the weights on the output layer as follows:*

$$\Delta_i w_{pq}(\delta) = -\eta \frac{\partial E^i}{\partial w_{pq}} + \mu \Delta_i w_{pq}(\delta - 1)$$

Step 8: *The network converged and successfully. Training is aborted.*

6.3. Results

Simulated plots of bioprocess system showing neural network applications has been obtained. Neural network has defined mechanism of adjusting the weights through learning or training. In our case the data is first normalized using upper and lower bound of the numerical value range and then the normalized data is used to serve as input for the neural network model.

$$N_X = \frac{(X - X_{min})}{(X_{max} - X_{min})} \tag{42}$$

Here, N_X denote the normalized concentration of data variable and subscripts *max* and *min* denote the maximum and minimum value of data variable in the input vector X respectively.

Simulated plot of bioprocess showing neural network applications with predicted and data point put for regression using the correlation method, **Figure 6.4**.

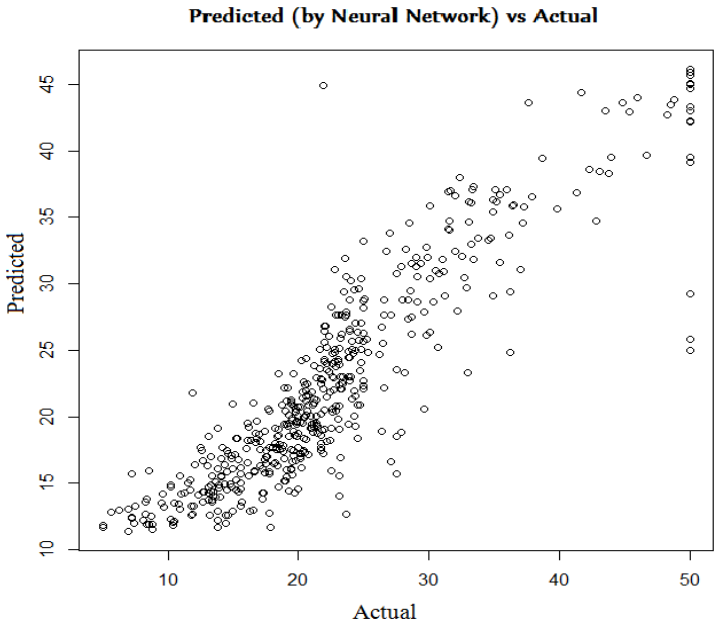


Figure 6.4. The typical actual and predicted data points distribution from ANN system

6.3.1 ANN approach for CCD used in Media optimization

The inputs of CCD for media component optimization was taken for neural network model, **Figure 6.5**, the simulation result provides for gradient descent, performance error minimization approach and regression plot for predected vs actual data, **Figure 6.6 (a)-(d)** respectively. The last plot of regression is compared with statistical method outcome. ANN has shown remarkable advantage over the later one. In **Table 5.5** of **chapter 5,** given the various inputs and statistically as well as ANN predicted values for media component optimization.

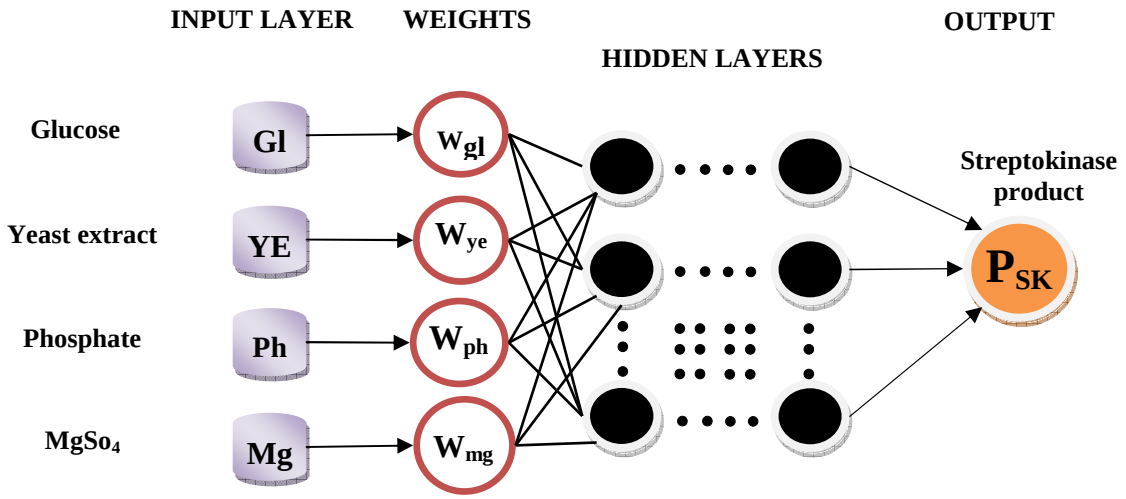


Figure 6.5. Neural network architecture using production medium components as input

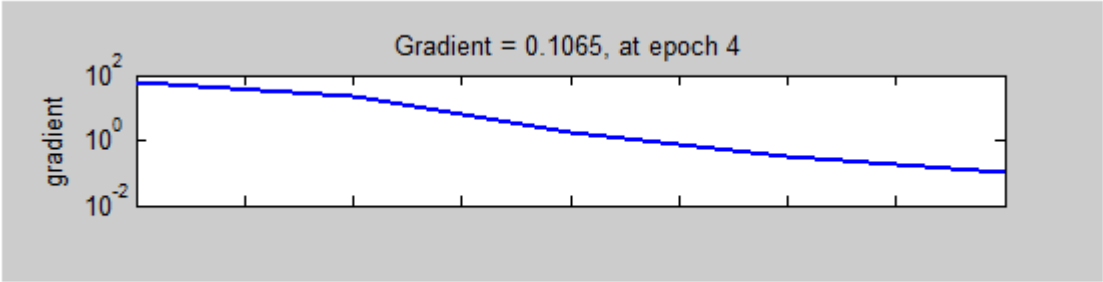


Figure 6.6(a). Showing the span of gradient function performance

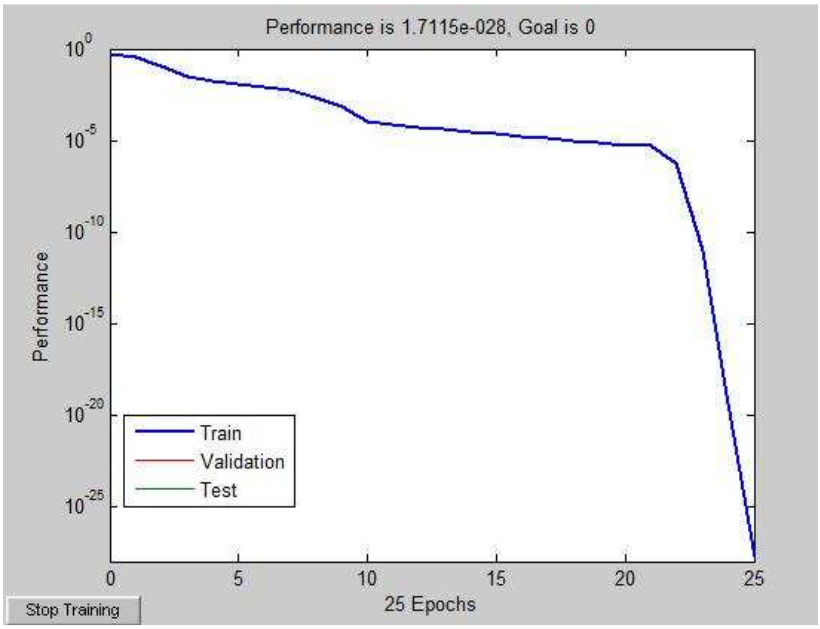


Figure 6.6(b). Showing performance of neural network via error minimization during training the media concentration data in ANN

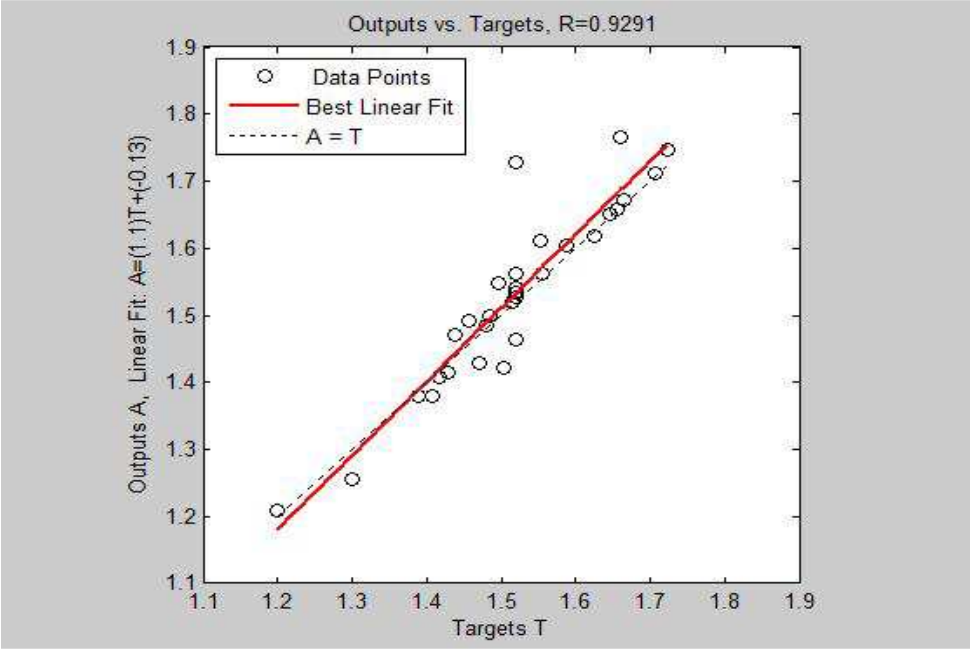


Figure 6.6(c). Plot of output vs target data obtained from ANN Model

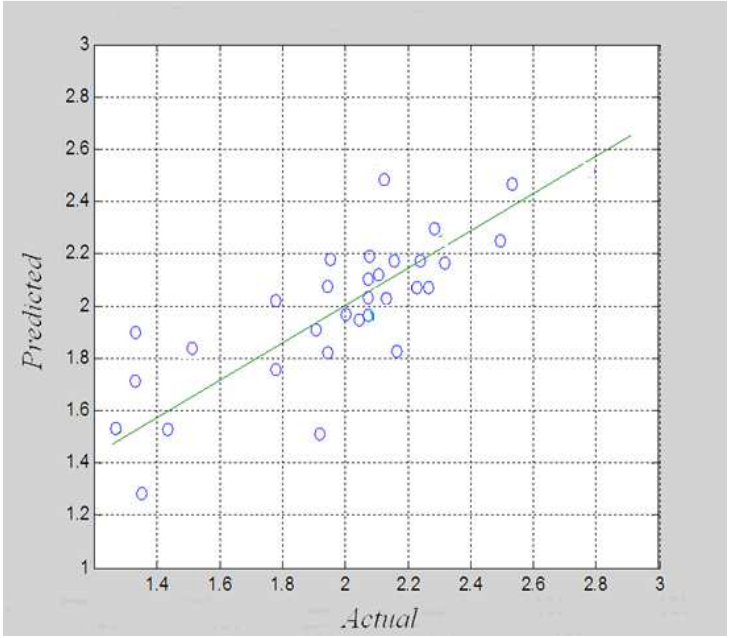


Figure 6.6(d). Regression plot obtained from statistical methodology, having correlation coefficient r^2 =0.7201

6.3.2. ANN approach for CCD used in Culture Condition Optimization

The input values of CCD for culture condition analysis was taken for neural network model, **Figure 6.7.** On simulation, result provides for gradient descent, performance error minimization approach and regression plot for predected vs actual data, **Figure 6.8 (a)-(d)** respectively. The last figure of regression is compared with statistical method outcome where ANN has shown better performance. The **Table 5.8** of **chapter 5,** provide the various inputs as well as statistically and ANN predicted values.

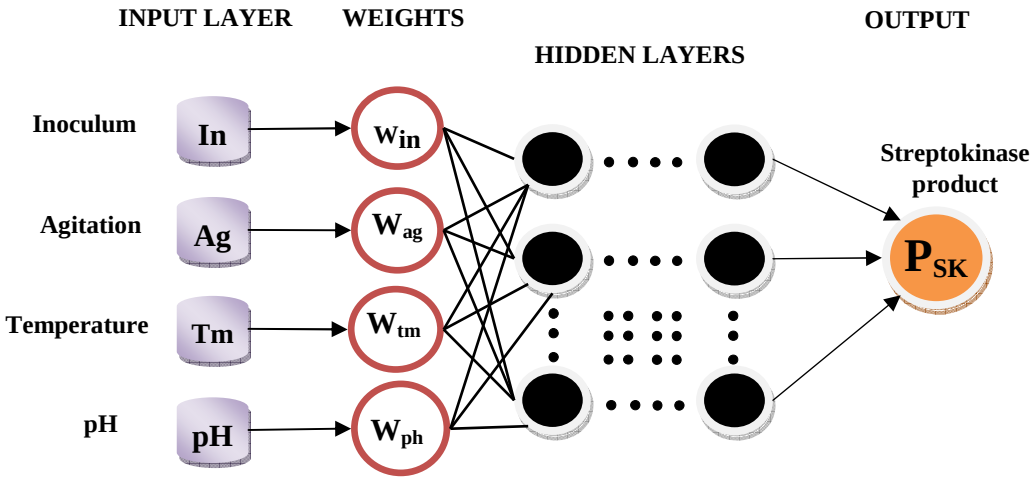


Figure 6.7. Use of Neural network in developing model using culture conditions parametrs

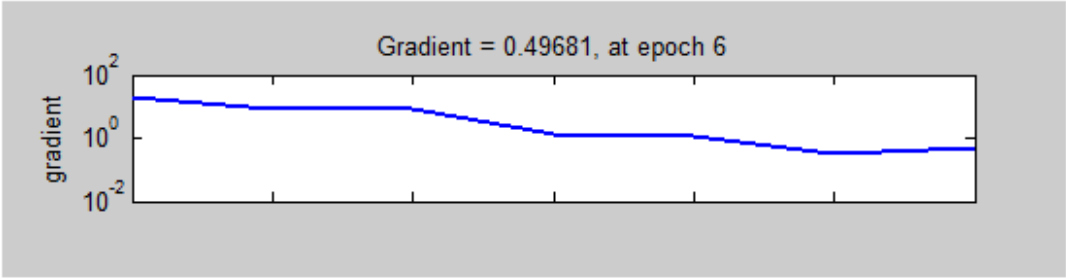


Figure 6.8(a). Gradient minimization function

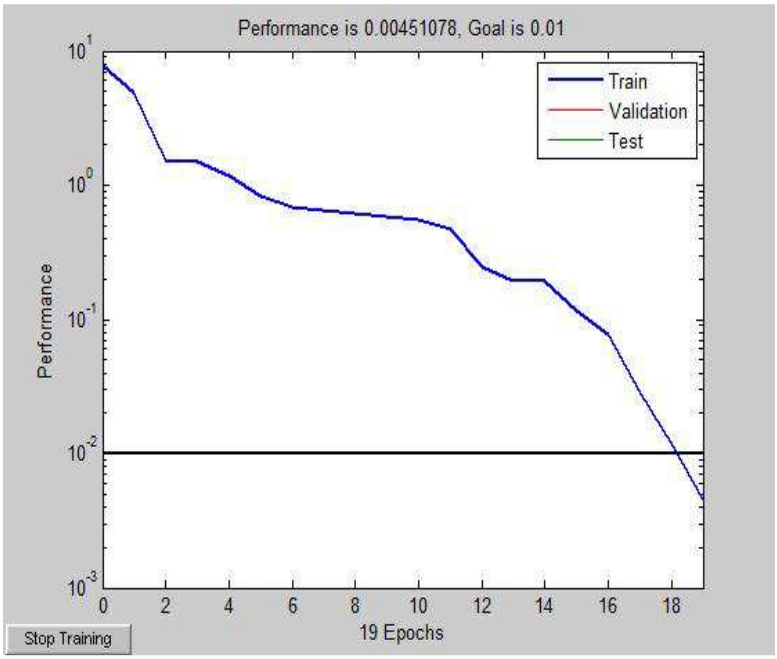


Figure 6.8(b). Performance of neural network with error minimization approach during training of culture condition parameter data in ANN

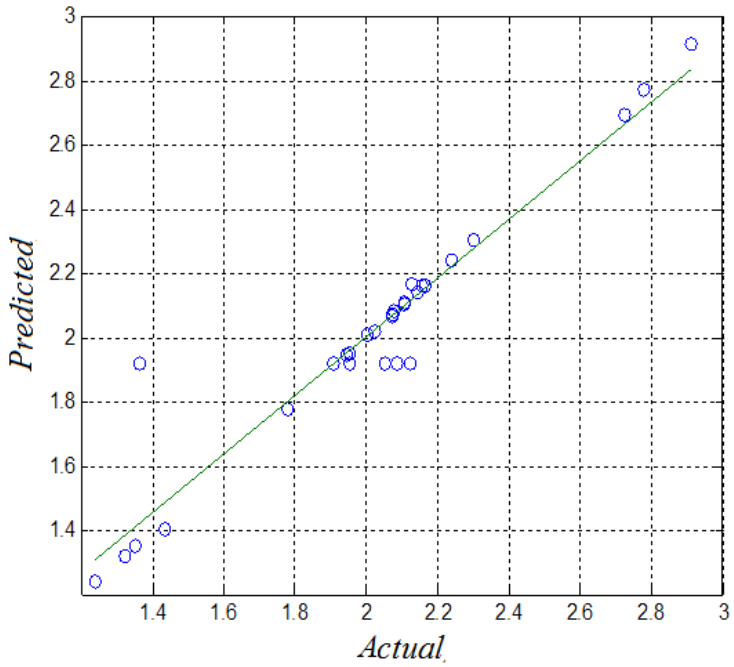


Figure 6.8(c). Regression plot for neural network results with coefficient $r^2 = 0.8632$

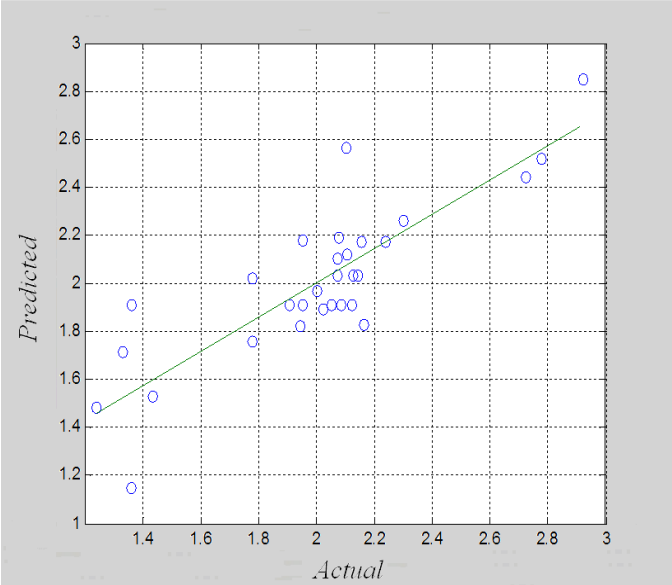


Figure 6.8(d). Regression plot for values obtained from statistical method, correlation coefficient r^2 =0.7102

6.3.3. ANN approach for predicting *N*+ cells population in population Dynamics model

The inputs for this model were obtained from population model dynamics model Eq. 8-17 (**Chapter 2**), the output vector of the respective parameters were given as input to neural network model, **Figure 6.9.**

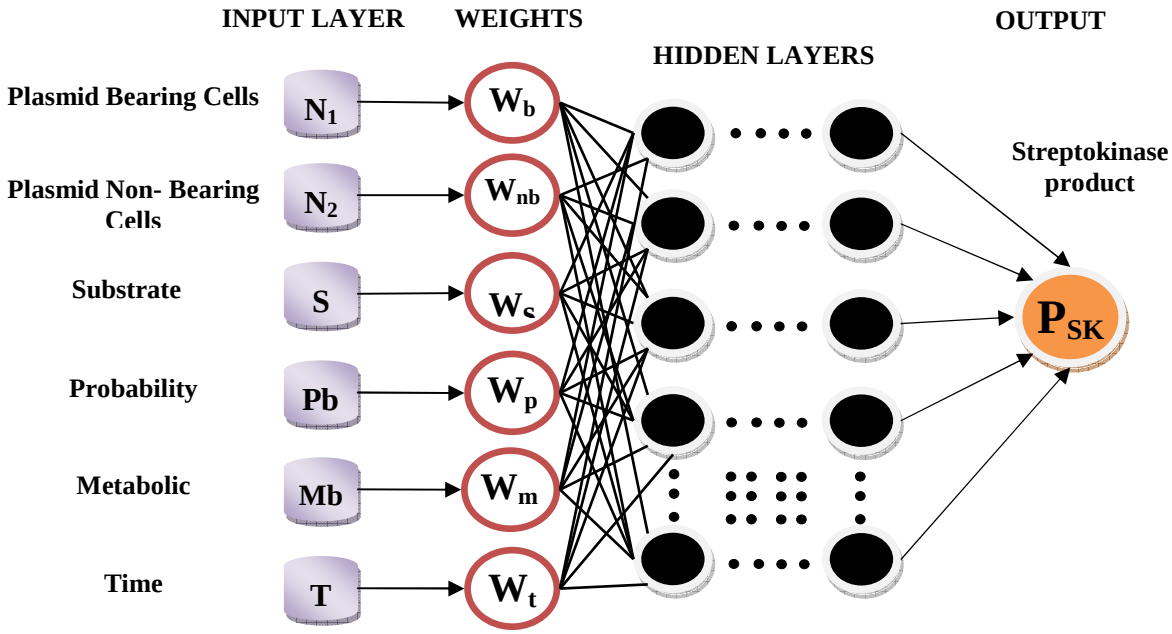


Figure 6.9. Showing input parameters of population dynamics model with wighted connections, hidden layer and output product, in neural network architecture

In this case also, result obtained in terms of gradient descent, performance error minimization approach and regression plot for predected vs actual data likewise, shown in **Figure 6.10 (a)-(d)** respectively. The last figure showed the plot of two methodology output i.e., the actual data is comparaed with simulated outcome from ANN, where ANN has shown high level of accuracy.

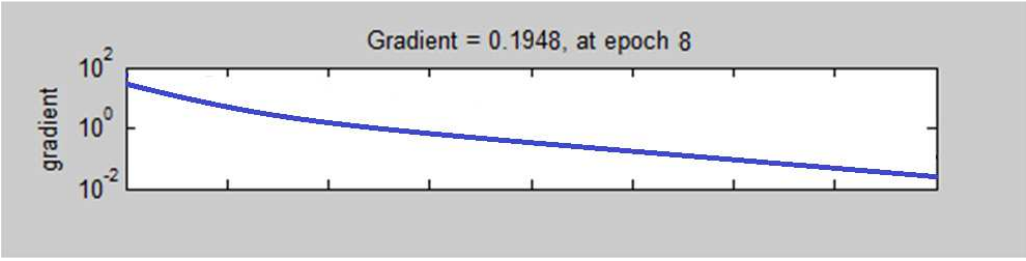


Figure 6.10(a). Gradient descent plot from ANN

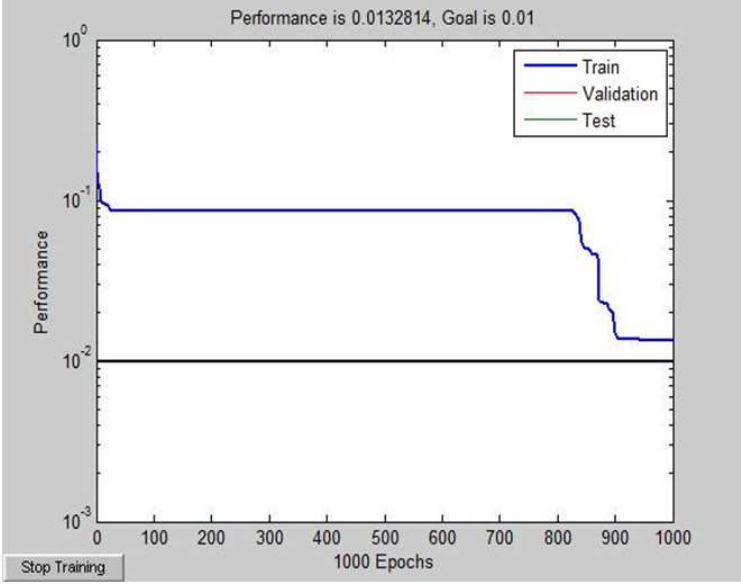


Figure 6.10(b). The performance efficiency of neural network in respect to error minimization during data training for population parameter model is represented

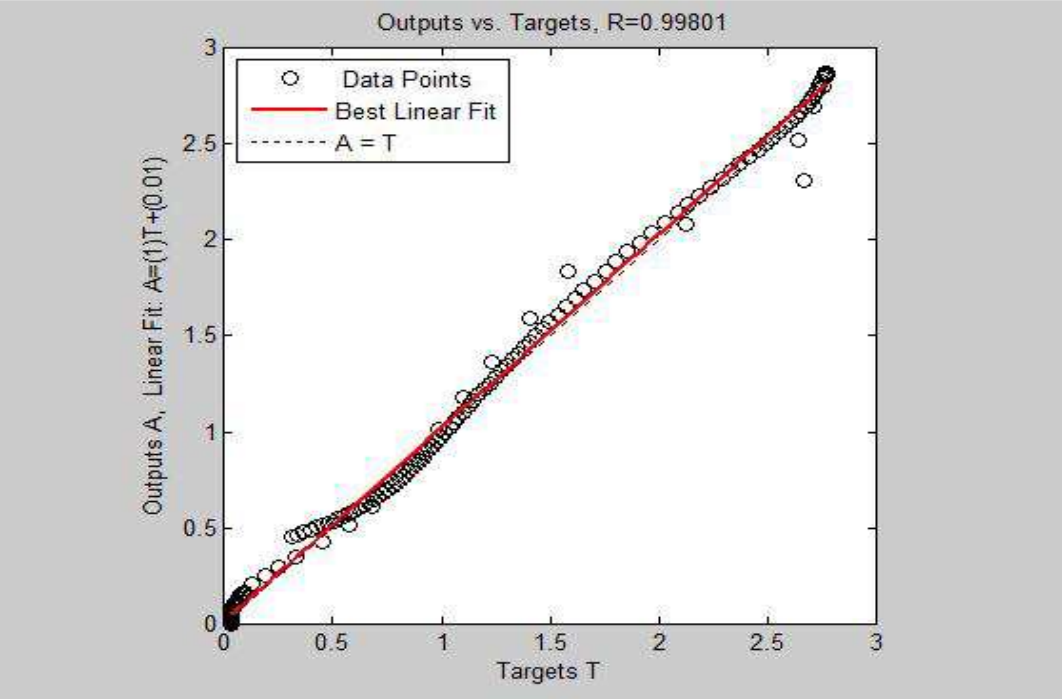


Figure 6.10(c). Showing the accuracy of neural network prediction, clearly presents the prdicted actual points of the data set. The correlation $\mathbf{r^2}$ in this case was found much closer to 1.0.

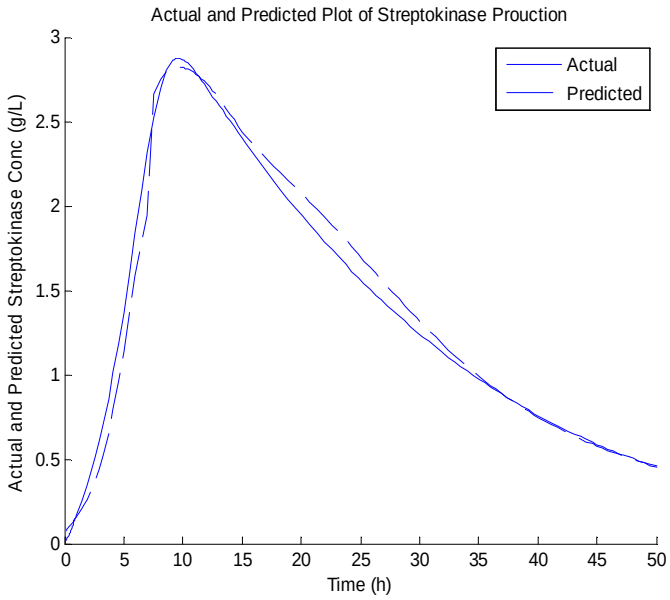


Figure 6.10(d). Data plot showing the actual (solid line) and predicted (dotted line) using feed-forward Neural network for product formation at dilution rate (D=0.12), on simulation of the model described in section A.

The various input data has been given in form of vector to neural network model, specified in *Appendix A2* together with ANN code for our bioprocess model.

6.4. Discussion

A supervision system has been configured for various facets of bioprocess field. The effort is to integrate the knowledge based systems, so neural network had been made to work with database mathematical models of the bioprocess. The process variable vectors were generated from the mathematical model simulation futher processed with neural network. The various process dynamics parameters once learned by the neural network after training, made the system could be supervised with previous interval database. The efficient on-line supervision of large scale recombinant streptokinase production had been considered and stages of the fermentation with output magnitute was taken into account.

It is evident that ANN model developed in reference to streptokinase production follow the actual data trend and satisfactory level of accuracy. In case of media composition and culture condition parameters modelling neural network has shown better prediction ability over ther classical statistical methodology. The characteristic ability of the neural network to learn complex type of non-linear input-output relation with requisite prior bioprocess knowledge base has been successfully used.

CHAPTER-7

CONCLUSION AND FUTURE PROSPECTS

7.1. Conclusion

Understanding the substantial role of vital parameters in streptokinase production to an extent provides a worthy approach to configure a machinery to evaluate the performance by reviewing the governing factors.

The population dynamics model is found to well depict the system behaviour with more inculcated real condition factors in comparison to the earlier established classical model systems. The plasmid copy number dynamics suggest to utilized the initial slot of span in fermentation process for better production magnitude and to neutralize the metabolite toxicity for rest of the span while copy number is having declining trend using weak alkaline solution. The use of GUI to simulate streptokinase selecting different set of parameters using variable dilution rate, substarte and inoculum concentration for taken time span made it feasible to observe the phenomena using desired level of input.

Four production media components were screened out by Plackett Burman methodology. Optimization of culture conditions was achieved to an extent in Central composite design technique, statistically model was found significant and lack of fit was not significant which is desirable. Four variables, viz., factors glucose, yeast extract, phosphate and $MgSO_4$, have shown there relevance in recombinant enzyme production, it would be due to the reason that glucose and yeast extract are prime carbon and nitrogen sources for proper cell growth, phosphate and $MgSO_4$ used to support ATP formation, protein synthesis at intracellular level. From the CCD analysis the updated combination of constituents has been obtained for production media while slightly enhanced level of pH, agitation and inoculum concentration is inferred from the CCD model for optimal production of streptokinase. Artificial Neural Network was also employed to predict the output and it has shown more accuracy in prediction in comparison to our statistical technique. The overall model supervision and output prediction using ANN approach is well performed and high level of precision.

The role of various parameters is identified using computational means, utilizing mathematical and statistical bioprocess algorithms. The representation of the dynamical system through modelling has its relevance in predicting the behaviour of the system on

changes made to initial conditions. In our proposed modelling strategy, new factors are taken into account to provide computational solution in different ways. The knowledge of interaction criteria of the various parameters governing streptokinase production made our task easier to configure a composite model. Screening vital media components statistically and then evaluating system structured and unstructured factors in model form would really ensure a high degree of precision in approach. On the other hand, the strategy to use variable antibiotic concentration prior supervising the bioprocess would be a novel approach to retain recombinant cell population for a greater span of time and thus enhancing the product yield. The structured model framework is developed to evaluate the subsequent performance and reliability of sub-cellular components in temporal fashion to give desired level of production of recombinant enzyme.

7.2. Future Prospects

The representation of fermentation process has several facets that are still required to be incorporated to improve the model. The area of enzyme engineering would be taken further to deal with streptokinase production implementing the existing usable information.

The proposed work in this thesis can be extended to evaluate batch and continuous fermentation for synthesizing other recombinant products on pilot scale using online process applications via controller system. The previously observed trends of parameter dynamics for crucial time span would be helpful to update and improve the fermentation strategy.

Use of micro-bioreactors to generate extensive data can be helpful to frame models that are more sophisticated, thereby it would be possible to test several parameters together utilizing only small amount of culture medium and test constituents to generate ample of output. A server with data-base can be further established installing bioprocess programmed tools and graphics to make available the use of online validated model facility for testing and organizing similar kind of work to the user.

The industrial level production of recombinant streptokinase needs to be carried out taking directions to collectively incorporate the mathematical as well as intelligence-based techniques. So training of model would be once required on previous knowledge database and then implementation can be made accordingly to avoid extra number of classical trials for optimization.

Several strategies of removing noise in carrying out fermentation process is under way, there is a need to evaluate other intelligence paradigm which may bring more accuracy in dealing with bioprocess related issues.

The emphasis also had been on structured model to understand the intra-cellular biochemical state and partitioning event of cell cycle. Likewise, the state of the entire cell population need to be assumed and more hybrid model as well as composite methods are desirable.

Furthermore, the study of the biological process happening within each cell at subcellular level may require to be incorporated together with genetic factors that affect the stability of the vector. There are several other unstructured features including metabolic pathways are to be considered to effectively evaluate and understand the unravelled causes of instability in fermentation process.

AUTHOR'S RESEARCH CONTRIBUTION

[1]. P. Kumar and S. Ghosh, "Population dynamics model for plasmid bearing and plasmid lacking cells for streptokinase production in continuous flow stirred tank bioreactor", International Journal of Engineering, Science and Technology (IJEST), **Multicraft Publisher**, Vol. 2, No. 5, pp. 118-127, 2010. (**3 Citations: viz. Elsevier, Taylor Francis and IIIS**)

[2]. P. Kumar and S. Ghosh, "Use of Simulation and Intelligence Based Optimization Approach in Bioprocess", Soft Computing for Problem Solving (SocProS-2013), vol. 2, **Springer Series, Advances in Intelligent and Soft Computing**, vol. 259, 2013.

[3]. P. Kumar and S. Ghosh, "Bioprocess System Dynamics Based Computational and Statistical Models", **Elsevier Procedia of Science and Technology,** IMS-2013, pp. 162-165, 2013.

[4]. P. Kumar and S. Ghosh, "Application of Neural Network and Genetic Algorithm Based Approaches to Bioprocess", **IEEE Xplore** published proceedings (ICICT), 2014. (Accepted)

[5]. P. Kumar and S. Ghosh, "Modelling of Bioprocess for Streptokinase Production Using Mechanistic and Neural Network Approaches", 3rd International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC), Vol. 1, pp. 161-166, March 25th-28th, 2012, International Institute of Informatics and Systemics **(IIIS)**, Orlando, Florida, USA.

[6]. P. Kumar and S. Ghosh, "Application of Computational Approaches to improve Streptokinase Production Process", Proceedings of World Congress on Biotechnology (Bright ICE Group) - 2012, May. 4th-6th, 2012, Hyderabad, INDIA.

[7]. P. Kumar and S. Ghosh, "Application of Computational Means to Identify Bioprocess Problems", 15th International Biotechnology Symposium and Exhibition (IBS-2012), pp. 66, September 16th-21st, 2012, Daegu, S. KOREA.

[8]. P. Kumar and S. Ghosh, "Use of Statistical and Computational Intelligence Based Optimization Technique to Increase Streptokinase Production", **J. Biotechnol. Biomater.,** 3rd World Congress on Biotechnology (**OMICS Group Conferences**), Vol. 2, Issue 6, pp. 244, September 13th -15th, 2012, Hyderabad, INDIA.

[9]. P. Kumar and S. Ghosh, "Estimation of Streptokinase Production during Fermentation Process Using Hybrid Approach", 1st International Science Congress-2011, pp. 63, Dec. 24th-25th, 2011, Indore, INDIA.

[10]. P. Kumar and S. Ghosh, "Modelling of Key Factors Governing E. coli Recombinant Strain Growth Dynamics in Bioreactor Environment", International Congress on Environmental Research (ICER-2011), pp.200, Dec. 15th -17th, 2012, Surat INDIA.

[11]. P. Kumar and S. Ghosh, "Application of Neural Network to improve Recombinant enzyme production", Proceedings of National Conference on Recent Trends in Mathematics and Statistics (CORTMAS-2012), pp. 29, March 12th-13th, 2012, Department of Mathematics, DDU University Gorakhpur, INDIA.

[12]. P. Kumar and S. Ghosh, "Statistical Optimization of E. coli Strain Growth in Submerged Culture Using Response Surface Methodology", CORTMAS-2012, pp. 30, March 12th-13th, 2012, Department of Mathematics, DDU University Gorakhpur, INDIA.

[13]. P. Kumar and S. Ghosh, "Enhanced Recombinant Enzyme Production using Computational Approach", Proceedings of National Conference on Global Challenges: Role of Science and Technology in giving their solutions (GCRSTS) – 2012, pp. 16, March 3th-4th, 2012, Bhiwani, INDIA.

[14]. P. Kumar and S. Ghosh, "Bioprocess Computational Models to Identify Role of Various Parameters in Streptokinase Production", Proceedings of National Conference on Biotechnology and Biodiversity, pp. 88-108, March 13th-14th, 2012, Rewa, INDIA. (**Young Scientist Award for Best Paper in the Conference**)

[15]. P. Kumar and S. Ghosh, "Use of Computational and Statistical Strategy to Enhance Recombinant Enzyme Production", Theoretical Biology, **Elsevie**r. **(**Communicated)

[16]. P. Kumar and S. Ghosh, "Statistical and Neural Network Modelling approaches to estimate Streptokinase Production in Batch and Continuous culture", International journal of Biomathematics, **World Scientific**. **(**Communicated)

APPENDIX - A1

GUI simulation code

function GUI()

```
% SIMPLE_GUI2 Select a data set from the pop-up menu, then
% click one of the plot-type push buttons. Clicking the button
% plots the selected data in the axes.

f = figure('Visible','off','Position',[360,500,450,285]);

  %  Construct the components.
  hsurf = uicontrol('Style','pushbutton','String','Antibiotics Addition',...
        'Position',[315,230,50,15],...
        'Callback',{@surfbutton_Callback});
  hmesh = uicontrol('Style','pushbutton','String','Bioprocess Simulation',...
        'Position',[315,200,50,15],...
        'Callback',{@meshbutton_Callback});
  hmesha = uicontrol('Style','pushbutton','String','Biopr Sim Together',...
        'Position',[315,170,50,15],...
        'Callback',{@meshbutton_Callbacka});
  hmeshb = uicontrol('Style','pushbutton','String','All Factors Sim',...
        'Position',[315,140,50,15],...
        'Callback',{@meshbutton_Callbackb});
  hmeshc = uicontrol('Style','pushbutton','String','Percent Pop Sim',...
        'Position',[315,110,50,15],...
        'Callback',{@meshbutton_Callbackc});
  hmeshd = uicontrol('Style','pushbutton','String','Batch Culture',...
        'Position',[315,80,50,15],...
        'Callback',{@meshbutton_Callbackd});
  hcontour = uicontrol('Style','pushbutton',...
        'String','Close',...
        'Position',[315,60,50,10],...
        'BackgroundColor','g',...
        'Callback',{@contourbutton_Callback});


  ha = axes('Units','Pixels','Position',[50,60,200,185]);
  align([hmesh,hmesha,hmeshb,hmeshc,hmeshd,hsurf,hcontour],'Center','None');

S.sl = uicontrol('style','slide',...
              'unit','pix',...
              'position',[390,40,80,20],...
              'min',0.01,'max',1,'val',0.50);
S.sll = uicontrol('style','slide',...
              'unit','pix',...
              'position',[520,40,80,20],...
              'min',5,'max',240,'val',100);
S.edd1 = uicontrol('style','edit',...
              'unit','pix',...
              'position',[260,40,120,20],...
              'fontsize',12,...
              'string','Dillution Rate');
S.ed = uicontrol('style','edit',...
              'unit','pix',...
              'position',[390,70,80,20],...
              'fontsize',12,...
              'string','0.50 per hr');
S.edd2 = uicontrol('style','edit',...
```

```
                'unit','pix',...
                'position',[610,40,120,20],...
                'fontsize',12,...
                'string','Time Duration');
S.edd = uicontrol('style','edit',...
                'unit','pix',...
                'position',[520,70,80,20],...
                'fontsize',12,...
                'string','100 hrs');
S.edd2aa = uicontrol('style','edit',...
                'unit','pix',...
                'position',[450,770,500,50],...
                'fontsize',20,...
                'BackgroundColor','Cyan',...
                'string','Bioprocess Model Dynamics Simulator');
S.sl2 = uicontrol('style','slide',...
                'unit','pix',...
                'position',[980,40,80,20],...
                'min',0.1,'max',10,'val',0.50);
S.sll2 = uicontrol('style','slide',...
                'unit','pix',...
                'position',[1100,40,80,20],...
                'min',1,'max',100,'val',100);
S.edd12 = uicontrol('style','edit',...
                'unit','pix',...
                'position',[850,40,120,20],...
                'fontsize',12,...
                'string','Inoculum Conc');
S.ed2 = uicontrol('style','edit',...
                'unit','pix',...
                'position',[980,70,80,20],...
                'fontsize',12,...
                'string','5 percent');
S.edd22 = uicontrol('style','edit',...
                'unit','pix',...
                'position',[1190,40,120,20],...
                'fontsize',12,...
                'string','Substrate Conc');
S.edd12 = uicontrol('style','edit',...
                'unit','pix',...
                'position',[1100,70,80,20],...
                'fontsize',12,...
                'string','50 g/L');
set([S.ed,S.sl,S.sll,S.edd,S.ed2,S.sl2,S.sll2,S.edd12],'call',{@ed_call,S});  % Shared Callback.

function [] = ed_call(varargin)
% Callback for the edit box and slider.
[h,S] = varargin{[1,3]};
   case S.ed
      L = get(S.sl,{'min','max','value'});
      E = str2double(get(h,'string'));
      if E >= L{1} && E <= L{2}
         set(S.sl,'value',E)
      else
         set(h,'string',L{3}) % User tried to set slider out of range.
      end
   case S.sl
      set(S.ed,'string',get(h,'value'))
      dilution_rate=get(h,'value');
      save dilution_rate
```

```
        disp ('Dilution rate changed');
    case S.edd
        L = get(S.sll,{'min','max','value'});
        E = str2double(get(h,'string'));
        if E >= L{1} && E <= L{2}
            set(S.sll,'value',E)
        else
            set(h,'string',L{3})
        end
    case S.sll
        set(S.edd,'string',get(h,'value'))
        time_duration=get(h,'value');
        save time_duration
        disp ('Duration fixed now');
    case S.ed2
        L = get(S.sl2,{'min','max','value'});  % Get the slider's info.
        E = str2double(get(h,'string'));  % Numerical edit string.
        if E >= L{1} && E <= L{2}
            set(S.sl,'value',E)  % E falls within range of slider.
        else
            set(h,'string',L{3}) % User tried to set slider out of range.
        end
    case S.sl2
        set(S.ed2,'string',get(h,'value')) % Set edit to current slider.
        inoculum_conc=get(h,'value');
        save inoculum_conc
        disp ('Inoculum concentration changed');
    case S.edd2
        L = get(S.sll2,{'min','max','value'});  % Get the slider's info.
        E = str2double(get(h,'string'));  % Numerical edit string.
        if E >= L{1} && E <= L{2}
            set(S.sll,'value',E)  % E falls within range of slider.
        else
            set(h,'string',L{3}) % User tried to set slider out of range.
        end
    case S.sll2
        set(S.edd12,'string',get(h,'value')) % Set edit to current slider.
        substrate_conc=get(h,'value');
        save substrate_conc
        disp ('Substrate amount fixed now');
     otherwise
end
end
  set([f,ha,hsurf,hmesh,hmesha,hmeshb,hmeshc,hmeshd,hcontour],...
  'Units','normalized');
  set(f,'Name','GUI for Bioprocess Dynamics')
%    % Move the GUI to the center of the screen.
  movegui(f,'center')
%    % Make the GUI visible.
  set(f,'Visible','on');

  function surfbutton_Callback(source,eventdata)
%    % Display surf plot of the currently selected data.
      bior_paper2_antibiotica();
  disp ('Antibiotics addition regulated successfully')
  disp ('Population Dynamics Model')
  end

  function meshbutton_Callback(source,eventdata)
      bior_paper2_2_a();
```

```
      disp ('Population Dynamics Model without Additional Antibiotics')
  end

function meshbutton_Callbacka(source,eventdata)
      bior_paper2_2aa();
      disp ('Population Dynamics Model, with and without Additional Antibiotics')
end

function meshbutton_Callbackb(source,eventdata)
      bior_paper2_2ba();
      disp ('Population Dynamics Model Showing All Factors Dynamics')
end

function meshbutton_Callbackc(source,eventdata)
      bior_paper2_2percenta();
      disp ('Percentage of Population Bearing and Lacking Cells Dynamics Model')
end

function meshbutton_Callbackd(source,eventdata)
      bior_paper2_2batch();
      disp ('Batch Culture Dynamics Model')
end

function popup_menu_Callback(hObject, eventdata, handles)
r=rand(1,10);
if r<=0.5
   value=0
else
   value=1
end
val = get(hObject,'Value');
switch val
case 1
   if val==0
      hobject=10;
   end
% % The user selected the first item
case 2
   if val==1
      hobject=50;
end
  function contourbutton_Callback(source,eventdata)
   disp 'Thanks for Using This Bioprocess Simulation Tool'
   clear all;
   close all;
   close();
  end
end
```

Appendix - A2

Population parameter vectors from model

t	$X1$	$X2$	S	q	M	P
0	0.7000	0	70.0000	0	0	0
0.0001	0.7000	-0.0000	69.9998	-0.0000	0.0001	0.0000
0.0002	0.7000	-0.0000	69.9996	-0.0000	0.0001	0.0000
0.0003	0.7001	-0.0000	69.9994	-0.0000	0.0002	0.0000
0.0004	0.7001	-0.0000	69.9992	-0.0000	0.0002	0.0001
0.0008	0.7002	-0.0000	69.9982	-0.0000	0.0005	0.0001
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2.6992	0.9794	-0.0108	67.5929	-0.0000	0.6138	0.1601
3.0720	0.9891	-0.0113	67.5137	-0.0000	0.6340	0.1654
3.4449	0.9990	-0.0117	67.4339	-0.0000	0.6543	0.1707
3.8177	1.0090	-0.0122	67.3537	-0.0000	0.6748	0.1761
4.1905	1.0146	-0.0125	67.3082	-0.0000	0.6864	0.1792
⋮	⋮	⋮	⋮	⋮	⋮	⋮
16.9858	5.9648	3.5652	7.1731	0.0007	16.0209	2.7230
17.4003	5.9032	3.6435	7.0292	0.0007	16.0576	2.6964
17.9097	5.8395	3.7201	6.9121	0.0007	16.0874	2.6687
18.4192	5.7710	3.7995	6.8090	0.0008	16.1137	2.6388
18.9286	5.7012	3.8776	6.7232	0.0008	16.1356	2.6082
⋮	⋮	⋮	⋮	⋮	⋮	⋮
46.3927	0.6243	9.1695	4.5108	0.0050	16.6997	0.2902
46.8920	0.6035	9.1912	4.5038	0.0051	16.7015	0.2806
47.3728	0.5834	9.2119	4.4995	0.0052	16.7026	0.2712
47.8537	0.5639	9.2320	4.4948	0.0052	16.7038	0.2622
48.3346	0.5447	9.2522	4.4873	0.0053	16.7057	0.2533
⋮	⋮	⋮	⋮	⋮	⋮	⋮
58.4767	0.3561	9.4481	4.4325	0.0061	16.7197	0.1657
58.8575	0.3439	9.4608	4.4281	0.0062	16.7208	0.1600
59.2383	0.3322	9.4727	4.4268	0.0063	16.7212	0.1546
59.6192	0.3208	9.4843	4.4249	0.0063	16.7216	0.1493
60.0000	0.3120	9.4936	4.4212	0.0064	16.7226	0.1452

Matlab Codes to generate a population model :

```
function bior_dilrates();
bior_dilrate1();
bior_dilrate2();
%.....................................................................
function bior_dilrate1();
% clear all
time=60;
[t,x]=ode45(@bior_dy221,[0,time],[0.7 0.00 70 0.00 0.00 0.00]);
%.....................................................................
figure(1);
plot(t,x(:,1),t,x(:,2));
title('Dynamics of Plasmid free and bearing cells at D rate 0.10 per hr');
xlabel('Time, h'),ylabel('Cell Population/Concentration, g/l');
legend('x1 (Pl Bearing)','x2 (Pl Non-Bearing)');
axis([0 time 0 15]);hold on;
%%
function dydt22a1 = bior_dy221(t,x);
e=2.71828;
x1=x(1);
x2=x(2);
s=x(3);
q=x(4);
m=x(5);
p=x(6);
d=0.10;
mum1=0.74; %for plasmid bearing cells
mum2=0.80; %for plasmid lacking cells
ks1=40;
ks2=20;
kp=0.0005;
kd=0.020;
yp=0.44;
s0=70; %For all other dilution rates
y=0.15;
ym=1.7;
m_th=0.75;
mp=0.010;
f1=mp-r+z;
f2=r+z;
if m > m_th
    mf = f1;
elseif m < m_th
    mf = f2;
    m1=0;
    m2=0;
end
mu1=mum1*s/(ks1+s);
mu2=mum2*s/(ks2+s);
dydt22a1 = [mu1*x1*(1 - q) - d*x1 - m1*mu1*x1 + r1* mu1*x1;...
          mu2*x2 + q*mu1*x1 - d*x2 + m1*mu1*x1 - m2*mu2*x2 - r1*mu1*x1;...
          d*(s0 -  s) - 1/y*( mu1*x1 + mu2*x2);...
          (1 - e^( -mf*(x2/(x1+x2))))* mu2;...
          ym*(x1*mu1 + x2*mu2) - d*m;...
         (yp*(mu1-q*kd)*x1-kp*p-d*p)];
```

```
%.....................................................................
%.....................................................................

function bior_dilrate2();
% clear all
time=60;
[t,x]=ode45(@bior_dy222,[0,time],[0.7 0.00 70 0.00 0.00 0.00]);
%.....................................................................
figure(2);
plot(t,x(:,1),'--',t,x(:,2),'--');
title('Dynamics of Plasmid free and bearing cells at D = 0.15 per hr');
xlabel('Time, h'),ylabel('Cell Population/Concentration, g/l');
legend('x1 (Pl Bearing)','x2 (Pl Non-Bearing)');
axis([0 time 0 15]);hold on;
%%
function dydt22a2 = bior_dy222(t,x);
e=2.71828;
x1=x(1);
x2=x(2);
s=x(3);
q=x(4);
m=x(5);
p=x(6);
d=0.15;
mum1=0.74; %for plasmid bearing cells
mum2=0.80; %for plasmid lacking cells
ks1=40;
ks2=20;
kp=0.0005;
kd=0.020;
yp=0.44;
s0=70;
y=0.15;
ym=1.7;
m_th=0.75;
mp=0.010;
f1=mp-r+z;
f2=r+z;
if m > m_th
    mf = f1;
elseif m < m_th
    mf = f2;
    m1=0;
    m2=0;
end
mu1=mum1*s/(ks1+s);
mu2=mum2*s/(ks2+s);
dydt22a2 = [mu1*x1*(1 - q) - d*x1 - m1*mu1*x1 + r1* mu1*x1;...
          mu2*x2 + q*mu1*x1 - d*x2 + m1*mu1*x1 - m2*mu2*x2 - r1*mu1*x1;...
          d*(s0 -  s) - 1/y*( mu1*x1 + mu2*x2);...
          (1 - e^( -mf*(x2/(x1+x2))))* mu2;...
          ym*(x1*mu1 + x2*mu2) - d*m;...
          (yp*(mu1-q*kd)*x1-kp*p-d*p)];
%.....................................................................
%.....................................................................
```

```
function nnprogram()

%Design for Training
clc; clear all;
P = [1 to n]; %Training Patterns (domain values)
T=[1 to n];
T=T';
net = newff([0 161;0 161;0 161;0 161;0 161],[5 1],{'tansig', 'purelin'});
% Input Vectors for training
% To plot the original data points and ANN output
Y = sim(net,P);
plot(P,T,P,Y,'o');
title('Data and Untrained Network Output');
%Train the network and plot the results
net.trainParam.goal=0.01;
net.trainParam.epochs = 100;
plot(P,T,'ko',X,Y);

%An alternative way to test training: postreg
Tout=sim(net,P); %Get network output for the training domain
[m,b,r]=postreg(T,Tout); %Performs a linear regression
T=T(:);
Tout=Tout(:);
x=[];y=[];
disp('S.No Actual Predicted');
disp('_____________________');
for i=1:length(T);
disp([sprintf('%d\t%0.4f\t%0.4f\n',round(i),T(i),Tout(i))]);
end;
title('Actual and Predicted Plot of Streptokinase Prouction');
xlabel('Time (h)'),ylabel('Actual and Predicted Streptokinase Conc (g/L)');
plot(t,T); plot(t,Tout,'--');
legend('Actual','Predicted');
x=x(:);
 [sprintf('%d\n',round(x)) T Tout];
([x A])
sprintf('%d \n%0.4f \n%0.4f\n',A)
```

- Training Session 1 explanation: Appearing in the first argument of *newff*. Notice also that the sizes of the domain and range sets are given as m × n, where m = dimension and n =number of data points.
- The arguments shown in *newff* are required; notice that the first argument gives the minimum and maximum values in the domain set. Normally, one uses *minmax*(P) in place of actually typing these values in. This also implicitly defines the number of nodes in the input layer (which is the domain dimension).
- The second argument in *newff*, defines the number of nodes in the hidden layer and in the output layer. This vector also implicity defines how many layers you have by counting the size of this vector. Thus, we can have as many layers as we wish.
- The last required argument is the type of transfer function. The *tansig* function is the inverse tangent function, which gives the same performance as our standard sigmoidal function. We will always use a linear layer in the output, so this argument is always *purelin*.
- The default training (optimization) method is combination of gradient descent and Newton's Method). In the next training/testing session, we will use a different method.

BIBLIOGRAPHY

Abrams, P., *The effect of density-independent mortality on the coexistence of exploitative competitors for renewing resources*, Am. Nat., **158(5)**: 459-470, (2001).

Ali, E., Ajbar, A.J. and Alhumaizi, K., *Dynamics of recombinant DNA cultures under time varying feed conditions*, Chem. Eng. Comm., **199(2)**: 1155-1168, (2012).

Ataai, M.M. and Shuler, M.L., *A Mathematical Model for Prediction of Plasmid Copy Number and Genetic Stability in Escherichia coli*, Biotechnol. Bioeng., **30(3)**: 389-397, (1987).

Bailey, J.E. and Ollis, D.F., *Biochemical Engineering Fundamentals*, Mc Graw-Hill Book Co., New York, **4**: 79-108, (1986).

Balagurunathan, B., Ramchandra, N.S. and Jayaraman, G., *Enhancement of stability of recombinant streptokinase by intracellular expression and single step purification by hydrophobic interaction chromatography*, Biochem. Eng. J., **39(1)**: 84-90, (2008).

Banerjee, S. and Debnath, M., *Continuous Production of 6-amino Penicillanic Acid (6-APA) by Agarose Immobilized Penicillin Acylaseina Packed Column Reactor*, Chem. Biochem. Eng., **21(3)**: 265-269, (2007).

Bezerra, M.A., Santelli, R.E., Oliveira, E.P., Villar, L.S. and Escaleira, L.A., *Response surface methodology (RSM) as a tool for optimization in analytical chemistry*, Talanta, **76(5)**: 965-977, (2008).

Bingle, L.E.H. and Thomas, C.M., *Regulatory circuits for plasmid survival*, Curr. Opin. Microbiol.**, 4(2):** 194-214, (2001).

Boros, I., Pósfai, G. and Venetianer, P., *High-copy-number derivatives of the plasmid cloning vector pBR322*, Gene, **30(1-3)**: 257-260, (1984).

Box, G.E.P. and Behnken, D.W., *Some New Three Level Designs for the Study of Quantitative Variables*, Technometrics, **2(4)**: 455-475, (1960).

Bruns, R.E., Scarminio, I.S. and Neto, B.D.B., *Statistical Design-Chemometrics*, Elsevier, Amsterdam, (2006).

Chauhan, B. and Gupta, R., *Application of statistical experimental design for optimization of alkaline protease production from Bacillus sp. RGR-14*, Proc. Biochem., **39(12)**: 2115-2122, (2004).

Chitra, S.P., *Use neural networks for problem solving*, Chem. Eng. Prog., **89(4)**: 44-52, (1993).

Choi, J.H. and Lee, S.Y., *Secretory and extracellular production of recombinant proteins using Escherichia coli*, Appl. Microbiol. Biotechnol., **64(4)**: 625-635, (2004).

Clarke, D.W., Mohtadi, C. and Tuffs, P.S., *Generalized predictive control-Part I. The basic algorithm*, Automatica, **23(2)**: 137-148, (1987).

Costa, E., Lucas, A., Aguado, J., A-vila, J.A., *Transferencia de materia en tanques agitados: burbujeo de gases en lı´quidos newtonianos y no newtonianos*, I. Turbinas de 6 paletas y difusor plano, An Quim, **78**: 387-392, (1982).

Costa, M.I.S. and Meza, M.E.M., *Coexistence in a chemostat: Application of a threshold policy*, Chemical Engineering Science, **61(10)**: 3400-3402, (2006).

Cserjan-Puschmann, M., Kramer, W., Durrschmid, E., Striedner, G. and Bayer, K., *Metabolic approaches for the optimisation of recombinant fermentation processes*, Appl. Microb. Biotechnol., **53(1)**: 43-50, (1999).

Deepak, V., Kalishwaralal, K., Ramkumarpandian, S., Babu, S.V., Senthilkumar, S.R. and Sangiliyandi, G., *Optimization of media composition for Nattokinase production by Bacillus subtilis using response surface methodology*, Bioresour. Technol., **99(17)**: 8170-8174, (2008).

Deka, D., Bhargavi, P., Sharma, A., Goyal, D., Jawed, M. and Goyal, A., *Enhancement of cellulase activity from a new strain of Bacillus subtilis by medium development and analysis with various cellulosic substrates*, Enzyme Res., **2011**: 1-8, (2011).

Desai, K.M., Vaidya, B.K., Singhal, R.S. and Bhagwat, S.S., *Use of an artificial neural network in modelling yeast biomass and yield of β-glucan*, Proc. Biochem., **40(5)**: 1617-1626, (2005).

Dhabhai, R., Chaurasia, S.P. and Dalai, A.K., *Effect of Pretreatment Conditions on Structural Characteristics of Wheat Straw*, Chem. Eng. Comm., **200(9)**: 1251-1259, (2013).

Di Massimo, C., Montague, G.A., Willis, M.J., Tham, M.T. and Morris, A.J., *Towards improved penicillin fermentation via artificial neural network*, Computers Chem. Eng., **16(4)**: 283-291, (1992).

Fredrickson, A.G., Megee III R.D. and Tsuchiya, H.M., *Mathematical modelling for fermentation processes*, Adv. Appl. Microbiol., **13**: 419-470, (1970).

Friehs, K., *Plasmid copy number and plasmid stability*, Adv. Biochem. Eng. Biotechnol., **86**: 47-82, (2004).

Friehs, K., Schügerl, K., In *DECHEMA Biotechnology Conferences 4B*, VCH Weinheim, 700-705, (1990).

Ganusov, V.V., Bril'kov, A.V. and Pechurkin, N.S., *Mathematical Modeling of Population Dynamics of Unstable Plasmid-bearing Bacterial Strains under Continuous Cultivation in a Chemostat*, Biophysics, **45** (5): 881-887, (2000).

Ganusov, W. and Brilkov, A.V., *Estimating the Instability Parameters of Plasmid-Bearing Cells.* I. *Chemostat Culture*, J. Theor. Biol., **219(2)**: 193-205, (2002).

Garcia-Ochoa, F. and Castro, E.G., *Estimation of oxygen mass transfer coefficient in stirred tank reactors using artificial neural networks*, Enzyme Microb Technol., **28(6)**: 560-569, (2001).

Gilmour, S.G., Response surface designs for experiments in bioprocessing, Biometrics, **62(2)**: 323-331, (2006).

Glassey, J., Montague, G.A., Ward, A.C. and Kara, B.V., *Enhanced supervision of Recombinant E. coli Fermentations via Artificial Neural Networks*, Proc. Biochem., **29(5)**: 387-398, (1994).

Grabherr, R., Nilsson, E., Striedner, G. and Bayer, K., *Stabilizing plasmid copy number to improve recombinant protein production*, Biotechnol. Bioeng., **77(2):** 142-147, (2002).

Greener, A. L., *Cloning host organisms*, US Patent, **5(552)**: 314, (1996).

Guardia, M.J. and Calvo, E.G., *Modelling of Escherichia coli growth and acetate formation under different operational conditions*, Enz. Microb. Technol., **29(6-7)**: 449-455, (2001).

Guo, H. and Chen, L., *Periodic solution of a chemostat model with Monod growth rate and impulsive state feedback control*, J. Theo. Biol., **260(4)**: 502-509, (2009).

Gupta, N., Gundampati, R.K. and Debnath, M., *Enhancement of D-amino Acid Oxidase by Trigonopsis variabilis*, APCBEE Procedia Elsevier, **2**: 115-119, (2012).

Haider, M.A., Pakshirajan, K., Singh, A. and Chaudhry, S., *Artificial Neural Network-Genetic Algorithm Approach to Optimize Media Constituents for Enhancing Lipase Production by a Soil Microorganism*, Appl. Biochem. Biotechnol., **144(3)**: 225-235, (2008).

Harmand, J., Rapaport, A., Dochain, D. and Lobry, C., *Microbial ecology and bioprocess control: oportunities and challenges*, J. Proc. Cont., **18(9)**: 865-875, (2008).

Heinmets, F., *Analysis of Cellular Growth Process in Biomathematics*, Marcel Dekker, New York, **1**: 157-184, (1969).

Henson, M.A., *Dynamic Modelling of Microbial Cell Populations*, Curr. Opin. Biotechnol., **14(5)**: 460-467, (2003).

Horiuchi, J.I., Kikuchi, S., Kobayashi, M., Kanno, T. and Shimizu, T., *Modelling of pH response in continuous anaerobic acidogenesis by an artificial neural network*, Biochem. Eng. J., **9(3):** 199-204, (2001).

Hoskins, J.C. and Himmelblau, D.M., *Artificial neural network models of knowledge representation in chemical engineering*, Comp. Chem. Eng., **12(9-10)**: 881-890, (1988).

Hotta, K., *Process Dynamics*, Baifukan, Tokyo (in Japanese), (1975).

Hsu, S.B. and Waltman, P., *Model of the effect of anti-competitor toxins on plasmid-bearing, plasmid free competition*, Taiwanese J. Mathematics, **6(1)**: 135-155, (2000).

Hur, W. and Yoon-Keun, C., *An artificial Neural Network for Biomass Estimation from Automatic pH Control Signal*, Biotechnol. Bioproc. Eng., **11(4)**: 351-356, (2006).

Hyoung, L.S., Chul-Kim, II, Bae, K.H. and Byun, S.M., *Enhanced production and secretion of streptokinase into extracellular medium in Escherichia coli by removal of 13 N-terminal amino acids*, Biotechnol. Lett., **19(2)**: 151-154, (1997).

Imanaka, T., Aiba, S., *A perspective on the application of genetic engineering: Stability of recombinant plasmid*, Ann. N. Y. Acad. Sci., **369**: 1-14, (1981).

Karim, M.N. and Rivera, S.L., *Artificial neural networks in bioprocess state estimation*, Adv. Biochem. Eng. Biotechnol., **46**: 1-33, (1992).

Kiewiet, R., Kok, J., Seegers, J.F.M.L., Venema, G. and Bron, S., *The Mode of Replication is a Major Factor in segregational Plasmid Instability in Lacctococcus lactis*, Appl. Environ. Microbiol., **59(**2): 358-364, (1993).

Kim, C.H., Lee, J.Y., Kim, M.G., Song, K.B., Seo, J.W., Chung, B.H., Chang, S.J. and Rhee, S.K., *Fermentation strategy to enhance plasmid stability during the cultivation of Escherichia coli for the production of recombinant levansucrase*, J. Ferment. Bioeng., **86(4)**: 391-394, (1998).

Klykov, S.P. and Kurakov, V.V., *A New Kinetic Structured Model for Cell Cultivation in a Chemostat*, BioProc. Internat., **10(9)**: 36-43, (2012).

Kramer, W., Elmecker, G., Weik, R., Mattanovich, D. and Bayer, K., *Kinetics studies for the optimization of recombinant protein formation*. Ann N Y Acad Sci., **782**: 323-333, (1996).

Krone, S.M., Lu, R., Fox, R., Suzuki, H. and Top, E.M., *Modelling the Spatial Dynamics of Plasmid Transfer and Persistence*, Microbiol., **153(8)**: 2803-2816, (2007).

Kumar S.S. and Gupta, R., *An extracellular lipase from Trichosporon asahii MSR 54: Medium optimization and enantioselective deacetylation of phenyl ethyl acetate*, Proc. Biochem., **43(10)**: 1054-1060, (2009).

Kumar, P. and Ghosh, S., *Population dynamics model for plasmid bearing and plasmid lacking cells for streptokinase production in continuous flow stirred tank bioreactor*, Int. J. Eng. Sci. Technol., **2(5):** 118-127, (2010).

Kurtanjek, Z., *Modelling and control by artificial neural networks in biotechnology*, Computers Chem Engg., Supplement 1, **18**: S627-S631, (1994).

Lee, S.B. and Bailey, J.E., *Genetically structured models for lac promoter-operator function in the Escherichia coli chromosome and in multicopy plasmids: Lac operator function*, Biotechnol. Bioeng., **26(11)**: 1372-1382, (1984).

Lee, S.Y. and Papoutsakis, E.T., *Metabolic Engineering*, Marcel-Dekker, USA, **2**: 13-58, (1999).

Lek, S. and Guegan, J.F., *Artificial neural networks as a tool in ecological modelling, an introduction*, Ecol. Model., **120(2-3)**: 65-73, (1999).

Lenski, R.E. and Bouma, J.E., *Effects of Segregation and Selection on Stability of Plasmid pACYC184 in Escherichia coli B*, J. Bacteriol., **169(11):** 5314-5316, (1987).

Lenski, R.E. and Hattingh, S.E., *Coexistence of two competitors on one resource and one inhibitor: a chemostat model based on bacteria and antibiotics*, J. Theor. Biol., **122(1)**: 83-93, (1986).

Lessard, P. and Beck, M.B., *Dynamic modelling of the activated sludge process: a case study*, Water Res. **27(6)**: 963-978, (1993).

Lin-Chao, S., Chen, W.T. and Wong, T.T., *High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II*, Mol. Microbiol., **6**(22): 3385-3393, (1992).

Linko, P. and Zhu, Y., *Neural network programming in bioprocess variable estimation and state prediction*, J. Biotechnol., **21(3)**: 253-269, (1991).

Lu, Z. and Hadeler, K.P., *Model of plasmid-bearing, plasmid-free competition in the chemostat with nutrient recycling and an inhibitor*, Math. Biosc., **148(2)**: 147-159, (1998).

Malke, H. and Ferretti, J.J., *Streptokinase: Cloning, expression, and excretion by Escherichia coli*, Proc. Natl. Acad. Sci. USA., **81(11)**: 3557-3561, (1984).

Malke, H., Steiner, K., Gase, K. and Frank, C., *Expression and regulation of the streptokinase gene*, Methods, **21(2)**: 111-124, (2000).

Massimo, C.D., Willis, M.J., Montague, G.A., Tham, M.T. and Morris, A.J., *Bioprocess model building using artificial neural networks*, Bioproc. Eng., **7(1-2)**: 77-82, (1991).

Mata-Alvarez, J., *A dynamic simulation of two-phase anaerobic digestion system for solid wastes*, Biotechnol. Bioeng., **30(7)**: 844-851, (1987).

Molleta, R., Verrier, D. and Albagnac, G., *Dynamic modelling of anaerobic digestion*, Water Res. **20(4)**: 427-434, (1984).

Monaghan, R.L., Gagliardi, M.M. and Streicher, S.L., In: A.L. Demain, J.E. Davies (Eds.), *Culture preservation and inoculums development, Manual of Industrial Microbiology and Biotechnology*, ASM Press, Washington D.C., 29-48, (1999).

Montogomery, D.C., *Design and Analysis of Experiments*, (fifth ed.) John Wiley and Sons, NY, (2001).

Moser, D.R. and Campbell, J.L., *Characterization and complementation of pMB1 copy number mutant: effect of RNA I gene dosage on plasmid copy number and incompatibility*. J. Bacteriol., **154(2)**: 809-818, (1983).

Muller, A.K., Rojo, F. and Alonso, J.C., *The level of the pUB110 replication initiator protein is autoregulated, which provides an additional control for plasmid copy number*, Nucleic Acids Res., **23(11)**:1894-900, (1995).

Nicoletti, M.C., Jain, L.C. and Giordano, R.C., *Computational Intelligence Techniques as Tools for Bioprocess Modelling, Optimization, Supervision and Control*, Comp. Intel. Tech. Bio. Model. SCI, **218**: 1-23, (2009).

Nielsen, J., *Modelling the growth of filamentous fungi*, Adv. Biotchem. Engng., **46**: 187-223, (1992).

Nielsen, J., Pedersen, A.G., Strudsholm, K. and Villadsen, J., *Modelling fermetations with recombinant microorganisms: formulation of a structured model*, Biotechnol. Bioengng., **37(9)**: 802-808, (1991).

Noack, D., Roth, M., Geuther, R., Muller, G., Undisz, K., Hoffmeier, C. and Gaspar, S., *Maintenance and Genetic Stability of Vector Plasmids pBR322 and pBR325 in Escherichia coli K12 strains Grown in a Chemostat*, Mol. Gen. Genet., **184(1)**: 121-124, (1981).

Palsson, B.O. and Joshi, A., *On the Dynamic Order of Structured Escherichia coli Growth Models*, Biotechnol. Bioeng., **29(6)**: 789-792, (1987).

Palsson, B.O. and Lightfoot, E.N., *Mathematical modelling of dynamics and control in metabolic networks. I. On michaelis-menten kinetics*, J. Theor. Biol., **111(2)**: 273-302, (1984*).*

Papagianni, M., Boonpooh, Y., Mattey, M. and Kristiansen, B., *Substrate inhibition kinetics of Saccharomyces cerevisiae in fed-batch cultures operated at constant glucose and maltose concentration levels*, J. Ind. Microbiol. Biotechnol., **34(4)**: 301-309, (2007).

Patnaik P.R., *Neural control of an imperfectly mixed fed-batch bioreactor for recombinant β-galactosidase*, Biochem. Eng. J., **3(2):** 113-120, (1999).

Patnaik, P.R., *A heuristic approach to fed-batch optimization of streptokinase production*, Bioproc. Eng., **13(2)**: 109-112, (1995).

Patnaik, P.R., *An Evaluation of a Neural Network for the Start-up phase of a Continuous Recombinant Fermentation Subject to Disturbances*, Biotechnol. Techniq., **9(9):** 691-696, (1995).

Patnaik, P.R., *An Integrated hybrid neural system for noise filtering, simulation and control of a fed-batch recombinant fermentation*, Biochem. Eng. J., **15(3)**: 165-175, (2003).

Patnaik, P.R., *Neural and Hybrid Neural Modelling and Control of Fed-Batch Fermentation for Streptokinase: Comparative Evaluation under Non-Ideal Conditions*, Canadian J. Chem. Eng., **82**(3): 599-606, (2004).

Patnaik, P.R., *Neural Optimization of Fed-batch Streptokinase Fermentation in a Non-ideal Bioreactor*, Canadian J. Chem. Eng., **80(5):** 920-926, (2002).

Patnaik, P.R., *Towards a comprehensive description of microbial processes through mechanistic and intelligent approaches*, BMBR, **4**(2): 029-041, (2009).

Pavlou, S., *Dynamic of chemostat in which one microbial population grows on multiple complementary nutrients*, Biotech. Bioeng., **30(3)**: 413-419, (1987).

Pelizer, L.H., Dalva, E., Danesi, G., Rangel, C.D.O., Sassano, C.E.N., Carvalho, J.C.M., Sato, S. and Moraes, I.O., *Influence of inoculum age and concentration in Spirulina platensis cultivation*, J. Food Eng., **56(4)**: 371-375, (2003).

Phillips, G.J., *New cloning vectors with temperature-sensitive replication*, Plasmid, **41**(1): 78-81, (1999).

Pinsach, J., Mas, C.D., Lopez-Santin, J., Striedner, G. and Bayer, K., *Influence of process temperature on recombinant enzyme activity in Escherichia coli fed-batch Cultures*, Enz. Microbiol. Technol., **43**(7): 507-512, (2008).

Qian, Y. and Tessier, J.C., *Modelling of a woodchip refiner using artificial neural network*, Chem. Eng. Technol., **18(5)**: 337-342, (1995).

Ramalingam, S., Gautam, P., Mukherjee, K.J. and Jayaraman, G., *Effects of post-induction feed strategies on secretory production of recombinant streptokinase in Escherichia coli*, Bioch. Eng. J., **33(1)**: 34-41, (2007).

Ranganath, M., Renganathan, S. and Gokulnath, C., *Identification of bioprocess using Genetic Algorithm*, Bioproc. Eng., **21(2)**:123-127, (1999).

Ray, A. and Skurray, R., *Stabilization of the cloning vector pACYC184 by insertion of F plasmid leading region sequences*, Plasmid, **11(3)**: 272-275, (1984).

Razali, F., Young, M.M., Scharer, J.M. and Glick, B.R., *Review: overexpression of protein under transcriptional regulation of lambda pL promoter system in Escherichia coli: consequences and bioprocess improvement approaches*, J. Chem. Nat. Res. Eng., **1**: 22-39, (2007).

Reinikainen, P. and Virkajarvi, I., *Escherichia coli Growth and Plasmid Copy Numbers in Continuous Cultivations*, Biotechnol. Lett., **11**(4): 225-230, (1989).

Reuss, M., *Stirred tank bioreactors, In Bioreactor system design, Asenjo JA and Merchuk JC (Eds) Marcel Dekker, New York*, 207-255, (1995).

Rouf, S.A., *Evaluation of process simulators for an integrated bioprocess design*, Report Univ. Waterloo, Ontario, Canada, (1999).

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., *Learning representations by back propagating errors*, Nature, **323**: 533-536, (1986).

Saksela, O., *Radial Caseinolysis in Agarose: A Simple Method for Detection of Plasminogen Activator in the Presence of Inhibitory Substances and Serum*, Anal. Biochem., **111(2)**: 276-282, (1981).

Samanta, S.K., Rani, M. and Jain, R.K., *Segregation and structural instability of recombination plasmid carrying goes for naphthalene degrading pathway*, Lett. Appl. Microbiol., **26(2)**: 265-269, (1998).

Sardonini, C.A. and Dibiasio, D., *A Model for Growth of Sacchromyces cerevisiae Containing a Recombinant Plasmid in Selective Media*, Biotechnol. Bioeng., **29(4)**: 469-475, (1987).

Schmidt, T., Friehs, K. and Flaschel, E., *Rapid determination of plasmid copy number*, J. Biotechnol., **49(1-3)**: 219-229, (1996).

Schubert, J., Simutis, R., Dors, M., Havlik, I. and Lbbert, A., *Bioprocess Optimization and Control: Application of Hybrid Modeling*, J. Biotechnol., **35(1)**: 51-68, (1994).

Seneta, E. and Tavare, S., *Some stochastic models for plasmid copy number*, Theor. Popul. Biol., **23(2)**: 241-256, (1983).

Sevella, B. and Bertalan, G., *Development of a MATLAB based bioprocess simulation tool*, Bioproc. Eng., **23(6)**: 621-626, (2000).

Sharma, U.K., Insights into plasmid biology, Curr. Sci., **64(5)**: 283-285, (1993).

Shimizu, H., Yasuoka, K., Uchiyama, K. and Shioya, S., *On-line fault diagnosis for optimal rice α-amylase production process of a temperature-sensitive mutant of Saccharomyces cerevisiae by an autoassociative neural network*, J. Fermen. Bioeng., **83(5):** 435-442, (1997).

Shioya, S., Shimizu, K. and Yoshida, T., *Knowledge-Based design and Operation of Bioprocess Systems*, J. Biosc. Bioeng., **87(3):** 261-266, (1999).

Shuler, M.L. and Domach, M.M., *Mathematical Models of the Growth of Individual Cells in Foundations of Biochemical Engineering: Kinetics and Thermodynamics in Biological* Systems, American Chemical Society Publications, **207**: 93-133, (1982).

Silveston, P.L., Budman, H. and Jervis, E., *Forced modulation of biological processes: A review*, Chem. Eng. Sci., **63(20)**: 5089-5105, (2008).

Simutis, R. and Lubbert, A., *Exploratory Analysis of Bioprocesses Using Artificial Neural Network-Based Methods*, Biotechnol. Prog., **13(4)**: 479-487, (1997).

Sriraman, K. and Jayaraman, G., *Enhancement of recombinant streptokinase production in Lactococcus lactis by suppression of acid tolerance response*, Appl. Microbiol. Biotechnol., **72(6)**: 1202-1209, (2006).

Stabury, P.F., Whittaker, A. and Hall, S.J., *Principles of Fermentation Technology (Chapter 6)*, Second ed., Butterworth-Heinemann, Jordon Hill, Oxford: 147-164, (1995).

Stephanopoulos, G. and Lapidus, G.R., *Chemostat Dynamics of Plasmid-Bearing, Plasmid-Free Mixed Recombinant Cultures*, Chem. Eng. Sci., **43(1)**: 49-57, (1988).

Stephens, M.L., Christensen, C. and Lyberatos, G., *Plasmid Stabilization of an Escherichia coli Culture through Cycling*, Biotechnol. Prog., **8(1)**: 1-4, (1992).

Stuebner, K., Boschke, E., Wolfe, K.H. and Langer, J., *Kinetic analysis and modelling of streptokinase fermentation*, Acta Biotechnol., **11(5)**: 467-477, (1991).

Summers, D.K., Beton, C.W. and Withers, H.L., *Multicopy plasmid instability: the dimer catastrophe hypothesis*, Mol. Microbiol, **8(6):** 1031-1038, (1993).

Syamsu, K., Impoolsup. A., Greenfield, P.F. and Mitchell, D.A., *Data analysis of plasmid stability in continuous culture of recombinant Saccharomyces cerevisiae*, Biotechnology Techniques, **6(5)**: 393-398, (1992).

Szewczyk, K.W., *The effect of dilution rate variation on the performance of continuous fermentation*, Bioproc. Eng., **6(1-2)**: 17-19, (1991).

Thangadurai, C., Suthakaran, P., Barfal, P., Anandaraj, B., Pradhan, S.N., Boneya, H.K., Ramalingam, S. and Murugan, V., *Rare codon priority and its position specificity at the 50 of the gene modulates heterologous protein expression in Escherichia coli*, Bochem. Biophys. Res. Comm., **376(4)**: 647-652, (2008).

Theophilus, R.F. and Ferreira, M.M.C., *Chemometrics II: spreadsheets for calculations of experimental design, a tutorial*, New Chemistry, **29(2)**: 338-350, (2006).

Thibault, J., Van-Breusegen, V. and Cheruy, A., *On-line prediction of fermentation variables using neural network*, Biotechnol Bioeng, **36(10)**: 1041-1048, (1990).

Tholudur, A. and Ramirez, W.F., *Optimisation of fed-batch bioreactors using neural network parameter function models*, Biotechnol. Prog., **12(3)**: 302-309, (1996).

Thomposon, M.L. and Kramer, M.A., *Modeling chemical processes using prior knowledge and neural networks*, Proc. Sys. Eng., **40(8):** 1328-1340, (1994).

Tomida, S., Hanai, T., Ueda, N., Honda, H. and Kobayashi, T., *Construction of COD simulation model for activated sludge process by fuzzy neural network*, J. Biosci. Bioeng., **88(2:)** 215-220, (1999).

Ungar, L.H., Powel, B.A. and Kamens, S.N., *Adaptive networks for fault diagnosis and process control*, Computers Chem. Eng., **14(4-5)**: 561-572, (1990).

Van Can, H.J., Te Braake, H.A., Bijman, A., Hellinqa, C. and Luyben, K.C., *An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances*, Biotechnol. Bioeng., **54(6)**: 549-566, (1997).

Vellanki, R.N., Potumarthi, R. and Mangamoori, L.N., *Constitutive Expression and Optimization of Nutrients for Streptokinase Production by Pischia pastoris Using Statistical Methods*, Appl. Biochem. Biotechnol., **158(1):** 25-40, (2009).

Vieira, S., Hoffman, R., *Estatística Experimental*, Atlas, Sao Paulo, (1989).

Wang, J., Zhao, L. and YU, T., *On-line Estimation in Fed-batch Fermentation Process Using State Space Model and Unscented Kalman Filter*, Chinese J. Chem. Eng., **18**(2): 258-264, (2010).

Wang, Z., Le, G., Shi, Y., Wegrzyn, G. and Wrobel, B., *A Model for Regulation of ColE1-like Plasmid Replication by Uncharged tRNAs in Amino Acid–Starved Escherichia coli Cells*, Plasmid, **47(2)**: 69-78, (2002).

Willis, M.J., Montague, G.A., Di Massimo, C., Tham, M.T. and Morris, A.J., *Artificial neural networks in process estimation and* control, Automatica, **28(6)**: 1181-1187, (1992).

Wong, S.L., Ye, R. and Nathoo, S., *Engineering and Production of Streptokinase in a Bacillus subtilis Expression-Secretion system*, Appl. Environ. Microbiol., **60**(2): 517-523, (1994).

Xiong, Y.H., Liu J.Z., Song, H.Y. and Ji, L.N., *Enhanced production of extracellular ribonuclease from Aspergillus niger by optimization of culture conditions using response surface methodology*, Biochem. Eng. J., **21(1)**: 27-32, (2004).

Yazdani, S.S. and Mukherjee, K., *Continuous-culture studies on the stability and expression of recombinant streptokinase in Escherichia coli*, Bioproc. Biosyst. Eng., **24(6)**: 341-346, (2002).

Yazdani, S.S. and Mukherjee, K.J, *Overexpression of streptokinase using a fed-batch strategy*, Biotechnol. Lett., **20(10):** 923-927, (1998).

Yet-Pole, I., Wen-Tengu, W. and Yung-Chuan, L., *Neural network modelling for on-line state estimation in fed-batch culture of L-lysine production*, The Chem Engg. J, **61(1)**: 35-40, (1996).

Yuan, S. and Zhang, T., *Dynamics of a plasmid chemostat model with periodic nutrient input and delayed nutrient recycling*, Nonlinear Analysis: Real World Applications, **13(5)**: 2104-2119, (2012).

Zabriskie, D.W. and Arcuri, E.J., *Factors influencing productivity of fermentations employing recombinant microorganisms*, Enz. Microb. Technol., **8(12)**: 706-717, (1986).

Zhang, Q., Reid, J.F., Litchfield, J.B., Ren, J. and Chang, S.W., *A prototype neural network supervised control system for Bacillus thuringiensis fermentations*, Biotechnol. Bioeng., **43(6)**: 483-489, (1994).