ISSN: 2320-1983 Print ISSN: 2320-1991 Online Volume 1 Issue 1

Applied Cell Biology

Regular Paper

ACB, 1(1), 2012 [16-21]

Generic feedback model distributed coding scheme by using computational biology approach

Bin Li*, Bing Qi, Liangrui Tang

School of Electric and Electronic Engineering, North China Electric Power University, Beijing, 102206, (CHINA)

E-mail: direfish@163.com

Received: 7th October, 2012 ; Accepted: 23rd November, 2012

ABSTRACT

A biology cell calculation algorithm is proposed in this paper to facilitate distributed topology generation and the routing procedure of each repressor monomer. In order to analyze the detail procedure, a mathematic model is presented and the optimum value is calculated by differential operation, and the sample mean and deviation is also presented. Heuristic algorithm is introduced to accelerate the calculation speed, and the coding scheme is designed based on the principle of generic positive feedback control. Consider the random disturbances effect, modified mathematic model is constructed and the interference effect is analyzed. The mathematic numerical results are presented to illustrated the performance of the proposed approach. ! 2012 Trade Science Inc. - INDIA

KEYWORDS

Distributed coding;
Bioinformatics;
Generic;
Promoter;
Signal processing.

INTRODUCTION

With the rapid development of the individuals, DNA molecules can be orderly they carry the genetic information by codon / anticodon subsystem into protein molecules that perform a variety of physiological and biochemical functions during completed life[1]. On the contents of the entire discipline in terms of computational biology ultimately based on the phenomenon in the life sciences and the law as an object of study, to solve biological problems as the ultimate goal, mathematics and computer just problem-solving tools and means. A wide range of research areas of computational biology into almost every area of modern biological research[2]. With the official launch of the human genome project, the massive amounts of data rapidly accumulated, it was found that relying on the traditional method of storage to deal with the explosive growth of data has been powerless, began to learn the tools of informatics methods by this appears the emerging discipline of bioinformatics[3].

Bioinformatics is mainly focused on the biology, information collection, storage, analysis and processing and visualization, etc., and computational biology focused primarily on the use of mathematical models and computational simulation techniques to study biological problems. Life Sciences has developed to the post-genomic era, that people face are vast amounts of data sets. In most cases, computational biology research requires pre bioinformatics research process, it is usually bioinformatics as a sub discipline of computational biology to investigate[4]. The use of large-scale and efficient theoretical models and numerical computational biology to identify represent the protein coding

Regular Paper

region of the genome sequence to decipher hidden in the nucleic acid sequence of the genetic laws, which deducted directly from the protein sequence prediction of protein three-dimensional structure and dynamics characteristics, with the study of biological macromolecular structure-function relationships and the interaction between biological macromolecules, and also the interaction of biological macromolecules[5]. Genome genetic language information release and regulation of transcriptional profiling and protein profiling data the analog life information flow process, thus recognizing the metabolism together with the laws of evolution to explore various aspects of human health and disease, so that new perspective from the genome sciences the progress of the human genome project results into the medical field possible.

The DNA to protein process is known as gene expression and the regulation of this process is the central topic of research in molecular biology at this stage[6]. We propose a novel distributed network cooperative approach with biology principle which we named it biology cell calculation algorithm (BCCA). The regulation of gene expression is mainly manifested in the several aspects, i.e., transcriptional regulation, differential processing of RNA transcript, and differential translation of RNA. Our article is rather brief since our aim is to integrate known models for biological processes with numerical techniques to solve these models for simulation and design purposes, rather than to give a broad introduction to either system theory or modeling itself. The genetic information of all living organisms, are stored the form of a gene in the DNA (or RNA) molecules within the cell.

DISTRIBUTED MATHEMATIC MODEL OF NEW BIOLOGY CODING SCHEMES

Genetic factor is the material basis of heredity, the specific nucleotide sequence of the genetic information in DNA or RNA molecules. Gene by copying the genetic information is passed to the next generation, offspring with parental similar traits. The promoter initiated transcription of a particular gene and is located near the gene they transcribe on the same strand and upstream. As promoters are typically immediately adjacent to the gene in question, positions in the promoter are designated relative to the transcriptional start site, where transcription of DNA begins for a particular gene[7]. Positive feedback can generate bistable status as shown in Figure 1, and in the absence of the external trigger signal is always in the original steady state. External input trigger signal under bistable flip from one stable state to another stable state, has a wide range of applications in the automation and control.

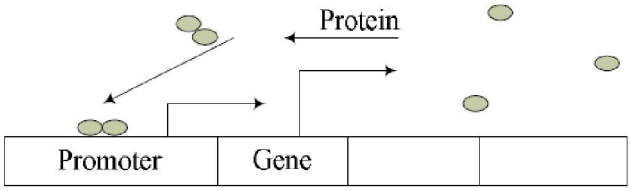


Figure 1 : Illustrations of positive feedback control

An open system which exchanges matter with the environment is usually changing with time until, if it is stable, it reaches a stationary state, or may be non-equilibrium state. According to above discussion, we propose a novel biological distributed coding approach[8]. The following coding diagram illustrated the detail procedure. The assumed per unit bandwidth of each communication link are, the network has two multicast request, and its source node, respectively, for the S1 and S2, the receiving node is D1 and D2, respectively. The network routing node network coding can be used for data forwarding. Can routing node U will receive packet conduct the XOR operation and send the combined results through UV, VD1, VD2 $b_1 \oplus b_2$., D1 and D2 of the two multicast request transmission rate can reach the maximum capacity, with their respective each uni-cast request to the sole use of the network reaches the same rate. Without the coding scheme, each request can only achieve a maximum rate of only 1.5 times uni-cast capacity.

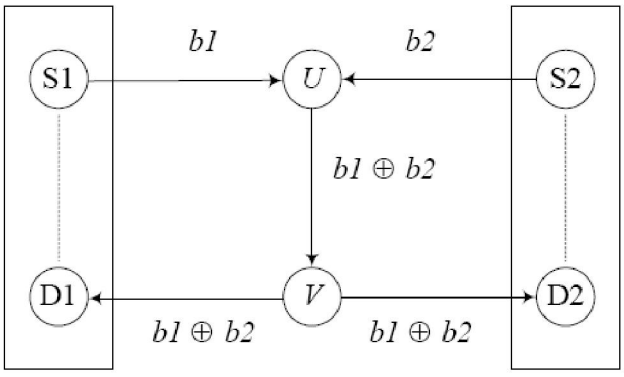


Figure 2 : Distributed encoding model with two-sources and two-sinks model

Regular Paper

The parameter value that lead to bistability shall be carefully investigated and additive external noise source will affect the repressor production[9]. Let x describe the repressor concentration within a colony of cells, in which the noise that shall be examined. Assume the transcriptional efficiency is α, σ_1 and σ_2 is the lower and upper bound of bistability.

Where α represents the inhibition of the relative ratio between the sub can improve the transcription rate and basal transcription rate, γ represents the ratio between the protein degradation rates and the rate of basal expression[3]. Let x represent the normalized number of repressor monomer, each cell type is stably inherited for many generations, and switching between the two types of cells occurs stochastically and rarely–roughly one switch in ten thousand cell divisions[10].

$$\frac{dx}{dt}=\frac{\alpha x^2}{1+(1+\sigma_1)x^2+\sigma_2 x^4}-\gamma x+1-\sigma\xi(t) \quad (1)$$

Where σ is an indicator of the perturbation strength, when α and γ changes, there is a balance status. Random disturbances affect the efficiency of gene transfer rate, the transcription efficiency α is random. The random fluctuations can generally be expressed as,

$$\frac{dx}{dt}=\frac{\alpha x^2}{1+(1+\sigma_1)x^2+\sigma_1 x^4}-\gamma x+1-\sigma\xi(t) \quad (2)$$

In which, ξ(t) fluctuated rapidly with zero expectation ($\overline{\xi}(t)=0$)

Based on the Euler theory, the differential equation can be expressed as,

$$\frac{dx(t)}{dt}=\lim_{\Delta t\to 0}\frac{x(t+\Delta t)-x(t)}{\Delta t} \quad (3)$$

Let h = Δt, t_k = t, t_k = t + h

Then we obtained,

$$x(t_{k+1})\approx x(t_k)+\frac{h\alpha x^2(t_k)}{1+(1+\sigma_1)x^2(t_k)+\sigma_2 x^4(t_k)}-h\gamma x(t_k)+h-\sigma h\xi(t_k) \quad (4)$$

Once we have an initial estimate, we update the parameter estimates using the data and the equations given below.

$$L(\theta)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}\exp\{-\frac{(x_i-\mu)^2}{2\sigma^2}\}=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right) \quad (5)$$

In order to achieve the maximum likelihood of the random samples quickly, we take logarithm to obtain,

$$\ln[L(\theta)]=\ln\left[\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\right]+\ln\left[\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right] \quad (6)$$

This simplified to,

$$\ln[L(\theta)]=-\frac{n}{2}\ln[2\pi]-\frac{n}{2}\ln[\sigma^2]-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad (7)$$

According to the Lagrange theorem, we take partial derivative of ln[L(θ)] with respect to expectation and standard deviation, then we get,

$$\frac{\partial}{\partial\mu}\ln L(\theta)=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \quad (8)$$

$$\frac{\partial}{\partial\sigma^2}\ln L(\theta)=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \quad (9)$$

Then the sample mean and deviation can be calculated by the estimator

$$\hat{\mu}=\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \quad (10)$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\frac{1}{n}\sum_{i=1}^{n}x_i)^2 \quad (11)$$

Therefore, we can get the expectation and variance of the sampling data[11]. It can be proved that it is an unbiased estimator for the mean value of the repressor production.

To implement the rapid calculation of the repressor production, we present an heurist algorithm as below:

Step 1: set initial value of x and get the original data input;

Step 2: sampling the input vector and discrete repressor concentration;

Step 3: calculating ϕ(*x*), and construct the difference coefficient;

$$\phi(x)=\frac{\alpha x^2(t)}{1+(1+\sigma_1)x^2(t)+\sigma_2 x^4(t)}-\gamma x(t)+1 \quad (12)$$

If there is not any random interference, particle will final stay at the equilibrium point (minimal energy) However,

Regular Paper

due to the interference of random noise, the particles can jump out of a balance point, to reach a point of balance this even if the random interference induced phase transition mechanism[12].

Step 4: Take estimation of the particle distribution function ξ(t), and substitute the expression into the following equation

$$\frac{dx}{dt} = \frac{\partial \phi(x)}{\partial x} - \sigma\xi(t) \quad \textbf{(13)}$$

Step 5: Keep continuous iteration until the calculation error is limited to the accepted level.

NUMERICAL RESULTS AND DISCUSSIONS

In the experiment, we assume that the state of a system is rigorously defined through the state variables of the system. For systems that have not reached their stationary state, the behavior with regards to time cannot be determined without knowing the initial conditions, or the values of the state variables at the start. The repressor production can be evaluated with different variance within [0,1], it can be seen from Figure 3 that for a deterministic variable, and the expectation is about 0.135. While the mean value will reduce to 0.125 under σ = 0.1 case, and 0.116 under σ = 0.2 case. There is an opposite trend appears, as long as the variance increase, the expectation decrease, i.e., 0.056 for unit variance. Hence, we can easily get the adaptive regulation for the network biology cell routing process, as the selection probability will automatically reduce with the increasing variance from statistical option.

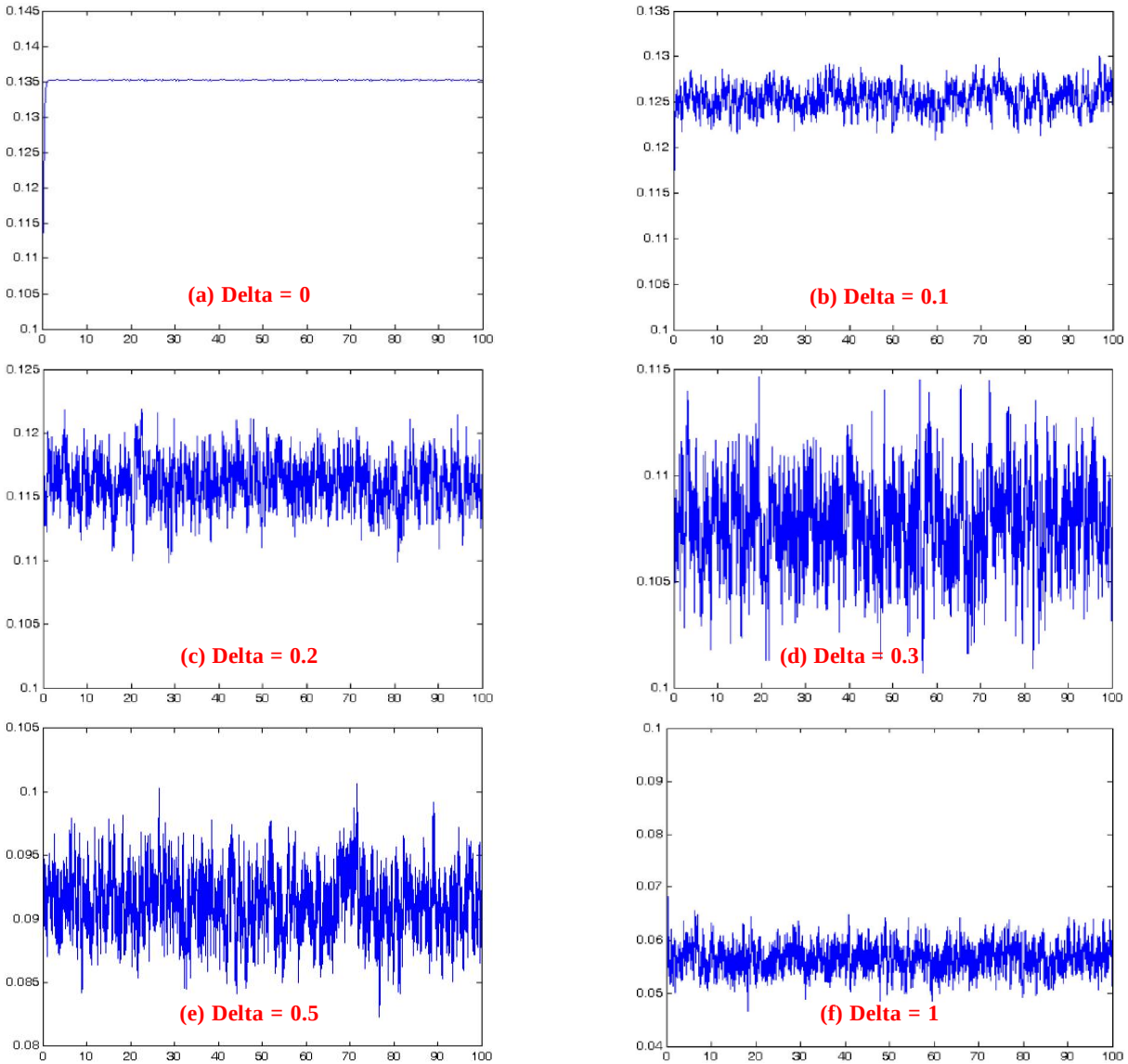


Figure 3 : Instantaneous value of the sampling value from statistical window

Regular Paper

To evaluate the efficiency of the sampling data, TABLE 1 present the results of confidence intervals for the parameter estimates with 95% confidence intervals. The period between lower bound and the upper bound is decreased from 2% to 0.4%. The mean value have two different trends compared with the initial value 0.01, it can be clear seen that the variance 0.3 is the boundary value

TABLE 1 : 95% confidence intervals evaluation

σ	Expectation	$Lower_{95\%}$	$Upper_{95\%}$
0	0.1349	0.1248	0.1462
0.1	0.1253	0.1231	0.1275
0.2	0.1162	0.1148	0.1176
0.3	0.1076	0.1067	0.1086
0.5	0.0916	0.0910	0.0922
1	0.0561	0.0559	0.0563

From the above demonstration topology, we can easily get the performance with the propose heuristic approach by using adaptive biology cell calculation algorithm. When the variance varies from 0 to 1, and deviation varies from -35% to 40%. According to the following graph at a very short distance from the start of the reactor, here we need to distinguish the parameters with reference model. Here the system in Figure 4 is designed with model reference adaptive system (MARS) for the non-linear timer varies system.

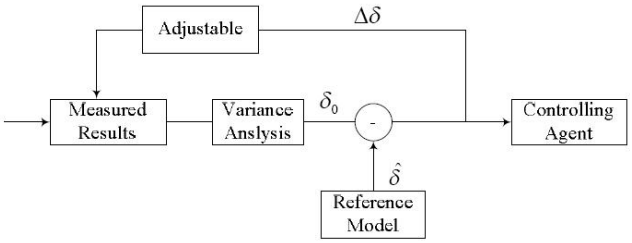
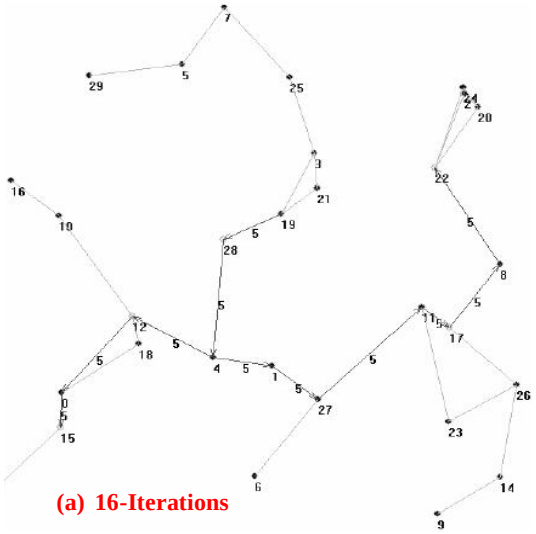
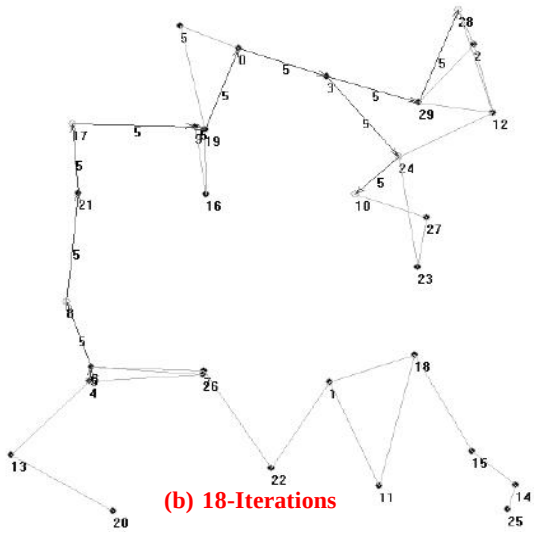


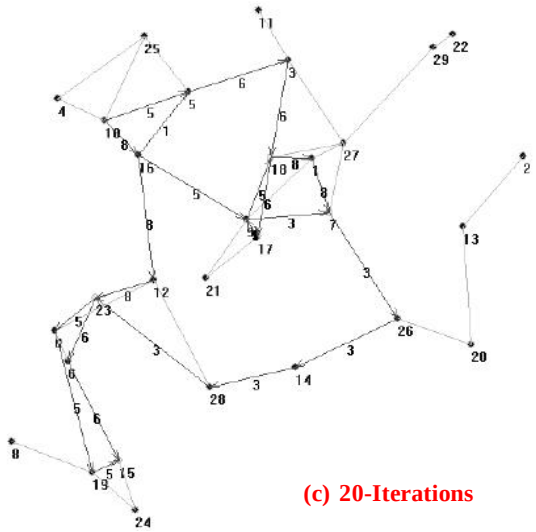
Figure 4 : Diagram of MRAS model with BCCA extensions



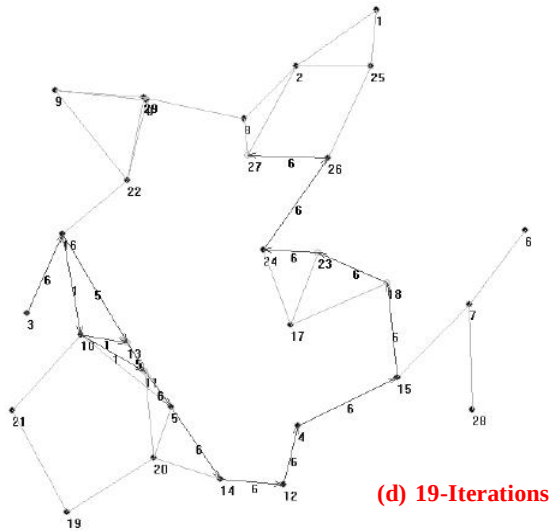
(a) 16-Iterations



(b) 18-Iterations



(c) 20-Iterations



(d) 19-Iterations

Figure 5 : Topology of each monocase distribution under different iterations

Regular Paper

TABLE 2 : Performance comparison between different approach with aspect of resource

	Network Capacity	BCCA -RES	DIJK -RES	PRIM -RES	Average Degree	Iteration Number
(a)	5.4286	11.837	11.98	13.262	3.095	16
(b)	5.7143	12.395	12.719	14.653	2.781	18
(c)	7.4	13.129	12.892	14.028	3.133	20
(d)	6.2	14.2	15.1	16.6	3.107	17

CONCLUSIONS

As soon as the biological process are turned into equations, the calculation efficiency shall be considered for practical use purpose. It is always usually solved numerically by using modern computer and large-scale tools. One central task for our proposed approach is to find the suitable mathematical model to describe the practical action in network and put them intelligently into different process through system theory. Various numerical solution techniques lends itself excellently to the system theory classification as well. To evaluate the complicated numerical codes, we often use heuristic algorithm instead large amount calculation and storage to accelerate the speed within the boundaries of the system, as well as the interaction between the system and its surrounding environment.

ACKNOWLEDGEMENT

This work was supported by "the Fundamental Research Funds for the Central Universities" (12QN10) and grants from the Major National Science and Technology Special Project (2010ZX03006-005-001)

REFERENCES

[1] Guy M.Goodwin, A.Richard Green; A behavioural and biochemical study in mice and rats of putative selective agonists and antagonists for 5-HT1 and 5-HT2 receptors, British Journal of Pharmacology, **84(3)**,19 Jul **(2012)**.

[2] Peter J.A.Cock, Tiago Antao, Jeffrey T.Chang; Biopython: Freely available python tools for computational molecular biology and bioinformatics, Bioinformatics, **25(11)**, 1422-1423 **(2009)**.

[3] P.Larranaga, B.Calvo, R.Santana, C.Bielza, J.Galdiano, I.Inza, J.A.Lozano, R.Armananzas, G.Santafe, A.Perez, A.Robles; Machine learning in bioinformatics, Briefings in Bioinformatics, **7(1)**, 86-112 **(2006)**.

[4] A.D.Lander, Q.Nie, F.Y.M.Wan; Membrane-associated non-receptors and morphogen gradients, Bull.Math.Biol., **69**, 33-54 **(2007)**.

[5] A.D.Lander, Q.Nie, F.Y.M.Wan; Do morphogen gradients arise by diffusion, Dev.Cell, **2**, 785-796 **(2002)**.

[6] D.T.Gillespie; The chemical Fangevin and Fokker-Planck equations for the reversible isomerization reaction, J.Phys.Chem.A, **106**, 5063-507 **(2002)**.

[7] F.M.Alhabdan, M.A.Abashar, S.Elnashaie; A flexible software package for industrial steam reformers and methanators based on rigorous heterogeneous models, Mathematical and Computer Modeling, **16**, 77-86 **(1992)**.

[8] A.E.Abasaeed, S.Elnashaie; A novel configuration for packed bed membrane fermentors for maximizing ethanol productivity and concentration, J.Membrane Science, **82(1-2)**, 75-82 **(1993)**.

[9] I.Jobses, G.Egberts, K.Luyben, J.A.Roels; Fermentation kinetics of zymomonas mobilis at high ethanol concentrations: Oscil lations in continuous cultures, Biotechnology and Bioengineering, **28**, 868-877 **(1986)**.

[10] A.Mahecha-Botero, P.Garhyan, S.Elnashaie; Bifurcation, stabilization, and ethanol productivity enhancement for a membrane fermentor, Mathematical and Computer Modelling, **41**, 391-406 **(2005)**.

[11] E.Wegman, D.Carr, Q.Luo; Visualizing multivariate data, in Multivariate Analysis: Future Directions, C.R.Rao, (Ed); The Netherlands: Elsevier Science Publishers, 423-466 **(1993)**.

[12] M.Wilk, R.Gnanadesikan; Probability plotting methods for the analysis of data, Biometrika, **55**, 1-17 **(1968)**.