Improved Differential Evolution Algorithm for Parameter Estimation to Improve the Production of Biochemical Pathway

Chuii Khim Chong, Mohd Saberi Mohamad , Safaai Deris, Mohd Shahir Shamsir, Yee Wen Choon, Lian En Chai

***Abstract* —This paper introduces an improved Differential Evolution algorithm (IDE) which aims at improving its performance in estimating the relevant parameters for metabolic pathway data to simulate glycolysis pathway for yeast. Metabolic pathway data are expected to be of significant help in the development of efficient tools in kinetic modeling and parameter estimation platforms. Many computation algorithms face obstacles due to the noisy data and difficulty of the system in estimating myriad of parameters, and require longer computational time to estimate the relevant parameters. The proposed algorithm (IDE) in this paper is a hybrid of a Differential Evolution algorithm (DE) and a Kalman Filter (KF). The outcome of IDE is proven to be superior than Genetic Algorithm (GA) and DE. The results of IDE from experiments show estimated optimal kinetic parameters values, shorter computation time and increased accuracy for simulated results compared with other estimation algorithms**

***Keywords*— Parameter Estimation, Differential Evolution Algorithm, Kalman Filter, Simulation.**

I. Introduction

The crucial step in the development of predictive models for cells or whole organisms is building dynamic models of biological systems. Such models can be regarded as the keystones of Systems Biology, ultimately providing scientific explanations of the biological phenomena [1]. Hence, one of the major challenges in the age of post-genomics is considered to be the understanding of dynamic metabolic behaviour of living cells [2]. Understanding of biological pathway's functions due to their complexity is difficult. Thus, not only we need to determine the components and their characteristics but also we need to focus on their continuous dynamic changes over time. One method to deal with this problem is to study the pathway as a network of biochemical reaction and subsequently model them as a system of ordinary differential equations (ODEs) [3, 4]. ODE based mathematical models can be implemented in various applications such as to simulate experiments before actual experiment is being performed, to study the phenomena that cannot be solved with experimentally, to aid in understanding the functions of a system etc. [5]. Design, analysis, optimization, and controlling of the biological system can be done with these ODEs. Different types of kinetic models such as Michaelis–Menten model or power law model are introduced with the purpose of studying the dynamic behavior of biological reaction systems [6]. Differential equations were used by scientists to simulate these dynamic changes in metabolic concentration but they require information which is related to the network structure and plethora of experimental data such as detailed kinetic rate laws, initial concentrations of metabolites and kinetic parameters [2]. Several models in metabolic networks modeling such as the threonine synthesis pathway in *Escherichia coli* have been developed by researchers [7].

The expert's proposition on dynamic model, how it is later fitted to the data, and how changes are taken into considerations if the predictions were not good enough are the process of modelling. Estimation of the parameters' value in the mathematical models for biochemical networks is typically done through minimization means [8]. Simulated result retrieved from simulation of the mathematical model with the aims to compare model results with the experimental data is called the forward problem. The inverse problem, on the other hand is the process where estimation of parameters of a mathematical model is done based on the measured observations [5]. This step is called parameter estimation and is one of the essential parts of model building. Without identifying the model parameters that define the data can cause inaccuracy in the conclusion [9]. Only some of these parameters in the model can be retrieved from experiments or

Manuscript received June 10, 2012. This work is financed by Institutional Scholarship MyPhd provided by the Ministry of Higher Education of Malaysia and Universiti Teknologi Malaysia with UTM GUP research grants (vot number: QJ130000.7123.00H67 and QJ130000.7107.01H29).

C. K. Chong is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (e-mail: ckchong2@live.utm.my)

M. S. Mohamad is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (corresponding author to provide phone: +60-7-553-3153; fax: +60-7-556-5044; e-mail: saberi@utm.my).

S. Deris Author is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (e-mail: safaai@utm.my).

M. S. Shamsir is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (e-mail: shahir@fbb.utm.my)

Y. W. Choon is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (e-mail: ywchoon2@live.utm.my).

L. E. Chai is with the Universiti Teknologi Malaysia, Skudai, 81310 Malaysia (e-mail: lechai2@live.utm.my).

from the previous works that have been done by other researchers and others have to be retrieved by comparing model results with experiments data [5]. Gathering data via experiments on genomic, proteomic, and metabolomic scales are growing generally in biological sciences. An accurate model building methods which can handle the high complexity is highly needed when the quality and the size of experimental data continue to grow rapidly [1]. Nevertheless, when the available data is noisy and sparse, i.e. widely and unevenly spaced in time, as is generally when measuring biological quantities at the cellular level makes the parameter estimation problem even more difficult to solve [10]. Noisy data can also occur when the collected results differ from each other and this is caused by the human error or apparatus limitation.

Parameter estimation (also known as model calibration) aims at finding the parameters of the parameters' value which give the best fit to a set of experimental data [1]. Biological data usually are nonlinear and dynamic. This problem is considered as a nonlinear programming (NLP) problem which generally known to be non-trivial and multimodal. Hence, traditional approach such as gradient-based or local optimization methods fail to provide optimal solutions. In order to overcome this limitation several state-of-the-art deterministic and stochastic global optimization methods are used by many researches [11]. The subsequent session is the explanation of few methods which include basic estimation approach and evolutionary algorithms.

In 1965, The Nelder-Mead algorithm (NM), also known as non-linear simplex method [12], is one of the best known algorithms for multidimensional unconstrained optimization without the need of derivatives information, which makes it appropriate for problems with non-smooth functions. NM is commonly used to solve parameter estimation problem which the function values are uncertain or in the cases where noise exists. It can also be implemented in problems with discontinuous functions which often occur in statistics and experimental mathematics. NM is very effective, particularly with a large number of parameters [13]. As a limitation of NM, where information regarding the convergence is very constrained and many of the iterations can run without a significant decrease of function values while the current results are still far from the optimal result. Besides that, the location of the initial seed for NM may affect convergence of the algorithm in the case of a function with more than one minimum.

Simulated annealing (SA) is another method which aimed at finding a better approximation to the global optimum in a large search space of a given function. SA is a generic probabilistic and metaheuristic approach and is implemented where the search space is discrete. One of the benefits of SA is its capability of not getting stuck in the local minima and the convergence is guaranteed in case of existence of large number of iterations [14, 15]. In addition, choosing the initial temperature or cooling schedule is challenging in SA. Furthermore, waste of computation time result by using too high temperature and using too low temperature would cause the reduction of quality of the search [14] and as a result, solving a complex system problem becomes very slow and uses more processor time [16]. Richard and his colleagues (2007) did use SA to estimate the relevant kinetic parameter in solving biochemical nonlinear parameter estimation problem. [17].

Genetic Algorithm (GA) is a subclass of evolutionary algorithms which is based on inheritance, mutation, selection, and crossover. Many scholars and researchers like Katera et al., 2004, and Donaldson and Gilbert (2008) used this algorithm to solve parameter estimation problem [9, 18]. The advantages of GA are its parallel search and searching efficiency [19] whereas finding local minima which may not be a true solution is considered as a disadvantage of genetic algorithm [20].

As a parallel search method, the Differential Evolution algorithm (DE) optimizes a problem by repeatedly trying to enhance a candidate solution with the goal of achieving the defined measure of quality. It is generally categorized as metaheuristic approach due to the fact that it works on no assumptions regarding the problem being optimized and can deal with substantial spaces of candidate solutions. The advantages of DE are considered to be high speed, efficiency, simplicity, and ease of use [21]. It was implemented by Moonchai Sompop *et al*. (2005) to enhance the production of bacteriocin, aspartate, beer, and cell process simulation by utilizing control and kinetic parameters [22]. DE shows to be very sensitive to control parameters: crossover constant (*CR*), population size (*NP*), and mutation factor (*F*) [23].

We proposed an improved Differential Evolution algorithm (IDE), a hybrid of DE and the Kalman Filter (KF), to solve the problems regarding the existence of noisy data that leads to low accuracy for estimated result and the increasing number of unidentified parameters which results in adding to the difficulties of the model in estimating the kinetic parameters. DE which is a stochastic-based approach, proved to be the best optimization algorithm out of the others. Stochastic-based approach is more appropriate to implement in the biological data in which they are usually non-convex and are easily trapped in local minimal [24]. Parameter estimation with DE is done without noisy data handling process. IDE takes advantage of KF which adds the feedback gathering feature from the noisy measurement to improve the performance of each output that was resulted by DE which provides higher accurate results. Biochemical pathways are regulatory pathway, signalling pathway, and metabolic pathway. Cell cycle pathway and aspartate biosynthesis pathway are the metabolic pathways which are the series of events that happened in a cell causing its division and duplication (replication) and synthesis aspartate, the essential amino acid. These are the symbolic pathways that are studied in this paper.

II. PROPOSED ALGORITHM

A. Experiment Setup

This paper proposes a hybrid of DE [25] and KF [26], which is an improved differential evolution algorithm (IDE). In parameter estimation, existing algorithms [22] merely implement DE whereas IDE implements a hybrid of DE and KF. Fig. 1 shows the details of the IDE. Kinetic parameters existed in the glycolysis pathway model for yeast [27] and Novak Tyson Cell Cycle in frog egg cell [28] go through IDE to estimate its optimal value. Fixed control parameter values used in this study are

i. population size, NP = 10,
ii. mutation factor, F = 0.5,
iii. crossover constant, CR = 0.9.

SBToolbox in Matlab 2008a and Copasi are the two main software implemented in this study. The mentioned metabolic pathways were collected from online database called Biomodel which is sustained by European Bioinformatics Institute (EMBL-EBI).

B. Improved Differential Evolution Algorithm (IDE)

In IDE, we added the process of updating the population as a new step that improved the conventional DE. This is a self-adapt approach. In conventional DE, the original population which is an m x n population matrix, is generated from the first generation (*Gen_1*) and continues until it reaches the maximum generation (*Gen_i*) in initialization process. m represents the number of generations and n represents the number of identifiable parameters. In evaluation process, the fitness function, J represented as

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

is applied to evaluate the fitness of each individual. X represents the state vector for measurement system, Y represents the state vector for simulated system, $\varnothing 0$ represents a set of original parameters, $\varnothing$ represents a set of estimated parameters, $X0$ represents the initial state, N=the ending index, and i=the index variable.

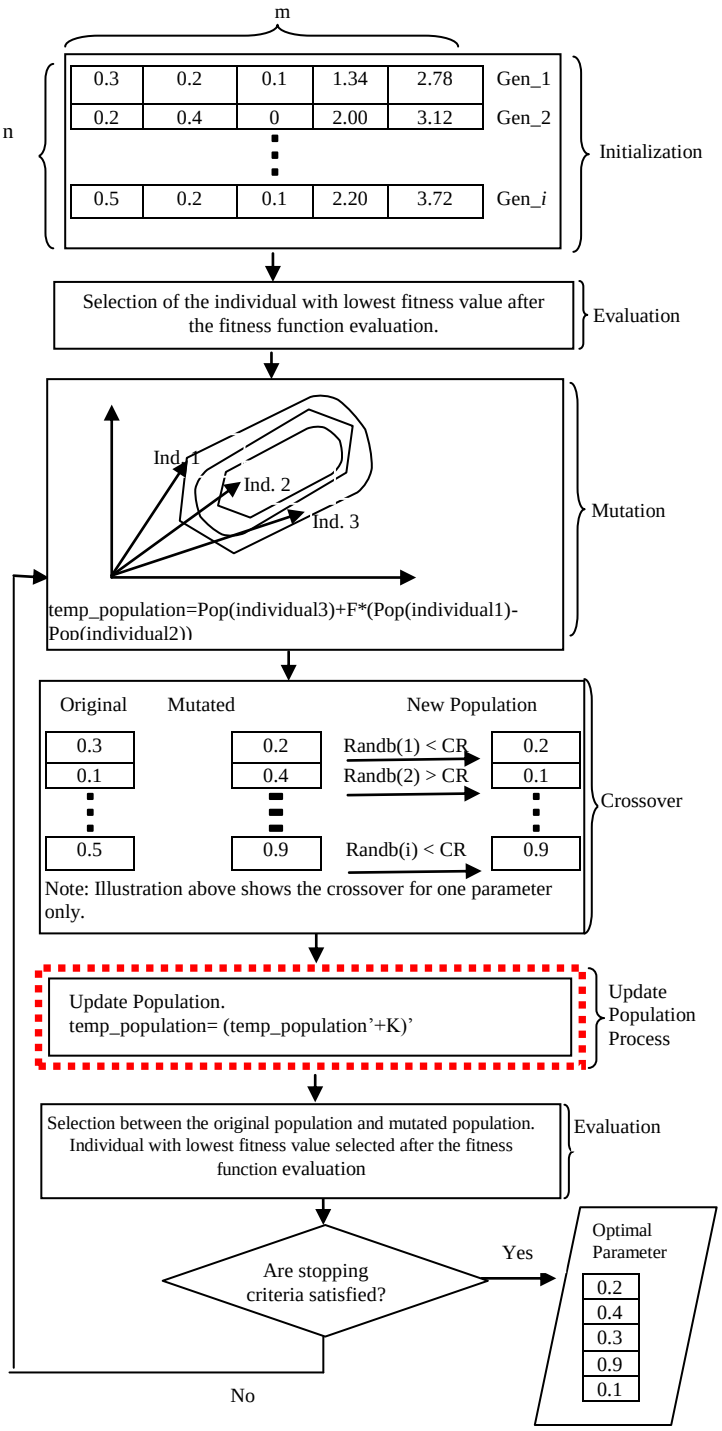
In mutation process, three individuals (*Ind1*, *Ind2* and *Ind3*) first being selected then treated with the formula showed in Fig 1. In the mutation section, *temp_population* represents the mutated population matrix, F represents the mutation factor, and *Pop* represents the original population matrix. The subsequent crossover process is mainly performed based on CR, which indicates crossover constant value, and *Randb(i)* which indicates i-th random evaluation of a uniform random number generator [0,1]. If the *randb(i)* value of the individual in mutated population is lower than the CR value then that individual becomes the individual for the resultant population of the crossover process and vice versa. This is followed by the updating process that is performed according to the Equation 2. This step updates the population, which is generated by the crossover process and it is based on the Kalman gain value K, retrieved from the Equation 3. The Kalman gain value from the Equation 3 takes into account the process noise covariance and measurement noise covariance. These noisy data values were obtained from the experiment and in this study the noisy data values used are 0.1. After handling the noisy data, the updated population once again undergoes the evaluation process and the whole process is repeated till the stopping criterion is met. The stopping criteria are set via predefined maximum loop values or when the fitness functions have converged. The updating population process is highlighted with the dotted box in Fig. 1 and is carried out according to the following formula.

$$temp\_population = (temp\_population' + K)' \quad (2)$$

$$K = P * H' * inv(H * P * H' + R) \quad (3)$$

Table 1
Pseudocode for IDE

Algorithm: IDE
BEGIN
STEP 1: Initialize population P based on D and evaluate it.
  WHILE (k<Max)
    FOR (i = 0 ; i < NP ; i++)
      STEP 1.1: Initialization
        Randomly select parents P [i_1], P [i_2], and P [i_3] where i, i_1, i_2, and i_3 are different.
      STEP 1.2: Mutation
        Create initial candidate C_1[i] = P [i_1] + F *(P [i_2] - P [i_3]).
      STEP 1.3: Crossover
        Create final candidate C[i] by crossing over the genes of P [i] and C_1[i] as follows:
        FOR (j = 0 ; j < NP ; j++)
          IF (U(0, 1) < CR)
            C[i][j] = C_1[i][j]
          ELSE
            C[i][j] = P [i][j]
        END-FOR
      STEP 1.4: Updating Population
        C[i] = inv(inv(C[i]) + K)
        K= P*H'*inv(H*P*H'+R)
      STEP 1.5: Evaluate C[i]
        IF (C[i] is better than P [i])
          P'[i] = C[i]
        ELSE
          P'[i] = P [i]
        END-IF
    END-FOR
    P = P'
  END-WHILE
END



Note: Updating population process is added after the crossover process to improve DE performance and it is highlighted with the dotted box.

Fig. 1. Schematic Overview of IDE.

Where
K = Kalman gain value,
H = observation matrix,
Q = process noise covariance,
D = number of the unknown parameters,
R = measurement noise covariance,
B = covariance of the state vector estimate,
H' = inverse of matrix H,
P = population of the current generation,
P' = the population to be formed for the next generation,
$C[i]$ = the candidate solution with population index i,
$C[i][j]$ = the j'th entry in the solution vector of $C[i]$,
N = the problem dimensionality,
$U(0, 1)$ = a uniformly distributed number between 0 and 1,
k = the scaling factor,
inv = the inverse function,
Max = maximum generation.

III. EXPERIMENTAL RESULT

Three estimation algorithms (GA, DE, and IDE) are compared in this study. Kinetic parameter values in Table 1and Table 2 are produced by the estimation algorithms and collected from literature review [27, 28]. Time series data for concentration of adenosine monophosphate (AMP) and Clycin were generated in order to evaluate the accuracy of each estimation algorithm. AMP and Clycin are significant metabolites. AMP acts as an energy regulator and sensor while Cyclin acts as a regulator for cell cycle. From the time series data, we calculate the average of error rate. The details of the accuracy measurement are discussed in this session.

Table 1.
Kinetic parameter values of IDE compared with GA and DE.

	Measurement kinetic parameter values[27]	**Simulated kinetic parameter values**		
Kinetic parameters		**GA**	**DE**	**IDE**
k9f	10	26.57	1.12	2.21
k9b	10	6.184	54.37	10.15

Note: Table shows the kinetic parameter values used in the calculation of average of error rate for metabolite AMP in Table 3.

Table 2
Kinetic parameter values of IDE compared with GA and DE.

Kinetic parameters	**Measurement kinetic parameter values [28]**	**Simulated kinetic parameter values**		
		GA	**DE**	**IDE**
k1	0.01	0.026	0.028	0.0102
k3	0.500	0.140	2.028	0.602
V2p	0.005	0	0.01	0.018
V2pp	0.250	0.069	0.658	0.347

Note: Table shows the kinetic parameter values used in the calculation of average of error rate for metabolite Cyclin in Table 4.

The simulated kinetic parameter values and measurement kinetic parameter values were replaced into the ordinary differential equations (ODEs) (Equation 4 and Equation 5) of AMP and Cyclin respectively.

$$\frac{dAMP}{dt} = -AMPflow - reaction9 \quad (4)$$

$$\frac{dCyclin}{dt} = R1 - R2 - R3 \quad (5)$$

Where
reaction_9= compartment * (k9f * *amp* * *atp* - k9b * power(*adp*,2)),
AMPflow=compartment * *amp* * *flow*,
compartment=constant value of 1,
amp=concentration of AMP,
pyr=concentration for PYR,
adp=concentration for adenosine diphosphate,
atp=concentration of adenosine triphosphate,
R1=+k1,
*R2=+k2*CYCLIN*,
R3=k3 CYCLIN*,
*k2=V2p+apcstar**(*V2pp-V2p*)
CYCLIN=concentration for cyclin,
apcstar=concentration of anaphase-promoting complex.

Time series data for concentration of AMP and Cyclin were ultimately produced from Equation 4 and Equation 5. The time series data contain measurement result, *y*, and simulated results *yi* for IDE, DE, and GA respectively. Error rate (*e*) and Average of error rate (*A*) are calculated according to Equation 6 and Equation 7 respectively.

$$e = \sum_{i=1}^{N} (y - yi)^2 \quad (6)$$

$$A = \frac{e}{N} \quad (7)$$

Table 3 and Table 4 show the average of error rate for AMP and Cyclin respectively.

Table 3.
Average of error rate for AMP.

Evaluation criteria	**GA**	**DE**	**IDE**
Average of error rate, *A*	0.000248	0.059148	0.000010

Note: Shaded column represents the best results.

Table 4.
Average of error rate for Cyclin.

Evaluation criteria	**GA**	**DE**	**IDE**
Average of error rate, A	1.156E-05	1.338E-05	0.001E-05

Note: Shaded column represents the best results.

For AMP (Table 3), IDE showed the lowest average of error rate with 0.000010. DE showed the worst performance with 0.059148 for the average of error rate. GA showed more moderate performance with average of error rate of 0.000248. However, for Cyclin (Table 4), IDE once again performed better than other estimation algorithms where average of error rate is 0.001E-05. The average of error rate for DE and GA are 1.338E-05 and 1.156 E-05 respectively. Lower average of error rate denotes that the simulated results are close to the measurement results and this shows the ability of Kalman filter to handle noisy data makes the IDE robust to noisy data.

Table 5 shows execution time of each estimation algorithm on a Core i5 PC with 4GB main memory. The result shows that DE required the longest time (6 minutes and 1 second and 9 minutes and 30 seconds) to find the optimal value for all kinetic parameters compared to IDE which took the shortest time (5 minutes and 35 seconds and 6 minutes 55 seconds). It is shown that IDE tends to use less computation time than DE and GA for glycolysis pathway and Novak Tyson Cell Cycle respectively.

Table 5.
Execution time of IDE compared with GA and DE.

Execution time (hh:mm:ss)	GA	DE	IDE
glycolysis pathway	00:05:42	00:06:01	00:05:35
Novak Tyson Cell Cycle	00:07:12	00:09:30	00:06:55

Note: Shaded column represents the best results.

Figure 2 shows the metabolite production graphs for the metabolites AMP and Cyclin based on the kinetic parameters that are collected from previous works [27, 28] and produced by IDE. The results showed that the kinetic parameters generated by IDE, enhanced the production rate where the dotted simulated lines (generated with the kinetics parameters that resulted by IDE) are moved to left when compared to the measurement lines (generated with the kinetics parameters that retrieved from experimental work).

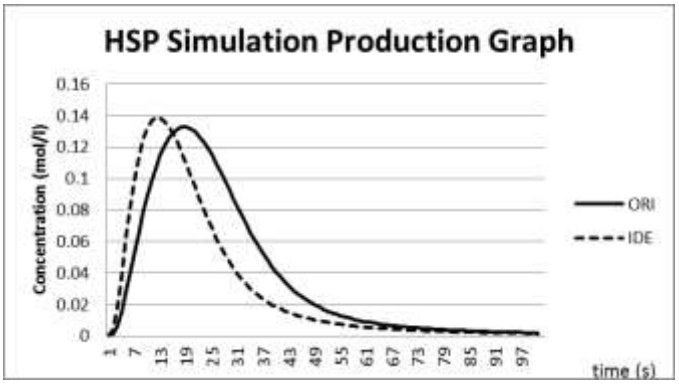


Fig. 2 (a). Production graph for metabolite HSP (ORI generated with the kinetic parameters that retrieved from experimental work and IDE generated with the kinetic parameters that was produced by IDE)

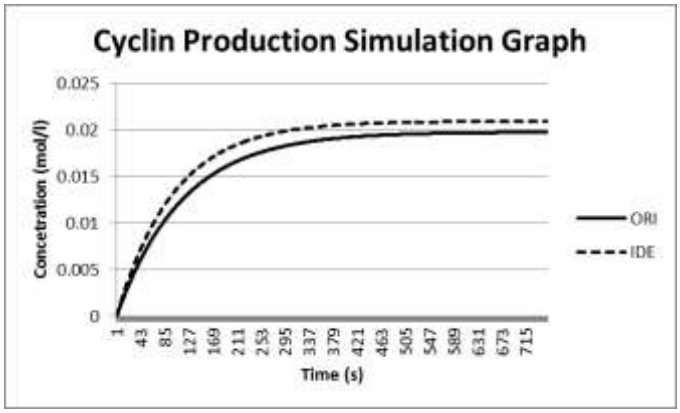


Fig. 2(b). Production graph for metabolite HSP (ORI generated with the kinetic parameters that retrieved from experimental work and IDE generated with the kinetic parameters that was produced by IDE)

Mean (*mu*) and standard deviation (STD) values are calculated according to the equation below.

$$mu = \frac{\sum_{i=1}^{N} e}{N} \tag{8}$$

$$S = \frac{\sum_{i=1}^{N} ((y - yi)^2 - mu)^2}{N - 1} \tag{9}$$

Table 6 shows the mean and STD values of fitness value for glycolysis pathway, and theronine biosynthesis pathway for 50 runs respectively. Fitness function implemented in this study is to minimize the difference between measurement results and simulated results. Based on the result from the table, STD values for metabolites AMP and Cyclin are 0.0992 and 0.0182. However, the mean for metabolites AMP and Cyclin are 0.0453 and 0.0027. The standard deviation is a measure of how widely values are scattered from the average value (the mean). The mean and STD values are close to 0 and this shows that results produced by IDE are consistent with low error rate. Other than that, it can also be analyzed that in the 50 runs simulation, the differences between each run are small as the STD values showed are close to the mean values which is close to 0. This deduces that IDE is a stable and reliable algorithm.

Table 6
Mean and standard deviation (STD) values of fitness value for glycolysis pathway, and Novak Tyson Cell Cycle for 50 runs.

	AMP	Cyclin
Mean	0.0453	0.0027
STD	0.0992	0.0182

According to Lillacci and Khammash (2010), to ensure that the final estimates are guaranteed to be statistically consistent with the measurements, chi-square test (X^2 test) as a statistical test is implemented. The degrees of freedom, *s* and confidence coefficient, γ implemented in this paper are 1 and 0.995. Interval estimates, σ^2 formed based on *s*, γ, and the formula found in Lillacci and Khammash (2010) is $0.0000393 < \sigma^2 < 9.550$. The hypothesis made here is that the simulated results are statistically consistent with the measurement results. X^2 value for metabolite HSP is 0.028956054 and metabolite Cyclin is 0.0000563 where both are appeared to be in between σ^2. Therefore, IDE passed the X^2 test, hypothesis accepted and the simulated results are proved to be statistically consistent with the measurement results.

IDE exhibits lesser computation time and possesses a higher accuracy when compared to both GA and DE. The implementation of DE that aims to estimate the relevant kinetic parameters and the additional of Kalman gain value which targets to handle the noisy data has improved the computational time and accuracy. Hence, the IDE, a stable and reliable estimation algorithm, which is a hybrid of DE and KF minimizes the computational time and also increases the accuracy between the simulated results and measurement results.

IV. Conclusion

In this paper, the experiment to compare the performances of three different estimation algorithms using glycolysis pathway data in yeast [27] and Novak Tyson Cell Cycle in frog egg cell [28] showed that an improved algorithm, IDE which is a hybrid algorithm of DE and KF performed the best with the shortest execution time and the lowest average of error rate. It successfully minimizes the high difficulty of the system in estimating the relevant kinetic parameters resulting in shorter computation time. The ability to handle noisy data has contributed to an improved accuracy of the estimated results. Besides that, IDE shows that it is a stable and reliable estimation algorithm by passing the chi square test (X^2 test) and showing the mean and STD value closer to 0 with 50 runs. In conclusion, IDE, a reliable algorithm is shown to be superior compared to both GA and DE in terms of computational time and accuracy. IDE can be generalized where it can be implemented in the areas which its data consists of noisy for example electrical and electronic engineering field [29].

DE shows to be very delicate to control parameters: population size (*NP*), crossover constant (*CR*), and mutation factor (*F*) [23]. Thus, for future work, self-adapting approach to these control parameters can be implemented to enhance the performance of the IDE. Moreover, additional steps can be added to the process of generating new populations with the aim of improving the performance of IDE.

Acknowledgment

Here we would like to take this opportunity to express our gratitude and appreciation to all the people who have given their heart whelming full support in making this paper a magnificent experience. To God the father of all, we thank him for the strength and wisdom that keeps us standing and for the

hope that keeps us believing that this work would be possible and more interesting. We also wanted to thank our family who inspired, encouraged and fully supported us for every trial that comes our way. To our colleagues who helped us ideas that needed for this work.

REFERENCES

[1] M. R. Fernandez, P. Mendes, and J. R. Banga, "A hybrid approach for efficient and robust parameter estimation in biochemical pathways," Biosystems, vol. 83, no. 2-3, 2006, pp. 248-265.

[2] R. S. Costa, D. Machado, I. Rocha, and E. C. Ferreira, "Hybrid dynamic modeling of Escherichia coli central metabolic network combining Michaelis–Menten and approximate kinetic equations," BioSystems, vol. 100, 2010, pp. 150-157.

[3] G. Koh, D. Hsu, and P. S. Thiagarajan, "Component-based construction of bio-pathway models: The parameter estimation problem," Theoretical Computer Science, vol. 412, 2011, pp. 2840-2853.

[4] P. Lecca, "Simulating the cellular passive transport of glucose using a time-dependent extension of Gillespie algorithm for stochastic π-calculus," Int. J. of Data Mining and Bioinformatics, vol. 4, 2007, pp.315-336.

[5] M. Ashyraliyev, Y. F. Nanfack, J. A. Kaandorp, and J. G. Blom, "Systems biology: parameter estimation for biochemical models," FEBS J. vol 276 no. 4, Feb 2009, pp.886-902.

[6] P. K. Liu, and F. S. Wang, "Hybrid differential evolution with geometric mean mutation in parameter estimation of bioreaction systems with large parameter search space," Computers and Chemical Engineering, Vol. 33, 200, pp. 1851–1860.

[7] C. Chassagnole, D. A. Fell, B. Rais, B. Kudla, and J. P. Mazat, "Control of the threonine-synthesis pathway in Escherichia coli: a theoretical and experimental approach," Biochem J., Vol. 356, No. 2, 2001, pp. 433-444.

[8] O. R. Gonzalez, C K, K. Jung, P. C. Naval Jr, and E. Mendoza, "Parameter estimation using Simulated Annealing for S-system models of biochemical networks," Bioinformatics, vol. 23, no. 4, 2007, pp. 480-486.

[9] S. Katare, A. Bhan, J. M. Caruthers, W. N. Delgass, and V. Venkatasubramanian, "A hybrid genetic algorithm for efficient parameter estimation of large kinetic models," Computers and Chemical Engineering, vol. 28, 2004, pp. 2569–2581.

[10] J. R. Porter, J. S. Burg, P. J. Espenshade, and P. A. Iglesias, "Identifying static nonlinear structure in a biological system using noisy, sparse data," Journal of Theoretical Biology, in press.

[11] C. G. Moles, P. Mendes, and J. R. Banga, "Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods," Genome Res., vol. 13, 2006, pp. 2467-2474.

[12] J. A. Nelder, and R. Mead, "A Simplex Method for Function Minimization, " The Computer Journal, vol. 7, no. 4, 1965, pp. 308-313.

[13] J. L. M. González, Parameter Optimisation and Adjustment by the Simplex Method (Nelder-Mead). Valladolid, Spain: Universidad de Valladolid, 2001.

[14] Back, T. (2009) Simulated Annealing [online]. LIACS Natural Computing Group Leiden University. Available from: http://natcomp.liacs.nl/NC/slides/sa.pdf.

[15] A.T. Kalai, and S. Vempala, "Simulated Annealing for Convex Optimization," Mathematics of Operations Research, vol. 31 no. 2, 2006, pp. 1–17.

[16] D. Handerson, H. S. Jacobson, and A. W. Johnson "The Theory And Practice Of Simulated Annealing," Theory and Practice, vol. 7, 2003, pp. 287–319.

[17] F. Richard, S. G. Harvey, and W. Gunter, "A Hybrid Optimization Approach to Parameter Estimation," in Proceedings of 17th European Symposium on Computer Aided Process Engineering (ESCAPE17), Bucharest, Romania, 2007, pp. 75-80.

[18] R. Donaldson, and G. David, "A Model Checking Approach to the Parameter Estimation of Biochemical Pathway," in Proceedings of 6th Conference on Computational Methods in Systems Biology (CMSB), Rostock-Warnemünde, Germany, 2008, pp. 269-287.

[19] M. J. Gao, J. Xu, J. W. Tian, and H. Wu, "Path Planning for Mobile Robot Based on Chaos Genetic Algorithm," in Proceedings of Fourth International Conference Natural Computation (ICNC), Jinan, China, 2008, pp. 409-413.

[20] S. F. Hwang, and R. S. He, "A hybrid real-parameter genetic algorithm for function optimization," Advanced Engineering Informatics, vol. 20, 2006, pp. 7–21.

[21] F. S. Wang, and J. P. Chiou, "Estimation of Monod model parameters by hybrid differential evolution," Bioprocess and Biosystems Engineering, vol. 24, 2001, pp. 109–113.

[22] S. Moonchai, W. Madlhoo, K. Jariyachavalit, H. Shimizu, S. Shioya, and S. Chauvatcharin, "Application of a mathematical model and Differential Evolution algorithm approach to optimization of bacteriocin production by Lactococcus lactis C7," Bioprocess and Biosystems Engineering, vol. 28, 2005, pp. 1–17.

[23] L. Feng, Y. F. Yang, and Y. X. Wang, "A New Approach to Adapting Control Parameters in Differential Evolution Algorithm," Lecture Notes in Computer Science, vol. 5361/2008, 2008, pp. 433-444.

[24] M. J. Dunlop, E. Franco, and R. M. Murray, "A Multi-Model Approach to Identification of Biosynthetic Pathways" in Proceedings of American Control Conference (ACC), New York, USA, 2007, pp. 1600 – 1605.

[25] R. Storn, and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces.," Journal of Global Optimization , vol. 11, no. 4 1997, pp. 341-359.

[26] J. A. Cetto, The Kalman Filter [online]. Institut de Robotica i Informatica Industrial, UPC-CSIC. Llorens i Artigas 4-6, Edifici U, 2a pl. Barcelona 08028, Spain. Available from: http://digital.csic.es/bitstream/10261/30069/1/doc1.pdf.

[27] K. Nielson, P. G. Sorensen, F. Hynne, and H. G. Busse, "Sustained oscillations in glycolysis: an experimental and theoretical study of chaotic and complex periodic behavior and of quenching of simple oscillations," Biophysical Chemistry, vol. 72, 1998, pp. 49-62.

[28] M. T. Borisuk, and J. J. Tyson, "Bifurcation analysis of a model of mitotic control in frog eggs," J Theor Biol, vol. 195, no. 1, 1997, pp. 69-85.

[29] U. Z. Ijaz, A. K. Khambampati, J. S. Lee, S. Kim, and K. Y. Kim, "Nonstationary phase boundary estimation in electrical impedance tomography using unscented Kalman filter," Journal of Computational Physics, vol. 227, no. 15, 2008, pp. 7089-7112.

Chuii Khim Chong was born in Malaysia, on March 22, 1986. She received the BSc degrees in Computer Science from Universiti Teknologi Malaysia, in 2010. She is now a PhD student at the Artificial Intelligence and Bioinformatics Research Group, Faculty of Computer Science and Information System, Universiti Teknologi Malaysia, Malaysia. Her research interests include evolutionary algorithms, database, and programming. She has published 2 international referred publications.

Dr. Mohd Saberi Mohamad was born in Malaysia, on February 03, 1980. He received the BSc and MSc degrees in Computer Science both from Universiti Teknologi Malaysia, in 2002 and 2005, respectively. He received the PhD degree in Intelligent Systems for Bioinformatics from Osaka Prefecture University in 2010. He is now a senior lecturer at Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, Malaysia. His journal papers in the field of bioinformatics using computational intelligence approaches have been published with the total number is more than 20 papers. He has interests in computational methods such as particle swarm optimizers, hybrid approaches, genetic algorithms, support vector machines, and neural networks.

Prof. Dr. Safaai Deris (M'98) was born in Malaysia on August 13, 1955. He is a Professor of Artificial Intelligence and Software Engineering at Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia. He received the M. E. degree in Industrial Engineering, and the Doctor of Engineering degree in Computer and System Sciences, both from Osaka Prefecture University, Japan, in 1989 and 1997, respectively. His recent academic interests include the application of intelligent techniques in scheduling and bioinformatics. He is also a reviewer for several refereed journals.

Dr. Mohd Shahir Shamsir received his PhD in Computer Science from the University of Exeter in 2005. He is currently a senior lecturer and IT Manager in Universiti Teknologi Malaysia. His research interests include Biodiversity

informatics, biodiversity databases, visualisation of data and data curation for biodiversity. Using Molecular Dynamics in studying protein folding, behaviour and conformations especially those related to conformational diseases. Other interests are in the area of pedagogy used in bioinformatics education.

Yee Wen Choon is a postgraduate student at the Artificial Intelligence and Bioinformatics Research Group, Faculty of Computer Science and Information System, Universiti Teknologi Malaysia. Her research interests include evolutionary algorithms, metabolic engineering, and programming. She has published 1 international referred publication.

Chai Lian En is a postgraduate student at the Artificial Intelligence and Bioinformatics Research Group, Faculty of Computer Science and Information System, Universiti Teknologi Malaysia. His current research interests involve modelling gene networks using statistical methods, including DBN, as well as analysis of cDNA microarray gene expression data.