Calibration of Algal Growth Model Using Multi-objective Genetic Algorithm

Devesh Prakash, Amuly Ratn, Sumit Kumar, and Purnendu Bose

Abstract—The calibration of a comprehensive mathematical model describing algal dynamics in batch reactors has been attempted. The mathematical model consisted of 29 variables and 40 parameters and was described by 29 stiff differential equations. Of the 40 parameters values, 23 values are not known with certainty and hence need to be adjusted to obtain a good fit between the model simulations and experimental data. Three sets of experimental data were available for model calibration. Each set of data consisted of time series on evaluation of seven variables. Objective was to manipulate the values of the adjustable parameters in the model such that model simulations fit all three sets of experimental data simultaneously with minimum error. The above proposition was formulated as a multi-objective optimization problem and solved using a genetic algorithm called NSGA-II. The code was implemented on MATLAB and Pareto-optimal sets of parameter values were obtained. Model simulations using optimized parameter values do provide a better fit to the experimental data as compared to fits that could be obtained through adjustment of parameter values by trial and error.

Index Terms—Algal growth model, genetic algorithm, multi-objective optimization, Pareto-optimal solution.

I. INTRODUCTION

A comprehensive mathematical model for describing algal growth dynamics in batch reactors has been formulated (Shriwastav, 2014) [1]. This model has 29 variables and 40 adjustable model parameters. Values of 17 of these parameters are either known with precision from literature reports or have been determined experimentally. The ranges of value of 13 other parameters are known from the literature, while values of 10 parameters are unknown. For any model to have practical utility and acceptance it is very important to determine proper values of all parameters in the model, i.e., the model must be calibrated Three sets of experimental data were available under different initial conditions, with each set containing the time series data of seven model variables. The objective is to determine a set of model parameters, such that model simulations using these parameters simultaneously fit all the above three sets of experimental data with minimum error. The model calibration was done previously by a trial and error procedure and a reasonable fit between model simulations and experimental data set was obtained. Present objective is to further fine-tune the model parameter values using a mathematical procedure such that a better fit between

experimental data and model simulations can be obtained.

The problem is obviously a multi-objective optimization problem with the objective that error in all three sets of experimental data be simultaneously minimized by manipulation of the adjustable parameter values. A number of mathematical tools are available to optimize a multi-objective problem. Genetic algorithm (GA) belongs to a set of optimization algorithm known as evolutionary algorithm (EA). The idea behind using an evolutionary algorithm like GA is that biological evolution has produced organisms, which are capable of living in almost every possible landscape available on this earth. So taking a cue from the nature, we can exploit the utility of natural evolution to do optimization. The GA has been massively used for solving optimization problems and is being investigated intensively to improve its utility. An advanced version of this technique known as Non- dominated Sorting Algorithm-II has been used in the present work [2]. In multi-objective optimization, it might not be possible to minimize all objectives simultaneously. Instead the objective is to find a set of non-dominating, i.e., Pareto optimal solutions. The parameter sets corresponding to the Pareto-optimal solutions provide the best fits between the model simulations and the experimental data [3].

In the present study, Pareto-optimal sets of parameter values have been obtained for the algal growth model using the NSGA-II algorithm. The code for implementing this algorithm is written and run in MATLAB. Model simulations using optimized parameter values do provide a better fit to the experimental data as compared to fits that could be obtained through adjustment of parameter values by trial and error. [4]

The main objectives of the present study were as follows,

- Formulation of a multi-objective optimization problem for simultaneously minimizing the error between model simulation and experimental data sets.
- Selecting and implementing an algorithm for solving the above problem.
- Checking the algorithm to ensure that it is fulfilling its objectives and that unique optimal solution sets are obtained after its application to the problem at hand.
- To demonstrate that model simulations using parameter values obtained after optimization showed superior fit to the experimental values as compared to the initial estimates of the parameters as obtained by Shriwastav (2014).

Devesh Prakash, Sumit Kumar, and Purnendu Bose are with the Indian Institute of Technology Kanpur, Uttar Pradesh, India (e-mail: pdevesh@iitk.ac.in, sumitkr@iitk.ac.in, pbose@iitk.ac.in).

Amuly Ratn was with the Indian Institute of Technology Kanpur, Uttar Pradesh, India (e-mail: amulyratn@gmail.com).

II. METHOD

A. Algal Growth Model Description

A mechanistic Algal Growth Model (AGM) was

formulated by Shriwastav (2014) for simulation of algal growth dynamics in a completely mixed batch system. The model consists of 29 variables as summarized in Table I and is described using 29 coupled and stiff ordinary differential equations (ODEs). The model also has 40 adjustable parameters, whose values have to be fixed through calibration using experimental data. The 40 parameters described above have been divided into three categories. Type I parameters (see Table II) are those whose values are known and universally accepted. [5] Type II parameters (see Table III) are those for which a range of values have been reported in the literature. The values of the Type III parameters (see Table IV) are mostly unknown, but they concern certain important biological processes incorporated in the above model.

B. Model Calibration Strategy

Three algal reactors were operated under different initial conditions for 5 days (Shriwastav, 2014). The values of the following 7 variables, A, DNO3, DIP, DOC, DIC, DO and pH were monitored at 12-hour intervals in all three reactors. At the outset, the AGM was run using the same initial conditions as the experiments and assuming certain values of Type II and Type III parameters. The objective was to fit the model simulation results to the experimental data obtained from all three experiments with the same set of parameters. The fit of the model simulations to the data was refined by adjusting the parameter values by trial and error. Reasonable agreements between experimental results and the model simulations could be obtained through this procedure. The parameter values obtained through this procedure are given in Tables III and IV [6].

A mathematical procedure involving multi-objective optimization using genetic algorithms was adopted to adjust the parameter values still farther and thus get a better fit of the model to the experimental data. Description of the mathematical procedure adopted for this purpose is given in the following sections.

C. Error in Model

Data is available for 7 variables for each experiment, i.e., 21 time series of data is available, with each time series consisting of 10 points. Let E_{ij} be the error corresponding to a particular time series, where i ($i = 1, 7$) denotes the variable and j ($j = 1, 3$) denotes the experiment. E_{ij} has been defined as $E_{ij} = \sum_{k=1}^{10}(y'_{ijk} - y_{ijk})^2$, where y_{ijk} is the kth experimental value in the time series and y'_{ijk} is the corresponding value obtained from model simulation [7].

All E_{ij} values calculated above are normalized as follows. For each i, the normalized error (N_{ij}), $j = 1…3$

$$N_{ij} = \frac{E_{ij}}{Max[E_{ij}]} \qquad (1)$$

Hence, all N_{ij} values are in the 0-1 range. Normalized error corresponding to each set of experimental data is obtained by summing the N_{ij} ($i = 1, 7$) values corresponding to each j, i.e., for $j = 1, 3 = \epsilon_j = \sum_{i-1}^{7} N_{ij}$.

D. Problem Formulation

Thus ε_j ($j = 1, 3$) are the three objective functions that must be simultaneously minimized by manipulating the parameters

shown in Tables III and IV.

Further for the purpose of optimization, the parameter values were constrained in the ranges indicated in Tables III and IV.

TABLE I: List of Variables in the Algal Growth Dynamics Model

S. No.	Symbol	Units	Description
1	A	mg L^{-1}	Algal biomass (as dry weight)
2	TN	moles L^{-1}	Total N in the system
3	DIN	moles L^{-1}	Bulk liquid phase inorganic N
4	AN	moles L^{-1}	Algal N
5	AIN	moles L^{-1}	Algal inorganic N
6	ANO$_2$	moles L^{-1}	Dissolved nitrite in algal cells
7	AON	moles L^{-1}	Algal organic N
8	DNO$_3$	moles L^{-1}	Bulk liquid phase dissolved nitrate (as N)
9	DNO$_2$	moles L^{-1}	Bulk liquid phase dissolved nitrite (as N)
10	TP	moles L^{-1}	Total P in the system
11	DIP	moles L^{-1}	Bulk liquid phase inorganic P
12	AP	moles L^{-1}	Algal P
13	AIP	moles L^{-1}	Algal inorganic P
14	AOP	moles L^{-1}	Algal organic P
15	H$_3$PO$_4$	moles L^{-1}	Bulk liquid phase dissolved H$_3$PO$_4$ (as P)
16	H$_2$PO$_4^-$	moles L^{-1}	Bulk liquid phase dissolved H$_2$PO$_4^-$ (as P)
17	HPO$_4^{2-}$	moles L^{-1}	Bulk liquid phase dissolved HPO$_4^{2-}$ (as P)
18	PO$_4^{3-}$	moles L^{-1}	Bulk liquid phase dissolved PO$_4^{3-}$ (as P)
19	TC	moles L^{-1}	Total C in the system
20	AC	moles L^{-1}	Algal C
21	DIC	moles L^{-1}	Bulk liquid phase dissolved inorganic C
22	DOC	moles L^{-1}	Bulk liquid phase dissolved organic C
23	CO$_2$(aq)	moles L^{-1}	Bulk liquid phase dissolved CO$_2$ (as C)
24	H$_2$CO$_3$	moles L^{-1}	Bulk liquid phase dissolved H$_2$CO$_3$ (as C)
25	HCO$_3^-$	moles L^{-1}	Bulk liquid phase dissolved HCO$_3^-$ (as C)
26	CO$_3^{2-}$	moles L^{-1}	Bulk liquid phase dissolved CO$_3^{2-}$ (as C)
27	O$_2$	moles L^{-1}	Bulk liquid phase dissolved O$_2$
28	O$_2$^{a}	moles L^{-1}	Dissolved O$_2$ in algal cells
29	H$^+$	moles L^{-1}	Bulk liquid phase H$^+$ ions

E. Solution Procedure

The multi-objective optimization problem described above

was solved using Non-dominated Sorting Genetic Algorithm-II (NSGA-II) developed by Deb *et al.*, (2002).

Genetic algorithms are a population-based stochastic global search technique inspired from the biological principals of natural selection and genetic recombination (Wang *et al.*, 2005). The algorithm used here for the optimization of parameters is a fast and elitist multi-objective Genetic Algorithm known as Non-dominated Sorting Genetic Algorithm-II (NSGA-II) (Deb *et al.*, 2002). [8] NSGA-II reduces computational complexity from O (MN^3) to O (MN^2), where M is the number of objectives and N is the population size as compared to other multi-objective evolutionary algorithms (Deb *et al.*, 2002). It also preserves the best solution found in each generation and carries it to the next generation, a phenomenon called elitism.

The basic algorithm for NSGA-II was developed by Deb *et al.*, (2002). In the present case, a parent population of size 'N', each containing a set of 23 parameter values is randomly generated, with each parameter taking a value in the range defined for that parameter. In the terminology of genetic algorithm, each parameter is a 'gene' and a complete set of parameters is a 'chromosome' and 'N' such 'chromosomes' constitutes the initial population. The initial population is sorted based on non-domination criteria into different fronts. A fitness or rank is assigned to each 'chromosome' in all the fronts. In addition to fitness, another parameter called crowding distance is also calculated for each 'chromosome'. High average crowding distance will result in better diversity in the 'chromosomes'. After sorting and computing crowding distance, the selection of individual is done for crossover and mutation operations using binary tournament selection with crowded-comparison operator. The crossover and mutation

operations create offspring population of size 'N'. In the present case, Simulated Binary Crossover (SBX) (Deb and Agrawal, 1994), with probability of 0.9 and Polynomial Mutation (Deb, 2001) with probability of 0.1 were used for crossover and mutation operations respectively. Also the distribution indices for crossover and mutation operators are taken as 10 and 50 respectively. The distribution indices determine degree of spread of the offspring population from the parent population (Deb and Beyer, 2001). Using the parent and offspring populations, the next generation is created based on the elitism principle. Successive generations are created as above, with the objective of ultimately obtaining 'Pareto-optimal' populations. Each 'chromosome' of the 'Pareto-optimal' population is considered to be 'equally good' solution to the optimization problem [9].

III. RESULT AND DISCUSSION

A. Calculation of Initial Error

The AGM was run using the same initial conditions as the experiments and assuming certain values of Type II and Type III parameters. The objective was to fit the model simulation results to the experimental data obtained from all three experiments with the same set of parameters. The fit of the model simulations to the data was subsequently refined by adjusting the parameter values by trial and error. Reasonable agreement between experimental results and the model simulations could be obtained through this procedure. The parameter values obtained through this procedure are given in Tables III and IV.

TABLE II: ALGAE MODEL PARAMETERS OF TYPE I

| S. No.	Parameter	Description	Value Used in Simulation	Unit
1.	K_{Ho}^{25}	Henry's Constant for O_2 at 25°C	1.3×10^{-3}	moles L^{-1} atm^{-1}
2.	K_{Lo}^{20}	Mass Transfer Coefficient of O_2 at 20°C	6.000	m day^{-1}
3.	K_{Hc}^{25}	Henry's Constant for CO_2 at 25°C	0.036	moles L^{-1} atm^{-1}
4.	K_{Lc}^{20}	Mass Transfer Coefficient of CO_2 at 20°C	5.358	m day^{-1}
5.	P_o	Partial Pressure of O_2 in Atmosphere	0.210	atm
6.	P_c	Partial Pressure of CO_2 in Atmosphere	3.8×10^{-4}	atm
7.	K_{c1}	$K_{c1} = [H_2CO_3]/[CO_2^{aq}]$	1.5×10^{-3}	--
8.	K_{c2}	$K_{c2} = [HCO_3^-].[H^+]/[H_2CO_3]$	3.6×10^{-4}	moles L^{-1}
9.	K_{c3}	$K_{c3} = [CO_3^{-2}].[H^+]/[HCO_3^-]$	4.7×10^{-11}	moles L^{-1}
10.	K_1	$K_1 = [H^+].[OH^-]$	1.0×10^{-14}	(moles L^{-1})^2
11.	K_{p1}	$K_{p1} = [H_2PO_4^-].[H^+]/[H_3PO_4]$	7.5×10^{-3}	moles L^{-1}
12.	K_{p2}	$K_{p2} = [HPO_4^{2-}].[H^+]/[H_2PO_4^-]$	6.2×10^{-8}	moles L^{-1}
13.	K_{p3}	$K_{p3} = [PO_4^{3-}].[H^+]/[HPO_4^{2-}]$	4.8×10^{-13}	moles L^{-1}
14.	f_H	Stoichiometry for H^+ consumption for algal growth	5.1×10^{-3}	mg H^+ mg^{-1} algae
15.	θ	Arrhenius temperature coefficient	1.070	--
16.	a	Specific interfacial area for mass transfer	0.752	m^{-1}

TABLE III: ALGAE MODEL PARAMETERS OF TYPE II

S. No.	Parameter	Description	Values in Literature	Range used for Optimization	Unit
1.	$(\mu_m)_a^{20}$	Maximum specific algal biomass growth rate at 20°C	0.200-2.500	0.200-2.500	day^{-1}
2.	$(k_r)_a^{20}$	Algal respiration rate constant at 20°C	0.050-0.300	0.050 - 0.300	day^{-1}
3.	$(\rho_{ap})_{max}^{20}$	Maximum uptake rate of P by algae at 20°C	0.050-0.200	0.050-0.200	mg P mg^{-1} algae d^{-1}
4.	$(q_{a0})_n$	Minimum weight fraction of N in an algal cell below which algal growth ceases	0.007-0.036	0.007-0.036	mg N mg^{-1} algae
5.	$(K_a)_n$	Attenuation coefficient for low bulk liquid phase nitrate concentration	1.4×10^{-6}-3.6×10^{-4}	1.4×10^{-6}-3.6×10^{-4}	moles L^{-1}
6.	$(K_a)_p$	Attenuation coefficient for low DIP concentration	3.2×10^{-8}-3.3×10^{-4}	3.2×10^{-7}-3.2×10^{-5}	moles L^{-1}
7.	$(\rho_{an})_{max}^{20}$	Maximum uptake rate of N by algae at 20°C	0.720	0.000-1.440	mg N mg^{-1} algae d^{-1}
8.	$(K_{aq})_n$	Attenuation coefficient for high AIN concentration inside algae cells	0.009	0.000-0.018	mg N mg^{-1} algae
9.	$(K_{aq})_p$	Attenuation coefficient for high AIP concentration inside algae cells	0.001	0.00-0.003	mg P mg^{-1} algae
10.	$(q_{a0})_p$	Minimum weight fraction of P in an algal cell below which algal growth ceases	0.001	0.000-0.002	mg P mg^{-1} algae
11.	$(K_{sa})_c$	Algae half saturation constant for DIC	1.3×10^{-5}	1.1×10^{-5}-1.5×10^{-5}	moles L^{-1}
12.	f	Fraction of carbon in algal biomass	0.400	0.000-0.800	mg C mg^{-1} algae
13.	$(K_s)_I$	Algae half-saturation constant for light	16.400-134.000 W m^{-2}	0.161 - 0.218 W L^{-1}	W L^{-1}

Next, the normalized error between the experimental values and corresponding model simulations was determined for all three set of experiments (see ε_j values in Table V). Further the total error was determined as $\epsilon = \sum_{j=1}^{3} \epsilon_j$. The value of ε was determined to be $j = 11.067$ (see Table V). Thus the objective of the mathematical procedure adopted was to further adjust the model parameter values such that the value of ε can be reduced further, thus obtaining a better fit between experimental results and the corresponding model simulations.

A. Testing of the Optimization Procedure

The multi-objective optimization was set up to minimize the ε_j ($j = 1, 3$) values simultaneously by manipulating the parameters shown in Tables III and IV. Further for the purpose of optimization, the parameter values were constrained in the ranges indicated in Tables III and IV.
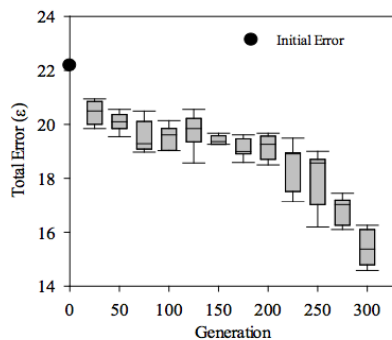
Scale: 2 unit= 1 error unit on Y-axis
Fig. 1. Evolution of total error (e) in successive generations population size = 10.

For the multi-objective optimization procedure, a population of size 10 was chosen. Each member of the population consisted of a randomly selected set of values for the 23 model parameters (see Tables III and IV), but constrained in the ranges defined for each parameter as given in Tables III and IV. Application of the multi-objective optimization procedure resulted in generation of successive population sets. The relevant statistics of total error (e) of the population sets (10 Nos.) generated in successive generations is shown in Fig. 1.

As expected, the total error (e) value is seen to diminish with successive generations Further, variations between corresponding population sets of successive generations must decrease and largely disappear as calculations are done for more and more generations. The metric used for this purpose is S_K, where, $S_K = \sum_{I=1}^{23} \sum_{J=1}^{N} (P_{i,j,k+1} - P_{i,j,k})^2$, where p_{ijk} is the parameter value for the ith parameter ($i = 1, 23$) of the jth member of population ($j = 1, N$) in the kth iteration.

The moving average of S_K, i.e., the average of S_K values of the preceding 15 generations to the kth generation is plotted against the kth generation in Fig. 2 (Hamilton, 1994; Wei, 1994). It is clearly seen that the variations in S_K decrease with number of generations and become nearly zero.

$$S_K = \sum_{I=1}^{23} \sum_{J=1}^{N} (P_{i,j,k+1} - P_{i,j,k})^2 \qquad (2)$$

Subsequently, the multi-objective optimization procedure was repeated with a population size of 20. The evolution of total error (e) and reduction in variability in population over successive generations (S_K) showed a similar trend in this case also these results were sufficient to conclude that the algorithm is working as intended.

B. Results

The parameter values obtained after parameter optimization with a population (size, $N = 10$) over 300 generations are presented in Appendix (Table VI). The parameters obtained through the initial calibration using trial and error procedure are also shown in the same table. The main conclusion to be drawn from these results is that the optimization procedure, in most cases, resulted in only a slight refinement of the parameter values obtained earlier through trial and error.

Error analysis of the optimization protocol is summarized in Appendix (Table VII). This clearly indicates that errors between experimental data and model simulations can be effectively reduced using the optimization protocol to adjust

model parameters.

The average parameter values and the corresponding standard deviations in each population are seen to be largely similar. This demonstrates that parameter values tend to converge to the same optimum range irrespective of the starting values for the optimization run or the population size.

TABLE IV: ALGAE MODEL PARAMETERS OF TYPE III

S. No.	Parameter	Description	Value Used in Simulation	Range used for Optimization	Unit
1.	$(N_R)_{min}$	Minimum *nitrate reductase* activity	0.070	0.059 - 0.080	--
2.	K_{NR}	Intrinsic rate of conversion of nitrate to nitrite within the algal cells	2.000	1.700 - 2.300	day^{-1}
3.	n	Empirical constant for *nitrate reductase* activity	3.000	2.550 - 3.450	--
4.	F	Maximum fraction of algal biomass that can be present as nitrite without triggering nitrite release from algal cell to the bulk phase	0.040	0.034 - 0.046	mg N mg^{-1} algae
5.	K_{NL}^{20}	Rate constant for nitrite release from algal cells to the bulk phase at 20ºC	0.055	0.046 - 0.063	day^{-1}
6.	k_{ex}^{20}	Rate constant for excretion of organic carbon by algae to the bulk liquid phase at 20ºC	5.0×10^{-5}	4.2×10^{-5} - 5.8×10^{-5}	moles L^{-1} d^{-1}
7.	K_{oa}	Oxygen uptake/excretion rate constant for algal cells	0.002	1.3×10^{-3} - 1.7×10^{-3}	moles L^{-1} d^{-1}
8.	q_o^{sa}	Maximum oxygen storage in algae cells	0.050	0.042 -0.058	mg O$_2$ mg^{-1} algae
9.	$(K_a)_r$	Attenuation coefficient on respiration for low oxygen levels	0.005	4.2×10^{-3} - 5.8×10^{-3}	mg O$_2$ mg^{-1} algae
10.	f_o	Stoichiometry of net algal oxygen production	1.240	1.054 - 1.426	mg O$_2$ released mg^{-1} algae
11.	α	Attenuation factor due to shading	0.002	0.002	mg^{-1} algae

(The range for the optimization of these parameters is taken as +/-15% of the original value used except for the last one, which was determined experimentally and is used as it is.)

TABLE V: INITIAL NORMALIZED ERRORS AND TOTAL ERRORS

Expt. No. (*j*)	A (*i* = 1) N_{1j}	(*i* = 2) N_{2j}	(*i* = 3) N_{3j}	(*i* = 4) N_{4j}	(*i* = 5) N_{5j}	(*i* = 6) N_{6j}	(*j*= 7) N_{7j}	$\epsilon_j = \sum_{i=1}^{7} N_{ij}$
1	0.813	0.422	1.000	1.000	0.389	0.368	0.904	4.896
2	0.040	0.860	0.703	0.299	0.371	0.343	0.220	2.836
3	0.504	0.524	0.145	0.173	0.585	0.404	1.000	3.335

C. Model Fit to Experimental Data

Experimental data on three experiments involving operation of algal reactors under different conditions were available. Corresponding model simulations using the AGM model, with model parameters refined by trial and error were also available. The model parameters were further refined using the optimization procedure to get a better model fit to the experimental data. To visualize the actual improvements in the model fit, the parameter values corresponding to Pop-10 of Appendix (Table VI) were used to simulate the experimental data respectively). In all cases, the model corresponding to the optimized parameters are seen to fit the experimental data more closely than the model simulation using parameters refined by trial and error. It may thus be concluded that the optimization strategy used for refining the model parameters of the AGM is a success and leads to closer fitting of the model simulations to the experimental data.

IV. SUMMARY AND CONCLUSIONS

An algal growth model (AGM) was formulated by Shriwastav (2014) to simulate algal growth in batch algal reactors. This model has 29 variables, which were described using 29 stiff ordinary differential equations. The model also had 40 adjustable parameters and values of 23 of these parameters were not known with sufficient accuracy. Experimental data on a batch algal reactor operation under eight initial conditions were available from Shriwastav

(2014). Under each of these conditions, values of 7 variables in the reactors were measured at 12-hour intervals over 5-days. The objective was to fit simulations of the AGM simultaneously to the experimental data sets by adjusting the 23 parameters. The parameters were initially adjusted by trial and error to obtain a reasonable fit of the model simulations to the experimental data (Shriwastav, 2014). The objective of the present optimization exercise was further refinement of the model parameters to obtain a better fit of the model simulations to the experimental data.
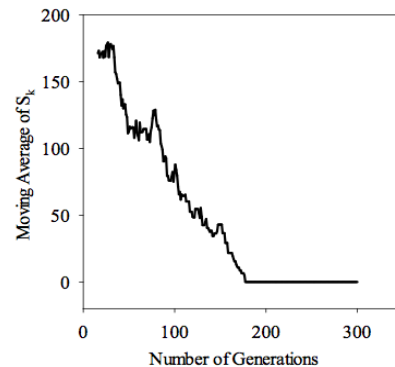

Fig. 2. Reduction in variability in populations (S_k) with successive generations with population size = 10.

In the present work, the model parameters were optimized by formulating a multi-objective optimization problem which simultaneously minimizes the error between model simulation and experimental data of all three experiments. A

genetic algorithm called NSGA-II was used to solve the optimization problem. The algorithm was implemented on MATLAB. The algorithm starts with a certain number of randomly selected sets (say 'N') of parameter values to be optimized. These population sets are modified after each generation until a Pareto-optimal sets of population are obtained. After implementation, the performance of the algorithm was checked as follows,

- Total error (e), is the sum of values of all three objective functions. The average value of e (over the 'N' population sets) declined progressively with the increase in the number of generations. This suggested that the algorithm was indeed adjusting the parameters such that error between model simulation and experimental data is reduced.

- The sum of the square of differences between the corresponding values of 'N' population sets in successive

generations declined and reduced to almost zero with the increase in the number of generations. This suggested that further improvement in the population sets was not possible.

- When two randomly selected population sets of size $N = 10$ were independently optimized as above, the final population sets obtained after 300 generations were similar.

- Comparison of two runs of the algorithm with $N = 10$ and 20 showed that the average values of parameters in the optimized population sets were similar.

Finally, as shown in Fig. 2, model simulations using parameter values obtained from the population set showed superior fit to the experimental values as compared to the parameters obtained by Shriwastav (2014) through adjustment by trial and error.

TABLE VI: PARAMETERS AFTER 300 GENERATIONS FOR POPULATION SIZE, $N = 10$: CASE I

S. No.	Parameter	Initial Value	Pop-1	Pop-2	Pop-3	Pop-4	Pop-5	Pop-6	Pop-7	Pop-8	Pop-9	Pop-10
1.	$(\mu_m)_a^{20}$	0.900	1.133	1.150	2.382	2.383	2.431	2.399	2.466	2.362	2.361	2.500
2.	$(k_r)_a^{20}$	0.100	0.266	0.264	0.239	0.220	0.246	0.243	0.249	0.235	0.226	0.235
3.	$(\rho_{ap})_{max}^{20}$	0.080	0.199	0.200	0.063	0.050	0.061	0.061	0.067	0.062	0.067	0.058
4.	$(q_{a0})_n$	0.01	0.017	0.016	0.034	0.031	0.033	0.033	0.034	0.032	0.033	0.033
5.	$(K_a)_n$	7.1×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	2.1×10^{-4}	2.3×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	1.9×10^{-4}	2.2×10^{-4}	2.2×10^{-4}	1.9×10^{-4}
6.	$(K_a)_p$	3.2×10^{-6}	1.4×10^{-6}	1.5×10^{-6}	2.4×10^{-6}	3.2×10^{-6}	3.6×10^{-6}	3.2×10^{-6}	3.10×10^{-6}	2.6×10^{-6}	4.8×10^{-6}	2.8×10^{-6}
7.	$(\rho_{an})_{max}^{20}$	0.600	0.201	0.180	0.455	0.467	0.487	0.475	0.480	0.420	0.447	0.544
8.	$(K_{aq})_n$	0.020	0.007	0.007	0.009	0.009	0.010	0.009	0.010	0.009	0.009	0.009
9.	$(K_{aq})_p$	1.0×10^{-3}	1.1×10^{-5}	2.4×10^{-5}	2.6×10^{-3}	2.6×10^{-3}	2.5×10^{-3}	2.6×10^{-3}	2.6×10^{-3}	2.4×10^{-3}	2.5×10^{-3}	2.6×10^{-3}
10.	$(q_{a0})_p$	0.003	0.002	0.002	0.001	0.001	0.002	0.002	0.002	0.001	0.002	0.001
11.	$(\kappa_{sa})_c$	1.0×10^{-5}	1.2×10^{-5}	1.2×10^{-5}	1.1×10^{-5}	1.2×10^{-5}	1.2×10^{-5}	1.2×10^{-5}	1.1×10^{-5}	1.2×10^{-5}	1.2×10^{-5}	1.1×10^{-5}
12.	f	0.251	0.361	0.325	0.659	0.783	0.698	0.708	0.696	0.766	0.761	0.672
13.	$(K_s)_I$	0.190	0.163	0.161	0.215	0.213	0.218	0.218	0.218	0.216	0.215	0.218
14.	$(\mu^{ir})^{mu}$	0.070	0.072	0.072	0.071	0.070	0.071	0.071	0.071	0.070	0.069	0.071
15.	K_{NR}	2.00	1.826	1.898	2.300	2.234	2.300	2.300	2.299	2.284	2.203	2.286
16.	n	3.00	3.212	3.240	2.572	2.608	2.583	2.574	2.579	2.550	2.565	2.577
17.	F	0.040	0.042	0.042	0.039	0.040	0.0401	0.039	0.040	0.040	0.041	0.040
18.	κ_{NL}^{20}	0.055	0.058	0.059	0.061	0.063	0.061	0.061	0.061	0.062	0.0623	0.061
19.	k_{ex}^{20}	5.0×10^{-5}	4.7×10^{-5}	4.7×10^{-5}	5.6×10^{-5}	5.8×10^{-5}	5.6×10^{-5}	5.6×10^{-5}	5.5×10^{-5}	5.5×10^{-5}	5.6×10^{-5}	5.6×10^{-5}
20.	κ_{sa}	1.5×10^{-3}	1.3×10^{-3}	1.3×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}	1.7×10^{-3}
21.	q_o^{sa}	0.050	0.043	0.045	0.055	0.054	0.055	0.054	0.054	0.054	0.054	0.055
22.	$(K_a)_r$	5.0×10^{-3}	5.7×10^{-3}	5.6×10^{-3}	4.3×10^{-3}	4.3×10^{-3}	4.2×10^{-3}	4.3×10^{-3}	4.2×10^{-3}	4.3×10^{-3}	4.3×10^{-3}	4.3×10^{-3}
23.	fo	1.240	1.090	1.112	1.341	1.334	1.353	1.357	1.353	1.337	1.333	1.363

TABLE VII: ERROR ANALYSIS: AFTER 300 GENERATIONS WITH POPULATION SIZE, $N = 10$: CASE I

Experiment No. (j)	Initial j	j Values after 300 iterations									
		Pop-1	Pop-2	Pop-3	Pop-4	Pop-5	Pop-6	Pop-7	Pop-8	Pop-9	Pop-10
1	4.896	5.975	5.994	1.602	4.698	1.598	1.531	1.671	1.532	1.507	1.707
2	2.836	1.308	1.380	5.458	3.698	5.210	4.321	4.729	3.929	3.836	4.252
3	3.335	4.895	4.902	3.453	0.738	3.638	3.567	2.067	0.842	0.805	1.001
4	1.995	1.173	1.163	4.650	4.723	4.445	4.472	5.355	5.356	5.392	5.393
5	3.624	2.126	2.123	0.381	0.536	0.397	0.454	0.416	0.638	0.593	0.361
6	1.903	0.875	0.859	0.605	1.162	0.651	0.756	0.737	1.410	1.138	0.756
7	1.838	1.231	1.219	0.402	0.553	0.404	0.432	0.421	0.597	0.543	0.396
8	1.764	0.658	0.651	0.639	1.065	0.686	0.727	0.705	1.071	0.980	0.721
Total Error ( )	22.191	18.241	18.291	17.190	17.173	17.029	16.260	16.101	15.385	14.794	14.587

support this work would not have met its desired end.

REFERENCES

[1] A. Shriwastav, "Algal-bacterial photo bioreactor for removal of nutrients from wate," Ph.D. dissertation, Dept. EEM, Indian Institute of Technology, Kanpur, India, 2014.

[2] C. Dai, M. Yao, Z. Xie, C. Chen, and J. Liu, "Parameter optimization for growth model of greenhouse crop using genetic algorithms," *Applied Soft Computing*, vol. 9, no. 1, pp. 13-19, 2009.

[3] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, pp. 1-34, 1994.

[4] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," *Machine Learning*, vol. 3, no. 2, pp. 95-99, 1988.

[5] N. Muttil and J. H. Lee, "Genetic programming for analysis and real-time prediction of coastal algal blooms," *Ecological Modelling*, vol. 189, no. 3, pp. 363-376, 2005.

[6] A. Ng and B. Perera, "Selection of genetic algorithm operators for river water quality model calibration," *Engineering Applications of Artificial Intelligence*, vol. 16, no. 5, pp. 529-541, 2003.

[7] G. J. Pelletier, S. C. Chapra, and H. Tao, "QUAL2Kw–a framework for modeling water quality in streams and rivers using a genetic algorithm for calibration," *Environmental Modelling & Software*, vol. 21, no. 3, pp. 419-425, 2006.

[8] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multi-objective evolutionary algorithms: Empirical results," *Evolutionary Computation*, vol. 8, no. 2, pp. 173-195, 2000.

[9] Q. Wang, "Using genetic algorithms to optimise model parameters," *Environmental Modelling & Software*, vol. 12, no. 1, pp. 27-34, 1997.

Devesh Prakash was born in Patna, India on January 10, 1994. As a graduate researcher at IIT Kanpur, he is pursuing his master's thesis under the guidance of Prof. Purnendu Bose. He has completed his B. Tech degree in civil engineering from IIT Kanpur, India.