Pre-Harvest Factors Optimization using Genetic Algorithm for Lettuce

I. Valenzuela[1], R. Baldovino[2], A. Bandala[1], and E. Dadios[2]
[1]*Electronics and Communications Engineering (ECE) Department, De La Salle University, 2401 Taft Avenue, Manila 0922 Philippines.*
[2]*Manufacturing Engineering and Management (MEM) Department, De La Salle University, 2401 Taft Avenue, Manila 0922 Philippines.*
ira_valenzuela@dlsu.edu.ph

***Abstract*—The agricultural sector is facing problems on crop development due to climate change and global warming. Crops such as rice, tomato, corn, lettuce, potato, wheat, soybeans and others are affected. Through analyzing the graphical representation of data, no optimum values are observed. In this study, the suitability of the genetic algorithm in finding the best condition for producing high quality lettuce crop was determined. The parameters that were optimized are the light intensity, temperature and CO_2. These parameters were essential preharvest factors for lettuce. The system selected the 50 fittest individuals based on the fitness score and then proceeds to the recombination process. A mutation has been applied to test if the solution is the global one. When the iterations had reached the required number of generation, the system stopped and gave the best condition for lettuce. Critical design on GA was done and the best fitness plot was obtained. The GA results showed that the optimum conditions for a high-quality lettuce crop needs a light intensity of 175.22296 µmol/m2/s, a temperature of 19.36228 ºC and a CO_2 level of 803.01855 ppm.**

***Index Terms*—Optimization; Genetic Algorithm; Lettuce; Photosynthesis Rate; Light Intensity; Temperature; CO_2.**

I. INTRODUCTION

Agricultural sector is facing problems on crop development due to climate change and global warming. Crops such as rice, tomato, corn, lettuce, potato, wheat, soybeans and others are affected. Light intensity, temperature and carbon dioxide (CO_2) are considered to be the main factors affecting the growth of a crop [1][2]. Crop growth also depends on the nutrients from the soil that contains Nitrogen (N), Phosphorous (P) and Potassium (K) [3], soil's temperature and humidity, diseases, leaf temperature, water level and transpiration rate. The use of greenhouse in a Controlled Environment Agriculture may help to maintain the temperature, humidity and photosynthetic active radiation [4][5]. Thus, this improves the crop production yield.

There are many researches that focus on the production and growth of crops. Liu et al. uses mathematical algorithms and optimization in cropland mask to protect the crops from pests and harmful fertilizers and improve the farming system [6]. Mathematical approach is a tedious process and requires a good foundation in higher computational analysis. Optimization tool is used to maximize the effectivity and functionality a design can take [7][8][9]. There are studies wherein a hybrid of neural network and genetic algorithm are used for optimization of plant growth [10][11]. Training the neural network needs actual data in order to produce a high efficiency network. This kind of approach are done for complex systems requiring many parameters and needs concrete data for optimization.

However, previous studies did not consider a simpler but effective approach in optimization. Also, considering a lot of parameters makes the system more complex thus requiring complex analysis. Besides not knowing the particular parameters that greatly affects the crop growth are observed in previous studies. This study aims to determine the suitability of the genetic algorithm in finding the best condition for producing high quality lettuce crop. Lettuce is one of the most important vegetable crop. It is often studied because of its high sensitivity to temperature. The parameters to be optimized are the light intensity, temperature and CO_2. These parameters are essential preharvest factors for lettuce. The knowledge on these parameters will help the farmers in growing and improving their yield crops. Also, this study can be utilized to people who are growing their crops in a controlled environment.

The paper is organized as follows: Section II discusses the design of the genetic algorithm. Section III explains on how the system will be optimized. In Section IV described the results obtained from the simulation and graphical analysis of the parameters. Finally, Section V gives conclusions.

II. ARCHITECTURE OF THE SYSTEM

The conceptual process flow used in the design is shown in Figure 1. Genetic algorithm has been used in this study because of its capability to find the optimal solution to a certain problem [12][13][14]. The evolutionary nature of this algorithm makes it appropriate for searching the best conditions that lettuce need in order to grow healthy.

In this study, the variables that will be optimized are light intensity, temperature and CO_2. Then, an initial random population will be created and the fitness score will be computed. The system will select the 50 fittest individuals based on fitness score and then proceeds to the recombination process. A mutation will be applied to test if the solution is global. When the iterations had reached the required number of generation, the system will stop and give the best condition for lettuce.

The algorithm takes as input the condition of the environment in which the lettuce is planted. It is designed to give the best condition on which the lettuce will grow healthy and provides good quality harvest. The output of the system is the best environmental temperature, light intensity and CO_2 levels.

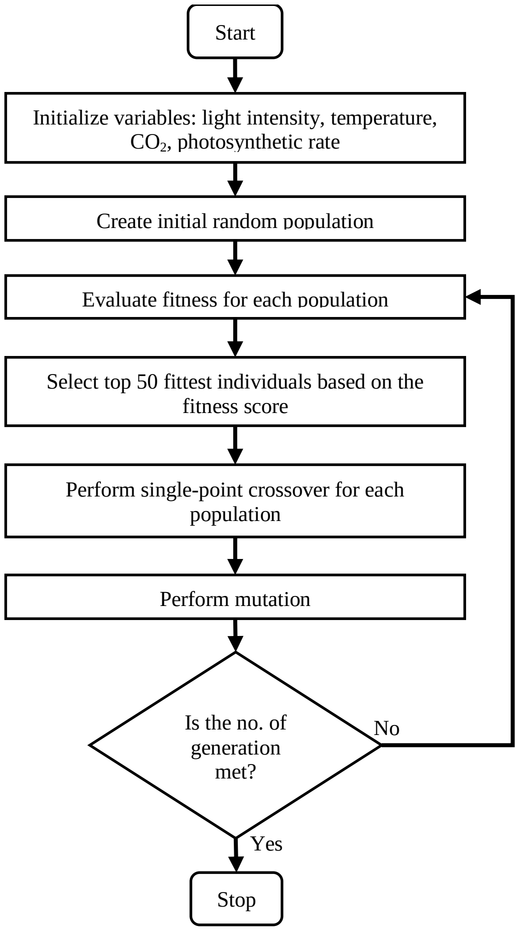

Figure 1: System process flow using GA

A. Chromosomes

The most important factors affecting the plant growth are the light intensity, temperature and CO_2 level. Many crops require an inductive photoperiod for the most rapid rate of development. The plants will delay the development but grows vegetative without this photoperiod.

The plant grows well when it is supplied by enough amount of sunlight containing the red and blue wavelengths. The plant does not need the green wavelength. It is reflected back by the plant and this is the reason why most of the leaf color is green. Also, CO_2 helps the plant to produce its own food. It helps to increase the dry matter content of the crop. Similarly, lettuce is grown in an environment with low temperature level. The temperature used in this study is the average daily temperature to determine aspects of crop growth.

Gene 1 (Light intensity)	Gene 2 (Temperature)	Gene 3 (CO_2)
100	19	800

Figure 2: Sample chromosome with its particular genes

Figure 2 displays the genetic components of a single gene. The unit used for measuring light intensity is μmol/m^2/s, temperature is ^{o}C and CO_2 level is parts per million (ppm).

The initial generation contains 50 chromosomes. These comprise different combinations of each gene with non-uniform distance. The data used were from the observed temperature, CO_2 level and light intensity during plant growth.

B. Fitness Function

In finding the best solution out of each of the numerous solutions that GA will generate, a fitness function is needed to quantify the effectiveness of each chromosome. In this study, a single-leaf photosynthesis model 0[1] has been used as fitness function. This model is selected due to its simplicity, allowing manipulation of CO_2, temperature and light response. This model is the photosynthetic rate (PR) as shown in Equation (1).

$$PR = \frac{1}{2}\theta\left\{\alpha L + P_{max} - \sqrt{(\alpha L + P_{max})^2 - 4\theta\alpha L P_{max}}\right\} \tag{1}$$

where:

PR : Photosynthetic rate
L : Instantaneous light (μmol/m^2/s)
P_{max} : Maximum rate of photosynthesis
α : Photosynthetic efficiency

The sharpness in the knee of the photosynthetic rate curve is represented by θ which has a typical value of 0.8. L is the instantaneous light inside the greenhouse at that time and is measured in μmol/m^2/s. P_{max} is the maximum rate of photosynthesis theoretically possible. It is dependent on the CO_2 level available inside the greenhouse as shown in Equation (2).

$$P_{max} = -0.116 + 1.136 * (1 - e^{-0.002 * CO_2}) \tag{2}$$

The α is described by Equation (3). It is the photosynthetic efficiency or the quantum yield of the plant, measured in parts per million (ppm), and is dependent on the current temperature.

$$\alpha = 0.0843 - (0.0003 * current\ temperature) - (0.0000341 * current\ temperature^2) \tag{3}$$

C. Selection

After the fitness values are computed, the best chromosomes in the entire population are selected. Nowadays, there are many selection techniques used in GA like roulette wheel, tournament selection, rank selection, Boltzman selection and steady-state selection [15][20][21]. The algorithm used for the selection is tournament. Tournament selection provides a type of selective pressure by holding a competition among individuals [14]. From the tournament, the best individual is the one with the highest fitness. The higher efficiency of GA in using this kind of selection process can lead to an optimal solution. This is because the fitness difference provides selection pressure and this makes the GA to improve the succeeding genes.

A penalty is added for the non-linear constraint algorithm in selecting the fittest chromosome. Top chromosomes including the parents and offspring will form the next generation. These all become parents and will undergo the reproduction. Other chromosomes that did not qualify will be discarded.

D. Crossover

Genetic recombination or crossover is the mating of partial solutions [12][13][14]. This advances the capability of the algorithm to approach and find the optimum solution [21][22][23]. Recombination is very important because it allows characteristics from two parents to be assorted [24]. There are different kinds of crossover techniques such as single point, two-point, scattered, heuristic, intermediate and arithmetic.

Single-point

Parent 1	100	19	800
Parent 2	150	23	950
Offspring 1	100	23	950
Offspring 2	150	19	800

Figure 3: Example of single-point crossover

For this study, single point crossover will be performed. The algorithm will replace either one of the genes. Figure 3 shows the sample single-point recombination.

E. Mutation

Mutation is the process of introducing genetic diversity in a given population to avoid the condition of chromosomes from becoming too similar to each other or local minima [12][13][14]. It also keeps the gene pool well stocked, and thus ensuring ergodicity. There are different techniques to introduce mutation. These techniques are Gaussian, adaptive feasible and uniform [16][17]. Figure 4 shows mutation in a given chromosome.

(a) (b)

Figure 4: Example for a) 8-bit string mutation and b) integer mutation

In this study, the mutation is done by applying the uniform process. It is a two-step technique wherein the algorithm selects a fraction of the vector entries of an individual for mutation. Then, it replaces each selected entry by a random number selected uniformly from the range for that entry.

III. OPTIMIZATION USING GENETIC ALGORITHM

In employing the genetic algorithm for the optimization of the preharvest factors, some parameters used has fixed values while others have varying values. The parameters with fixed values are the three optimization variables specifically light intensity, temperature and CO_2, the number of generation which is equal to 100, and the initial population size of 50. On the other hand, the parameters with varying values are the mutation rate, the selection rate, and crossover rate.

The relationship of the temperature on photosynthetic rate is shown in Figure 5a. High photosynthetic rate is observed when the temperature is between 22-23ºC. In Figure 5b and 5c, light intensity and CO_2 amount has varying effects on photosynthetic rate. It does not affect the photosynthetic rate much. Thus, temperature affects most in the increase of photosynthetic rate.

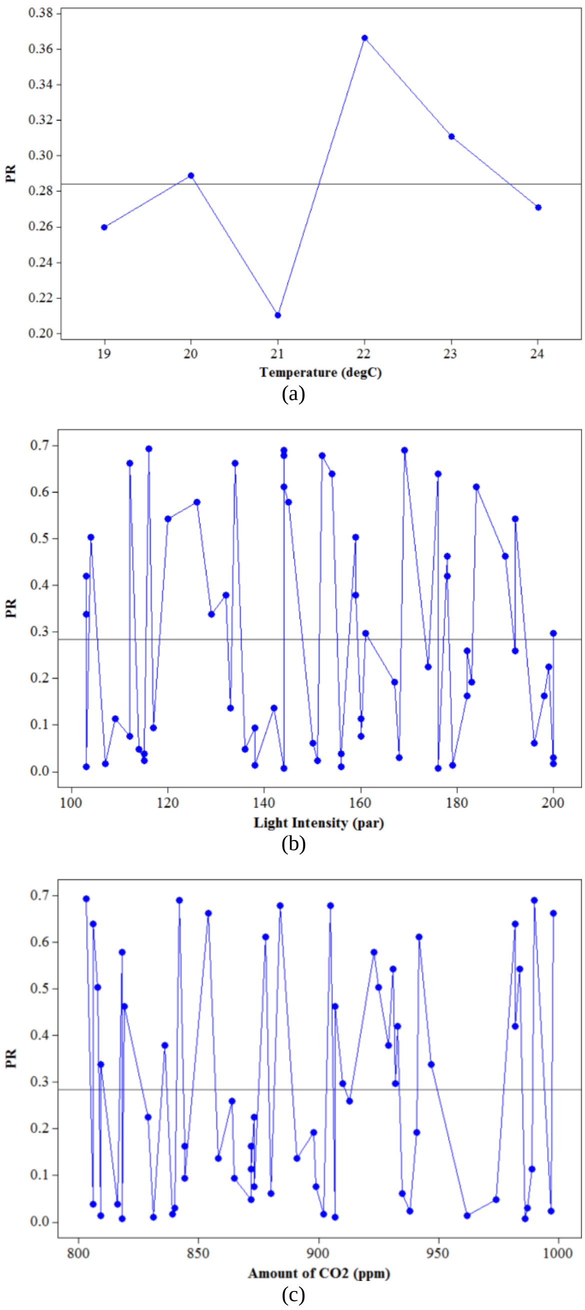

Figure 5: Plot showing (a) the effect of temperature on photosynthetic rate, (b) the effect of light intensity on photosynthetic rate, (c) the effect of amount of CO_2 on photosynthetic rate

In Figure 6, the interactions of the two variables to photosynthetic rate are shown. It is noticeable in Figure 6a that higher photosynthetic rate is achieved when the values at the middle of values of light intensity and temperature. The mid value of light intensity and highest value of CO_2 gives a high value of photosynthetic rate as shown in Figure 6b. When the temperature is low and the value of CO_2 is high, the photosynthetic rate is also high (see Figure 6c). Due to the varying effects of these parameters on photosynthetic rate, GA is needed in determining the optimum values of each.

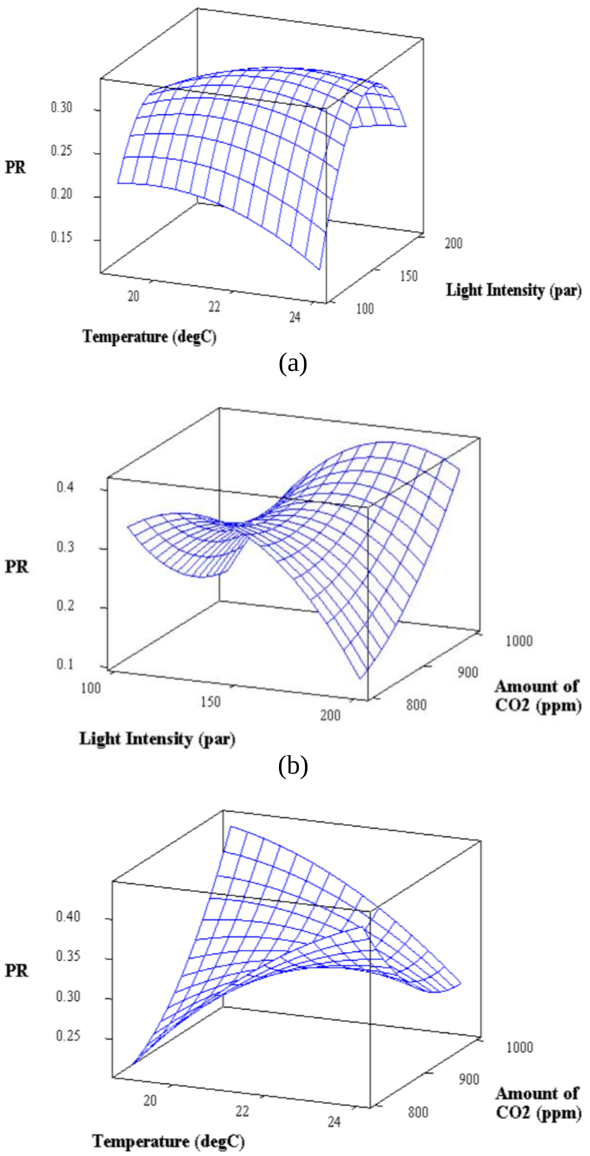

Figure 6: Plot showing (a) interaction of light intensity and temperature and its effect on photosynthetic rate, (b) interaction of light intensity and amount of CO_2 and its effect on photosynthetic rate, and (c) interaction of temperature and amount of CO_2 and its effect on photosynthetic rate.

The main objective of this study is to determine and evaluate the appropriateness of GA in finding the optimum conditions for the preharvest factors of a lettuce. GA was designed to produce the optimal values of light intensity, temperature and CO_2 level for the lettuce growth. After several explorations, the design that gives the optimal values for preharvest factors are selection rate = 0.2, crossover rate = 0.9 and mutation rate = 0.01. The best fitness plot showing the average costs per population and the best individuals are shown in Figure 7. In this plot, the fitness value for the best individual in each population is approximately 800 and the population's average fitness values is between 600 and 800.

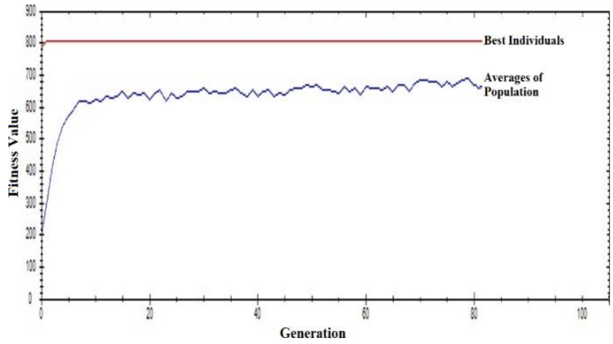

Figure 7: Best fitness plot

Table 1 shows the optimum values obtained using GA for each pre-harvest factor. These values are acceptable because it is between their tolerable ranges. The range of values that can give high photosynthetic rate are the following: for light is between 100 to 200 μmol/m^2/s, for temperature is between 19 to 24 ºC and for CO_2 is between 800 to 1000 ppm [18][19].

Table 1
Results of GA

Preharvest Factor	Optimum Value
Light Intensity	175.22296
Temperature	19.36228
CO_2	803.01855

IV. Conclusion

This study presented the suitability of applying the genetic algorithm in determining the optimum values of light intensity, temperature and CO_2. These factors are needed to be determined because it greatly affects the lettuce growth. Through graphical analysis, the optimum values cannot be determined. The graphs show varying effects of the parameters on photosynthetic rate. That is why GA has been applied in this study. Critical design on GA was done and the best fitness plot was obtained. The GA results showed that the optimum conditions for a high-quality lettuce crop needs a light intensity of 175.22296 μmol/m^2/s, a temperature of 19.36228 ºC and a CO_2 level of 803.01855 ppm.

Acknowledgment

The authors would like to thank the Department of Science and Technology (DOST) – Engineering Research and Development for Technology (ERDT) for funding this study.

References

[1] Thornley, J. (2002). Instanteneous canopy photosynthesis: analytical expression for sun and shade leaves based on exponential light decay down the canopy and an acclimated non-rectangular hyperbola for leaf photosynthesis. *Annals of Botany*, 451-458.

[2] Wang, Y., & Leuning, R. (1998). A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: model description and comparison with a multi-layered model. *Agriculture and Forest Meterorology*, 89-111.

[3] Kaur, B., & Owusu, R. K. (2015). Inverse problems and data fusion for crop production applications targeting optimal growth - fertilization.

[4] Pal, B. B., Chakraborti, D., & Biswas, P. (2009). A genetic algorithm based hybrid goal programming approach to land allocation problem for optimal cropping plan in agricultural system.

[5] Zheng, Z., & Wang, Y. (2016). Research on the relationship among the growth period environmental factors of tomato under the condition of mulched drip irrigation in greenhouse.

[6] Liu, J., Huffman, T., & Shang, J. (2016). Estimation of crop yield in regions with mixed crops using different cropland masks and time-series MODIS data.

[7] Bureerat, S., & Cooper, J. (1998). Evolutionary methods for the optimizations of engineering systems. IEEE *Colloquium on Optimisation and Control: Methods and Application* (pp. 1-10). IET.

[8] Chaturvedi, S., Pragya, P., & Verma, D. H. (2015). Comparative analysis of particle swarm optimization, genetic algorithm and krill herd algorithm. IEEE *International Conference on Computer, Communication and Control* (pp. 1-7). IEEE.

[9] Fidrysiak, B. A., & Przewozniczek, M. W. (n.d.). Towards finding an effective way of discrete problems solving: the particle swarm optimization, Genetic Algorithm and Linkage Learning Techniques Hybridization. (pp. 1-9). IEEE.

[10] Bijay Baran Pal, D. C. (2009). a genetic algorithm based hybrid goal programming approach to land allocation problem for optimal cropping plan in agricultural system.

[11] T. Morimoto, T. t. (1993). Growth optimization of plant by means of the hybrid system of genetic algorithm and neural network.

[12] Haupt, R. L., & Haupt, S. E. (2004). *Practical genetic algorithm.* New Jersey: John Wiley and Sons, Inc.

[13] Mitchell, M. (1999). *An introduction to genetic algorithms.* London: MIT Press.

[14] Sivanandam, S., & Deepa, S. (2008). *Introduction to genetic algorithms.* New York: Springer.

[15] Kuncheva, L., & Jain, L. (2000). Designing classifier fusion system by genetic algorithm. *IEEE Transactions on Evolutionary Computation*, 351-373.

[16] Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification.* New York: John Wiley and Sons, Inc.

[17] Muhlenbein, & Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm. *IEEE Trans. on Evolutionary Computation*, 25-49

[18] Brechner, M., & Both, A. (n.d.). *Hydrophonic lettuce handbook.* Cornell University Press.

[19] Vriend, R. (2014). *Lettuce Production Guidelines.* Enkhuizen, Netherlands: Enza Zaden Export B.V

[20] Bedruz, R. et al. (2016). Philippine vehicle plate recognition using image thresholding and genetic algorithm. *IEEE Region 10 Conference.* IEEE.

[21] Uy, A.C.P. et al. (2015). Machine vision for traffic violation detection system through genetic algorithm. *International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE.

[22] Quiros, A. R. F. et al. (2015). A genetic algorithm and artificial neural network based approach for the machine vision of plate segementation and character recognition. *International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE.

[23] Uy, A.C.P. et al. (2016). Automated traffic violation apprehension system using genetic algorithm and artificial neural network. *IEEE Region 10 Conference.* IEEE.

[24] Baldovino, R.G.B. et al. (2017). GA Optimization of coconut sugar cooking process: a preliminary study using stochastic universal sampling (SUS) technique. *2017 9th ICCAE*.