A SPECIES COMPETE-DIE OUT (SCD) ALGORITHM MODEL FOR IMPROVING THE PERFORMANCES OF EVOLUTIONARY COMPUTATION IN GREENHOUSE

MENG YAO[1], CHUN-NI DAI[1], MING LI[1], ZHU-JIE XIE[2], CHUN-HONG CHEN[2]

[1]School of Information Science & Technology, East China Normal University Shanghai 200026, PRC
[2]Shanghai Key Lab of Protected Horticultural Technology, Shanghai 201106, PRC
E-MAIL: {myao, mli}@ee.ecnu.edu.cn, ming_lihk@yahoo.com

Abstract:

Precision agriculture and quality one have gained attentions in modern research of agriculture due to the importance of crops to human being [19-23]. Owing to the fact that there are many uncertain factors for crop systems [24], the research of crops in greenhouses plays a role in this regard and evolution algorithms are paid attention to, see e.g. [25]. This paper aims at introducing a new model with respect to species compete-die out algorithms (SCD) for evolutionary computation. It can be used to release the limitation of evolutionary computation of premature convergence so as to optimize the parameters of the dry matter production and distribution simulation models of cucumber in glasshouse. The testing results based on the present model are compared to that based on the simple evolutionary algorithms under the same situations, exhibiting that it has advantages in stability and astringency.

Keywords:

Species compete-die out algorithms; evolutionary algorithms; premature convergence; optimization; greenhouse

1.   Introduction

Control issue is tough in greenhouse due to its nature in non-linearity and multi-variables [1,2]. Methods to deal with optimal control problems are rich, such as gradient method [3], simulated annealing [4,5], and so on [6,16,17]. These methods have heuristic qualities and exhibit respective advantages but they might be not enough to solve the hard problem of optimal control in greenhouse [5].

Conventionally, evolutionary algorithms (EA) or genetic algorithms (GA) are used for solving this problem because of their two remarkable qualities: implicit parallelism and global search [7-12]. Nevertheless, EA or GA has an inherent limitation in that it is relatively easy for the solutions to fall into premature convergence [5,13,14,18]. In fact, the basic cause of such premature of EA or GA comes from the explicit fact that selection pressure makes high fitness individuals reproduce quickly and thus supplant low-fitness individuals, some of which may be more promising, but not yet fully exploited [13,14]. For this reason, we propose a species compete-die out (SCD) algorithm model for evolutionary computation.

The present model applies clustering bases on niching method [15] and simulates the phenomena in nature and society: the social and history continually advance forward with the force of combination of independent evolution in every species and overall competition among all species. Furthermore, in order to improve the performance of local search and accelerate convergence efficiency, the SCD model adaptively adjusts fitness function and mutation step in the process of search. Our experiments regarding SCD truly reach global solutions, exhibiting that SCD is apparently superior to simple EA in stability and astringency.

The rest of paper is organized as follows. Section 2 introduces the research background, including the growth models of cucumber in greenhouse. Section 3 elaborates our model designed to optimize the parameters in the growth models. Section 4 evaluates the present model by comparing the result we got with that of simple EA. Conclusions are given in Section 5.

2.   Research background

2.1.   The growth models of cucumber in greenhouse

The daily growth speed of cucumber can be calculated by

$$GCR = \frac{[0.682DTGA - RM(T_O)Q_{10}^{(T_{av}-T_o)/10}](1-a)}{CR_{lv}CP_{lv} + CR_{st}P_{st} + CR_{rt}CP_{rt} + CR_{so}CP_{so}} \quad (1)$$

where W_{lv}, W_{st}, W_{rt}, and W_{so} are corre sponding weighs of biomass of leaf, stem, root and storage organ, respectively, Q_{10} is the temperature coefficient of respiration, T_{av} is the daily average temperature, T_o is the standard temperature 25°, α is the content number of mineral and other element in plant (in this work, it equals to 5%), $RM(T_0)$ is the ratio of maintain respiration when the temperature is 25° and can be calculated by

$$RM(T_o) = mc_{lv} \times W_{lv} + mc_{st} \times W_{st} + mc_{rt} \times W_{rt} + mc_{so} \times W \quad (2)$$

where mc_{lv}, mc_{st}, mc_{rt} and mc_{so} are maintain respiration coefficients of leaf, stem, root and storage organ, respectively, $DTGA$ is the daily total photosynthetic rate of the whole canopy and can be counted by

$$DTGA = (\sum_{i=1}^{3} \sum_{j=i}^{5} P[i][j] \times WGUSS[j] \times LAI \times WGUSS[i]) \times DL$$

$$(3)$$

where DL is the day length, LA I is the leaf area index of canopy, $WGUSS[i]$ and $WGUSS[j]$ are the Gauss weighted coefficients, $P[i][j]$ is the photosynthetic rate of a single leaf and can be described by

$$P[i][j] = AMAX \times [1 - EXP(-EFF \times I_{l,a}[i][j] / AMAX)]$$

$$(4)$$

where $AMAX$ is the photosynthetic rate when the intensity of light is saturated and equals to 44, EFF is the original using efficiency of absorbed light and equals to 0.015, $I_{l,a}[i][j]$ is the instantaneous PAR and can be described as follow

$$I_{l,a}[i][j] = k\rho[i]I_o[i]\exp(-kDIS[j]LAI) \quad (5)$$

where

$$I_o[i] = 0.5Q \times SINB[i](1 + 0.4SINB[i]) / DSINBE$$

$$(6)$$

$$\rho[i] = 1 - 2(1 - \sqrt{1-\sigma}) / [(1 + \sqrt{1+\sigma})(1 + 2SINB[i])]$$

$$(7)$$

where $I_o[i]$ is the instantaneous PAR on top of coronal, $\rho[i]$ is the reflectivity of canopy, k is the extinction coefficient of canopy and is equal to 0.7, $DIS[i]$ is the Gauss distance, LAI is the leaf area index, Q is the daily total solar radiation, $DSINBE$ is the efficient proof of the angle when it is relatively small, σ is the diffusion coefficient of the visible light of a single leaf and equals to 0.2, $SINB[i]$ is the sine function of solar height angle and is equal to
$SINB$ $[i] =$

$$\max\{0, SSIN + CCOS \cdot \cos[\frac{\pi(24 + 0.5DL \cdot DIS[i])}{12}]\}$$

$$(8)$$

where DL is the day length, $DIS[i]$ is the Gauss distance.

2.2. Experimental introduction

In our experiments the object function is described shortly by $GCR = y = f(X, \Psi)$, where X is the vector of independent variable: $X = [x^{(1)}, x^{(2)}, ..., x^{(n)}]^T$, Ψ is the parameter vector: $\Psi = [\sigma, K, AMAX, EFF, Q_{10}]^T$. Each parameter in this vector is coded by floating number to represent a bit of the chromosome. Thus the original fitness function can be calculated by:

$$F(\Psi) = \min \sum_{i=1}^{n} |y_i - f(X_i, \Psi)| \quad (9)$$

where n is the total number of independent variable.

3. The spices compete-die out algorithms model

3.1. Evolutionary phenomena in nature and society

At the very beginning of evolution our ancestors all lived within their tribes and the competition and evolution was taken place only in the inner, so every tribe can develop independently. The result was the development between tribes was rather uneven. In order to take the limit resources in nature, the strong spices used to commit aggression against the weak. Then some tribes generally died out, some were assimilated, some were getting more and more stronger, which made the whole human history advancing continually. Also, animals have the same history as human being.

It is rather important for the ancient men to be left in an unattached space to develop themselves before attending the competition of the whole society. Once they have no opportunity to grow fully, some of which, whose fitness are usually lower, will die out easily. In this case, the nature will lose many categories of genotype and the diversity cannot be maintained availably. So in our new model the thought of niche is used to protect all kinds of genotype of individuals.

3.2. The spices compete-die out algorithms model

Inspired by the evolutionary phenomena and the hierarchical fair competitions in nature and society, we propose the spices compete-die out algorithms model (SCD) for evolutionary algorithms. Simple evolutionary algorithms (sEA) usually result in premature convergence or just arriving local optimal solution. The main reason is

that the selection pressure often makes the high-fitness individuals reproduce largely and then result to a high proportion of individuals with similar genes. But they are not surely the best in the whole space of the search.

So in our model, before the start of evolution all individuals are divided into some species according to the fitness of the first generation individuals. At the first several generations, evolutions only take place in the inner of species, not between the species, which can provide every individual some opportunity to develop fully without outer disturbance. After several generations of inner evolutions, competitions between different species begin, which results in to the final evolution and development of the whole species. The structure of our model is showed in Figure 1.
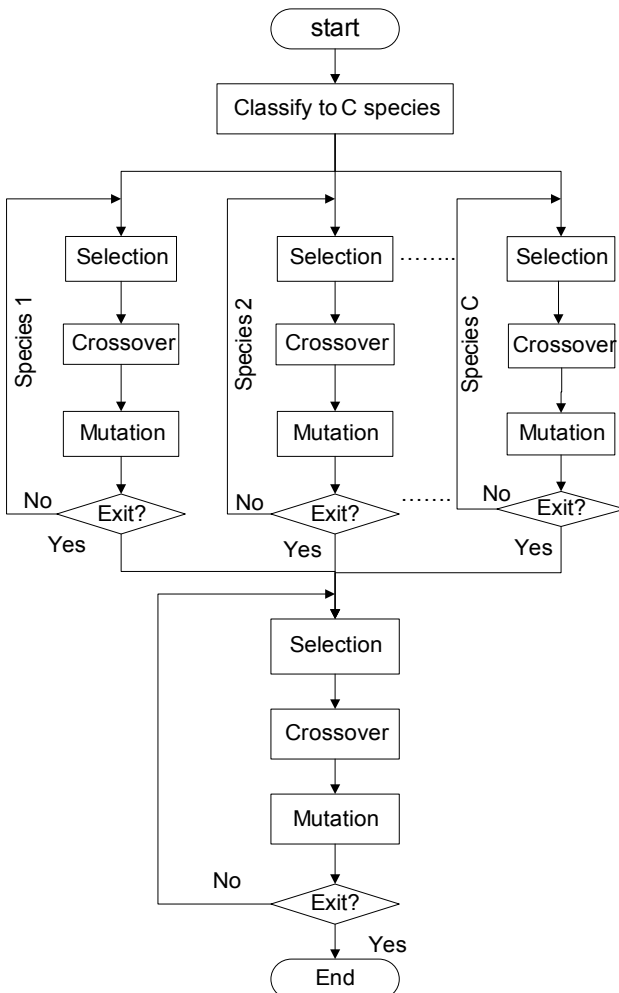

Figure 1. Structure of the SCD model

It is important for individuals to evolve independently

in the inner species at the start of evolution because unattached development can ensure excellent genotypes in different levels of fitness can develop well and be inherited fully. In this case the apparent low-fitness but maybe promising individuals can be protected effectively and thus the diversity of revolution can be maintained. Once the unlimited expand of some good genotypes at the beginning can be greatly avoided, so can the premature convergence.

The method of classification is inspired by the phenomenon of hierarchical fair competition often seen in societal and biological systems. According to the original fitness, all individuals are divided into a number of species to have unattached development. The individuals having close fitness values are divided into the same class. Such a kind of hierarchical fair competition can effectively protect individuals with different level of fitness in one population and maintain the diversity of genotype.

Chromosome structure of individual V_i in SCD is described as follows

$$[U_{\min}^{(1)}, U_{\max}^{(1)}] \ [U_{\min}^{(2)}, U_{\max}^{(2)}] \ldots [U_{\min}^{(n)}, U_{\max}^{(n)}]$$

σ_{i1}	σ_{i2}	$\cdots\cdots$	σ_{in}

where σ_{ij} represents the jth bit of chromosome of individual V_i, n is the dimension of V_i, $[U_{\min}, U_{\max}]$ is the range of value of σ_{ij}. We employ floating code in SCD to take advantage of direct representation of genotype of chromosome and avoiding Hamming distance. In addition, in the use of floating code we can make use of fine-tuning to effectively improve the local search of optimization.

The choice of selection strategy of evolutionary algorithms is rather important. In the new model we employ two different strategies. When evolution happened in inner species, we choose the ranking selection, which can gives more selective pressure towards the optimum when the fitness of the population are similar. When individuals of every species all take part in the competition and evolution, we choose the conservation of some best strategy and the $\mu + \lambda$ selection strategy. The combination of the two strategies can avoid being destroyed by the operator of crossover and mutation and thus can save the excellent genotype to maintain the diversity through out the whole finding of the solution.

In order to improve the performance of SCD, we modify the original fitness function to indirectly adjust the selection pressure

$$F_i'(\hat{\Psi}) = \frac{F_i(\hat{\Psi})}{D_i S_i} \qquad (10)$$

In the above expression,

$$S_i = \frac{1}{\sum_{\forall j \in \Gamma k} l_{i,j}} = \frac{1}{\sum_{\forall j \in \Gamma k} \sum_{s=1}^{n} \frac{|\sigma_{is} - \sigma_{js}|}{U_{\max}^{(s)} - U_{\min}^{(s)}}} \qquad (11)$$

$$D_i = \sum_{\forall j \in \Gamma k} \exp[-\frac{(\sum_{s=1}^{n} \frac{|\sigma_{is} - \sigma_{js}|}{U_{\max}^{(s)} - U_{\min}^{(s)}})^2}{(\alpha/2)^2}] \qquad (12)$$

where Γ_k represents the kth species, V_i is an individual in Γ_k, V_j represents the other individuals in Γ_k, n denotes then dimension of vector V_i, D_i is the density function of V_i, α is the radius. It can be concluded that the bigger D_i is, the more individuals there are in the range of α. In this case because the chromosome of the nearby individuals is rather similar with individuals V_i, the fitness of V_i then should be reduced correspondingly. By this way we can not only protect the diversity of chromosome in the process of search but also accelerate the speed of constringency.

In SCD model, we put forward the adjustment of the mutation step called local fine-tuning. In order to save the excellent genotype locating in the promising area or next to the global solution, we reduce the mutation step when searching in the promising area to improve the final global solution. The method is as follows

$$\sigma_{ik}' = \begin{cases} \sigma_{ik} + b[1 - r \cdot (1 - \frac{F(V_i)}{F_{\max}})^\lambda] & rnd(2) = 0 \\ \sigma_{ik} - a[1 - r \cdot (1 - \frac{F(V_i)}{F_{\max}})^\lambda] & rnd(2) = 1 \end{cases} \qquad (13)$$

where σ_{ik} represent the kth mutating loca of the individual V_i, $[U_{\min}^{(k)}, U_{\max}^{(k)}]$ is the range of σ_{ik}, $a = \sigma_{ik} - U_{\min}^{(k)}$, $b = U_{\max}^{(k)} - \sigma_{ik}$, r is a stochastic numeral between zero and one, λ is a parameter between two and five, $F(V_i)$ is the fitness of V_i and $F_{\max}$ is the maximal fitness.

4. Validation and analysis

We used SCD model and sEA respectively on the growth models of cucumber in greenhouse to optimize the five parameters: σ, $K[i]$, *AMAX* and *EFF*, Q_{10}. The experiment results are listed in Table 1. In this table, mean denotes the average value of corresponding parameter in our experiments. The maximum generation of SCD and sEA in our test is 200 generations.

From the above table, we can see that the SCD is superior to sEA in stability because the final standard

variance of each parameter in SCD is far less than that of sEA. In our tests when sEA evolved at the fortieth generation, the fitness couldn't, in fact, be improved, which indicated that the arithmetic was in stagnation and probably in premature convergence of local optimum. On the other hand, the spending time of simple EA was about fifty-eight seconds, while SCD about two hundred and ten seconds, which could be explained by the fact that SCD has introduced some complex operators such as the classification according to the fitness and the local fine-tuning. Actually the SCD model has better convergence efficiency and stability on the cost of more time spending. The convergence of SCD is showed as Figure 2 and sEA as Figure 3.
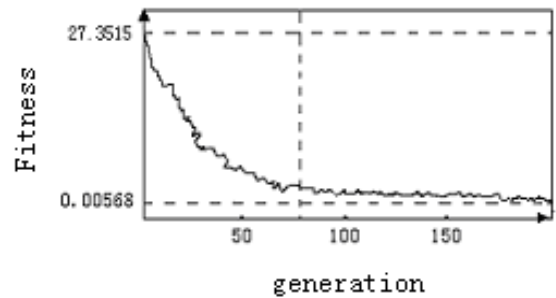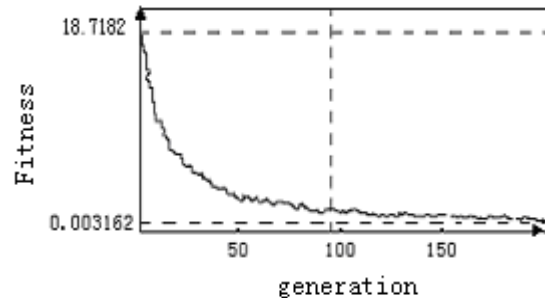

Figure 2. Convergence of sEA in the first experiment

Figure 3. Convergence of SCD in the first experiment

5. Conclusions

We have explained our SCD model and applied it to optimize some parameters of the growth model of cucumber. Validation of the present model has been demonstrated by comparison with simple EA. The results exhibit that it has considerable advantages in stability and convergence to resolve optimal problem in greenhouse.

Table 1. Experimental results

Parameter	Theoretic value	SCD		sEC	
		Mean	Variance	Mean	Variance
σ	0.2	0.2003874	0.0013998	0.2166730	0.0492665
K	0.7	0.6999238	0.0014915	0.6490705	0.1522802
$AMAX$	44	43.839866	0.0612765	40.325800	6.5379088
EFF	0.015	0.0149999	0.0000144	0.0154443	0.0016266
Q_{10}	1.4	1.3986007	0.0089206	1.3604547	0.1191753

References

[1] K. Metselaar, "Auditing predictive models: a case of study in crop growth", PhD thesis, Wageningen agricultural university, The Netherlands, 1999.

[2] I. Seginer, G. Straten and F. Buwalda, "Nitrate concentration in greenhouse lettuce: a modelling study", Acta Hort. (ISHS) 456, pp.189-198, 1998.

[3] J. D. Faco, "A Nonlinear Conjugate reduced-gradient method for solving bounded nonlinear optimal control problems on parallel computers", http://dmawww.epfl.ch/roso.mosaic/ismp97/ismp_abs_1152.html

[4] E. Aarts and J. Korst, "Simulated annealing and boltzmann machine", John Wiley and Sons, 1989.

[5] D. Adler, "Genetic algorithm and simulated annealing: a marriage proposal", Proc., IEEE Int. Conf. On Neural Networks, pp. 1104-1109, 1993.

[6] W. E. Bosarge, O. G. Johnson, "Numerical properties of the Ritz-Trefftz algorithm for optimal control", Communications of the ACM, Vol. 14, No. 6, pp. 402-406, 1971.

[7] T. Krink, R. K. Ursem, and B. Filipic, "Evolutionary algorithms in control optimization: the greenhouse problem", Proc., the Genetic and Evolutionary Computation, Conference GECCO-2001, San Francisco, California, 2001.

[8] H. Pohlheim and A. Heiβner, "Optimal control of greenhouse climate using a short time climate model and evolutionary algorithms", Proc., 3rd IFAC/ISHS Workshop on "Mathematical and Control Applications in Agriculture and Horticulture", pp. 113-118, 1997.

[9] H. Pohlheim and A.Heiβner, "Optimal control of greenhouse climate using real-world weather data and evolutionary algorithms", GECCO'99-Proc., the Genetic and Evolutionary Computation Conference, San Francisco, CA, pp. 1672-1677, 1999.

[10] H. Pohlheim and A.Heiβner, "Optimal control of greenhouse climate using genetic algorithms", MENDEL'96 2nd Int. Conf. on Genetic Algorithms, Technical University of Brno, Czech Republik, pp. 112-119, 1996.

[11] R. K. Ursem, B. Filipic and T. Krink, "Exploring the performance of an evolutionary algorithm for greenhouse control", Proc., 24th Int. Conf. on Information Technology Interfaces, pp. 429-434, 2002.

[12] S. Mardle and S. Pascoe, "An overview of genetic algorithms for the solution of optimization problems", http://econltsn.ilrt.bris.ac.uk/cheer/ch13_1/ch13_1p16.htm

[13] J. J. Hu and E. Goodman, "The hierarchical fair competition (HFC) model for parallel evolutionary algorithms", Proc., the 2002 Congress on Evolutionary Computation: IEEE CEC2002, Honolulu, Hawaii, 2002.

[14] J. J. Hu, E. D. Gooddman, K. Seo, and M. Pei, "Adaptive hierarchical fair competition (ahfc) model for parallel evolutionary algorithms", Proc., the Genetic and Evolutionary Computation Conference, GECCO 2002, New York, Morgan Kaufmann Publishers, pp. 772-779, 2002.

[15] F. Streichert, G. Stein, H. Ulmer and A. Zell, "A clustering based niching method for evolutionary algorithms", http://www-ra.informatik.uni-tuebingen.de/publikationen/2003/streichert03aclustering.pdf.

[16] M. Li, "An iteration method to adjusting random loading for a laboratory fatigue test", International Journal of Fatigue, Vol. 27, No. 7, pp. 783-789, July 2005.

[17] M. Li, "An optimal controller of an irregular wave maker", Applied Mathematical Modelling, Vol. 29, No.1, pp. 55-63, Jan. 2005.

[18] M. Yao, H. Y. Meng, and et al, "Towards improvement in locating of underground tomb relics using EM

radar", Signals and Genetic Algorithm 2001 Genetic and Evolutionary Computation Conference, pp. 493-498, July 2001.

[19] J. Blahovec and M. Kutilek, "Physical Methods in Agriculture: Approaches to Precision and Quality", Kluwer Academic Publishers, New York, 2002.

[20] D. K. Benbi and R. Nieder, Eds., "Handbook of Processes and Modeling in the Soil-Plant System", The Haworth Press, New York, 2003.

[21] K. D. Vohnout, Ed., "Mathematical Modeling for System Analysis in Agricultural Research", Elsevier Science, Amsterdam, The Netherlands, 2003.

[22] K. R. Reddy and H. F. Hodges, "Climate Change and Global Crop Productivity", CABI Publishing, Wallingford, UK, 2000.

[23] H. D. Groote and O. Traore, "The cost of accuracy in crop area estimation", Agricultural Systems, Vol. 84, No. 1, pp. 21-38, April 2005.

[24] A. Shrestha, Ed., "Cropping Systems: Trends and Advances", Food Products Press, Binghampton, NY, USA, 2003.

[25] D. G. Mayer, B. P. Kinghorn, and A. A. Archer, "Differential evolution–an easy and efficient evolutionary algorithm for model optimization", Agricultural Systems, Vol. 83, No. 3, pp. 315-328, Mar. 2005.