26

Evolutionary Algorithms in Modelling of Biosystems

Rosario Guzman-Cruz, Rodrigo Castañeda-Miranda, Juan García-Escalante, Luis Solis-Sanchez, Daniel Alaniz-Lumbreras, Joshua Mendoza-Jasso, Alfredo Lara-Herrera, Gerardo Ornelas-Vargas, Efrén Gonzalez-Ramirez and Ricardo Montoya-Zamora
Universidad Autonoma de Zacatecas, Universidad Politécnica del Sur de Zacatecas
México

1. Introduction

Globally there is a land area of 13000 million ha, of which the area used for agricultural production is 1500 million ha (Naciones Unidas sobre el Desarrollo de los Recursos Hídricos en el mundo [NUDRHM], 2006), that is 12% of the total surface. Because the productivity of the agricultural sector is reduced, in general, the intensive production in controlled environments is an alternative to increase productivity in the aforementioned sector. For this reason, the use of greenhouses for agricultural production has increased in recent years. In the world, currently some 265,000 hectares are cultivated in greenhouses. Asia, with more than 138,000, represents at this moments the leading world power in intensive production under plastic cover. In the second place, with 95,000 ha, the Mediterranean basin is placed. Northern Europe (16,000 ha) and the American Continent (15,600 ha) share the rest. Holland remains the maximum exponent in agriculture under high-tech cover, especially so that which refers to crystal greenhouses.

The crop inside greenhouse allows to obtain productions of better quality and higher yields, at any time of the year, while it permits extending the cycle of farming, enabling production in the most difficult times of the year and getting better prices. Due to the benefits that provide crops inside greenhouse, crops should have a controlled microclimate using different systems to regulate environmental conditions, limiting the excesses and filling shortcomings in terms of crop needs, in order to take place in every moment the optimal conditions of different stages of plant development. In addition to the advantages that production has in controlled environments it is important to reduce production costs, such as heating costs, irrigation and fertilization. To achieve this it is necessary to have a system that controls the environment inside the greenhouse that allows improving production and crop quality, product quality, time of the production process and improving production costs. Moreover, the system should allow the farmer to have sufficient information about requirements and changes observed over time in crop development. These control systems are based on mathematical models that describe the behavior of variables, such as air temperature, the concentration of carbon dioxide (CO_2) and absolute humidity. Therefore, to improve the productivity of the crop under greenhouse it is necessary to study the process

of crop growth as well as in quantitative form moreover, the climatic environment in which the crop develops. The modeling allows having a quantitative form of the interaction of simultaneous processes (Lopez et al., 2006).

If the model is not too complex in terms of state variables it can be used to design control systems for example optimal control strategies (Seginer & Ioslovich, 1998a; Van Henten, 1994; Tap, 2000) for the best growth, production and crop quality.

Due to the advantages that production has in controlled environments, mathematical models have an essential role in the description of the main state variables of the system. Hence the importance of having a model that adequately describes the climatic conditions inside the greenhouse. That is, the problem consists of calibrating or fitting a model that describes the internal environment of the greenhouse, i.e., the problem consists in finding the set of parameters that make the difference between estimated values by the model and real values be minimal.

Then the problem of calibrating a model is reduced to a problem of search and optimization of parameters involved in the model which is formulated as a function of multivariable nonlinear optimization, but these types of problems can have multiple local optimal solutions. Some local optimization methods have been used to adjust the model parameters. However, it is necessary to use a method that allows escaping from local optima and finding the global optimum. So, for the optimization of the parameters it is possible to use a global search method such as evolutionary algorithms.

Evolutionary Algorithms (EA) techniques inspired by the theory of evolution, such as genetic algorithms (GA), Evolutionary Strategies (ES), Evolutionary Programming (EP) and Differential Evolution (DE) that are useful to solve the problem of optimization of the parameters of a mathematical model. Once the mathematical model is adjusted, the behavior of the main variables involved in the greenhouse environment can be known and, on the basis of this, take the appropriate decisions to operate the systems installed in it, such as the ventilation system, mesh shading, etc. to keep appropriate levels of temperature, humidity and CO_2 levels and in this way increase crop yields and improve final product quality.

2. Models for climate prediction

The crop production under greenhouse is influenced by the climate inside, so it has to keep state variables that characterize it within a certain range of values (CIDEIBER, 2005).

To study a model for the greenhouse production it should identify of the variables that interfere with the greenhouse climate dynamics such as state or control variables (indoor air temperature, soil temperature, internal CO_2 concentration, internal absolute humidity), external or disturbance variables (outside air temperature, external relative humidity, external CO_2 concentration, radiation, wind speed) and control variables: natural ventilation, heating). It is important to take into account that any part of the greenhouse can act as an individual body or volume. When the flow of physical quantities, such as, energy or mass is considered, the conservation law is valid for any volume.

For the prediction of environmental variables it is necessary to adjust the climate greenhouse model. Consequently, the methodology used for this adjustment is to a) defining the structure of a mathematical model that describes the climate inside a greenhouse, b) conduct a sensitivity analysis of parameters involved in the model to identify the parameters that affect more the state variables, c) making calibration of parameters identified using AE and finally e) doing the validation of the mentioned model, Figure 1.

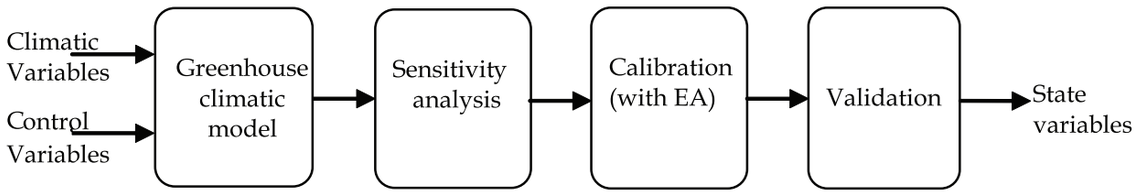


Fig. 1. Diagram of the development process for the greenhouse climate model in last years, empirical and mechanistic mathematical models have been developed to predict climate and crop growth under greenhouses.

In last years there have been developed empirical and mechanistic mathematical models to predict climate and crop growth under greenhouses.

The process of crop production in greenhouses is complex because it depends on crop growth and external climatic conditions and the design of the greenhouse. A tool to improve management of the greenhouses is mathematical models to predict the behavior of climate variables. There are complex simulation models with a relatively large number of state variables that describe a system; however, these cannot be implemented in control systems. For which it is required the use of simple models with the minimum possible number of variables for optimization and control systems. Thus, the model studied in this research is a model for optimization and control purposes.

Generally there are three types of mathematical models: white-box models, black box and gray box. White box models are deterministic and explanatory models of a system (Thornley & Johnson, 2000; France & Thornley, 2006). Typically, a set of ordinary differential equations is defined to describe the behavior of the system state variables, variables that represent the properties or attributes of the system that is under consideration. White box models are more appropriate to express hypotheses in mathematical form and thus provide a quantitative description and explanation of the most important processes occurring in a system. Black box models are direct descriptions of data and provide direct observable relationships between the variables in a system without any explanation of the underlying mechanisms. They are a powerful means to describe and summarize data. Grey box models are used when some physical understanding is available, but several parameters stay to determine the observed data.

2.1 White box models

The development of a white-box mathematical model requires sufficient understanding of the physical, chemical and biological processes that occur in a system and its use demands a proper validation. Explanatory models can be static or dynamic. The simulation of crop growth and development involves two processes construction of mathematical models and numerical solution of the set of equations that describe the behavior of the system, through the use of a digital computer.

Dynamic simulation models are based on the assumption that the state of a system can be quantified and that changes in the state can be described by mathematical equations, equations of rate of change or differential equations. A model of growth and development of crop includes several components: state variables, differential equations, parameters and inputs. Normally a state variable is a variable that appears in the accumulation term of a dynamic balance of mass or energy. A state variable is a variable that can be quantified (at least conceptually) and it allows knowing the behavior of the system at all future instant in

time. The differential equations represent the change velocity of the state variables (López Cruz, 2004). Crop growth or climate in a greenhouse is described by a dynamic model, represented by a nonlinear ordinary differential equation as follows:

$$\dot{x} = f(x,u,\theta,t), \qquad x(t_b) = \beta \tag{1}$$

Where $x \in \Re^n$ are the state variables, $u \in \Re^m$ are the control inputs, $\theta \in \Re^q$ are time-invariant parameters, β is the vector of initial conditions and t denotes the time (Van Henten & Van Straten, 1994).
From the 60's to date, models for crop growth have been developed. The most known results are the model of crop growth SUCROS (a Simple and Universal Crop Growth Simulator) and LINTUL (Light Interception and Utilization) which are generic models (Bouman et al, 1996) i.e., they can be adapted to any crop year (Hernández Hernández, 2009).
Moreover, in the last 30 years mechanistic mathematical models have been applied in modeling of greenhouse climate. Two of the first proposals were made in 1983 by Boy and Udink Ten Cate. Both models consider equations for the temperature of the greenhouse warming effect and opening windows.

2.2 Black box models

The problem of identifying a given system consists of given certain inputs, $u(t)$, and outputs, $y(t)$, of n dynamic system:

$$u^t(t) = \left[u(1),u(2),...,u(t)\right] \tag{2}$$

$$y^t(t) = \left[y(1),y(2),...,y(t)\right] \tag{3}$$

It is looked for the relationship between past observations $\left[u^{t-1},y^{t-1}\right]$ and future outputs $y(t)$:

$$y(x) = g\left(u^{t-1},y^{t-1}\right) + v(t) \tag{4}$$

The term $v(t)$ considers the fact that the output $y(x)$ will not be an exact function of past data, $v(t)$ is described as a random noise signal. However, a goal should be that $v(t)$ be as small as possible so that you can think that $g\left(u^{t-1},y^{t-1}\right)$ is a good prediction of $y(x)$ given past data. Eq. (4) models general dynamic systems in discrete time. Static systems can be viewed as particular cases of dynamical systems.
Now, the problem consists of finding a function g in (4). So, it has to look for a family of functions. This family of functions can be parameterized by a vector of parameters θ with finite dimension:

$$g\left(u^{t-1},y^{t-1},\theta\right) \tag{5}$$

Parameterize the function g with a vector θ of finite-dimension is usually an approximation. To find a good parameterization, it is necessary to decide on a structure and having a set of data collected $\left[u^N,y^N\right]$, the quality of θ can be evaluated by adjusting the model and the registered data:

$$\sum_{t=1}^{N}\left\|y(t)-g\left(u^{t-1},y^{t-1},\theta\right)\right\|^{2} \tag{6}$$

The standard and the real way to achieve or try to achieve minimum in θ may differ but many schemes of system identification follow this concept.
Now the family structure models (5) is too general and useful to write g as a concatenation of two mappings: one that takes the increasing number of past observations u^t, y^t and maps them in a vector $\varphi(t)$ of fixed dimension and one that maps this vector to space of outputs:

$$g\left(u^{t-1},y^{t-1},\theta\right)=g\left(\varphi(t),\theta\right) \tag{7}$$

where

$$\varphi(t)=\varphi\left(u^{t-1},y^{t-1}\right) \tag{8}$$

This vector is called vector regression and its components are called regressors. The regressor vector can be parameterized as:

$$\varphi(t)=\varphi\left(u^{t-1},y^{t-1}\right) \tag{9}$$

Which can be denoted as $\varphi(t,\eta)$. Sometimes $\eta=\theta$, that is the regression vector depends on all model parameters (Sjöberg, 1995).
Black box linear models for dynamic systems can be described with one $g(\cdot,\cdot)$ selected to be a linear mapping, $\theta^T\varphi(t)$. Regressor selection $\varphi(t)$ specifies whether the model is an ARX, ARMAX, a state space model, etc. Consider the same type of regressors but using a nonlinear mapping given black-box nonlinear models.
The selection of a nonlinear mapping (5) is separated into two partial problems for dynamic systems:

1. How to choose the regression vector $\varphi(t)$ of inputs and past outputs.
2. How to choose the nonlinear mapping $g(\varphi)$ from regressor space to output space

2.3 Gray box models

A gray box model corresponds to a combination of a black box model with a white box model. This model is also named semi-physique or based on knowledge, because in it all the knowledge of the process is introduced (or part of it) and additionally those unknown parameters are estimated through system measurements (Moreno and Acuña, 2005).

3. Evolutionary algorithms

As mentioned above, mathematical models have a set of parameters that need to be adjusted so that the data estimated by the model is more closely to the observed data. This problem of parameter estimation is also known as calibration parameters problem.
The task of calibrating a model is formulated as an optimization problem, so that different methods of solution have been proposed. Local or global methods of optimization can be used to solve the calibration problem. Some researchers have used local optimization

methods (Tap, 2000, Linker et al., 2004) or manual calibration (Van Henten, 1994) to adjust the parameters of a greenhouse climate model. A method for calibrating a mathematical model is using a nonlinear multivariable optimization function, but it is well known that these optimization problems may have optimal local solutions. Such problems are known as multimodal (Eiben & Smith, 2003). However, in recent years it has increased the use of global optimization methods to solve these kind of problems (Michalewicz, 1994) due to the advantages of getting global optimal possible solutions. For this reason, there is a need to develop algorithms based on this methodology to be applied to greenhouse climate. According to Eiben and Smith (2003), evolutionary algorithms (EA) could be an excellent alternative to provide a response to the challenge of achieving automated solution methods for problems more complicated and faster. EA are stochastic search methods that include genetic algorithms (GA), Evolutionary Strategies (ES), Evolutionary Programming (EP) and Differential Evolution (DE) (Michalewicz, 1994).

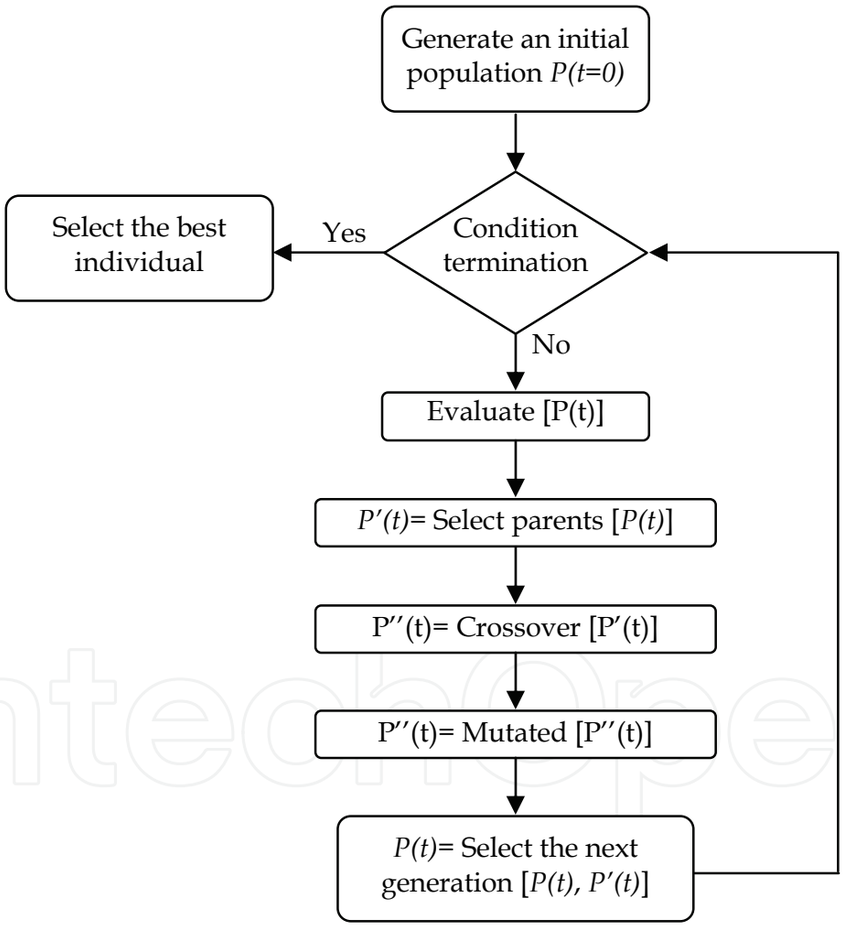


Fig. 2. Diagram of a general scheme of an evolutionary algorithm.

The structure of any AE is the same (Eiben & Smith, 2003), Fig. 3.3. The differences between evolutionary techniques consist of the type of selection operator, mutation and crossover applied to find the optimal value of the parameters to calibrate.

In general, it takes five basic components to implement an EA that solves a any problem (Michalewicz, 1992):

1. A representation of potential solutions to the problem.
2. One way to create an initial population of potential solutions (this is usually done randomly, but it can also be used deterministic methods).
3. An evaluation function that plays the role of the environment, qualifying the solutions produced in terms of its "fitness."
4. Genetic operators that alter the composition of the offspring (normally using the crossing and mutation).
5. Values for the various parameters used by the genetic algorithm (population size, cross and mutation probability, maximum number of generations, etc.).

3.1 Genetic algorithms

The traditional representation used in GA to encode a set of solutions that is the binary scheme, i.e. it is a string formed by zeros and ones (Eiben and Smith, 2003). However, to solve the problem of setting parameters of a greenhouse climate model to a series of real values is used to represent a candidate solution to a problem, i.e. a vector $(p_1,...,p_n)$ where $p_i \in R$ and n is the number of parameters that are needed to be calibrated.

3.2 Parental choice

N is the population size P. The parent selection is determined by tournament (Coello, 2007) which consists of:

1. Shuffle to the individuals in the population.
2. Choose k individuals in the population and comparing them with the basis of their adaptability (typically k = 2).
3. The fittest individual is the winner of the tournament. "
4. The population is shuffled k times to select N parents

3.3 Crossing

The cross is an operator that forms a new individual by combining parts of each parent. In this case, the type of employed crossing is crossbreeding in two points where both parents are employed to generate a new individual and crossover points are chosen randomly (Coello, 2007). Be $p,\ p' \in P$ and $i, j \in \{1,2,...,n\}$ random crossing points, so parents.

$$p = \left(p_1,...,p_{i-i},p_i,p_{i+1},...,p_{j-1},p_j,p_{j+1},...,p_n\right) \tag{10a}$$

and

$$p' = \left(p'_1,...,p'_{i-i},p'_i,p'_{i+1},...,p'_{j-1},p'_j,p'_{j+1},...,p'_n\right) \tag{10b}$$

Generate the children

$$h = \left(p_1,...,p_{i-i},p_i,p'_{i+1},...,p'_{j-1},p_j,p_{j+1},...,p_n\right) \tag{11a}$$

And

$$h' = \left(p'_1,...,p'_{i-i},p'_i,p_{i+1},...,p_{j-1},p'_j,p'_{j+1},...,p'_n\right) \tag{11b}$$

3.4 Mutation

The type of mutation is not applied uniformly (Michalewicz, 1996), where the position of the individual that will be altered is selected randomly. Given $p=(p_1,...,p_{,n})$, this is changed by the individual mutant $p=(p_1,...,p_i',...,p_s)$ where

$$p_i' = \begin{cases} p_i + \Delta(t, U - p_i), & \text{si } R = 0 \\ p_i - \Delta(t, p_i - L), & \text{si } R = 1 \end{cases} \tag{12}$$

$p_i \in [L,U]$, $R \in \{0,1\}$ are chosen randomly and

$$\Delta(t,y) = y * \left(1 - r^{\left(1-\frac{t}{T}\right)^b} \right) \tag{13}$$

Where $r \in [0,1]$ is a random, T is the maximum number of generations and b is the degree of non-uniformity of the mutation (Michalewicz use b = 5).

3.5 Survival selection

The survival selection is used in the replacement based on age, that is, on each generation parents are replaced by children who now represent the current population.

3.6 Evolutionary strategies

The type of representation used to represent a candidate solution for a problem is real.

3.7 Parental choice

Parents selection is random with uniform distribution of the population of N individuals.

3.8 Crossing

The type of crossing used is discrete where each element of the generated child is obtained randomly from the elements of parents with an equal probability for both parents. That is, given $p,\ p' \in P$ where $p=(p_1,...,p_n)$ and $p'=(p'_1,...,p'_n)$, a child $h=(h_1,...,h_n)$ is created with

$$h_i = \begin{cases} p_i & \text{si } R = 0 \\ p'_i & \text{si } R = 1 \end{cases} \tag{14}$$

where $i \in \{1,...,n\}$ and $R \in \{0,1\}$ are chosen randomly

3.9 Mutation

In this case, the mutation is not correlated with n step sizes is used which consist on a given $p=(p_1,...,p_n)$, p is extended in n step sizes thus resulting $p=(p_1,...,p_n,\sigma_1,...,\sigma_n)$. Then, the mutated individual is $p'=(p'_1,...,p'_n)$ where

$$\begin{aligned} p_i' &= p_i + \sigma_i' \cdot N_i(0,1), \\ \sigma_i' &= \sigma_i \cdot e^{\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)} \end{aligned} \tag{15}$$

with $\tau' = \frac{1}{\sqrt{2n}}$ and $\tau = \frac{1}{\sqrt{2\sqrt{n}}}$

3.10 Survival selection

Let N be the number of elements of the population and M the number of children generated. The selection of survival used is deterministic of type (N, M) in which after creating M children and calculating its adaptability, the N best are chosen to move to the next generation (Eiben & Smith, 2003).

3.11 Evolutionary programming

The type of representation used to represent a candidate solution to a problem is real.

3.12 Parental choice

Deterministic selection is used in which each parent generates exactly one child via mutation.

3.13 Crossing

Evolutionary programming is a technique in which there is no crossing between individuals of the population.

3.14 Mutation

A Gaussian perturbation is applied. That is, given $p = (p_1,...,p_n,\sigma_1,...,\sigma_n)$ is transformed into $p' = (p'_1,...,p'_n,\sigma'_1,...,\sigma'_n)$ where

$$\begin{aligned} p_i' &= p_i + \sigma_i' \cdot N_i(0,1), \\ \sigma_i' &= \sigma_i \cdot (1 + \alpha \cdot N(0,1)) \end{aligned} \tag{16}$$

With $\alpha \approx 0.2$.

3.15 Survival selection

The survival selection is deterministic of type (N + N) in which competition between each pair is through tournament after creating N children of the N individuals in the population. Therefore, let P 'represents the total population that includes parents and children, then each individual $p \in P'$ is compared with other q individuals selected randomly. In each comparison, if p is better than its opponent a gain is assigned to p. The N individuals with the greatest values of gain are selected to be parents of the next generation. Typically, q = 10 is recommended (Eiben & Smith, 2003).

3.16 Differential evolution

Representation of a candidate solution to a problem in ED is of real type. Thus that to solve the problem of setting parameters of a greenhouse climate model a series of real values is used to represent a candidate solution for a problem, i.e. a vector $p_j = (p_{j1},...,p_{jn})$ where j = 1, ..., N, $p_i \in R$ N is the number of elements of the population and n is the number of parameters that need to be calibrated (Price and Storn, 2005).

The initial population is chosen randomly if nothing is known about the problem. As a rule, it is assumed a uniform distribution for random decision making. In ED different strategies can be adopted (Price and Storn, 2005). The main idea behind the ED is a new scheme to generate vectors. The ED generates these new vectors when adding the weight difference between two members of the population vectors to a third vector member. If the fitness of the resulting vector is less than the population member chosen then the new vector replaces the vector with which it was compared. This vector to be compared can be (though not necessarily is) part of the generation process. In addition, the best individual is evaluated in each generation G for keeping track of progress during which it is done minimization.
There are different variations in ED, in this case, it is applied the classic strategy ED/random/1/bin in the calibration problem.

3.17 Selection

The selection of vectors to disturb is random.

3.18 Mutation

The type of mutation is randomly applied which consists of generating N vectors of the form:

$$v = p_1 + F \cdot (p_2 - p_3) \tag{17}$$

where $p_1, p_2, p_3 \in \mathrm{P}$ are different, $F \in [0,2]$ is a real constant factor that controls the amplification of the differential variation between the vectors p_2 y p_3 .

3.19 Crossing

The crossing applied is binomial type which consists of combining the previously mutated vector $v = (v_1, v_2, ..., v_n)$ with a target vector $p = (p_1, p_2, ..., p_n) \in \mathrm{P}$ to generate a test vector $p' = (p'_1, p'_2, ..., p'_n)$ where

$$p'_j = \begin{cases} v_j & \text{if } r_j \le CR \\ p_j & \text{other case} \end{cases} \tag{18}$$

with $r_j \in [0,1]$ random and $CR \in [0,1]$ is the constant of crossing, a parameter that increases the diversity of individuals in the population.

3.20 Survival selection

For the selection of individuals that move to the next generation a simple selection one to one is used where the trial vector competes with the target vector. For the case of calibration in which minimization of J (p), the vector with the lowest value in the objective function goes to the next generation.

4. Case study and results

Table 1 shows a summary of the techniques used in each evolutionary algorithm. . To each algorithm, the parent selection, the crossover and the mutation were chosen based on the advantages that these have over the others (Michalewicz, 1996; Eiben and Smith, 2003; Coello, 2007).

	Genetic Algorithms	Evolutionary Strategies	Evolutionary Programming	Differential Evolution
Representation	Real-valued	Real-valued	Real-valued	Real-valued
Parent selection	Deterministic: by mean tournament	Uniform random	Deterministic (each parent creates one offspring)	Random
Recombination	2-point crossover	Discrete	None	Binomial
Mutation	Non uniform	Uncorrelated with n step sizes	Gaussian perturbation	Random
Survival Selection	Generational	Determination: (μ,λ)	Determinations: $(\mu+\mu)$	Elitist

Table 1. Comparative table of the main techniques of the evolutionary algorithms

4.1 Study cases

Some general statistics are used to analyze the obtained results, such as the correlation coefficient, r, which establishes the association degree between the measured and simulated variable. It is defined by means

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \tag{19}$$

Where x_i y y_i are measured data and estimated data, respectively, on the time i. making sure that $-1 \le r \le 1$.

The standard percentage error of the prediction (%SEP) which establishes the dispersion degree between the measured and simulated variable.

$$\%\ ESP = \frac{100}{\bar{x}}\sqrt{\frac{\sum_{i=1}^{n}\left(x_i - y_i\right)^2}{n}} \tag{20}$$

The efficiency coefficient (E) and the average relative variance (ARV). These estimators are used to determine the way in which the model can explain the total variance of the data (Ríos-Moreno et al., 2006).

$$E = \frac{S_{obs} - S}{S_{obs}} \tag{21}$$

and

$$\Delta(t,y) = y * \left(1 - r^{\left(1-\frac{t}{T}\right)^b}\right) \tag{22}$$

where

$$S_{obs} = \sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2 \tag{23}$$

and

$$S = \sum_{i=1}^{n} \left(x_i - y_i \right)^2 \tag{24}$$

For a perfect relation, r and E should be near 1 and the values of %SEP and ARV near 0.

4.2 White box model

The model proposed by Tap (2000) was considered, one of the most complete and available in the literature, which is associated a differential equation to each one of the state variables that are: air temperature (T_i), soil temperature (T_s) CO_2 concentration (C_i) and absolute humidity inside (V_i).

The equation of the greenhouse air temperature indicates that the variation of temperature inside the greenhouse is proportional to the heat exchange ventilation, heat input due to the heating system, to the exchange through the roof and walls, to heat exchange with the deep soil, to heat input by radiation, to the evaporative heat loss due to transpiration as well as the interchange due to condensation on the roof of the greenhouse. So, the rate of change of air temperature (T_i) with regards to time is described by:

$$C_g \frac{dT_i}{dt} = k_v \left(T_o - T_i \right) + H + k_r \left(T_o - T_i \right) + k_s \left(T_s - T_i \right) + \eta G - \lambda E + \frac{\lambda}{\varepsilon + 1} M_c \tag{25}$$

Where C_g is the greenhouse heat capacity expressed in J $^{o}C^{-1}$ m^{-2}, k_v is the coefficient of heat transfer vent W $^{o}C^{-1}$ m^{-2}, T_o is the outside temperature in ^{o}C, H is the heat input for heat, k_r is the coefficient of heat of ventilation from the cover in W $^{o}C^{-1}$ m^{-2}, k_s is the coefficient of heat transfer from the ground in W $^{o}C^{-1}$ m^{-2}, η is the solar efficiency factor, G is the radiation in W m^{-2}, E is the crop transpiration, λ is the energy of vaporization of water in J g^{-1}, ε heat resistance from the cover between inside and outside and M_c is the condensation of water on the cover.

Then a more detailed description of each one of the parameters that interfere in equation (25), which represents the rate of change of air temperature inside the greenhouse. So, it defines the energy of vaporization of water such as:

$$\lambda = l_1 + l_2 T_i \tag{26}$$

where l_1 and l_2 are the coefficients of energy of vaporization.

The heat transfer coefficient ventilation k_v is a function of ventilation rate Φ_v

$$k_v = M_{air} c_p \Phi_v \tag{27}$$

where M_{air} is the air density and c_p the specific heat of air.

The relationship for the natural ventilation flow through windows Φ_v, does not differ between the opening of barlovento windows (r_{ww}) and sotavento (r_{wl}), when in fact these

have a very different influence on ventilation (from Jong, 1990). Thus, the way to generalize this relationship is:

$$\Phi_v = \left(\frac{\sigma r_{wl}}{1 + \chi r_{wl}} + \zeta + \xi r_{ww} \right) w + \psi \tag{28}$$

Where σ, χ, ς, ξ and ψ are parameters that determine the ventilation rate and w represents wind speed. This relationship was developed by Van Henten (1994), and it is based on work by Jong (1990).

Another important factor that interferes with both the air temperature change and in the absolute humidity is the crop transpiration E, which is calculated by an adopted version of the Penman-Monteith transpiration. The Penman-Monteith model was chosen because it does not need the leaf temperature as input. In this sense, it is saved a state variable, while the model is still reasonably accurate (Jolliet & Bailey, 1992).

$$E = \frac{sn\eta G + \rho_a c_p D_g g_b}{\lambda \left(s + \gamma \left(1 + \frac{g_b}{g} \right) \right)} \tag{29}$$

The Penman-Monteith equation assumes that the leaf area index (LAI) is one that is the crop in the greenhouse is conceived as a big leaf. Basically transpiration is generated by a contribution of the net short-wave radiation absorbed nηG and a contribution due to the vapor pressure deficit D_g, γ is the apparent psychometric constant and g is the conductance of the leaf. In Eq. (30), the slope of the curve of saturated steam pressure (s) can be approximated by the polynomial

$$s = s_1 T_i^2 + s_2 T_i + s_3 \tag{30}$$

Where s_1, s_2 and s_3 are polynomial parameters, $\rho_a c_p$ represents the heat capacity of air volume, gb is the conductance of the leaf. All these quantities are assumed constant. The vapor pressure deficit of air D_g, is calculated as the difference between saturated vapor pressure p_g^* a T_i (Hanan, 1998)

$$p_g^* = a_1 e^{\frac{a_2 T_i}{a_3 + T_i}} \tag{31}$$

And air vapor pressure to the prevailing water content p_g

$$p_g = \Lambda \left(T_i + T_0 \right) V_i \tag{32}$$

Where Λ is the constant of pressure that can be obtained from the ideal gas laws, T_0 is used to convert T_i ^oC to Kelvin and V_i the concentration of water vapor in the air. So

$$D_g = p_g^* - p_g \tag{33}$$

The conductance of the leaf is related to short-wave radiation and the concentration of CO_2 through the regression equation

$$g = g_1\left(1 - g_2 e^{-g_3 G}\right)e^{-g_4 C_i} \tag{34}$$

Where from g_1 to g_4 are regression coefficients.
A detailed modelling of the transport of water vapor to the cover would require a description of the cover temperature, humidity inside the greenhouse, the possible re-evaporation of condensation, and the present quantity of condensation on the cover. To calculate the condensation are needed humidity inside the greenhouse (V_i) and the cover temperature (T_c). To keep the model as simple as possible, the temperature of the cover will be calculated based on the relationship obtained by Bakker (1995), which describes the temperature of the cover as an algebraic average between the outdoor temperature and indoor temperature

$$T_c = \frac{\varepsilon}{\varepsilon + 1} T_o + \frac{1}{\varepsilon} T_i \tag{35}$$

In this sense, the influence of the heat capacity of the cover and short-wave radiation from the sky are ignored. So, the approximation (3.8) is allowed, as Bakker (1995) showed that even instantly condensation can be incorrect, an average of condensation on the basis of a day is correct. Condensation on the greenhouse cover takes place when the roof temperature (T_c) is below the dew point (dew point) from the air of the greenhouse. Introducing the index of humidity as the mass of water vapor per humid air mass unit, condensation takes place when the percentage of humidity under pressure of saturated steam on the cover (W_c^*) is lower than the percentage of air humidity (W_g). The superscript * indicates that the amount considered is saturated vapor pressure. T_0 complete W_c^*, first the pressure of saturated steam at T_c is calculated according to Eq. (35) (replacing T_i by T_c). The vapor pressure in the air of the greenhouse (p_g) is calculated according to Eq. (32). When it is known the vapour pressure, humidity ratio W can be calculated as follows

$$W = \frac{\omega p}{p_{atm} - p} \tag{36}$$

Where ω is the parameter of the ratio of humidity and p_{atm} is the atmospheric air pressure. W_g or W_c^* can be calculated substituting p_g or p_c for p, respectively.
Used W_c^* and W_g, the rate of condensation (M_c) is calculated as follows

$$M_c = \begin{cases} m_1 |T_i - T_c|^{m_2} \left(W_g - W_c^*\right), & if \quad W_g > W_c^* \\ 0, & if \quad W_g \le W_c^* \end{cases} \tag{37}$$

Where $m_1 | T_i - T_c | m_2$ is the mass transfer coefficient, m_1 and m_2 are the parameters of mass transfer coefficients. In the process of condensation, water and energy are transported simultaneously. At the moment that water condenses, the energy of condensation is released into the surrounding environment. This is why condensation is part of the temperature equation (25) and the humidity equation (40).
In the case of the equation for soil temperature it is indicated that it is proportional to the heat exchange between the surface layer and the ambient temperature and the heat exchange between the surface layer and the deep soil. Thus, the rate of change of soil temperature (T_s) respect to time is:

$$C_s \frac{dT_s}{dt} = -k_s \left(T_s - T_i\right) + k_d \left(T_p - T_s\right) \tag{38}$$

Where C_s is the heat capacity of soil in J $\cdot^o$C^{-1} $\cdot$m^{-2}, k_d coefficient of heat transfer from layer to layer of soil in W $\cdot^o$C^{-1} $\cdot$m^{-2}, T_p is the temperature of the deep layer soil in °C.

For the CO_2, concentration, the equation establishes that the change in CO_2 concentration is proportional to CO_2 exchange with the outside, to the injection of CO_2 to increase of CO_2 by respiration as well as the reduction of CO_2 by photosynthesis. Thus, the rate of change of CO_2 concentration (C_i) with respect to time is described by:

$$\frac{V_g}{A_g}\frac{dC_i}{dt} = \Phi_v \left(C_o - C_i\right) + \varphi_{inj} + R - \mu P \tag{39}$$

V_g/A_g is the average height of the greenhouse in m, Φ_v is the ventilation flow in m $\cdot$s^{-1}, C_o is the external CO_2 concentration in g $\cdot$m^{-3}, φ_{inj} is the CO_2 injection flow g $\cdot$s^{-1} $\cdot$m^{-2}, R is the respiration of the crop in g $\cdot$s^{-1} $\cdot$m^{-2}, P is the photosynthesis of the crop in g $\cdot$s^{-1} $\cdot$m^{-2} and μ is the molar fraction of CO_2 and CH_2O.

For humidity, the equation indicates that the change of humidity inside the greenhouse is proportional to the increase of humidity by transpiration, to the exchange of humidity by respiration as well as the loss of humidity from condensation on the cover of the greenhouse. Thus, the rate of change of humidity in the greenhouse (V_i) respect to time is:

$$\frac{dV_i}{dt} = \frac{A_g}{V_g}\left(E - \Phi_v \left(V_i - V_o\right) - M_c\right) \tag{40}$$

Where V_o is the outdoor humidity, in kg $\cdot$m^{-3}.

As mentioned before, measurements were made from climate variables involved in the model that describes the greenhouse environment, among them it is found the outdoor relative humidity. However, the model has as one of its inputs to external absolute humidity, so it is estimated as absolute humidity (AH) from air temperature (T) and relative humidity (RH) by means of the following expression (Hanan, 1998):

$$HA = \frac{18e_1}{29\left(1 - e_1\right)} \tag{41}$$

where

$$e_1 = \frac{HR}{100}\frac{e_2 / 18}{1/29 + e_2/18} \tag{42}$$

$$e_2 = 3.77 \times 10^{-3} 2.965 \times 10^{-4} T + 5.2 \times 10^{-6} T^2 + 3.7 \times 10^{-7} T^3 \tag{43}$$

Moreover doing the estimation of the four state variables which has the mathematical model, it is determined from the internal relative humidity because it is most used the unit in practice, calculated by the following equation (Hanan, 1998):

$$HA = 100\frac{p_g}{p_g^*} \tag{44}$$

where p_g is the vapor pressure of air in the greenhouse and p_g^* is the saturated vapor pressure of air in the greenhouse.
In order to calibrate a mathematical model properly, it is important to carry out a sensitivity analysis (Van Henten, 2003), which evaluates the relative importance of input variables and model parameters on the evolution over time of the model's state variables (Saltelli et al., 2000). Then model calibration is performed. Finally, a validation process is completed, by comparing simulation results using parameter values obtained from calibration, with measurements which were not used during model calibration.
To achieve a good fit of a greenhouse climatic model, it is necessary to find suitable values for the parameters in the model. In order to find the parameters that most affect the state variables, it is necessary to do a sensitivity analysis of the model.
The time evolution of the sensitivity model (S(t)) of all states with respect to all parameters is defined as

$$S(t) = \frac{\partial x(t)}{\partial p} \tag{45}$$

where $x(t) \in R^n$ is the evolution time of the state vector, $p \in R^m$ is the parameter vector and t denotes time (Tap, 2000). Thus, S(t) is a $n \times m$ dimensional matrix, where every element represents the evolution time of the derivative of one state by one parameter. Eq. (46) is an approximation of the time evolution of

$$\frac{dS}{dt} \approx \frac{\partial f}{\partial x} S + \frac{\partial f}{\partial p}, \qquad S(t_0) = 0 \tag{46}$$

Where the right-hand side of the output equation is defined by f:

$$\frac{dx(t)}{dt} = f(x, u, p) \tag{47}$$

Where u is the input vector (controls) which does not depend on p. To be able to make a good comparison between the different sensitivities the relative sensitivity S_{rel} is calculated, as given by the following equation:

$$S_{rel} = S\frac{p}{y} = \frac{\partial y}{\partial p}\frac{p}{y} \tag{48}$$

A drawback is that this can yield numerical problems when x is very small or equal to zero. Parameter ranking is done on the basis of

$$I = \int_{t_0}^{t_f} |S_{rel}| \, dt \tag{49}$$

where t_0 is the initial time and t_f the final time.
The following step is the process of calibration of the model which consists of altering parameters to obtain a better fit between the simulated and measured data. An appropriate method to make the calibration is to use a non-linear multivariable optimization function (Tap, 2000), to minimize the sum of square errors (J):

$$J(p) = \sum_{h=1}^{L}\sum_{i=1}^{M}\sum_{j=1}^{N} w_h\left(\overline{y_h}(t_i, p) - y_{hj}(t_i)\right)^2; \qquad (50)$$

$$p^* = \arg\min J(p)$$

Where w_h is the relative weight of each output , $y_h(t_i, p)$ is the simulated output, y_h in time t_i, $y_{hj}(t_i)$ is the j-th repetition of the measurement y_h in time t_i, L is the number of outputs, M is the number of real measurements (time), N number of repetitions in each real measurement (time), p is the parameter set of calibration and p* are the parameters that reduce J(p) to a minimum. The weights w_h determine the relative importance of the different outputs in Eq. (50). These were calculated by normalization of the output vector to avoid problems with the units of the state variables.

4.3 Black box model

In Other control the air temperature inside of a closed environment it Can Be Modeled by auto regressive models with external inputs (ARX) and by auto regressive moving average models with external inputs (ARMAX) considering the inputs and outputs Are Measured by sensors. For a system with one input and one output (SISO) the model is Given by Eq. (1) (Ljung, 1999, Aguado & Martinez, 2003; Ljung, 2005)

$$\begin{aligned} &y(t) + a_1 y(t-1) + \ldots + a_{n_a} y(t-n_a) = \\ &\quad b_1 u(t-n_k) + \ldots + b_{n_b} u(t-n_k-n_b+1) + e(t) + c_1 e(t-1) + \ldots + c_{n_c} e(t-n_c) \end{aligned} \qquad (51)$$

Where y(t) is the output of the ARX and ARMAX models for t = 1, t-1,… t-na; u(t) is the input for t = t-nk, t-nk -1,…, t-nk-nb +1; na is the number of samples passed in the time of the output; nk is the delay time of the input u(t), e(t) is the white noise associated with the input and t is discrete time.

To evaluate the temperature inside of a closed environmental using ARX and ARMAX models, more input variables are required, so the models have multiple inputs and one output (MISO). The structures ARX and ARMAX for MISO systems are defined by Eq. (2) and (3), respectively

$$A(q)y(t) = B(q)u(t-n_k) + e(t) \qquad (52)$$

Where A(q), B(q) are matrices and C(q) is a vector, all defined by Eq. (53) –(54):

$$\text{A(q):}\ 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a} \qquad (53)$$

$$\text{B(q):}\ 1 + b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \qquad (54)$$

And the operator q^{-1} is the backward shift operator

$$q^{-1}u(t) = u(t-1) \qquad (55)$$

For a system in which the number of inputs is given by n_y and the number of outputs by nu, A(q) and B(q) are n_y by n_y and n_u by n_u matrices, respectively, whose elements are polynomials in the shift operator q-m (with m any natural number). The entries $a_{ij}(q)$ and $b_{ij}(q)$ of the matrices A(q) and B(q), respectively, can then be written as

$$a_{ij}(q) = \delta_{ij} + a_{1_{ij}} z^{-1} + \cdots + a_{n_{a_{ij}}} z^{-n_{a_{ij}}} \tag{56}$$

and

$$b_{ij}(q) = b_{1_{ij}} z^{-n_{k_{ij}}} + \cdots + b_{n_{b_{ij}}} z^{-n_{k_{ij}} - n_{b_{ij}} + 1} \tag{57}$$

Where δ_{ij} represents the Kronecker symbol.
From the above it is clear that the ARX structure for a given system can be defined by means of the number of poles n_a, the number of zeros n_{b-1} and the number of time delays n_k. The matrices A(q) and B(q) are determined by means of off-line parameter identification methods (Uchida-Frausto et al., 2003).
To achieve a good fit of a greenhouse climatic model, it is necessary to estimate the suitable values for the coefficients implicated in the autoregressive model. That is, for each structure of ARX and ARMAX models we need to obtain the coefficients $a_1,...,a_{n_a}$ and $b_1,...,b_{n_b}$ and the order of the model given by the values the parameters n_a, n_b and nk, based on the information provided by the inputs and outputs in order to get the best fit between the measured values and the estimated values by the model.
To determine the coefficients of ARX and ARMAX models that better fit the simulated to the measured data. A method to make it is to minimize the sum of square errors (J):

$$\begin{aligned} J(p) &= \sum_{i=1}^{N} \left(\bar{y}(t_i, p) - y(t_i) \right)^2 \\ p^* &= \arg\min J(p) \end{aligned} \tag{58}$$

Where $\bar{y}(t_i, p)$ is the simulated output, y in time t_i, $y(t_i)$ is the measurement y in time t_i, N is the number real measurement (time), p is the parameters set (coefficients of the model) and p* are the parameters that reduce J(p) to a minimum.
In the current work, the minimization of Eq. (58) is a non-linear multivariable optimization problem that can be solved by using evolutionary algorithms, such as: GAs and EP since they are global optimization methods. The structure of any EA is the same (Eiben & Smith, 2003), as is shown in Fig. 1. Differences among evolutionary techniques consist of the kind of selection, mutation and crossover operators applied to find the optimum value of the parameters for the optimization function. In this case, the kind of selection, the crossover and the mutation used for each EA is presented in Table 1.

4.4 Gray box model

In a gray box model it is made the integration of a white box model and a black box model. The workout of the gray box model is done with the indirect strategy used in research as Acuña et al. (1999), Dimitri (1992), Thomson and Kramer (1994) y Thornley & Johnson (2000). In figure 3 is presented the indirect training diagram.
The objective of a gray box model training consist of determining the ARX model parameters which approximate the existing and unknown relation between and a priori of some white box model parameters and the relevant resultant variables. Since the ARX

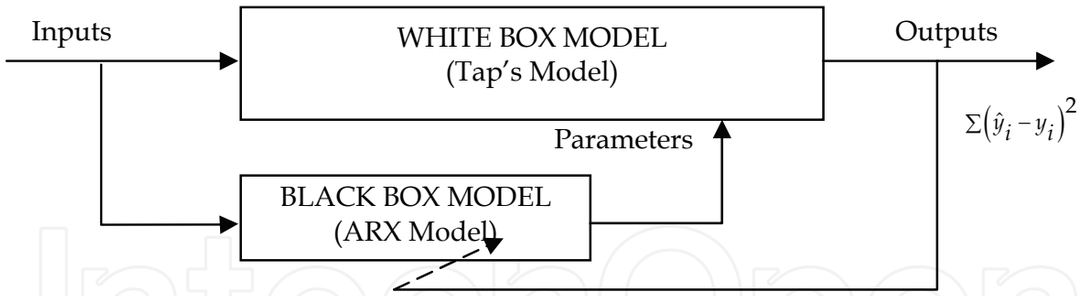


Fig. 3. Indirect training diagram in a gray box model.

model parameters are unknown, the output of the white box model (measurable outputs) is minimized. The objective function is:

$$J(p) = \frac{1}{2}\sum_{i=1}^{N}\left(y(t_i) - y_{cg}(t_i, p)\right)^2$$
$$p^* = \arg\min J(p) \tag{59}$$

Where $y(t_i, p)$ is the simulated output by the gray box model, $y(t_i)$ is the measurement y in time t_i, N is the number real measurement (time), p is the set of parameters (coefficients of the model) and p* are the parameters that reduce J(p) to a minimum.

The objective function gives a consistent real value in the sum of the square error between obtained values by the model and the expected outputs, with which the optimization method determines the ARX model parameters.

In this case, the White box model used is Tap's (2000). However, for the natural ventilation flow through windows Φ_v, Eq. (28), and the saturated steam p_g^*, Eq. (31), and Ti (Hanan, 1998) would be estimated by a black box model, this is, with an ARX model. Due to Eq. (28) and (31) depend on parameters σ, χ, ς, ξ, ψ, a_1, a_2 and a_3 that according to the sensitivity analysis these turn out to be the most sensitive.

5. Results

5.1 White box model

Initial values of model parameters were taken from Tap (2000). Table 2 shows the statistics corresponding to the results of the simulation before model calibration. Figure 4 shows the results obtained before the calibration for the simulation of a week's data.

A local sensitivity analysis was carried out using measured climate data in order to select the most sensitive model parameters to be estimated during the calibration process. As result of this analysis, the parameters related to the opening of the windows and to the evapo-transpiration of the plant were very sensitive. That is, the inside air temperature and the absolute humidity were determined to be most affected by the saturation vapour pressure parameters (a1, a2 and a3) and the ventilation rate parameters (σ, χ, ς, ξ and ψ). The number of parameters could be reduced by considering a smaller sensitivity index. However, in this case, 8 parameters were considered because there was a considerable difference between the effect of these 8 and that of the rest.

Calibration of parameters was made by means of the four evolutionary techniques (global search method): GAs, ES, EP and ED.

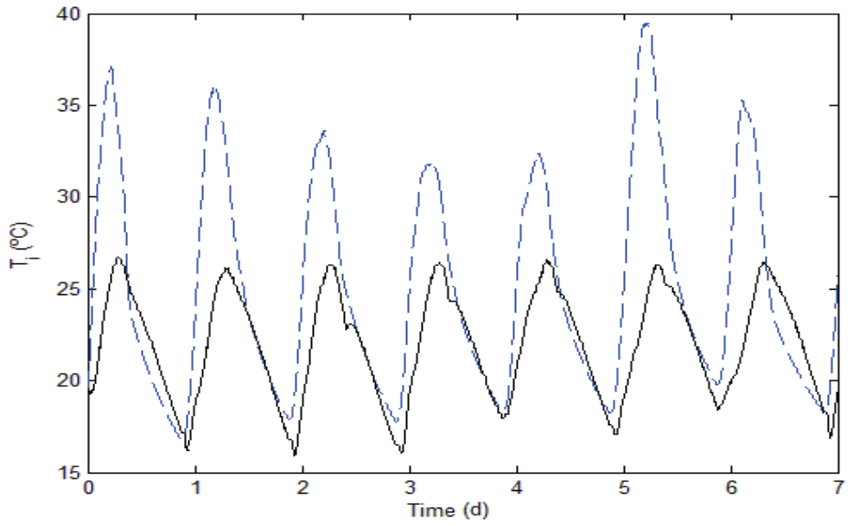


Fig. 4. Indirect Measured (solid line) and simulated (dotted line) air temperature before calibration.

	Method	r	E	%SEP	AVR
Before calibration	Original values	0.6643	-2.3902	24.5262	3.3902
After calibration	GAs	0.6422	-0.1936	14.5527	1.1936
	EP	0.6785	-0.5266	16.4581	1.5266
	ES	0.6398	-1.345	20.3982	2.345
	DE	0.6874	-0.2708	15.0159	1.2708

Table 3. Statistical results of greenhouse air temperature before calibration and after calibration of a white box model using evolutionary algorithms.

The statistics of the results of the simulation using the parameters given by GAs, ES, EP and DE are shown in Table 3. The obtained results of the simulations for a week with the GAs,

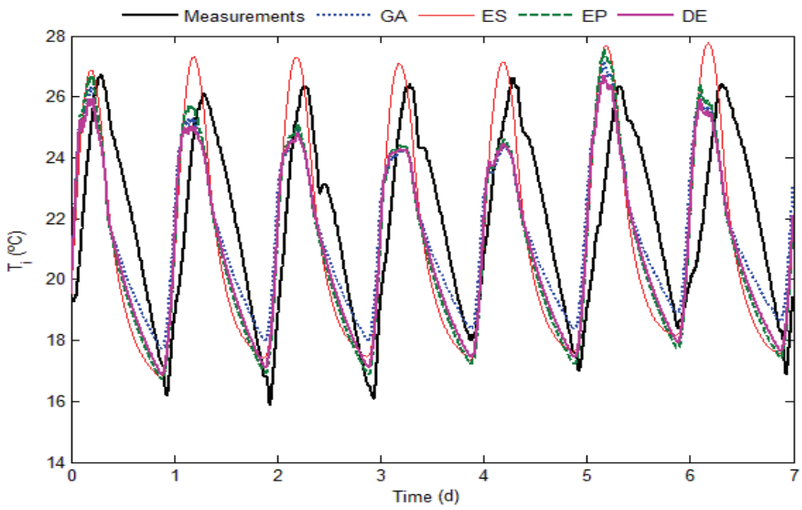


Fig. 5. Measured (solid line) and simulated air temperature after calibration with GAs, ES, EP and DE.

ES, EP and DE parameters are shown in the Figure 5. Results for Evolutionary Algorithms are obtained from 10 runs and the selection was made considering the parameters that minimized the error better between measured and estimated data.
Results show a minor error in the model's predictions for the air temperature within the greenhouse after calibration. Although there is not an increase in the correlation (r), for T_i, for each calibration method results between results obtained before and after calibration with GA and EP, unlike the other two methods, Table 3. Furthermore, it can be observed that the efficiency coefficient (E) is negative in all cases, suggesting that average use values of observed data is better than the estimates obtained. However, percentage standard error of the prediction (% SEP) changed from 24 to 14 (Table 3) for the temperature T_i , when GA are used. Finally, average relative variance (ARV) decreased from 3.39 to 1.19 (Table 3) for the temperature T_i when GA is used. In general terms there is an improvement when GA is used.

5.2 Black box model

An ARX model with structure na = 1 (one output variable), for nb=2 (four input variables) and for nk = 1 was evaluated.

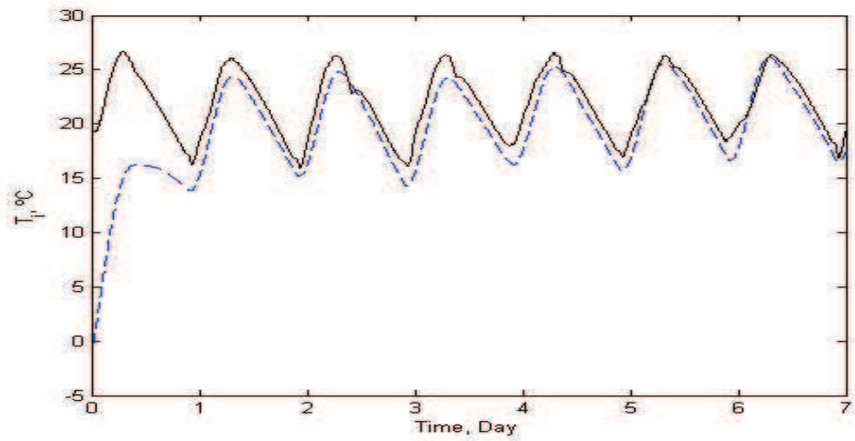


Fig. 6. Measured (solid line) and simulated (dotted line) air temperature by means of an ARX model

Table 4 shows the statistical results obtained when 25% of the data is evaluated to estimate the ARX model. Fig. 6 shows the behavior of the ARX model with structure na=1, nb=[2 2 2 2]. Looking at Table 4 it can be observed that a better fit occurs when the identification parameter is performed with GAs. Correlation r takes the value 0.86 and the efficiency (E) obtains value of 0.0.44, percentage standard error of the prediction (%SEP) is 9.94 and the average relative variance (ARV) is 0.55.

Method	r	E	%SEP	AVR
ARX	0.6620	-0.8402	18.0704	1.8402
GAs	0.8691	0.4429	9.9427	0.5571
EP	-0.0058	-0.5386	16.5233	1.5383
ES	-0.0029	-1.3235	20.3051	2.3235
DE	0.7735	-0.2726	11.3607	0.7274

Table 4. Statistical results of greenhouse air temperature before calibration and after calibration of a black box model using evolutionary algorithms.

According to the results, when the parameter identification of an ARMAX structure is performed by means of GAs there is a better fit of the simulated data to the measured data when t 25% of the data are used to estimate and 75% of the data is used to validate the model.

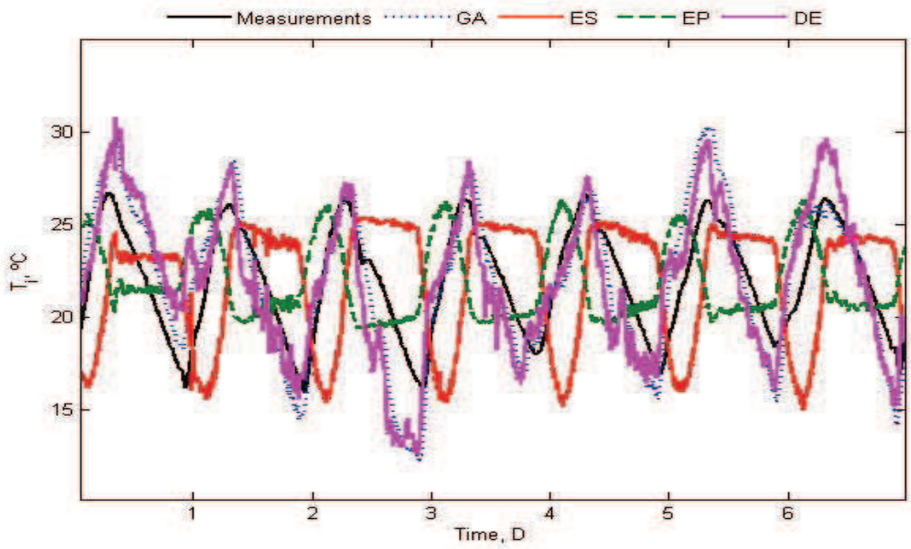


Fig. 7. Measured (solid line) and simulated air temperature using GAs, ES, EP and DE

5.2 Gray box model

In this case the white box model applied was that of Tap (2000), where natural ventilation flow through windows Φ_v and the saturated vapour p_g^* are estimates through an ARX structure. The results of this simulation are shown in Fig. 8 where 50 % of the data was used for estimating the parameters.

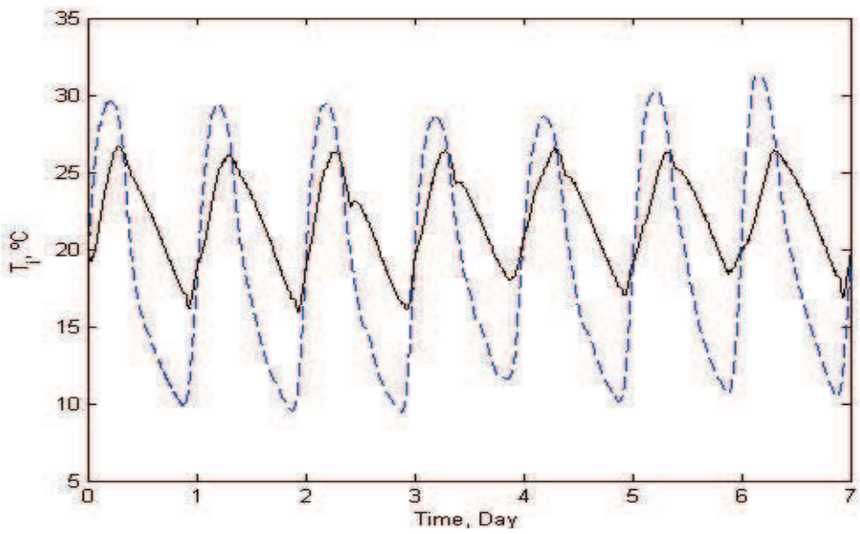


Fig. 8. Measured (solid line) and simulated (dotted line) air temperature by means of a gray box model

The statistical results of the parameter identification through GA, ES, EP, and DE are shown in Table 5. According to these results there is a better fit of the estimated data through the gray box model when parameter identification is done using DE. Fig. 9 shows the results of each of the simulations

Method	r	E	%SEP	AVR
Gray box model	0.7196	-3.1438	27.0180	4.1438
GAs	0.7186	-32.1576	76.4268	33.1576
EP	0.6787	0.3450	10.7419	0.6550
ES	0.671	0.2778	11.3200	0.7222
DE	0.9015	0.7064	7.2173	0.2936

Table 5. Statistical results of greenhouse air temperature before calibration and after calibration of a gray box model using evolutionary algorithms

The obtained results of the simulations for a week with the GAs, ES, EP and DE parameters are shown in the Fig. 5.

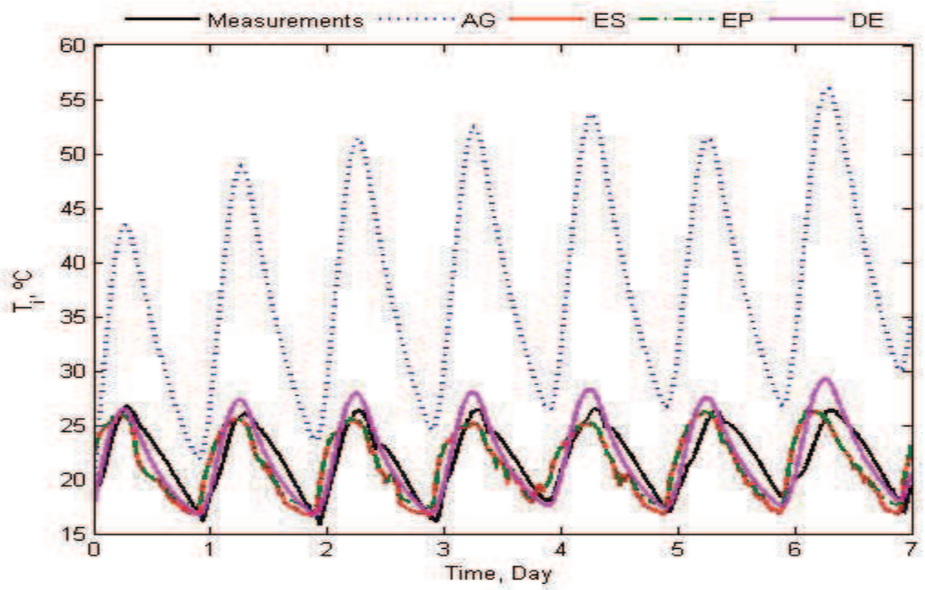


Fig. 9. Measured (solid line) and simulated air temperature using GAs, ES, EP and DE

6. Conclusions

Climate conditions were measured in a greenhouse located in the central region of Mexico. To perform the calibration, 4 global evolutionary algorithms (GAs, ES, EP and DE) were applied and the estimations obtained by the model using the parameter values given with the different methods were compared in order to ascertain which method was more effective.

The climate model gave best predictions for the air temperature within the greenhouse when using the parameter values obtained by means of the GA when a white box model was used.

In the same way, results obtained with GAs show that this method is more effective than the others methods to find parameters for auto regressive models to predict air temperature inside of a greenhouse with structure na = 1, nb = [2 2 2 2], and nk = [1 1 1 1] and the data group 25%:75%. The advantage in this case consists in the use of small sample data (25%) can give a better estimation that the traditional method (square least) used to the parameter identification of an ARX model.

For gray box model, the prediction of air temperature inside of a greenhouse is better when the parameter identification is done by means of the DE method.

To this point, the model can be used to design and development of algorithms of control. Likewise, the model can be integrated with a physiological model to get a production process model of a greenhouse.

7. Acknowledgment

This research was partially supported by CONACyT. Many thanks to Cindy Esparza of the Translation Edition Office of UPSZ for the language revision.

8. References

Acuña, G.; Cubillos, F.; Thibault, J. & Latrille, E. (1999). *Comparison of methods for training grey-box neuronal networks models*. Computers and Chemicals Eng., Supplement, 561-564

Aguado, B. A. & Martínez, I. M. (2003). *Automática y Robótica: Identificación y Control Adaptativo*, Prentice Hall Ed.

Bakker, J. C. (1991). *Analysis of humidity effects on growth and production of glasshouse fruit vegetables*, PhD thesis, Wageningen Agricultural University, The Netherlands

Bot, G. P. A. (1983). *Greenhouse climate: form physical processes to a dynamic model*, PhD thesis, Wageningen Agricultural University, The Netherlands

Bouman, B. A. M.; Van Keulen, H.; Van Laar, H. H. & Rabbinge, R. (1996). *The 'School of De Wit' crop growth simulation models: A pedigree and historical overview*, Agric. Sys., Vol. 52, Nos.2/3, pp. 171-198

CIDEIBER. (October 2005). México, actividades del sector primario, sector agrícola vegetal, Available from Internet: http:// www.cideiber.com/infopaises/Mexico/ Mexico-04-01.html

Coello, C. A. (2007). *Introducción a la computación evolutiva*, Notas de curso, Departamento de computación, CINVESTAV-IPN, México

De Jong, T. (1990). *Natural ventilation of large multi-span greenhouses*, PhD thesis, Wageningen Agricultural University, The Netherlands

Dimitri, C. (1992). *A Hybrid Neural Network-First Principles Approach to Process Modeling*, Universidad de Pennsylvania, AIChE Journal, Vol. 38

Eiben, A. E. & Smith, J. E. (2003). *Introduction to Evolutionary Computing*, Springer, Berlin,

France, J. & Thornley, J. (2006). *Mathematical Models in Agriculture*, CABI

Hanan, J. (1998). GREENHOUSES: Advanced Technology for Protected Horticulture, volumen 1, CRC, first edition

Hernández-Hernández, F. (2009). Simulación del crecimiento y desarrollo del chile pimiento (Capsicum Annuum L.) para mejorar su producción en condiciones controladas, Tesis Maestria, Universidad Autónoma de Querétaro

INU. (2006). *El agua, una responsabilidad compartida*. 2º Informe de las Naciones Unidas sobre el desarrollo de los recursos hídricos en el mundo

Jolliet, O. & Bailey, B. J. (1992). The effect of climate on tomato transpiration in greenhouses: measurements and models comparison, Agricultural and Forest Meteorology, 58:43–62

Linker, R.; Seginer, I. & Buwalda, F. (2004). *Description and calibration of a dynamic model for lettuce grown in a nitrate-limitingenvironment*, Mathematical and Computer Modelling, 40(9–10), 1009–1024

Ljung, L. (1999). System Identification, Theory for the user, Prentice Hall Ed

Ljung, L. (2005). System Identification Toolbox for Use with MATLAB, The Matworks Inc

López-Cruz, I.; Ramirez-Arias, A. & Rojano-Aguilar, A. (2004). Análisis de sensibilidad de un modelo dinámico de crecimiento para lechugas (Lactuca sativa) cultivadas en invernadero, Agrociencia 38 (6): 613-624

López, I. L.; Rojano, A.; Ramírez, A. & Bonilla, M. (2006). *Modelos matemáticos para el crecimiento y desarrollo de cultivos: concepto y metodología*, Notas, Universidad Autónoma de Chapingo

Michalewicz, Z. (1992). *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, second edition

Michalewicz, Z. (1994). *Evolutionary computation techniques for nonlinear programming problems*, International Transactions in Operational Research, 1(2), 223–240

Michalewicz, Z. (1996). Genetic Algorithms + Data Structures = Evolution Programs, Springer, USA

Moreno Quintana, V. & Acuña Leiva, G. (2005). Estimacion de parámetros utilizando modelos de caja gris con redes neuronales de base radial, 13° encuentro chileno de computación

Price, K. V. & Storn, R. M. (2005). Differential Evolution: A practical approach to global optimization, Springer, Germany

Ríos-Moreno, G. J.; Trejo-Perea, M.; Castañeda-Miranda, R.; Hernández-Guzmán, V. M. & Herrera-Ruiz, G. (2006). *Modelling temperature in intelligent buildings by means of autoregressive models*, Automation in Construction, 16(5), 713–722

Saltelli, A.; Chang, K. & Scott, M. (2000). *Sensitivity Analysis*, J. Wiley & Sons, Chichester, England, 475 p

Seginer, I. & Ioslovich, I. (1998). A single state variable model to mimic tomgro for control applications, Acta Horticulturae 456: 93-100

Sjöber, J.; Zhang, Q.; Ljung, L.; Benveniste, A.; Deylon, B.; Glorennec, P.Y.; Hjalmarsson, H. & Juditsky, A. (1995). *Nonlinear black-box modeling in system Identification: a unified overwied*, Automatica, 31:1691-1724

Tap, F. (2000). *Economics-based optimal control of greenhouse tomato crop production*, Phd Thesis, Wageningen Agricultural University, Wageningen The Netherlands, 127 P

Thibault, I.; Acuña, G.; Perez, R.; Jorquera, H.; Molin, E. & Agosin, E. (1999). *A hybrid representation approach for modeling complex dynamic bioprocesses*, Bioprocesses Engineering 22

Thomson, M. & Kramer, M. (1994). Modeling Chemical Processes Using Prior Knowledge and Neural Networks, MIT, Process System Eng., Vol. 40

Thornley, J. H. M. & Johnson, I. R. (2000). *Plant and crop modelling, a mathematical aapproach to plant and crop physiology*, The Blackburn Press, New Jersey, USA., 669 P

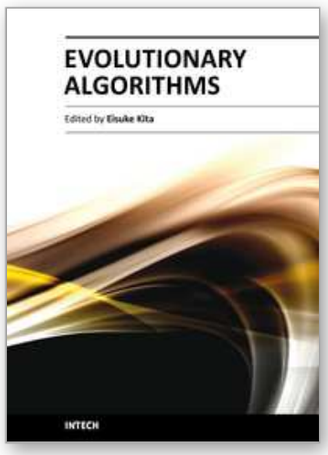
Uchida Frausto, H.; Pieters, J. G. & Deltour, J. M. (2003). *Modelling greenhouse temperature by means of auto regressive models*, Biosystems Engineering, 84, 147-157

Udink-Ten-Cate, A. J. (1983). *Modeling and (adaptive) control of greenhouse climates*, PhD thesis, Wageningen Agricultural University, The Netherlands

Van-Henten, E. J. (1994). *Greenhouse climate management: an optimal control approach*, PhD Thesis, Wageningen Agricultural University, The Netherlands

Van Henten, E. J. & Van Straten. (1994). *Sensitivity analysis of a dynamic growth model of lettuce*, J. Agric. Engng. Res., 59, 19-31

Van-Henten, E. J. (2003). Sensitivity Analysis of an Optimal Control Problem in Greenhouse Climate Management, Biosystems Engineering, 85:335-364



Evolutionary Algorithms
Edited by Prof. Eisuke Kita

ISBN 978-953-307-171-8
Hard cover, 584 pages
Publisher InTech
Published online 26, April, 2011
Published in print edition April, 2011

Evolutionary algorithms are successively applied to wide optimization problems in the engineering, marketing, operations research, and social science, such as include scheduling, genetics, material selection, structural design and so on. Apart from mathematical optimization problems, evolutionary algorithms have also been used as an experimental framework within biological evolution and natural selection in the field of artificial life.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Rosario Guzman-Cruz, Rodrigo Castañeda-Miranda, Juan Garcĩa- Escalante, Luis Solis-Sanchez, Daniel Alaniz-Lumbreras, Joshua Mendoza-Jasso, Alfredo Lara-Herrera, Gerardo Ornelas-Vargas, Efrén Gonzalez-Ramirez and Ricardo Montoya-Zamora (2011). Evolutionary Algorithms in Modelling of Biosystems, Evolutionary Algorithms, Prof. Eisuke Kita (Ed.), ISBN: 978-953-307-171-8, InTech, Available from: http://www.intechopen.com/books/evolutionary-algorithms/evolutionary-algorithms-in-modelling-of-biosystems

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen