

Document downloaded from:

<http://hdl.handle.net/10251/125201>

This paper must be cited as:

Herrero Durá, JM.; Blasco, X.; Martínez Iranzo, MA.; Ramos Fernández, C.; Sanchís Saez, J. (2007). Non-linear Robust Identification of a Greenhouse Model using Multi-objective Evolutionary Algorithms. *Biosystems Engineering*. 98(3):335-346.
<https://doi.org/10.1016/j.biosystemseng.2007.06.004>



The final publication is available at

<http://doi.org/10.1016/j.biosystemseng.2007.06.004>

Copyright Elsevier

Additional Information

Non-linear Robust Identification of a Greenhouse Model using Multiobjective Evolutionary Algorithms

J.M. Herrero; X. Blasco; M. Martínez; C. Ramos; J. Sanchis

*Department of Systems Engineering and Control, Polytechnic University of
Valencia, Camino de vera s/n, 46022 Valencia, Spain;
e-mail of corresponding author: juaherdu@isa.upv.es*

Abstract

This paper presents the non-linear modelling, based on first principle equations, for a climatic (temperature & humidity) model of a greenhouse where roses are to be grown using hydroponic methods, and the fitting of its parameters (15 in all) based on real data collected for the summer period. To do so, a procedure for estimating a set of non-linear models Θ_P (Pareto optimal) when several optimisation criteria are considered simultaneously within a multiobjective optimisation context is presented. A new multiobjective evolutionary algorithm, ϵ^λ -MOGA, has been designed to converge towards $\hat{\Theta}_P^*$, which is a reduced but well distributed representation of Θ_P since good convergence and distribution of the Pareto front $\mathbf{J}(\Theta_P)$ is achieved by the algorithm. *A posteriori*, $\hat{\Theta}_P^*$ can be used as the basis to choose an optimal model that offers a good relationship among the different optimality criteria that have been established, as has been shown by the results obtained in the identification and validation of the greenhouse model presented in this paper.

1 Introduction

Greenhouses were originally controlled without using mathematical models, which obviously led to anything but optimal performance. In recent years, the use of modelling and identification techniques has made it possible to begin working on the development and application of systems with more sophisticated control strategies (Young *et al.*, 1993; Nielsen & Madsen, 1996). The problems involved in controlling greenhouses are strongly dependent on the geographical area; solutions that are valid in some regions must be adapted or changed in order to fit others. More particularly, in Mediterranean countries the high levels of radiation and the high temperatures and humidity during the summer period are factors that differentiate these regions from others in Northern Europe. Until now, many of the controllers designed for greenhouses have been associated to a single control variable, *i.e.* temperature, and this has given rise to monovariate controllers. Under the abovementioned summertime conditions in the Mediterranean regions, this control is altogether insufficient and must be complemented with humidity control (Baille *et al.*, 1994), which then creates a need for multivariate controllers.

The multivariate process thus defined is also of a non-linear nature and it is influenced by biological processes that make it very complicated to develop a suitable mathematical model which would permit gaining access to very significant information, not only as far as control of the greenhouse is concerned but also regarding the design of the greenhouse itself. Several alternatives arise when faced with a problem of this magnitude. One of them is to deal with the process as if it were a black box without using information *a priori* and fit a neuronal network-type model (Seginer *et al.*, 1994; Linker *et al.*, 1998) or

a fuzzy set (Ehrlich *et al.*, 1996). One important drawback of using this type of techniques is the lack of a physical relation among the model parameters and the fundamental magnitudes of the crop, which turns them into barely comprehensible models. Another alternative consists in modelling the well-known physical and physiological phenomena that occur within a greenhouse by formulating first principle equations based on mass and energy balances (Boulard & Baille, 1993; Boulard *et al.*, 1996). In this case the model parameters do have a physical meaning, but problems arise when attempts are made to adjust these parameters.

Thus, on the one hand, obtaining reliable models implies having access to equations based on first principles that are sufficiently representative of the processes that take place inside greenhouses. On the other hand, it also means that a technique to adjust the parameters is needed in order to achieve a maximum reduction of the identification error (IE), that is to say, the discrepancies between the real outputs from the processes that occur inside the greenhouse and those that would be obtained from simulations with the proposed models.

The IE can be produced by systematic (bias) and/or random (variance) errors. The former are produced mainly by the non-modelled dynamics, that is to say, they are related to the structure of the selected model, while the latter are produced by the presence of measurement noise in the data from the experiment.

One function that is commonly used to minimise the IE is the 2-norm, which facilitates optimisation when the models are linear because the parameters are obtained directly and no numerical optimisation methods are required (L. Pronzalo, 1997).

The main advantage of using the 2-norm as a function can be lost when the model is not linear, since there is no longer any guarantee that the optimal parameters will be obtained explicitly. Furthermore, parameters estimation can turn out to be inadequate when the data include outliers, since the estimations could be heavily distorted by the effect of such atypical values (Zhang, 1996). Another equally important problem stems from the fact that the quadratic functions distort the influence of the errors because errors below unity are underestimated, whereas those above unity are overestimated.

This and other reasons related to the desired performance have led to the appearance of other optimality criteria in the literature (Zhang, 1996), for example, the minimisation of the 1-norm function, the ∞ -norm, Fair, Turkey, the median, and so forth. In any case, either because the model is not linear or because the type of function used is not 2-norm, the traditional optimisers may be unsuitable. For instance the SQP (Sequential Quadratic Programming) which is a variant of the Gauss-Newton optimisation method. This would be due to the fact that the function to be optimised is perhaps not convex and/or it may have local minima.

Furthermore, it is possible to consider the possibility of establishing several optimality criteria at the same time, which would be able to increase the quality of the estimated model. Thus, processes identification is stated as a multiobjective optimisation problem (MOP).

The criteria to be optimised will usually be in conflict, as there is no single solution that optimises all of them. Hence, there will be a set of solutions (in principle all equally valid) known as a **Pareto optimal set**. An intermediate solution will have to be selected *a posteriori* from this set so that it offers a

good compromising relation among the previously specified criteria. Obtaining the Pareto optimal set avoids having to repeat the optimisation process if there are any changes made to the preferences regarding the optimisation criteria, since only the solution selection would have to be repeated.

One very interesting alternative in resolving MOPs is based on the use of Evolutionary Algorithms (EAs), which allow several elements of the Pareto optimal set to be generated at the same time, in parallel and in a single run. This is made possible thanks to the populational nature of EAs. A number of authors have developed different operators or strategies for converting the original EAs into MOEAs that converge towards the Pareto optimal set and are diverse enough to be able to characterise it. The good results obtained with MOEAs together with their capacity to handle a wide variety of problems with different degrees of complexity explain why they are being used increasingly more frequently; indeed they are currently one of the branches where most progress is being made within the EAs field (Fonseca, 1995; Zitzler, 1999; Coello *et al.*, 2002; Alander, 2002; Coello *et al.*, 2005).

The rest of the paper is organised as follows: in section two we present the proposed multiobjective identification procedure is presented, as well as an alternative method for selecting an optimal compromise model based on the Pareto optimal solutions, the ϵ^2 MOGA evolutionary algorithm developed for solving multiobjective optimisations and the non-linear climatic model of the greenhouse in state space obtained from the first principles. In section three the experimental details and materials are presented, while in section four results of parameters identification of climate model by means of the multiobjective robust identification approach presented in this study is shown. The paper finishes, in section five, with a discussion of the most important conclusions

and contributions of the work.

2 Theoretical considerations

2.1 Multiobjective robust identification approach

The technique is based on the acceptance of an initial model structure that is obtained from the *a priori* knowledge we have about it. Some or all of its parameters are unknown at this point, since the objective is to determine them. It is common practice to represent the (linear or non-linear) process model by means of a series of first-order differential equations that can be obtained from physical principles.

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), \theta) \quad (1)$$

$$\hat{\mathbf{y}}(t, \theta) = g(\mathbf{x}(t), \mathbf{u}(t), \theta) \quad (2)$$

where $f(\cdot), g(\cdot)$ are the non-linear functions of the model; $\theta \in D \subset R^L$ is the column vector of unknown model parameters; $\mathbf{x}(t) \in R^n$ is the column vector of model states; $\mathbf{u}(t) \in R^m$ is the column vector of model inputs; and $\hat{\mathbf{y}}(t, \theta) \in R^l$ is the column vector of model outputs.

The aim is for the model behaviour to match that of the real process as well as possible.

The process behaviour can be characterised by means of experiments, while the model one can be obtained from simulating it by applying the same input

signals as those used in the experiment on the process.

Typically this objective is achieved by minimising a function where, throughout the experiment, differences between the process outputs and the model are penalised (see *Fig. 1*).

[Fig. 1 about here.]

Thus, the identification error matrix $\mathbf{E}(\theta)$ with dimensions $(l \times N)$ is defined as

$$\mathbf{E}(\theta) = \mathbf{Y} - \hat{\mathbf{Y}}(\theta) \quad (3)$$

where $\mathbf{Y} = [\mathbf{y}(t_1) \dots \mathbf{y}(t_N)]$ are the process output measurements ($\mathbf{y}(t) \in \mathcal{R}^l$ is the column vector of the process outputs) when the inputs $\mathbf{U} = [\mathbf{u}(t_1) \dots \mathbf{u}(t_N)]$ are applied to the process. ; and $\hat{\mathbf{Y}}(\theta) = [\hat{\mathbf{y}}(t_1, \theta) \dots \hat{\mathbf{y}}(t_N, \theta)]$ are the simulated outputs (the outputs of the model are calculated by integrating Eqn (1)) when the same inputs \mathbf{U} are applied to the model. $t_i, i \in [1 \dots N]$ are the different points in time in which the process outputs are sampled and those of the model are calculated. N is the number of samples of each output and input from the experiment. It is assumed that the interval between samples is constant $t_i = i \cdot T_s$, T_s being the sampling period or interval between samples.

The identification error vector $\mathbf{e}_j(\theta)$ for output $j \in [1 \dots l]$ is defined as row j of the identification error matrix $\mathbf{E}(\theta)$ and the identification error $e_j(t_i, \theta)$ for output $j \in [1 \dots l]$ and sample $i \in [1 \dots N]$ are defined as element ji of the identification error matrix $\mathbf{E}(\theta)$.

Therefore, the optimality criteria $J_i(\theta)$ will be established as identification error functions for a particular output j

$$J_i(\theta) = f(\mathbf{e}_j(\theta)), \quad i \in A := [1, 2, \dots, s], \quad (4)$$

When several identification error functions are considered simultaneously, model identification can be seen as a multiobjective optimisation problem

$$\min_{\theta \in D} \mathbf{J}(\theta), \mathbf{J}(\theta) = \{J_1(\theta), J_2(\theta), \dots, J_s(\theta)\} \quad (5)$$

Consequently, there is no single optimal model and a Pareto optimal set Θ_P (solutions where none dominate the others) must be found (see *Fig. 2*). Pareto dominance is defined as follows.

[Fig. 2 about here.]

A model θ^1 , with function value $\mathbf{J}(\theta^1)$ dominates another model θ^2 with function value $\mathbf{J}(\theta^2)$, denoted by $\mathbf{J}(\theta^1) \prec \mathbf{J}(\theta^2)$, if and only if

$$\forall i \in A, J_i(\theta^1) \leq J_i(\theta^2) \wedge \exists k \in A : J_k(\theta^1) < J_k(\theta^2).$$

Therefore the Pareto optimal set Θ_P , is given by

$$\Theta_P = \{\theta \in D \mid \neg \exists \tilde{\theta} \in D : \mathbf{J}(\tilde{\theta}) \prec \mathbf{J}(\theta)\}. \quad (6)$$

It is not easy to find a mathematical expression for the line or hypersurface formed by the Pareto front; in fact in most real cases it is impossible.

Since determining $\mathbf{J}(\Theta_P)$ is computationally unworkable because, in most cases, it is a set of infinite points, it is sometimes enough to determine a set of points $\hat{\Theta}_P^* \subset \Theta_P$, so that $\mathbf{J}(\hat{\Theta}_P^*)$ characterises $\mathbf{J}(\Theta_P)$ in a suitable manner (see *Fig. 2*).

Although Θ_P is a unique set, $\hat{\Theta}_P^*$ will not be because for that to be the case

we should have that $\hat{\Theta}_P^* = \Theta_P$.

The $\hat{\Theta}_P$ algorithm will be used to characterise ϵ -MOGA and the solution thus obtained will be a finite set of optimal models spread along the Pareto front, $\hat{\Theta}_P^*$. ϵ -MOGA is characterised, among other things, by its capacity to capture the ends of the Pareto front, and thus $\hat{\Theta}_P^*$ will contain the optimal models $\hat{\theta}^{J_i}$ of each J_i considered on an individual basis.

$$\hat{\theta}^{J_i} = \arg \min_{\theta \in D} J_i. \quad (7)$$

From the information generated by these optimal models it is possible to determine the ideal point

$$\mathbf{J}^{ideal} = \{J_1(\hat{\theta}^{J_1}), \dots, J_s(\hat{\theta}^{J_s})\} \quad (8)$$

and from there an optimal compromise model $\hat{\theta}^{ideal}$ can be chosen from the J_i functions that have been proposed, such as minimising a norm:

$$\hat{\theta}^{ideal} = \arg \min_{\theta \in \hat{\Theta}_P^*} \|\mathbf{J}(\theta) - \mathbf{J}^{ideal}\| \quad (9)$$

The optimal model $\hat{\theta}^{ideal}$ will depend on $\hat{\Theta}_P^*$ and this set, in turn, will depend on the optimality criteria that are chosen. An optimality criterion will be considered to be robust, when faced with the existence of incorrect measurements due to the sensors failings, if it filters out inappropriate information and generates a satisfactory model (L. Pronzalo, 1997). An example would be the estimator that minimises the median of the identification error module.

$$\hat{\theta}_{med} = \arg \min_{\theta \in D} \text{med}_i |e(t_i, \theta)|. \quad (10)$$

This estimator implicitly eliminates 50 % of data, which makes it very robust to incorrect measurements but perhaps a little too conservative. Its main drawback is that it could reject data that cover an important part of the response, with the subsequent loss of identifiability. It is for this reason that another, less conservative, possibility is to minimise the third quartile

$$\hat{\theta}_{Q3} = \arg \min_{\theta \in D} \text{Q3}_i |e(t_i, \theta)| \quad (11)$$

which would mean that only 25 % of data is removed. An extreme case would be that of the estimator that minimised the ∞ -norm where no data would be removed; this is why this estimator is not considered to be robust to wrong measures.

Both the median function and the one using the third quartile make optimisation difficult because it is not possible to find a direct expression of them. They are therefore going to need powerful optimisers.

2.2 ϵ -MOGA algorithm

The ϵ -MOGA variable (ϵ -MOGA) is an elitist multiobjective evolutionary algorithm based on the concept of ϵ -dominance (Laumanns *et al.*, 2002). ϵ -MOGA makes it possible to obtain an ϵ -Pareto set, $\hat{\Theta}_P^*$, that converges towards the Pareto optimal set Θ_P in a distributed manner and utilises limited memory resources. It also adjusts the limits of the Pareto front dynamically and

prevents the solutions belonging to the ends of the front from being lost.

For this reason, the objective function space is split up into a fixed number of boxes. For each dimension, n_box_i cells ϵ_i width,

$$\epsilon_i = (J_i^{max} - J_i^{min})/n_box_i.$$

This grid preserves the diversity of $\mathbf{J}(\hat{\Theta}_P^*)$ since one box can be occupied by only one solution. This fact prevents the algorithm from converging towards just one point or area inside the function space (*Fig. 3*).

[Fig. 3 about here.]

The concept of ϵ -dominance is defined as follows. For a model θ , $box_i(\theta)$ is defined by

$$box_i(\theta) = \left[\frac{J_i(\theta) - J_i^{min}}{J_i^{max} - J_i^{min}} \cdot n_box_i \right] \forall i \in [1 \dots s] \quad (12)$$

Let $\mathbf{box}(\theta) = \{box_1(\theta), \dots, box_s(\theta)\}$. A model θ^1 with function value $\mathbf{J}(\theta^1)$ ϵ -dominates the model θ^2 with function value $\mathbf{J}(\theta^2)$, denoted by $\mathbf{J}(\theta^1) \prec_\epsilon \mathbf{J}(\theta^2)$, if and only if

$$\mathbf{box}(\theta^1) \prec \mathbf{box}(\theta^2) \vee (\mathbf{box}(\theta^1) = \mathbf{box}(\theta^2) \wedge \mathbf{J}(\theta^1) \prec \mathbf{J}(\theta^2)) .$$

Hence, a set $\hat{\Theta}_P^*$ is ϵ -Pareto if and only if $\forall \theta^1, \theta^2 \in \hat{\Theta}_P^*, \theta^1 \neq \theta^2$

$$\hat{\Theta}_P^* \subseteq \Theta_P \wedge (\mathbf{box}(\theta^1) \neq \mathbf{box}(\theta^2)) \quad (13)$$

Now the ϵ^\nearrow -MOGA algorithm to obtain an ϵ -Pareto front $\mathbf{J}(\hat{\Theta}_P^*)$, which is a well-distributed sample of the Pareto front $\mathbf{J}(\Theta_P)$, will be described. The algorithm, which adjusts the width ϵ_i dynamically, is composed of three

populations (see *Fig. 4*).

[Fig. 4 about here.]

- (1) Main population $P(t)$ explores the searching space D during the algorithm iterations (t). Population size is $Nind_P$.
- (2) Archive $A(t)$ stores the solution $\hat{\Theta}_P^*$. Its size $Nind_A$ can be variable and will never be higher than

$$Nind_max_A = \frac{\prod_{i=1}^s n_box_i + 1}{n_box_{max} + 1} \quad (14)$$

donde $n_box_{max} = \max_i n_box_i$.

- (3) Auxiliary population $G(t)$. Its size is $Nind_G$, which must be an even number.

The pseudocode of the ϵ^{λ} -MOEA algorithm is given by

1. $t := 0$
2. $A(t) := \emptyset$
3. $P(t) := ini_random(D)$
4. $eval(P(t))$
5. $A(t) := store_{ini}(P(t), A(t))$
6. while $t < t_max$ do
7. $G(t) := create(P(t), A(t))$
8. $eval(G(t))$
9. $A(t+1) := store(G(t), A(t))$
10. $P(t+1) := update(G(t), P(t))$
11. $t := t+1$
12. end while

The main steps of the algorithm are detailed as follows:

Step three. $P(0)$ is initialised with $Nind_P$ individuals (models) that have been randomly selected from the searching space D .

Step four and eight. Function **eval** calculates function value (Eqn (8)) for each individual in $P(t)$ (step four) and $G(t)$ (step eight).

Step five. Function **store_{ini}** checks individuals of $P(t)$ that might be included in the archive $A(t)$ as follows:

- (1) Non-dominated $P(t)$ individuals are detected, Θ_{ND} .
- (2) Function space limits are calculated from $J(\Theta_{ND})$.
- (3) Individuals in Θ_{ND} are analyzed, one by one, and those that are not ϵ -dominated by individuals in $A(t)$, will be included in $A(t)$.

Step seven. With each iteration, the function **create** creates $G(t)$ as follows:

- (1) Two individuals are selected randomly, θ^{p1} from $P(t)$, and θ^{p2} from $A(t)$.
- (2) A random number $u \in [0 \dots 1]$ is produced.
- (3) If $u > P_{c/m}$ (probability of crossing/mutation), θ^{p1} and θ^{p2} are crossed over by means of the extended linear recombination technique.
- (4) If $u \leq P_{c/m}$, θ^{p1} and θ^{p2} are mutated using random mutation with Gaussian distribution and then included in $G(t)$.

This procedure is repeated $Nind_G/2$ times until $G(t)$ is filled up.

Step nine. Function **store** checks, one by one, which individuals in $G(t)$ must be included in $A(t)$ on the basis of their location in the function space (see *Fig. 5*). Thus $\forall \theta^G \in G(t)$

- (1) If θ^G lies in the area $Z1$ and is not ϵ -dominated by any individual from $A(t)$, it will be included in $A(t)$ (if its box is occupied by an individual not ϵ -dominated too, then the individual lying farthest away from the $box(\theta^G)$ will be eliminated). Individuals from $A(t)$ which are

ϵ -dominated by θ^G will be eliminated.

- (2) If θ^G lies in the area $Z2$ then it is not included in the archive, since it is dominated by all individuals in $A(t)$.
- (3) If θ^G lies in the area $Z3$, the same procedure is applied as was used with the function \mathbf{store}_{ini} but now applied over a population $P'(t) = A(t) \cup \theta^G$, that is, $\mathbf{store}_{ini}(P'(t), A(t) = \emptyset)$. In this procedure new function limits and ϵ_i widths could be recalculated.
- (4) If θ^G lies in the area $Z4$, all individuals from $A(t)$ are deleted since all of them are ϵ -dominated by θ^G . θ^G is included and function space limits are $\mathbf{J}(\theta^G)$.

[Fig. 5 about here.]

Step 10. Function **update** updates $P(t)$ with individuals from $G(t)$. Every individual θ^G from $G(t)$ is compared with an individual θ^P that is randomly selected from among the individuals in $P(t)$ that are dominated by θ^G . If $\mathbf{J}(\theta^G) \prec \mathbf{J}(\theta^P)$ then θ^G replaces θ^P . θ^P will not be included in $P(t)$ if there is no individual in $P(t)$ dominated by θ^G .

Finally, individuals from $A(t)$ compound the solution $\hat{\Theta}_P^*$ with the multiobjective minimization problem.

2.3 Greenhouse model

For some time now, agricultural engineers have been working to perfect models of the physical and physiological processes that take place inside greenhouses based on mass and energy balances, including the biological behaviour of plants. In (Stanghellini & de Jong, 1995) there is a groundbreaking study on the description of a model of the humidity in a greenhouse that is based

on obtaining a first principles non-linear model of the humidity by defining the balance of condensation, ventilation and transpiration flows. In this last case, the Penman-Monteith equation (Monteith, 1973) is used to incorporate the saturation and radiation deficit measurements so that they can be evaluated. This model is still utilised today to design the ventilation systems in greenhouses (Seginer, 2002). The humidity model is complemented by energetic balance models at different levels. Again, a first principles equation is constructed to include the balance of thermal flows associated with the ventilation, convection, conduction and latent heat due to transpiration by the plants (Baille *et al.*, 1994; Jolliet & Bailey, 1992) that define the temperature evolution. Equations can be defined for the evolution of the temperature in each greenhouse, depending on their different volumes and floor areas, and the interactions among them. The model can vary in its complexity depending on the number of volumes selected, which gives rise to a higher or lower number of differential equations (Blasco, 1999; Rodríguez, 2002).

In this study, the greenhouse is considered to be a volume of air that is delimited by the walls, the roof and the floor thus establishing two subsystems, namely, the volume of air and the floor, this latter acting as a thermal mass (Albright *et al.*, 1985). The state variables that describe the climatic behaviour are temperature \hat{T}_i ($\hat{\cdot}$ is used for the output variables of the model) and humidity $\hat{H}R_i$ (or absolute humidity H_i) in the air and the floor temperature T_m (called the thermal mass temperature).

The water mass balance and air energy balance establish the first two equations of state and the third is set by the energy balance over the thermal mass.

$$\rho v_i \frac{dH_i}{dt} = F_v + C_{sat}(E + fog), \quad (15)$$

$$v_i \rho c_p \frac{d\hat{T}_i}{dt} = Q_s - Q_{cc} + Q_m - Q_v - C_{sat}(Q_e + Q_n). \quad (16)$$

$$A_i C_m \frac{dT_m}{dt} = Q_{sm} - Q_m - Q_f. \quad (17)$$

where: the inside temperature \hat{T}_i and the temperature of the thermal mass T_m are in $^{\circ}C$; the absolute inside humidity H_i is in Kg_{H_2O}/Kg_{air} ; the volume of the greenhouse v_i is given in m^3 and the area A_i in m^2 ; the density of the air ρ is in Kg_{air}/m^3 ; specific heat of the air c_p is in $JKg^{-1}^{\circ}C^{-1}$; the air saturation coefficient C_{sat} is dimensionless and the heat capacity of the thermal mass C_m is given in $Jm^{-2}^{\circ}C^{-1}$.

The flows in the mass balance are as follows (all given in Kg_{H_2O}/s): F_v , renovation flow due to the opening of a window; E , crop evapotranspiration, which is estimated from the Penman-Monteith equation (Monteith, 1973) and has important non linearities, and fog, which is the water that comes from the fog system.

The energy balance terms are (all in W): Q_s , solar energy supplied to the air; Q_{cc} , energy exchanges due to conduction and convection; Q_m , energy exchanges with the thermal mass; Q_e , energy losses due to crop evapotranspiration; Q_n , energy losses due to fogging and Q_v , energy exchange due to ventilation.

The energy balance terms are (all in W): Q_m , energy exchanges between the thermal mass and the inside air; Q_{sm} , energy stored by the thermal mass during the day and Q_f , losses into the ground.

Figure 6 shows a model diagram, from an I/O point of view.

[Fig. 6 about here.]

The output variables are: inside humidity \hat{HR}_i in % and inside temperature \hat{T}_i in °C. The input variables that can be fitted are: window opening control $MV_\alpha \in [0, 100]\%$; heating control $MV_w \in [0, 100]\%$ and fog control $MV_{fog} \in [0, 100]\%$.

Measurable disturbances are: solar radiation S_o in $W \cdot m^{-2}$; outside temperature T_o in °C; outside humidity HR_o in % and wind speed V in m/s .

As can be seen in Appendix A, the model has a large number of parameters. Some of them are easy to determine, for example, the volume and the area of the greenhouse, but others, such as the maximum stomatal conductance are not such simple matters. The complexity of the model and the large number of unknown parameters, together with the fact that some of them vary over time, make them difficult to fit.

3 Materials and experimental details

3.1 Planning the experiments

The operating conditions of a greenhouse are potentially imposed by the effects of disturbances, mainly due to solar radiation and the outside temperature. These disturbances follow a typical behaviour that is repeated daily and which depends largely on the time of year. Thus, it is impossible to give a single set of parameters $\hat{\theta}$ that permits reproducing the greenhouse behaviour throughout

a whole year. The model will therefore be adjusted to cover the dynamics to be found in the summer period (without distinguishing between day and night-time) because this is the time of year when it is most interesting to control the climate inside the greenhouse. At this time of year the heating is not used, and so this actuator is not taken into account.

Hence, the trials will last for a multiple 24-hour period. The more days are used, the better the model will represent the period in question. A large number of days, however, makes the simulations costly to perform and the time required to fit the model increases considerably. This is why, in this case and as a compromise, two (non-consecutive) days were chosen, and the input signals used were those that are normally utilised in the day-to-day operation of the greenhouse. The sampling period was 15 seconds, which is more than enough to capture the dynamics of the processes taking place inside the greenhouse.

For the identification task, data from the 11th and 15th June 2002 were used, and data collected on 20th June, 28th July, 22nd August and 8th September 2002 were used for validation.

Moreover, the data provided by the solar radiation and wind speed sensors have to be processed because, due to the physical nature of the devices and also the magnitudes they are measuring, they introduce noise that has to be eliminated in order to be able to make proper use of the information in the identification task. In this case, a first-order low-pass digital filter was used to filter both solar radiation and wind speed.

3.2 Multiobjective identification details

Before moving on to the multiobjective identification that will be obtained by the Pareto front, the following aspects need to be dealt with: the model adaptation; selection of the parameters to be identified; the procedure for establishing the initial conditions and the optimality criteria to be applied.

With regard to the model adaptation, in the particular case of the greenhouse climatic model (*Fig. 6*), the state Eqn (15), (16) & (17) are adapted directly to the generic equation (1).

In relation to the parameters selection, for the case of hydroponic cultivation of roses in a greenhouse, the candidate set of parameters to be estimated (θ) is associated, on the one hand, with the specific growing of rosebushes and, on the other, with parameters that are associated with different heat transfer constants and reference temperatures inside the greenhouse. The meaning of each parameter to be identified, together with its adjustment range, can be consulted in Notation section. These approximate ranges have been estimated from previous analytical studies, and as a result the searching area is drastically reduced. Thus, adapting the generic problem in state variables to the greenhouse model results in:

$$\theta = [gws_{max} \ gws_{min} \ k \ L \ gwb \ \tau \ a \ G_o \ Ac \dots \dots C_m \ h_m \ T_{ref} \ \alpha_m \ k_a \ fog_{max}]^T, \quad (18)$$

$$\mathbf{u}(t) = [MV_\alpha \ MV_{fog} \ S_o \ T_o \ HR_o \ V]^T, \quad (19)$$

$$\hat{\mathbf{y}}(t) = [\hat{T}_i \ \hat{HR}_i]^T, \quad (20)$$

$$\mathbf{x}(t) = [H_i \ \hat{T}_i \ T_m]^T. \quad (21)$$

The states will be initialised using the real variables measurements. The first state variable H_i can therefore be initialised directly from the value of the outputs $T_i(0)$ and $HR_i(0)$ in the initial moment, as indicated in Appendix A (Eqn (A.7) & (A.8)). The second state variable \hat{T}_i is at the same time an output variable, and it can therefore be initialised with the value this output has in the initial moment $T_i(0)$. Initialisation of the third state variable T_m is not so straightforward because there is no sensor to measure it. The initial value of $T_m(0)$ will be estimated using information about the inputs and outputs in the initial moment and the Eqn (16) for the energy balance in the air that was evaluated in that moment. Since the simulations are started at night (and hence $S_o = 0$) and without activating the fogging system (which is logical), the equation would result as follows:

$$v_i \rho c_p \frac{dT_i(0)}{dt} = -Q_{cc}(0) + Q_m(0) - Q_e(0) - Q_v(0).$$

which, on expansion, would give:

$$\begin{aligned}
v_i \rho c_p \frac{dT_i(0)}{dt} &= -A_i A_c (T_i(0) - T_o(0)) \\
&+ A_i h_m (T_m(0) - T_i(0)) - \lambda E(0) \\
&- \rho c_p AV(0)(a\alpha + G_o)(T_i(0) - T_o(0)).
\end{aligned}$$

It can be assumed that $\frac{d\hat{T}_i(0)}{dt} = 0$ since it varies very little at night (for the summer days that are going to be used in the identification $dT_i(0)/dt \approx 0.1 \cdot 10^{-3} \text{ Cs}^{-1}$), so that $T_m(0)$ would be obtained in the following way:

$$\begin{aligned}
T_m(0) &= \frac{1}{A_i h_m} \left(A_i A_c (T_i(0) - T_o(0)) + \lambda E(0) \right. \\
&\left. + \rho c_p AV(0)(a\alpha + G_o)(T_i(0) - T_o(0)) \right) + T_i(0),
\end{aligned}$$

where:

$$E(0) = \frac{A_i 2L\rho c_p D_i(0) gwb}{\left(\Delta + \gamma \left(1 + \frac{gwb}{gws_{min}} \right) \right) \lambda}.$$

To determine $\hat{\Theta}_p^*$ the minimisation of the third quartile of the identification error of the two greenhouse outputs are going to be considered independently. By so doing, it will become quite clear just how powerful EAs are when it comes to optimising complex functions.

$$J_1(\theta) = \text{Q3}_i |e_1(t_i, \theta)| \tag{22}$$

$$J_2(\theta) = \text{Q3}_i |e_2(t_i, \theta)| W \tag{23}$$

where

$$\begin{aligned} e_1(t_j, \theta) &= T_i(t_j) - \hat{T}_i(t_j, \theta) \\ e_2(t_j, \theta) &= HR_i(t_j) - \hat{H}R_i(t_j, \theta) \end{aligned}$$

and W is a diagonal weighting matrix with:

$$W_{jj} = \begin{cases} -0.02 * \hat{H}R_i(t_j) + 2.2 & \text{if } \hat{H}R_i(t_j) \in [60 \dots 100] \% \\ 1 & \text{if } \hat{H}R_i(t_j) \in [0 \dots 60] \% \end{cases} . \quad (24)$$

The purpose behind using parameter W is to weight any errors that are produced for higher levels of humidity, thus giving them a lower relative importance. The reason for this is related with the commercial humidity sensor that is used, since its accuracy drops notably for these values (due to the condensation that is usually produced on the sensor because of the lack of ventilation). *Fig. 7* shows the relationship between W_{jj} and $\hat{H}R_i(t_j)$. It can be observed that the errors produced for relative humidity values of around 100% are weighted by 0.2; that is to say, they have a lower relative importance than those that occur for humidity values of 60%.

[Fig. 7 about here.]

Next, the following multiobjective optimisation problem is posed

$$\min_{\theta \in D} \mathbf{J}(\theta) = \{J_1, J_2\}. \quad (25)$$

The parameters of the ϵ -MOGA algorithm that were chosen are as follows:

the searching space D is determined by the ranges that were established for each of the 15 parameters of the vector θ , which can be consulted in Notation section; $t_{max} = 60000$; $n_box_1 = n_box_2 = 70$; $Nind_P = 100$; $Nind_G = 4$ and $P_{c/m} = 0.1$.

4 Results and discussion

Figure 8 shows the Pareto front $\mathbf{J}(\hat{\Theta}_P^*)$ that is obtained as a solution to the proposed multiobjective optimisation problem.

[Fig. 8 about here.]

The ideal point \mathbf{J}^{ideal} can be obtained by analysing the Pareto front using the following minimum levels:

$$J_1^{min} = \min J_1(\theta) = 0.618 \text{ } ^\circ C$$

$$J_2^{min} = \min J_2(\theta) = 2.355 \text{ } \%$$

by which

$$\mathbf{J}^{ideal} = \{J_1^{min}, J_2^{min}\} = \{0.618, 2.355\},$$

and one possible ideal model could be obtained as:

$$\begin{aligned} \hat{\theta}^{ideal} &= \arg \min_{\theta \in \hat{\Theta}_P^*} \|\mathbf{J}(\theta) - \mathbf{J}^{ideal}\|_2 = \\ &= [0.010013, 0.0036737, 0.65611, 0.6123, 0.041075, \\ &\quad 0.30059, 0.0011483, 0.0019854, 18.883, 1.4658e^5, \\ &\quad 2.5027, 17.339, 0.11426, 1.9999, 0.0047946], \end{aligned} \tag{26}$$

$$\mathbf{J}(\hat{\theta}^{ideal}) = \{0.845, 2.564\}. \quad (27)$$

Figure 9 shows the response of the $\hat{\theta}^{ideal}$ model with the data utilised in the identification process, while *Fig. 10* shows the response for the days chosen for model validation.

[Fig. 9 about here.]

[Fig. 10 about here.]

As can be seen, the model successfully reproduces the greenhouse behaviour in an appropriate manner, both for the days used in the identification and those utilised in the validation (see Table 1).

[Table 1 about here.]

5 Conclusions

A multiobjective evolutionary algorithm, ϵ -MOGA, based on the concept of ϵ -dominance was developed for robust identification of non-linear processes. Identification is set out as a multiobjective optimisation problem and ϵ -MOGA estimates the non-linear model set $\hat{\Theta}_P^*$ by assuming the simultaneous existence of several optimality criteria. $\mathbf{J}(\hat{\Theta}_P^*)$ results in a well-distributed sample of the optimal Pareto front $\mathbf{J}(\Theta_P)$. With this procedure all kinds of optimality criteria can be used and all sorts of processes can be identified provided that their outputs can be calculated by model simulation (differentiability with respect to the unknown parameters is not necessary).

In this paper a non-linear model of a greenhouse used for roses hydroponic cultivation using first principles has been presented, and the ϵ -MOGA has been used to adjust the 15 unknown parameters of the greenhouse model with data from the summer period in the Mediterranean area. The optimality criteria that were selected for use were the minimisation of the third quartile of the identification error of the two outputs in the model. Being able to establish independent optimality criteria for each output is another advantage of the multiobjective approach because it avoids having to make a decision, *a priori*, about the relative importance of the fit of the inside humidity and temperature.

It has been shown that the $\hat{\theta}^{ideal}$ model, determined *a posteriori* when the Pareto optimal set is available, is a good solution that offers a compromise among the different optimality criteria that have been established.

Acknowledgments

Partially supported by MEC (Spanish government) and FEDER funds: projects DPI2005-07835 and DPI2004-8383-C03-02

References

Alander J (2002). An indexed bibliography of Genetic Algorithms & Pareto and constrained optimization. Tech. rep., Dpt. of Information Technology, University of Vaasa

- Albright D; Seginer I; Marsh L; Oko A** (1985). In situ thermal calibration of unventilated greenhouses. *Journal of Agricultural Engineering Research*, **31**, 265–281
- Baille M; Baille A; Delmon D** (1994). Microclimate and transpiration of a greenhouse rose crop. *Agric. Forest Meteorol*, **71**, 83–87
- Blasco F** (1999). Control predictivo basado en modelos mediante la incorporación de técnicas de optimización heurística. aplicación a procesos no lineales y multivariables [Model based predictive control using heuristic optimization techniques. Application to non-linear and multivariable proceses.] Ph.D. thesis, Universidad Politécnica de Valencia, Valencia
- Boulard T; Baille A** (1993). A simple greenhouse climate control model incorporating effects of ventilation and evaporative cooling. *Agricultural and Forest Meteorology*, **65**, 145–157
- Boulard T; Draoui B** (1995). Natural ventilation of greenhouse with continuous roof vents: Measurements and data analysis. *Journal of Agricultural Engineering Research*, **61**, 27–36
- Boulard T; Draouiand B; Neirac F** (1996). Calibration and validation of a greenhouse climate control model. *Acta Horticulture*, **406**, 49–61
- Coello C; Toscano G; Mezura E** (2005). Current and future research trends in evolutionary multiobjective optimization. *Information Processing with Evolutionary Algorithms. Industrial Applications to Academic Speculations*, 213–231
- Coello C; Veldhuizen D; Lamont G** (2002). *Evolutionary algorithms for solving multi-objective problems*. Kluwer Academic Publishers
- Ehrlich H; Kühne M; Jäkel J** (1996). Development of a fuzzy control system for greenhouses. *Acta Horticulturae*, **406**, 463–470
- Fonseca C** (1995). Multiobjective genetic algorithms with application to con-

- trol engineering problems. Ph.D. thesis, Dpt. of Automatic Control and Systems Engineering. University of Sheffield
- Joliet O; Bailey B** (1992). The effect of climate on tomato transpiration in greenhouses: measurements and models comparison. *Agric. Forest. Meteor.*, **58**, 43–62
- Pronzalo E; Walter E** (1997). Identification of parametric models from experimental data. Springer
- Laumanns M; Thiele L; Deb K; Zitzler E** (2002). Combining convergence and diversity in evolutionary multi-objective optimization. *Evolutionary Computation*, **10**(3), 263–282
- Linker R; Seginer I; Gutman P** (1998). Optimal CO₂ control in a greenhouse modelled with neural networks. *Computers and Electronics in Agriculture*, **19**, 289–310
- Monteith J** (1973). Principles of environmental physics. Contemporary biology. Edward Arnold Ed., UK
- Nielsen B; Madsen H** (1996). Identification of transfer functions for control of greenhouse air temperature. *Journal of Agricultural Engineering Research*, **60**, 25–34
- Rodríguez F** (2002). Modelado y control jerárquico de crecimiento de cultivos en invernadero. [Modelling and hierarchical control of crop growing in greenhouse.] Ph.D. thesis, Universidad de Almería, Almería
- Seginer I** (2002). The Penman-Monteith evapotranspiration equation as an element in greenhouse ventilation design. *Journal of Agricultural Engineering Research*, **82**(4), 423–439
- Seginer I; Boulard T; Bailey J** (1994). Neural network models of the greenhouse climate. *Journal of Agricultural Engineering Research*, **59**(3), 203–216

- Stanghellini C; de Jong T** (1995). A model of humidity and its applications in a greenhouse. *Agricultural and Forest Meteorology*, **76**, 129–148
- Young P; Chotai A; Tych W** (1993). Identification, estimation and true digital control of glasshouse system. *The Computerized Greenhouse*, 3–50
- Zhang Z** (1996). Parameter estimation techniques: A tutorial with application to conic fitting. *Image and Vision Computing Journal*, **15**(1), 59–76
- Zitzler E** (1999). Evolutionary algorithms for multiobjective optimization: Methods and applications. Ph.D. thesis, Swiss Federal Institute of Technology Zurich

Notation

Range of possible value are indicated for identifiable parameters. Exact value are indicated for constant or know parameters.

A Windows area, 130 m^2

A_c Loss coefficient of conduction and convection, [2, 20]

A_i Greenhouse surface area, 240 m^2

a Constant for renewal volumetric flow, [0.0005, 0.01]

C_m Thermal mass heat capacity, [100000, 300000] $J \text{ }^\circ\text{C}^{-1} \text{ m}^{-2}$

c_p Air heat capacity, $1003 \text{ J Kg}^{-1} \text{ }^\circ\text{C}^{-1}$

C_{sat} Air saturation coefficient, dimensionless

D_i Air water vapor deficit, KPa

E Crop evapotranspiration, Kg_{H_2O}/s

F_v Water rate in the air renewal flow, Kg_{H_2O}/s

fog Water rate of fog system, Kg_{H_2O}/s

fog_{max} Maximum water rate of fog system, [0.001, 0.005] Kg_{H_2O}/s

F_v Water rate in the air renewal flow, Kg_{H_2O}/s

G Renewal air flow, m^3/s

G_o Losses of renewal air flow, [0.0005, 0.01]

gwb Boundary-layer conductance, [0.001, 0.05] m/s

gws Stomatal conductance, m/s

gws_{max} Maximum stomatal conductance, [0.01, 0.03] m/s

gws_{min} Minimum stomatal conductance, [0.0001, 0.005] m/s

h_m Conductivity coefficient between air and thermal mass, [1,20] $\text{W m}^{-1} \text{ }^\circ\text{K}^{-1}$

H_i Inside absolute humidity, Kg_{H_2O}/Kg_{air}

H_o Outside absolute humidity, Kg_{H_2O}/Kg_{air}

H_{sat} Absolute saturation humidity, Kg_{H_2O}/Kg_{air}
 $\hat{H}R_i$ Inside relative humidity, %
 HR_o Outside relative humidity, %
 k Extinguishing coefficient of radiation, [0.1, 0.7]
 k_a Conductivity coefficient between thermal mass and ground, $[0.5, 10] W m^{-1} \circ K^{-1}$
 L Leaves area index, $[0.5, 2] m_{leaves}^2/m_{ground}^2$
 MV_α Windows opening manipulated variable, %
 MV_{fog} Fog system manipulated variable, %
 MV_W Heating system manipulated variable, %
 P Atmospheric pressure, 98.1 KPa
 $psat$ Saturation pressure, KPa
 Q_{cc} Energy exchange by conduction and convection phenomena, W
 Q_e Energy loss due to crop evapotranspiration, W
 Q_f Energy loss through ground, W
 Q_m Energy exchange with thermal mass, W
 Q_n Energy loss by nebulization, W
 Q_s Solar energy supplied to air volume, W
 Q_{sm} Energy stored by the thermal mass during the day, W
 Q_v Energy exchange due to window ventilation, W
 Rn Solar radiation absorbed by the crop, W/m^2
 S_o Solar radiation, W/m^2
 \hat{T}_i Inside temperature, $^{\circ}C$
 T_m Thermal mass temperature, $^{\circ}C$
 T_o Outside temperature, $^{\circ}C$
 T_{ref} Ground temperature at reference depth, [10, 20] $^{\circ}C$
 V Wind speed, m/s
 v_i Greenhouse volume, 850 m^3

- W Energy from heating system, W
- W_{max} Maximum power of heating system, $5000 W$
- z_{ref} Reference depth, $6 m$
- α Opening window angle, $^\circ$
- α_m Rate of absorbed heat by thermal mass, $[0.01, 0.3]$
- α_{max} Maximum window angle, 12°
- Δ Slope of water vapour saturation, $KPa/^\circ C$
- γ Psychrometric constant, $0.066 KPa/^\circ C$
- λ Latent heat of vaporization, J/Kg
- ρ Air density, $1.25 Kg_{air}/m^3$
- τ Transmission coefficient of the greenhouse, $[0.3, 0.9]$

A Appendix

Complementary equations of Greenhouse model.

Opening window angle:

$$\alpha = \frac{MV_\alpha}{100} \alpha_{max} \quad (A.1)$$

Water rate of fog system:

$$fog = \frac{MV_{fog}}{100} fog_{max} \quad (A.2)$$

Energy from heating system:

$$W = \frac{MV_W}{100} W_{max} \quad (A.3)$$

Water rate in the air renewal flow:

$$F_v = \rho G (H_o - H_i) \quad (A.4)$$

Renewal air flow (Boulard & Draoui, 1995):

$$G = AV(a\alpha + G_o) \quad (\text{A.5})$$

Air saturation coefficient:

$$C_{sat} = \begin{cases} 1 & H_i < H_{sat} \\ 0 & H_i = H_{sat} \end{cases} . \quad (\text{A.6})$$

Absolute to relative humidity conversion:

$$HR = \begin{cases} 100 & HR > 100 \\ HR & HR \leq 100 \end{cases} ,$$

$$HR = \frac{100H \cdot P}{0.611psat(T)}, \quad (\text{A.7})$$

$$p_{sat}(T) = 0.61 \left[1 + 1.414 \sin(5.82e^{-3}T) \right]^{8.827} . \quad (\text{A.8})$$

Depending on different cases, (T, HR, H) corresponds to the inside $(\hat{T}_i, \hat{H}R_i, H_i)$ or the outside (T_o, HR_o, H_o) of the greenhouse. It also makes it possible to calculate the saturation absolute humidity H_{sat} corresponding to $HR = 100\%$.

Crop evapotranspiration (Monteith, 1973):

$$E = \frac{A_i(\Delta Rn + 2L\rho c_p D_i gwb)}{\left[\Delta + \gamma \left(1 + \frac{gwb}{gws} \right) \right] \lambda}, \quad (\text{A.9})$$

$$\Delta = p_{sat}(\hat{T}_i + 0.5) - p_{sat}(\hat{T}_i - 0.5), \quad (\text{A.10})$$

$$Rn = (1 - e^{kL})\tau So, \quad (\text{A.11})$$

$$D_i = p_{sat}(\hat{T}_i) \left[1 - \frac{\hat{H}R_i}{100} \right], \quad (\text{A.12})$$

$$\lambda = (3.1468 - 0.002365(\hat{T}_i + 273))10^6, \quad (\text{A.13})$$

$$gws = gws_{min} + (gws_{max} - gws_{min}) \cdot \left[1 - \exp\left(-\frac{\tau So}{160}\right) \right] g_D, \quad (\text{A.14})$$

$$g_D = \begin{cases} \frac{0.39}{0.029 + D_i} & D_i \geq 0.361 \\ 1 & D_i < 0.361 \end{cases}.$$

Solar energy supplied to air volume:

$$Q_s = A_i \tau So \quad (\text{A.15})$$

Energy exchange by conduction and convection phenomena:

$$Q_{cc} = A_i A_c (\hat{T}_i - T_o) \quad (\text{A.16})$$

Energy loss due to crop evapotranspiration:

$$Q_e = \lambda E \quad (\text{A.17})$$

Energy exchange due to window ventilation:

$$Q_v = \rho c_p G (\hat{T}_i - T_o) \quad (\text{A.18})$$

Energy loss by nebulization:

$$Q_n = \lambda f o g \tag{A.19}$$

Energy exchange between thermal mass and inside air:

$$Q_m = A_i h_m (T_m - \hat{T}_i) \tag{A.20}$$

Energy stored by the thermal mass during the day:

$$Q_{sm} = \alpha_m Q_s \tag{A.21}$$

Energy loss through ground:

$$Q_f = A_i k_a \left(\frac{T_m - T_{ref}}{z_{ref}} \right) \tag{A.22}$$

Figure captions

Fig. 1. The parameter identification problem by optimising the optimality criteria $J(\theta)$. $\hat{\theta}$, optimal parameters; $\mathbf{u}(t)$, model/process inputs; $\mathbf{y}(t)$, process outputs; $\hat{\mathbf{y}}(t)$, outputs model and $\mathbf{e}(t)$, identification error

Fig. 2. A solid line represents the Pareto optimal set Θ_P and the Pareto front these optimals produce $\mathbf{J}(\Theta_P)$ for an MOP with two dimensions. Squares are used to represent a possible sample from the Pareto optimal set $\hat{\Theta}_P^*$ and the distributed sample from the Pareto front produced by these optimals $\mathbf{J}(\hat{\Theta}_P^*)$

Fig. 3. The concept of ϵ -dominance. ϵ -Pareto Front $\mathbf{J}(\hat{\Theta}_P^*)$ in a two-dimensional problem. J_1^{\min} , J_2^{\min} , J_1^{\max} , J_2^{\max} , limits space; ϵ_1 , ϵ_2 box widths; and n_box_1 , n_box_2 , number of boxes for each dimension

Fig. 4. ϵ -MOEA algorithm structure. $P(t)$, the main population; $A(t)$, the archive; $G(t)$ the auxiliary population

Fig. 5. Function space areas (Z) and limits (J). (a) two-dimensional case (b) three-dimensional case

Fig. 6. Greenhouse climatic model

Fig. 7. $W_{jj} = f(\hat{H}R_i(t_j))$. $\hat{H}R_i(t_j)$ inside relative humidity at sample t_j ; W_{jj} , element jj of diagonal weighting matrix W

Fig. 8. Pareto front $\mathbf{J}(\hat{\Theta}_P^*)$. $\mathbf{J}(\hat{\theta}^{ideal})$ is highlighted by (*)

Fig. 9. Real (solid line) and simulated (dotted line) inside humidity and inside temperature (top and bottom, respectively) with $\hat{\theta}^{ideal}$ for the 11th and 15th June 2002

Fig. 10. Real (solid line) and simulated (dotted line) inside humidity and inside temperature (top and bottom, respectively) with $\hat{\theta}^{ideal}$ for 20th June, 28th July, 22nd August and 8th September 2002

Table 1
Optimality criteria $\mathbf{J}(\hat{\theta}^{ideal})$ obtained with $\hat{\theta}^{ideal}$ for the validation data

<i>Date</i>	$\mathbf{J}(\hat{\theta}^{ideal})$
20-6-02	{0.75, 3.64}
28-7-02	{1.43, 3.06}
22-8-02	{0.99, 3.95}
8-9-02	{1.06, 2.40}

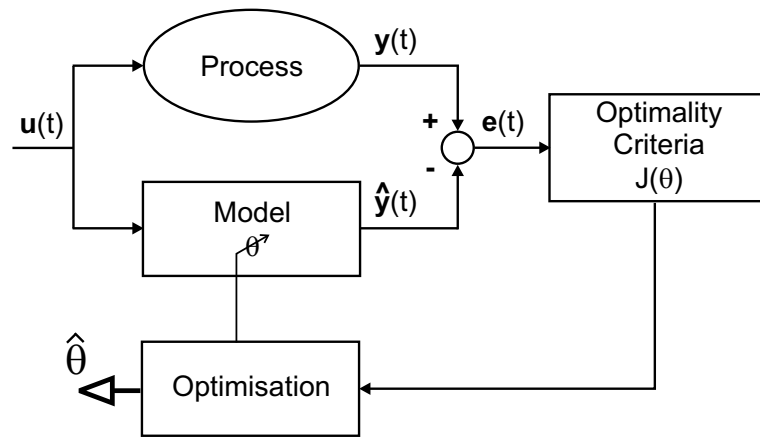


Fig. 1. The parameter identification problem by optimising the optimality criteria $J(\theta)$. $\hat{\theta}$, optimal parameters; $\mathbf{u}(t)$, model/process inputs; $\mathbf{y}(t)$, process outputs; $\hat{\mathbf{y}}(t)$, outputs model and $\mathbf{e}(t)$, identification error

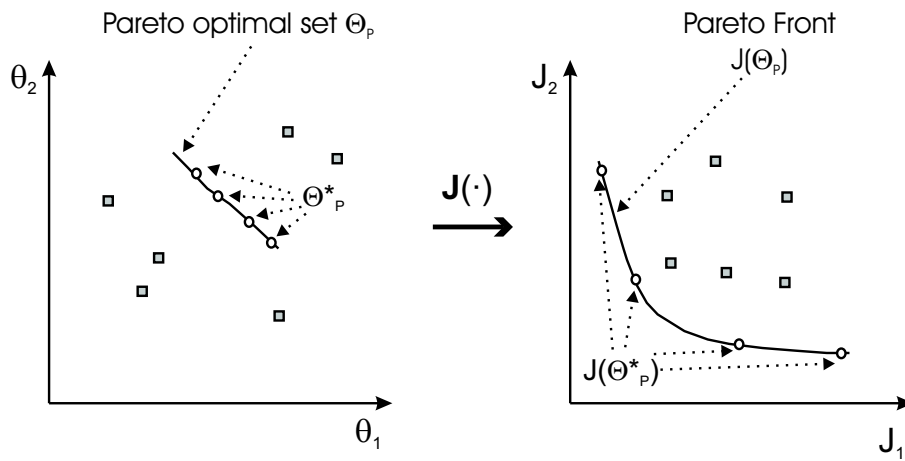


Fig. 2. A solid line represents the Pareto optimal set Θ_P and the Pareto front these optimals produce $J(\Theta_P)$ for an MOP with two dimensions. Squares are used to represent a possible sample from the Pareto optimal set $\hat{\Theta}_P^*$ and the distributed sample from the Pareto front produced by these optimals $J(\hat{\Theta}_P^*)$

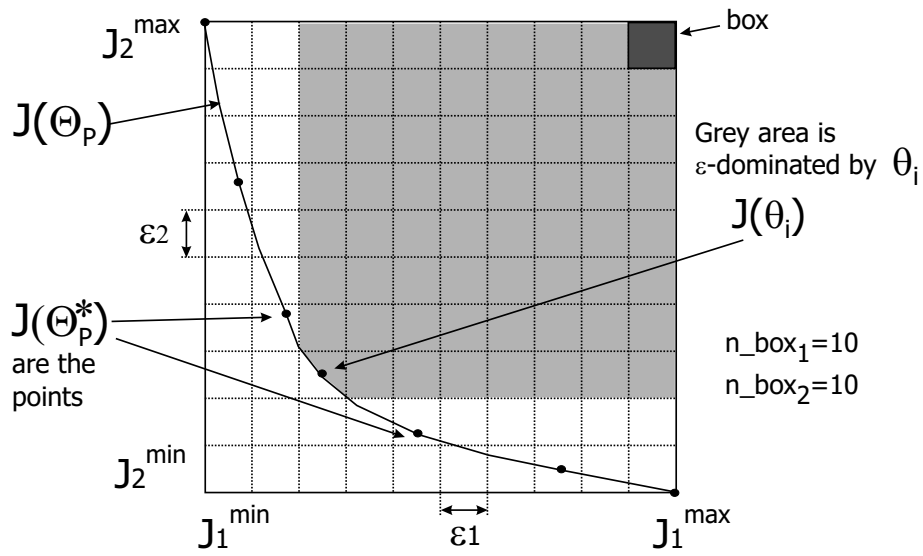


Fig. 3. The concept of ϵ -dominance. ϵ -Pareto Front $J(\hat{\Theta}_P^*)$ in a two-dimensional problem. J_1^{\min} , J_2^{\min} , J_1^{\max} , J_2^{\max} , limits space; ϵ_1 , ϵ_2 box widths; and n_{box_1} , n_{box_2} , number of boxes for each dimension

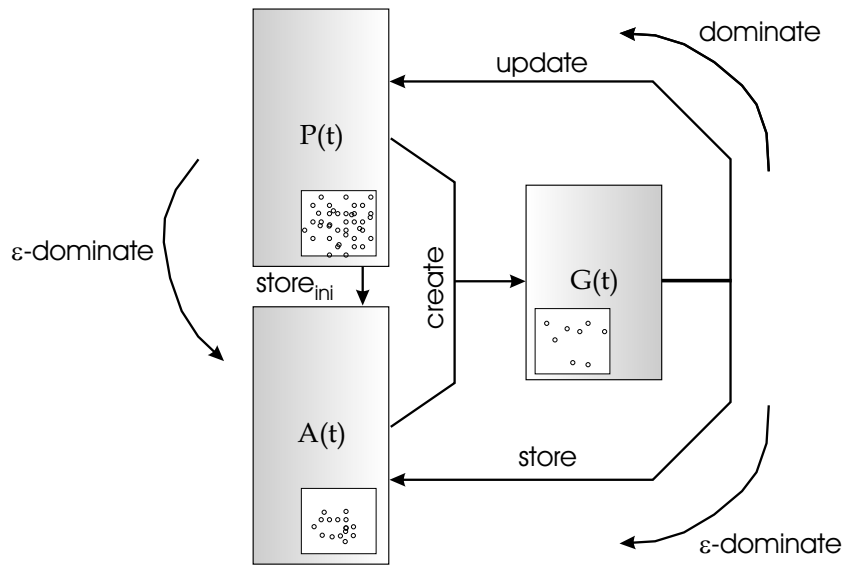


Fig. 4. ϵ -MOEA algorithm structure. $P(t)$, the main population; $A(t)$, the archive; $G(t)$ the auxiliary population

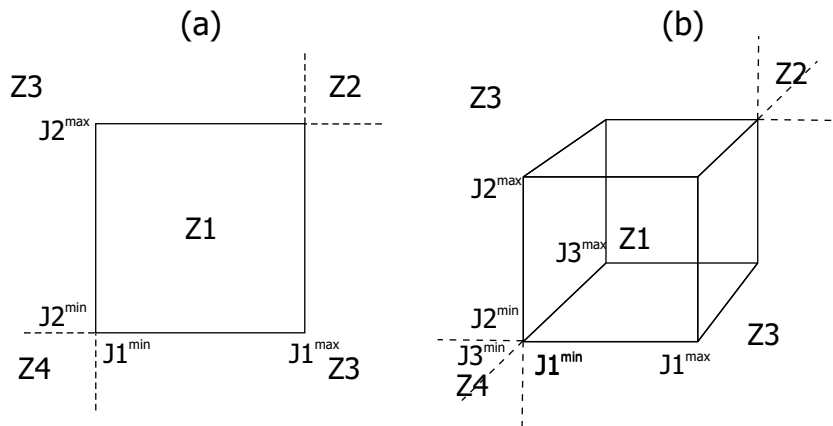


Fig. 5. Function space areas (Z) and limits (J). (a) two-dimensional case (b) three-dimensional case

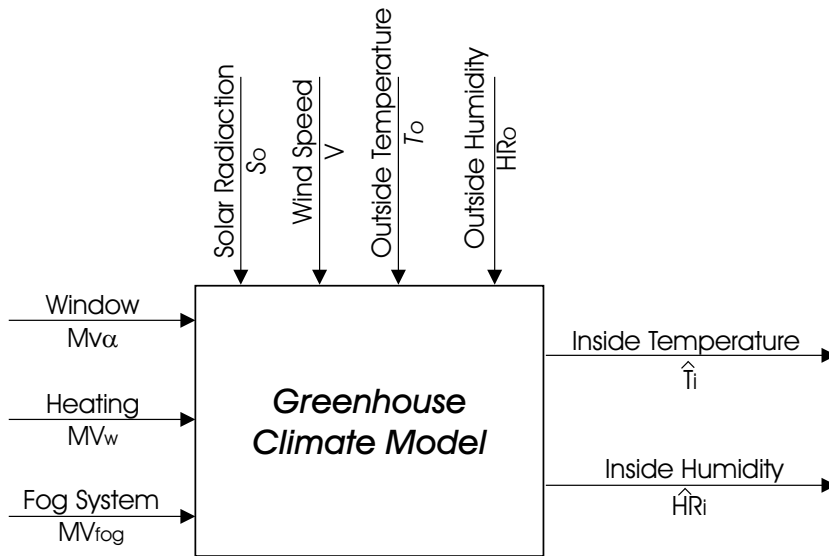


Fig. 6. Greenhouse climatic model

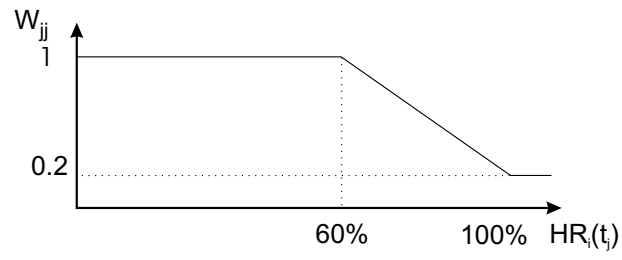


Fig. 7. $W_{jj} = f(\hat{HR}_i(t_j))$. $\hat{HR}_i(t_j)$ inside relative humidity at sample t_j ; W_{jj} , element jj of diagonal weighting matrix W

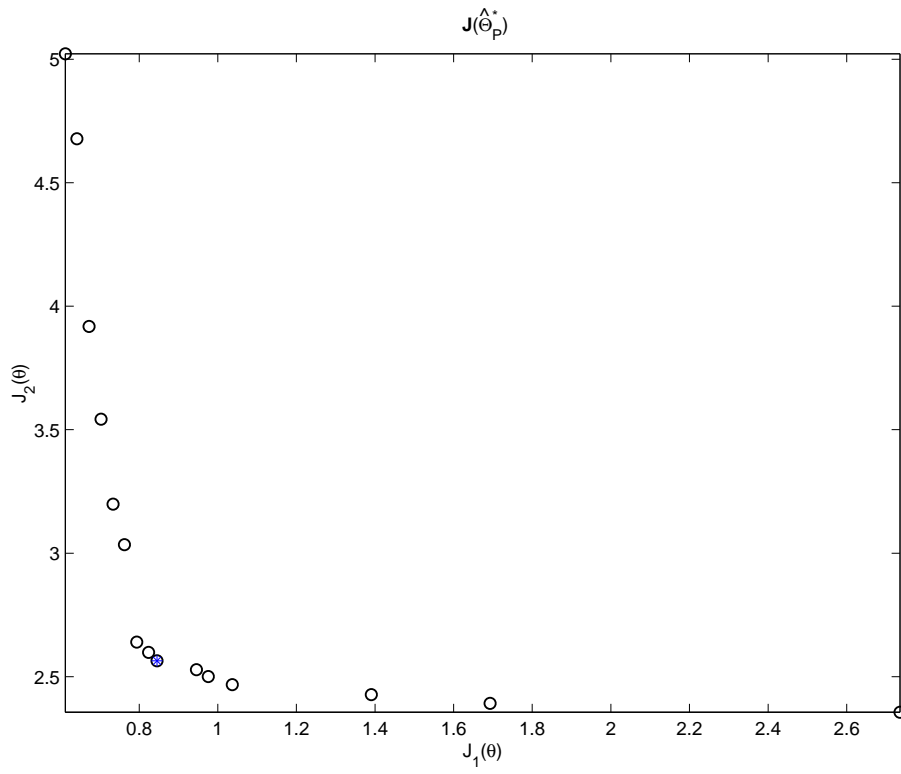


Fig. 8. Pareto front $\mathbf{J}(\hat{\Theta}_P^*)$. $\mathbf{J}(\hat{\theta}^{ideal})$ is highlighted by (*)

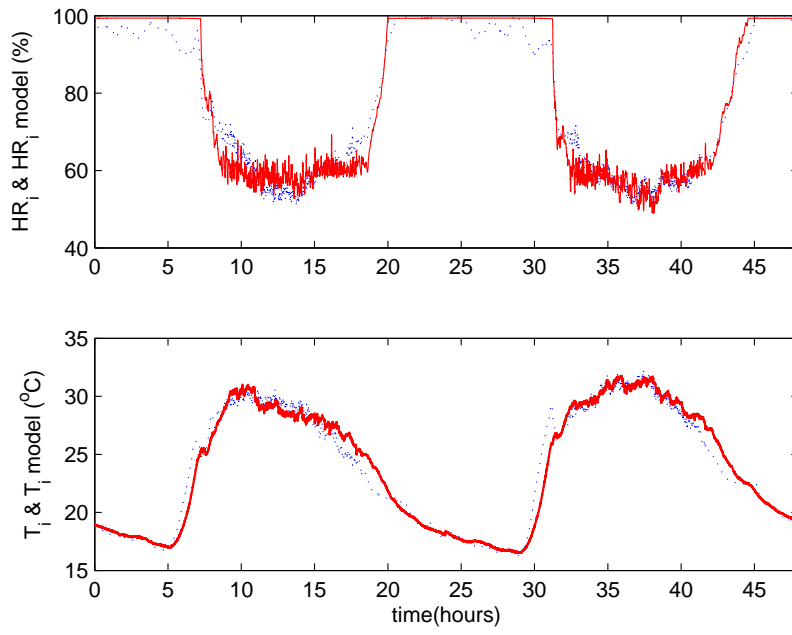


Fig. 9. Real (solid line) and simulated (dotted line) inside humidity and inside temperature (top and bottom, respectively) with $\hat{\theta}^{ideal}$ for the 11th and 15th June 2002

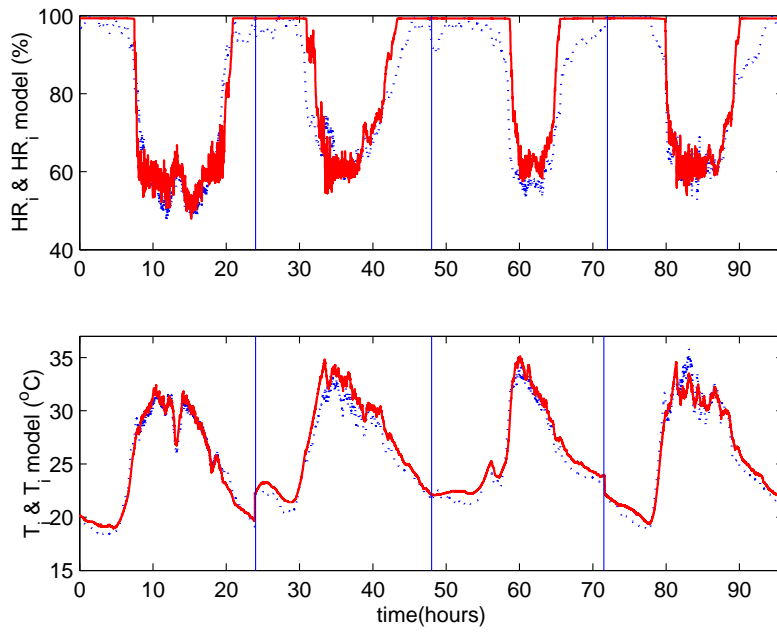


Fig. 10. Real (solid line) and simulated (dotted line) inside humidity and inside temperature (top and bottom, respectively) with $\hat{\theta}^{ideal}$ for 20th June, 28th July, 22nd August and 8th September 2002