# GREENHOUSE CLIMATE MODELS: AN OVERVIEW

***José Boaventura Cunha, jboavent@utad.pt***

*UTAD – Universidade de Trás-os-Montes e Alto Douro, Dep. Engenharias*
*CETAV – Centro de Estudos e de Tecnologias do Ambiente e Vida*
*5001-911 Vila Real, Portugal*

**Abstract:** Greenhouse climate and crop models are essential for improving environmental management and control efficiencies. In this paper, are described several types of models that could be used to simulate and predict the greenhouse environment, as well as the tuning methods to compute their parameters. This study focuses on the dynamical behaviours of the inside air temperature, humidity and carbon dioxide concentration models and their domains of application. Linear and non-linear models will be covered, focusing on issues such as: physical models, black-box models, and neural networks models. Several experiments will be presented to illustrate the performance of each model in the simulation and prediction of the greenhouse climate. The models are described as functions of the outside climate, the control actions performed and the transpiration and photosynthesis responses of the plants. The data used to compute the simulation models were acquired in an experimental greenhouse using a sampling time interval of 1 minute. The greenhouse is automated with several actuators and sensors that are connected to an acquisition and control system based on a personal computer.

**Keywords:** Crop models, Greenhouse climate models, Neural networks, System identification

## 1. Introduction

To improve the management and control of a greenhouse climate, an adaptive climate control strategy must be used to compute the optimal control signals used for a defined cost performance function. The adaptive *PID* – Proportional Integral Derivative controller structure showed in *figure 1* could be used for this purpose. The adaptive controller needs the use of greenhouse climate models to predict future outputs based on past and current inputs, the expected control actions and the predicted weather inputs. An optimiser module computes the proportional, integral and derivative gains, in order to minimize or maximize the specified cost function according to the constraints imposed. Afterwards, these parameters are sent to the *PID* closed controlled loop.

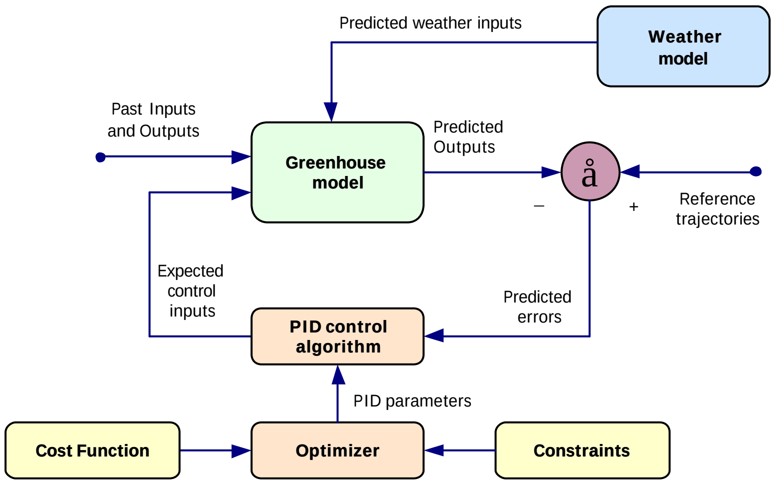


**Figure 1. Structure of the climate adaptive PID controller**

Normally, the optimisation is performed to maximise a cost function which has a positive term related with the expected crop economic value and a negative term related with the operation costs of the climate conditioning equipment. Also, constraints for the physical and physiological values of the actuators, the environment and the crop are applied. This paper describes possible implementations for the models used by the controller.

## 2. Greenhouse climate models

Simulation models to describe the dynamic behaviour of the air temperature and humidity and dioxide concentration inside the greenhouses have been published in several studies. These models could be based on energy and mass flows equations (Boulard et al., 1993, Bot, 1991), or derived by using a system identification approach using linear and non-linear techniques, such as the recursive least squares algorithms and neural networks to tune the parametric models (Boaventura Cunha et al., 2000; J. P. Coelho et. al, 2002).

### *2.1. Physical based Models*

Physical models describe the flow and mass transfers generated by the differences in energy and mass content between the inside and outside air, or by the control or exogenous energy and mass inputs, as showed in equations 1 and 2 (Bakker et.al., 1994).

$$\frac{dT_i}{dt} = \frac{1}{C_{aph}}\left(q_{in,h} - q_{out,h} + p_h\right) \qquad (1)$$

$$\frac{dc_m}{dt} = \frac{1}{V}\left(q_{in,m} - q_{out,m} + p_m\right) \qquad (2)$$

where $T_{ag}$ is the air temperature, $C_{aph}$ the thermal capacity, $q_{in,h}$ and $q_{out,h}$ the energy inflow and outflow, $p_h$ the energy production per unit of time, $c_m$ the mass concentration, $q_{in,m}$ and $q_{out,m}$ the mass inflow and outflow and $p_m$ the mass produced per unit of time referred to the greenhouse volume $V(m^3)$.

In the previous equations, the transport mechanisms for conduction, convection and radiation are implicit. For instance, the heat flux from inside to outside due to ventilation, which is a term of $q_{out,h}$ in equation1, is described by the following equation:

$$q_{vent,h} = q_v r c_p (T_i - T_o) \qquad (3)$$

where $q_v$ is the volumetric flux trough the windows, $rc_p$ the volumetric specific heat, $T_i$ the inside temperature and $T_o$ the outside temperature.

The drawback of this methodology is that the development of these models are difficult to tune in practice, since they use a large number of parameters and physical variables. Moreover, when properly tuned, they can only provide good predictions over short future time horizons, since the greenhouse-crop system is time variant.

### 2.2. Black-box Linear Parametric Models

This method is based on experimentation where the input $u$ and output $y$ signals from the system to be identified, *figure 2*, are recorded and subjected to data analysis in order to infer a model. This procedure is known as system identification (Ljung, 1987). In this case, the output signal vector $y$ is formed by the measurements of the inside air temperature and humidity, $T_i$ and $RH_i$. These models must be related to the external influences of the outside weather conditions, as well to the control signals performed on the greenhouse actuator equipment.

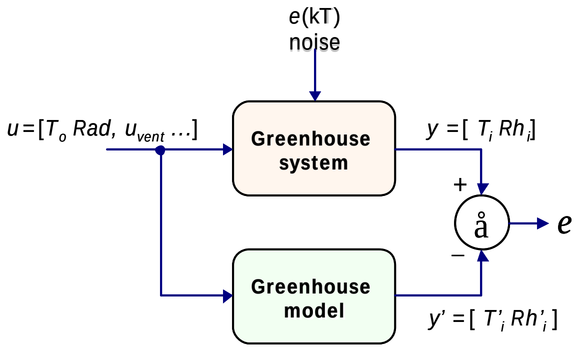


**Figure 2. System and model representation**

In this figure, $e$, $y'$ and $e$ denote a noise signal, the model output and the prediction error. The model inputs $u$ are formed by the measurements of the outside temperature $T_o$, Solar irradiation *Rad*, ventilator and heater control signals $u_{vent}t$ and $u_{heater}$, among other relevant measurements.

The system identification procedure could be developed assuming that the greenhouse climate can be described as a linear system around a particular operating point. In this way, linear parametric ARX models can be employed to describe the dynamics of the greenhouse climate system (Boaventura Cunha et al., 1997; 2000, Ljung, 1987). Previous work has shown that the second-order ARX model equations 4 and 5 describe the dynamics of the air temperature and relative humidity well.

$$T_{is}(kT) = \frac{\left[ B_{1,t} \; B_{2,t} \; B_{3,t} \; B_{4,t} \right]}{1 + a_{1,t}q^{-1} + a_{2,t}q^{-2}} \begin{bmatrix} T_i(kT) - T_o(kT) \\ Rad(kT) \\ (T_{pipe} - T_i) \times u_{heat}(kT) \\ (T_i - T_o) \times u_{vent}(kT) \end{bmatrix} \tag{4}$$

$$Rh_{is}(kT) = \frac{\left[ B_{1,h} \;\; B_{2,h} \;\; B_{3,h} \;\; B_{4,h} \right]}{1 + a_{1,h}q^{-1} + a_{2,h}q^{-2}} \begin{bmatrix} (Rh_i - Rh_o) \times u_{vent}(kT) \\ T_i(kT) \\ Rh_i(kT) - Rh_o(kT) \\ (T_{pipe} - T_i) \times u_{heat}(kT) \end{bmatrix} \tag{5}$$

where: $a_i$ denotes the denominator parameters of the transfer functions, $B_i$ the polynomials in the delay operator, $T_o$ the outside air temperature, *Rad* the outside solar radiation, $T_i$ and $Rh_i$ the measured inside air temperature and humidity, $T_{pipe}$ the temperature of the heating pipes and $u_{vent}$, $u_{heat}$ the ventilation and heating inputs.

However, the simulations computed with this second approach, for sets of data that were not used to compute the models parameters, are more sensitive to mismatches compared to the physical models. This is due to the fact that the ARX black-box models are a great simplification of the entire system. Therefore, the parameters are time-varying. As an example, the process of solar irradiation conversion to heat varies throughout the day and the year, since the Sun elevation and the optical properties of the cover varies in time.

To overcome this difficulty, recursive estimation algorithms must be implemented to compute the time-varying parameters of the transfer functions of equations 4 and 5, (Aström et al., 1989). Since the parameters are slow time-variant and the signals have no excitation for long periods of time, an estimator that forgets the information only in the directions in which new information is gathered must be used. The following recursive least squares algorithm with exponential forgetting, equations 6.1 to 6.3 can be employed with this objective, (Boaventura Cunha et al., 2000; Salgado et al., 1988),

$$q(k) = q(k-1) + K(k)(y(k) - j^{T}(k)q(k-1)) \tag{6.1}$$

$$K(k) = \frac{P(k-1)j(k)}{u(k) + j(k)^{T}P(k-1)j(k)(1-a(k)u(k))} \tag{6.2}$$

$$P(k) = P(k-1) - \frac{P(k-1)j(k)j(k)^{T}P(k-1)}{(u(k)^{-1} - a(k))^{-1} + j(k)^{T}P(k-1)j(k)} \tag{6.3}$$

where $q$ denotes the estimated model parameters (coefficients of polynomials $A$ and $B$) and $K$ and $P$ are the gain and covariance matrixes. For this estimator, the estimations converge to values such that $P$(k)=$aI$, where typically $0.0001 < a < 0.001$, and $j$(k) and $u$(k) are the regression vector and the estimate of the variance of the residuals.

## *2.3. Black-box Non-linear Parametric Models*

Artificial neural networks are collections of mathematical models that reproduce some of the observed properties of biological nervous systems. The key element of the ANN is the structure of the information processing system. This system is composed of a large number of highly interconnected processing elements that are analogous to neurons and are coupled together with weighted connections that are analogous to synapses.

Non-linear autoregressive models are potentially more powerful than linear ones because they can model more complex underlying characteristics of the data. There are a broad number of ANNs topologies. Among the most widespread are feedforward networks. In this paper, a multilayer perceptrons (MLP) network with a hyperbolic tangent (tanh) activation function is used. These types of structure have proved to be universal aproximators (Hornik et al., 1989). This means that they can approximate any reasonable function $f$ with a subjective accuracy given by:

$$f(u) = \left( \sum_{j=1}^{k} v_{jl} t \left( \sum_{i=1}^{n} w_{ij} \cdot u_i - q_j \right) - q_l \right), l = 1...m \tag{7}$$

where $t$ is the tanh function, k is the number of hidden units, $v_{jl}$ and $w_{ij}$ are weights, $q_i$ are biases and $u$ the data vector.

In this work, the non-linear function $f$ is estimated based on data samples using the Lavenberg-Marquardt optimisation technique. The Lavenberg-Marquardt is the standard method for minimization of mean square error criteria, due to its rapid convergence properties and robustness (Marquardt, 1963). Neural networks have several major drawbacks. They require large numbers of data samples due to their large number of degrees of freedom. Problems such as over-fitting and sub-optimal minima may occur more severely than in the linear case. Also, this method requires a large computation time for training, i.e. for learning the system behaviour, which restricts its application to real-time implementations.

## 3. Results and conclusions

The inside air temperature and humidity simulation models were identified using the described approaches for a greenhouse located in the north of Portugal. The greenhouse has a floor area of 210m$^2$, covered with 200μm polyethylene film. Several actuators and sensors are installed and connected to an acquisition and control system based on a Personal computer and a data acquisition and control card (*PCL-818*, from Advantech) using a sampling interval of 1 minute.

In figure 3, the simulation results achieved with the physical model for the air temperature and relative humidity are shown. The model parameters were computed off-line using the data of the month of January 2000, and the simulations were performed for a validation data set of the first 6 days of March of the same year.

Table 1 shows the performance results of the physical and parametric models for the pure simulation and the 60 step ahead predictions, which corresponds to predictions in the future time horizon of 60 minutes.

In this case, the data sets used to compute the models parameters and the simulations were the same of the previous case and the parametric models were estimated off-line using a least squares algorithm and a neural network. The criteria performance used is the root mean squared errors, $RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{n} E_k^2}$, with N being the size of the data samples and Ek the error between simulated and measured values.

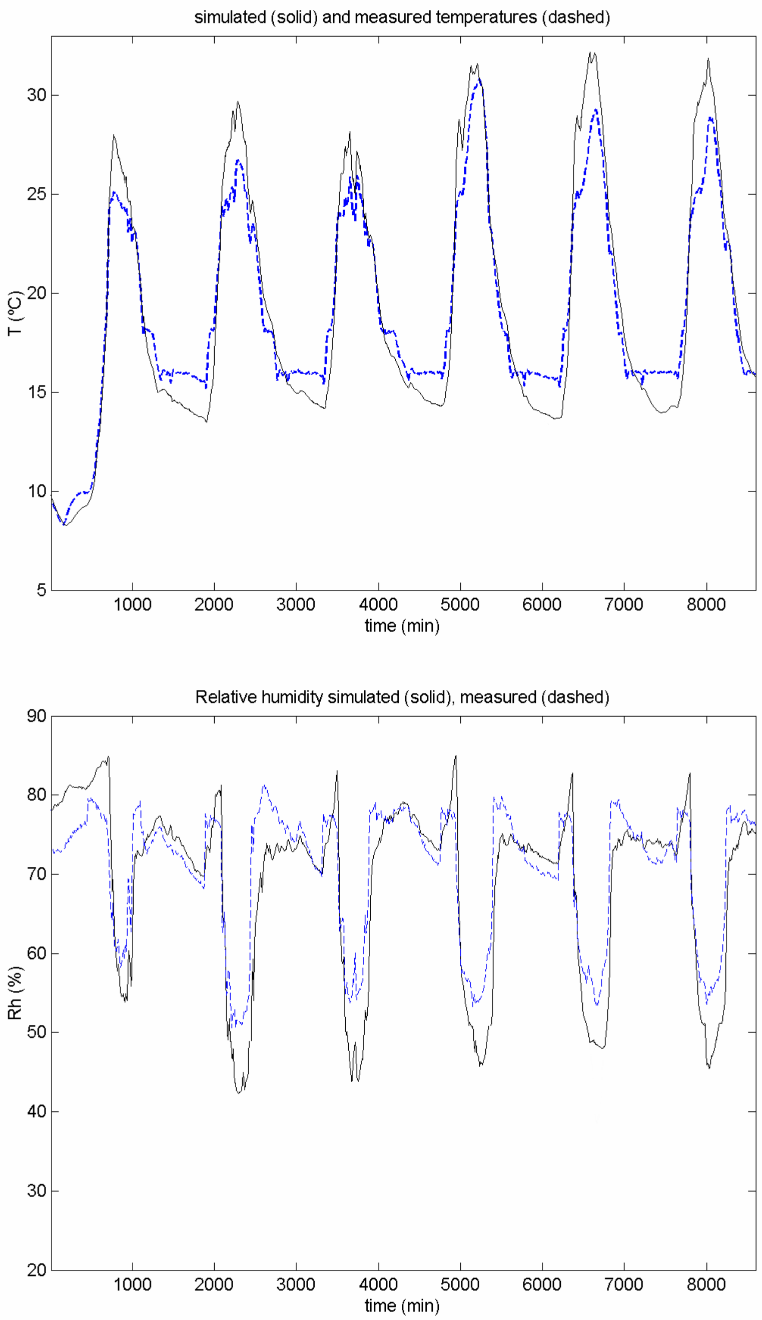


**Figure 3. Measured and simulated temperatures (top) and *relative humidity* (bottom)**

Table1 – Performance of the air temperature and relative humidity models for the data set validation

| Model | Air temperature | | Air relative humidity | |
|---|---|---|---|---|
| | *RMSE* (simulation) | RMSE (60 min prediction) | RMSE (simulation) | RMSE (60 min prediction) |
| Physical | 1.4063 | 0.1928 | 4.9401 | 0.6914 |
| ARX | 2.0022 | 0.2231 | 5.2810 | 0.9933 |
| Neural Network | 1.8309 | 0.2198 | 5.1772 | 0.9614 |

This study has drawn several comparisons between physical, linear and non-linear modelling techniques applied to simulate the inside greenhouse climate. It was observed from various simulations that non recursive physical models give better results than other off-line methods, when are used validation data sets periods distant in time from the data sets used to compute the models. However, computation times for ARX models are much lower, and if recursive estimation techniques are applied, the results achieved for the short time prediction horizon, from 1 to 60 minutes, are better with these models. The use of neural networks models has the major drawback of requiring a large computation time for training, which restricts their application to real-time implementations.

***REFERENCES***

Aströrn, K. J. And B. Wittenmark (1989). Adaptive control, Addison - Wesley, Massachusetts.

Bakker, J. C., Bot, G.P.A., Challa, H., N.J. van de (1994). Greenhouse climate control, Wageningen Pers Publishers, Wageningen, 1994, 384pp.

Bot, G.P.A. (1991). Physical modelling of greenhouse climate. Proceedings of the IFAC/ISHS Workshop, pp: 7-12.

Boulard, T. and A. Baille, (1993). A simple greenhouse climate control model incorporating effects on ventilation and evaporative cooling. Agricultural and Forest Meteorology, 65, pp:145-157.

Hornik K., Stinchcombe M. and White H. 1989. Multi-Layer Feedforward Networks are Universal Aproximators, Neural Networks 2, 359-366.

J. Boaventura Cunha., C. Couto and A.E.B. Ruano, (1997). Real-time parameter estimation of dynamic temperature models for greenhouse environmental control. Control Eng. Practice, Vol. 5, N.10, pp. 1473-1481.

J. Boaventura Cunha, Carlos Couto, A.E.B. Ruano (2000). A greenhouse Climate Multivariable Predictive Controller, Acta Horticulturae N. 534, ISHS, pp:269-276.

J. P. Coelho, J. Boaventura Cunha, P. B. de Moura Oliveira (2002). Solar radiation Prediction methods applied to improve greenhouse climate control, World Congress of Computers in Agriculture and Natural Resources, 13-15 March, 2002, pp:154-161.

Ljung, L. (1987). System identification - theory for the user, PTR Prentice-Hall, New Jersey.

Marquardt, D. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM J. Appl. Math. 11, pp. 164-168.

Salgado, Mario E., Graham C. Goodwin and Richard H. Middleton, (1988). Modified least squares algorithm incorporating exponential resetting and forgetting. Int. J. Control, 21, 477-491