View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Ghent University Academic Bibliography

Biosystems Engineering (2003) **84** (2), 147–157
doi:10.1016/S1537-5110(02)00239-8
IT—Information Technology and the Human Interface
Available online at www.sciencedirect.com
SCIENCE DIRECT®

Modelling Greenhouse Temperature by means of Auto Regressive Models

H. Uchida Frausto[1]; J.G. Pieters[1]; J.M. Deltour[2]

[1] Ghent University, Biosystems Engineering, Coupure links 653, Ghent 9000, Belgium; e-mail of corresponding author: Jan.Pieters@rug.ac.be
[2] Faculté universitaire des Sciences agronomiques de Gembloux, Unité de Physique des Biosystèmes, Avenue de la Faculté 8, 5030 Gembloux, Belgium

(Received 16 January 2002; accepted in revised form 3 October 2002)

In this study, it was investigated to what extent linear auto regressive models with external input (ARX) and auto regressive moving average models with external input (ARMAX) could be used to describe the inside air temperature of an unheated, naturally ventilated greenhouse under Western European conditions. Outside air temperature and relative humidity, global solar radiation, and cloudiness of the sky were used as the input variables. Firstly, different models were built for the first and middle week of each season. The models were suitable to describe the greenhouse temperature evolution satisfactorily, except for the ventilation periods, apparently due to the non-linear effect of ventilation strategies. It was also observed that ARX models performed better than ARMAX models. None of the input variables could be omitted from models for a complete year. It was found that the application of a single model structure for a complete year required frequent retuning. Retuning when the goodness of fit falls below a pre-set threshold, proved to be more efficient than retuning at fixed time intervals in maintaining high accuracy.
© 2003 Silsoe Research Institute. All rights reserved
Published by Elsevier Science Ltd

1. Introduction

The greenhouse climate is a very complex system in which the variables highly depend on the outside climate conditions and on the greenhouse design, while most of them are inter-dependent through heat and mass transfer phenomena. Furthermore, the canopy is not a passive element in the greenhouse microclimate system. Via several control mechanisms, plants can modify to some extent the greenhouse climate by changing their heat, vapour and carbon dioxide exchange rates. In actively controlled greenhouses, the climate also depends on manually or automatically installed control strategies. It is clear that the scientific understanding of the greenhouse climate mechanism has advanced a lot by the availability of computers that allow to simulate the climate by means of static and dynamic greenhouse climate models. Whereas the pioneering works (for instance, the model of Takakura *et al.*, 1971) were rather simple by modern standards, the actual deterministic models allow the greenhouse climate dynamics to be studied in a detailed way.

Consequently, the calibration and validation of such models requires many parameters to be determined experimentally. For most climate control purposes, the application of such models is rather difficult and time-consuming, and success is dependent on the accuracy of some parameters (Udink ten Cate, 1987). Consequently, in recent decades several black and grey box models have been developed (Challa, 1981; Tantau, 1985). Black box models merely describe statistically what happens with a given input over a limited range (Hanan, 1998). This implies that they are fully based on empiricism. Grey box models have a structure that is based partly on physical, chemical or biological laws (like deterministic models) and partly on empiricism. Both black and grey box models often show relatively accurate high-frequency responses, which make them especially suitable for control purposes (Udink ten Cate, 1987). Some of these models have the advantage that they do not require explicit evaluation of transfer coefficients or model formulation, which is, for instance the case with neural networks. Using this technique, Seginer *et al.* (1994) were able to model the greenhouse climate satisfactorily. Auto regressive methods allow models to be built using experimentally obtained input–output relations from the system (*i.e.* input and output

 © 2003 Silsoe Research Institute. All rights reserved
Published by Elsevier Science Ltd

Notation

a	direction coefficient obtained from linear regression
a_{ij}	model parameter (with i and j any natural number)
A	matrix containing the model parameters a_{ij}
b	offset coefficient obtained from linear regression
b_{ij}	model parameter (with i and j any natural number)
B	matrix containing the model parameters b_{ij}
c_I	model parameter (with i any natural number)
C	vector containing the model parameters c_i
e	error
G	goodness of fit, %
h_e	outside air relative humidity, %
n	number of samples
n_a	number of poles
n_b	number of zeros increased by one
n_c	order of the measured error
n_k	time delay in the discrete time domain
p_{cl}	cloudiness of the sky
q_s	global solar ration flux, W m^{-2}
r^2	coefficient of determination of linear regression
s	square mean error
T_e	outside air temperature, °C
T_i	inside air temperature, °C
t	discrete time
u	input signal
y	output signal
z	z-transform operator

Greek letters

δ	Kronecker symbol
σ	standard deviation

Subscripts

i,j	any natural number
m	simulated output data
o	original output data

values obtained from 'field' experiments) using techniques of parameter identification. Applications of this technique in horticultural engineering include the greenhouse climate model of Boaventura *et al.* (1996) for Mediterranean climate conditions and the plant response model of Boonen *et al.* (2000).

The objective of this study was to investigate the outside climate variables which must at least be included in linear auto regressive models simulating the inside air temperature of greenhouses under Western European conditions and to what extent the most appropriate model structure and retuning frequency vary throughout the year.

2. Theoretical considerations

Although climate characteristics are continuous variables, they are measured and registered at time steps which give measured climate data, a discrete character. In this discrete domain, the dynamic greenhouse system can be modelled in several ways, one of which is by means of linear auto regressive relations between the discrete output $y(t)$ and the discrete input $u(t)$, such as auto regressive models with external input (ARX) and auto regressive moving average models with external input (ARMAX). Assuming a single input–single output system, the following expressions can be used to describe this relationship (Ljung, 1999):

$$y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) \quad (1)$$

where: $y(t)$ is the output signal at (discrete) time t; $u(t-n_k)$ is the input signal at (discrete) time t-n_k; $e(t)$ is the perturbation or even any not measurable input in the system (noise); a_{n_a}, b_{n_b}, c_{n_c} are the model parameters; n_a, n_b, n_c indicate the order of the respective polynomials (output, input, perturbation) and n_k is time delay from input to output.

For ARMAX, Eqn (1) is often represented as

$$A(z)y(t) = B(z)u(t-n_k) + C(z)e(t) \quad (2)$$

and for ARX as

$$A(z)y(t) = B(z)u(t-n_k) + e(t) \quad (3)$$

where the matrices $A(z)$ and $B(z)$ and the vector $C(z)$ are given by

$$A(z): 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a} \quad (4)$$

$$B(z): b_1 z^{-1} + \cdots + b_{n_b} z^{-n_{b+1}} \quad (5)$$

$$C(z): 1 + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c} \quad (6)$$

and z^{-1} is the backward shift operator:

$$z^{-1} u(t) = u(t-1) \quad (7)$$

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

$$a_{ij}(z) = \delta_{ij} + a_{1_{ij}} z^{-1} + \cdots + a_{n_{aij}} z^{-n_{a_{ij}}} \quad (8)$$

and

$$b_{ij}(z) = b_{1_{ij}} z^{-n_{k_{ij}}} + \cdots + b_{n_{b_{ij}}} z^{-n_{k_{ij}} - n_{b_{ij}} + 1} \tag{9}$$

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 time delay n_k. The definition of the ARMAX structure additionally requires the order of the measured error n_c to be known. The matrices $A(z)$, $B(z)$ and $C(z)$ are determined by means of off-line parameter identification methods.

3. Procedures

3.1. *Data sets*

Instead of using measured values for the outside climate data and the inside air temperature, data sets were composed using outside climate data of the Belgian typical reference year (Dogniaux *et al.*, 1978) and simulated values for the inside air temperature. The outside climate variables included in the data set were the outside air temperature T_e in °C, the outside air relative humidity h_e in %, the global solar radiation flux density q_s in W m^{-2}, and the dimensionless cloudiness of the sky p_{cl}. The simulated values for the inside air temperature T_i in °C were obtained by means of the Gembloux Dynamic Greenhouse Climate Model (GDGCM), to which the data of the typical reference year were fed. It can be discussed whether the use of simulated data instead of measured data is appropriate to study the behaviour of ARX and ARMAX models. It can be expected that ARX and ARMAX models obtained in this way will fit somewhat better to the values simulated by the GDGCM than to values that would be measured in a real greenhouse. However, use of simulated data obtained by means of an accurate deterministic model (see below for the accuracy of the GDGCM) allows a better understanding of the ARX and ARMAX model behaviour, since the data used to build and validate the ARX and ARMAX models can be studied in a detailed way by means of the deterministic model. Since in this study, the behaviour of ARX and ARMAX models was studied, rather than their absolute accuracy, simulated data were adopted.

The GDGCM is a semi-one-dimensional greenhouse climate model, describing the energy and mass exchanges between seven internal layers (four soil layers, one vegetation layer, one inside air layer, one cover), which form the system, and three external layers (subsoil, outside air, and sky) which constitute, together with the solar radiation, the boundary conditions. For each of the layers, heat loss or gain by solar radiation, far-infrared radiation, conduction, convection, and latent heat is described mathematically. Furthermore, a mass transfer equation for vapour is considered. The model also allows to simulate the effect of control procedures, like ventilation, heating, *etc*. Simulated values of the inside air temperature are typically closer than 0·5°C to measured values. More details on the model and its validation can be found in de Halleux *et al.* (1985), Nijskens *et al.* (1991), Pirard *et al.* (1994) and Pieters and Deltour (1997).

An unheated glasshouse equipped with a proportionally controlled natural ventilation system was assumed. When the inside air temperature was below the ventilation set-point temperature, an air renewal rate of 0·2 h^{-1} was adopted. When the inside air temperature was 3°C or more above the ventilation set-point temperature, the air renewal rate was set at 40·2 h^{-1}. When the inside air temperature was between 0 and 3°C above the ventilation set-point temperature, the ventilation rate was determined by linear interpolation between 0·2 and 40·2 h^{-1}. The ventilation set-point temperatures for day-time and night-time—determined by an astronomical clock—were fixed at 20 and 18°C, respectively. A linear light-dependent increase of the day-time ventilation set-point temperature was provided from a global solar energy flux density of 200 W m^{-2} on and with a maximum of 2°C (*i.e.* the maximum ventilation set-point temperature was 22°C) for a solar energy flux density of 500 W m^{-2} or higher. Two transition periods between night- and day-time, in which intermediate set-point temperatures were imposed, were introduced to avoid abrupt temperature changes in the greenhouse. Greenhouse construction and cladding parameters were obtained from Pirard *et al.* (1993) and Pollet and Pieters (2000) and are summarised in Table 1. Simulations were carried out for every minute of a complete year, while data were output for every 5 min.

3.2. *Structure definition and selection*

All four input variables of the outside climate (namely air temperature, relative humidity, global solar radiation and cloudiness of the sky) used in the GGDCM were used in the ARMAX and ARX structures. From the validation results described in Pirard *et al.* (1993) and from other literature (von Zabeltitz, 1986; Bakker *et al.*, 1995), it is well documented that inclusion of these four variables in greenhouse climate models enables accurate results for the inside air temperature to be obtained.

Table 1
Values for the Gembloux Dynamic Greenhouse Climate Model parameters (after Pirard *et al.*, 1993; Pieters & Pollet, 2000)

Soil characteristics							
Thermal conductivity (four layers), $W\,m^{-1}\,K^{-1}$		700	1950	1900	1900		
Layer thickness (four layers), m		0·05	0·15	0·30	0·70		
Specific mass (four layers), $kg\,m^{-3}$		1300	1450	1600	1650		
Heat capacity (four layer), $J\,kg^{-1}\,K^{-1}$		3350	1250	1250	1200		
Reflectance for solar radiation, dimensionless				0·85			
Emittance for far-infrared radiation, dimensionless				0·40			
Subsoil layer thickness, m				8·8			
Characteristic length of the floor, m				1001			
Construction characteristics							
Latitude , °N				50·78			
Length of the greenhouse, m				48·00			
Width of the greenhouse, m				44·80			
Number of spans, dimensionless				14			
Height to the eaves, m				3·10			
Height to the ridge, m				3·75			
Emittance of the cover surface, dimensionless				0·90			
Cover transmittance for far-infrared radiation, dimensionless				0·0			
Dry cover transmittance for beam radiation (angles of incidence of 0, 15, 30, 45, 60, 75 and 90°), dimensionless	0·846	0·846	0·841	0·822	0·773	0·552	0·00
Wet cover transmittance for beam radiation (angles of incidence of 0, 15, 30, 45, 60, 75 and 90°), dimensionless	0·842	0·840	0·804	0·735	0·635	0·475	0·00
Dry cover reflectance for beam radiation (angles of incidence of 0, 15, 30, 45, 60, 75 and 90°), dimensionless	0·074	0·074	0·089	0·138	0·187	0·328	1·00
Wet cover reflectance for beam radiation (angles of incidence of 0, 15, 30, 45, 60, 75 and 90°), dimensionless	0·078	0·080	0·126	0·225	0·325	0·405	1·00
Cover absorptance for diffuse solar radiation, dimensionless				0·065			
Dry cover transmittance for diffuse solar radiation, dimensionless				0·710			
Wet cover transmittance for diffuse solar radiation, dimensionless				0·652			
Frame transmittance for solar radiation, dimensionless				0·86			
Cladding heat capacity per unit surface area, $J\,kg^{-1}\,m^{-2}$				8000			
Maximum equivalent water film thickness, mm				0·040			
Vegetation characteristics							
Reflectance for solar radiation, dimensionless				0·22			
Canopy attenuation coefficient, dimensionless				0·6093			
Characteristic length of the leaves, m				0·20			
Emittance for far infrared radiation, dimensionless				0·95			
Heat capacity, $J\,kg^{-1}\,K^{-1}$				4180			
Air characteristics							
Specific mass, $kg\,m^{-3}$				1·25			
Specific heat , $J\,kg^{-1}\,m^{-3}$				1256			
Latent heat of condensation for water vapour, $kJ\,kg^{-1}$				2437			
Air speed near the cover, $m\,s^{-1}$				0·30			
Lewis number, dimensionless				0·89			

For the selection of the most appropriate structures of the ARX and ARMAX models, the maximum number of poles n_a and the maximum time delay n_k were set at 4. The maximum number of zeros was set at 5 (i.e. n_b was set at 4), while the order of the measured error n_c was limited to 1. Since the sampling period was 5 min, this implies that data being gathered 20 min before could be included in the model. From the literature (Bakker *et al.*, 1995), this seems to be a reasonable time delay to take into account thermal inertia effects on the inside air temperature.

Taking into account the possible combinations of coefficients and variables to be included, this implies that for each set of data 4096 structures had to be elaborated and tested for their output accuracy for the ARMAX and the ARX models. From these 4096 model structures, the model showing the best goodness of fit (Ljung, 2000) was selected. The goodness of fit is defined as

$$G = \left(1 - \frac{\sqrt{\sum_{i=1}^{n} (y_{m,i} - y_{o,i})^2}}{\sqrt{\sum_{i=1}^{n} \left(y_{o,i} - \frac{1}{n}\sum_{k=1}^{n} y_{o,k}\right)^2}} \right) \times 100 = \left(1 - \frac{s}{\sigma}\right) \times 100 \qquad (10)$$

where: y_m is the output of the ARMAX or ARX model; y_o represents the original output data (in this case, the temperature simulated by means of the GGDCM); n is the number of samples; s is the square mean error of modelled versus original output and σ is the standard deviation of the original system output.

Apart from the goodness of fit, the suitability of the different model structures was also assessed by means of the coefficient of determination obtained through linear regression of the modelled results on the original data for the inside air temperature. The regression coefficients allowed to further assess the agreement between modelled and original temperatures.

The parameter identification for the ARMAX and ARX structures was carried out using some special features of the commercially available software package MATLAB (Ljung, 2000). This procedure was repeated for the first and middle week of each season. Validation of the models was carried out using the data set of the following week. It was then investigated to what extent the model structures and parameters (coefficients) differed from each other for the several periods, *i.e.* to what extent the model structure or parameter set depended on the specific data set used for its construction.

3.3. *Analysis of hierarchy of variables*

For control purposes, the quality of the adjustment is only one of the selection criteria to be used. Simplicity of the model is also to be considered. Therefore, it was tried to simplify each of the selected structures by leaving out one of the variables and the subsequent retuning of the model. Since the four variables were omitted one by one, this implies that four new models were obtained. The goodness of fit and the regression parameters were subsequently compared with the performance results of the original model structure. In this way, the impact of each input variable over the output variable could be determined and a hierarchy of the input variables could be established for each season and middle season.

3.4. *Construction of general model structure*

Because of the simplicity requirement, it is not convenient to use different model structures for each season and each middle season. Therefore, it was tried to define a model structure that gives sufficient accuracy throughout a complete year. To this end, each of the previously obtained structures was tuned for all periods and tested for its goodness of fit and regression parameters. Furthermore, a general model structure was built by selecting all or most of the components of the separate seasonal model structures. This means that terms for which the corresponding coefficient (parameter) was zero in all or most of the seasonal models were left out of the general model structure. The general model was subsequently compared with the other model structures with respect to its performance characteristics.

As a second step in this optimisation process, the retuning frequency, needed to maintain a predefined goodness of fit or coefficient of determination for the regression of simulated on original inside air temperatures, was investigated. Two strategies were tested. The first strategy consisted in retuning the general model at fixed time intervals. Each time, the model was retuned based on the results of the preceding time interval and used for the prediction of the temperature during the following time interval. This means that the tuning period was always as long as the prediction period. On the one hand, longer tuning periods tend to produce more accurate regression results for the parameter estimation. On the other hand, longer prediction periods tend to lead to worse estimates, because of the shift from the tuning period, which leads to the use of a model under conditions that are different from the conditions for which it was tuned. From this, it is clear that some optimal retuning interval can be found. Therefore, three different intervals were tested, namely 7, 14 and 30 days. The second strategy was based on a tolerance criterion. Retuning was carried out only after it was observed that the performance of the model had fallen below a predefined threshold value.

4. Results and discussion

4.1. *Selection of seasonal models*

Table 2 gives the selected models for the first and middle week of each season. It can be observed that all selected models included all four outside climate variables. Except for the ARMAX model for the middle spring week and the ARX model for the middle summer week, it was also found that in most cases, inclusion of values of the outside climate data which were older than 15 min did not improve the model performance. In most cases, outside climate data older than 10 min were not considered. Consequently, the resulting models were

Table 2

Model structures and typical parameter identification results for the seasonal models giving the inside air temperature (T_i) as a function of outside air temperature (T_e), outside air relative humidity (h_e), global solar radiation flux (q_s), cloudiness of the sky (p_{cl}) and perturbation (e) in the discrete time (t) domain, using the z-transform operator (z^{-1}); a_1 to a_4, b_{11} to b_{43} and c_1 are regressive coefficients as defined in the model structure

$$T_i(t) = \frac{[T_e(t) \;\; h_e(t) \;\; q_s(t) \;\; p_{cl}(t)]}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4}} \begin{bmatrix} b_{11} z^{-1} + b_{12} z^{-2} + b_{13} z^{-3} \\ b_{21} z^{-1} + b_{22} z^{-2} + b_{23} z^{-3} + b_{24} z^{-4} \\ b_{31} z^{-1} + b_{32} z^{-2} + b_{33} z^{-3} \\ b_{41} z^{-1} + b_{42} z^{-2} + b_{43} z^{-3} \end{bmatrix} + \frac{1 + c_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4}} e(t)$$

	a_1	a_2	a_3	a_4	b_{11}	b_{12}	b_{13}	b_{21}	b_{22}	b_{23}	b_{24}	b_{31}	b_{32}	b_{33}	b_{41}	b_{42}	b_{43}	c_1
ARMAX																		
Winter	−0·952	0	0	0	0	0·035	0	0	$2·35 \times 10^{-4}$	0	0	0	$2·88 \times 10^{-4}$	0	0·107	0	0	0·669
Midwinter	−0·947	0	0	0	0	0·037	0	0	$-8·74 \times 10^{-5}$	0	0	$3·21 \times 10^{-4}$	0	0	0·168	0	0	0·745
Spring	−0·931	0	0	0	0	0·045	0	0	0	$6·64 \times 10^{-4}$	0	$5·77 \times 10^{-4}$	0	0	0·167	0	0	0·485
Midspring	−2·086	0·986	0·302	−0·201	0·422	−0·826	0·404	0	−0·031	0·061	−0·030	−0·031	0·061	−0·030	0·309	−0·581	0·272	−1·005
Summer	−0·994	0	0	0	0·611	−0·611	0	−0·061	0·061	0	0	$4·35 \times 10^{-5}$	$-7·89 \times 10^{-6}$	0	0·100	−0·028	0	0·435
Midsummer	−1·875	0·880	0	0	0·003	0	0	0	$2·95 \times 10^{-4}$	0	0	0	$1·75 \times 10^{-5}$	0	0·021	0	0	0
Autumn	−0·958	0	0	0	0	0·021	0	0	0·002	0	0	$2·47 \times 10^{-4}$	0	0	0·120	0	0	0·671
Midautumn	−0·964	0	0	0	0·295	−0·277	0	−0·020	0·020	0	0	$1·17 \times 10^{-4}$	$2·33 \times 10^{-4}$	0	0·098	0·077	0	0·835
ARX																		
Winter	−0·948	0	0	0	0	0·037	0	0	$2·62 \times 10^{-4}$	0	0	$3·02 \times 10^{-4}$	0	0	0·116	0	0	0
Midwinter	−0·927	0	0	0	0	0·051	0	0	$-1·11 \times 10^{-4}$	0	0	$4·45 \times 10^{-4}$	0	0	0·217	0	0	0
Spring	−0·923	0	0	0	0	0·050	0	0	$7·10 \times 10^{-4}$	0	0	$6·23 \times 10^{-4}$	0	0	0·195	0	0	0
Midspring	−0·957	0	0	0	0	0·025	0	0	0·001	0	0	$2·56 \times 10^{-4}$	0	0	0·154	0	0	0
Summer	−0·990	0	0	0	0·746	−0·740	0	−0·083	0·083	0	0	$2·87 \times 10^{-4}$	$-2·67 \times 10^{-4}$	0	0·185	−0·131	0	0
Midsummer	−0·966	0	0	0	0	0·015	0	0	0	0	0·002	$1·70 \times 10^{-4}$	0	0	0·131	0	0	0
Autumn	−0·956	0	0	0	0	0·021	0	0	0·003	0	0	$2·73 \times 10^{-4}$	0	0	0·145	0	0	0
Midautumn	−0·959	0	0	0	0	0·019	0	0	$3·93 \times 10^{-4}$	0	0	$4·03 \times 10^{-4}$	0	0	0	0·197	0	0

ARMAX: auto regressive moving average model with external input; ARX: auto regressive model with external input.

rather compact. This should be considered an advantage since this facilitates their use in climate control applications. These results are in agreement with the typical thermal inertia effects met in greenhouses (Bot, 1989; Bakker *et al.*, 1995; Hanan, 1998). The slightly higher 'memory effect' for the relative humidity with respect to the other variables has to be explained by the fact that vapour exchange in greenhouses is a somewhat slower phenomenon than heat exchange.

With respect to the several coefficients, it is found that in most models similar coefficients do not differ a lot among the several seasonal tunings. This implies that the influence of the several outside climate variables does not vary a lot throughout the year. From Table 2 it can be observed, however, that the influence of the relative humidity, which is usually low, is relatively larger in autumn. This might be a consequence of the high humidity and the higher humidity variations during that period of the year in a maritime climate. The fact that the highest absolute values for the relative humidity were found for summer, might be somewhat misleading. For both the ARX and the ARMAX models, the values for b_{21} and b_{22} obtained for summer were almost exactly the same but opposite in sign. Consequently, since relative humidity changes are usually very small when considered over 5 min intervals, the effect of the relative humidity is small, whereas the effect of changes in relative humidity is much more important.

The influence of solar radiation can be seen to be relatively less important with respect to inside air temperature changes. This is in contrast with the very important role that plays solar radiation in the heat balance of the greenhouse. The rather low coefficients are to be explained by the fact that solar radiation has only an indirect effect on inside air temperature, since it is not directly absorbed by the air. In fact, solar radiation needs first to be absorbed by the plants, the soil and the construction, which subsequently exchange part of the heat gained in that way with the air through convection. This mechanism obviously damps the effect of short-term variations of the solar radiation level on the inside air temperature.

In order to assess the accuracy of the several ARMAX and ARX models, the simulated results were compared with the original data for the inside air temperature. *Figure 1* (a and b) shows an example for both winter period models. The difference between the simulated and the original inside air temperatures is also plotted in these figures. From this figure, it can be observed that the average deviation of the simulated results was negligible, while the standard deviation was 0.5°C for both the ARMAX and ARX models. The extreme deviations were situated between −1·9 and 1·7°C and were linked with abrupt changes in the temperature evolution. Furthermore, the deviations between simulated and original values showed a similar course for both models, so the goodness of fit of both models was almost the same, namely 83%.

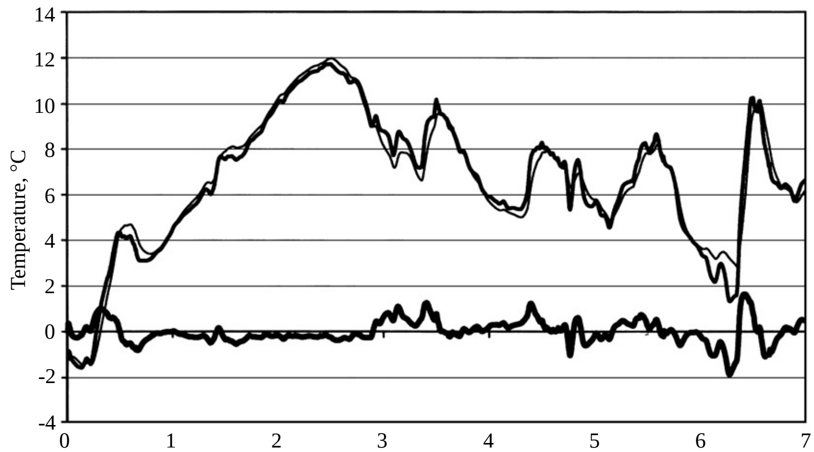
The analysis results of the linear regression of the simulated results on the original data confirmed these good performances, as can be deduced from Table 3. Table 3 gives the linear regression results for all 16 models. It can be observed that both winter models had a coefficient of direction close to 1 (namely 0·98) and an offset close to 0 (namely about 0·15), while the coefficient of determination was higher than 0·97.

When comparing the regression analysis results for the other seasonal models, it can be observed that most models gave good results, although the performances from middle spring to middle summer were clearly worse than those from autumn until early spring. Comparing the results from Tables 2 and 3 shows that the most complicated model structures had the lowest performances. The poor performance of these models clearly show the inability of these model structures to take into account the (non-linear) effect of the ventilation control strategy, although this strategy fully depends on the greenhouse and the outside climate behaviour. Inclusion of the ventilation strategy or its effect in the ARMAX models could thus result in better performances. Finally, the results of Table 3 clearly illustrate that the ARX models performed mostly better than the ARMAX models.

4.2. *Hierarchy of variables*

The performances of the model structures with three input variables were compared with those with four input variables. Depending on the period for which the model was valid, the omission of one of the outside climate inputs lowered the coefficient of determination by −0·0005 to 0·0963 for the ARMAX models and by 0·0004 to 0·0388 for the ARX models. In most cases, it was found that omission of the outside air temperature had the greatest impact on the coefficient of determination, followed by the global solar radiation and the cloudiness of the sky, while the outside relative humidity had the lowest impact.

Thus, the effect of omitting the relative humidity from the models was examined. Table 4 shows the analysis results of the linear regressions of the modelled results on the original data. From these results and the results for the models with four inputs, it is clear that the omission of the relative humidity from the



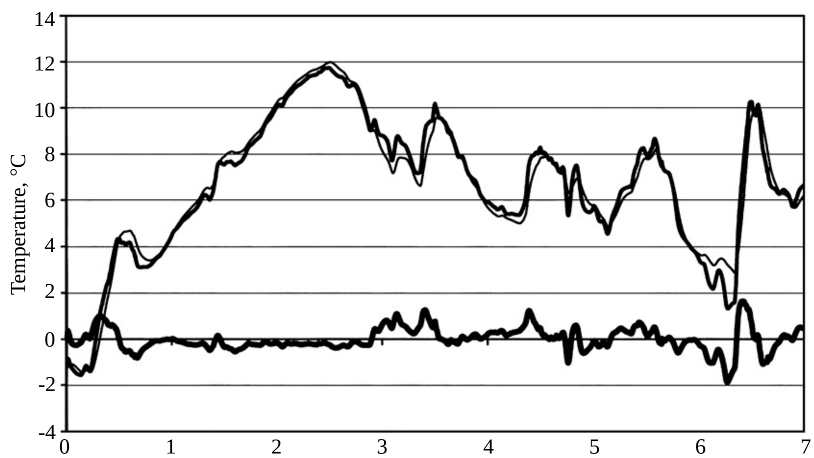


Fig. 1. Original data versus results of the (a) auto regressive moving average model with external input and the (b) auto regressive model with external input and the difference between both for the first week of winter; ▬▬, original; ——, model; ▬▬, residuals

Table 3
Direction coefficient a, offset b and determination coefficient r^2 resulting from the linear regression of the ARMAX and ARX results on the original data; models with four inputs: outside air temperature, outside air relative humidity, global solar radiation flux and cloudiness of the sky

	ARMAX			*ARX*		
Season	*a*	*b*	r^2	*a*	*b*	r^2
Winter	0·977	0·145	0·971	0·978	0·144	0·971
Middle winter	0·945	0·320	0·957	0·961	0·240	0·965
Spring	0·989	0·075	0·988	0·991	0·063	0·989
Middle spring	1·077	0·345	0·871	0·898	1·449	0·918
Summer	0·819	3·194	0·880	0·885	1·950	0·928
Middle summer	0·558	9·176	0·576	0·786	3·692	0·833
Autumn	0·893	1·567	0·918	0·915	1·350	0·922
Middle autumn	0·980	0·188	0·967	0·976	0·232	0·965

ARMAX: auto regressive moving average model with external input; ARX: auto regressive model with external input.

model leads to a negligible loss of accuracy for most models, *i.e.* for most periods. The loss of accuracy for the ARMAX models for summer, mid-summer and autumn, however, clearly shows that the relative humidity could not be omitted from all structures. The ARX models show a similar behaviour, but here

Table 4

Direction coefficient a, offset b, and determination coefficient r^2 resulting from the linear regression of the results from the models with three inputs (outside air temperature, global solar radiation flux and cloudiness of the sky) on the original data and the difference Δr^2 between the determination coefficients for the four inputs models and the three inputs models

	ARMAX				*ARX*			
Season	*a*	*b*	r^2	Δr^2	*a*	*b*	r^2	Δr^2
Winter	1·001	−0·064	0·967	0·004	1·004	−0·068	0·968	0·003
Middle winter	0·914	0·549	0·956	0·001	0·937	0·423	0·965	0·001
Spring	1·042	−0·487	0·981	0·007	1·051	−0·539	0·986	0·003
Middle spring	1·051	0·748	0·853	0·018	0·949	0·597	0·879	0·039
Summer	0·688	5·044	0·797	0·086	0·845	2·241	0·900	0·027
Middle summer	0·452	11·57	0·293	0·283	0·786	3·378	0·660	0·173
Autumn	0·878	1·419	0·729	0·189	0·898	1·104	0·738	0·184
Middle autumn	0·993	0·051	0·968	−0·001	0·993	0·049	0·966	−0·001

ARMAX: auto regressive moving average model with external input; ARX: auto regressive model with external input.

again their performance was much better than for the ARMAX models. As a result, it was concluded that a general model for a complete year must include all four input variables.

4.3. *General model structure*

4.3.1. *Selection and performance*

Despite the fact that the most suitable model structures did not differ a lot throughout the year (see Table 2), it was found that none of the seasonal models could be used as such to model the greenhouse temperature for a whole year. It was found that for almost all models, the goodness of fit fell below 65% about 2 months after the tuning date (data not shown).

Using the common characteristics of the model structures for the first and middle week of each season, the following structure was selected and tested for its usefulness during a complete year.

$$T_i(t) = \frac{\left[\,T_e(t) \quad h_e(t) \quad qs(t) \quad p_{cl}(t)\,\right]}{1 - a_1 z^{-1}} \begin{bmatrix} b_{11}z^{-1} + b_{12}z^{-2} \\ b_{21}z^{-1} + b_{22}z^{-2} + b_{23}z^{-3} \\ b_3 z^{-1} \\ b_4 z^{-1} \end{bmatrix} + \frac{1 + c_1 z^{-1}}{1 - a_1^{-1}} e(t) \quad (11)$$

For the ARX model the coefficient c_1 was zero. Since it was already found that the relative humidity could not be omitted for all seasons, it was considered a necessary input variable in the general model.

Table 5 gives the performance characteristics of the general model that was tuned separately for the first and middle week of each season. Although the coefficients of determination were in the range from 0·845 to 0·985 for the ARMAX models and from 0·903 to 0·987 for the ARX models, the coefficient of direction and the offset were not close to 1 and 0, respectively, at least for the

Table 5

Direction coefficient *a*, offset *b*, and determination coefficient r^2 resulting from the linear regression of the general structure model results on the original data and goodness of fit *G* with retuning after the first and midweek of each season

	ARMAX				*ARX*			
Season	*a*	*b*	r^2	*G, %*	*a*	*b*	r^2	*G, %*
Winter	0·968	0·196	0·967	82·3	0·979	0·150	0·973	83·8
Middle Winter	0·954	0·267	0·956	80·3	0·961	0·231	0·960	81·0
Spring	0·993	0·050	0·985	89·1	0·990	0·072	0·987	89·4
Middle Spring	1·032	−0·099	0·907	67·0	0·896	1·459	0·903	69.6
Summer	0·817	3·255	0·872	64·7	0·881	2·036	0·924	72·8
Middle Summer	0·809	4·243	0·845	55·8	0·890	1·868	0·938	75·0
Autumn	0·789	2·526	0·944	66·9	0·877	1·939	0·949	76·9
Middle Autumn	0·973	0·236	0·964	81·7	0·976	0·225	0·964	81·8
Average	0·917	1·333	0·930	73·5	0·931	0·998	0·950	78·8

ARMAX: auto regressive moving average model with external input; ARX: auto regressive model with external input.

period from summer to autumn. Therefore, it was concluded that the introduction of a single model structure which was tuned only once for a complete year was not a suitable option.

4.3.2. *Tuning frequency*

Figure 2 gives the evolution of the goodness of fit and the coefficient of determination resulting from the regression of modelled on original results throughout a complete year for two retuning frequencies, namely once per 7 and once per 30 days. From this figure, it is clear that a fixed retuning frequency resulted in highly varying performances of the general structure model, with the 30 days retuning interval giving rise to somewhat better results. As was to be expected, the performance was again best in winter and worst in summer. This implies that the introduction of a pre-defined minimum level of accuracy is more meaningful from a climate control point of view. *Figure 3* gives the resulting goodness of fit and the moments of tuning for the general structure ARX model for the whole spring season, for which the pre-defined minimum goodness of fit was set at 70%. It can be seen that 26 retunings were needed, 25 of which were situated in the second halve of the period under consideration. For several days, the goodness of fit was below the pre-defined minimum value. This is to be

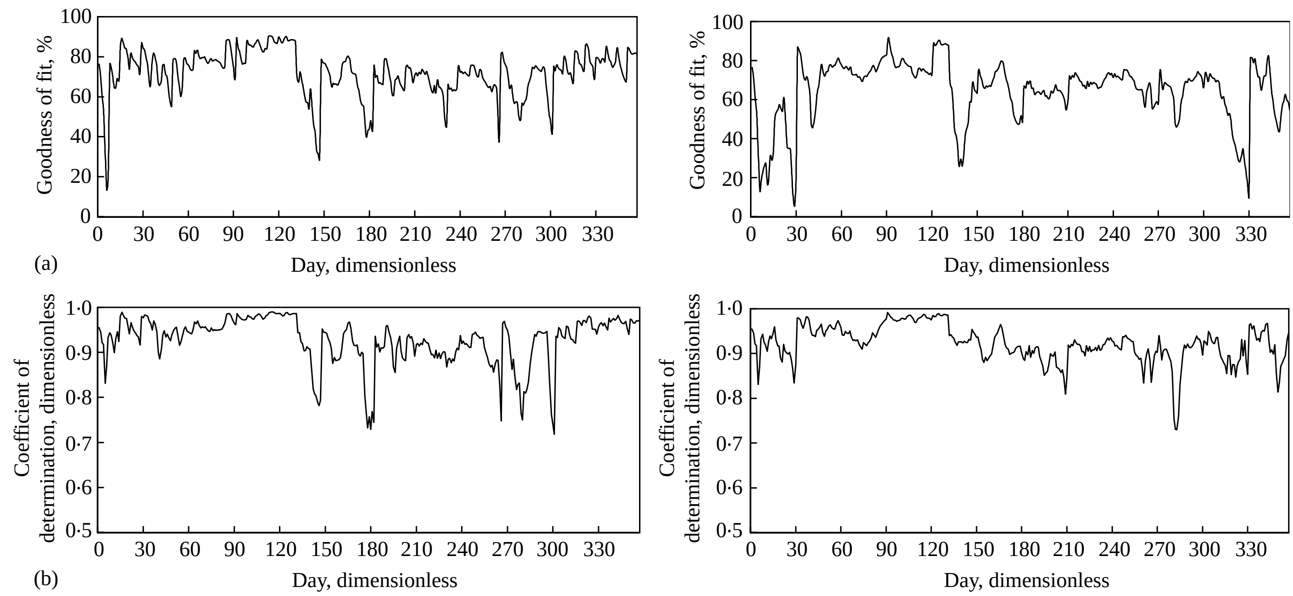


Fig. 2. Goodness of fit and coefficient of determination resulting from the regression of the results of the auto regressive model with external input on the original data obtained for a retuning period of (a) 7 and (b) 30 days over a complete year, starting on 1 December

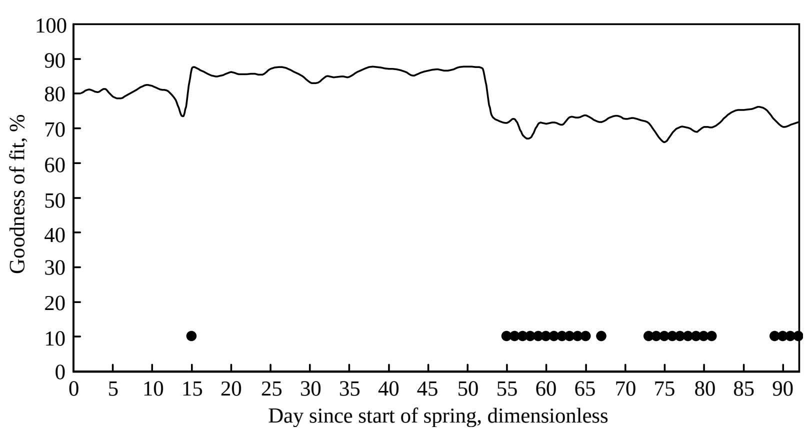


Fig. 3. Goodness of fit of the general structure auto regressive model with external input for the complete spring season with retuning each time the goodness of fit falls below 70%; ●, moment of retuning

explained by the fact that in these periods, even a daily retuning did not allow the pre-defined value to be obtained. When considered over a complete year, it was found that this situation is mainly restricted to the period from middle spring to autumn, in which ventilation plays a major role.

5. Conclusions

In this study, it was investigated to what extent linear auto regressive models with external input (ARX) and linear auto regressive moving average models with external input (ARMAX) could be used to describe the inside air temperature of an unheated, naturally ventilated greenhouse under Western European conditions. Four input variables were considered, namely outside air temperature, outside air relative humidity, global solar radiation and cloudiness of the sky. Firstly, separate models were built for the first and middle week of each season. It was found that these linear regression models were suitable to describe the greenhouse behaviour during most of the year, except for the ventilation periods, due to the fact that the behaviour of the greenhouse becomes highly non-linear when control strategies are imposed. Inclusion of the ventilation effect in the models should thus be envisaged. It was also observed that ARX models performed better than ARMAX models. Although the inclusion of the outside air humidity in the several seasonal models did not contribute significantly to the accuracy of the results during most of the year, it was observed that none of the input variables could be omitted from models, which have to be applied over a complete year. It was found that the application of a single-model structure for a complete year required frequent retuning during the year in order the model to remain sufficiently accurate. For this purpose, retuning when the goodness of fit falls below a pre-set threshold proved to be more efficient than retuning at fixed time intervals.

Acknowledgements

Hugo Uchida Frausto thanks CONACyT for its financial support (contract 124965) as well as Dr Gilberto Herrera Ruiz (Universidad Autónoma de Querétaro, Mexico) for his collaboration in this project.

References

Bakker J C; Bot G P A; Challa H; Van de Braak N J (1995). Greenhouse climate control. An integrated approach. Wageningen Pers, Wageningen

Boaventura Cunha J; Ruano A E B; Couto C (1996). Identification of greenhouse climate dynamic models. Proceedings of the Sixth International Conference on Computers in Agriculture, Cancun, Mexico, pp. 161–171

Boonen C; Joniaux O; Janssens K; Berckmans D; Lemeur R; Kharoubi A; Pien H (2000). Modelling dynamic behavior of leaf temperature at three-dimensional positions to step variations in air temperature and light. Transactions of the ASAE, **43**(6), 1755–1766

Bot G P A (1989). Greenhouse simulation models. Acta Horticulturae, **245,** 315–325

Challa H (1981). Some remarks concerning the use of models in greenhouse climate research. Acta Horticulture, **107,** 117–120

de Halleux D; Deltour J M; Nijskens J; Nisen A; Coutisse S (1985). Dynamic simulation of heat fluxes and temperatures in horticultural and low emissivity glass-covered greenhouses. Acta Horticulturae, **170,** 91–96

Dogniaux R; Lemoine M; Sneyers R (1978). Année-type moyenne pour le traitement de problèmes de captation d'énergie solaire. [Typical reference year for treatment of solar energy captation problems.] Misc. sér. B, no. 45. Institut Royal Météorologique de Belgique, Bruxelles

Hanan J J (1998) Greenhouses. Advanced Technology for Protected Horticulture. CRC Press LLC, Boca Raton

Ljung L (1999). System Identification. Theory for the User. Prentice-Hall, Upper Saddle River

Ljung L (2000). System Identification Toolbox User's Guide. The MathWorks, Inc., Natick

Nijskens J; Deltour J M; Coutisse S; Nisen A (1991). Sensitivity study of greenhouse climate dynamic model. Bulletin des Recherches Agronomiques de Gembloux, **26**(3), 389–410

Pieters J G; Deltour J M (1997). Performances of greenhouses with the presence of condensation on cladding materials. Journal of Agricultural Engineering Research, **68**(2), 125–137, doi: 10.1006/jaer.1997.0187

Pirard G; Vancayemberg F; Deltour J (1993). Rapport d'activités de décembre 1993. [Research activities report of December 1993.] Centre d'Etude de la Régulation Climatique des Serres (I.R.S.I.A), Gembloux

Pirard G; Vancayemberg F; Deltour J (1994). Rapport d'activités de décembre 1994. [Research activities report of December 1994.] Centre d'Etude de la Régulation Climatique des Serres (I.R.S.I.A), Gembloux

Pollet I V; Pieters J G (2000). Condensation and radiation transmittance of greenhouse cladding materials, part 3: results for glass plates and plastic films. Journal of Agricultural Engineering Research, **77**(4), 419–428, doi: 10.1006/jaer.2000.0628

Seginer I; Boulard T; Bailey B J (1994). Neural network models of the greenhouse climate. Journal of Agricultural Engineering Research, **59**(3), 203–216, doi: 10.1006/jaer.1994.1078

Takakura T; Jordan K A; Boyd L L (1971). Dynamic simulation of plant growth and environment in the greenhouse. Transactions of the ASAE, **14**(5), 964–971

Tantau H-J (1985). Analysis and synthesis of climate control algorithms. Acta Horticulturae, **174,** 375–380

Udink ten Cate A J (1987). Analysis and synthesis of greenhouse climate controllers. In: Computer Applications in Agricultural Environments (Clark J A; Gregson K; Saffell R A, eds), pp 1–19. Butterworths, London

von Zabeltitz C (1986). Gewächshaüser: Planung und Bau. [Greenhouses: planning and construction], Verlag Eugen Ulmer. Stuttgart