

MODELOS ARX PARA PREDECIR LA TEMPERATURA DEL AIRE DE UN INVERNADERO: UNA METODOLOGÍA

ARX MODELS FOR PREDICTING GREENHOUSE AIR TEMPERATURE: A METHODOLOGY

Irineo L. López-Cruz¹, Abraham Rojano-Aguilar¹, Waldo Ojeda-Bustamante² y Raquel Salazar-Moreno¹

¹Ingeniería Agrícola y Uso Integral del Agua. Universidad Autónoma Chapingo. 56230. Chapingo, Estado de México. (ilopez@correo.chapingo.mx). ²Instituto Mexicano de Tecnología del Agua. Jiutepec, Morelos.

RESUMEN

Se presenta un procedimiento para obtener un modelo dinámico lineal de autoregresión con variables exógenas (ARX) para predecir el comportamiento de la temperatura del aire en el interior de un invernadero. Los ARX son modelos matemáticos dinámicos derivados de la teoría de Identificación de Sistemas. Las variables de entrada del modelo fueron temperatura del aire, radiación solar, velocidad del viento afectando el área de ventilación del invernadero y humedad relativa, cuantificadas en una estación meteorológica ubicada a 700 m de un invernadero en Chapingo, Estado de México. La variable de respuesta fue la temperatura del aire dentro del invernadero. Se tomaron muestras de las variables de entrada y de salida del modelo cada 5 min durante un ciclo de cultivo. Para determinar la estructura del mejor modelo, se evaluaron hasta 100 000 modelos ARX usando el criterio de información y el error de predicción final de Akaike. La bondad del ajuste entre las temperaturas simuladas y observadas, y el análisis residual, indicó que modelos ARX de segundo orden o superior predicen en forma adecuada el comportamiento de la temperatura dentro del invernadero.

Palabras clave: Modelos ARX, modelos de caja negra, modelos dinámicos lineales, temperatura del invernadero.

INTRODUCCIÓN

Los invernaderos mexicanos dependen principalmente de la ventilación natural como mecanismos de control atmosférico, aunque son insuficientes para mitigar los efectos negativos de la temperatura y la humedad ambiental generadas en su interior, en un intervalo fuera de los valores óptimos para el desarrollo de los cultivos. Hay grandes fluctuaciones día/noche de los valores de la temperatura y la humedad relativa (López *et al.*, 2004, Costa y Giacomelli, 2005); además, los niveles normales de CO₂ disminuyen, especialmente durante las horas de mayor incidencia de radiación solar. Por tanto, para

ABSTRACT

A procedure is presented for obtaining a dynamic linear model of auto-regression with exogenous variables (ARX) for predicting the behaviour of the air temperature inside a greenhouse. The ARX are dynamic mathematical models derived from the theory of Systems Identification. The input variables of the model were air temperature, solar radiation, wind velocity affecting the ventilation area of the greenhouse and relative humidity, quantified in a meteorological station located 700 m from a greenhouse in Chapingo, State of México. The response variable was the air temperature inside the greenhouse. Samples were taken of the input and output variables of the model every 5 min during a crop cycle. To determine the structure of the best model, as many as 100 000 ARX models were evaluated using the information criteria and final prediction error of Akaike. The adjustment between the simulated and observed temperatures, and the residual analysis, indicated that ARX models of second degree or above, adequately predict the behaviour of the temperature inside the greenhouse.

Key words: ARX models, black box models, dynamic linear models, greenhouse temperature.

INTRODUCTION

Mexican greenhouses depend mainly on natural ventilation as atmospheric control mechanism, although they are insufficient to mitigate the negative effects of the temperature and atmospheric humidity generated inside it, in a range outside the optimum values for the development of crops. There are great daytime/night-time fluctuations of the values of temperature and relative humidity (López *et al.*, 2004; Costa and Giacomelli, 2005). Furthermore, the normal levels of CO₂ are reduced, especially during the hours of highest incidence of solar radiation. Therefore, for an adequate management of the crop, artificial climate of the greenhouse environment is required with systems of heating, forced ventilation, fogging, refrigeration and enrichment of the atmosphere with CO₂.

Recibido: Octubre, 2005. Aprobado: Diciembre, 2006.

Publicado como ARTÍCULO en *Agrociencia* 41: 181-192. 2007.

un buen manejo del cultivo se requiere climatización artificial del ambiente del invernadero con sistemas de calefacción, ventilación forzada, nebulización, refrigeración y enriquecimiento de la atmósfera mediante CO₂.

Una herramienta para mejorar el manejo de los invernaderos son los modelos matemáticos precisos y simples para predecir el comportamiento de las variables atmosféricas. En los últimos 20 años se ha estudiado el clima del invernadero usando modelos matemáticos, físicos o mecanicistas (Bot, 1983; Deltour *et al.*, 1985; Wang y Boulard, 2000), o modelos empíricos del tipo caja negra, como los de autoregresión con variables exógenas (ARX) (Uchida-Frausto *et al.*, 2003; Boaventura-Cunha, *et al.*, 1996). El desarrollo de modelos basados en el balance de masa y energía del ambiente en un invernadero requiere tiempo, ya que deben ser calibrados y validados antes de su optimización y control. Los modelos autoregresivos tienen la ventaja de ser generados rápidamente, ya que se obtienen experimentalmente al establecer las relaciones entrada-salida del sistema mediante técnicas de identificación de parámetros (Ljung, 1999).

Uchida-Frausto *et al.* (2003) estudiaron modelos ARX para predecir la temperatura del aire de un invernadero. Estos autores generaron estructuras de modelos para cada estación del año y encontraron que un modelo genérico con una calidad de predicción aceptable durante todo el año, siempre requiere recalibración o reajuste. Además, no utilizaron datos climáticos reales sino simulados, ni emplearon el análisis residual para evaluar la calidad de predicción de los modelos ARX. Por tanto, en la presente investigación se evaluó la metodología para generar modelos ARX y predecir el comportamiento de la temperatura del aire de un invernadero mexicano con ventilación natural.

MATERIALES Y MÉTODOS

Modelos de autoregresión con variables exógenas (ARX)

Un modelo general ARX (Ljung, 1999; Ljung 1988-2005; Aguado-Behar y Martínez-Iranzo, 2003) con una entrada y una salida se describe mediante la ecuación en diferencias lineales:

$$y(t) + a_1 y(t-1) + \dots + a_{na} y(t-na) = b_1 u(t-nk) + b_2 u(t-nk-1) + \dots + b_{nb} u(t-nk-nb+1) + e(t) \quad (1)$$

donde, $y(t)$ es la salida del modelo ARX para $t=t, t-1, \dots, t-na$; $u(t)$: es la entrada del modelo ARX modelo para $t=t-nk, t-nk-1, \dots, t-nk-nb+1$; na es el número de pasos de tiempo de la salida en el pasado; nb es el número de pasos de tiempo de la entrada en el pasado; nk es el retardo de la entrada $u(t)$ con respecto a la salida $y(t)$; $e(t)$ es ruido blanco asociado con la variable de salida.

A tool for improving the management of greenhouses are precise and simple mathematical models for predicting the behaviour of the atmospheric variables. In the past 20 years, greenhouse climate has been studied using mathematical, physical or mechanical models (Bot, 1983; Deltour *et al.*, 1985; Wang and Boulard, 2000), or empirical or black box models, such as the autoregression models with exogenous variables (ARX) (Uchida-Frausto *et al.*, 2003); Boaventura-Cunha *et al.*, 1996). The development of models based on the balance of mass and energy of the atmosphere of a greenhouse requires time, given that they must be calibrated and validated prior to their optimization and control. The autoregressive models have the advantage of being generated rapidly, as they are obtained experimentally when establishing the input-output relationships of the system through techniques of identification of parameters (Ljung, 1999).

Uchida-Frausto *et al.* (2003) studied ARX models to predict the air temperature of a greenhouse. These authors generated structures of models for each season of the year and found that a generic model with an acceptable quality of prediction throughout the year always requires recalibration or readjustment. Furthermore, they did not use real climatic data, but rather simulated data, nor did they employ the residual analysis to evaluate the prediction quality of the ARX models. Therefore, in the present study an evaluation was made of the methodology to generate ARX models and to predict the behaviour of the air temperature of a Mexican greenhouse with natural ventilation.

MATERIALS AND METHODS

Auto-regression models with exogenous variables (ARX)

A general ARX model (Ljung, 1999; Ljung 1988-2005; Aguado-Behar and Martínez-Iranzo, 2003) with an input and output is described by the equation in linear differences:

$$y(t) + a_1 y(t-1) + \dots + a_{na} y(t-na) = b_1 u(t-nk) + b_2 u(t-nk-1) + \dots + b_{nb} u(t-nk-nb+1) + e(t) \quad (1)$$

where $y(t)$ is the output of the ARX model for $t=t, t-1, \dots, t-na$; $u(t)$: is the input of the ARX model for $t=t-nk, t-nk-1, \dots, t-nk-nb+1$; na is the number of time steps of the output in the past; nb is the number of time steps of the entrance in the past; nk is the delay of the input $u(t)$ with respect to the output $y(t)$; $e(t)$ is the blank noise associated with the output variable.

An ARX model for estimating the temperature of a greenhouse requires more than one input variable, thus the ARX model should be adjusted with a single input and a single output (SISO) in an ARX

Un modelo ARX para estimar la temperatura de un invernadero requiere más de una variable de entrada por lo que se debe ajustar el modelo ARX con una sola entrada y una sola salida (SISO) en un modelo ARX con múltiples entradas y múltiples salidas (MIMO). Un modelo ARX MIMO, con entradas y salidas está definido por la ecuación:

$$A(q) y(t) = B(q)u(t) + e(t) \quad (2)$$

donde $A(q)$ es una matriz de dimensiones $ny \times ny$; sus elementos son polinomios en el operador de retraso $q^{-1} y(t) = y(t-1)$.

$$A(q) = I_{ny} + A_1 q^{-1} + \dots + A_{na} q^{-na} \quad (3)$$

I_{ny} es la matriz identidad. En forma matricial:

$$A(q) = \begin{bmatrix} a_{11}(q) & a_{12}(q) & \dots & a_{1ny}(q) \\ a_{21}(q) & a_{22}(q) & \dots & a_{2ny}(q) \\ \dots & \dots & \dots & \dots \\ a_{ny1}(q) & a_{ny2}(q) & \dots & a_{nyny}(q) \end{bmatrix} \quad (4)$$

donde los elementos a_{kj} son polinomios en el operador q^{-1} .

$$a_{kj}(q) = \delta_{kj} + a_{kj}^1 q^{-1} + \dots + a_{kj}^{na_{kj}} q^{-na_{kj}} \quad (5)$$

$$\delta_{kj} = \begin{cases} 1 & \text{si } k = j \\ 0 & \text{de lo contrario} \end{cases}$$

La matriz $B(q)$ está formada como $A(q)$ pero tiene dimensiones:

$$ny \times nu, B(q) = B_0 + B_1 q^{-1} + \dots + B_{nb} q^{-nb} \quad (6)$$

$$B(q) = \begin{bmatrix} b_{11}(q) & b_{12}(q) & \dots & b_{1nu}(q) \\ b_{21}(q) & b_{22}(q) & \dots & b_{2nu}(q) \\ \dots & \dots & \dots & \dots \\ b_{ny1}(q) & b_{ny2}(q) & \dots & b_{nynu}(q) \end{bmatrix} \quad (7)$$

$$b_{kj}(q) = b_{kj}^1 q^{-nk_{kj}} + \dots + b_{kj}^{nb_{kj}} q^{-nk_{kj} - nb_{kj} + 1} \quad (8)$$

Para establecer la estructura de un modelo ARX se requiere determinar los coeficientes $a_{11}, a_{12}, a_{ny1}, \dots, a_{ny2}, \dots, a_{nyny}$ y $b_{11}, b_{12}, \dots, b_{1nu}, \dots, b_{ny1}, b_{ny2}, \dots, b_{nyny}$, así como el orden del modelo dado por los valores de los parámetros na, bn y nk , a partir de la información de las variables de entrada y salida que permitan obtener el mejor ajuste entre la predicción del modelo y la variable de salida medida.

Identificación de la temperatura en un invernadero

Paso 1.

Diseño del experimento para obtener datos de las variables de entrada y salida: las variables de entrada fueron: temperatura del

model with multiple inputs and multiple outputs (MIMO). A MIMO ARX model with inputs and outputs is defined by the equation:

$$A(q) y(t) = B(q)u(t) + e(t) \quad (2)$$

where $A(q)$ is a matrix with dimensions $ny \times ny$; its elements are polynomials in the delay operator $q^{-1} y(t) = y(t-1)$.

$$A(q) = I_{ny} + A_1 q^{-1} + \dots + A_{na} q^{-na} \quad (3)$$

I_{ny} is the identity matrix. In matricial form:

$$A(q) = \begin{bmatrix} a_{11}(q) & a_{12}(q) & \dots & a_{1ny}(q) \\ a_{21}(q) & a_{22}(q) & \dots & a_{2ny}(q) \\ \dots & \dots & \dots & \dots \\ a_{ny1}(q) & a_{ny2}(q) & \dots & a_{nyny}(q) \end{bmatrix} \quad (4)$$

where the elements a_{kj} are polynomials in the operator q^{-1} .

$$a_{kj}(q) = \delta_{kj} + a_{kj}^1 q^{-1} + \dots + a_{kj}^{na_{kj}} q^{-na_{kj}} \quad (5)$$

$$\delta_{kj} = \begin{cases} 1 & \text{si } k = j \\ 0 & \text{de lo contrario} \end{cases}$$

The matrix $B(q)$ is formed like $A(q)$, but its dimensions are:

$$ny \times nu, B(q) = B_0 + B_1 q^{-1} + \dots + B_{nb} q^{-nb} \quad (6)$$

$$B(q) = \begin{bmatrix} b_{11}(q) & b_{12}(q) & \dots & b_{1nu}(q) \\ b_{21}(q) & b_{22}(q) & \dots & b_{2nu}(q) \\ \dots & \dots & \dots & \dots \\ b_{ny1}(q) & b_{ny2}(q) & \dots & b_{nynu}(q) \end{bmatrix} \quad (7)$$

$$b_{kj}(q) = b_{kj}^1 q^{-nk_{kj}} + \dots + b_{kj}^{nb_{kj}} q^{-nk_{kj} - nb_{kj} + 1} \quad (8)$$

To establish the structure of an ARX model, it is necessary to determine the coefficients $a_{11}, a_{12}, a_{ny1}, \dots, a_{ny2}, \dots, a_{nyny}$ and $b_{11}, b_{12}, \dots, b_{ny1}, b_{ny2}, \dots, b_{nyny}$, as well as the order of the model given by the values of the parameters na, nb and nk , from the information of the input and output variables that make it possible to obtain the best fit between the prediction of the model and the measured output variable.

Identification of the temperature of a greenhouse

Step 1.

Experimental design to obtain data of the input and output variables: the input variables were: air temperature ($u_1(t)$, °C), measured at 2.5 m height, solar radiation ($u_2(t)$, W m⁻²), wind velocity ($u_3(t)$ ($u_3(t)$, m s⁻¹ measured at 2.5 m height and relative

aire ($u_1(t)$, °C) medida a 2.5 m de altura, radiación solar ($u_2(t)$, $W m^{-2}$), velocidad del viento ($u_3(t)$, $m s^{-1}$) medida a 2.5 m de altura y humedad relativa (HR) ($u_3(t)$, %), medidas fuera del invernadero. La variable de salida fue la temperatura del aire ($y(t)$, °C) dentro del invernadero medida a 1.5 m de altura. El tiempo de muestreo fue 5 min.

Las variables meteorológicas temperatura, humedad relativa y radiación solar se tomaron en un invernadero de 1080 m² (27 m anchura×40 m longitud), 19° 29' N, 98°, 53' O, a una altitud de 2244 m, en el campo experimental Tlapeaxco de la Universidad Autónoma Chapingo, Estado de México, México. La orientación (azimut) del invernadero es 48° E con respecto al eje Norte-Sur. La altura de las paredes laterales del invernadero es 2.1 m. El invernadero tiene una sola cubierta de plástico calibre 700 ultravioleta y ventanas laterales (170 m²) con mallas anti-insectos. El área total del invernadero está dividida en tres naves con una ventana cenital cada una y el área total de ventilación cenital tiene 60 m². La temperatura y la HR del aire se regulan mediante ventilación natural (cierre y apertura de ventilas) y un sistema de calefacción directa. El sistema de calefacción no operó durante el experimento.

Las variables de entrada y salida se midieron de abril a julio de 2004 (Figura 1) mediante dos estaciones meteorológicas automáticas marca Davis Weather modelo Standard GroWeather Station (Standard Groweather station, Davis Instruments Corp, USA). Una se ubicó en el centro del invernadero y la otra a 700 m de distancia dentro de la estación meteorológica del Campo experimental Montecillo de la Universidad Autónoma Chapingo.

De acuerdo con la operación normal de un invernadero comercial, la ventilación natural fue través de las ventilas laterales y cenitales, abiertas las 24 h del día. Sin embargo, en los modelos ARX hubo dos estrategias de ventilación: ventilas abiertas todo el día (siempre abiertas), y apertura de ventilas sólo de 8:00 a 18:00 h (cierre y apertura de ventilas).

Paso 2.

Selección del tipo de modelo. En esta investigación sólo se estudiaron modelos ARX.

Paso 3.

Elección del criterio de ajuste. Para determinar los valores óptimos de na , nb y nk se usaron el error de predicción final de Akaike (FPE) y el criterio de información de Akaike (AIC). Estos criterios utilizan una función de pérdida mínima en combinación con un número mínimo de parámetros estimados (Ljung, 1999; Nelles, 2001):

$$FPE = \left(\frac{1 + \frac{d}{N}}{1 - \frac{d}{N}} \right) V \tag{9}$$

$$AIC = \log \left(V \left(1 + 2 \frac{d}{N} \right) \right) \tag{10}$$

humidity (RH) ($u_3(t)$,%), measured outside the greenhouse. The output variable was the air temperature ($y(t)$, °C) inside the greenhouse measured at 1.5 m height. The sampling time was 5 min.

The meteorological variables of temperature, relative humidity and solar radiation were taken in a greenhouse of 1080 m² (27 m width×40 m length), 19° 29' N, 98°, 53' W, at an altitude of 2244 m, in the experimental field Tlapeaxco of the Universidad Autónoma Chapingo, State of México, México. The orientation (azimut) of the greenhouse is 48° E with respect to the North-South axis. The height of the lateral walls of the greenhouse is 2.1 m. The greenhouse has a single plastic cover calibre 700 ultraviolet and lateral windows (170 m²) with anti-insect screens. The total area of the greenhouse is divided into three naves with a zenithal window in each one, and the total area of zenithal ventilation is 60 m². The air temperature and RH are regulated by means of natural ventilation (closing and opening of vents) and a direct heating system. The heating system did not operate during the experiment.

The input and output variables were measured from April to July of 2004 (Figure 1) through two Davis Weather automatic meteorological stations, model Standard GroWeather Station (Standard GroWeather Station, Davis Instruments Corp, USA). One was located in the center of the greenhouse and the other at 700 m distance within the meteorological station of the Montecillo experimental field of the Universidad Autónoma Chapingo.

According to the normal operation of a commercial greenhouse, natural ventilation was through the lateral and zenithal vents, which

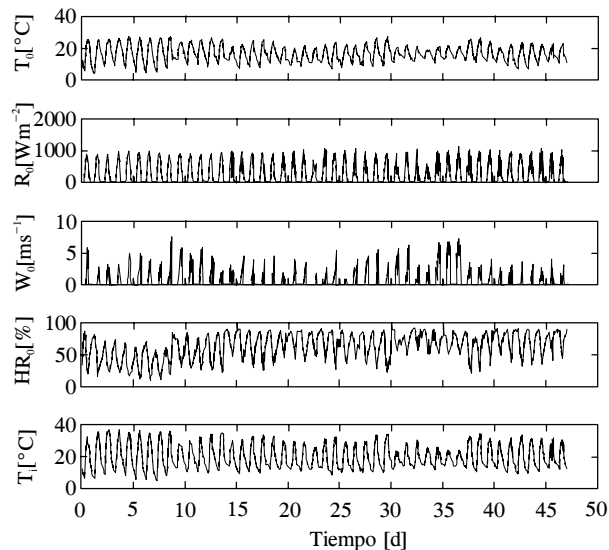


Figura 1. Variables meteorológicas: temperatura (T_o), radiación solar (R_o), humedad relativa (HR_o), velocidad del viento (W_o) externas y temperaturas dentro del invernadero (T_i), medidas de abril a julio de 2004, en Chapingo, México.

Figure 1. Meteorological variables: external temperature (T_o), solar radiation (R_o), relative humidity (HR_o), wind velocity (W_o), and temperatures inside the greenhouse (T_i), measured from April to July of 2004, in Chapingo, México.

donde V es una función de pérdida, típicamente un ajuste cuadrático de la estructura; d es el número de parámetros estimados; N es el número de datos.

Paso 4.

Selección de los subconjuntos de datos para estimar y validar los modelos ARX. Se evaluaron los subconjuntos de datos 90%:10%, 80%:20%, 77%:23%, 66.67%:33.33% y 50%:50% para estimar y validar los modelos. Se empleó la herramienta de identificación de sistemas (System Identification Toolbox) disponible en el software Matlab (Ljung, 1998-2005), para estimar la estructura de los modelos y para su validación. Se evaluaron 100 000 modelos ARX, ya que el valor máximo para na (una variable de salida) y para nb (cuatro variables de entrada) fue 10 y para nk uno, por tanto, el número de combinaciones (estructuras de modelos) resultante es 10^5 .

Paso 5.

Evaluación de los modelos. Se usó el nivel ajuste (%) entre temperaturas observadas y simuladas, así como el análisis residual para evaluar la calidad predictiva del modelo.

RESULTADOS Y DISCUSIÓN

En los Cuadros 1 a 5 se muestran los resultados obtenidos al evaluar los subconjuntos de datos para estimar y validar los modelos ARX usando las dos estrategias de ventilación natural.

Para los datos de 90%:10% y 80%:20% se observó sobreajuste en el conjunto de datos de validación. Ésto se muestra en las últimas dos columnas de los Cuadros 1 y 2, donde los porcentajes de ajuste sobre el conjunto de estimación son menores que los obtenidos sobre el conjunto de validación.

Para los datos de 66.67%:33.33% y 50%:50% los porcentajes de ajuste sobre el conjunto de estimación son mayores que los del conjunto de validación. Ésto se presenta en las últimas dos columnas de los Cuadros 3 y 4.

El análisis residual mostró que tanto la autocorrelación de los errores como la correlación cruzada entre residuales y variables de entrada fueron satisfechas usando 90%:10% y 80%:20%. Ésto se muestra en las Figuras 2A a 6A, donde se presenta el comportamiento para el modelo con estructura, $na=2$, $nb=[2 \ 1 \ 2 \ 1]$, $nk=[1 \ 1 \ 1 \ 1]$ (cuarto renglón del Cuadro 1) uno de los mejores modelos obtenidos empleando 90%:10%. Las funciones de correlación cruzada están dentro del intervalo de confianza de 99%, y la auto correlación de los residuales es $corr_{xx}(0)=1$ y cercana a cero para otros valores de retraso. Estos modelos no son aceptables debido al sobre ajuste.

were open 24 h a day. However, there were two ventilation strategies in the ARX models: vents open all day (always open), and vents open only from 8:00 to 18:00 h (closing and opening of vents).

Step 2.

Selection of the type of model. In the present investigation, only ARX models were studied.

Step 3.

Selection of the adjustment criterion. To determine the optimum values of na , nb and nk , the Akaike final prediction error (PFE) was used along with the Akaike information criterion (AIC). These criteria were used in a minimum loss function in combination with a minimum number of estimated parameters (Ljung, 1999; Nelles, 2001):

$$FPE = \left(\frac{1 + \frac{d}{N}}{1 - \frac{d}{N}} \right) V \tag{9}$$

$$AIC = \log \left(V \left(1 + 2 \frac{d}{N} \right) \right) \tag{10}$$

where V is a function of loss, typically a quadratic adjustment of the structure; d is the number of estimated parameters; N is the number of data.

Step 4.

Selection of the subgroups of data for estimating and validating the ARX models. The data subgroups of 90%:10%, 80%:20%, 77%:23%, 66.67%:33.33% and 50%:50% were evaluated to estimate and validate the models. The systems identification tool (System Identification Toolbox) available in the Matlab software (Ljung, 1998-2005) was used to estimate the structure of the models and for their validation. An evaluation was made of 100,000 ARX models, given that the maximum value for na (one exit variable) and for nb (four entrance variables) was 10 and for nk one, therefore, the resulting number of combinations (structures of models) is 10^5 .

Step 5.

Evaluation of the models. The adjustment level (%) was used between observed and simulated temperatures, as well as the residual analysis to evaluate the predictive quality of the model.

RESULTS AND DISCUSSION

Tables 1 to 5 show the results obtained when evaluating the subgroups of data to estimate and evaluate

Cuadro 1. Mejores modelos ARX obtenidos usando 90% de los datos para estimación y 10% para validación del modelo.
Table 1. Best ARX models obtained using 90% of the data for estimation and 10% for validation of the model.

	Estructura									FPE ($\times 10^{-2}$)	AIC	Función de pérdida ($\times 10^{-2}$)	Ajuste en estimación (%)	Ajuste en validación (%)
	<i>na</i>			<i>nb</i>			<i>nk</i>							
SA	3	1	2	10	2	0	0	0	0	2.22	-3.80	2.21	77.75	80.26
SA	3	1	2	7	2	0	0	0	0	2.22	-3.80	2.21	77.74	80.25
SA	3	1	2	3	2	0	0	0	0	2.22	-3.80	2.21	77.74	80.22
SA	2	2	1	2	1	1	1	1	1	2.47	-3.69	2.47	72.93	74.55
CA	2	2	1	2	1	1	1	1	1	2.48	-3.69	2.48	70.32	73.33
CA	2	2	1	9	1	1	1	1	1	2.48	-3.69	2.48	70.36	73.49
CA	3	1	2	3	2	0	0	0	0	2.23	-3.80	2.22	75.74	79.89
CA	3	1	2	5	2	0	0	0	0	2.23	-3.80	2.22	75.73	79.88

na: número de retrasos de la variable temperatura dentro del invernadero.

nb: número de retrasos para las variables de entrada temperatura externa, radiación solar, humedad relativa y velocidad del viento.

nk: retardo entre las variables de entrada y salida.

SA: ventilas siempre abiertas.

CA: cierre y apertura de ventilas.

FPE: error de predicción final de Akaike.

AIC: criterio de información de Akaike.

Cuadro 2. Mejores modelos ARX obtenidos usando 80% de datos para estimación y 20% para validación del modelo.
Table 2. Best ARX models obtained using 80% of the data for estimation and 20% for validation of the model.

	Estructura									FPE ($\times 10^{-2}$)	AIC	Función de pérdida ($\times 10^{-2}$)	Ajuste en estimación (%)	Ajuste en validación (%)
	<i>na</i>			<i>nb</i>			<i>nk</i>							
SA	10	3	10	6	4	0	0	0	0	1.95	-3.93	1.94	79.64	80.26
SA	10	3	8	3	4	0	0	0	0	1.96	-3.93	1.95	79.90	79.98
SA	3	2	8	3	2	0	0	0	0	1.98	-3.91	1.98	79.48	78.81
SA	10	1	8	7	1	1	1	1	1	2.24	-3.79	2.23	70.27	72.15
SA	10	1	8	2	1	1	1	1	1	2.24	-3.79	2.23	70.48	72.16
SA	3	1	2	1	1	1	1	1	1	2.28	-3.77	2.28	71.47	71.31
CA	3	1	2	1	1	1	1	1	1	2.29	-3.77	2.29	68.72	69.68
CA	10	1	8	6	1	1	1	1	1	2.24	-3.79	2.23	68.85	71.33
CA	3	2	8	2	2	0	0	0	0	1.99	-3.91	1.98	77.87	78.27
CA	10	3	9	6	4	0	0	0	0	1.95	-3.93	1.94	78.52	79.86
CA	10	3	10	7	4	0	0	0	0	1.95	-3.93	1.94	78.51	79.86

na: número de retrasos de la variable temperatura dentro del invernadero.

nb: número de retrasos para las variables de entrada temperatura externa, radiación solar, humedad relativa y velocidad del viento.

nk: retardo entre las variables de entrada y salida.

SA: ventilas siempre abiertas.

CA: cierre y apertura de ventilas.

FPE: error de predicción final de Akaike.

AIC: criterio de información de Akaike.

En las Figuras 2B a 6B se muestra el comportamiento del modelo con estructura $na=3$, $nb=[2\ 2\ 1\ 1]$ y $nk=[0\ 0\ 0\ 0]$ (segunda fila del Cuadro 4), uno de los mejores modelos obtenidos empleando 50%:50%. Las funciones de correlación cruzada entre residuales y las variables de entrada, predominantemente están fuera de los intervalos de confianza. Por tanto, los subconjuntos de datos óptimos están entre 80%:20% y 66.67%:33.33%. Después de analizar varias

the ARX models using the two natural ventilation strategies.

For the data of 90%:10% and 80%:20%, over-fit was observed in the group of validation data. This is shown in the last two columns of Tables 1 and 2, where the percentages of fit of the estimation group are lower than those obtained in the validation group.

For the data of 66.67%:33.33% and 50%:50%, the percentages of fit of the estimation group are higher

Cuadro 3. Mejores modelos ARX obtenidos usando 66.66% de datos para estimación y 33.33% para validación del modelo.
Table 3. Best ARX models obtained using 66.6% of the data for estimation and 33.33% for validation of the model.

	Estructura									FPE ($\times 10^{-2}$)	AIC	Función de pérdida ($\times 10^{-2}$)	Ajuste en estimación (%)	Ajuste en validación (%)
	<i>na</i>			<i>nb</i>			<i>nk</i>							
SA	5	9	3	1	6	0	0	0	0	1.90	-3.96	1.89	81.79	75.96
SA	2	4	2	1	4	0	0	0	0	1.98	-3.91	1.98	80.57	75.68
SA	2	4	2	1	2	0	0	0	0	1.99	-3.91	1.98	80.48	75.68
SA	2	8	2	2	5	1	1	1	1	2.14	-3.84	2.13	78.02	69.98
SA	2	6	2	2	1	1	1	1	1	2.15	-3.83	2.15	77.39	69.72
CA	10	9	9	3	4	0	0	0	0	1.85	-3.98	1.84	81.69	70.33
CA	10	4	9	1	4	0	0	0	0	1.88	-3.96	1.87	80.87	74.08
CA	2	4	2	1	2	0	0	0	0	2.00	-3.91	1.99	79.61	72.54
CA	2	7	2	10	5	1	1	1	1	2.15	-3.83	2.14	76.54	64.39
CA	2	8	2	2	5	1	1	1	1	2.14	-3.84	2.14	76.88	65.00
CA	2	6	2	1	1	1	1	1	1	2.16	-3.83	2.15	76.28	65.47

na: número de retrasos de la variable temperatura dentro del invernadero.
nb: número de retrasos para las variables de entrada temperatura externa, radiación solar, humedad relativa y velocidad del viento.
nk: retardo entre las variables de entrada y salida.
 SA: ventilas siempre abiertas.
 CA: cierre y apertura de ventilas.
 FPE: error de predicción final de Akaike.
 AIC: criterio de información de Akaike.

Cuadro 4. Mejores modelos ARX obtenidos usando 50% de datos para estimación y 50% para validación del modelo.
Table 4. Best ARX models obtained using 50% of the data for estimation and 50% for validation of the model.

	Estructura									FPE ($\times 10^{-2}$)	AIC	Función de pérdida ($\times 10^{-2}$)	Ajuste en estimación (%)	Ajuste en validación (%)
	<i>na</i>			<i>nb</i>			<i>nk</i>							
SA	10	2	2	1	1	0	0	0	0	1.77	-4.03	1.762	80.43	76.38
SA	3	2	2	1	1	0	0	0	0	1.79	-4.02	1.785	80.16	76.57
SA	3	10	2	4	10	0	0	0	0	1.88	-3.97	1.866	79.27	71.02
SA	3	7	2	2	1	1	1	1	1	1.89	-3.96	1.889	78.07	70.78
SA	3	7	2	1	1	1	1	1	1	1.89	-3.96	1.890	78.12	71.18
CA	10	2	8	1	1	0	0	0	0	1.75	-4.04	1.739	79.65	75.52
CA	3	2	2	1	1	0	0	0	0	1.79	-4.01	1.792	79.30	74.65
CA	3	10	2	8	10	1	1	1	1	1.88	-3.97	1.866	78.30	67.07
CA	3	7	2	2	1	1	1	1	1	1.89	-3.96	1.890	77.09	67.14
CA	3	7	2	1	1	1	1	1	1	1.89	-3.96	1.891	77.15	67.62

na: número de retrasos de la variable temperatura dentro del invernadero.
nb: número de retrasos para las variables de entrada temperatura externa, radiación solar, humedad relativa y velocidad del viento.
nk: retardo entre las variables de entrada y salida.
 SA: ventilas siempre abiertas.
 CA: cierre y apertura de ventilas.
 FPE: error de predicción final de Akaike.
 AIC: criterio de información de Akaike.

combinaciones, se encontró que el mejor subconjunto de datos que evita el sobreajuste en el conjunto de validación y satisface la prueba de los residuales fue 77% para estimación y 23% para validación de los modelos (Cuadro 5).

Dado que el objetivo es obtener el mejor modelo para predecir la temperatura dentro del invernadero, se empleó el procedimiento de validación cruzada para

than those of the validation group. This is shown in the last two columns of Tables 3 and 4.

The residual analysis showed that both the auto-correlation of the errors and the cross correlation between residuals and input variables were satisfied using 90%:10% and 80%:20%. This is shown in Figures 2A to 6A, where the behaviour for the model with structure is presented, $na = 2$, $nb = [2 \ 1 \ 2 \ 1]$,

Cuadro 5. Mejores modelos ARX obtenidos usando 77% de los datos para estimación y 23% para validación del modelo.
Table 5. Best ARX models obtained using 77% of the data for estimation and 23% for validation of the model.

	Estructura									FPE ($\times 10^{-2}$)	AIC	Función de pérdida ($\times 10^{-2}$)	Ajuste en estimación (%)	Ajuste en validación (%)
	na			nb			nk							
SA	10	4	10	2	5	0	0	0	0	1.90	-3.95	1.89	80.31	78.94
SA	10	3	9	1	2	0	0	0	0	1.93	-3.95	1.92	79.87	79.10
SA	3	4	2	1	2	0	0	0	0	1.95	-3.93	1.95	80.22	78.23
SA	2	10	7	2	5	1	1	1	1	2.16	-3.83	2.15	76.06	74.06
SA	2	8	7	2	1	1	1	1	1	2.17	-3.82	2.17	75.67	73.33
SA	2	3	2	1	1	1	1	1	1	2.23	-3.79	2.23	74.45	72.00
CA	10	3	10	8	5	0	0	0	0	1.91	-3.95	1.89	78.92	77.20
CA	10	4	9	2	4	0	0	0	0	1.90	-3.96	1.89	79.59	77.82
CA	10	3	9	2	2	0	0	0	0	1.91	-3.95	1.90	79.10	77.60
CA	2	3	2	1	1	1	1	1	1	2.24	-3.79	2.24	72.39	69.49
CA	2	8	7	5	1	1	1	1	1	2.18	-3.82	2.17	73.25	70.72
CA	2	10	7	5	5	1	1	1	1	2.17	-3.83	2.15	73.62	71.60

na: número de retrasos de la variable temperatura dentro del invernadero.
 nb: número de retrasos para las variables de entrada temperatura externa, radiación solar, humedad relativa y velocidad del viento.
 nk: retardo entre las variables de entrada y salida.
 SA: ventilas siempre abiertas.
 CA: cierre y apertura de ventilas.
 FPE: error de predicción final de Akaike.
 AIC: criterio de información de Akaike.

elegir de este subconjunto el modelo con el mejor ajuste y que además satisficiera la prueba de los residuales. El mejor modelo con base en el criterio de información de Akaike, el error de predicción final de Akaike y grado de ajuste fue el que tiene la estructura $na=10$,

$nk=[1\ 1\ 1\ 1]$ (fourth line of Table 1) one of the best models obtained employing 90%:10%. The cross correlation functions are within the confidence interval of 99%, and the auto-correlation of the residuals is

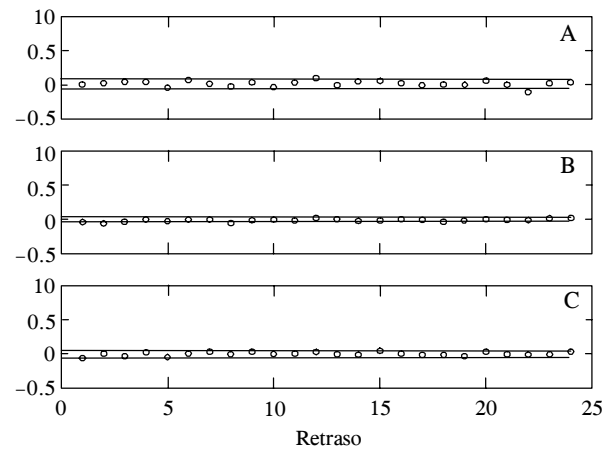


Figura 2. Funciones de autocorrelación de los residuales para los modelos ARX. (A) modelo de segundo orden $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) modelo de tercer orden $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) modelo de segundo orden $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.
Figure 2. Functions of auto-correlation of the residuals for the ARX models. (A) second order model $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) third order model $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) second order model $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

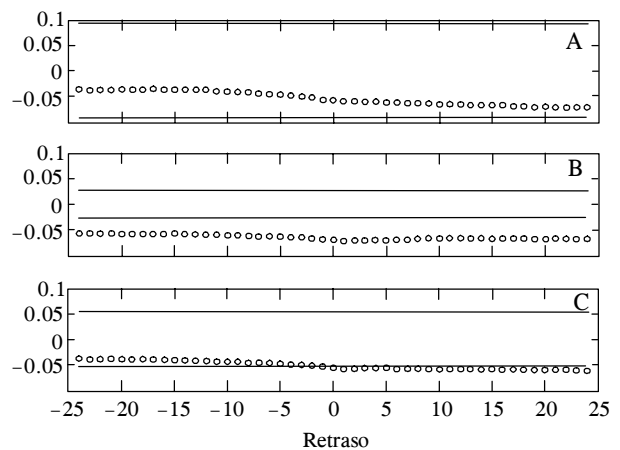


Figura 3. Funciones de correlación cruzada de los residuales contra la temperatura fuera del invernadero. (A) modelo de segundo orden $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) modelo de tercer orden $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) modelo de segundo orden $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.
Figure 3. Functions of cross correlation of the residuals against the temperature outside the greenhouse. (A) second order model $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) third order model $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) second order model $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

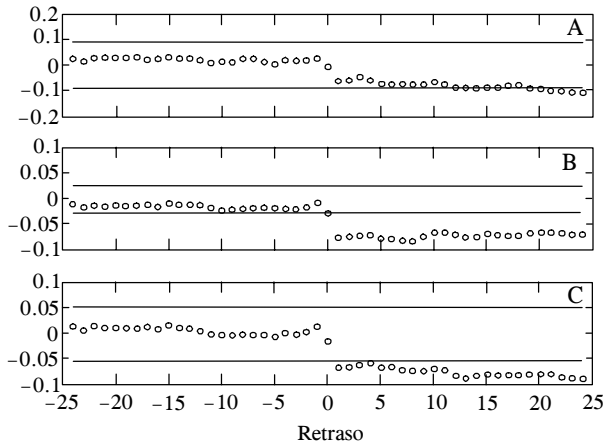


Figura 4. Funciones de correlación cruzada de los residuales contra la radiación solar fuera del invernadero. (A) modelo de segundo orden $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) modelo de tercer orden $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) modelo de segundo orden $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

Figure 4. Functions of cross correlation of the residuals against the solar radiation outside of the greenhouse. (A) second order model $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) third order model $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) second order model $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

$nb=[4\ 10\ 2\ 5]$, $nk=[0\ 0\ 0\ 0]$ (primer renglón Cuadro 5), aunque el modelo con estructura $na=3$, $nb=[4\ 2\ 1\ 2]$,

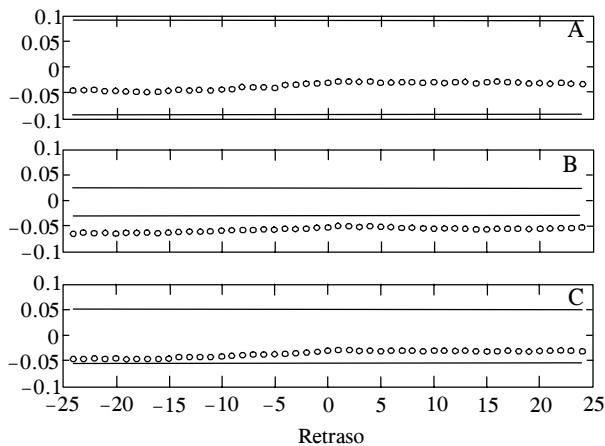


Figura 5. Funciones de correlación cruzada entre los residuales y la velocidad del viento fuera del invernadero. (A) modelo de segundo orden $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) modelo de tercer orden $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) modelo de segundo orden $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

Figure 5. Functions of cross correlation between the residuals and the wind velocity outside of the greenhouse. (A) second order model $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) third order model $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) second order model $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

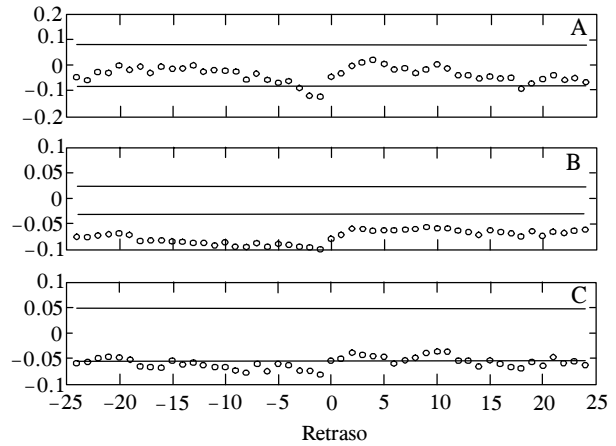


Figura 6. Funciones de correlación cruzada entre los residuales y la humedad relativa fuera del invernadero. (A) modelo de segundo orden $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) modelo de tercer orden $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) modelo de segundo orden $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

Figure 6. Functions of cross correlation between the residuals and the relative humidity outside of the greenhouse. (A) second order model $na=2$, $nb=[2\ 1\ 2\ 1]$, $nk=[1\ 1\ 1\ 1]$; (B) third order model $na=3$, $nb=[2\ 2\ 1\ 1]$, $nk=[0\ 0\ 0\ 0]$; (C) second order model $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$.

$corr_{xx}(0)=1$ and close to zero for other values of delay. These models are not acceptable due to over-fit.

Figures 2B to 6B show the behaviour of the model with structure $na=3$, $nb=[2\ 2\ 1\ 1]$ and $nk=[0\ 0\ 0\ 0]$ (second row of Table 4), one of the best models obtained employing 50%:50%. The functions of cross correlation between residuals and the input variables, are predominantly outside the confidence intervals. Therefore, the optimum data subgroups are between 80%:20% and 66.67%:33.33%. After analyzing various combinations, it was found that the best data subgroup that avoids over-fit in the validation group and that satisfies the proof of the residuals was 77% for estimation and 23% for validation of the models (Table 5). Given that the objective is to obtain the best model for predicting the temperature inside the greenhouse, the cross validation procedure was used to select from this subgroup the model with the best fit and which would satisfy the test of the residuals. The best model based on the Akaike information criterion, the Akaike final prediction error and degree of fit was that having the structure $na=10$, $nb=[4\ 10\ 2\ 5]$, $nk=[0\ 0\ 0\ 0]$ (first line Table 5), although the model with structure $na=3$, $nb=[4\ 2\ 1\ 2]$, $nk=[0\ 0\ 0\ 0]$ (third line Table 5) presents a similar degree of fit. As from the point of view of the Control Theory (Kirk, 2004; Ljung, 1999) it is important to know the predictive quality of the model, the value of nk should be other than zero.

$nk=[0\ 0\ 0\ 0]$ (tercer reglón Cuadro 5) presenta un grado de ajuste similar. Como desde el punto de vista de Teoría de Control (Kirk, 2004; Ljung, 1999) es importante conocer la calidad predictiva del modelo, el valor de nk debe ser diferente de cero. Por tanto, se evaluaron los mejores modelos cuando el retardo entre las salidas y las entradas es la unidad. Los modelos de décimo y tercer orden aunque son los mejores de acuerdo con los criterios establecidos, no son los seleccionados ya que no hay retardo entre las variables de entrada y la salida ($nk=[0\ 0\ 0\ 0]$).

En el Cuadro 5 se observa que aun el modelo de segundo orden con estructura $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$ tiene un comportamiento aceptable si se considera tanto la calidad de su ajuste al conjunto de datos de validación, como el análisis residual que se muestra en las Figuras 2c a 6c. Si se considera el criterio FPE, la diferencia entre los mejores modelos reportados en el Cuadro 3 es muy pequeña, ya que en ambos conjuntos de datos las varianzas fueron 2.16×10^{-6} y 2.57×10^{-6} . Si se considera el error (función de pérdida), las varianzas fueron 2.19×10^{-6} y 2.63×10^{-6} .

Esto es importante al elegir el mejor modelo ya que si el comportamiento es bastante similar entre ellos y algunos presentan una estructura compleja y otros simple, desde el punto de vista de la Teoría de Control se prefiere el modelo con la estructura más simple en cuanto a los valores de na , nb y número de parámetros. Así se facilita el proceso de diseño de un sistema de control.

La estructura completa del modelo ($na=2$, $nb=[3\ 2\ 1\ 1]$), $nk=[1\ 1\ 1\ 1]$) se presenta en la siguiente ecuación:

$$T_i(t) = \frac{[T_o(t)R_o(t)H_o(t)V_o(t)]}{1 - 1.6211q^{-1} + 0.6349q^{-2}} \begin{bmatrix} 0.07457q^{-1} & 0.04322q^{-2} & -0.1079q^{-3} \\ 0.0003737q^{-1} & -0.0001219q^{-2} & \\ -0.01186q^{-1} & & \\ 0.0007896q^{-1} & & \end{bmatrix} \quad (11)$$

De acuerdo con los valores de los coeficientes, la variable más importante es la temperatura fuera del invernadero, seguida por la humedad relativa, velocidad del viento y radiación solar. Estos resultados, a excepción de la velocidad del viento, están de acuerdo con lo reportado por Uchida-Frausto *et al.*(2003). En la presente investigación, la nubosidad no fue analizada.

Therefore, the best models were evaluated when the delay between the outputs and the inputs is the unit. Although the tenth and third order models are the best according to the established criteria, they are not selected, because there is no delay between the input and output variables ($nk = [0\ 0\ 0\ 0]$).

In Table 5, it can be observed that even the second order model with structure $na=2$, $nb=[3\ 2\ 1\ 1]$, $nk=[1\ 1\ 1\ 1]$ has an acceptable behaviour if one considers both the quality of its fit to the group of validation data and the residual analysis shown in Figures 2c to 6c. If the FPE criterion is considered, the difference among the best models reported in Table 3 is very small, as in both groups of data the variances were 2.16×10^{-6} and 2.63×10^{-6} .

This is important when selecting the best model, because if the behaviour is very similar among them and some present a complex structure while others present a simple one, from the point of view of the Control Theory the preferred model is the one with the simpler structure with respect to the values of na , nb and number of parameters. Thus the design process of a control system is facilitated.

The complete structure of the model ($na=2$, $nb=[3\ 2\ 1\ 1]$), $nk=[1\ 1\ 1\ 1]$ is presented in the following equation:

$$T_i(t) = \frac{[T_o(t)R_o(t)H_o(t)V_o(t)]}{1 - 1.6211q^{-1} + 0.6349q^{-2}} \begin{bmatrix} 0.07457q^{-1} & 0.04322q^{-2} & -0.1079q^{-3} \\ 0.0003737q^{-1} & -0.0001219q^{-2} & \\ -0.01186q^{-1} & & \\ 0.0007896q^{-1} & & \end{bmatrix} \quad (11)$$

According to the values of the coefficients, the most important variable is the temperature outside of the greenhouse, followed by relative humidity, wind velocity and solar radiation. These results, except for wind velocity, agree with what was reported by Uchida-Frausto *et al.* (2003). In the present investigation, nubosity was not analyzed.

The matrix of covariances associated with the estimation of the coefficients a_{11} , a_{12} , $a_{ny1}, \dots, a_{ny1}, \dots, a_{ny2}, \dots, a_{nyny}, \dots; b_{11}$, b_{12} , $b_{ny1}, \dots, b_{ny2}, \dots, a_{nyny}, \dots;$ makes it possible to know the standard deviations, which provide information of the uncertainty of the estimation of each parameter. The standard deviations associated with each of the parameters of the ARX model are presented in the matrix M_{sd} (12), where the first row

La matriz de covarianzas asociada con la estimación de los coeficientes $a_{11}, a_{12}, any, \dots, a_{ny1}, \dots, a_{ny2}, \dots, a_{nyny}, \dots; b_{11}, b_{12}, bny_1, \dots, b_{ny2}, \dots, a_{nyny}, \dots$; permite conocer las desviaciones estándar, las cuales proporcionan información sobre la incertidumbre de la estimación de cada parámetro. Las desviaciones estándar asociadas con cada uno de los parámetros del modelo ARX se presentan en la matriz M_{sd} (12), donde la primera hilera contiene las desviaciones estándar de los coeficientes correspondientes a la temperatura del aire del invernadero y las siguientes hileras muestran las desviaciones pertenecientes a las variables de entrada: temperatura, radiación, velocidad del viento y humedad relativa cuantificadas fuera del invernadero.

$$M_{sd} = \begin{bmatrix} 0 & 0.0081 & 0.0079 \\ 0 & 0.0114 & 0.0177 & 0 \\ 0 & 0.0000 & 0.0000 & 0 \\ 0 & 0.0015 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 \end{bmatrix} \quad (12)$$

La matriz M_{sd} (12) muestra que los coeficientes del modelo ARX son determinados con precisión y aquellos asociados con la temperatura fuera del invernadero presentan mayor incertidumbre.

Al usar un modelo ARX como parte de un algoritmo de control de la temperatura del aire de un invernadero para predecir escenarios, es importante que en su proceso de obtención se determine el tamaño óptimo de datos para estimación y validación. De acuerdo con nuestros resultados se puede predecir sin recalibrar el modelo hasta cerca de un 23% de los datos no usados para generar la estructura. De otra manera, como mostró Uchida-Frausto et al. (2003), será necesario recalibrar periódicamente el modelo o usar algoritmos recursivos para determinar nuevas estructuras ante nuevos escenarios (Ljung, 1999; Aguado-Behar y Martínez-Iranzo, 2003). Los modelos ARX fueron similares con y sin implementación de cierre y apertura de ventilas, tanto en su estructura como en su calidad predictiva. Esto sugiere que la ventilación natural es insuficiente para regular en forma apropiada la temperatura dentro del invernadero.

CONCLUSIONES

Se propuso y evaluó una metodología para generar modelos ARX de la temperatura del aire dentro de un invernadero a partir de las variables externas temperatura, radiación solar, humedad relativa y velocidad del viento. El procedimiento propuesto permite obtener

contains the standard deviations of the coefficients corresponding to air temperature of the greenhouse and the following rows show the deviations belonging to the input variables: temperature, radiation, wind velocity and relative humidity quantified outside of the greenhouse.

$$M_{sd} = \begin{bmatrix} 0 & 0.0081 & 0.0079 \\ 0 & 0.0114 & 0.0177 & 0 \\ 0 & 0.0000 & 0.0000 & 0 \\ 0 & 0.0015 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 \end{bmatrix} \quad (12)$$

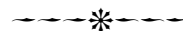
The matrix M_{sd} (12) shows that the estimation of the coefficients of the ARX model are determined with accuracy and those associated with the temperature outside the greenhouse present greater uncertainty.

When an ARX model is used as part of an algorithm of control of the air temperature of a greenhouse to predict scenarios, it is important that in its obtainment process the optimum size of data for estimation and validation is determined. According to our results, up to nearly 23% of the data not used to generate the structure can be predicted without recalibration of the model. Otherwise, as demonstrated by Uchida-Frausto et al. (2003), it will be necessary to recalibrate the model periodically or use recursive algorithms to determine new structures when there are new scenarios (Ljung, 1999; Aguado-Behar and Martínez-Iranzo, 2003). The ARX models were similar with and without implementation of closing and opening of vents, both in their structure and in their predictive quality. This suggests that natural ventilation is insufficient for an appropriate regulation of the temperature inside the greenhouse.

CONCLUSIONS

A methodology was proposed and evaluated to generate ARX models of the air temperature inside a greenhouse from the external variables of temperature, solar radiation, relative humidity and wind velocity. The proposed procedure makes it possible to obtain ARX models with a fit of more than 70% over the group of data of estimation and validation. The models also satisfy the residual analysis and the estimated coefficients are accurate.

—End of the English version—



modelos ARX con un ajuste mayor a 70% sobre el conjunto de datos de estimación y validación. Los modelos satisfacen también el análisis residual y los coeficientes estimados son precisos.

LITERATURA CITADA

- Aguado-Behar, A., y M. Martínez-Iranzo. 2003. Identificación y Control Adaptativo. Pearson Educación. Madrid. España. 285 p.
- Boaventura-Cunha, J., A. E. B. Ruano, and C. Couto. 1996. Identification of greenhouse climate dynamic models. Sixth International Conference on computers in agriculture. Cancún, México. pp: 161-171.
- Bot, G. P. A. 1983. Greenhouse climate: from physical processes to a dynamic model. PhD Thesis. Wageningen. The Netherlands. 240 p.
- Costa, P., and G. Giacomelli. 2005. Los planes del éxito, agricultura protegida: productividad basada en el nivel tecnológico. *Productores de Hortalizas* (Febrero 2005): 48-52.
- Deltour, J., D. de Halleux, J. Nijskens, S. Coutisse, and A. Nisen. 1985. Dynamic modelling of heat and mass transfer in greenhouses. *Acta Horticulturae* 174: 119-126.
- Kirk, D. 2004. *Optimal Control Theory. An Introduction*. Dover Publications INC. New York, USA. 452 p.
- López, C. I. L., B. vanThoor, y A. Rojano A. 2004. Modelos dinámicos mecanicista y empírico para predicción de la temperatura en un invernadero Mexicano. VI Congreso Latinoamericano y del Caribe de Ingeniería Agrícola. Noviembre 22-24. San José Costa Rica. 14 p.
- Ljung, L. 1999. *System Identification. Theory for the User*. Prentice Hall, New Jersey, USA. 609 p.
- Ljung L. 1988-2005. *System Identification Toolbox for Use with MATLAB*. The Matworks INC. Massachusetts, USA. 408 p.
- Ljung, L., and T. Glad. 1994. *Modeling of Dynamic Systems*. Prentice Hall. New Jersey, USA. 361 p.
- Nelles, O. 2001. *Nonlinear System Identification*. Springer. Berlin, Germany. 785 p.
- Uchida-Frausto, H., J. G. Pieters, and J.M. Deltour. 2003. Modelling greenhouse temperature by means of auto regressive models. *Biosystems Eng.* 84(2): 147-157.
- Wang, S., and T. Boulard. 2000. Predicting the microclimate in a naturally ventilated plastic house in a Mediterranean climate. *J. Agric. Eng. Res.* 75: 27-38.