SIMULTANEOUS THREE CLIMATIC VARIABLE PREDICTIONS IN A GREENHOUSE

PREDICCIÓN SIMULTÁNEA DE TRES VARIABLES CLIMÁTICAS EN UN INVERNADERO

Raquel Salazar Moreno[1*]; Abraham Rojano Aguilar[1]; Uwe Schmidt[2]; Christian Huber[2]

[1]**Universidad Autónoma Chapingo. Carretera México-Texcoco, km 38.5, Chapingo, Estado de México. C. P. 56230, MÉXICO.**
[2]**Institute for Horticultural Sciences, Humboldt-University Berlin. Lentzeallee 55 - 14195 Berlin. GERMANY**
Corre-e: raquels85@yahoo.com.mx (*Autor responsable)

ABSTRACT

The complexity of the greenhouse crop production, together with the environmental variables interaction, has encouraged the development of various models to predict and simulate variables in order to manage the greenhouse in an efficient way. One of the models that has been used successfully in prediction is the Artificial Neural Networks (ANN). The purpose of this paper was to predict the greenhouse inside conditions based on the outside environmental conditions. A data set of 19,960 values was used from the experimental greenhouse naturally ventilated, without CO_2 enrichment, at Humboldt University of Berlin, from August 16 to October 24, 2007 and include the external variables: solar radiation, air temperature, wind velocity, relative humidity, and carbon dioxide concentration, heat transfer by heating system and ventilation opening; and the internal variables: air temperature, relative humidity and CO_2 concentration. A three layer artificial neural network was trained and tested and validated using back conjugate gradient back propagation algorithm with a hyperbolic tangent function and momentum algorithm. The predicted values obtained from the ANN model were close to the measured values. These results showed that ANN model learnt the behavior and interactions between all variables. The aid of the ANN model developed was the simultaneous prediction of temperature, relative humidity and CO_2 concentration inside of the greenhouse, which will be helpful in checking the accuracy of sensor readings.

Additional key words: greenhouse, neural networks, prediction, environmental conditions.

RESUMEN

La complejidad de la producción en invernadero, junto con la interacción de las variables ambientales, ha alentado el desarrollo de varios modelos para predecir y simular variables con el propósito de manejar el invernadero en una forma eficiente. Uno de los modelos que sido usado con éxito en la predicción es el de Redes Neuronales Artificiales (RNA). El propósito de este trabajo fue predecir las condiciones internas de un invernadero con base en las condiciones externas ambientales. Se usó un conjunto de datos de 19,960 valores obtenidos entre el 16 de agosto y el 24 de octubre de 2007 en el invernadero experimental de la Universidad de Berlín, el cual era ventilado naturalmente y no estaba enriquecido en CO_2. Las variables externas incluidas fueron radiación solar, temperatura del aire, velocidad del viento, humedad relativa, concentración de dióxido de carbono, transferencia de calor por sistema de calefacción y abertura de ventilación, en tanto que las variables internas consideradas incluyeron temperatura del aire, humedad relativa y concentración de CO_2. Se desarrolló y evaluó una RNA de tres estratos y se validó por medio de un algoritmo de propagación de gradiente conjugado con una función tangencial hiperbólica y un algoritmo de momentum. Los valores pronosticados obtenidos con el modelo RNA fueron semejantes a los valores medidos. Estos resultados mostraron que el modelo RNA aprendió el comportamiento y las interacciones entre todas las variables. El apoyo obtenido con el modelo RNA desarrollado fue la predicción simultánea de la temperatura, la humedad relativa y la concentración de CO_2 en el interior del invernadero, lo cual ayudará en la verificación de la precisión de las lecturas de los sensores.

Palabras clave adicionales: invernadero, redes neuronales, predicción, condiciones ambientales.

Recibido: 31 de mayo de 2009
Aceptado:24 de junio de 2009

INTRODUCCIÓN

Some of the most critical environmental parameters affecting plant growth in the greenhouse are air temperature, relative humidity and carbon dioxide (CO_2). The way environment affects plant growth is not necessarily straightforward and the effect of one parameter is moderated by the others (Stanghellini and van Meurs, 1992). Air temperature is the factor of primary importance to most growers. Regulating temperature has a direct influence on relative humidity and CO_2 levels. (Portree, 1996). However, relative humidity of the air inside the greenhouse is also determined by the temperature of the inside boundary layers. Another environmental factor that is necessary for growth is carbon dioxide. It is required for photosynthesis; a carbon dioxide deficiency can be a limiting factor to growth. For instance, if moist air within the greenhouse is not removed, high relative humidity (greater than 70 %) will cause excessive condensation and problems with leaf mold infection of tomatoes. Regulating these environmental factors to the benefit of the greenhouse crop is part of good greenhouse management. The quantity and quality in greenhouse crop production can be realized only if the growth factors are managed so they are available to the plant at the right time with desired amounts (Portree, 1996).

Greenhouse climate and crop response models have been used successfully to predict the greenhouse microclimate and crop responses. Therefore, there has been increasing interest in the direct use of models in environmental control. The potential benefit has been identified in theoretical analyses and demonstrated in controlled experiments but there has been less progress in the practical realization of optimal control. This in part reflects the fact that research is still revealing shortcomings in the quantitative understanding of some processes.

Artificial Neural Network (ANN) technology is a form of artificial intelligence that “learns” by processing representative data patterns through its internal architecture. ANN technology, because of its empirical nature, is sometimes erroneously referred to as an “advanced” type of regression analysis. What distinguishes ANN technology from regression is Kolmogorov’s Theorem (Hecht-Nielsen, 1987), which guarantees that any continuous function can be represented by a three layer feed-forward neural network with ‘n’ elements in the input layer, ‘2n+1’ elements in the hidden layer, and ‘m’ elements in the output layer. In addition, unlike regression, which treats all output variables independent of each other, the presence of common arcs in the ANN architecture allows it to identify important inter-relationships that may exist between output variables.

Seginer (1997) used a Neural Network model to optimize environmental control and it was found that an NN model relieves the difficulty in developing a physical model which relates air ventilation rate to outside climate conditions. Linker *et al*. (1998) extended the use of Neural Network models by not only fitting models to experimental data, but also by using them to optimize the greenhouse operation. A data based greenhouse model was developed and incorporated in an optimization scheme to compute optimal temperature and CO_2 concentration set points as a function of the changing weather. This study indicated that the models used not only fitted the data well but also produced reasonable optimization results to control the greenhouse environment.

The objective of this work was to investigate the use of Artificial Neural Networks as a prediction tool to control the greenhouse environment. The benefit of this model is the simultaneous prediction of the three important environmental variables: temperature, relative humidity and CO_2 concentration. The model can be used as a predicting tool for inside environmental conditions based on the outside environmental conditions which will provide the potential to operate the greenhouse climate in an optimal way.

MATERIALS Y METHODS

Data were collected at the experimental Venlo-style glass greenhouse at Humboldt University of Berlin from August 16 to October 24, 2007. The greenhouse was 9.6 m width, 12 m long, 4.1 m height until ridge; the ratio of the roof vent opening to roof area was 50 %, floor space was 115.2 m^2 with 128 tomato plants (8´18). This greenhouse was the control greenhouse for different experiments at Institute for Horticultural Sciences at Berlin. Only natural ventilation was applied when inside temperature was above 27^{o}C and heating system was applied for temperatures below 18 ^{o}C.

Sensors for all the inside environmental variables were located at the center of the greenhouse at a height of 1.5 m. The outside air temperature was between 2.5 ^{o}C and 22.8 ^{o}C; relative humidity varied from 40 % to 97.3 % in the study period.

A multi-layered perceptron (MLP) network was used in this work and is shown in Figure 1. This typically has three layers. Nodes in the input layer represent the input variables used for predicting the output variables of interest. The variables used in the input layer were: outside average air temperature (T, ^{o}C), relative humidity (RH, %), external CO_2 concentration (CO_2, ppm), wind velocity (WV), solar radiation (SR), heat transfer by heating system (HT) and vents opening left (VeL) and right (VeR). The variables used in the output layer were: inside air temperature, inside relative humidity and inside CO_2 concentration. Each ANN layer consists of individual nodes (elements), and the nodes were interconnected across layers by special (usually non-rational) non-linear transfer functions, expressed in terms of the nodal input variables and connection weights.

Heat transfer by heating system was calculated as follows:

$$HT = U\ A\,(T_{prom} - T_{\text{int}}) \tag{1}$$

Where *HT* is heat transfer by heating system, *U* is the heat transfer coefficient for steel (25 W/m^{2}K), T_{int} is air temperature inside of the greenhouse, T_{prom} was taken as $\left[(H_{in} + H_{out})/2\right]$,

Table 1. Sample of measured and predicted values for temperature, relative humidity and CO_2 for October 25, 2007.

Time	Tm	Tp	RHm	RHp	CO_2m	CO_2p	Time	Tm	Tp	RHm	RHp	CO_2m	CO_2p
00:00	19.4	19.2	79.8	75.5	446	450.4	13:00	19.5	20.1	89.4	82	416	421.9
01:00	19.6	19.6	80.4	74.9	446	455.6	14:00	19.2	18	89	79.3	416	473
02:00	19.2	18.7	81	74.4	477	456.3	15:00	18.9	18.8	88.7	78.4	416	433.1
03:00	19	18.9	81	75.3	475	447.2	16:00	20.3	20.8	83.4	77.7	406	394.3
04:00	19.5	19.5	76.7	72.7	464	446.4	17:00	19.9	20.8	83.5	76.3	416	395.7
05:00	19.5	18.4	78	74.5	477	453.6	18:00	19.6	20.3	83	75	416	407.2
06:00	19.8	19.8	78.1	74.8	422	440.8	19:00	19.2	19.8	82.4	74.3	416	416.3
07:00	20	20.8	81.9	75.3	459	430.4	20:00	18.7	18.5	83.8	74.3	398	430.6
08:00	19.7	20.4	83.7	75.7	447	439.5	21:00	19.7	19.7	66.6	66.8	416	408.6
09:00	20	20.5	85.3	77.8	416	438.6	22:00	18.9	18.8	80.4	75.2	447	426
10:00	19.3	19	86.3	75.7	416	470.2	23:00	19.6	19.7	78.7	73.8	447	411
11:00	20.8	21.1	84	79.7	416	433.9	23:55	18.8	18.9	82.1	75.4	477	427.1
12:00	19.8	20.5	87.8	81.6	416	434.4	MSE	MSEtemp=0.357		MSERH=65.05		MSECO2=2135.34	

being H_{in} is water temperature in the pipe entering the greenhouse and H_{out} is water temperature in the pipe leaving the greenhouse. Likewise, A is total heating area (B+R; 46.77 m^2), where B is the heating pipe area in the bottom (26.58 m^2) and R the heating pipe area in the roof (17.18 m^2).

In our study, a hyperbolic function was used as shown by Equation (2) where variable z is expressed in terms of the input variables, which when multiplied by the connection weights w_{ik} generated the hidden nodes described in Equation (3),

$$f(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{2}$$

$$\begin{aligned} h_1 &= w_{11} f(x_1) + ... + w_{41} f(x_4),.... \\ h_k &= w_{1k} f(x_1) + .. + w_{4k} f(x_4) \end{aligned} \tag{3}$$

Again, the application of the hyperbolic function to hidden nodes and the connection weights that relates hidden nodes with the output nodes $\overline{w_{ik}}$ generates Equation (4), but now a bias unit w_2 was added for numerical stability.

$$\begin{aligned} y_1 &= \overline{w}_{11} f(h_1) + .. + \overline{w}_{k1} f(h_k) + w_1; \\ y_2 &= \overline{w}_{12} f(h_1) + .. + \overline{w}_{k2} f(h_k) + w_2 \end{aligned} \tag{4}$$

Equation (5) generates a multivariate nonlinear optimization problem which is solved using a conjugate gradient back-propagation method where the decision variables are the weights.

$$Min\ E = \frac{1}{Tn} \sum_{t=1}^{T} \sum_{j=1}^{n} (y_j^t(data) - y_j^t(ANN\ results))^2 \tag{5}$$

where E is mean square error, T is total number of data, n is number of output variables, y_j refers output variables, x_i represents input variables, h_k points out hidden nodes, w_{ik} expresses weights between input and hidden nodes and $\overline{W_{ik}}$ weights between hidden and output nodes.

The above procedure was applied for different number of nodes in the hidden layer until a minimum mean square error (E) was achieved, thereby producing the optimal architecture for the neural network. Intermittent ANN “validation” was performed during training to avoid problems of over-training. That is, periodically during training, the network was validated with a different data set, and this process was repeated until the validation error began to increase. At this point, ANN training was finished and the ANN was then tested with a third data set to evaluate how effectively it had learned to generalize system behavior.

A total number of 19,960 data points were available. Among them, 50 % were used for training, 25 % for verification and 25 % for testing (Swingler, 2001). During training, data patterns were processed through the ANN using Equations (3) and (4), and the connection weights were adaptively adjusted until a minimum acceptable error between the ANN predicted output and the actual output in Equation (5) was achieved. It was identified at that point that the ANN had “learned” to predict the system behavior of interest (i.e. values of output variables) in response to the values of the input variables.

RESULTS AND DISCUSSION

After training, validation and testing, the optimal ANN architecture was found (Figure 1), which has only one hidden layer with nine nodes. The real and predicted variables using the 19,960 patterns data set are shown in Figure 2, where it can

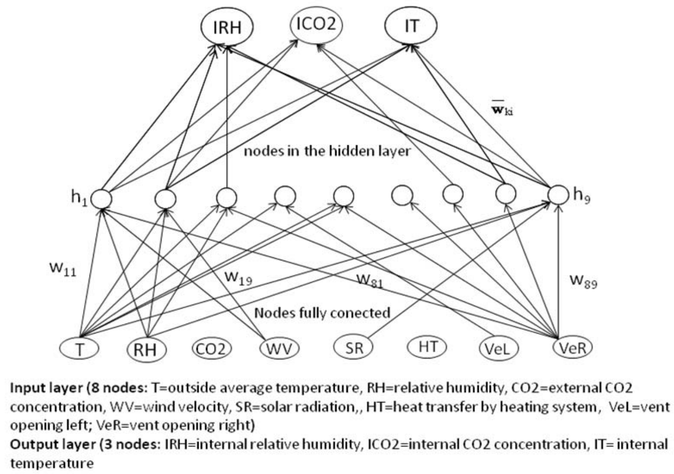

Figure 1. Artificial Neural Network architecture.

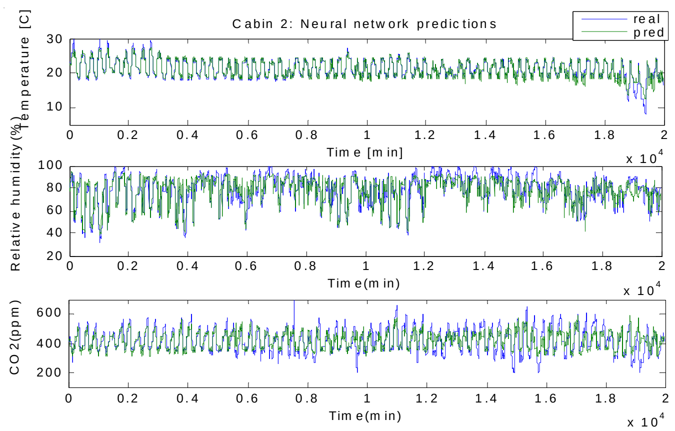

Figure 2. Real and predicted variables using the 19,960 patterns data set.

be seen that measured and predicted data followed the same tendency. Also, a linear regression was performed between measured data (*T*, *RH* and CO_2 concentration) and predicted variables (output variables from ANN) for the complete data set (19960 patterns). The regression coefficients were as good as 0.96 and 0.91, for *T* and *RH*, respectively, and less so as 0.72 for CO_2.

To evaluate the performance of the ANN model developed for prediction purposes in the greenhouse, a 1000 new data set from October 24, 2007 to October 28, 2007 was used to feed the ANN. Figure 3 has the comparison between measured and predicted variables over time. ANN predictions were relatively close to measured values. The ANN model was able to predict the three internal variables (air temperature, relative humidity and CO_2) simultaneously, given a set of input variables. Table 1 shows a sample of the measured and predicted values for one day. The mean square error (MSE) in the three cases was calculated, the prediction performance for air temperature (*T*) was the best (MSEtemp=0.357). For the case of relative humidity (*RH*) the predicted values were acceptable (MSE_{RH}=65). The worst case was for CO_2 prediction with an MSE_{CO2}=2135 and this big value was due to the drop in the CO_2 measurements occurred on October 25 between 6:05–6:50 am. The biggest difference between measured and predicted values was 142 at 6:30 am. Figure 4 depicts the measured and predicted values for this new data set. Sensors were very sensible to disturbances inside of the greenhouse; therefore we need to take out any outliers in the data used for ANN predictions. For the three cases: temperature, relative humidity and CO_2 it was not able to find a pattern for under and upper predictions because sometimes they occurred in the day time and in other cases during the night time.

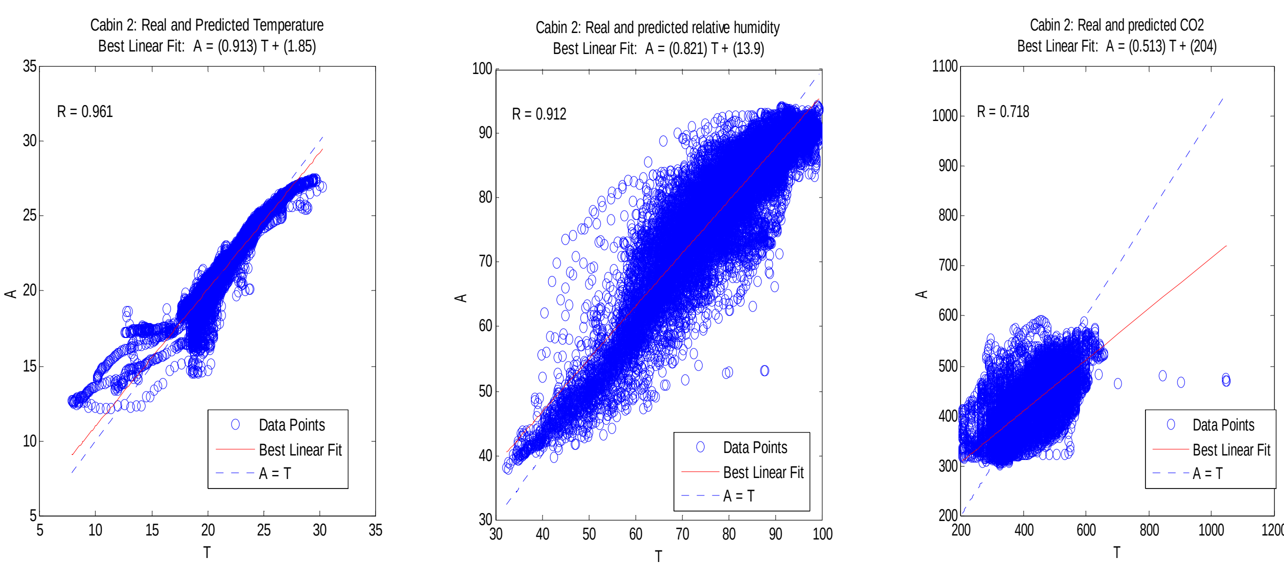

Figure 3. Linear regression between real and predicted variables using different data set for input in ANN.

CONCLUSIONS

A Neural Network model was successfully developed for climate prediction purposes in an experimental greenhouse naturally ventilated with tomato crop, located at Humboldt University of Berlin. The ANN was able to predict temperature and relative humidity with a good prediction performance; however CO_2 predictions were not accurate estimated because of sensor readings outliers. This work showed the benefits of using ANN as a predictive tool, specifically the aid of this study is the simultaneous prediction of three important environmental variables inside of the greenhouse: temperature, relative humidity and CO_2 concentration, based on the outside environmental variables, which could be helpful when there are a lot of disturbances inside of the greenhouse so it is better to use the prediction than the sensors readings to control greenhouse environment.

LITERATURE CITED

Hecht-Nielsen R. 1987. Counterpropagation networks: Proceedings of the International Conference on Neural Networks, IEEE Press, New York v. II. pp: 19-31.

Linker R.; Seginer I.; Gutman P.O. 1998. Optimal CO_2 control in a greenhouse modelled with neural networks. Computers and Electronics in Agriculture 19: 289-310.

Portree J. 1996. Greenhouse vegetable production guide for commercial growers. Province of British Columbia Ministry of Agriculture, Fisheries and Food.

Seginer I. 1997. Some artificial neural network applications to greenhouse environmental control. Computers and Electronics in Agriculture 18: 167-186.

Stanghellini C.; Van Meurs W.Th.M. 1992. Environmental control of greenhouse crop transpiration. Journal of Agricultural Engineering Resources 51: 297-311.

Swingler K. 2001. Applying Neural Networks: A Practical Guide. Academic Press. Third Edition. San Diego, California. pp: 110-113.