Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT
Neurocomputing 56 (2004) 423 – 428
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Letters

Modelling greenhouse temperature using system identification by means of neural networks

Hugo Uchida Frausto*, Jan G. Pieters

Biosystems Engineering, Ghent University, Coupure Links 653, B-9000 Gent, Belgium

Abstract

An NNARX system is proposed for modelling the internal greenhouse temperature as a function of outside air temperature and humidity, global solar radiation and sky cloudiness. The models show a good performance over a complete year using only two training periods, 1 week in winter and one in September. Finding the balance between the number of neurons in the hidden layer of the NNARX system and the number of iterations for model training is found to play an important role in obtaining this good performance.

Keywords: Auto-regressive model; Recurrent neural network; Backpropagation

1. Introduction

The greenhouse microclimate provides the plants with good environmental conditions for growing, one of which is the inside air temperature. This temperature is the result of complex and interactive heat and mass exchanges between the inside air and the several elements of the greenhouse (construction, vegetation, etc.) and the outside boundaries (outside air, sky, solar radiation). Over the last decades, numerous deterministic greenhouse climate models have been built. In general, these models have a high degree of complexity with lots of parameters that have to be determined by calibration. In contrast to deterministic models, black box models do not suffer from the need to determine appropriate values for lots of parameters. These models cannot be used to investigate the internal functioning of the system, but they can be very helpful for climate control purposes, certainly when combined with automated parameter

* Corresponding author. Tel.: +32-9-264-6190; fax: +32-9-264-6235.
E-mail addresses: hugo.uchida@ugent.be (H. Uchida Frausto), jan.pieters@ugent.be (J.G. Pieters).

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2003.08.001

identification techniques. In a previous paper [6], the authors presented applications of the linear auto-regressive models ARX (auto-regressive model with external input) and ARMAX (auto-regressive moving average model with external input) for the simulation of the inside air temperature in greenhouses. The major drawback of such models, however, is their inability to perform well over long periods, because of the lack of adaptation. This implies that frequent retuning is necessary, which greatly reduces the practical use of such models in greenhouse climate control. In this paper it will be investigated to what extent combining an ARX model with Neural Network architectures, resulting in a so-called NNARX, can improve model performance as a basis for greenhouse climate control.

2. Methods and procedures

2.1. Data sets

The inside air temperature data used for training and validation of the model were not measured, but determined via simulation using the extensively investigated and validated deterministic Gembloux Dynamic Greenhouse Climate Model (GDGCM). More information on this model can be found in [5], while the simulation circumstances and the characteristics of the unheated, ventilated greenhouse, used in the simulations, can be found in [6]. The GDGCM was fed with data from the Belgian typical reference year [1], namely the outside air temperature, the outside air relative humidity, the outside global solar radiation flux density, and the cloudiness of the sky. In this way, data for the inside air temperature and the outside climate were obtained for every 5 min of a complete year.

2.2. Model description

The model structure was a multi-layer perceptron network with only one hidden layer containing neurons with a hyperbolic tangential activation function and an output layer with one neuron with a linear activation function. The input to this structure was the vector containing the regressors of an ARX model. The governing overall equation, giving the inside air temperature T_i, can then be written as

$$\hat{T}_\mathrm{i}(w,W) = \sum_{j=0}^{q} W_j \tanh_j\left(\sum_{l=1}^{m} w_{jl}\varphi(t)_l + w_{j0}\right) + W_0, \tag{1}$$

where w represents the $n_h \times (n_i + 1)$ matrix that contains the weights from the inputs to the hidden layer and W represents the vector that contains the $n_i + 1$ weights (where the 1 is due to the bias) from the hidden to the output layer and with $\varphi(t)_l$ the vector that contains the regressors of the ARX model, q the number of internal neurons, and m the number of input variables. The idea is to select the regressors based on inspiration from linear system identification and then determine the best possible network architecture with the given regressors as inputs. More details on NNARX models can be found in [4].

For the determination of the number of neurons in the hidden layer, two different criteria were used: the sum of the number of inputs and the number of outputs or twice the square root thereof. Based on the results of [6], the inputs to the model were the four previous (5 min) samples of the four outside climate data and the inside air temperature. This implies that the number of neurons in the hidden layer was 20 or 8, according to both criteria, respectively. The model was implemented under MATLAB using a toolbox described in [3].

2.3. Model training and validation

The NN training method used was the backpropagation algorithm. Two different data sets were used for training: those of the first week of winter and those of the second week of September (late summer). These two periods were chosen because the former corresponds to the first week of the data set that is rather easy to simulate, while the latter period was found to be very difficult to model accurately using an ARX or ARMAX model without NN architecture. The number of iteration steps in the training was increased in steps of 100 until an acceptable goodness of fit was obtained. The goodness of fit was defined as

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

where y_m is the output of the NNARX model; y_o represents the original output data; s is the square mean error of modelled output versus original data (results of the GDGCM); and σ is the standard deviation of the original system output [2]. From Eq. (2), it is clear that the highest goodness of fit that can be obtained is 100%, while the lowest value is $-\infty\%$. For the validation of the model, the data of the other weeks of the year were used.

3. Results and discussion

Fig. 1 shows the original data versus the NNARX results for both the winter and September trained models. From this figure, it can be directly deduced that for both periods, using only eight neurons in the hidden layer of the NNARX resulted in an unacceptably poor performance. It is clear that in September, when ventilation is very important for climate control during hot days, the use of eight neurons resulted in a model which could not take into account the effect of ventilation, coming in action from 19°C on. When comparing the results for both training periods, no large differences in performance can be found on this figure. This model behaviour was also reflected by the values for the goodness of fit. When using eight neurons, even after 800 iterations the goodness of fit was lower than 40%, while the ones for the 20 neurons models

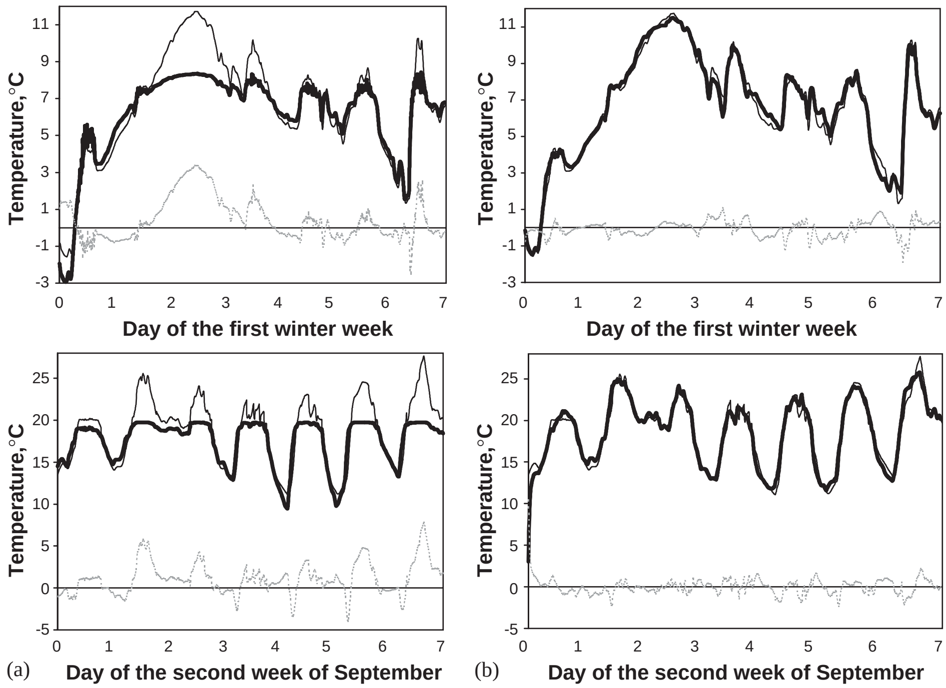

Fig. 1. Original data versus results of the NNARX model trained with data of the first winter week and the second week of September using (a) eight neurons in the hidden layer and (b) twenty neurons in the hidden layer; —, original; —, model; ...., residuals.

resulted in a goodness of fit of more than 75% for the winter and the September trained model after 500 and 200 iterations, respectively. Consequently, models with only eight neurons in the first hidden layer were not investigated any longer.

Results for the validation of the models trained during winter and during the first week of September are shown in Fig. 2. For a complete year, daily averages of the absolute values of the differences between simulated results and original data are shown. From this figure, it can be concluded that the models performed the best in the periods during which they were trained, as was to be expected. However, looking at the complete year, it is observed that the September trained model performed much better than the winter trained model. This is most likely to be explained by the fact that in September, ventilation plays an important role in climate control, while in winter its role is limited. As a result, training the model with September data enables the NNARX to take the non-linear ventilation effects into account. This also explains why the addition of a NN architecture to the ARX model allows much better performances to be obtained, with much less need to retune the models. As a matter of fact, it is possible to keep the daily average absolute simulation error smaller than about 1°C during most of the year by using the September trained model from day 80 to day 350 and the winter trained model during the other days.

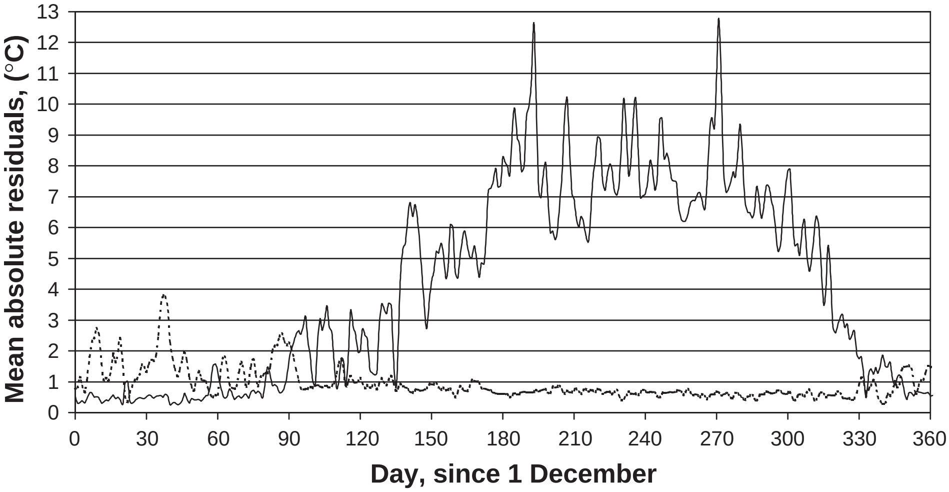

Fig. 2. Daily mean of the absolute value of the differences between original data and the results of the NNARX trained with data of (—) the first winter week and (.......) the second week of September using 20 neurons in the hidden layer for a complete year, starting on 1 December.

In general, it can be concluded that the NNARX system allows fairly reliable models to be built, which could be further investigated for inclusion in climate control. It should be kept in mind, however, that such models can only be applied as long as the conditions are similar to those of the training period. Under exceptional circumstances, such as snow storms or other situations which lead to a change in the thermal characteristics of the greenhouse, the model might give erroneous results. Use of such models for greenhouse climate control will thus necessarily imply the simultaneous inclusion of feedback mechanisms.

4. Conclusions

An NNARX system was proposed for modelling the internal greenhouse temperature. The models showed a good performance over long periods without the need of frequent retuning the parameters, as indicated by the simulation results and the goodness of fit. The number of neurons in the hidden layer of the NNARX system was found to play an important role in obtaining this good performance. Introduction of such models in climate control systems needs further investigation, amongst others to find a balance between the number of neurons in the hidden layer and the number of iterations in the training procedure.

Acknowledgements

Hugo Uchida Frausto thanks CONACyT (Consejo Nacional de Ciencia y Tecnologia, Mexico) for its financial support (contract 124965).

References

[1] R. Dogniaux, M. Lemoine, R. Sneyers, 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, 1978.
[2] L. Ljung, System Identification Toolbox User's Guide, The MathWorks Inc., Natick, 2000.
[3] M. Nørgaard, Neural network based system identification toolbox, Technical Report, 00-E-891, Department of Automation, Technical University of Denmark, 2000.
[4] M. Nørgaard, O. Ravn, N.K. Poulsen, L.K. Hansen, Neural Networks for Modelling and Control of Dynamic Systems, Springer, London, 2000.
[5] J.G. Pieters, J.M. Deltour, Performances of greenhouses with the presence of condensation on cladding materials, J. Agric. Eng. Res. 68 (2) (1997) 125–137.
[6] H. Uchida Frausto, J.G. Pieters, J.M. Deltour, Modelling greenhouse temperature by means of auto regressive models, Biosystems Eng. 84 (2) (2003) 147–157, doi: 10.1016/S1537-5110(02)00239-8.

Hugo Uchida Frausto was born in Mexico City on October 3rd, 1972. He graduated from the Monterrey Institute of Technology (ITESM) Campus Queretaro in 1995 with a Bachelor of Science in Electronic and Communications Engineering. He obtained his M.Sc. in Automatic Control and Automation from the Universidad Autonoma de Queretaro (UAQ) in 2000. He entered in the Ph.D. Program at Ghent University in the Faculty of Agricultural and Applied Biological Sciences in 2000 where he is currently researcher. His research interests include: greenhouse climate modelling and control, renewable energy resources, digital image processing by mathematical morphology, machine vision, neuronal coding.

Jan Pieters graduated from the Ghent University in 1991 with a M.Sc. in Agricultural Engineering. In 1995 he obtained his Ph.D. in Applied Biological Sciences at the same university. In 1996–1997 he was a visiting scientist at the Centre for Environmental Biotechnology and the Civil and Environmental Engineering Department of the University of Tennessee–Knoxville (USA). From 1999 on, he is professor in Biosystems Engineering, with teaching and research activities in the field of agricultural, environmental and food engineering.