Hybrid MLP-RBF Model Structure for Short-Term Internal Temperature Prediction in Greenhouse Environments

Peter Eredics, Tadeusz P. Dobrowiecki
Budapest University of Technology and Economics, Department of Measurement and Information Systems, Budapest, Hungary
eredics@mit.bme.hu, dobrowiecki@mit.bme.hu

***Abstract*—A wide variety of greenhouse temperature models have been proposed in the literature in the previous years. This paper proposes a hybrid modeling method incorporating a multilayer perceptron neural network and a radial basis function neural network aimed to be more accurate on input regions not covered by training data. The results show that the proposed method has better performance compared to the original physical-neural hybrid model if the input values are not far from the input range of the values used for training.**

I. INTRODUCTION

A. Greenhouse modeling

Greenhouses are widely known structures in the field of vegetable and ornamental plant production. The transparent walls allow energy input from solar radiation while the warmed air is kept inside the house. Because of this greenhouses are ideal environments for plants requiring warmer temperatures than the outside climate.

Today large industrial greenhouses are all equipped with actuators aimed to regulate the internal environmental parameters: in most cases air temperature and humidity are the controlled parameters [1][2], but in modern high-tech greenhouses the CO_2 concentration; the amount of solar or artificial radiation; the control of fogging; the control of the irrigation, etc. is also used [3][4].

Control solutions designed for greenhouse systems have a wide spectrum. From the very simple set point computers available on the market to the very sophisticated control solutions presented in the literature there are several control methodologies applied to this field. Mahaman et al. proposed a rule-based expert system for greenhouse control and also for pest management [5]. Lafont et al. applied fuzzy methodology to the problem based on physical models [6]. Sigrimis et al. also used fuzzy rules to control irrigation in a greenhouse [7]. There is active research on greenhouse control from the optimal control perspective: Van Straten applied optimal control methods for traditional greenhouses [8], while Speetjens et al. extended the solution for the most modern energy saving Watergy greenhouse concept [9]. Neural network based modeling has a great tradition in the field of greenhouse modeling, both for prediction and for control purposes [10][11].

There are many more important greenhouse control approaches along the ones detailed above. The common property of almost all such control solution is the fact that a model of the system is needed for the control. The format and the requirements against the model highly varies with the different implementations, but it can be concluded, that greenhouse models play key role in solving control related problems.

II. THE HYBRID PHYSICAL-RBF GREENHOUSE MODEL

A. Physical Models vs. Black-Box Models

Physical model applications have a long history in the field of greenhouse modeling. If all physical quantity can be measured in the greenhouse and if both the model structure and its parameters are accurate, then a physical model can be a very accurate predictor. In theory based on the heat transfer equations and by measuring all necessary parameters in the greenhouse such a model could be crated [12]. Unfortunately in practice a model representing all physical relation in the thermal equations would be extremely complex. Furthermore there is no real chance to measure all necessary parameters with enough precision and without disturbing noise. These circumstances limit the applicability and accuracy of physical models in practice.

Black-box neural models are also widely applied for greenhouse modeling. In this case the model structure does not have to be determined in advance, the neural network learns the model structure from the data samples during the training process. This property ensures great flexibility in the application of black-box models, as such models are able to follow changes in the greenhouse itself (e.g. new plants arrive, new equipment is set up, etc.) In theory if the number of training samples is large enough, and the complexity of the network is not restrictive the black-box model can be very accurate. Unfortunately in practice the number of training samples is limited by the number of measurements available and the complexity of the network must be limited to avoid over fitting (learning the noise disturbing the measurements). These are the practical limitations of black-box models.

Comparing physical models and black-box models in the field of greenhouse modeling does not result in a clear winner. The precision of physical models is limited by their structure and the number of data available, while on the other hand a low complexity but accurately constructed physical model structure is likely to avoid large prediction errors even if it is not accurate in most of the data points. Physical models also have good performance on input regions not present in the training data, as some of the knowledge about the problem is initially built into the model structure at the time of the design.

Black-box models are just the opposite: as there is no built-in knowledge in the model structure, black box models can be very accurate on regions of the input where training data is available. But on the other hand input regions not represented during the training process can result in very large errors.

B. *Hybrid Physical-Neural Modeling*

The hybrid physical-RBF model was first applied to greenhouse modeling problems by Linker et al [13]. The goal was to mix the good attributes of physical models and black-box neural models in a hybrid structure. Hybrid modeling has two possible implementations as presented in Fig. 1. [14]

Serial hybrid modeling (Fig.1A) means that the black-box component of the system is responsible for learning the parameter of the physical model. This way the structure of the physical model is fixed, while the parameters are approximated by a black-box model with high nonlinearity if necessary. The drawback of this approach is the fact, that the output error of the model must be transformed back through the inverse physical model of the system in order to train the black-box model.

In case of parallel hybrid modeling (Fig.1B) the black-box model is used to learn the prediction error of the physical component. This way the black-box component of the system works as a correction of the physical model when necessary. Linker et al. used parallel hybrid modeling in [13]. The physical model (see equation (1) where T_i is the internal temperature of the greenhouse; T_o is the outside temperature; S_o is the outdoor solar radiation; H is the heating flux; Q is the ventilation flux; U is the overall heat transfer coefficient of the cover; σ and c are the air density and specific heat; λ is the solar heating efficiency parameter) had intentionally very low complexity to ensure acceptable predictions over the whole input space.

$$T_i = \frac{T_o+(\lambda S_o+H)}{(\,pcQ+U\,)} \tag{1}$$

A radial basis function (RBF) neural network was selected for the black-box component as this network only provides prediction where training data was available – on other segments of the input space the RBF network returns 0 by default [15]. Fig. 2. shows the model structure of a RBF neural network. The inputs of the network are shown on the left in Fig. 2. The radial basis functions are placed in the single hidden layer of the network. Each radial basis function is a real-valued function of the input space, where the value of the function only depends on the distance of the input from the so called center of the function. The most often used basis function is the Gaussian:

$$\phi_n(r) = e^{-(\epsilon r)^2}. \tag{2}$$

The output layer of the network is linear, as described by

$$y(x) = \sum_{i=1}^{n} w_i \phi(\|x - x_i\|). \tag{3}$$

The result of the finite basis size of the functions is the fact that RBF networks can learn local patterns. Compared to other neural network architectures this property makes the RBF ideal candidate for hybrid modeling, as this behavior makes it possible to learn the output error on one part of the input space without changing the previously learned errors on other parts of the input space.

C. *Introducing the Hybrid MLP-RBF Model*

The main aim of this paper is to improve the accuracy of the hybrid greenhouse model on previously unseen input samples. The proposed method is based on the hybrid Physical-RBF model by Linker et al. but the physical model has been replaced by a simple Multilayer Perceptron (MLP) neural network. The network structure of the MLP is intentionally kept simple (the number of neurons in the hidden layer is small) to avoid exact fitting of the known samples, thus mimicking the loose fitting of the physical model.

Application of an MLP neural network instead of a physical model has two main advantages. First the model structure does not have to be specified in advance, as the structure is created by the neural network learning process. In practice physical models should be created one by one for all possible greenhouse actuator setup, considering all possible inputs and relations. In case of a black-box MLP model the choice of input variables and the number of hidden layer neurons are the only decision to make, the model structure is generated dynamically. The second advantage of the MLP model compared to the physical (in most cases linear) model is its capability of nonlinear modeling. This attribute ensures that the MLP network can provide better approximation than a fixed structure physical model.

On the other hand the application of an MLP model instead of a physical model introduces some limitations as well: While simple physical models are highly reliable even far from the working point of the system, an MLP model cannot give such guarantees far from its training samples.

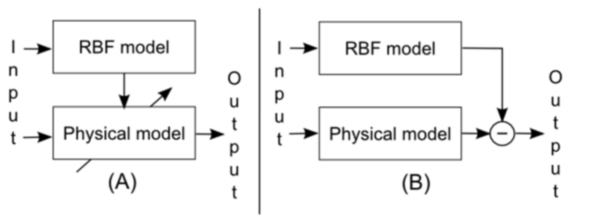

Figure 1. Possible hybrid modeling implementations: A) serial; B) parallel;

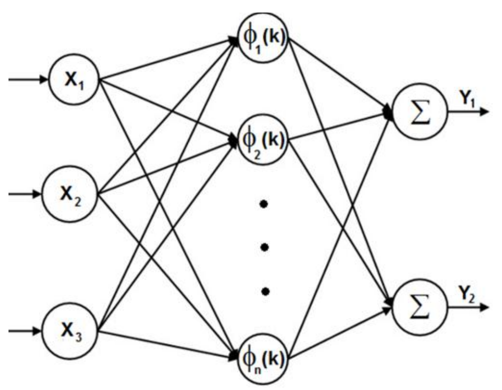

Figure 2. The radial basis function neural network structure

III. Measurement and Test Scenario

The comparison of the original Physical-RBF method and the proposed MLP-RBF method was done with measurement data from a small-scale laboratory greenhouse [16]. The laboratory environment was chosen as data source as it eliminates several weather related disturbance of real greenhouses, thus results are easier to compare.

A 2052 minutes long measurement sequence was recorded with one minute time resolution. The external infrared lamps (simulating the solar radiation) were controlled between 0% and 100% in 10% steps in a “30 minutes on – 30 minutes off” pattern. This measurement sequence was repeated two times. Fig. 3. shows the lamp control signal and the internal temperature of the small scale greenhouse. During the measurements external and internal temperature and the lamp control signal was recorded at every time step.

The data was used in a test setup illustrated by Fig. 3. The input data was split into a train set (90%) and a test set (10%). These train and test sets were used for both the original and for the proposed method.

The implemented physical model was a simplification of Eq. 1. where the unused actuator effects (e.g. heating) were left out. The MLP model had the following inputs: the internal temperature, the internal-external temperature difference and the solar radiation. The number of neurons in the hidden layer of the MLP was intentionally chosen to be low (it was set to 3), to avoid good fitting in order to have generalization performance similar to the physical model.

Both methods were used for one step ahead prediction of the temperature change inside the greenhouse. For both methods the data sets have been normalized to have all variables in the range of [0 1]. Table I. summarizes the normalization parameters in the form of offset and scaling.

IV. Results

A. General Prediction Performance

Both methods have been simulated 10 times with a random training and testing subset of the measured small-scale greenhouse data. The mean average errors (MAE) on the test sets were as follows.

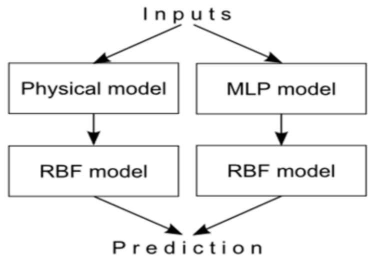

Figure 3. The test implementation of the two methods – the RBF correction model had the same implementation for both methods to eliminate disturbances in the results originating from that module

TABLE I.
NORMALIZATION PARAMETERS OF THE INPUT AND OUTPUT VARIABLES

Variable	Offset	Scaling
Internal temperature (°C)	22.5	8.75
Internal-external temperature difference (°C)	- 6.625	6.5
Solar radiation (%)	0	100
Output: predicted temperature change (°C)	- 0.25	0.5

The physical model had the highest MAE with a 0.0465 value without any RBF correction applied. This result is not a surprise as the model structure was intentionally kept simple to ensure good generalization even for input regions with very few or no training samples. After applying the RBF correction (the method of Linker et al.) has 0.0429 mean average error. This result is 10.4% better compared to the pure physical model.

If the simple MLP model is applied to the data alone (with 3 neurons in the hidden layer), the model has a MAE value of 0.0448. After applying the RBF correction module the MAE value drops to 0.0428. This is a 4.6% correction provided by the RBF neural network.

Comparing the results of the hybrid Physical-RBF and

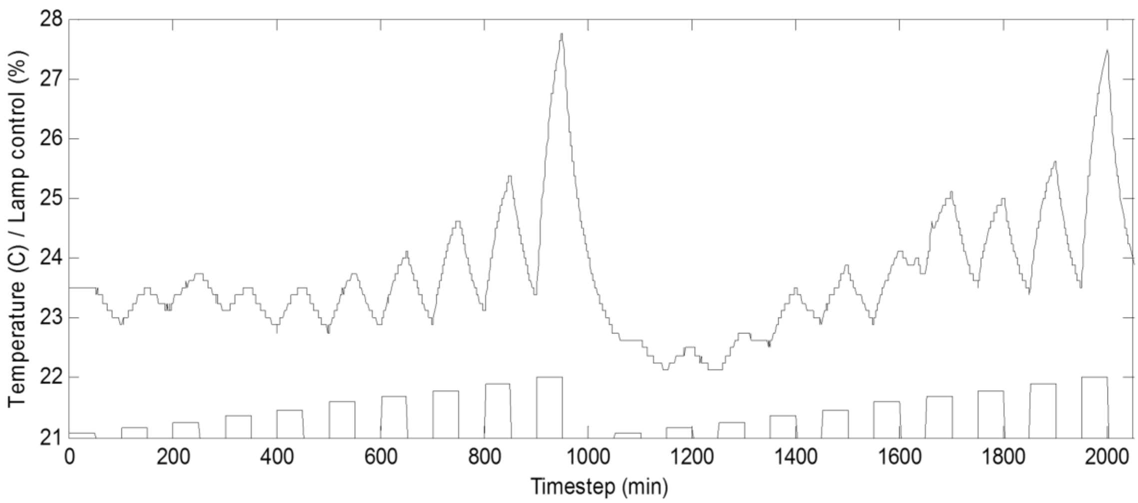

Figure 3. Internal temperature and lamp control signal recorded in the small scale experimental greenhouse

the MLP-RBF model structure there is no significant difference in the result of the two models, both have approximately the same predictive power.

B. Prediction Performance in Regions without Training Data

The aim of the proposed method was to increase prediction performance on input ranges without training data available. To measure this performance the training and testing data set was filtered systematically: several experiments were done by selecting a disjoint set of train and test samples. The following results were obtained by training both models with samples where the solar radiation was less or equal to 30% or equal to 50% or more. The test samples where all the measurements where the radiation was set to 40%.

In this scenario the Physical-RBF model had a MAE value of 0.0559, while the MLP-RBF model had 0.0434 prediction error. Both values are higher than the ones in the previous section, but it was expected as the train and the test set was selected from disjoint input regions. The comparison of the two methods shows that the MLP-RBF model has significantly better performance (23% gain in this scenario) on such input regions. The reason for that is the better approximation capability of the MLP model based on its nonlinear modeling capability, its flexible structure and the higher number of parameters to be tuned.

It has to be noted that in an extreme case when the missing values are selected from the beginning or the end of the variable value range (e.g. 0% or 100% radiation value), the performance of the MLP-RBF model drops significantly. In case of such out of the range values the larger the distance from the known data samples, the worse the MLP-RBF model performance is, so in such cases the application of the original method might be still beneficial.

V. Conclusion

The proposed MLP-RBF hybrid model structure was aimed to increase internal temperature prediction performance in case if some input ranges are not covered with training samples. The method is based on the Physical-RBF model structure of Linker et al. but the simple physical model of the greenhouse was replaced with a low-complexity MLP neural network.

Both methods were tested on real measurement data originating from a small-scale laboratory greenhouse. The general prediction performance of the two methods were proven to be equal: both methods had better accuracy compared to the pure physical or the pure MLP models, but the overall mean absolute error showed no significant differences. In case of input ranges not covered by training data the proposed MLP-RBF model was proven to be more accurate. In special cases when the missing input data was originating from the beginning or the end of the range of the affected parameter the results were not such appealing: the more distant the input data is from the known input values, the larger error the MLP-RBF model had. In cases when such new data samples can be expected the original Physical-RBF model has better performance. If such outlier values are not likely the proposed MLP-RBF model was proved to be 4-28% more accurate than the original implementation.

Acknowledgment

This research was partially supported by the ARTEMIS JU and the Hungarian National Development Agency (NFÜ) in frame of the R3-COP (Robust & Safe Mobile Co-operative Systems) project.

The results discussed above are partly supported by the grant TÁMOP - 4.2.2.B-10/1--2010-0009.

References

[1] Caponetto, R., L. Fortuna, G. Nunnari, L. Occhipinti, and M.G. Xibilia. Soft Computing for Greenhouse Climate Control. IEEE Transactions On Fuzzy Systems, 8, no. 6, pp. 753 –760, December 2000.

[2] Zhang, Zhuhong. Multiobjective Optimization Immune Algorithm in Dynamic Environments and Its Application to Greenhouse Control. Applied Soft Computing 8, no. 2. pp. 959–971, March 2008.

[3] Linker, R., I. Seginer, and P.O. Gutman. Optimal CO2 Control in a Greenhouse Modeled with Neural Networks. Computers and Electronics in Agriculture 19, no. 3, pp. 289–310, March 1998.

[4] Perdigones, A., J.L. García, A. Romero, A. Rodríguez, L. Luna, C. Raposo, and S. de la Plaza. Cooling Strategies for Greenhouses in Summer: Control of Fogging by Pulse Width Modulation. Biosystems Engineering 99, no. 4, pp. 573–586, April 2008.

[5] Mahaman, B.D, H.C Passam, A.B Sideridis, and C.P Yialouris. DIARES-IPM: a Diagnostic Advisory Rule-based Expert System for Integrated Pest Management in Solanaceous Crop Systems. Agricultural Systems 76, no. 3. pp. 1119–1135, June 2003.

[6] Lafont, F., and J.-F. Balmat. Optimized Fuzzy Control of a Greenhouse. Fuzzy Sets and Systems 128, no. 1, pp. 47–59, May 2002.

[7] Sigrimis, N.A., K.G. Arvanitis, and G.D. Pasgianos. Synergism of High and Low Level Systems for the Efficient Management of Greenhouses. Computers and Electronics in Agriculture 29, no. 1–2, pp. 21–39, October 2000.

[8] Van Straten, Gerrit. Acceptance of Optimal Operation and Control Methods for Greenhouse Cultivation. Annual Reviews in Control 23, no. 0, pp. 83–90, 1999.

[9] Speetjens, S.L., J.D. Stigter, and G. van Straten. Towards an Adaptive Model for Greenhouse Control. Computers and Electronics in Agriculture 67, no. 1–2, pp. 1–8, June 2009.

[10] Seginer, Ido. Some Artificial Neural Network Applications to Greenhouse Environmental Control. Computers and Electronics in Agriculture 18, no. 2–3, pp. 167–186, August 1997.

[11] Fourati, Fathi, and Mohamed Chtourou. A Greenhouse Control with Feed-forward and Recurrent Neural Networks. Simulation Modelling Practice and Theory 15, no. 8, pp. 1016–1028, September 2007.

[12] G.P.A. Bot. Greenhouse Climate : from Physical Processes to a Dynamic Model. PhD dissertation. Wageningen : Bot, 1983.

[13] Linker, Raphael, and Ido Seginer. Greenhouse Temperature Modeling: a Comparison Between Sigmoid Neural Networks and Hybrid Models. Mathematics and Computers in Simulation 65, no. 1–2, pp. 19–29, April, 2004.

[14] Thompson, M. L. and M. A. Kramer. Modelling chemical processes using prior knowledge and neural networks. American Institute of Chemical Engineers Journal, 40, no. 8, pp. 1328–1340, 1994.

[15] J. Park, I. Sandberg: Universal approximation using radial-basis function networks. Neural Comput., 3, pp. 246–257, 1991.

[16] P. Eredics, P. Györke, T.P. Dobrowieck, B. Pataki. A Small-Scale Experimental Greenhouse for Computing Thermal Models, 13th IEEE International Symposium on Computational Intelligence and Informatics. Budapest, Hungary, pp. 219-223, 2012.