Hybrid Knowledge Modeling for an Intelligent Greenhouse

P. Eredics* and T.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—**The quality of control provided by greenhouse control systems can be improved by applying model based and intelligent control. To this aim a good model of a greenhouse is needed. For a large variety of industrial or recreational greenhouses the derivation of an analytical model is not feasible, due to the large amount of identification data and expertise needed. A black-box model of the whole system is neither feasible, as it would require a huge number of teaching samples. The only way to tackle this problem is the decomposition of the greenhouse system using hybrid modularized models, where each module represents a relatively loosely coupled component of the system. This paper discusses such decomposition and a model under development.**

I. Introduction

Greenhouses are construction structures widely used in vegetable production and for growing ornamental plants and flowers. Solar radiation passing through the transparent walls and roofs is essential for photosynthesis, and supplements the heating system in the cold season. In hot weather other actuators, like roof vents, shading systems, exhaust fans or evaporative cooling may be also used to avoid overheating. In most modern greenhouses these automated actuators are operated by some kind of control system.

Despite a huge leap in informatics, apart of special purpose applications and studies, control systems for greenhouses have not changed much in the last years: each actuator is usually individually controlled based on set-points and actual measurements [1]. This traditional control design has following major drawbacks:

(1) The adjustment of set-points depends strongly on the expertise of the greenhouse owner.

(2) The control system is reactive, without predicting the future state of the greenhouse, it is impossible to control effectively for a more extended time horizon.

(3) The adjustment of the actuators is not synchronized (all are set independently from each other), resulting in possible oscillations in the control and poor maintenance of the greenhouse climate.

One way to overcome these limitations is to increase the level of intelligence of the system [2].

II. Intelligent Control for an Intelligent Greenhouse

The goal of all greenhouse control solutions is to provide suitable environmental conditions for the plants. The traditional form of set-point based control leans heavily on human intelligence. It assumes that the greenhouse operator has the necessary know-how and a good understanding of the interactions between the plant biology, plant grow economics, and the control system itself. Only then is the operator able to choose optimal set-point combinations. Unfortunately with several actuators finding the optimal control configuration intuitively without serious theoretical modeling and computing is impossible, so the control performance in this case cannot be maximal. Yet the greenhouse operator is fully aware of the needs of the plants, the control system therefore should operate on this available information rather than on the unreliable set-point values. Recapitulating, instead of accepting set-points, an intelligent control system should expect and accept global control goals, not influenced by imprecise assumptions made by human operators in set-point selection. For a greenhouse the control goals can be set e.g. in the form of target parameter zones.

The concept of control goal as the direct control information makes the human interaction easier, because the knowledge intensive transformation from goals to control actions is left to the control system. This transformation can be implemented with predictive modeling, which solves also the second problem of traditional control, namely its reactiveness. Predictive modeling means in our case making assumptions of the actuator settings, and predicting the future thermal states of the greenhouse. These thermal states anticipated over a given time span can be then evaluated with respect to the control goals. The costs of the actuator setting (e.g. the cost of running window opening motors or of turning on heating) and the deviation from the goals can be fused together into a numerical cost function. By computing the minimal value of this cost, e.g. by trying all different actuator configurations, makes it possible to find the most appropriate actuator settings at any time. Setting the actuators that way is thus a rational decision based on the overall goals of the greenhouse maintenance. Unfortunately some actuator configurations cannot be kept unaltered for the whole prediction time span (e.g. configurations with heating or misting turned on) so the predictive modeling scheme described above is not directly applicable.

Predictive modeling solves in part the missing synchronization of the actuators, but the common problem of traditional control, i.e. swinging controlled variable, still remains. The possible control loops, repeatedly setting and resetting the actuators cannot be avoided this way. All these problems can be handled by (AI) planning. Instead of examining fixed actuator configurations for the whole prediction length, control plans can be used to allow changes in the actuator states at any time. The quality of the plans (costs of the deviation from the goals and costs of the actuators yielding the total cost value) can be evaluated by a straightforward extension of the ideas discussed earlier.

Similar problem setting can be found in the HVAC (Heating, Ventilation and Air-Conditioning) control in buildings [3]. The predictive HVAC control is based on a comfort index well applicable to the human. The thermal modeling is simple, and the people can leave the building by any means if uncomfortable. The plants however cannot tell, or walk out, but will fade beyond return. So in a seemingly smaller greenhouse environment the situation is more complicated. We have only a rough idea what the comfortable climate would really mean. As a remedy we must use better models, but then the great structural and temporal variability of the targeted greenhouses chips in.

We expect from the concept of intelligent control to provide solutions to all principal limitations of the traditional greenhouse control, with better environmental conditions for the plants and lower costs for the owners. Its realization requires on the other hand much more computational power and resources for the control system. Luckily unused computational power is already present in many contemporary greenhouses in the form of data loggers. What is not readily present is the know-how to make farseeing plans and to transform them into suitable short-term control parameters.

III. The Requirements and the Experimental Testbed

The necessary basis for an intelligent control is the precise model of the greenhouse conditioned on measurements collected from a real greenhouse environment [4]. Traditional control systems record only a single temperature measurement within the greenhouse and optionally one more from outside implicitly assuming a single compartment homogeneous model, whatever the greenhouse structure or the actual plant 'filling' is. This is of course far from accurate, especially when the interaction between the internal climate and the external weather is at stake. Accurate modeling requires large amount of data to be collected from strategically selected locations. Luckily quasi-homogenous thermal zones can be localized in most of the greenhouses. Fig. 1 shows typical zones in a simplified industrial greenhouse. For efficient thermal modeling temperature measurements from every zone are essential. In special-purpose zones (e.g. on the desks under cover) more measurement locations might be also reasonable. By increasing the number of measurement locations the greenhouse operator may also specify more detailed control goals for the different areas of the house.

The real-life object behind the simplified model is a medium size 100 m^2 greenhouse equipped for the experiments with a measurement and control system [5]. This greenhouse has 18 desks; therefore 18 temperature measurements are available from Zone-1. Zone-2 can be separated into 2 compartments in case of low utilization, i.e. 2 temperature measurement points are set up here. All other zones are measured at a single location. In Zone-2 and Zone-4 radiation measurements are also recorded (see Table 1). It is also beneficial to use weather forecasts (available on the internet) applicable to the location of the greenhouse to obtain external temperature and predicted cloud coverage values.

IV. Modeling the Greenhouse

The principal physical parameter influencing the well-being of the plants and the expenses of the owner is the internal temperature. The goal of thermal modeling is thus to predict the future temperature values for all important measurement location (zones). The natural inputs to such model are: (1) the latest measured temperature and radiation values; (2) online weather forecast; and (3) the actual configuration of the actuators. The required outputs are the predicted temperature values for all measurement locations within the prediction horizon.

During model type selection the most important properties are adaptivity, modularity, transparency and coverage. The model has to be adaptive to follow changes caused by internal alterations (e.g. growing plants, different utilization of the greenhouse through the year) and external effects (e.g. varying radiation energy input for different time periods or variable shading effects of the trees around the greenhouse depending on the seasons).

Although in some cases (mainly in intentionally designed lab environments, or large scale industrial greenhouses) control based on involved analytic parametric models is possible (e.g. [6]), we aimed at practical situations, where the greenhouse infrastructure is heavily constrained by economical reasons, the character of the production could fluctuate on a weekly basis, and where there is a pronounced lack of expertise in control solutions. In such cases the only feasible solution is to found the control on the adaptive black-box models, where our main concern will be to couple them by

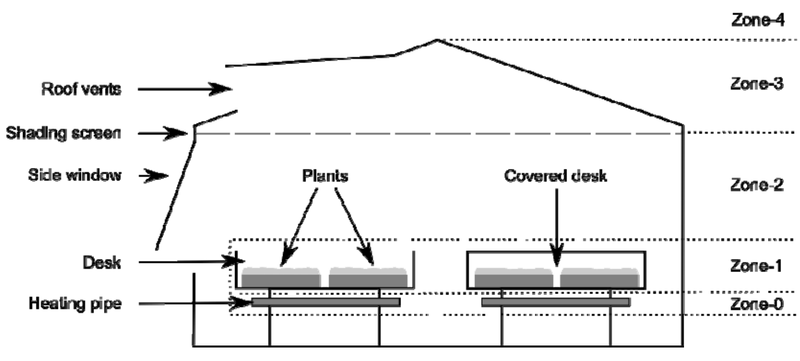

Figure 1. Simplified thermal zone structure of an industrial greenhouse

TABLE I.
ZONES IN THE EXPERIMENTAL GREENHOUSE

Zone	Principal information	Temperature measurement locations	Radiation measurement locations
0	Heating pipe temperature	1	-
1	Desk temperatures	18	-
2	In-house air temperature below the shading screen	2	1
3	In-house air temperature above the shading screen	1	-
4	External temperature and online weather data	1+1 online	1+1 online

learning with the compartments and structures evident within the system.

Another consideration is the modularity of the model. Monolithic model of the whole system would be a comfortable solution, but due to the large number of inputs (24 temperature measurements, 6 actuator states and 3 radiation values) and the complexity of the problem this is practically unfeasible, i.e. the decomposition of the whole model into modules cannot be avoided.

As mentioned before the aim of the thermal model is to predict the thermal state of the desks and other internal zones. To this aim the predictive model of the external weather is also expedient as it should help considerably in following weather trends in the internal predictions. Due to the one-way casual relation (the external weather can affect the greenhouse, but not the other way round) this problem and the model can be treated separately. This yields two completely different problems to be solved, namely the external weather prediction and the internal thermal modeling (see Fig. 2). The granularity of decomposition could be raised even further, as the resulting modeling problems are easier to handle. The two modules of Fig. 2 decouple the system along the walls of the greenhouse. This is the original approach of the traditional greenhouse control, dealing with only these two system components. In our case both compartments can be decomposed further as shown in the proposed model decomposition in Fig. 3, with expected result of a more accurate modeling. Module-A and B are involved in the external weather and temperature forecasts. Module-C is the heating system while Module-D represents the whole house except the desks. Module-E and F are responsible for the prediction of the state and the temperature of the desks itself.

A. External Near-House Weather Model

Online weather forecasts have the advantage of providing directly full forecasts along with current measurements. Unfortunately their precision at the actual location of the greenhouse is usually not acceptable. Therefore Module-A is responsible for the prediction of the locally recorded external weather (temperature and radiation) based on the local measurements. For this problem we used time-series mining of the earlier measurements to produce reliable predictions for a few hours ahead [7]. The model is implicitly present in the recorded time series and the prediction is based on finding similar trends from the past. After finding such situations in the past, their subsequent measurements can be used as prediction candidates to the present. The weighted average of the predictions (weights are chosen as a function of the similarity with the examined situation) yields a good forecast for the forthcoming hours. The model initialization is done simply by collecting measurements and the adaptation is automatically assured by the model derivation.

B. Online Missing Data Restoration Model

Although online weather data from the internet is not accurate enough for local temperature modeling, online trends predictions are quite reliable, thus using them as a model input looks beneficial. On the other hand because of the need of internet connection this input has lower reliability then other locally measured data. Thus during normal operation Module-B is monitoring the weather forecast from the online source. If such data is available and trustworthy (variables are in range and changes are acceptable) then it is simply forwarded to Module-D (Global Greenhouse Model). However in case of network outage or data corruption the data has to be restored from the local measurements, because Module-D cannot work without this input. For temperature a simple model can be built to approximate the relation between online and locally recorded temperatures. The cloud coverage cannot be handled similarly, because local radiation sensors provide no data at night. Although the former forecasts of the online source can be used for a short time, after a few hours this data can only be restored with large errors as the average of the formerly recorded values. This setup ensures that Module-D has always all the input values it needs.

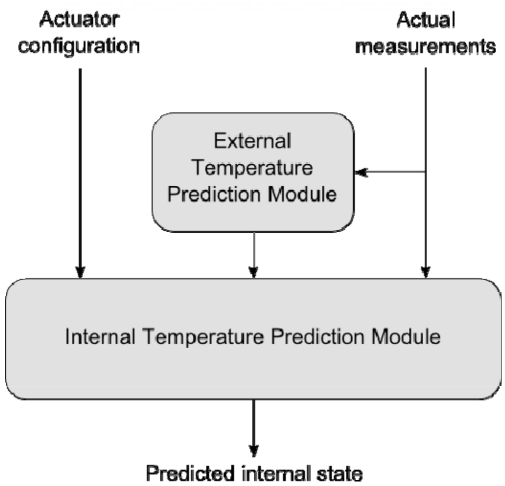

Figure 2. Initial decomposition step of the global thermal model

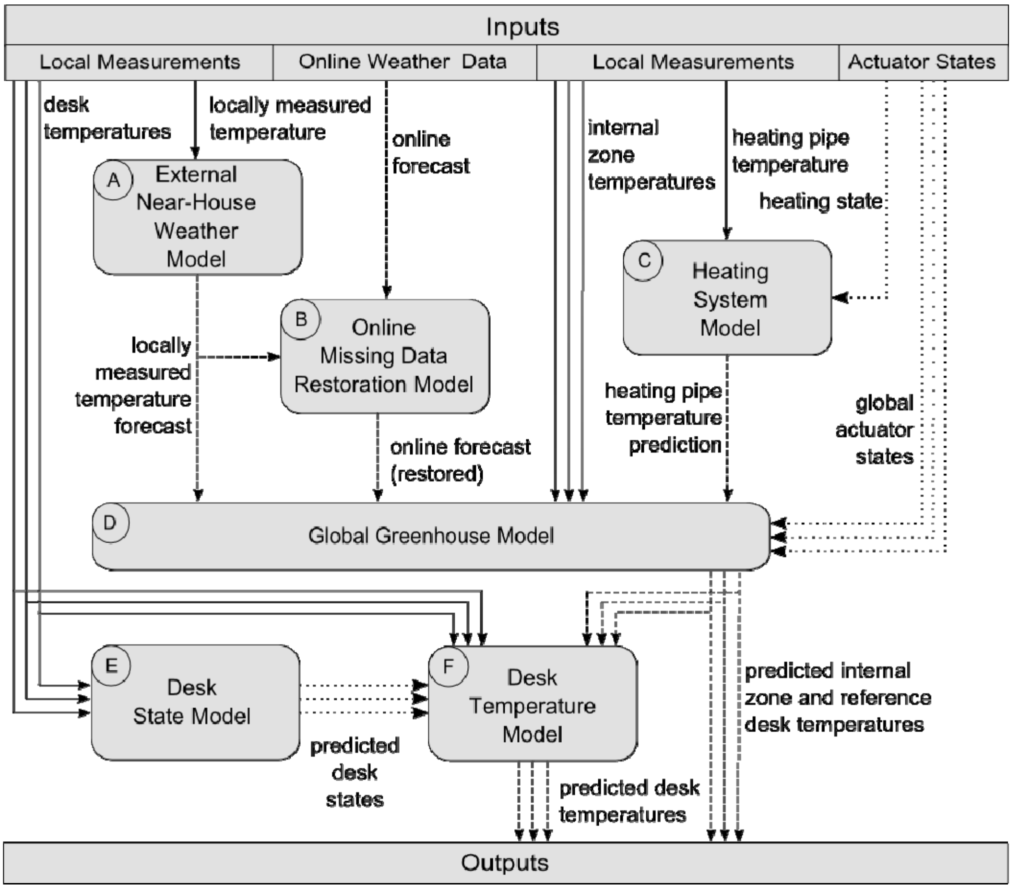

Figure 3. The proposed decomposition of the global greenhouse model into 6 modules related to the thermal substructures of the greenhouse (solid lines: measurements; dotted lines: states; dashed lines: predictions)

C. Heating System Model

Module-C predicts the heating pipe temperature, based on the current internal temperature, heat pipe temperature and the heating control signal. This model can be easily separated from other modules, as the pipe temperature is mainly determined by its control signal. This module can be implemented by a simple learning system (e.g. a neural network), where training data can be obtained from a few experiments with the heating system.

D. Global Greenhouse Model

In the proposed model decomposition Module-C and Module-D are responsible for all internal zones, except for the desks: at this level the desks are represented by a single reference desk as this approach reduces the number of inputs significantly without notably affecting the modeling performance. The reference desk is used later to calculate the complete state of Zone-1. The application of the reference desk also makes the design flexible to handle less equipped greenhouses.

Module-D can be realized also as a neural network with the following inputs: (1) locally measured temperature forecast from Module-A; (2) online weather forecast (original or restored by Module-B, if necessary); (3) the latest measurements from all the zones; (4) the predicted temperature of the heating system; (5) the latest actuator configuration. As the outputs Module-D produces predictions for the reference desk, Zone-2 and Zone-3.

E. Desks State Model

As a special condition the desks in the experimental greenhouse can be covered for the protection of sensitive plants. Unfortunately such state of the desks (covered or not) is not recorded in any way, although this circumstance heavily determinates the relation between the desk and its environment. This calls for the application of separate Module-E to predict the state of every single desk. Its inputs are the temperatures measured close to the plants on every desk (below the cover if the desk is covered) along with the Zone-2 recordings. This module is special, because due to the rare state changes of the desks and its independence of the automated actuators this model can be recomputed off-line (updated for example every night). This module can be implemented as a decision tree.

F. Desk Temperature Model

Module-F is responsible for predicting the temperature of the desks (air temperature close to the soil level) based on (1) the current measurements; (2) the predicted state of the desks; (3) the predictions for the reference desk and (4) the predictions for Zone-2. This model has 20 temperature and 18 state inputs with 18 prediction outputs, and can be also realized as a neural network. The model could be further decomposed into 18 separate components for all the desks, but then the problem would be the modeling of the coupling between closely placed desks.

The modeling of such cross-dependence is left to the learning of the neural network.

V. RESULTS

The hybrid greenhouse model based on the decomposition outlined in this paper is currently under development, and will be tested on the recorded data from the experimental greenhouse. Module-A (External Near-House Weather Model) and Module-C (Heating System Model) have been already implemented, while implementation of the other 4 modules is still in progress based on the measurements and the outputs of the existing modules.

Time-series mining applied to the local temperature prediction produces an average absolute error below 2.5 Celsius degrees in the first 4 hours (see Fig 4). This time span is critical in the current application, because this is the reasonable length for the forward planning of the actions. The error is lower than the average absolute error of the online forecast and it is acceptable for the current purpose.

The behavior of the heating system depends mainly on its control signal: when the heating is turned on, the temperature of the pipe is rising almost constantly, while the temperature with heating turned off is more dependent on the air temperature. To handle these different behaviors 2 separate neural models are built for both cases. The switching between the two models is determined by the heating control signal of the executed plan. Both neural models have 8 neurons in the hidden layer and their common inputs are the latest measured temperature values. The cooling pipe model has an additional input representing the time since the heating was deactivated. In Fig. 5 the result of a 4 hours long prediction example is shown.

VI. CONCLUSION

Accurate modeling of a greenhouse is essential for the evolution of greenhouse control towards high quality intelligent systems. Modeling such a complex system cannot be done by simply dismissing the problem structure and applying global black-box modeling, because the model complexity, the necessary large amount of training data and training time may become prohibitive. The proposed decomposition of the modeling problem uses 6 modules for handling different thermal processes related to different structures within the greenhouse. Dedicated modules are predicting the parameters of the external weather. Separate modules are used to model the global interior processes of the greenhouse. Finally yet two other components determine the state and calculate the prediction for each desk. Models implemented so far include the external local temperature prediction and the prediction of the heating pipe temperature. Recent work aims at the implementation of the remaining modules allowing the development of the hybrid knowledge model based intelligent control for the whole greenhouse.

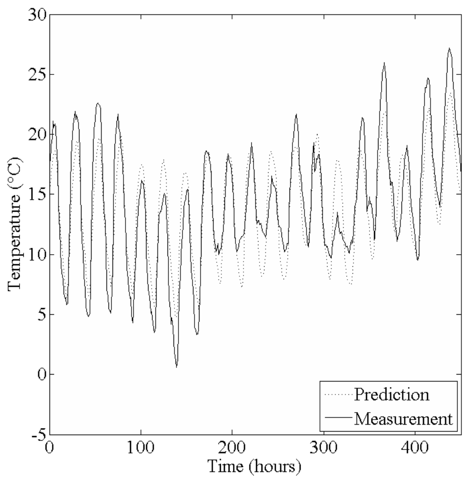

Figure 4. Predicting the external temperature 4 hours ahead with time-series mining applied on a 440 hours validation dataset

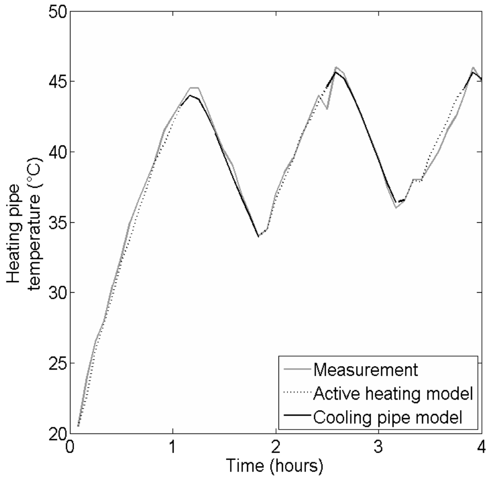

Figure 5. Predicting the heating pipe temperature with switching between 2 independent models

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Hungarian Fund for Scientific Research (OTKA).

REFERENCES

[1] Zazueta, F.S., Bucklin, R., Jones, P.H., Smajstrla, A.G.: Basic Concepts in Environmental Computer Control of Agricultural Systems. Agricultural and Biological Engineering Dept, Institute of Food and Agricultural Sciences, University of Florida, 2008.

[2] Blasco, X., Martineza, M., Herreroa, J.M., Ramosa, C., Sanchisa, J.: Model-based predictive control of greenhouse climate for reducing energy and water consumption. Computers and Electronics in Agriculture, pp. 49–70, Elsevier Science Publishers B. V. Amsterdam, The Netherlands, 2007.

[3] [HVAC] R.Z. Freirea, G.H.C. Oliveirab, N. Mendes: Predictive controllers for thermal comfort optimization and energy savings. Energy and Buildings, vol. 40, pp 1353-1365, 2008.

[4] Cunha J.B. Greenhouse Climate Models: An Overview. In: European Federation for Information Technology in Agriculture Conference, Debrecen, Hungary, 2003.

[5] P. Eredics: Measurement for Intelligent Control in Greenhouses, 7th International Conference on Measurement, pp 178-181, Smolenice Castle, Slovakia, 2009.

[6] G.P.A. Bot: Physical modeling of greenhouse climate. In: IFAC/ISHS workshop held in Matsuyama, Japan, 1991.

[7] P. Eredics: Short-Term External Air Temperature Prediction for Intelligent Greenhouse by Mining Climatic Time Series. In: 6th IEEE International Symposium on Intelligent Signal Processing, pp. 317–322, Budapest, Hungary,2009.