Neural Network Models for Predictive Climate Control in Intelligent Buildings.

Bertil Thomas
Dept. of Electrical and Computer Engineering.
Chalmers Lindholmen University College.
P.O.Box 8873
SE-402 72 Göteborg, Sweden
bertil@chl.chalmers.se

Mohsen Soleimani-Mohseni
Dept. of Building Services Engineering,
Chalmers University of Technology,
402 96 Göteborg, Sweden
mohsen@chl.chalmers.se

Abstract - **The problem how to identify black-box prediction models for model predictive control (MPC-control) of the indoor climate in buildings is discussed. A number of identification experiments have been carried out in a building and different prediction models, such as linear ARX- and ARMAX-models and non-linear neural network models, have been identified based on these experiments. For the models, different input signals have been used, such as the outdoor and indoor temperature, heating power, time of day and sun radiation. It is shown that non-linear neural network models give more accurate temperature predictions than the linear models, and that the predictions are improved when using more input signals to the models. In particular, it is shown that for non-linear models, but not for linear, the predictions are improved when using the time of day as an input signal. This shows that a non-linear combination of sun-radiation and time of day is important when predicting the indoor temperature of buildings.**

Keywords: **Model Predictive Control, temperature prediction, identification, neural networks, Building Automation Systems.**

I. INTRODUCTION

Today, there are many ways to control the indoor climate in buildings such as schools, offices, hospitals, residential buildings and supermarkets. A number of control concepts have been proposed and examined during the last 30 years [1, 2]. Yet, there is no common agreement which is the most efficient way of controlling the climate in different types of buildings. There is a need for a control system, which easy to take into service, easy to handle and which gives a comfortable indoor climate at a low energy cost. The authors of this paper believe that Model Predictive Control (MPC-control) will become a useful control strategy when designing intelligent Building Automation systems (BA-systems) of the future.

Model Predictive Control is a control strategy where the calculation of the control signals is based on predictions of future values of the controlled variables [3, 4]. In MPC-control of buildings it is of great importance to use accurate prediction models with ability to predict the indoor climate at least 15-30 minutes ahead. Since the indoor climate (temperature, air quality and humidity) of a building is affected by a lot of variables, it is important that the predictions are based on many input signals in order to get the required accuracy. Some variables which affect the indoor climate are directly connected to the outdoor climate, such as the outdoor temperature, solar radiation, outdoor humidity and wind. Other variables are connected to the activity inside the building, such as the number of people inside the building, their current activities, the electrical power use, ventilation etc [5].

Some advantages of MPC-control when compared to other control methods for HVAC-control in BA-systems are the following:

- The method is not restricted to linear processes or linear models. It can be used both for linear and non-linear processes and the prediction models may be linear or non-linear. The neural network prediction models can perhaps be trained with data, which is already gathered in the BA-system of a building.

- The MPC-control method is well suited for control of processes with measurable disturbances. Since all measurable disturbances can be included in the prediction model, there is no need for a separate calculation of the feed-forward parts of the controllers. This will simplify the calculation of feed-forward controllers.

- The method is well suited for processes and control systems with upper and lower limitations on what signal levels the actuators can produce, (for example systems with fans, control valves and radiators). Such limitations can be taken account for in the performance criteria when calculating the control signals.

- The method can be used for multivariable systems with many input- and output signals.

In this report, the problem how to identify black-box prediction models for combined feed-forward and feedback MPC-control of the temperature in buildings is discussed. A number of identification experiments have been carried out in a small building and different linear and non-linear prediction models have been identified based on these experiments. Neural network models have proven to be successful non-linear black-box model structures in many prediction applications and they have attracted a growing interest in recent years [6, 7, 8, 9]. There are also other non-linear model structures that can be used for prediction, such as radial basis function networks, wavelet networks, recurrent networks, graybox models and fuzzy models. However, in this report, we will concentrate on feed forward neural network models for this problem.

II. IDENTIFICATION EXPERIMENT

The identification experiments were carried out in an experimental building, with the inner dimensions 4.6x2.4 m. The building is heated by an electrical radiator with adjustable power, it has windows in three different directions and it is surrounded by a number of trees (Fig. 1.). Two temperature sensors were used to measure the temperatures inside and outside the building:

- Indoor temperature sensor. This was placed in the middle of the building at 1.7 m height.

- Outdoor temperature sensor. This was placed in the shadow at the north of the building.

In addition to the two temperature sensors above, a third sensor was used to give an estimate of the sun radiation intensity. This temperature sensor was placed outside the building, directly in the sunlight, and the vital part of the sensor was covered with a black cylinder of rubber. When the sun radiation is strong, this sensor gives a higher temperature value than the outdoor temperature sensor discussed above. The difference between the two temperatures has been used as an estimate of the sun radiation intensity.

The total length of the identification experiment was 624 hours (26 days and nights) and it was carried out between the middle of August and middle of September 2003. All variables, such as the temperature values of the three temperature sensors and the electrical power of the radiator were measured with a sampling interval of 15 minutes (that is, 96 temperature readings were collected every 24 hour, and a total of 2496 readings for every sensor). The sampling interval is about 10% of the measured time constant for the building. The first 1440 readings (16 days and nights) were used as training data (estimation data) for the prediction models, while the last 1056 readings (11 days and nights) were used as test data (validation data) for verification of the models. The electrical power was adjusted manually between 0 and 2000 W during the experiment.

A. Variables Used

$T_o(k)$	Outdoor temperature, shadow (^{O}C)
$T_i(k)$	Indoor temperature, measured (^{O}C)
$T_s(k)$	Outdoor temperature, sun (^{O}C)
$P(k)$	Radiator power (W)
τ	Time of day (h)
$T_P(k)$	Predicted indoor temperature (^{O}C)
k	Sampling number

The following values were measured:

$\max(T_o) = 22.9°C$
$\min(T_o) = 1.3°C$
$\max(T_i) = 37.7°C$
$\min(T_i) = 6.1°C$
$\max(T_s) = 39.5°C$

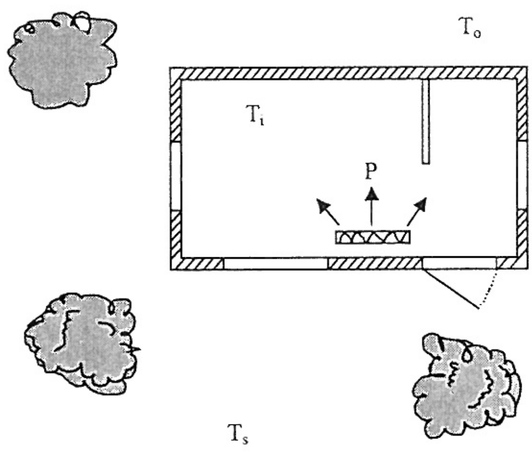

Fig. 1. The building and its surroundings.

$\min(T_s) = -0.6°C$
$\bar{T}_o = 13.83°C$ (Average outdoor temperature)
$\bar{T}_i = 17.9°C$ (Average indoor temperature)
$\bar{P} = 141W$ (Average electrical power use)

Fig. 2-3 show parts of the temperature recordings for the building studied.

B. Models

Using the identification experiment discussed above, a number of linear and non-linear models has been identified to predict the indoor temperature of the building 15 or 30 minutes ahead. The models differ regarding the number of input signals and regarding the order and structure of the models. In the simplest non linear models, only three input variables were used (T_i, T_O and P), while four (T_i, T_O, P and T_S) or five inputs (T_i, T_O, P, T_S and τ) were used in the more complex models. The same input variables were used for the identified ARX-models.

First order one-step-ahead prediction models:

$$T_P(k+1) = f[T_i(k), T_o(k), P(k)] \tag{1}$$

$$T_P(k+1) = f[T_i(k), T_o(k), T_s(k), P(k)] \tag{2}$$

$$T_P(k+1) = f[T_i(k), T_o(k), T_s(k), P(k), \tau(k)] \tag{3}$$

First order two-step-ahead prediction models:

$$T_P(k+2) = f[T_i(k), T_o(k), P(k)] \tag{4}$$

$$T_P(k+2) = f[T_i(k), T_o(k), T_s(k), P(k)] \tag{5}$$

$$T_P(k+2) = f[T_i(k), T_o(k), T_s(k), P(k), \tau(k)] \tag{6}$$

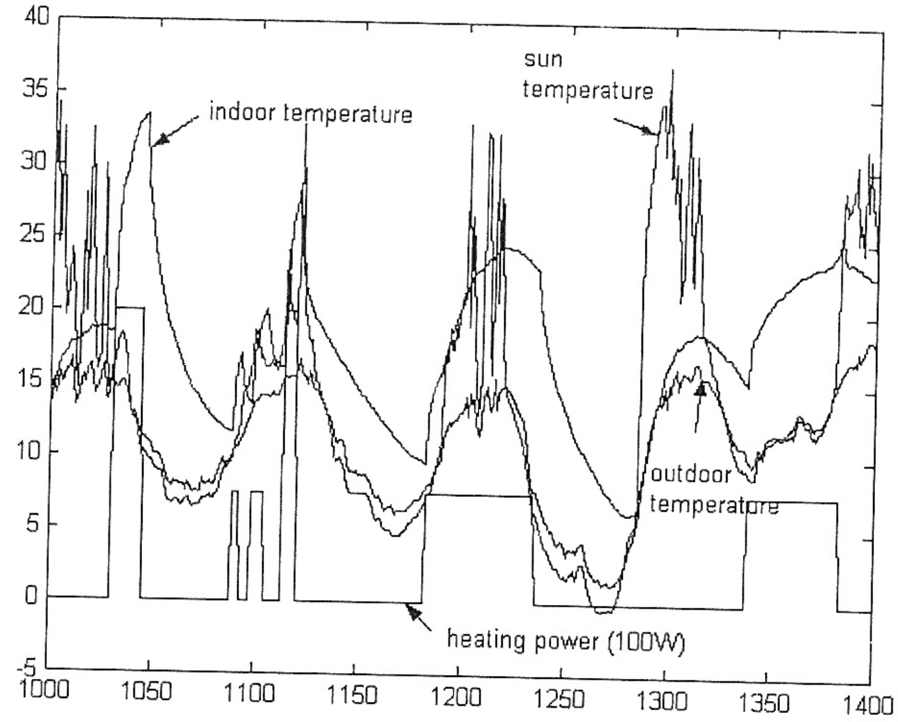

Fig. 2. A part of the temperature recordings (training-data) for the building studied. The sampling number is on the horizontal axis

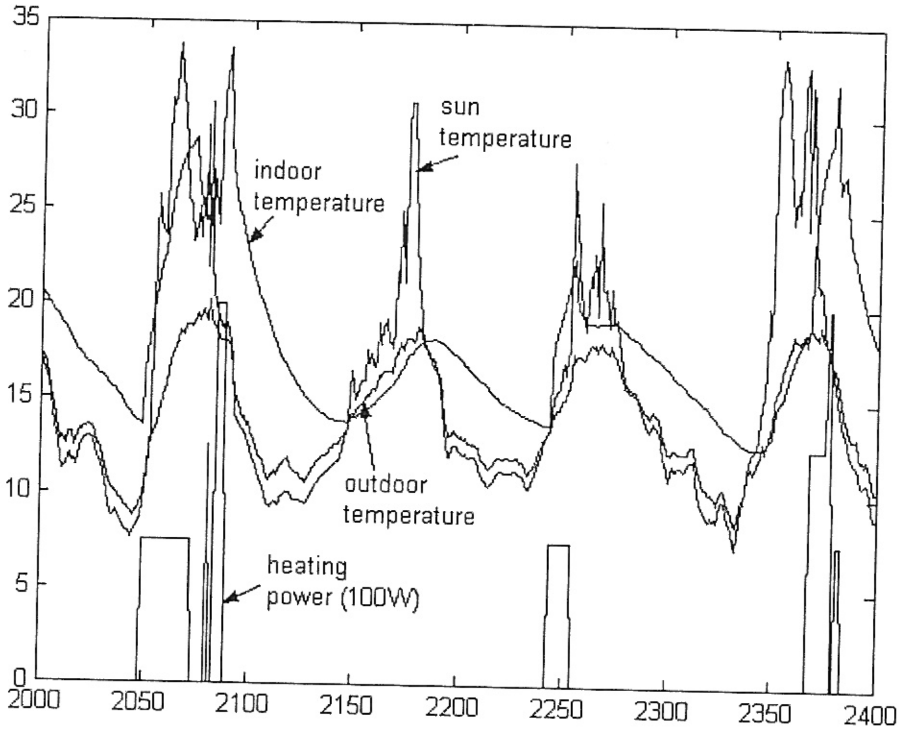

Fig.3. A part of the temperature recordings (test-data) for the building studied. The sampling number is on the horizontal axis

Second order one-step-ahead prediction models:

$$T_P(k+1) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ P(k), P(k-1) \end{bmatrix} \quad (7)$$

$$T_P(k+1) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ T_s(k), T_s(k-1), P(k), P(k-1) \end{bmatrix} \quad (8)$$

$$T_P(k+1) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ T_s(k), T_s(k-1), P(k), P(k-1), \tau(k) \end{bmatrix} \quad (9)$$

Second order two-step-ahead prediction models:

$$T_P(k+2) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ P(k), P(k-1) \end{bmatrix} \quad (10)$$

$$T_P(k+2) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ T_s(k), T_s(k-1), P(k), P(k-1) \end{bmatrix} \quad (11)$$

$$T_P(k+2) = f\begin{bmatrix} T_i(k), T_i(k-1), T_o(k), T_o(k-1), \\ T_s(k), T_s(k-1), P(k), P(k-1), \\ \tau(k) \end{bmatrix} \quad (12)$$

In the table above, $f[\cdot]$ denotes a linear or a non-linear function. In the linear ARX-case, $f[\cdot]$ denotes a linear difference-equation function of the following type:

$$\begin{aligned} T_P(k+1) &= a_1 T_i(k) + a_2 T_o(k) + a_3 T_s(k) + \\ &+ a_4 P(k) + a_5 \tau(k) + a_6 \end{aligned} \quad (13)$$

The parameters of the ARX-models have been identified using the ordinary least squares method. Other linear models, such as OE- (Output Error), BJ-(Box-Jenkin) and ARMAX-models [8] have been identified using an iterative Gauss-Newton algorithm. For all linear models, the parameters have been identified by minimizing the purely quadratic prediction error criteria (*MSE*) below. In the non-linear case, feed-forward multi-layer neural network models with one hidden layers of sigmoid (logsig) neurons and one linear output neuron have been used. The parameters (i e synaptic weights and thresholds) of the neural networks models have been calculated using the Levenberg-Marquardt algorithm [10]. Other algorithms such as the scaled conjugate gradient method have been examined as well, but the results using these methods were not as accurate as for the Levenberg-Marquardt algorithm and are therefore not presented here.

C. Performance Criteria

In order to compare the predicting accuracy for the different models, the following two performance measures have been calculated for all identified linear and non-linear models:

Mean Square Error

$$MSE = \frac{1}{n}\sum_{k=2}^{n+1}\left[T_P(k) - T_i(k)\right]^2 \quad (14)$$

Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{k=2}^{n+1}\left|T_P(k) - T_i(k)\right| \quad (15)$$

where $T_P(k)$ is the predicted indoor temperature (^{O}C) and $T_i(k)$ is the measured indoor temperature (^{O}C). The parameter n is the number of examples. The mean absolute error has been calculated both for training data and for test data.

III. RESULTS

The result of the identification experiment for the non linear neural network models is dependent on the initial random values for the synaptic weights in the network. Therefore, the result will in general not be the same in two different trials even if the same training examples have been used. In this report, we will only present the best result obtained after 10 trials with the same input-output-data for each model. As the "best" result, the one that gave the smallest mean absolute error *MAE* for test data was chosen. The networks used are feed forward neural networks with 10 neurons in the middle layer and one output neuron. The number of network inputs was between 3 and 9 dependent on the number of input signals and the order of the models.

In addition to the networks discussed above, we have also examined networks with more than two layers of neurons and networks with more than 10 neurons in the middle layer. However, we could not observe any significant improvement when using larger networks, so only the results for x-10-1 networks are presented here. In general, 100 epochs of training was used in all learning experiments. This number of epochs was chosen because in a number of trials, it was shown that the prediction result for test-data could in general not be further improved by using more than 100 epochs.

In the tables I-III, the result using first and second order ARX-models are presented together with the result using non-linear neural network prediction models.

TABLE I

Result for first order one-step-ahead prediction models

Input signals	Performance criteria	ARX model	Neural network model
T_O, T_i and P	*MSE*	0.232	0.1069
	MAE (training)	0.2639	0.1618
	MAE (test)	0.2769	0.1994
T_O, T_i, P and T_S	*MSE*	0.232	0.0836
	MAE (training)	0.2630	0.1449
	MAE (test)	0.2755	0.1637
T_O, T_i, P, T_S and τ	*MSE*	0.230	0.0512
	MAE (training)	0.2638	0.1260
	MAE (test)	0.2748	0.1458

TABLE II

Result for second order one-step-ahead prediction models

Input signals	Performance criteria	ARX model	Neural network model
T_O, T_i and P	*MSE*	0.0591	0.03833
	MAE (training)	0.1268	0.1051
	MAE (test)	0.1183	0.1070
T_O, T_i, P and T_S	*MSE*	0.0558	0.0370
	MAE (training)	0.1168	0.1007
	MAE (test)	0.1047	0.1000
T_O, T_i, P, T_S and τ	*MSE*	0.0557	0.0210
	MAE (training)	0.1167	0.0956
	MAE (test)	0.1048	0.0992

TABLE III

Result for second order two-step-ahead prediction models

Input signals	Performance criteria	ARX model	Neural network model
T_O, T_i and P	*MSE*	0.6318	0.5995
	MAE (training)	0.336	0.2971
	MAE (test)	0.3087	0.2698
T_O, T_i, P and T_S	*MSE*	0.6229	0.5985
	MAE (training)	0.3157	0.2902
	MAE (test)	0.2792	0.2524
T_O, T_i, P, T_S and τ	*MSE*	0.6223	0.4894
	MAE (training)	0.3161	0.2830
	MAE (test)	0.2782	0.2484

From the results presented above, the following observations can be done:

- For all cases studied above, the non-linear neural network models give better temperature predictions than the linear ARX-models. For first order models, the difference is larger than for second order models. One example is that a first order non-linear one-step-ahead prediction model with five input signals gives *MAE(test)*=0.1458, while the linear model gives *MAE(test)*=0.2748.

- The predictions are better for models with four input signals than for models with three input signals. This holds true both for linear and non-linear models, although the improvement is more significant for non-linear models. One example is that a first order non-linear one-step-ahead prediction model with three input signals gives *MAE(test)*=0.1994, while the same type of model with four input signals gives *MAE(test)*=0.1637.

- For the neural network models, but not for the ARX-models, the predictions are further improved when using the time of day as a fifth input signal. One example is that a second order, non-linear, two-step-ahead prediction model with four input signals (T_O, T_i, P and T_S) gives *MSE*=0.5985, while the same type of model with five inputs (T_O, T_i, P , T_S and τ) gives *MSE*=0.4894. This property of the non-linear model can probably be explained by the fact that a non-linear combination of sun-radiation and time of day is important when predicting the indoor temperature. Some hours at daytime the sun is shining directly into the windows of the building. This gives a quick rise of the indoor temperature. Other hours at daytime the sun is shadowed by trees. That is, the sun radiation is affecting the indoor temperature very different depending on the time of day. However, only a non-linear model can map this non-linear combination of time of day and sun radiation in a good way.

- The predictions for test data are about as good as the predictions for training data. For second order two-step-ahead prediction models, the result for test data is even better than for training data. This shows that the obtained networks have a good ability to generalize.

In addition to the experiments and results presented in the tables above, other types of linear models such as Output-Error (OE-), Box Jenkins (BJ-) and ARMAX-models, as well as higher order ARX-models have also been identified using the same data as above. For higher order ARX-models we obtained the MSE-values according to table 4 below. The table shows that higher-order ARX-models are not significantly better than second order ARX-models. This indicates that a more complex noise-model is not required [8]. The BJ- and ARMAX-structures gave almost the same MSE-values and MAE-values for test data as the ARX-models when using the same order of the models. The difference between the MAE-values is less than 0.5%. However, the MAE-value for test data and ARMAX-models are increasing when using higher model orders than 2, while the best ARX-model is of fourth order. Some results when using higher order ARX-models as well as ARMAX- and BJ- models are given in table 4. The results using OE-models were clearly worse than the result for ARX-models and are therefore not further discussed here.

The conclusion regarding the ARMAX-, BJ- and OE-models is that they will not make any significant improvement when compared to the ARX-model when predicting the indoor temperature of the building. These models will not change the conclusions regarding the difference between neural network models and linear (ARX) models.

TABLE IV:
Result for one-step-ahead ARX-, ARMAX- and BJ- prediction models of different orders when using T_O, T_i, P and T_S as input signals

Order of the model	*MSE* for ARX-models.	*MSE* for ARMAX-models.	*MSE* for BJ-models.
1	0.2320	0.2100	0.09902
2	0.0558	0.05495	0.05468
3	0.05460	0.05431	0.05480
4	0.05436	0.05382	0.05322
5	0.05422	-	-
6	0.05385	-	-
7	0.05367	-	-
8	0.05356	-	-

IV. CONCLUSIONS

In this paper, the problem how to identify black-box prediction models for MPC-control of the indoor climate in intelligent buildings have been discussed. For the building studied, it was shown that non-linear neural network models trained by the Levenberg Marquardt algorithm gave more accurate temperature predictions than linear regression ARX-models using the least squares method. It was also shown that the predictions are improved when using more input signals to the models. In particular, it is shown that for non-linear neural network models, but not for linear models, the predictions are further improved when using the time of day as a fifth input signal.

V. REFERENCES

[1] Underwood, *HVAC Control Systems*, E & F Spoon Publishing Company, 1999.

[2] Nilsson, Per-Erik and the Commtech Group, *Achieving the Desired Indoor Climate, Energy Efficiency Aspects of System Design*, Studentlitteratur, Lund, Sweden, 2003.

[3] Prett, D. M. and Garcia, C. E. *Fundamental Process Control*, Butterworths, USA, 1988.

[4] Maciejowski, J. M, *Predictive Control with Constraints*, Prentice Hall & Pearson Education Limited, Essex, England, 2002.

[5] Mohsen Soleimani-Mohseni, *Feed-forward control and dynamic modelling in temperature control of buildings*, Document D-65, Dep. of Building Services Engineering, Chalmers University of Technology, Sweden, 2002.

[6] Jonas Sjöberg, *Non-Linear System Identification with Neural Networks*, Dissertation no. 381, Linköping University, Sweden, 1995.

[7] Nelles, O. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*, Springer Verlag, Berlin, 2001.

[8] Lennart Ljung, *System Identification – Theory for the User*, 2nd edition, Prentice-Hall, Upper Saddle River, N. J, 1999.

[9] Haykin, S, *Neural Networks: A Comprehensive Foundation,* Macmillan College Publishing company, 866 Third Ave. N. Y., 1994.

[10] Howard Demuth and Mark Beale, *Neural Network Toolbox for use with Matlab*, The Math Works Inc, User's Guide, Natick, USA, 2000.