Building Simulation 2005

Ninth International IBPSA Conference
Montréal, Canada
August 15-18, 2005

DEVELOPMENT OF SINGLE-ZONE PREDICTIVE EQUATIONS USING LINEAR REGRESSION FOR ADVANCED CONTROLLERS SYNTHESIS

Roberto Z. Freire[1], Gustavo H. C. Oliveira[1] and Nathan Mendes[2]
[1]Industrial and Systems Engineering Graduate Program – PPGEPS
[2]Thermal Systems Laboratory – LST
Pontifical Catholic University of Paraná
Curitiba/PR/Brazil - Zip Code 80215-901
e-mail: {roberto.freire;gustavo.oliveira;nathan.mendes}@pucpr.br

ABSTRACT

Regression equations can be used for predicting temperature and relative humidity in single-zone buildings. This paper presents a methodology for computing the parameters of such kind of models based on weather data and room air hygrotermal conditions. In addition, the computed model can be very effective on the synthesis of advanced controllers applied to HVAC systems. A building hygrothermal simulation tool has been used to generate the data (heating signal, outdoor temperature and relative humidity, total solar radiation, indoor temperature and relative humidity). Validation procedures have shown very good agreement between the regression model and the simulation tool for winter and summer periods in Brazilian cities.

INTRODUCTION

Since the world-wide energy crisis in the 70's, building energy simulation programs have been developed in USA and Europe in order to reduce the energy consumption especially on HVAC (Heating, Ventilation and Air Conditioning) systems. In general, the main objectives of obtaining models for thermal analysis in offices, residential buildings and shopping malls are: i) estipulate better indoor climate conditions for the occupants; ii) avoid the energy waste to decrease the HVAC equipment operating cost and iii) simulate buildings interacting with HVAC equipment.

A classification usually found in the literature is the analytical, semi-empirical or empirical models. The analytical models depend on many parameters and are based on parametric equations. In a similar way, the semi-empirical models depend on test data and less parameters, combining analytical and empirical calculations. Simulation programs, such as DOMUS (Mendes et al., 2003a) and ASTECCA (Mendes et al., 2003b), have analytical and empirical models for HVAC systems.

Regression equations are empirical models widely used on identification process where a mathematical model is required. In thermal systems, a linear regression can be used on the identification process and functional relationships between the variables are provided.

Over the last decades, a large number of models have been developed in order to understand building behaviors submitted to different climate conditions. Based on Givoni (1999) method, Givoni and Krüger (2004) have presented the results of the application of formulae to predict daily indoor temperatures in three monitored low-cost houses in Curitiba, Brazil. In (Papst and Lamberts, 2001), a thermal performance analysis based on Givoni's regression model has been presented. Thermal performance comparisons between three residential buildings have been made. Givoni and Krüger (2003) have presented predictive regression equations to evaluate maximum, average and minimum temperatures of specific houses occupied by different families.

Virk and Loveday (1994) have presented a model in which the multivariable stochastic identification technique is applied to a full-scale test room with dedicated heating, ventilation and air conditioning plant. This model could predict indoor temperature of a test-zone in cold climates.

Accurate models are extremely useful for a variety of purpose, namely prediction, control, reliability aspects and system management. Combining the HVAC and building envelope study, much more detailed analyses can be done. In this way, advanced control techniques, which can be used in HVAC context (Freire et al., 2005), are usually based on a model that describes the process output signal behavior related to the input signals. An example is the Model Predictive Control Technique, which control laws are based on models constructed with linear regressions.

The present paper is focused on a methodology, known as system identification, using linear regression, for a HVAC integrated to a building model. The aim is obtaining a model that describes the HVAC system and building behavior in a structure that is useful for advanced control law synthesis.

This paper describes in the next section the simulation tool used to produce data for the identification process. Then, the building identification method is described followed by the model estimation, validation. Finally, an application example of such methodology for advanced control law synthesis is discussed and conclusions about this research are addressed.

SIMULATION TOOL AND ENVIRONMENT

The building simulation tool used in the identification procedure is the ASTECCA software (Mendes et al., 2003b). ASTECCA is a computational environment based on the Matlab/Simulink for building hygrothermal performance analysis with automatic control. This simulation tool contains mathematical models for buildings, HVAC (Heating, Ventilation and Air Conditioning) systems, sensors, weather data and control algorithms.

The toolbox ASTECCA used due to its user-friendly interface provided by the Matlab/Simulink features and also by the fast configuration of inputs, outputs and parts of the different subsystems included in the building and equipments, such as HVAC systems.

ASTECCA has been developed using analytical and semi-empirical models. The analytical part determines the envelope effects in simulations, e.g., heat transfer between surfaces of the zone. On the other hand, semi-empirical models describe HVAC systems based on collected data and some equations, which parameters needed to be calculated according to its configuration before simulation. Figure 1 shows the environment thermodynamic model used to generate data. The building mathematical model is described in terms of non-linear state-space based equations, with a lumped approach for the room air governing equations – energy and mass balances.

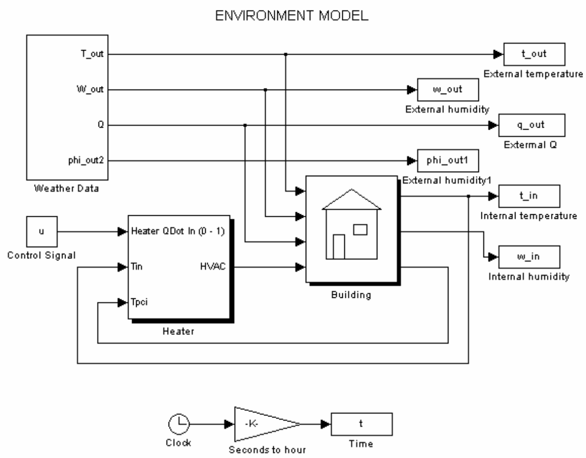

Figure 1. ASTECCA – Environment model used to generate data.

By using the ASTECCA environment, it is possible to verify the main signals behavior related to the building, e.g., indoor temperature and relative humidity, for different kinds of inputs. These inputs can be used all together in the simulation or individually. They are, typically, the outdoor climate variables and the HVAC control signal.

In the ASTECCA environment, the building simulation can be performed by using all input signals simultaneously or individually. The latter characteristic is very useful when the analysis of each perturbation needs to be accomplished.

The building is defined as follows. Assume a room with $5.0m \times 5.0m \times 2.5m$ of length, width and height respectively. Walls are configured in 3 layers, mortar ($0.02m$ of thickness), brick ($0.10m$) and mortar ($0.02m$). The basic dry-basis material properties are given in Table 1. The external and internal convection coefficients have been fixed at 5 W/m^2K. The reflectance of the ground in front of the wall has been ignored and 0.35 solar absorptivity of the external surface of the wall has been used. The maximum power of the heating system has been set to 1.5 kW. An external ventilation rate of 0.29 *ach* has been considered.

Table 1
Thermophysical properties

PROPERTY	BRICK	MORTAR
ρ_0 [kg/m^3]	1900	2050
λ [W/m-K]	1.11	1.96
c [J/kg-K]	920	950

The input signals used on the identification procedure are the climate variables, that is, outdoor temperature, outdoor relative humidity and total solar radiation, and the control signal applied to the heating system. The output signals are the indoor temperature and relative humidity.

IDENTIFICATION METHOD

The system identification is a theory where models are constructed from observed data. In the black box system identification approach, a pair of input/output data is collected from the system and, by means of an optimization procedure, the best model that fits the collected data are computed. When a single pair of input/output data is used, one obtains a SISO (Single-Input-Single-Output) identification procedure and if more than one kind of input and output data are involved, a MIMO (Multiple-Input-Multiple-Output) procedure is performed.

A thermal system as an indoor environment could be defined as a MIMO (Multiple-Input-Multiple-Output) system. It could be divided in subsystems that relate all the input variables to each output, i.e., the indoor temperature and relative humidity. The input signals used on the identification process are heating power, outdoor temperature, outdoor relative humidity and total solar radiation.

The MIMO system identification procedure follows the flow depicted in Figure 2 (Ljung, 1999) (Johansson, 1993). This procedure, in the context of the present paper, is discussed as follows.

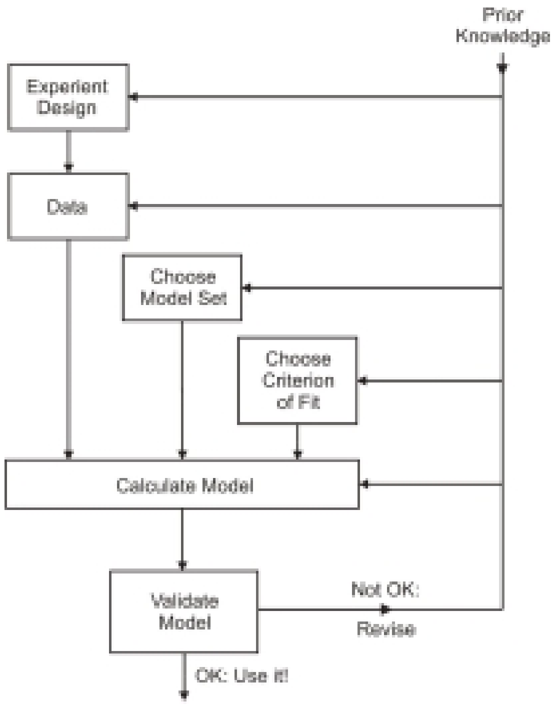

Figure 2. The system identification loop.

Experiment Design

First, the pairs of input and output data are collected. Data used on the identification process are obtained by applying a TRY (Test Reference Year) weather data in the ASTECCA simulation tool together with a random signal in the HVAC equipment and by measuring their effects on the indoor hygrothermal conditions.

In this case that we do not know the values of the parameters in θ, but that we have recorded inputs and outputs over a time interval $1 \leq k \leq N$ from software ASTECCA as shown on equation below:

$$Z^N = \{u(1), T_{EXT}(1), H_{EXT}(1), S_{EXT}(1), y(1), ..., u(N), T_{EXT}(N), H_{EXT}(N), S_{EXT}(N), y(N)\} \quad (1)$$

Y(k) can be $Y_H(k)$ in the relative humidity case (in *%*) or $Y_T(k)$ in the temperature case (in *°C*). $T_{EXT}(k)$, $H_{EXT}(k)$ and $S_{EXT}(k)$ are the outdoor temperature (in *°C*), outdoor relative humidity (in *%*) and total solar radiation (in *W/m²*) respectively. *u(k)* is the input signal applied on the HVAC system that varies from 0 to 1, representing the turn off and turn on states, respectively.

Model Structure Choice

Second, a model structure is selected. Here it has been assumed a couple of linear MISO (Multiple-Input-Single-Output) ARX (Auto Regressive Exogenous) models, since two outputs are involved, i.e., the indoor temperature and relative humidity, as follows:

$$\begin{aligned} & y(k) + a_1 y(k-1) + ... + a_n y(k-n) = \\ & + b_1 u(k-1) + ... + b_m u(k-m) \\ & + c_1 T_{EXT}(k-1) + ... + c_p T_{EXT}(k-p) \\ & + d_1 H_{EXT}(k-1) + ... + d_r H_{EXT}(k-r) \\ & + e_1 S_{EXT}(k-1) + ... + e_s S_{EXT}(k-s) + \zeta(k) \end{aligned} \quad (2)$$

where k is the discrete time instant, $k \in \Re$. $\zeta(k)$ is a non correlated random sequence, zero mean and variance σ.

For more compact notation, the following vectors are introduced:

$$\theta = \begin{bmatrix} a_1 \\ \vdots \\ a_n \\ b_1 \\ \vdots \\ b_m \\ c_1 \\ \vdots \\ c_p \\ d_1 \\ \vdots \\ d_r \\ e_1 \\ \vdots \\ e_s \end{bmatrix} \qquad \varphi(k) = \begin{bmatrix} -y(k-1) \\ \vdots \\ -y(k-n) \\ u(k-1) \\ \vdots \\ u(k-m) \\ T_{EXT}(k-1) \\ \vdots \\ T_{EXT}(k-p) \\ H_{EXT}(k-1) \\ \vdots \\ H_{EXT}(k-r) \\ S_{EXT}(k-1) \\ \vdots \\ S_{EXT}(k-s) \end{bmatrix} \quad (3)$$

where θ is the parameter vector and $\varphi(k)$ is the data vector. Therefore, for each output signal, Equation (2) can be rewritten as the following linear regression equation:

$$y(k) = \varphi^T(k)\theta + \zeta(k) \quad (4)$$

Model structures such as (Equation 4) that are linear in θ are known as linear regressions. The best prediction of the output at time instant *k*, i.e., $\hat{y}(k \mid \theta)$, can be computed by the following prediction equation:

$$\hat{y}(k \mid \theta) = \varphi^T(k)\theta \quad (5)$$

Parameter Computation

Third, an approach is then to select θ, in Equation (1) through (5) so as fit the calculated values $\hat{y}(k\,|\,\theta)$ as well as possible to the measured outputs by the least squares method:

$$\min_{\theta} V_N(\theta, Z^N) \tag{6}$$

where

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{k=1}^{N}\left(y(k) - \varphi^T(k)\theta\right)^2 \tag{7}$$

We shall denote the value of θ that minimizes equation (7) by $\hat{\theta}_N$:

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta, Z^N) \tag{8}$$

Since V_N is quadratic in θ, we can find the minimum value easily by setting the derivative to zero:

$$0 = \frac{d}{d\theta}V_N(\theta, Z^N) = \frac{2}{N}\sum_{k=1}^{N}\left(\varphi(k)y(k) - \varphi^T(k)\theta\right) \tag{9}$$

which gives

$$\hat{\theta}_N = \left[\sum_{k=1}^{N}\varphi(k)\varphi^T(k)\right]^{-1}\sum_{k=1}^{N}\varphi(k)y(k) \tag{10}$$

Once the vectors $\varphi(k)$ are defined, the solution can be easily found by modern numerical software.

Model Validation

Finally, it is important to perform the model validation exercise. The purpose is to verify if the estimated model fulfills the approximation requirements for a certain application. Usually the major objective is to obtain a minimal complexity model that lies in such requirements.

To deal with such a problem, different pairs of data are generated. One of them is used in the model parameters computations (by using Equation 9) and the others to validate the model, that is, to check if the model can reproduce the system behavior in different conditions from the one that it was computed.

Several tests can be used in model validation, one of them is to quantify the model error, by means of the MSE (Mean Square Error) test. The MSE test is given by:

$$MSE_{T,H} = \frac{\sum_{k=1}^{N}\left(Y_{T,H}(k) - \hat{Y}_{T,H}(k)\right)^2}{N} \tag{11}$$

The residual analysis comprises tests over the difference between the actual and the predicted data (the residuals) to check the presences of unmodeled dynamics. A common test in such a context is the autocorrelation one. If the process and noise model estimates correctly, it can be shown that the residuals autocorrelation function is an impulse signal (Billings and Voon, 1986; Johansson, 1993).

MODEL ESTIMATION

In the identification process, 70000 ("N" on Eq. 1) pairs of input/output samples have been used, for a 12-day sample period. Data have been collected between day 2 and day 12 in order to reduce initial condition effects. This procedure is necessary to assure good response accuracy under the presence of different weather conditions.

Two regression models (1st – indoor air temperature and 2nd – indoor relative humidity) have been obtained by using a MISO identification process and are described by Equations (12) and (13):

$$\begin{aligned}
&(1 - 2.954\,q^{-1} + 2.909\,q^{-2} - 0.9546\,q^{-3})Y_T(k) = \\
&+ (475.6\times10^{-6}q^{-1} - 7.001\times10^{-6}q^{-2} \\
&- 468.1\times10^{-6}q^{-3})u(k) \\
&+ (1.151\times10^{-3}q^{-1} - 2.3\times10^{-3}q^{-2} \\
&+ 1.149\times10^{-3}q^{-3})T_{EXT}(k) \\
&+ (-4.855\times10^{-6}q^{-1} + 8.385\times10^{-6}q^{-2} \\
&- 3.515\times10^{-6}q^{-3})H_{EXT}(k) \\
&+ (-1.842\times10^{-9}q^{-1} + 3.514\times10^{-9}q^{-2} \\
&+ 1.912\times10^{-9}q^{-3})S_{EXT}(k)
\end{aligned} \tag{12}$$

$$\begin{aligned}
&(1 - 2.955\,q^{-1} + 2.910\,q^{-2} - 0.9552\,q^{-3})Y_H(k) = \\
&+ (-1.402\times10^{-5}q^{-1} + 2.108\times10^{-7}q^{-2} \\
&+ 1.379\times10^{-5}q^{-3})u(k) \\
&+ (-11.81\times10^{-5}q^{-1} + 23.7\times10^{-5}q^{-2} \\
&- 11.89\times10^{-5}q^{-3})T_{EXT}(k) \\
&+ (9.782\times10^{-6}q^{-1} - 4.987\times10^{-8}q^{-2} \\
&- 9.686\times10^{-6}q^{-3})H_{EXT}(k) \\
&+ (3.425\times10^{-10}q^{-1} - 7.868\times10^{-13}q^{-2} \\
&- 3.436\times10^{-10}q^{-3})S_{EXT}(k)
\end{aligned} \tag{13}$$

All input signals used on the model identification process comes from the TRY weather file of Curitiba, Brazil, during the period between January 1st and January 12th. Figure 3 shows the external data considered for Curitiba that has been applied to the model. Initial conditions of 20°*C* for the indoor air temperature and of 62.8% for the indoor relative humidity have been set.

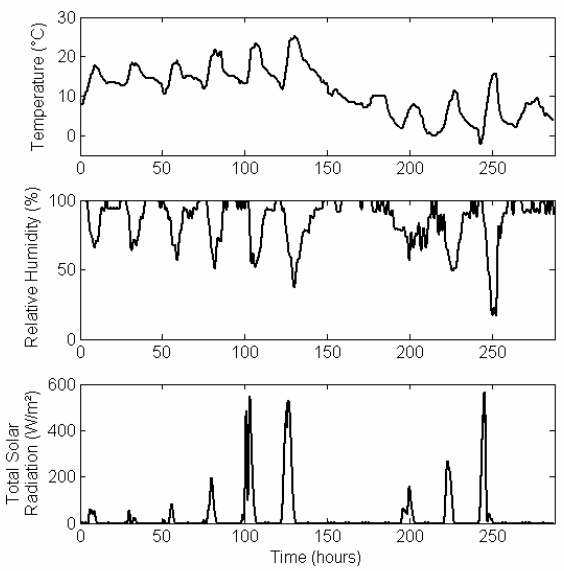

Figure 3. External temperature, relative humidity and total solar radiation for the simulation period in Curitiba – Brazil (July, 1st – 12th).

The MSE for Equations (12) and (13) are presented below:

$$MSE_T = 2.162\times10^{-7} \tag{14}$$

$$MSE_H = 3.244\times10^{-10} \tag{15}$$

The autocorrelation function of the temperature signal prediction (estimation data for Curitiba, Brazil, during the winter) residuals is shown in Figure 4. It can be noticed it represents impulse like signal, meaning that the estimate is adequate.

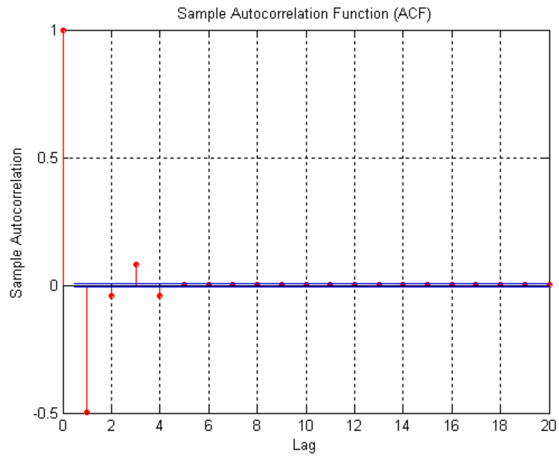

Figure 4. Autocorrelation function of the temperature signal prediction residuals (Curitiba, during the winter).

MODEL VALIDATION

The validation procedure starts applying a different weather data (January, 1st – 12th Curitiba TRY weather file) rather than the one used in the identification process (July, 1st – 12th Curitiba TRY weather file, Fig. 3), in order to verify the system behavior in different conditions. This assures that the identified system can be generalized and provide a good response even for other input signals. Therefore, to validate the obtained model, a simulation procedure has been carried out comparing results in terms of room air temperature and relative humidity using the building simulation software ASTECCA and the responses from the identified model (Eqs. 12 – 13).

First the TRY weather data file for Curitiba has been used for the 12-day winter period and a sample time of *10* seconds. In addition, an ON-OFF controller has been considered, which set-point temperature was 26°*C* ± 2°*C* actuating during the 10 last days. Initial conditions of 20°*C* for the indoor air temperature and of 62.8% for the indoor relative humidity have been applied to both models.

The results presented in Figure 5 show a very good agreement between the identified model and the ASTECCA building simulation tool. The control signal presented in Figure 5 has been multiplied by the heater power, which is 1.5 kW.

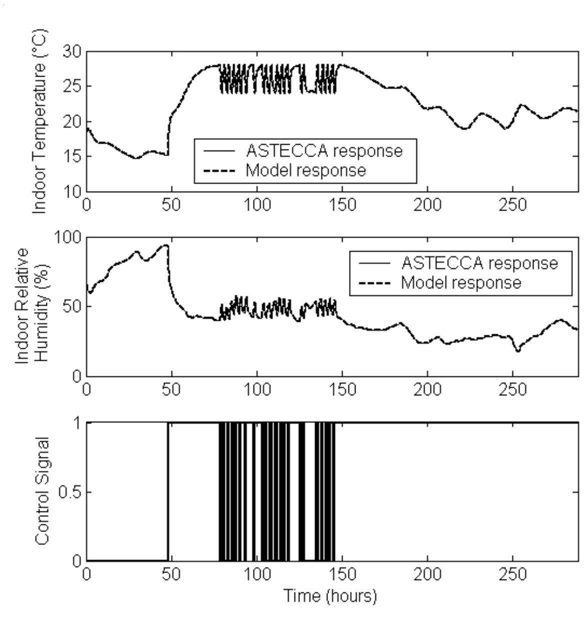

Figure 5. Indoor air temperature, relative humidity - comparison between ASTECCA simulation and model response for Curitiba – Brazil (July, 1^{st} – 12^{th}).

As shown above, the model has been identified by using as input signal a winter period of Curitiba, which is the coldest capital in Brazil. Despite the good results presented in Figure 5, simulation results for two hot and humid Brazilian cities are shown in order to confirm the model accuracy for different weather conditions.

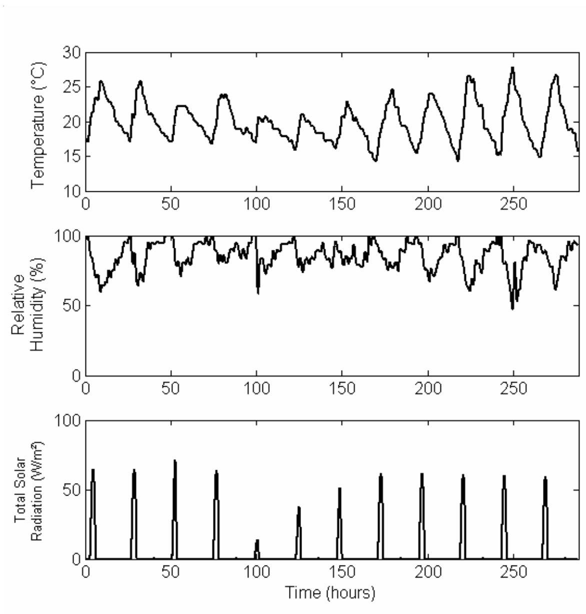

Figure 6. External temperature, relative humidity and total solar radiation for the simulation period in Rio de Janeiro – Brazil.

In this way, the 12 first days of January, using TRY weather files of Rio de Janeiro and Belém (Figures 6 and 7), have been considered to compare the identified model and the simulation tool, using the same parameters presented before, except for the heater signal which has been set 0. In this way, the model response has been obtained as a pure function of outdoor conditions (temperature, humidity and global solar radiation).

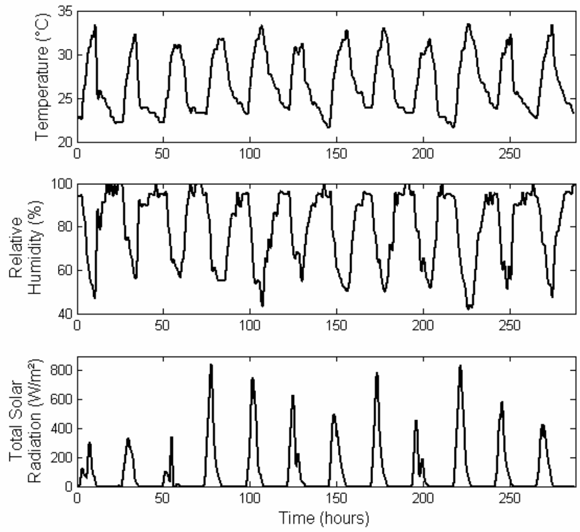

Figure 7. External temperature, relative humidity and total solar radiation for the simulation period in Belém – Brazil.

Figures 8 and 9 illustrate the results in terms of indoor air temperature and relative humidity, using both the ASTECCA simulation tool and the predictive equations. These results, for Rio de Janeiro and Belém, show that the use of the predictive equations provide very precise results regardless the weather data file, unless temperatures go below *0°C* as frosting/melting phenomenon was neglected for this analysis in Brazil. Table 2 shows the MSE values for both cities.

Table 2
Mean Square Errors

	RIO	**BELEM**
Temperature	1.886×10^{-7}	3.208×10^{-10}
Relative humidity	1.772×10^{-7}	3.476×10^{-10}

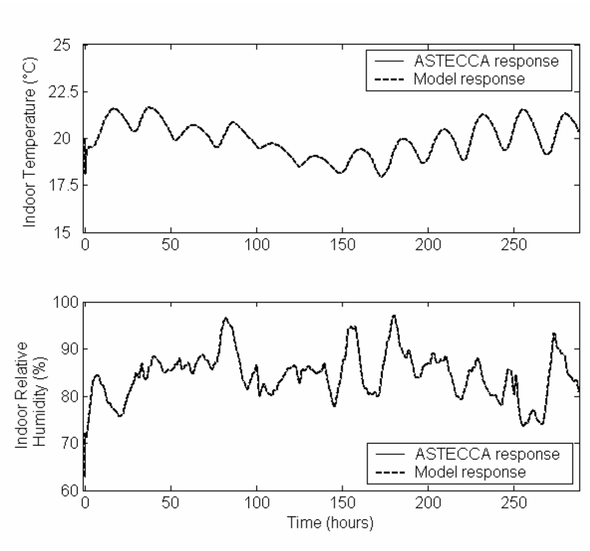

Figure 8. Indoor air temperature and relative humidity - comparison between ASTECCA simulation and model response for Rio de Janeiro – Brazil.

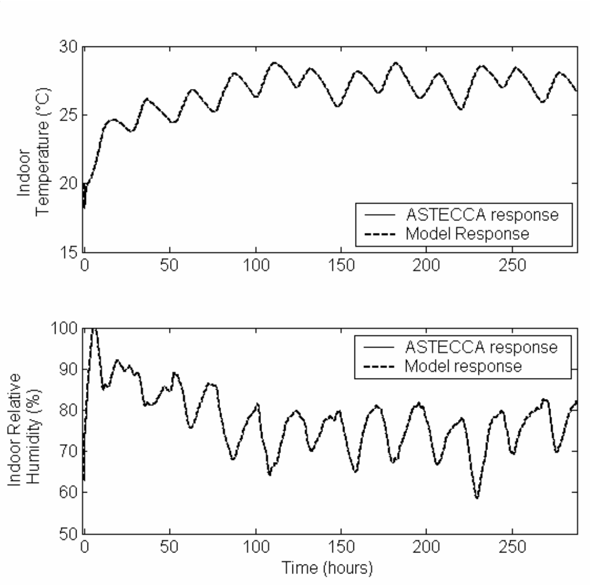

Figure 9. Indoor air temperature, relative humidity - comparison between ASTECCA simulation and model response for Belém – Brazil.

The autocorrelation function of the temperature signal prediction (validation data for Rio de Janeiro, Brazil, during the summer) residuals is shown in Figure 10. It can be noticed it represents impulse like signal, meaning that the estimate is satisfactory.

APLICATION EXAMPLE

Control Law Synthesis

The methodology for developing building thermal models presented in the previous sections can be applied in the synthesis of advanced controllers, for instance the MBPC (Model Based Predictive Control) algorithm.

Model predictive controllers are defined by the following steps: *(i)* first, a model is used to compute the predicted process output. *(ii)* Second, a cost function describes the closed-loop performance of the system and *(iii)* then this cost function is minimized in relation to the future control signals. Finally, *(iv)* the first of these control signals is applied to the process (receding horizon strategy). Several MBPC algorithms have been proposed based on this scheme and the main difference between them is the strategy used in each step described above (Clarke, 1994).

The building modeling strategy present here is related with the first step of MBPC algorithm. That is, in order to control the building indoor temperature and humidity conditions, it is necessary to supply the MBPC algorithm with a building model which has the same structure as the one illustrated by equations (12) and (13). Therefore, this is an interesting application of the methodology presented in the previous sections.

Moreover, a thermal comfort control law synthesis has been successfully tested by Freire *et al.* (2005), using two MISO (Multiple-Input-Single-Output) models, similar to the ones given by equations (12) and (13). The algorithm presented in their work uses two building equations to predict the PMV (Predicted Mean Vote) (Fanger, 1974) and, according to the predicted results, the control law can generate the signal to the HVAC device that optimizes the thermal comfort.

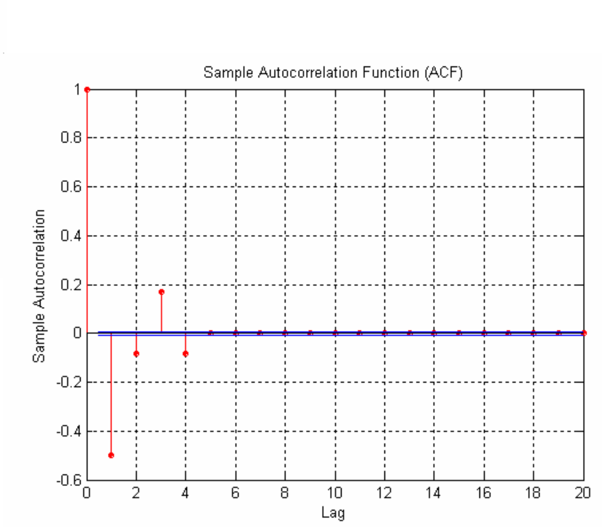

Figure 10. Autocorrelation function of the temperature signal prediction residuals. (Rio de Janeiro, during the summer).

CONCLUSION

This paper has described the development of predictive equations based on an identification process, using linear regression with collected data from simulations performed with a building simulation tool. Four input disturbances - heating power signal, external temperature, relative humidity and total solar radiation have been used to identify two models; one for predicting indoor air temperature and the other one for predicting indoor air relative humidity.

The model has been identified for a winter period of the city of Curitiba and results in terms of room air temperature and relative humidity have shown the model high accuracy for three simulation samples, using the TRY weather files of three Brazilian cities: Curitiba, Rio de Janeiro and Belém.

Further research will be focused on the use of the developed predictive equations to the synthesis of advanced control strategies for HVAC systems.

ACKNOWLEDGMENTS

The authors thank CAPES – *Coodenação de Aperfeiçoamento de Pessoal de Nível Superior* - of the Ministery of Education of Brazil for support of this work.

REFERENCES

Billings, S. A., Voon, W. S. F. 1986. Correlation Based Model Validity Tests for Non-Linear Models. International Journal of Control. Vol. 44. N° 1. pp. 235 – 244.

Clarke, D. W. 1994. Advances in Model Based Predictive Control. Oxford University Press.

Fanger, P. O. 1974. Thermal Comfort. McGraw-Hill Inc.. New York, USA.

Freire, R. Z., Oliveira, G. H. C., Mendes, N., 2005. Thermal Comfort Based Predictive Controllers for Building Heating Systems. Proc. of the 16th IFAC World Congress (IFAC'05), Prague, Czech Republic. (In press).

Givoni, B., 1999. Minimum Climatic Information Needed to Predict Performance of Passive Buildings in Hot Climates. Proc. of the Sixteenth International Passive and Low Energy Architecture Conference (PLEA'99), Brisbane, Australia. pp. 197 – 202.

Givoni, B., Krüger, E. L., 2003. An Attempt to Base Prediction of Indoor Temperatures of Occupied Houses on their Thermo-Physical Properties. Proc. of the Eighteenth International Passive and Low Energy Architecture Conference (PLEA'03), Santiago, Chile.

Johansson, R., 1993. System Modeling and Identification. Prentice Hall.

Krüger, E. L., Givoni, B., 2004 Predicting thermal performance in occupied dwellings. Energy and Buildings, Elsevier-Ireland, v. 36, n. 3, pp. 301 – 307.

Ljung, L., 1999. System Identification: Theory for the User. Prentice-Hall PTR 2nd Edition.

Mendes, N., Oliveira, R. C. L. F., Santos G. H., 2003a. Domus 2.0: A Whole-Building Hygrothermal Simulation Program. Proc. of the Eighth International Conference on Building Performance Simulation (IBPSA'03), Eindhoven The Netherlands.

Mendes, N., Oliveira, G. H. C., Araújo, H. X. and Coelho, L. S., 2003b. A Matlab-Based Simulation Tool for Building Thermal Performance Analysis. Proc. of the Eighth International Conference on Building Performance Simulation (IBPSA'03), Eindhoven The Netherlands.

Papst, A. L., Lamberts, R., 2001. Thermal Performance Evaluation of Three Houses in Florianópolis, South of Brazil. Proc. of the Eighteenth International Conference on Passive and Low Energy Architecture (PLEA'01), Florianópolis, Brasil. pp. 293 – 297.

Virk, G. S., Loveday, D. L., 1994. Model-Based Control for HVAC Applications. Proc. of the Third IEEE Conference on Control Applications, Glasgow. pp. 1861 – 1866.