Application of Dynamic Building Inverse Models to Three Occupied Residences Monitored Non-Intrusively

T.A. Reddy, Ph.D.

ABSTRACT

Three occupied residential houses were monitored in a non-intrusive manner during the fall of 1987 and periods identified during which the houses were under free-floating mode. These then served as the common data base on which three different dynamic inverse model formulation approaches, namely the thermal network, time series, and time derivative models, were subsequently assessed in terms of their ability to yield sound building envelope physical parameters and to the related effort involved in the identification process. For all three models, the internal air temperature was taken as the response variable while ambient air temperature, solar radiation and whole-house electricity consumption (which in all three houses was deemed representative of the internal heat generation) were assumed to be the independent variables. Four different thermal network configurations were tried out on the data in terms of parameter identifiability and consistency. In the framework of the time series and the time derivative models, the effects of choosing different orders and alternative statistical procedures are discussed. The predictive power, i.e., the tracking ability of the various models and their variants is also assessed. Despite the fact that parameter estimates are generally difficult to identify from a measurement protocol such as the present, we find the time series approach to be promising. Consequently, in terms of practical applications, we speculate that intelligent controls could be designed to incorporate algorithms based on this formulation, thereby augmenting on-line modeling and control capabilities of energy use in buildings.

INTRODUCTION

An important tool for energy-related studies is a model that is able to predict the heating and cooling energy requirements of a building based on its physical description, the weather, and the usage behavior of the resident. The conventional wisdom that as one included more and more detail in modeling the physical phenomena one could get closer to reality, led to the development of several large computer models and algorithms of simulation (for example, DOE [1981]). There is no arguing the fact that these are essential for design purposes, i.e., before the building is actually built. This includes not only designing the building and the equipment but also performing sensitivity studies in order to estimate and optimize long-term energy use. However, when researchers tried to apply these computer codes to existing buildings in an effort to understand and subsequently decrease this energy consumption, inadmissible differences in energy use between the predicted and observed values were frequently found. These differences were difficult, if not impossible, to account for or to systematically include in the computer model, since one was faced with the formidable task of trying to decide which of the numerous input values to modify and by how much (Hsieh et al. 1989).

The above quagmire could be attributed to a number of causes: differences between actual construction and that specified by the blue-prints; improper physical understanding of phenomena such as infiltration, building and ground thermal interaction, thermal bridges, or convective loops inside walls (Nisson and Dutt 1985; Kasuda 1985); or simply due to resident behavior. In an effort to reconcile such differences, attempts have been made to adapt a discipline called system parameter identification (colloquially referred to as the inverse approach) to the realm of energy use in buildings. This approach as applied to a physical system is basically a technique whereby a reduced, i.e., an aggregated or macroscopic, system model (albeit simplified) is identified from performance records of the system while in actual operation (Sinha and Kuszta 1983).

Work on dynamic inverse model formulations has been going on for almost a decade. A recent paper by Rabl (1988) has classified the various types of dynamic inverse model formulation approaches to predict building energy use, laid down the basic theory, and drawn attention to the common underlying features of these models and how the building physical parameters could be deduced from the coefficients of the regression terms. The principal reason for undertaking this particular study was to assess the strengths and weaknesses of the various models and how they fare when applied to actual measured data from three different residential houses in the Princeton, NJ, area. The data have been collected under "lived-in" conditions as non-intrusively as possible during the swing seasons of the year. These correspond to situations where parameter estimation is likely to be most uncertain given that the signals will be relatively "weak," while being "noisy" at the same time.

DESCRIPTION OF RESIDENCES AND INSTRUMENTATION (Norford et al. 1987)

All three are ranch-style houses of frame construction, located in the Princeton area, NJ, and all are about 30 years old. They could be considered typical American houses, i.e., probably of lightweight construction, no planned solar strategies, common landscaping, and not super-insulated. Also all three have central air-conditioning. Moreover, all these houses have no major source of internal heat generation other than that used by the electrical appliances since all three have electric cooking ranges. However, House 1 does have an oil-fired water heater.

The instrumentation packages of all the houses were not identical but instead included some sensors unique to each house. Each house was outfitted with a microcomputer and data-logging hardware; Houses 1 and 2 had powerline carrier systems and House 3 had a hardwire system. The powerline carrier in House 1 was marginally satisfactory due to electrical interference on the house wiring, while the systems in Houses 2 and 3 were quite reliable. Each house also received at least two indoor temperature sensors, an outside temperature sensor, and a pulse-generating whole-house utility meter, all of which were connected to the data-logging hardware. Hourly average temperatures and total electricity use were recorded for all three houses; for Houses 1 and 2, the latter was done every 10 minutes. Solar radiation data on a horizontal surface were measured nearby at one location only.

Periods of data when the air-conditioner system was not being operated were identified for all three houses. These fell essentially during August and September and corresponded to periods during which the houses were under free-floating mode. Intervals during which windows were known to be open or fans operated were also removed. This finally resulted in 290 hourly observations for House 1, 562 for House 2, and 486 for House 3.

For purposes of this study, the internal air temperature (T_i) was taken s the response variable while ambient dry-bulb air temperature (T_a), solar radiation (Q_S) (in kWh/m^2 measured on a horizontal surface) and whole-house electricity consumption (i.e., the internal energy Q_A in kWh/h) were assumed to be the regressor variables.

Though it relates to only a portion of the data for House 3, Figure 1 illustrates the type of variation in T_i, T_a, Q_S and Q_A encountered during the present study. We found that for all three houses, differences between T_i and T_a were on the order of 6^o to 8^oC on the average, which, considering the values of the overall heat loss coefficients found (discussed later in this paper), resulted in the heat loss through the building shell to be the most important source of heat flow. Except for the solar loads, which were small, we think that all other heat flows were generally large enough not to introduce excessive problems related to ill-determined parameter identification.

It should be pointed out that by choosing the swing seasons one can overcome the added complexity in having to include the effect of the heating or cooling device in the modeling equations. On the other hand, it is generally difficult to accurately predict building energy use during these periods. Small differences between ambient and internal air temperatures could very easily lead to large instabilities in the system since small loads (both internal and external), not well understood from the available data, could have a large impact. Moreover, the type of measurement protocol followed enhances the instability associated with parameter estimation since, while being more representative of the "real lived-in" environment of the resident, it does not enable controlled experiments to be performed (wherein one could artificially induce extreme situations of building operation that enable more accurate and robust parameter identification).

One of the objectives of this study was to evaluate the extent to which the various inverse models are appropriate under such a measurement protocol (as would be the case when on-line identification is done). Stated otherwise, there is bound to be noise in all modeling approaches, but which of these nevertheless yield the most stable parameter estimates?

BRIEF DESCRIPTION OF INVERSE MODEL FORMULATION APPROACHES

The inverse dynamic model formulation approaches appropriate for building dynamic energy use have been classified and described by Rabl (1988). We have adopted this classification and will limit ourselves to a brief review and presentation of only those formulae relevant to our subsequent analysis. The various model formulation approaches could essentially be viewed as belonging to one of two possible groups: analogue models or gray-box models.

Analogue Models

These basically involve approximating the building in terms of a thermal network using the electric analogue. The basic set of differential equations describing the dynamic behavior of the network is then laid down and manipulated so that it is recast into a form suitable for regression against actual data. Although multi-zones in buildings can be treated, the parameter identification process (which is basically a statistical regression) gets progressively poorer as more parameters are included in the model. In this study, we will limit our discussion to single-zone buildings only. Depending on how one goes about manipulating the set of equations, we can distinguish two approaches: modal analysis (Carter and de Villiers 1987; Neveu et al. 1986) and the equivalent thermal parameter method (ETP). Only the latter will be addressed in the present study.

The ETP method (Sonderegger 1977a) normally requires that the basic set of differential equations be manipulated such that nodes whose temperatures are not being explicitly measured are removed by substitution from the final set of equations. This is then discretized using an appropriate finite-difference scheme prior to regressing against actual data. This manipulation is often tedious and only simplified networks containing a few nodes can be treated with confidence. Moreover, one has to guard against the problem of overdetermined (i.e., the number of regression coefficients is more than the number of physical parameters of the network) and underdetermined cases. In the latter case, one is forced to simplify the network, while a constrained regression using Lagrangian multipliers (Stocker 1980) is the proper way of dealing with the former. Despite the effort entailed, we have selected this approach for evaluation. This is primarily because the network representation is widely understood and appreciated not only among those involved in building energy modeling but also by practitioners in the field.

Gray-Box Models

We have suggested the use of the terminology gray-box as against the more conventional appellation black-box (Mikradakis et al. 1983) in order to stress the distinction that the latter is one whose internal characteristics are completely unknown and so "opaque" that no physical interpretation can be assigned to the regression coefficients of the inverse model(as is usually the case in the non-physical sciences). The gray-box models, on the other hand, are defined as those that have enough internal constraints based on physical laws built into them that coefficients determined by the inverse model permit explicit identification of the building system parameters.

There are basically two different model formulation approaches of gray-box models: the time series and the time derivative models, both of which assume that most phenomena influencing the thermal response of the building are linear in the driving terms. Note that though this assumption is not rigorously correct due to effects such as infiltration, shading, and night insulation, the assumption of linearity is, in most instances, probably a good one (Rabl 1988).

The Time Series Model. Also referred to as the "transfer function" method, following Box and Jenkins (1976), the time series model assumes a linear functional relationship between the present and lagged values of the response variable and those of the driving terms. Since it combines regression analysis in a time-series framework it can be equally referred to as the MARMA (multivariable auto-regressive moving average) method (Mikradakis et al. 1983).

We shall refer to the order of the model as designating the number of lag terms of each of the four variables (T_i, T_a, Q_A, and Q_S) appearing in the model. Expressing them in their discretized form, i.e., temperatures as mean values and energies as sums over the time interval Δt (taken in this study as one hour), a 2222 model will imply:

$$T_i(n)-T_i(n-1) = a_2[T_i(n-1)-T_i(n-2)]$$
$$+ b_o[T_a(n)-T_i(n-1)]+b_1[T_a(n-1)-T_i(n-1)]$$
$$+ b_2[T_a(n-2)-T_i(n-1)]$$
$$+ C_oQ_A(n)+C_1Q_A(n-1)+C_2Q_A(n-2)$$
$$+ d_oQ_S(n)+d_1Q_S(n-1)+d_2Q_S(n-2) \qquad (1)$$

where <u>n</u> is an integer representing time increments such that $\underline{t} = \underline{n.\Delta t}$.

Expressions for the steady-state coefficients (L and A_S) and that of the time constant τ are assembled in Table 1.

<u>The Time Derivative Model</u>. This approach has been developed to its present form by Rabl (1988), though certain efforts have been made previously in this direction (for example, Hammerstein [1984]). It is based on a Taylor series expansion of the time series model and correlates the time derivatives of the response variable in a linear canonical form to those of the driving or exogenous variables. Subsequently, since data at discrete time intervals are available, a finite difference scheme is chosen to replace the derivatives by differences prior to the statistical analysis.

Limiting ourselves to the set of four variables used previously, the thermal performance of a building under free-floating mode, in the framework of a 1111 model, is given by:

$$T_i(t)-T_a(t) = -\alpha_1\mathring{T}_i(t) + \beta_1\ \mathring{T}_a(t)$$
$$+ \gamma_o\ Q_A(t) + \gamma_1\ \mathring{Q}_A(t) \qquad (2)$$
$$+ \delta_o\ Q_S(t) + \delta_1\ \mathring{Q}_S(t)$$

where $\underline{\dot{X}\ (t)}$ represents the first derivative of <u>X</u> with time <u>t</u>.

Expressions for the steady-state building physical parameters are also given in Table 1.

<u>Past Studies</u>

There exist in the published literature a number of studies that deal with the inverse problem associated with dynamic building energy use. In the framework of thermal networks, numerous studies have been reported, both with actual data (Sonderegger 1977b; Wilson et al. 1984; Achterbosch et al. 1985; Griffith 1985; Cleary 1987; Persily 1982) as well as with computer experiments (Taylor and Pratt 1988). There have also been several studies that have used the time series approach (Crawford and Woods 1985; Subbarao 1985; Van Hatten and Norlen 1987; Parkanen 1988) to both small residences as well as large office buildings.

There are only two studies which, to our knowledge, have attempted to apply different inverse model formulation approaches to the same data set: that of Hammerstein (1984), who monitored five occupied residential houses in Sweden for two months prior and two months after major retrofits to the building shells were made; and that by Norford et al. (1985) and Rabl (1988) to data monitored from a large commercial building near Princeton, NJ.

Perhaps the one group that has had the most experience with the inverse problem approach is the Buildings Energy Group (to name a few, Subbarao [1985, 1988] of SERI. Their general philosophy is to perform rigidly controlled experiments on a short-term basis and to determine the building parameters from a disaggregation of heat flows entering the energy balance. Audit descriptions are used to obtain initial estimates of the parameters, which are then modified or "tuned" in view of actual measured building performance. Such tests have been performed on test cells, actual residences, and commercial buildings.

ANALYSIS AND DISCUSSION

The ETP Approach

In the present study, we have chosen the simple 2R1C generic network and the four variants selected are shown in Figure 2. T_S represents the building shell temperature. The networks referred to as RC1, RC2, and RC3 represent internal storage-dominated structures, while RC4 corresponds to a shell-dominated structure. Note that the choice of RC4 is perhaps inappropriate since the prevalent thinking seems to be that most residences have low-mass building shells; this selection was done deliberately in order to gauge the sensitivity of the ETP model formulation approach as a whole. The only distinction between RC1 and RC2 is the effective location of the node at which solar radiation is assumed to impinge. While RC2 is identical to the configuration chosen by Hammerstein (1984), RC3 is the one selected by Sonderegger (1977).

Table 2 lists the pertinent regression equations and the constraints that ought to be performed to remove the effect of over-determination. In this entire study we chose not to over-burden the regression by imposing additional constraints on the model coefficients (for example, relating to the sign of the values) so as to make the building parameters physically plausible. These conditions are listed in Reddy (1989). Note the similarity of all four functional forms, especially those of RC1, RC2, and RC4. The effect of the clamp temperature (Sonderegger 1977) (probably representative of the influence of the basement) in RC3 not only introduces an intercept but also an extra term containing T_i not present in the other expressions.

Table 3 assembles the values of the building physical parameters as well as the corresponding values of the time constants. The following observations are made:

i) There is remarkably small variation in the adjusted coefficient of determination, $\bar{R}^2$, from one network to another for a particular house. In this regard, these results by themselves are incapable of suggesting not only which network is most suitable but also which generic type of network ought to be chosen.

ii) The estimated parameter sets differ from network to network, some parameters like C (or τ) exhibiting lesser variation, which is to be expected since certain parameters are better defined from the data than others.

iii) For House 3, all four parameter sets except for RC3 seem meaningful and close to each other, except for the value of H_2, which is large and probably ill-defined from data.

iv) For House 2, RC2 and RC4 parameter sets seem close (except for H_2).

v) For House 1, except for RC3, either H_1 or H_2 generally is negative (and therefore unphysical), while A_S, C, and τ seem fairly consistent.

vi) Network RC3 seems inappropriate for Houses 1 and 2 while its unsuitability for House 3 is less obvious.

vii) The fact that one of the parameters is negative does not necessarily imply that the particular network is unsuitable to the specific house: multicollinearity could be one reason. It could be that due to the nature of the data, certain parameter estimates become ill-defined and the estimates would then be incorrect; a physical parameter turning out to be negative is a blatant manifestation of this effect. We noticed that for House 1 the term $\dot{Q}_A$ was very small, and, consequently (though this is not entirely proper) we redid the regression using the equations for RC1 while omitting the $\dot{Q}_A$ term. The parameter estimates, listed under RC1b in Table 3, now seem reasonable.

Figure 3 illustrates how well the various models track the measured values of T_i for House 2. In order to quantify the tracking ability, RMSE values for all three houses have also been computed; they are assembled in Table 4. RMSE values over a period of a week to 10 days are of the order of 1^oC or less. Although the parameter sets for each model were estimated based on the entire data set, we have chosen to include and discuss graphs illustrating the tracking ability during certain portions of the monitoring period only.

The most noteworthy observation is that the choice of one or another network does not seem to affect the overall tracking ability of the model and consequently the tracking ability alone does not seem to be a sufficiently sensitive criterion to enable grading or even discerning the more physically meaningful model among competing network configurations.

The Time Series Approach

A number of sets of building physical parameters estimated by stepwise regression using different orders of the models as well as variants in the procedure of statistical regression (explained below) are shown in Table 5.

The first three columns have been found by fitting all the terms of the function simultaneously. For all three houses, parameters from 2222 and 3333 models are close to each other, while the 1111 model shows more variation.

Figure 4 illustrates the predictive ability of the three models for House 3. The RMSE values are assembled in Table 6. We notice that the three models not only have more or less identical RMSE values, but are also close to one another. As noted previously vis-a-vis the ETP variants, the predictive ability does not seem to be a sensitive enough criterion to grade different identified parameter sets.

We have performed a preliminary stability check on the set of building parameters estimated, by which we imply studying the sensitivity of the regressed values if only certain portions of the data were used for parameter identification. There are basically two ways of doing this. One could simply truncate the data and determine the set of parameters for each of the two periods separately. We would then be looking at how these parameter sets vary seasonally and also how well the models predict outside the regression period. Although this is indeed what is ultimately of interest, the more basic analysis would be to simply break up the original data set into two subsets that do not have any inherent periodic trends in them and estimate the parameters from each subset separately. The break-up could be done in several ways, though; the one adopted here is to assign the first, third, fifth . . . i.e., all odd-numbered rows of hourly measurements to one sub-set and all even-numbered rows to another, thereby creating two subsets (which we shall call A and B). The results of the regression for House 2 using a 2222 model are included in the last column of Table 5 and show no bias with respect to the parameter estimates given in column 2. Though this does not constitute a complete test of stability, the results are indeed encouraging in that the two parameter sets are fairly close.

The Time Derivative Approach

According to Rabl (1988), the greatest advantage of the time derivative approach to the parameter identification process is that it allows different time scales to be mixed. For example, one can integrate Equation 2 from time t_2 to t_1, these times being chosen such that based on physical considerations certain terms drop out. For example:

i) (t_2-t_1) can be taken as 24 hours, in which case the terms capturing the transient behavior of the system may become very small. The terms drop out and we have the steady-state case.

ii) If t_2 is taken as 6 p.m. and t_1 as 6 a.m., the treatment of solar radiation is greatly simplified, since $\mathring{Q}_S$ can essentially be taken as zero and Q_S is easily determined from knowledge of the daily total radiation.

Thus the time derivative approach enables one to go from the steady state to the transient behavior in a systematic manner. One could proceed successively by first determining the steady-state parameters from longer time scales wherein the steady-state behavior of the building is more pronounced, and subsequently identifying the transient terms using hourly time intervals. The RMSE values of these two subsets (called 2222A and 2222B) are also shown in Table 6.

We have attempted to identify the building parameters using four different variants, the results of which are summarized in Table 7.

A. Steady-State Analysis. A 24-hour time interval was chosen spanning 6 a.m. to 6a.m. This is probably a better choice than the conventional midnight-to-midnight period since the former is a more natural period in terms of resident behavior cycles.

The steady-state heat loss coefficient and the effective solar aperture are listed under column A. The results are generally poor since both Houses 1 and 3 have negative A_S values and abnormally low L values. This could be attributed to having neglected the transient terms, i.e., proper steady state is not reached even over a 24-hour period. This is especially true of Houses 1 and 3. On the other hand, for House 2, which is a low mass building, the estimate of L seems right.

B. Using 12 Hourly Time Intervals: Successive Procedure. As discussed previously, we now choose time scales of 12 hours (6 a.m.-6 p.m. and 6 p.m.-6 a.m.). Next, we retain the coefficients of the steady state terms (i.e., γ_0 and δ_0) and using the data set pertaining to hourly time intervals perform a stepwise regression of the residuals vs. the remaining terms (which accounts for the transient behavior). The building parameters obtained are listed under column B of Table 7.

C. Using 12 Hourly Time Intervals: Stepwise Procedure. Using data corresponding to time scales of 12 hours, a stepwise regression, with the first derivative of the solar terms (i.e., δ_1 of Equation 2) set to zero, was done in order to identify the significant parameters. The parameters are listed under column C of Table 7.

D. Using Hourly Time Intervals. Equation 2 was assumed on an hourly basis, and a stepwise regression was performed, the results of which are given under column D of Table 7.

Figure 5 illustrates the tracking ability of variants B, C, and D for House 2 when used to predict building response (i.e., T_i) , while Table 8 gives corresponding tracking RMSE errors.

We can draw the following conclusions:

i) All three procedures seem to yield more or less similar sets of building parameters for House 2. However, for House 1 and House 3, except perhaps for L, both A_S and τ values are very different.

ii) Procedure D does not seem satisfactory. Values identified are unphysically small (for Houses 1 and 3), and the $\overline{R}^2$ values are appreciably lower while the RMSE values are larger. Moreover, its tracking ability is generally poorer. This does seem surprising, since Equation 2 is, in fact, identical to that used in RC2 of the ETP approach, which had yielded good results. This could perhaps be due to the fact that the degrees of freedom are different for both approaches: RC2 performed a constrained regression while the time derivative did not.

iii) It is difficult to choose the better procedure between B and C since neither is systematically superior; in terms of tracking ability and summary statistics they are similar. However, from Table 8 we note that, except for House 2, the corresponding physical parameter sets (L, τ, and A_S) are markedly different.

INTERCOMPARISON AND CONCLUDING REMARKS

The sets of building parameters for all three residences identified by the three different model formulation approaches that seem most physically realistic and consistent with each other are assembled in Table 9, while Figure 6 illustrates their tracking ability for House 2. We find that the standard errors of the parameter estimates are large: about 30% to 40% for L and τ and up to 100% for A_S. The latter is not surprising since the houses are not strongly influenced by solar radiation, while solar radiation is cross-correlated with T_i. The standard errors for the other parameters are larger than what one could obtain from controlled experiments on the house but this may not be excessively large for certain types of applications; for example, shaving of electric peaks by preconditioning the building.

While the selection of parameter sets was subjective (especially those of the ETP method), it nevertheless enables certain observations to be made. Both the ETP and time series approaches seem to be fairly consistent, both in the parameter estimates as well as in their tracking ability (about 1°C or less over a prediction period of one week to 10 days). On the other hand, the time derivative approach fares badly as compared to the other two approaches, despite its flexibility in allowing different time scales to be mixed.

In the framework of the present study, which is limited to only three houses as well as to the type of simplified non-intrusive experimental protocol followed, we draw the following broad conclusions:

a) The time derivative approach seems to have poor parameter identifying capability. Though this may have to do with the quality of measured data at hand, we contend that the problem could be more fundamental in nature. The physical parameters L, A_S and τ are computed in terms of individual regression coefficients. Given the numerous sources of errors, multiple regression will assign improper coefficients to such individual parameters, the problem being compounded by the presence of lag variables. Unless effects like multicollinearity and data errors are kept to a minimum, it seems unrealistic to assume that the statistical determination will be robust and that the physical parameter set thereby computed sound.

b) The ETP approach, though appealing as a result of its physical correspondence, does have inherent drawbacks that limit its applicability to the inverse problem approach. One has to decide on a specific network, manipulate the governing equations into a suitable form, and remove effects arising due to under-or over-determination before the regression can be performed. Our analyses show that the identification process may often not be sensitive enough to distinguish between competing networks and even between alternate generic configurations. Moreover, as was the case for the time derivative approach, individual regression coefficients may be unstable and consequently so will the building parameters. Another drawback inherent in this approach is the lack of a systematic procedure by which the structure of the modeling equation can be progressively modified following the dictates of statistical parsimony.

c) The time series approach has inherent features that overcome, to some extent, the drawbacks of the previous two approaches. It requires very little effort in model manipulation prior to regressing the data, yet offers a systematic and progressive means of reaching the most parsimonious order of the model. (This procedure can be likened to an interactive procedure as against a search procedure implied by the ETP approach.) Also, the building parameters are estimated from sums of several regression coefficients pertaining to one exogenous parameter and its lag variables. Though the danger of the influence of errors is not entirely removed, the magnitude will be lower. That the parameter sets identified are more stable and consistent than previous approaches could be attributed to the above inherent advantages of the time series approach. In other words, all models are bound to have noise associated with using them in order to identify building physical parameters. However, under similar conditions of data collection, the time series approach has been found to yield most consistent results most conveniently.

The ultimate objective of this study was not to identify or develop either models for building energy performance prediction or the most appropriate related experimental protocol (as was done in Subbarao [1988]). This study was undertaken in order to evaluate different models when applied to the same data set gathered under "adverse" conditions of building operation and data collection. This involved (i) learning the advantages and disadvantages of using one or the other in terms of effort, and (ii) gauging the stability and consistency of the parameter estimates. The data set was collected unobtrusively during the period when the building was under free-floating mode: conditions which would most adversely affect the stability of the building physical parameters identified. Despite this, the results of this study suggest that the time series modeling approach could be useful for certain practical applications, for example, on-line identification. Appropriate algorithms could be built into an "intelligent" controller, which would certainly be an important step for energy conservation and comfort in buildings.

ACKNOWLEDGMENTS

The author has benefited from stimulating discussion and suggestions from several of his colleagues, most notably A. Rabl, K. Subbarao, J.S. Haberl and P. Komor. Valuable criticism preferred by J.F. Kreider & L.K. Norford is also acknowledged, as is that of the anonymous reviewers of this paper. The painstaking effort involved in setting up and maintaining the instrumentation to collect data for the three residences was done by L.K. Norford and G.V. Spadaro. This research was funded by the New Jersey Gas and Electric Utilities and the New Jersey Department of Commerce, Energy and Economic Development under the New Jersey Energy Conservation Laboratory.

NOMENCLATURE

A_S = mean effective solar aperture including both the conversion from horizontal surface as well as optical effects

a, b, c, d = weighting factors for the time series model

C = effective heat capacity

H = thermal conductance

L = steady-state overall heat loss coefficient of building

N = number of terms in the time series or time derivative models

n = integer indicating present time interval

Q_A = total internal heat generation in building

Q_S = global solar radiation on a horizontal surface per unit area

R = thermal resistance

$\bar{R}^2$ = coefficient of determination adjusted for number of degrees of freedom

RMSE = root mean square error

T_a = ambient dry bulb temperature

T_C = clamp temperature of building

T_i = interior air temperature of building

t = time

α, β, γ, δ = coefficients of the time derivative model

Δt = time interval or time step

τ = time constant

REFERENCES

Achterbosch, G.G.J.; de Jong, P.P.G.; Krist-Spit, C.E.; van der Muelen, S.F.; Verberne, J. 1985. "The development of a convenient thermal dynamic building model." Energy and Buildings, Vol.8, p. 183.

Box, G.E.P.; and Jenkins, G.M. 1976. Time series analysis: forecasting and control. New York: Holden Day.

Carter, C.; and de Villiers, J. 1987. Principles of passive solar building design. New York: Pergamon Press.

Cleary, P.G. 1987. "Which types of analysis can be carried out with EMS data? An examination of data from an apartment building." ASHRAE Transaction, Vol. 93, Part 2.

Crawford, R.R.; and Woods, J.E. 1985. A method for deriving a dynamic system model from actual building performance data." ASHRAE Transaction, Vol. 91, Part 2.

DOE, 1981. DOE-2 engineers manual, Version 2.1A, D.A. York and C.C. Cappiello (Eds.). Washington, DC: Department of Energy.

Griffith, J.E. 1985. "Determination of thermal time constants in residential housing." ASHRAE Transaction, Vol. 91, Part 2B, pp. 1450-1459.

Hammerstein, S. 1984. "Estimation of energy balances for houses," Bulletin M84.18, National Swedish Institute for Building Research, December.

Hsieh, E.H.; Norford, L.K.; Socolow, R.H.; and Spadaro, G.V. 1989. "Calibrated computer models to track building energy use: the role of the tenant and operator decisions." Submitted to Energy and Buildings, January.

Kasuda, T. 1985. "Heat transfer in buildings," chap. 9, Handbook of Heat Transfer Applications, W.M. Roshenow, J.P. Hartnett and E.N. Ganic (Eds.), 2nd Edition. New York: McGraw-Hill.

Neveu, A.; Bacot, P.; and Regas, R. 1986. "Modeles d'evolution thermique des batiments: conditions pratiques d'identification, Revue Generale de Thermique, No.296-297, p. 413.

Nisson, J.D.N.; and Dutt, G.S., 1985. The superinsulated home book. New York: John Wiley.

Norford, L.K.; Rabl, A.; Socolow, R.H.; and Spadaro, G.V. 1985. "Monitoring the energy performance of the Enerplex office buildings: results of the first year of occupancy." Princeton PU/CEES Report No. 203.

Norford, L.K.; Rabl, A.; and Spadaro, G.V. 1987. "Task I.B.2: a pretest of night cooling strategies," Internal Milestone Report, CEES, Princeton University, November.

Mikradakis, S.; Wheelwright, S.C.; and McGee, V.E., 1983. Forecasting: methods and applications, 2nd edition. New York: John Wiley and Sons.

Parkanen, J. 1988. "Prediction and fault detection of building energy consumption using multi-input, single-output dynamical model." Proceedings of the International Symposium Energy Options for the Year 2000, Vol. 3, Energy and Buildings, p. 57, September.

Persily, A.K. 1982. "Understanding air infiltration in houses." Ph.D. thesis, Mechanical and Aerospace Eng. Dept., Princeton University.

Rabl, A. 1988. "Parameter estimation in buildings: methods for dynamic analysis of measured energy use." ASME J. of Solar Energy Eng., Vol. 110.

Reddy, T.A. 1989. "Identification of building parameters using dynamic models: analysis of three occupied residences monitored non-intrusively, PU/CEES Report No. 236, Princeton University.

Sinha, N.K.; and Kuszta, B. 1983. Modeling and identification of dynamic systems. New York: Van Nostrand Reinhold Co.

Sonderegger, R.C. 1977. "Modeling residential heat load from experimental data: the equivalent thermal parameters of a house." International Conference on Energy Use Management, p.183, Tucson, AZ.

Sonderegger, R.C. 1978. "Diagnostic tests determining the thermal response of a house." ASHRAE Transaction , Vol. 84, Part 1, pp. 691-702. Also, "Dynamic models of house heating based on equivalent thermal parameters." Ph.D. thesis, Aerospace and Mechanical Sciences Dept., Princeton University, 1977.

Stoecker, W.F. 1980. Design of thermal systems. New York: McGraw-Hill.

Subbarao, K. 1985. "Building parameters and their estimation from performance monitoring." SERI/TP-253-2661, Solar Energy Research Institute, Golden, Colorado.

Subbarao, K., 1988. "PSTAR-primary and secondary terms analysis and renormalization: A unified approach to building energy simulations and short-term monitoring." SERI/TR-254-3175, Solar Energy Research Institute, Golden, Colorado.

Taylor, Z.T.; and Pratt, R.G. 1988. "The effect of model simplifications on equivalent thermal parameters calculated from hourly building performance data." Proceedings of the 1988 ACEEE Summer Study on Energy Efficiency in Buildings, Vol. 10, p. 268.

Van Hatten, D.; and Norlen, U. 1987. "Development and application of an evaluation method for solar buildings." Publication unknown.

Wilson, N.W.; Colborne, W.G.; and Ganesh, G., 1984. Determination of thermal parameters for an occupied house." ASHRAE Transaction, Vol. 90, Part 2.

Table 1. Equations for the building physical parameters in terms of regression coefficients (Rabl 1988).

	Steady State Bldg. Parameters L	A_S	Effective Bldg. Heat Capac. C	Time Constant τ
Time Series Approach	$\frac{\sum_{k=o}^{N_i} a_k}{\sum_{k=o}^{N_A} c_k}$	$\frac{\sum_{k=o}^{N_S} d_k}{\sum_{k=o}^{N_A} c_k}$	$L(\tau_1 + \tau_2)$	(2222 model) $\Delta t. \left[\ln \left(\frac{-a_1 \pm \sqrt{a_1^2 - 4 a_2}}{2 a_2} \right) \right]^{-1}$
Time Derivative Approach	$\frac{\alpha_o}{\gamma_o} - \frac{\beta_o}{\gamma_o}$	$\frac{\delta_o}{\gamma_o}$	$L \tau$	(1111 model) $\frac{\alpha 1}{\alpha o}$

Table 3. Values of the building physical parameters estimated from regression using different configurations of thermal networks. The values of the time constants have also been computed from the building parameters estimated and are listed so as to enable easier comparison with gray-box models.

	RC1 (RC1b)	RC2	RC3	RC4
House 1				
H_1 (kW/°C)	-0.18 (0.34)	0.37	0.42	0.36
H_2 (kW/°C)	48.6 (15.2)	-93.7	-93.1	-67.0
H_3 (kW/°C)	- -	-	0.50	-
A_S (m^2)	2.4 (1.6)	2.5	1.0	2.3
C (kWh/°C)	17.7 (16.2)	17.9	14.5	17.5
τ (h)	48.3 (47.6)	48.8	15.5	48.2
$\bar{R}^2$	0.32 (0.33)	0.32	0.36	0.33
House 2				
H_1 (kW/° C)	0.95	0.53	-2.9	0.50
H_2 (kW/°C)	21.6	36.1	-50.5	-13.0
H_3 (kW/°C)	-	-	-0.45	-
A_S (m^2)	16.8	8.0	-13.0	7.9
C (kWh/°C)	21.6	10.7	36.2	10.4
τ (h)	23.8	20.6	11.6	20.7
$\bar{R}^2$	0.75	0.75	0.77	0.74
House 3				
H_1 (kW/°C)	0.46	0.45	0.32	0.48
H_2 (kW/°C)	71.2	120.5	148.2	192.8
H_3 (kW/°C)	-	-	0.10	-
A_S (m^2)	6.6	6.3	5.0	7.1
C (kWh/°C)	21.3	20.6	15.4	22.4
τ (h)	46.6	45.9	36.5	46.6
$\bar{R}^2$	0.52	0.52	0.52	0.52

Table 2. Final form of the regression equations for the four electrical networks

(a) RC1

$$\dot{T}_i = a_o\dot{T}_a + b_1(T_a - T_i) + c_o\dot{Q}_A + c_1Q_A + d_1Q_S$$

where

$$a_o = \frac{R_2}{R_1(1 + R_2/R_1)}, \quad b_1 = \frac{1}{C\ R_1(1 + R_2/R_1)}$$

$$c_o = \frac{R_2}{(1 + R_2/R_1)}, \quad c_1 = \frac{1}{C(1 + R_2/R_1)}, \quad d_1 = \frac{A_S}{C(1 + R_2/R_1)}$$

with

$$\frac{a_o}{b_1} = \frac{c_o}{c_1}$$

(b) RC2

$$\dot{T}_1 = a_o\dot{T}_a + b_1(T_a - T_i) + c_o\dot{Q}_A + c_1Q_A + d_o\dot{Q}_S + d_1Q_S$$

where

$$a_o = \frac{R_2}{R_1(1 + R_2/R_1)}, \quad b_1 = \frac{1}{C\ R_1(1 + R_2/R_1)}, \quad c_o = \frac{R_2}{(1 + R_2/R_1)}$$

$$c_1 = \frac{1}{C(1 + R_2/R_1)}, \quad d_o = \frac{A_S\ R_2}{(1 + R_2/R_1)}, \quad d_1 = \frac{A_S}{C(1 + R_2/R_1)}$$

with

$$\frac{ao}{b_1} = \frac{c_o}{c_1} = \frac{d_o}{d_1}$$

(c) RC3

$$\dot{T}_i = k + a_o\dot{T}_a + b_1\ (T_a - T_i) + b_2T_i + c_o\dot{Q}_A + c_1Q_A + d_o\dot{Q}_S + d_1Q_S$$

where

$$k = \frac{T_c}{R_3}, \quad a_o = \frac{R_2}{R_1 . R^*}, \quad b_1 = \frac{1}{C\ R_1\ .\ R^*}$$

$$b_2 = -\frac{1}{C\ R_3R^*}, \quad c_o = \frac{R_2}{R^*}, \quad c_1 = \frac{1}{C\ R^*}$$

$$d_o = \frac{A_S\ R_2}{R^*}, \quad d_1 = \frac{A_S}{C\ R^*}$$

$[\ R^* = (1 + R_2/R_1 + R_2/R_3)$

with

$$\frac{a_o}{b_1} = \frac{c_o}{c_1} = \frac{d_o}{d_1}$$

(d) RC4

$$\dot{T}_i = b_1\ (T_a - T_i) + c_o\dot{Q}_A + C_1Q_A + d_1Q_S$$

where

$$b_1 = \frac{1}{C\ R_1}, \quad c_o = R_2$$

$$c_1 = \frac{R_2}{C}\ (1/R_1 + 1/R_2), \quad d_1 = \frac{A_S}{C}.$$

Table 4. Root Mean Square Errors (in ^{o}C) between actual and model predicted values of T_i using the ETP model formulation approach.

Period	House	No of hours	RC1	RC2	RC3	RC4	RC1b
09/22-09/30	1	216	0.7	-	-	0.7	0.8
09/11-09/17	2	144	1.0	0.9	-	0.9	-
09/22-10/1	2	228	0.9	0.8	-	0.8	-
08/27-09/2	3	156	2.1	2.1	1.8	2.1	-
09/23-10/1	3	216	1.0	1.0	1.0	1.0	-

Table 5. Sample values of the building parameters estimated by stepwise regression using the time series model approach. A functional form without intercept was selected and all parameters fitted simultaneously.

Model	1111	2222	3333	2222 (stability check)	
House 1					
L ($kW/^{o}C$)	0.27	0.43	0.35		
A_S (m^2)	0.3	3.0	2.2		
τ_1 (h)	51.1	50.3	45.3		
τ_2 (h)	-	0.56	0.55		
C ($kWh/^{o}C$)	13.8	21.9	16.0		
House 2					
L ($kW/^{o}C$)	0.36	0.31	0.30	0.25	0.35
A_S (m^2)	4.3	3.1	2.9	2.8	3.3
τ_1 (h)	20.7	19.8	20.0	23.5	18.6
τ_2 (h)	-	1.03	0.97	1.16	0.97
C ($kWh/^{o}C$)	7.5	6.5	6.3	6.2	6.8
House 3					
L ($kW/^{o}C$)	0.44	0.42	0.46		
A_S (m^2)	6.2	5.5	6.8		
τ_1 (h)	45.6	43.5	46.0		
τ_2 (h)	-	0.82	1.39		
C ($kWh/^{o}C$)	20.1	18.6	21.8		

Table 6. Root Mean Square Errors between actual and model predicted values of T_i using the time series model.

	Period	Hours	RMSE (^{o}C)		
			2222T	2222A	2222B
House 1	09/22-09/30	216	0.7	-	-
House 2	09/11-09/17	144	0.7	1.1	0.6
	09/22-10/1	228	0.8	0.7	1.0
			1111	2222	3333
House 3	08/27-09/2	156	2.1	2.1	2.0
	09/23-10/1	216	1.0	1.0	1.1

Table 7. Values of the building physical parameters estimated using different variants of the time derivative approach.

Model	A -	B 1110	C 1110	D 1111
	Relevant variant			
House 1				
L (kW/^{o}C)	0.04	0.31	0.23	0.33
A_S (m^2)	-4.7	0.0	-2.1	-0.96
τ (h)	-	22.6	32.1	6.9
$\bar{R}^2$	-	0.772	0.770	0.518
RMSE (^{o}C)	-	2.20	2.17	3.53
House 2				
L (kW/^{o}C)	0.20	0.25	0.25	0.25
A_S (m^2)	0.28	1.8	1.8	0.93
τ (h)	-	8.1	8.1	9.9
$\bar{R}^2$	-	0.932	0.931	0.777
RMSE (^{o}C)	-	1.27	1.27	2.49
House 3				
L (kW/^{o}C)	0.03	0.19	0.20	0.18
A_S (m^2)	-5.1	0.0	0.23	-0.67
τ (h)	-	17.7	24.7	11.3
$\bar{R}^2$	-	0.815	0.810	0.758
RMSE (^{o}C)	-	3.02	3.05	3.57

Table 8. Root Mean Square Errors (^{o}C) between actual and model predicted values of T_i using the time derivative model formulation approach.

	Period	Hours	Variant		
			B	C	D
House 1	09/22-09/30	216	2.3	1.5	2.5
House 2	09/11-09/17	144	0.8	1.1	0.5
	09/22-10/1	228	1.0	1.6	1.4
House 3	08/27-09/2	156	2.8	2.1	2.7
	09/23-10/1	216	1.6	1.6	1.7

Table 9. Building parameter sets identified by the three approaches which seem most physically realistic and consistent.

	ETP	Time Series	Time Derivative
House 1	RC1b	3333	Variant C
L (kW/°C)	0.34	0.35	0.23
A_S (m^2)	1.6	2.2	-2.1
C (kWh/°C)	16.2	16.0	7.4
τ_1 (h)	47.6	45.3	32.1
τ_2 (h)	-	0.55	-
$\bar{R}^2$	0.322	0.366	-
RMSE (°C) Model	0.15	0.17	
Tracking	0.8	0.7	1.5
House 2	RC2	2222	Variant C
L (kW/°C)	0.53	0.31	0.25
A_S (m^2)	8.0	3.1	0.93
C (kWh/°C)	10.7	6.5	2.5
τ_1 (h)	20.6	19.8	9.9
τ_2 (h)	-	1.03	-
$\bar{R}^2$	0.745	0.819	-
RMSE (°C) Model	0.18	0.16	-
Tracking	0.8-0.9	0.7-0.8	1.1-1.6
House 3	RC1	3333	Variant C
L (kW/°C)	0.46	0.46	0.20
A_S (m^2)	6.6	6.8	0.23
C (kWh/°C)	21.3	21.8	4.94
τ_1 (h)	46.6	46.0	24.7
τ_2 (h)	-	1.39	-
$\bar{R}^2$	0.521	0.582	-
RMSE (°C) Model	0.16	0.15	-
Tracking	1.0-2.1	1.1-2.0	1.6-2.1

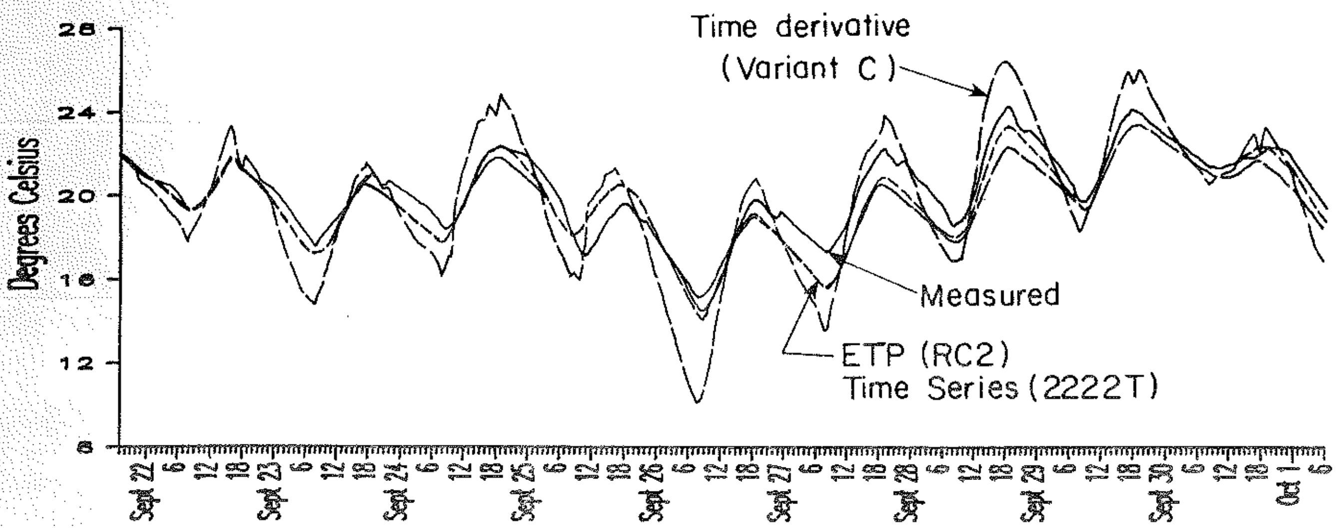

Figure 6. Tracking ability with respect to T_i of different approaches (House 2)