= CEES #215

Reprinted from February 1988, Vol. 110, Journal of Solar Energy Engineering

Parameter Estimation in Buildings: Methods for Dynamic Analysis of Measured Energy Use

A. Rabl[1]
Princeton University,
Center for Energy and
Environmental Studies,
Princeton, NJ 08544

Dynamic analysis of energy data can help improve the efficiency of buildings in several ways: evaluation of proposed modifications of a building or its operation (e.g., changes in thermostate setpoints); verification of performance on the basis of short-term measurements (corrected for weather); diagnostics and optimal control of HVAC equipment (real-time comparison of actual and predicted performance can be a powerful diagnostic tool). For this purpose one would like a simple building model whose parameters can readily be adjusted by a statistical fit to the data. This paper reviews the available methods: thermal networks, modal analysis, differential equations, ARMA (autoregressive moving average) models, Fourier series, and calibrated computer simulations. The basic models can be applied in several ways, differing in choice of dependent variable, number of coefficients, statistical criterion, time step, finite differencing scheme, and implementation as linear or nonlinear algorithm. The relation between the various approaches is examined. It is shown how the results of each of these methods can be presented in a standardized format that maximizes their physical interpretation, in terms of time constants and admittances (including heat loss coefficient and solar aperture). A general proof is given that the effective heat capacity equals the heat loss coefficient multiplied by a sum of time constants. The methods are tested with data from an office building. Special attention is focused on difficulties, due to air exchange or solar gains, that are likely to arise in practice.

1 Introduction

There are many different methods and models for the analysis of energy use in buildings. No single one is universally the best. Rather, the choice in a given situation depends on what one wants to calculate and what input data are available.

To begin we note a basic distinction in the goals of different models: the *forward problem* and the *inverse problem*. The designer of a building is concerned with the forward problem: he has the description of the building and he wants to calculate its peak and average loads.

By contrast, once a building has been in use the energy consumption is known, from utility bills if not from an energy management system. At that point the relevant questions are:

(*a*) How does the consumption compare with design predictions (and in case of discrepancies, are they due to anomalous weather, to unintended thermostat settings, to malfunctioning systems, or to other causes)?

(*b*) How would the consumption change if thermostat settings or ventilation rates are changed (in other words: what is the cost of comfort)?

(*c*) How much could be saved by retrofits of building shell or equipment?

(*d*) If retrofits are implemented, can one verify that the savings are due to the retrofit and not to other causes, e.g., the weather?

(*e*) How can one optimize control and operation of the HVAC equipment? (A model that compares measured and predicted behavior under actual operating conditions can be a valuable diagnostic tool).

All of these questions involve the inverse problem: given actual performance data for a building, how much can one learn about its physical characteristics? One could, of course, try to answer these questions by going back to the blueprints of the building and repeating the analysis that was performed at the design stage. But there are several difficulties. The process requires much labor, assuming that the original blueprints can be found at all. There is the possibility of errors or misinterpretations (Diamond et al., 1985). Materials as actually installed are often different from the bulk properties reported in the literature. And the builder may not have followed all the original specifications.

The inverse problem forms the subject of the present paper. It is a classical problem that arises in many fields from engineering to economics, and it is usually known as estimation or system identification problem (see e.g., Gelb, 1974; de

[1] Current Address from February 1988 to January 1989: Centre d'Energetique, Ecole des Mines, 60 bd. St.-Michel, 75272 Paris, CEDEX 06, France.

Contributed by the Solar Energy Division for publication in the JOURNAL OF SOLAR ENERGY ENGINEERING. Manuscript received by the Solar Energy Division, June, 1987.

Larminat and Thomas, 1975; and the journal *Inverse Problems* published by IOP, 1985). The building is considered as a black box whose characteristics are to be inferred from measured temperature and energy data.

A model for the inverse problem has to meet requirements very different from the forward problem. For the inverse problem the number of adjustable parameters should be small because the information content of the data is very limited, being collected under fairly repetitive conditions and subject to errors. Also, one needs a systematic procedure for choosing the parameters and fitting them to the data. But the models determined by the inverse process can be much simpler than those used in the forward direction. In particular, as we show in Section 5.5, most buildings can be approximated by a single zone if they are thermostatically controlled to have uniform temperature. Therefore, we present only the equations for a single zone.

To see which models might be suitable for the inverse problem, we list in Table 1.1 the principal types of building models, separated according to time resolution into steady state models (part *a*) and dynamic models (part *b*). Steady state models are much simpler, but they are generally limited to monthly, weekly, or daily averages. There are many situations where dynamic models are preferable or required: warmup and cooldown; peak loads; rapid monitoring; diagnostics; and optimal control (for the application of control theory to buildings, see Letherman, 1981, and ICBEM 1987). Since the inverse model for the steady state situation is well developed (the PRISM model of Fels et al., 1986), we address only dynamic models.

The requirements of the inverse problem are natural for the following models: thermal networks (Sonderegger, 1977 and 1978), modal analysis (Bacot et al., 1984), differential equations, ARMA (autoregressive moving average) models (Subbarao, 1985), and Fourier analysis (Shurcliff 1984; Subbarao, 1984).

The existing mainframe computer simulation programs like BLAST and DOE2.1, are awkward for the inverse problem. One can try to calibrate a computer simulation by comparing its output with measured data and adjusting some of the input, such as air exchange, internal heat gains, thermostat setpoints, and shading coefficients. The results can be good, as shown by the example we cite (Hsieh, 1988). However, this so-called calibrated computer program approach lacks a systematic procedure for deciding which of the innumerable input parameters should be adjusted and by how much. Furthermore, the process is labor intensive. It would be ill suited as control algorithm for a thermostat or an energy management system. On the other hand, BLAST and DOE2.1 have the advantage of being able to model the HVAC equipment as well. Hence the most powerful approach would be a hybrid computer simulation program that contains subroutines for system identification. The structure of the proposed new simulation program, the Energy Kernel System, would be superbly suited for that (DOE-2 User News, 1986).

As the survey in Table 6.1 shows, many articles have already been written on the subject of the inverse problem, but each is limited to the favorite method of its author. The present paper attempts to provide an overview of the different methods, elucidating how they are related to each other. Particular attention is focused on their physical interpretation. The methods are tested with data from an office building. The large scatter in the results highlights the danger of compensating errors: even with wrong parameters a model may give a good fit to the data. We discuss ways of reducing the difficulties and suggest directions for future research.

2 Description of Models

2.1 Thermal Networks. Thermal networks offer perhaps the conceptually simplest approach. The building is approximated by a simple network with a few resistances and capacitances. Several possible networks are shown in Fig.2.1. They can describe various features of a building, and it is instructive to discuss them briefly.

A dynamic model must contain at least one resistance and

Nomenclature

A = solar aperture, including opaque surfaces [m^2] ($=A_{sol}/A_{aux}$ at $\omega=0$)

$a_{aux}(k)$, $a_{ext}(k)$, $a_{int}(k)$, $a_{sol}(k)$ = coefficients of ARMA series

A_{aux}, A_{ext}, A_{int}, A_{sol} = admittances, equations (4.6)–(4.9)

c_p = specific heat [kJ/kg°]

C = effective heat capacity [kWh/°C]

$\mathbf{C}$ = matrix of heat capacities for modal analysis

$\mathbf{D}$ = matrix characterizing sensitivity to driving terms in modal analysis

$\mathbf{K}$ = matrix of heat conductivities for modal analysis

L = steady state heat loss coefficient [kW/°C] ($=A_{int}/A_{aux}$ at$\omega=0$).

m = flow rate of air (ventilation and infiltration) [kg/s].

n = $t/\Delta t$

N_{aux}, N_{ext}, N_{int}, N_{sol} = -1 + number of terms in ARMA series or differential equation

Q_{aux} = total nonsolar heat input to building (including or excluding heat content of air exchange, depending on whether admittances are defined as total or as conductive terms) [kW]

Q_{sol} = measure of solar radiation used for definition of A and A_{sol} (in this paper hemispherical horizontal radiation is used [kW/m2]

R = thermal resistance [C/kW]

t = time

T_{int} = interior temperature of building [°C]

T_{ext} = exterior temperature [°C]

$\mathbf{U}$ = driving functions for modal analysis

$\mathbf{Y}$ = output of modal analysis

$\mathbf{X}$ = eigenvectors for modal analysis

α = coefficients of differential equation with subscripts, e.g., $\alpha_{int,1}$ = coefficient of $\dot{T}_{int}$

Δt = time step

τ = time constant

ϕ = phase

ω = frequency

Table 1.1 Methods for thermal analysis of buildings

A * under the "forward" and "inverse" columns indicates applicability to forward problem (calculating energy performance from building characteristics) and inverse problem (determining characteristics of building from energy consumption data). The references refer either to the original work or to a good recent summary.

a) STEADY STATE METHODS
Not for peak loads (with possible exception of ASHRAE TC 4.7)

	forward	inverse	comments
degree day method [ASHRAE 1985]	*		The simplest. Based on fixed reference temperature 18.3 C. Can go quite wrong for commercial or superinsulated buildings.
variable base degree day method [ASHRAE 1985]	*		Variable reference temperature. Can be good approximation for annual consumption.
bin method [ASHRAE 1985]	*		Input: hours in each 2.8 C (5 F) bin of ambient temperature. More flexible than var. base deg. day method: can model temperature dependent features, weekends, etc.
PRISM [Fels 1986]		*	Needs data for energy use (several periods/year) and for daily average ambient (no T_{int}). Finds reference temperature, and heat loss coefficient divided by heater efficiency. Best for weather correction.
ASHRAE TC 4.7	*		modified bin method, with cooling load factors etc. to account for some transient effects and determine peak loads.

b) DYNAMIC METHODS

	forward	inverse	comments
thermal network [Sonderegger 1977]	*	*	In forward direction no limit on complexity of network. For inverse problem network must be simple, with equival. thermal parameters.
response factor series [Stephenson and Mitalas 1967]	*		Tabulated results for building components [ASHRAE 1985] useful for calculation of peak loads.
Fourier analysis [Shurcliff 1984]	*	*	Calculates response to sinusoidal (constant plus diurnal) input. Can be combined with calculation in time domain.
ARMA model [Subbarao 1985]		*	Coefficients lack direct physical interpretation, but that can be provided with time constants and admittances.
BEVA [Subbarao 1985]	*	*	Combination ARMA + Fourier methods. Loads calculated in time domain.
modal analysis [Bacot et al. 1984]	*	*	Diagonalization of the differential equations for the building. For inverse problem building is approximated by small number of modes.
differential equation [Eq.2.10 of this paper]		*	Approximates building by linear diff. eq. Order and coefficients adjusted by data. Can be integrated analytically. Much flexibility for fitting, prediction and control.
computer simulation (e.g. DOE 2.1, BLAST)	*	*	Very detailed. Potentially the most accurate method. Also models HVAC equipment. Requires much expertise and labor for coding the input.
hybrid methods	*	*	Computer simulation, plus diff.eq. or ARMA. To be developed.

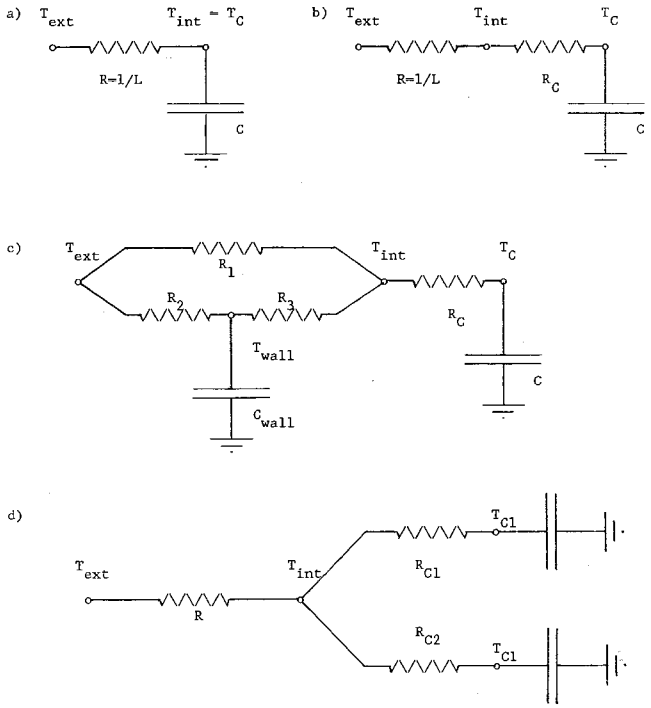


Fig. 2.1 Examples of simple thermal networks for buildings: L = total heat loss coefficient. (a) Simplest possible model with heat capacity (1R1C); (b) with resistance between interior and heat capacity (2R1C); (c) with two heat capacities, C in interior, C_{wall} in envelope, L given by $L = 1R_1 + 1/(R_2 + R_3)$; (d) with two interior heat capacities, C_1 (interior wall and floor) and C_2 (furniture), with different couplings R_{C1} and R_{C2}.

one capacitance. Thus the simplest possibility is the 1R1C network in Fig. 2.1(*a*). It has features that are somewhat unrealistic. The capacitance is assumed to always have the same temperature as the interior air. In reality almost all of the heat capacity resides in the solid parts of the building, not the interior air, and there is always some thermal resistance between the air and the rest of the building.

One can account for that by adding a resistance R_c between the capacitance C and the interior air temperature T_{int} as in the 2R1C network of Fig. 2.1(*b*). This added resistance can have an important effect in reducing the peak instantaneous thermal loads during warmup, because the temperature of the air responds to changes of the thermostat setpoints much faster than the bulk of the building. This 2R1C network has been especially popular (Sonderegger, 1977; Norford et al., 1984; Wilson et al., 1985).

Fig. 2.1(*b*) assumes that all the heat capacity is in the interior of the building. While this is obviously true for interior air, for furniture and for interior floors and walls, it does not apply to the building envelope. In buildings with brick walls, the heat capacity of the envelope can have a significant effect in buffering the diurnal pulse of exterior temperature and solar radiation. To account for that one can use the network of Fig. 2.1(*c*). Here the envelope is characterized by two paths: massless through R_1 (good approximation for air exchange and for windows) and massive through $R_2 + R_3$ (for heavy walls).

There are many other possibilities. For example, Fig. 2.1(*d*) might be used to account for two different masses inside the building, e.g., furniture and interior walls.

It is important to realize that for a successful dynamic simulation one may have to choose numerical values of the resistances and capacitances that are quite different from the static properties of the components. This is easy to understand for the case of a thick interior wall in Fig. 2.1(*b*). Since the conductivity of the wall is finite, it takes some time for a thermal pulse to penetrate into the wall. If a sinusoidal temperature variation is applied to the surface, the temperature response in the interior is damped, and the damping increases with frequency and with depth. This implies that the storage capacity of the wall decreases as the frequency increases. A thick wall is more complicated than a single isothermal mass.

In practice the driving forces of buildings follow a pattern that is dominated by a single frequency, the 24 hour cycle. At a single frequency it is permissible to treat a thick wall by a single isothermal point mass (Carslaw and Jaeger, 1959). This is the basis for the concept of *equivalent thermal parameters*, as studied by Sonderegger (1977). For example, for a brick wall with a thickness of 0.1 m with diurnal driving forces the effective heat capacity is very roughly half the static heat capacity. If one uses simple networks as in Fig. 2.1, one must therefore interpret the parameters as equivalent thermal parameters.

To use a thermal network for calculations, one writes down the differential equations (first order) that describe the energy balance at each node. For example, the simplest network, Fig. 2.1(*a*) with a single heat input Q in the interior, is described by the equation

$$Q=\frac{(T_{\text{int}}-T_{\text{ext}})}{R}+C\frac{dT_{\text{int}}}{dT} \tag{2.1}$$

For data fitting one can use the finite difference approximation (see Section 2.3) or an integrated version (see Section 3.5).

Thermal networks can be a valuable guide for one's intuitive understanding of buildings. But there are drawbacks. In the forward direction one first has to calculate the numerical values of the equivalent thermal parameters (unless one uses a very large number of low-mass nodes). This can be tedious since no ready made software is available for that purpose. In the inverse direction one lacks a systematic procedure for deciding which network to use for a given building. The process of casting the differential equations in a form suitable for data fitting is quite laborious. And with some networks, e.g., 2R1C, an ambiguity arises from the fact that the number of coefficients in the linear regression exceeds the number of physical parameters: they are overdetermined.

For these reasons we recommend a different approach, more powerful, yet simpler. Its results do not obviously correspond to networks (or rather, as Subbarao, 1984, has shown, they imply networks where some parameters may be negative), but that is not objection: the goal is not to find networks but to develop models that are as useful and as accurate as possible and that convey a maximum of physical information about a building.

2.2 The ARMA Model. To develop a general modeling approach, we begin by assuming only the most general properties that the thermal processes in a building must satisfy. In view of the narrow range of temperatures involved, one can suppose a linear approximation for all heat transfer equations. Furthermore, the processes must obey causality in the sense that the present value of the interior temperature T_{int} must be a unique function of the past values of T_{int}, exterior temperature T_{ext}, auxiliary heat input Q_{aux} and solar heat input Q_{sol}. The most general relation satisfying these two constraints, linearity and causality, is an integral equation of the form

$$\int_0^\infty dt'\,[a_{\text{int}}(t')\,T_{\text{int}}(t-t')-a_{\text{ext}}(t')\,T_{\text{out}}(t-t')$$
$$-a_{\text{aux}}(t')Q_{\text{aux}}(t-t')-a_{\text{sol}}(t')Q_{\text{sol}}(t-t')]=0 \tag{2.2}$$

where a_{int}, a_{ext}, a_{aux} and a_{sol} are functions that depend only on the building. They are called *transfer functions*. The choice for their signs is arbitrary; the notation simply follows Subbarao (1985). (Some authors have extended the integrals from $-\infty$ to $+\infty$ with the understanding that the transfer functions are defined to vanish for negative arguments).

The auxiliary heat input Q_{aux} is defined as the total nonsolar heat input (net furnace output, lights, occupants). This grouping of heat inputs may appear somewhat arbitrary but it is a convenient choice for most buildings. The relation between actual solar heat gain and the measure of Q_{sol} (e.g., direct or total radiation, on horizontal or on tilted surface) depends on the solar aperture which is usually not known but rather to be determined from the data.

In practice the temperatures and energies are known only at discrete time intervals. Hence it becomes necessary to approximate the integral by a series of discrete time steps with interval Δt

$$\sum_{k=0}^{N_{\text{int}}} a_{\text{int}}(k)\,T_{\text{int}}(n-k)-\sum_{k=0}^{N_{\text{ext}}} a_{\text{ext}}(k)\,T_{\text{ext}}(n-k)$$
$$-\sum_{k=0}^{N_{\text{aux}}} a_{\text{aux}}(k)\,Q_{\text{aux}}(n-k)-\sum_{k=0}^{N_{\text{sol}}} a_{\text{sol}}(k)\,Q_{\text{sol}}(n-k)=0, \tag{2.3}$$

the time argument being indicated by integers, according to

$$t=n\,\Delta t$$

This is the *ARMA* (*autoregressive moving average*) *model*, in the notation of Box and Jenkins (1976); it has also been called "generalized response factor series" by Subbarao (1985). The response factor series of Mitalas and Stephenson (1967) is a special case, corresponding to a single term for T_{int} (i.e., $N_{\text{int}}=0$). A look at numerical values, tabulated, for example, in the ASHRAE Handbook of Fundamentals (ASHRAE, 1981), shows that in that case one needs to include a fairly large number of terms, easily more than twenty for each driving force if one wants reasonable accuracy. While that is acceptable in the forward direction, it would not be practical for the inverse problem.

However, the need for a large number of terms in equation (2.3) is avoided if one allows more than one term in the sum over T_{int}. Then the previous values of T_{int} contain some of the information that would otherwise have to be supplied by a long series in the other variables. Often one can obtain good accuracy even when each of the N is as small as 1.

Of the total number of parameters in equation (2.3), only $N_{\text{int}}+N_{\text{ext}}+N_{\text{aux}}+N_{\text{sol}}+2$ are independent, because two variables can be eliminated by the following arguments. Obviously the overall scale of equation (2.3) does not matter, and one can arbitrarily set one of the coefficients equal to unity. Furthermore, equation (2.3) must hold under steady state conditions. But when all T and Q are constant, the first law of thermodynamics requires that

$$L(T_{\text{int}}-T_{\text{ext}})=Q_{\text{aux}}+A\,Q_{\text{sol}} \tag{2.4}$$

where

$$L=1/R$$

is the *steady state heat loss coefficient* of the building, and the *steady state solar aperture* A is defined such that

$$\text{solar heat gain}=A\,Q_{\text{sol}} \tag{2.5}$$

when Q_{sol} is the chosen measure of solar radiation. Since under steady state conditions equations (2.3) and (2.4) must become identical for any T and Q, one finds immediately the constraint

$$\sum_{k=0}^{N_{\text{int}}} a_{\text{int}}(k)=\sum_{k=0}^{N_{\text{ext}}} a_{\text{ext}}(k). \tag{2.6}$$

Furthermore, L and A are given by

$$L=\frac{\sum_{k=0}^{N_{\text{int}}} a_{\text{int}}(k)}{\sum_{k=0}^{N_{\text{aux}}} a_{\text{aux}}(k)} \quad \text{and} \quad A=\frac{\sum_{k=0}^{N_{\text{sol}}} a_{\text{sol}}(k)}{\sum_{k=0}^{N_{\text{aux}}} a_{\text{aux}}(k)} \tag{2.7}$$

2.3 Differential Equation. The terms in equation (2.2) have the form

$$\int_0^\infty dt'\,a(t')\,T(t-t').$$

Expanding T in a Taylor series one obtains

$$\int_0^\infty dt'\ a(t')\ T(t-t') = \sum_{n=0}^{\infty} \alpha_n T^{(n)}(t) \tag{2.8}$$

where the coefficients α_n are defined by

$$\alpha_n = \int_0^0 dt'\ a(t')\ \frac{(-t')^n}{n!} \tag{2.9}$$

Therefore, a building can be described by a differential equation of the form

$$\begin{aligned}&\alpha_{\text{int},0} T_{\text{int}}(t) - \alpha_{\text{ext},0} T_{\text{ext}}(t) - \alpha_{\text{aux},0} Q_{\text{aux}}(t)\\ &- \alpha_{\text{sol},0} Q_{\text{sol}}(t) + \alpha_{\text{int},1} \dot{T}_{\text{int}}(t) - \alpha_{\text{ext},1} \dot{T}_{\text{ext},1}(t)\\ &\quad - \alpha_{\text{aux},1} \dot{Q}_{\text{aux}}(t) - \alpha_{\text{sol},1} \dot{Q}_{\text{sol}}(t)\\ &\qquad + \text{higher order terms} = 0\end{aligned} \tag{2.10}$$

with constant coefficients α, to be determined by fitting the data. In practice only the lowest order terms will be kept. Analogous to the constraint of equation (2.6) we have

$$\alpha_{\text{int},0} = \alpha_{\text{ext},0} \tag{2.11}$$

The connection with L and A will be discussed in equation (3.3) below.

When data are given only at discrete time intervals, one possible approach is to replace the derivatives by finite differences (another approach will be described in Section 3.5). There are several possible schemes for doing so, and the choice affects the form of the series. For instance, with a second order equation the following two schemes appear natural

the first

$$\begin{aligned}&X(t) \to [X(n-1) + X(n-2)]/2\\ &\dot{X}(t) \to [X(n-1) - X(n-2)]/\Delta t\\ &\ddot{X}(t) \to [X(n-1) - X(n-1) - X(n-2) + X(n-3)]/(2\Delta t^2)\end{aligned} \tag{2.12}$$

and the second

$$\begin{aligned}&X(t) \to X(n-1)\\ &\dot{X}(t) \to [X(n) - X(n-2)]/(2\Delta t)\\ &\ddot{X}(t) \to [X(n) - 2X(n-1) + X(n-2)]/\Delta t^2\end{aligned} \tag{2.13}$$

In any case, having selected a particular finite difference approximation, one can interpret the ARMA model as a differential equation, and vice versa. The number of terms in the series depends on the differencing scheme, and its increases with the order of the differential equation. Suppose, for example, that a differential equation contains the 0th, 1st, and 2nd derivatives of a function $X(t)$ with arbitrary coefficients as

$$\alpha_{X,0} X(t) + \alpha_{X,1} \dot{X}(t) + \alpha_{X,2} \ddot{X}(t) \tag{2.14}$$

Then the corresponding terms in the ARMA model are

$$\sum_{k=0}^{N_X} a_X(k)\ X(n-k) \tag{2.15}$$

and N_x is 3 for equation (2.12) and 4 for equation (2.13). The correspondence between the α and the a is obtained by inserting the differencing scheme and equating coefficients between equation (2.14) and (2.15). For the scheme of equation (2.12) the result is readily found to be

$$\begin{aligned}a_X(0) &= \alpha_{X,2}/(2\Delta t^2) = a_X(3)\\ a_X(1) &= \alpha_X/2 + \alpha_{X,1}/\Delta t - \alpha_{X,2}/(2\Delta t^2)\\ a_X(2) &= \alpha_X/2 - \alpha_{X,1}/\Delta t - \alpha_{X,2}/(2\Delta t^2).\end{aligned} \tag{2.16}$$

This matching of terms can be carried out for any N_X or for any number of derivatives. Nonetheless, in an actual data fit the results of ARMA series and differential equation can be quite different because different variables are regressed against each other.

2.4 Modal Analysis. The differential equations for a thermal network can be written in the form

$$C_n\ \dot{T}_n = \sum_{k=1}^{N} K_{nk}\ T_k + \sum_{i=1}^{M} D_{ni} U_i \tag{2.17}$$

where C_n is the heat capacity of the n'th node and K_{nk} is the effective conductance between the n'th and the k'th nodes. The T_n are the temperatures of the nodes, and the U_i are the driving terms, i.e., heat inputs and external temperatures that are applied to the system.

One can interpret equation (2.17) as a matrix equation, with a diagonal matrix $\mathbf{C}$ whose elements are the heat capacities. The capacities of some nodes may be so small that one can neglect them for some calculations; for instance, with the network of Fig. 2.1(*b*) one often neglects the heat capacity of the T_{int} node because the capacity of the air is very small compared the mass of the building. But strictly speaking, all C_n are greater than zero. Hence the inverse of the matrix $\mathbf{C}$ exists. Furthermore, for the applications considered here the effective conductances are symmetric

$$K_{nk} = K_{kn} \tag{2.18}$$

Being symmetric, the matrix $\mathbf{C}^{-1}\ \mathbf{K}$ can be diagonalized, and it has real eigenvalues. The corresponding eigenvectors are the modes of the system, hence the name modal analysis. Designating by $\mathbf{P}$ the matrix that relates the vector of node temperature $\mathbf{T}$ to the eigenvectors $\mathbf{X}$ according to

$$\mathbf{T} = \mathbf{P}\ \mathbf{X}, \tag{2.19}$$

one can write equation (2.17) in the form

$$\dot{\mathbf{X}} = \mathbf{E}\mathbf{X} + \mathbf{F}\mathbf{U}, \tag{2.20}$$

with

$$\mathbf{E} = \mathbf{P}^{-1}\mathbf{C}^{-1}\mathbf{K}\mathbf{P} \text{ and } \mathbf{F} = \mathbf{P}^{-1}\mathbf{C}^{-1}\mathbf{D} \tag{2.21}$$

$\mathbf{E}$ is the diagonal matrix of eigenvalues, and its elements are

$$E_i = -1/\tau_i \text{ where}$$

the τ_i are the time constants.

The solution of equation (2.20) is

$$\begin{aligned}X_k(t) = \exp(-t/\tau_k)\,[X_k(t_0) + \sum_{i=1}^{M} \int_{t_0}^{t}\\ \exp(t'/\tau_k)\ F_{ki}\ U_i(t')\ dt']\end{aligned} \tag{2.22}$$

where the $X_k(t_0)$ are the initial conditions. Usually one does not measure the eigenvectors themselves; rather the output of interest is a linear combination (or a vector of linear combinations),

$$\mathbf{Y}(t) = \mathbf{G}\mathbf{X}(t) + \mathbf{H}\mathbf{U}(t) \tag{2.23}$$

For a single-zone building $\mathbf{Y}$ is the scalar T_{int}. The last term, $\mathbf{H}\mathbf{U}(t)$, has been added to allow for the possibility of direct coupling between driving term and output $\mathbf{Y}$.

Modal analysis can be used both in the forward and in the inverse direction (Carter, 1979; Bacot, Neveu and Sicard, 1984). It can provide a rapid solution because, once the matrix $\mathbf{C}^{-1}\ \mathbf{K}$ has been diagonalized, the solution is explicitly given by equations (2.22) and (2.23). Furthermore, one frequently finds that only a few dominant modes are important, and that the others can be neglected without serious error. In the inverse direction the goal is to determine the parameters of the model, i.e., the initial conditions $X_k(t_0)$ and the matrix elements of $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{H}$. This is done by varying the parameters to minimize the discrepancy (which can be defined by least square or by some other criterion) between the measured values of $\mathbf{Y}(t)$ and the values calculated according to equation (2.23). The process is nonlinear because of equation (2.22). A software package of this type has been prepared

by Bacot (1985); it is based on the model analysis and uses the search algorithm of Marquardt (1963). One potential limitation of the nonlinear approach of Bacot (1985) is the need for a fairly long uninterrupted data set (at least three times the longest time constant) because the initial conditions $X_k(t_0)$ are among the quantities that must be identified.

Modal analysis involves N coupled first order equations for a system with N modes. To establish the connection with the methods of Sections 2.2 and 2.3, we note that a system of N coupled first order equations is equivalent to a single differential equation of the n'th order. It is instructive to demonstrate this fact for the simplest cases, with one single output variable $Y=T$, and taking $\mathbf{H} = 0$ in equation (2.23). For $N = 1$ the matrices **E**, **G** and **X** reduce to scalars. Multiplying equation (2.20) by G and substituting $T = GX$, we obtain

$$\dot{T}=ET+GFU. \tag{2.24}$$

This is just a different way of writing the differential equation of the 1R1C network, equation (2.1).

For $N = 2$ we pass from modal analysis to differential equation by eliminating X_1 and X_2 in favor of $\dot{T}$ and T by means of two linear equations. For the first equation we take equation (2.23) directly, in this case

$$T=\sum_{k=1}^{2} G_k X_k. \tag{2.25}$$

For the second equation we combine the derivative of equation (2.25) with equation (2.20), with the result

$$\dot{T}=\sum_{k=1}^{2} G_k E_k X_k+\mathbf{G F U}. \tag{2.26}$$

Solving the last two equations for the G_kX_k we find

$$G_1X_1=\frac{E_2T-\dot{T}+\mathbf{G F U}}{E_2-E_1} \text{ and } G_2X_2=\frac{E_1T-\dot{T}+\mathbf{G F U}}{E_1-E_2}. \tag{2.27}$$

Next we take the second derivative of T

$$\ddot{T}=\mathbf{G}\ddot{\mathbf{X}} \tag{2.28}$$

and eliminate the derivatives of X by repeated application of equation (2.20)

$$\ddot{T}=\mathbf{GE}\dot{\mathbf{X}}+\mathbf{GF}\dot{\mathbf{U}}$$

$$=\mathbf{GE}^2\mathbf{X}+\mathbf{GEFU}+\mathbf{GF}\dot{\mathbf{U}} \tag{2.29}$$

Since **E** is diagonal, only the combinations $G_1 X_1$ and $G_2 X_2$ appear, and they can now be replaced by equation (2.27). After some straightforward algebra one obtains the result

$$\ddot{T}-(E_1+E_2)\dot{T}+E_1E_2T=$$

$$-(E_1+E_2)\mathbf{GFU}+\mathbf{GEFU}+\mathbf{GF}\dot{\mathbf{U}} \tag{2.30}$$

This shows that the modal analysis with $N = 2$ modes implies a differential equation of $N = 2$nd order. Conversely, if one starts with the differential equation, equation (2.30), one can solve it to find $\dot{T}$ and T. Then one can obtain G_1X_1 and G_2X_2 by means of equation (2.27). Thus one can explicitly derive the two modal equations

$$G_k\dot{X}_k=G_kE_kX_k+\sum_{i=1}^{M} G_kF_{ki}U_i \text{ for } k=1 \text{ and } 2. \tag{2.31}$$

(For $N=2$ and $\dim Y=1$, only the combination G_kX_k can be identified from the data, not the **G** and **X** separately). Hence the second order differential equation is equivalent to the two first order equations of the modal analysis. The argument can be generalized to higher orders. Marchio (1986) has developed a linear identification algorithm for the modal analysis, based on equation (2.30). The latter is a special case of the general differential equation equation (2.10), where the format of the driving terms is matched to the modal equations.

To conclude Section 2, we note that the four methods described above are different ways of representing a differential equation; in that sense they are equivalent. The most general equations are equation (2.3) for the ARMA model and equation (2.10) for the differential equation, and we take them as a basis for the rest of the paper. The most important differences between different methods lie not in the underlying equations but in the way they are applied for data fitting, as we shall see. Sections 3.5 and 4.3 will highlight the power of the differential equation which arises from the fact that it can be integrated and solved analytically.

3 The Identification Process

The problem of identification consists of first choosing a model and then determining the coefficients of the model by fitting the data. There are several different ways of doing this, differing for example in the choice of the independent variable, in the way of incorporating constraints, in the criterion for goodness of fit, and in the implementation as a linear or a nonlinear algorithm.

3.1 Choice of Independent Variable. For the analysis of buildings one will usually choose either the interior temperature T_{int} or the heating energy Q_{aux} (although other choices are conceivable, for instance the first derivative of T_{int}). The results can be quite different, as the following simple example shows: suppose the thermostat has maintained the interior temperature at a constant value during the period of data collection. Choosing T_{int} as a dependent variable, one would find a perfect fit with the "model" T_{int} = constant, in other words, one would not learn anything about the parameters of the real system. Q_{aux} would be a much better choice in that case. The dependent variable should show significant variation in response to the driving forces.

3.2 Constraints. In some models some of the coefficients are not independent. For example, the coefficients of the ARMA series of equation (2.3) must satisfy the two constraints discussed in Section 2.2. Different ways of incorporating these constraints can affect the values of the coefficients, but not the physical results (i.e., admittances, time constraints, predicted temperatures, etc.)

To be specific, let us take $T_{int}(n)$ as dependent variable of the ARMA series, with the condition $a_{int}(0)=1$ for the second constraint. The first constraint, equation (2.6), can be included if we replace $a_{int}(N_{int})$ in equation (2.3) by

$$a_{int}(N_{int})=\sum_{k=0}^{N_{ext}} a_{ext}(k)-\sum_{k=0}^{N_{int}-1} a_{int}(k) \tag{3.1}$$

The result is

$$T_{int}(n)=T_{int}(n-N_{int})-\sum_{k=1}^{N_{int}-1} a_{int}(k)[T_{int}(n-k)-T_{int}(n-N_{int})]$$

$$+\sum_{k=0}^{N_{ext}} a_{ext}(k)[T_{ext}(n-k)-T_{int}(n-N_{int})]$$

$$+\sum_{k=0}^{N_{aux}} a_{aux}(k)Q_{aux}(n-k) \tag{3.2}$$

$$+\sum_{k=0}^{N_{sol}} a_{sol}(k)Q_{sol}(n-k)$$

If instead one wants to use the energy $Q_{aux}(n)$ as dependent

variable, the constraint $a_{aux}(0) = 1$ is more convenient than $a_{int}(0) = 1$.

3.3 Statistical Criterion. Traditionally data fitting has been done according to the least squares criterion, i.e., the parameters of the model are varied until the sum of squared differences between data and calculated values is a minimum. The least squares method is very sensitive to anomalous data points, so-called outliers, that may fall outside the range of validity of the model. Recently robust fitting procedures have been developed that de-emphasize the influence of outliers, for instance by minimizing the sum of absolute values of the differences rather than the sum of squares (see, e.g., Hogg, 1979).

3.4 Linear and Nonlinear Algorithms. To explain the distinction between linear and nonlinear data fitting, let us consider equation (3.2). This equation can be used with the *linear* least squares method. One simply inserts *measured* values for *all* the variables and obtains the coefficients $a(k)$ via the standard formulas.

But there is another possibility. Given the driving functions T_{ext} and Q_{aux} (n) as well as one initial value of T_{int}, one can *calculate* the future values of T_{int}. The latter are obtained by repeated application of equation (3.2), time step by time step, and thus they are *nonlinear* functions of the coefficients $a(k)$. Following this approach, a nonlinear regression technique is required; in practice this means some type of numerical search. Nonlinear algorithms have been described by Bacot (1985) and by Subbarao (1984).

Why should one bother with the complications of a nonlinear method if equation (3.2) can be resolved with linear least squares? The interest of the nonlinear method lies in its greater sensitivity. While the linear method regresses against the most recent time period only, the nonlinear method goes all the way back to the initial values. A wrong guess for one or more coefficients may not show up very clearly in the linear method because at each new time step one uses a new set of measured values; the predictions involve only one time step. In the nonlinear approach, by contrast, one tests predictions over extended intervals; hence the increased sensitivity.

As a consequence, with the linear method one risks finding a set of coefficients that are physically wrong but nonetheless appear to make good predictions. While this danger is always present, it should be lower with the nonlinear method.

3.5 Choice of Time Step and Steady State Limit. The finer the time resolution of the data, the smaller the error introduced by the finite difference approximation. Also, only those features of a model can be identified that are associated with time constrants comparable to or shorter than the time step of the data. Hence short time steps may appear desirable. On the other hand, one can enhance the reliability of data by averaging them over extended periods. This can be very useful with the differential equation where the averaging can be done in a way that reduces roundoff errors to a minimum.

Integrating the differential equation (equation (2.10)) from t_1 to t_2 one obtains

$$\alpha_{int,0}\int_{t_1}^{t_2} T_{int}dt - \alpha_{ext,0}\int_{t_1}^{t_2} T_{ext}dt - \alpha_{aux,0}\int_{t_1}^{t_2} Q_{aux}dt$$

$$- \alpha_{sol,0}\int_{t_1}^{t_2} Q_{sol}dt + \alpha_{int,1}\,[T_{int}(t_2) - T_{int}(t_1)]$$

$$- \alpha_{ext,1}\,[T_{ext}(t_2) - T_{ext}(t_1)] - \alpha_{aux,1}\,[Q_{aux}(t_2) - Q_{aux}(t_1)]$$

$$- \alpha_{sol,1}\,[Q_{sol}(t_2) - Q_{sol}(t_1)] \qquad (3.3)$$

$$+ \text{ higher order terms} = 0$$

This form is mathematically equivalent to the differential equation. With this form there will be *no finite difference errors if one has data for both the end points and for the averages of the variables.*

For data fitting equation (3.3) offers some interesting advantages. By judicious choice of t_1 and t_2 one can minimize the influence of certain terms, thereby making the determination of the remaining terms more reliable. For example, uncertainties due to solar gains can be reduced by setting t_1 and t_2 equal to sunrise and sunset. Then only the daily total solar radiation is needed, for which data are much more readily available than on an hourly basis. Furthermore, the troublesome time-of-day variation of the solar aperture is washed out. And the derivative terms of the solar flux can be neglected near sunrise and sunset. Note that data sets with different time intervals can be mixed because the coefficients α are independent of the time step (by contrast to the ARMA coefficients); this is one of the conveniences of the differential equation.

Since the driving forces are quasi-periodic, the role of transients can be minimized by choosing time intervals that are either very long or where initial and final conditions are nearly the same. Sometimes that can be an acceptable approximation even for 24-hour periods. In the steady state limit it is easy to see how equation (3.3) reduces to equation (2.4); there are three variables: $(T_{int} - T_{ext})_{av}$, $(Q_{aux})_{av}$, and $(Q_{sol})_{av}$, averaged from t_1 to t_2, and two parameters:

$$L = \frac{\alpha_{int,0}}{\alpha_{aux,0}} = \frac{\alpha_{ext,0}}{\alpha_{aux,0}} \text{ and } A = \frac{\alpha_{sol,0}}{\alpha_{aux,0}}. \qquad (3.4)$$

The steady-state analysis can be reduced to two variables if one rewrites equation (2.4) in the form

$$\frac{(Q_{aux})_{av}}{(T_{int} - T_{ext})_{av}} = L - A\,\frac{(Q_{sol})_{av}}{(T_{int} - T_{ext})_{av}} + \text{ transients.} \qquad (3.5)$$

This offers the convenience of a two-dimensional plot, where L and A can be found immediately as intercept and slope of the straight line fit.

The fact that the coefficients α do not change with time interval suggeests that the most powerful approach is a sequential procedure: first determine L and A from the steady-state analysis, then insert these values into equation (3.3), using a finer time resolution to identify the higher order terms.

4 Physical Interpretation of Parameters

The coefficients $a(k)$ of the ARMA series contain all the information of the model, and they permit prediction of future behavior. However, their physical interpretation is not transparent, and they can change drastically as the order of a fit is increased. Here we show how to extract quantities that have a more direct physical interpretation: admittances (including L and A), time constants, and heat capacity. We also show how these quantities can be obtained from the coefficients of the differential equation. We recommend admittances and time constants as a standardized way of presenting the results for any building model.

4.1 Admittances and Fourier Analysis. Let us consider the response of the ARMA series to a pure sinusoidal input at frequency ω. After transients have died out, the response of a linear system must also be sinusoidal with the same frequency (see, e.g., Carslaw and Jaeger, 1959). To be specific, let us take exterior temperature in the form

$$T_{ext}(n) = T_{ext,\omega}\exp\,(i\,\omega\,n\,\Delta t) \qquad (4.1a)$$

with complex amplitude

$$T_{ext,\omega} = |T_{ext,\omega}|\,\exp\,(i\,\phi_{ext,\omega}), \qquad (4.1b)$$

with analogous expressions for the thermal inputs

$Q_{aux}(n) = Q_{aux,\omega} \exp(i\ \omega\ n\ \Delta t)$ with

$$Q_{aux,\omega} = |Q_{aux,\omega}| \exp(i\ \phi_{aux,\omega}) \tag{4.2}$$

and

$Q_{sol}(n) = Q_{sol,\omega} \exp(i\ \omega\ n\ \Delta t)$ with

$$Q_{sol,\omega} = |Q_{sol,\omega}| \exp(i\ \phi_{sol,\omega}) \tag{4.3}$$

Then the resulting (nontransient) interior temperature must take the form

$T_{int}(n) = T_{int,\omega} \exp(i\ \omega\ n\ \Delta t)$ with

$$T_{int,\omega} = |T_{int,\omega}| \exp(i\ \phi_{int,\omega}) \tag{4.4}$$

Inserting equations (4.1)–(4.4) into the ARMA series, equation (2.3), one can readily solve for the amplitude and phase of the response. It is convenient to write the answer in the form

$$T_{int,\omega}\ A_{int,\omega} = T_{ext,\omega}\ A_{ext,\omega} + \\ + Q_{aux,\omega}\ A_{aux,\omega} + Q_{sol,\omega}\ A_{sol,\omega} \tag{4.5}$$

where

$$A_{int,\omega} = \sum_{k=0}^{N_{int}} a_{int}\ (k) \exp\ (-i\ \omega\ k\ \Delta t), \tag{4.6}$$

$$A_{ext,\omega} = \sum_{k=0}^{N_{ext}} a_{ext}\ (k) \exp\ (-i\ \omega\ k\ \Delta t), \tag{4.7}$$

$$A_{aux,\omega} = \sum_{k=0}^{N_{aux}} a_{aux}\ (k) \exp\ (-i\ \omega\ k\ \Delta t), \tag{4.8}$$

and

$$A_{sol,\omega} = \sum_{k=0}^{N_{sol}} a_{sol}\ (k) \exp\ (-i\ \omega\ k\ \Delta t), \tag{4.9}$$

The A are complex *admittances*. The ratio of their absolute values determines the amplitude, their phase difference determines the phase lag of the response. For instance, if $Q_{aux}(\omega) = 0 = Q_{sol}\ (\omega)$ then the interior temperature has amplitude

$$|T_{int,\omega}| = |T_{ext,\omega}|\ |A_{ext,\omega}|/|A_{int,\omega}| \tag{4.10}$$

and phase

$$\phi_{int,\omega} = \phi_{ext,\omega} + \text{phase of } A_{ext,\omega} - \text{phase of } A_{int,\omega} \tag{4.11}$$

which means that T_{int} lags T_{ext} by the phase difference

$$\phi_{ext,\omega} - \phi_{int,\omega} = \text{phase of } A_{int,\omega} - \text{phase of } A_{ext,\omega}. \tag{4.12}$$

The (steady-state) heat loss coefficient L and the (steady state) solar aperture A are a special case, corresponding to zero frequency. Setting $\omega = 0$, it is easy to see how equations (4.6)–(4.9) reduce to equation (2.7).

Thus we see that in their combination as admittances, equations (4.6)–(4.9), the coefficients $a(k)$ of the ARMA series have a clear *physical interpretation as amplitude and phase of the response to sinusoidal driving functions*. This fact provides a test of the validity of a fit: for a good fit the admittances should not change significantly if the number of coefficients $a(k)$ is increased further.

Equation (4.5) can also be considered as a relation between Fourier transforms. In the limit where $\Delta t \to 0$ and the number of coefficients $N \to \infty$, the A's of equations (4.6)–(4.9) become the Fourier transforms of the $a(t)$, and equation (4.5) follows directly when the convolution theorem is applied to equation (2.2) and the T_ω, etc. are interpreted as Fourier transforms of $T(t)$. As we have defined the admittances, equation (4.5) follows rigorously from the ARMA series equation (2.3). But the latter is an approximation, valid only as long as all the frequencies are small compared to $1/\Delta t$. In any case, in order to predict the behavior of a building, the time domain will usually be more convenient since the driving forces are not periodic enough to justify a calculation in the frequency domain. As we see it, the principal benefit of the Fourier coefficients (admittances) lies in providing a systematic way of presenting the results and of testing the convergence of the identification process.

In passing we note that Subbarao (1984) used thermal networks containing negative quantities in order to interpolate the admittances from one frequency to another. That is necessary for the forward problem if one takes the admittances at zero and at diurnal frequency as a starting point. But if one starts, as we do here, with the coefficients $a(k)$, then equations (4.6)–(4.9) interpolate automatically between frequencies, without recource to thermal networks.

The admittances can also be obtained directly from the coefficients α of the differential equation. Inserting equations (4.1)–(4.4), with t instead of $n\Delta t$, into the differential equation (2.10), one sees equation (4.5) emerge again, if one identifies the admittances as

$$A_{int,\omega} = \sum_{k=0}^{N_{int}} \alpha_{int,n}\ (i\ \omega)^n \tag{4.13}$$

$$A_{ext,\omega} = \sum_{k=0}^{N_{ext}} \alpha_{ext,n}\ (i\ \omega)^n \tag{4.14}$$

$$A_{aux,\omega} = \sum_{k=0}^{N_{aux}} \alpha_{aux,n}\ (i\ \omega)^n \tag{4.15}$$

and

$$A_{sol,\omega} = \sum_{k=0}^{N_{sol}} \alpha_{sol,n}\ (i\ \omega)^n \tag{4.16}$$

In buildings two frequencies dominate: zero and the diurnal cycle. To characterize the response, one needs for each driving force one parameter at zero frequency and two parameters at diurnal frequency. These are the parameters of Subbarao's BEVA model. Eight of them are independent because of equation (2.6).

4.2 Time Constants The preceding section has dealt with the nontransient response to sinusoidal driving functions. The transient behavior can be characterized by time constants, defined by the response to a sudden change of T_{int}, while all driving forces are held constant. As far as the change is concerned, we can set the driving forces equal to zero. Then the ARMA series reduces to

$$a_{int}(0)\ T_{int}(n) + a_{int}(1)\ T_{int}\ (n-1) \\ + \ldots + a_{int}(N)\ T_{int}(n-N) = 0. \tag{4.17}$$

The solution must have exponential form, with time constant τ,

$$T_{int}(n) = \exp\ (-n\ \Delta t/\tau). \tag{4.18}$$

Inserting equation (4.18) into (4.17) we obtain

$$a_{int}(0) + a_{int}(1) \exp\ (\Delta t/\tau) \\ + \ldots a_{int}(N) \exp\ (N\Delta t/\tau) = 0. \tag{4.19}$$

This is an algebraic equation of order N in $\exp\ (\Delta t/\tau)$ and it has N solutions. Specifically for $N = 1$ there is only one time constant, given by

$$\exp\ (\Delta t/\tau) = -\frac{a_{int}(0)}{a_{int}(1)}. \tag{4.20}$$

(This relation was mentioned by Crawford and Woods, 1985, but it is correct only for $N = 1$.) For a second order model the two time constants are the two solutions of the quadratic equation

$$\exp(\Delta t/\tau)=\frac{-a_{\text{int}}(1)+-\sqrt{a_{\text{int}}(1)^2-4\,a_{\text{int}}(0)\,a_{\text{int}}(2)}}{2a_{\text{int}}(2)}. \quad (4.21)$$

We also state the corresponding expressions for the differential equation. We insert equation (4.18), with t instead of $n\,\Delta t$, into the differential equation (2.10). Cancelling the exponential we are left with an algebraic equation of order N in τ. According to the fundamental theorem of algebrac this equation can be written in terms of the N roots τ_i as

$$(\tau-\tau_1)\,(\tau-\tau_2)\,(\tau-\tau_3)\ldots(\tau-\tau_N)=0. \quad (4.22)$$

Comparison of coefficients yields

$$\begin{aligned}\frac{\alpha_{\text{int},1}}{\alpha_{\text{int},0}} &= \tau_1+\tau_2+\ldots+\tau_N\\ \frac{\alpha_{\text{int},2}}{\alpha_{\text{int},0}} &= \sum_{\substack{i,j=1\\ i\neq j}}^{N}\tau_i\tau_j \qquad (4.23)\\ &\ldots\\ \frac{\alpha_{\text{int},N}}{\alpha_{\text{int},0}} &= \tau_1\tau_2\tau_3\ldots\tau_N.\end{aligned}$$

4.3 Effective Heat Capacity. While the heat capacity of a single node in a thermal network is obvious, the effective capacity of a complex system demands some thought. Different parts have different temperatures and some are closer to the outside than others. How much of the heat stored in the exterior walls of a building can be considered "in the building?" Heat storage is intrinsically a transient heat flow phenomenon. The separation of heat storage and heat loss requires care because the heat loss coefficient has been defined only for steady state conditions. The following argument may appear tortuous, but it does avoid any ambiguities.

First, we remove unnecessary complications by setting solar gains equal to zero and exterior temperatures constant, writing $T=T_{\text{int}}-T_{\text{ext}}$. Then we note how many variables must be specified at time t to characterize the state of a one-zone building that is described by an N'th order differential equation in T: in addition to $T(t)$ one needs all the derivatives up to $T^{(n-1)}(t)$. The most natural definition of effective heat capacity is the heat stored in a building when it has reached steady state conditions, that means

$$T\neq 0 \text{ and } T^{(1)}=0,\ T^{(2)}=0,\ \ldots\ T^{(N-1)}=0. \quad (4.24)$$

To bring the building to this state, we begin at $t=-\infty$ with T and all derivatives equal to zero, adding heat $Q(t)$ until equation (4.24) is satisfied at $t=0$. Thereafter we set $Q=0$, allowing T to decay back to 0 at $t=\infty$. Integrating the differential equation from $0-$ (i.e., $0-\epsilon$ where $\epsilon>0$ is infinitesimal) to ∞ and dividing by $\alpha_{\text{aux},0}$ to convert to energy units we obtain, with $Q(0-)=LT(0-)$ and all derivatives of Q equal to zero at $0-$,

$$\frac{\alpha_{\text{int},0}}{\alpha_{\text{aux},0}}\int_{0-}^{\infty}T(t)\,dt=\frac{\alpha_{\text{int},1}}{\alpha_{\text{aux},0}}T(0-)-\frac{\alpha_{\text{aux},1}}{\alpha_{\text{aux},0}}Q(0-). \quad (4.25)$$

It is tempting to interpret the left-hand side as the heat flow out of the building during this time, which in turn must equal the heat stored at $t=0$, i.e., the product of heat capacity and $T(0)$. Thus the heat capacity C would be

$$C=\frac{\alpha_{\text{int},1}}{\alpha_{\text{aux},0}}-\mathrm{L}\frac{\alpha_{\text{aux},1}}{\alpha_{\text{aux},0}}. \quad (4.26)$$

By virture of equations (3.4) and (4.23) this is equivalent to

$$C=L\left[(\tau_1+\tau_2+\ldots+\tau_N)-\frac{\alpha_{\text{aux},1}}{\alpha_{\text{aux},0}}\right] \quad (4.27)$$

The term $\alpha_{\text{aux},1}/\alpha_{\text{aux},0}$ could be interpreted as sum of time constants for Q.

To prove that this interpretation is indeed correct, despite the lack of steady state conditions, we also look at the integral from $-\infty$ to 0–

$$\frac{\alpha_{\text{int},0}}{\alpha_{\text{aux},0}}\int_{-\infty}^{0-}T(t)\,dt+\frac{\alpha_{\text{int},1}}{\alpha_{\text{aux},0}}T(0-)=\int_{-\infty}^{0-}Q(t)dt. \quad (4.28)$$

The right-hand side represents the total heat input and by the first law it must equal the total heat loss from $-\infty$ to $+\infty$. There is no question about interpreting the integral from $-\infty$ to $+\infty$, for any Q that vanishes identically at $-\infty$ and $+\infty$

$$\frac{\alpha_{\text{int},0}}{\alpha_{\text{aux},0}}\int_{-\infty}^{\infty}T(t)\,dt=\int_{-\infty}^{\infty}Q(t)\,dt \quad (4.29)$$

as heat loss because initial and final state are the same. If we can choose a heat input $Q(t)$ in such a way that $T(-T)=T(t)$, then it follows by symmetry that equation (4.25) is indeed the heat loss from 0 to ∞. To find the required $Q(t)$ we have to solve the differential equation

$$\alpha_{\text{aux},0}Q+\alpha_{\text{aux},1}\,Q^{(1)}+\ldots+\alpha_{\text{aux,Naux}}\,Q^{(\text{Naux})}=f(t)$$

over the interval $-\infty$ to $0-$, where $f(t)$ is a function that is determined by $T(t)$ during the decay. A solution exists as shown by any book on linear differential equations with constant coefficients. For $N_{\text{aux}}\geq 1$ this $Q(t)$ may differ from the one used for equation (4.25), but that is all right since at this point we only need to justify the interpretation as heat loss. That completes the proof.

A word about the interpretation of our result for C. It is only as exact as the coefficients of the differential equation. In practice the latter are determined by means of data obtained under normal diurnal cycling, a condition that does not give enough time for a heat pulse to penetrate all parts of a massive building. Thus C is an effective heat capacity; it is a measurement of the heat storable under normal cycling. For slower cycling C would increase, and so would the time constants of the data fit. This illustrates to what extent the parameters of a model are dependent on the frequencies represented in the data.

5 Difficulties

In principle the methods outlined above are straightforward, but the practical implementation faces many pitfalls. Without any claim to completeness we discuss the most common difficulties; further problems will become apparent in the following section. Some arise from incomplete or erroneous data, others from faults of the model (e.g., the modeling of solar gains, or the assumption that model parameters are independent of time).

5.1 Variable Air Exchange Rate. In most buildings the air exchange rate varies, with weather and/or operating mode. This means that the most important parameter, the heat loss coefficient L, contains a variable component. As an indication of the kind of error introduced by assuming a constant heat loss coefficient, suppose that air exchange accounts for half of the total heat loss and that the exchange rate varies by + − 50 percent; then the total heat loss coefficient varies by + − 25 percent. Such numbers are typical.

If one has data for the air exchange rate m, one can subtract the term

$$Q_{\text{air}}=(m\,c_p)\,(T_{\text{int}}-T_{\text{ext}}),$$

$$\text{with } c_p=\text{specific heat of air}, \quad (5.1)$$

from the total heat input and then the analysis is straightforward. But how can one get such data? With current technology it involves an SF_6 tracer gas apparatus that is expensive and labor intensive. A new and simpler technique, based on perfluorocarbon tracers, has recently been developed

(Dietz et al., 1985). Well suited for average measurements over extended periods, it does not yet appear to be simple enough for routine real time measurements in hourly intervals. In the foreseeable future it is better to try and get by without hourly air exchange data. That is facilitated by the trend in new construction towards increased air tightness coupled with mechanical ventilation.

In buildings without mechanical ventilation, the air exchange rate increases with wind speed v and with $(T_{int} - T_{ext})$, in approximately linear fashion. It would be straightforward to augment the differential equation or the ARMA series with terms proportional to $v\ (T_{int} - T_{ext})$ and $(T_{int} - T_{ext})^2$. The nonlinearity introduced by the latter term is benign and would not cause any fundamental difficulties. As a simpler compromise one could assume one value for summer and another one for winter.

For office buildings with mechanical ventilation during occupied periods, the data of Persily and Grot (1985) suggest that the air exchange rate is likely to be fairly constant, within perhaps twenty percent, in the heating and in the cooling season. But the economizer mode or open windows can easily increase the rate by a factor of five, with much irregularity and variation from one building to another (for open windows, see data in Section I.13 of NJECL, 1986). During unoccupied periods the air exchange rate is likely to be lower than the occupied value, dependent on wind velocity and temperatures. These considerations suggest that one should select data without economizer or open windows, where it may be a fair approximation to assume two values of the air exchange rate, one when the building is occupied and one when it is not. However, in many buildings the economizer is used at temperatures as low as 0°C. Hence it would be desirable for the model to include the economizer.

That turns out to be quite simple, provided the ventilation system functions according to the standard economizer design. In that design the supply fan provides a constant flow rate m_{sup} at constant temperature T_{sup}. For the latter a value around 15°C is typically chosen to prevent the building from overheating (exact values depend on specifics of the building). Over an exterior temperature range from $T_{ext,min}$ (about 0°C) to 15°C one maintains T_{sup} at the desired value by mixing varying amounts of outside air with return air from the building. This mode of operation is called economizer because no air conditioning is required (by contrast with the older design with constant outside air flow rate). A simple energy balance at the junction of return air, outside air and supply air yields the outside air flow rate

$$m = m_{sup}\ (T_{int} - T_{sup})/(T_{int} - T_{ext}) \quad \text{for } T_{ext,min} < T_{ext} < T_{sup} \tag{5.2}$$

with

$$T_{ext,min} = T_{int}\ (1 - m_{sup}/m_{min}) + T_{sup} m_{sup}/m_{min} \tag{5.3}$$

where m_{min} is the minimum required for maintaining acceptable air quality. These equations assume that the return air is at T_{int}, a very good approximation. We have compared this equation with measured air exchange rates for the Enerplex Office Buildings (described in Section 6) and found good agreement (within twenty percent most of the time). Equation (5.2) can easily be included in the model of a building.

To account for the physics one should keep all coefficients of the model constant except for the air exchange term. With the differential equation that is done automatically if one allows only the term α_{int} $(= \alpha_{ext}$, of course) to vary between periods. In the ARMA all a would be affected, but one can impose consistency between the data fits by adding a term $\Delta m c_p\ (T_{int} - T_{ext})$ to Q_{aux} for one of the two periods. By varying Δm one can find the best value by trial and error.

5.2 Correlated Data. The independent variables should not be correlated with each other. In this regard weather data are often problematic: exterior temperature tends to be correlated with solar radiation. This can produce effects such as the following which was observed when using the 1R1C network to fit data of the Enerplex office building for unoccupied summer periods (Norford et al., 1985). The solar input raises T_{int} above T_{ext}, and for the heat capacities involved, the difference $(T_{int} - T_{ext})$ was large at night, small during the day. In this situation the statistical analysis for certain fits found a strong negative correlation between solar radiation and $(T_{int} - T_{ext})$: the product of heat loss coefficient and solar aperture was negative, clearly unphysical. Similar problems were encountered by Curtiss (1986) when applying these methods to solar pond data.

In principle it is possible to eliminate the effect of partial correlations among data by redefining the variables according to the Gram-Schmidt orthogonalization process of linear algebra. That is beyond the scope of the present paper.

5.3 Solar Radiation. The definition of the solar aperture is straightforward if all windows of a building face the same direction and if one has data for the solar radiation incident on the plane of the windows. But in practice only the hemispherical radiation (i.e., sum of direct and diffuse radiation) on the horizointal surface will be known. Hence the usual approach is to define the solar aperture relative to the hemispherical radiation on the horizontal surface by assuming equation (2.5) for the solar gain. In that case the solar aperture is likely to vary with time of year.

For a more accurate approach one would have to separate the contributions of direct and diffuse radiation for each of the window orientation. Direct and diffuse radiation are rarely measured, but they could be estimated by means of standard correlations (Rabl, 1985). Trying to identify the separate apertures for each window orientation separately would probably not be feasible on the basis of performance data alone: the number of parameters would become excessive. More promising would be a hybrid approach: measure the window areas for each orientation directly, and use these numbers to calculate Q_{sol} as an equivalent building-weighted measure of solar radiation (K. Subbarao and J. M. Gordon, personal communications). Another solution has been mentioned in Section 3.5: if one integrates the differential equation between sunrise and sunset, the time of day-variation of the solar aperture is averaged out.

In this paper we have only tried the simplest model, i.e., equation (2.5) based on hemispherical radiation on the horizontal. Even though the resulting solar aperture is not well defined, the solar gains are also relatively small in most buildings (perhaps ten to thirty percent of the total heat input) in winter. To that extent the analysis of solar gains may be difficult and unimportant at the same time.

5.4 Sky and Air Temperature. So far we have treated T_{ext} as if it were a single well defined quantity. In reality the surrounding air, the ground, and the sky may all be at different temperatures, with different heat transfer coefficients for each. In clear weather the difference between sky and air temperature can exceed 30°C. Thus the choice of T_{ext} can become problematic, especially in summer when the uncertainty becomes large compared to the difference between interior and exterior air temperatures. The dilemma is similar to that of the solar aperture: an accurate model would require too many parameters.

5.5 Multiple Zones. At the design stage (forward problem) the analysis of separate zones is crucial, because the HVAC system must be designed to have sufficient capacity in each zone. But for the inverse problem differences between zones are as difficult to analyze as they are irrelevant. After all, if the HVAC equipment is operating properly it will maintain

Table 6.1 Tests of dynamic methods for inverse problem

a) with measured data

AUTHOR	METHOD	BUILDING
Sonderegger [1977, 1978]	network (2R1C)	townhouse
Pryor and Winn [1982]	network (2R2C)	passive solar house
Janssen [1982]	network (1R1C)	house
Forrester and Wepfer [1984]	ARMA	office building
Norford et al. [1985]	network (2R1C) ARMA	office building
Crawford and Woods [1985]	ARMA	house
Subbarao [1985]	ARMA	test cell, and house
Wilson et al. [1985]	network (2R1C)	house
Norlen et al. [1986]	network (3R2C)	test cell

b) with calculated data

Bacot [1985]	modal analysis nonlinear	passive solar house
Marchio [1986]	modal analysis linear	passive solar house
Ryan [1986]	networks	house

Comments:

Pryor and Winn [1982], Crawford and Woods [1985], and Norlen et al. [1986] have presented an efficient algorithm for on-line identification, improving the estimates as more data are added, with minimal computation.

the entire building at more or less the same temperature. To the extent that the interior temperature is uniform, there is only a single zone, as far as inverse models are concerned. Even the occurrence of simultaneous heating and cooling does not matter, if one has separate measurements for each.

The smaller the temperature differences between zones, the smaller the heat flows between them and the more difficult their analysis. Only in the case of semi-conditioned spaces, such as basements or artria, does it make sense to attempt a zonal analysis (see Subbarao, 1985 for an example). In the case of the atria of the Enerplex Buildings a zonal analysis has not been successful (Burch et al., 1986).

6 Test With Data

To see how well the various methods work in practice, one must test them. Table 6.1 lists several such tests reported in the literature, grouped according to measured or synthetic data. While measured data from real buildings are the ultimate criterion, it is difficult to sort out the effect of data errors from the effect on inadequacies of the model. Therefore, one needs tests with data that have been calculated (computer experiments, in other words). That seems to be the only practical way of carrying out systematic tests of inverse models under controlled conditions. Tests with measured data play a complementary role, checking the realism of the assumptions.

Evidently there have been numerous tests already, but almost all are limited to one method at a time. In presenting an overview of all methods, we would like to report a more comprehensive evaluation. Also, the previous tests have dealt with simple buildings (test cells or houses). As we are interested in the uncharted regime of commercial buildings, an exploratory analysis with measured data is appropriate at this point, and most instructive, as we shall see.

6.1 The Enerplex Office Buildings. We have chosen data for the Enerplex Office Buildings, a set of two nearly identical office buildings of 12,000 m^2 floor area in Princeton (some 100 km south of New York). Completed in 1984, they have been instrumented with about 100 data channels, recorded in hourly intervals (Norford et al., 1984, 1985). The instrumentation includes a tracer gas (SF_6), apparatus for measuring the air exchange rate. Due to the difficulties of obtaining tracer gas data, that apparatus was turned on only during certain periods of several weeks each. As is to be expected, not all the channels are recording correctly all of the time (Bloch, 1980). After checking the data for completeness (in the sense of including all the variables needed for our analysis) and consistency, we are left with about 14 days of data (with 15 gaps) for the South Building and 19 days (with 10 gaps) for the North Building. Due to gaps the longest continuous interval is 6 days. The data are for winter 1985/86. The North Building is fully occupied and its interior temperature ranges from about 19 to 25°C. In the South Building T_{int} is lower (about 16 to 20°C) because only one of the three floors is fully occupied and conditioned.

Despite the large number of data channels we remain in the dark about certain pieces of equipment that we could not afford to monitor continuously. One such item is the computer room which has its own dedicated chiller. Since the temperature in the computer room is close to that of the rest of the building, we have assumed that the chiller removes exactly the heat generated by the computer. Their operation is fairly constant as far as we could tell from spot measurements with a portable power meter. So we directly subtract their consumption as a constant rate (60 kW) from the total thermal input to the building (400 to 800 kW in winter). There is a cafeteria in the building, and the air removed by the kitchen exhaust is warmer than the average of the building; thus the heat loss is larger than implied by the tracer gas air exchange data. Unfortunately, we cannot monitor the exhaust flows and temperatures of the kitchen. Additional uncertainty stems from lights and heaters in atrium and entrance ways (most of whose heat does not go into the building), from a large satellite antenna, and from outdoor lighting (60 kW).

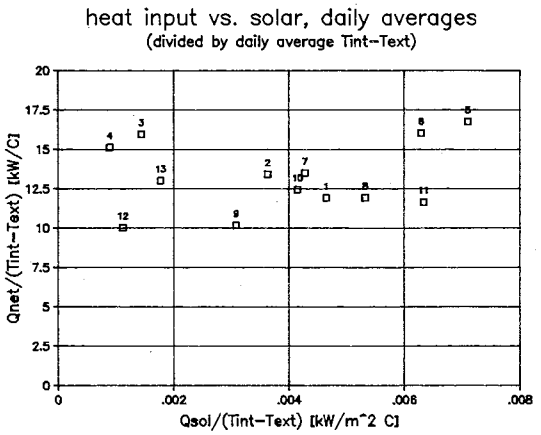
Considering the rms fluctuations about the most probable values, we estimate that the energy input has an error of about 30 kW. Compared to the total, that is only 5 percent. But the relative error doubles after we subtract the heat loss due to air exchange, taking Q_{aux} as

$$Q_{aux} = Q_{tot} - (m\ c_p)\ (T_{int} - T_{ext}). \tag{6.1}$$

Since the scale of L is set essentially by Q_{aux}, this implies an absolute error around 2 kW/C for the observed magnitude of L. The only other error of comparable importance comes from the air exchange which has an uncertainty of about + − 0.1 per hour (Persily, personal communication). Noting that for the Enerplex Buildings 1 air change per hour corresponds to 15 kW/C, and combining the errors quadratically, we conclude that the total error in L is about 2.5 kW/C, or 15 to 20 percent.

One of the lessons is that accurate data for the thermal input to a large commercial building may be hard to come by, even in buildings with electric resistance heat (the North Building). The problem is further exacerbated for heat pumps (the South Building) or furnaces where one has to measure flow rate and temperature difference. Often the electric wiring is not laid out according to end use, and it may be difficult to disaggregate the energy use and monitor flows to the outside of a building.

All the results reported in the following excluded the air exchange terms as per equation 6.1 (with one clearly marked exception, the summer 1984 data). Therefore, L is the conductive part of the heat loss coefficient. Similarly the time constant is based on the conductive part only. The total heat loss coefficient is the sum of conductive and air exchange terms, the latter varying from about 6 to 22 kW/C; the corresponding time constant would be about half the one for conduction alone. We have subtracted the air exchange term in order to minimize uncertainties; that is feasible for a research project, but for routine applications it will be better to include the air exchange, along the lines discussed in Section 5.1.



day	T_{int}-T_{ext} [C]	Q_{aux} [kW]	Q_{sol} [kW/m^2]	ΔT_{int} [C]	ΔT_{ext} [C]
1	20.9	248	.097	-.33	-3.47
2	19.8	265	.072	.02	-3.90
3	14.2	226	.020	-.20	1.02
4	17.2	261	.015	-.43	-5.07
5	17.1	286	.121	-.05	-4.47
6	17.6	282	.111	.41	5.12
7	22.0	296	.094	-.47	-8.21
8	25.9	309	.138	.56	-.58
9	24.1	245	.074	-.88	8.81
10	16.4	203	.068	.16	-1.91
11	20.2	235	.128	.99	-.48
12	22.7	227	.025	-.20	-.20
13	20.1	262	.036	.26	.87

Fig. 6.1 Steady state analysis of Enerplex North Building, winter 1985/86. Numbers are listed underneath figure, including changes of interior and exterior temperatures ΔT_{int} and ΔT_{ext} (value at 24:00 minus value at 0:00). Days are consecutive, except for gaps after day 2 and day 7.

6.2 Steady State Analysis. We begin the data analysis by applying the steady state limit, equation (3.5), to the data of the North Building (there are more quasi steady-state periods in the data set for the North than for the South Building). The data are plotted in Fig. 6.1 with labels for each of the 13 days. If transients were negligible and if the solar aperture were well defined, the data should fall on a straight line with intercept L and slope $-A$. Unfortunately, the scatter is too large for a clear picture to emerge.

To see whether the outliners, days 5, 6, and 12, can be explained in terms of transients, we have listed under Fig. 6.1 not only the daily averages but also the changes (from 0:00 to 24:00) of T_{int} and T_{ext}. The effective heat capacity remains uncertain even after the dynamic analyses reported in Tables 6.2 and 6.3, but it must be on the order of 500 kWh_t/C. Then a 1°C rise in T_{int} during 24 hours corresponds to an average rate of 500 kWh/24 h = 21 kW, and Q_{aux} would have to lie that much above the straight line trend of equation (3.5). In this data set T_{int} rarely changes more than 0.5°C and never more than 1°C, hence the asociated scatter should be only around five percent.

Changes in T_{ext} can also have an effect because of the heat capacity of the exterior skin of the building. For an estimate we recall from equation (4.10) that the change in T_{int} produced by a diurnal change ΔT_{ext} can be found by multiplying ΔT_{ext} by the admittance ratio A_{ext}/A_{int}. The latter is quite uncertain, but the numbers in Table 6.2 suggest an absolute value around 1/4. Thus the transient effects from variations of exterior temperature in Fig. 6.1 could contribute as much as + − 40 kW. If T_{ext} rises, the point should lie above the trend line because the exterior skin is effectively colder than implied by the daily average of T_{ext}.

Whereas day 6 is consistent with this explanation, day 5 should be far on the opposite side of the trend line, and day 12 should be very close to it. Maybe one could blame uncontrolled infiltration of air through the mass of the building envelope (Norford and Persily, 1987). Depending on the direction of the air flow, the storage of heat in the envelope could be enhanced or reduced; we have no data for that.

6.3 Dynamic Analysis. For the dynamic analysis we have tested various orders of the ARMA model and of the differential equation (equation (2.10)) with finite differencing according to equation (2.12). Results are summarized in Table 6.2, both for T_{int} and for Q_{aux} as dependent variable. Note that the 1R1C network corresponds to the 1000 fits of the differential equation. The most egregious feature of these results is the enormous scatter, far outside the range implied by the standard errors of the statistical analysis: L ranges from 16 to 24 kW/C, A from 518 to 2987 m^2, and τ from 13.5 to 118.6 hr (for the higher order models only the largest time constant is listed). Large scatter has also been found by Marchio (1987) who analyzed these data with modal analysis, using both the linear (Marchio, 1986) and the nonlinear (Bacot, 1985) algorithms (the nonlinear version did not converge to reasonable values for this case).

As emphasized by Subbarao (1985), a necessary condition for a good fit is the stasbility of the physical parameters (admittances and time constants) as more coefficients are added to the model. While that condition was well satisfied by the data of Subbarao [1985], the trends in Table 6.2 are confusing and in some cases unphysical.

Even worse, there seems to be no clear criterion for selecting the best fit. We have listed both R^2 (adjusted for degrees of freedom; Draper and Smith, 1981) and standard errors. For the latter there are two possibilities: the standard error during the fit (T_{int} calculated from measured T_{int} at previous hour), and the standard error in the predictive mode (all T_{int} calculated from initial T_{int} at $t=0$). The difference between these two standard errors corresponds to the difference between linear and nonlinear fitting discussed in Section 3.4.

The differential equation, as used for Table 6.2, seems to lead to large standard errors, and worst of all, exponential divergence for the second order fit in T_{int}. That is a consequence of the symmetric finite differencing scheme that we have chosen: it causes the prediction for T_{int} to depend most strongly on the coefficient that is least well determined (the one for the highest derivative). A wrong sign leads to divergence. The results of Table 6.2 were derived several months ago, for the preprint version of this paper, before we had recognized the full potential of the approach outlined in Section 3.5. Unexpectedly caught by a tight deadline for getting the final version to the publisher, we have not had the time to try this new approach with the Enerplex data. We did, however, use it successfully with data measured in a house during August and September: despite the small temperature differences we obtained results that were very reasonable and consistent.

An interesting feature of the results is revealed by plotting L versus A, in Fig. 6.2. The dynamic results are plotted as points. The steady state analysis fixes only a linear constraint; the two lines shown correspond to two different steady state periods. There is a linear trend, implying that a fit can compensate too high a solar gain by too high a heat loss while maintaining good agreement for T_{int} and Q_{aux}. From all the evidence, including Table 6.3, the solar aperture in winter appears to be in the range of 400 to 1000 m^2. The corresponding values of L lie between 15 and 18 kW/C. Had we omitted the solar gains altogether, L would have turned out 10 to 20 percent lower, corresponding to the $A = 0$ intercept of the trend line.

The picture becomes even more perplexing when we consider what values for the heat loss coefficient L had been found in previous analyses of the Enerplex Buildings, summarized here in Table 6.3. At the design stage the prediction

Table 6.2 Summary of fits to measured data for Enerplex South office building (floor area 12,000 m^2) winter 1985/86; 345 hourly data points, with 15 gaps. Heat of air exchange has been subtracted before fitting, thus heat loss coefficient L and time constant τ include conduction only. τ is largest time constant only. A = steady state solar aperture, equation (2.5).

TYPE OF FIT			QUALITY OF FIT				PARAMETERS					
Method	depend. variab.	Order of fit[a] i e a s	R^2 (adjusted)	standard error fit[b] [C]	[kW]	test[c] [C]	0 frequency L [kW/C]	A [m^2]	diurnal frequency[d] A_{int}/A_{aux} [kW/C;°]	A_{ext}/A_{aux} [kW/C;°]	A_{sol}/A_{aux} [m^2;°]	τ [hr]
Dif.Eq.	T_{int}	1 0 0 0	0.931	5.4		2.6	17.7	668	75;76°			15.7
		1 1 1 1	0.934	5.3		2.3	17.9	751	81;69°	18;-20°	746;- 3°	17.0
		2 2 2 2	0.941	5.0		bad[e]	18.2	786	80;68°	14;-23°	788;- 5°	16.1
	Q_{aux}	1 0 0 0	0.917		90.6		16.0	518				14.5
		2 1 1 1	0.919		89.6		16.2	563	62;67°	17;-17°	561;- 5°	14.0
		2 2 2 2	0.926		85.7		16.4	569	64;69°	14;-22°	593;- 5°	14.0
ARMA	T_{int}	2 0 0 0	0.878	0.13			19.5	1,685	367;86°			70.4
		2 1 1 1	0.910	0.11		1.1	22.6	2,636	499;52°	17;-74°	2,319;- 8°	118.6
		2 2 2 2	0.914	0.11		1.2	24.0	2,987	362;43°	12; 79°	1,803;-18°	118.5
	Q_{aux}	2 0 0 0	0.913		93.5	2.0	16.6	615	94;68°			21.5
		2 1 1 0	0.989		32.6	3.6	19.2	1,315	29; 9°	9; 87°	531;-59°	13.5
		2 2 2 2	0.991		30.3	2.6	18.7	1,175	41;19°	11; 58°	573;-35°	21.0

notes:
(a) numbers in these colums indicate order of differential equation or - 1 + number of terms in ARMA model, i for T_{int}, e for T_{ext}, a for Q_{aux} and s for Q_{sol}.
(b) rms error of dependent variable during fit.
(c) rms error of T_{int} in predictive mode.
(d) ratio of the admittances defined in Eqs.4.6 - 4.9.
(e) exponential divergence.

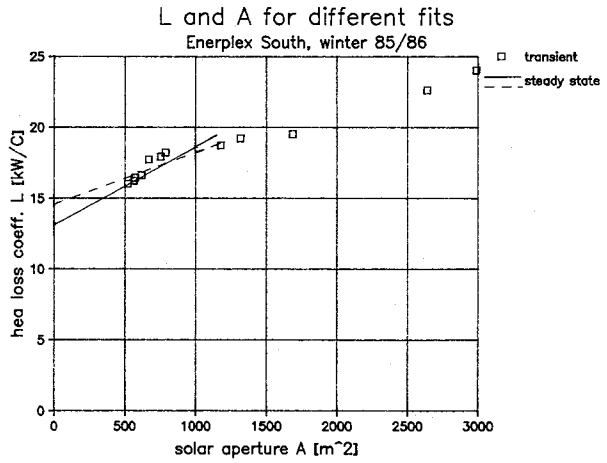


Fig. 6.2 Steady state heat loss coefficient L and solar aperture A for Enerplex South Building, winter 1985/86, as obtained by different methods. Steady state analysis yields only a constraint, shown by straight lines for two different data periods. Transient points are from dynamic analysis reported in Table 6.2.

was L = 10.0 kW/C, based on dimensions and materials in the blueprints. The first data, for the South Building, summer 1984, were quite consistent with that because they included a small amount of air exchange (not measured but estimated to contribute between 0.15 to 0.4 kW/C); taking errors and air exchange into account, that would put the conductive L between 8 and 12 kW/C. While the intervening January and February 1985 numbers for the South Building are unreliable because the data period is too brief, the result for winter 1985/86 seems firm enough to be taken seriously. Thus there is a curious development: between 1984 and 1986 the conductive L for the South Building appears to have increased from 10 (+ −2) to 16 (+ −2.5) kW/C. For the North Building the starting point is later and higher (14 kW/C, if corrected for solar, during winter 1984/85), but there also appears to be an increase to 16 kW/C (data in Fig. 6.1, as well as independent analyses by Marchio, 1987, Rabl and Subbarao, not presented here).

Increasing air leakage cannot be invoked as an explanation since its effect has been measured and removed from these data. As there have been no retrofits of the building shell, we can imagine only two possible explanations. One is that the occupants might keep the venetian blinds more closed than in the beginning, thus reducing the solar aperture; the data fit could misinterpret that as increased heat loss. The other possibility is gradual accumulation of moisture in the roof. Such moisture could act like a heat pipe, evaporating below and condensing above. While we have seen some external moisture damage of the roof, we cannot tear the roof apart to verify this hypothesis.

6.4 Errors. A comment about errors: There are several possible estimates, none entirely satisfactory. While random errors reflect themselves in the standard errors provided by the statistical analysis, systematic errors cause a bias that remains hidden. There are errors in the data and there are errors in the model; for neither can we separate systematic and random contributions. In the interest of keeping this paper within almost finite limits, we have not listed all coefficients and their standard errors for all the fits. For the Enerplex Buildings the standard errors vary enormously from one coefficient and one fit to another, ranging from a few to over one hundred percent. The leading coefficients of the differential equations (i.e., the lowest order) tend to have smaller errors than the ARMA coefficients. For instance, the standard error of the heat loss coefficient is around ten percent for the differential equation. In the ARMA model the physical quantities are combinations of coefficients whose signs alternate in practice; as a result the total relative error can easily become absurdly large when the standard errors are combined according to the usual rules. A proper determination of the errors seems to require a numerical search as complicated as the nonlinear fitting procedures. By contrast, the differential equation can yield the errors of the physical quantities directly.

6.5 Calibrated Computer Simulation. Finally we present in Fig. 6.3 a comparison between predicted and measured Q_{aux} for the $N_{int} = N_{ext} = N_{aux} = N_{sol} = 2$ ARMA fit of the North Building. The lower set shows Q_{aux} without air exchange, i.e., according to equation (6.1), as fit to the data. The upper set was obtained from the lower set by restoring the air exchange contribution and the electricity consumed outside the building. In other words, the upper set shows the total elec-

Table 6.3 Previous analyses of Enerplex Buildings; L = conductive part of heat loss coefficient (except for set 2, note *a*)

DATA PERIOD (BUILDING)	METHOD	PARAMETERS: 0 frequency L [kW/C]	0 frequency A [m²]	diurnal frequency A_{int}/A_{aux} [kW/C;°]	diurnal frequency A_{ext}/A_{aux} [kW/C;°]	diurnal frequency A_{sol}/A_{aux} [m²;°]	heat capac. [kWh/C]	τ [hr]	REFERENCE
1) -- (South)	theory (from (blueprints of building)	10.0	400 (summer) 600 (spring,fall) 900 (winter)				500		[Norford et al. 1984]
2) summer 1984 weekends (South)	network (2R,1C)	12.0[a]	200				700	65	[Norford et al.
	ARMA	13.7[a]	239	255;78°	22; 45°	248; 10°			1985]
3) fall 1985 (South)	direct measure of radiation through windows		250 (fall)						[Proto-papas, 1985]
4) Jan,Feb 1985 23 days (North)	steady state	12.2[b]							[Norford et al. 1985]
1 night (South)	steady state	12.1							
8 days (South)	ARMA	10.8[c]	35	96;76°	13; 19°	90;- 4°		34	

Notes:
a) includes air exchange term (unknown but probably between 1.5 and 4 kW/C).
b) lower bound, because no data to separate solar contribution; solar could increase that value to 14 kW/C.
c) uncertain because data for only 8 hours per day, daytime.

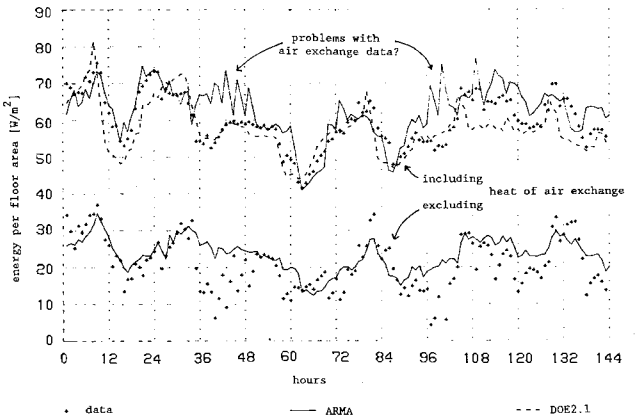


Fig. 6.3 ARMA fit and data for Enerplex North Building, 00:00 Friday 31, Jan-24:00 Wednesday 5 Feb. 1985. Upper set includes, lower set excludes heat of air exchange. Upper set also shows calibrated computer simulation with DOE2.1 (Hsieh, 1988). I think E. Hsieh for making the results of her work available and T. Stein for preparing the plot.

tricity consumption of this building (it is heated by electric resistance). Superimposed is the result of a calibrated computer program (Hsieh, 1988). The latter uses DOE2.1c, carefully adjusting air infiltration and ventilation (average rates only), shading coefficients, thermostate schedules, and operating schedules. The data cover six consecutive days at the end of January.

By and large both predictions seems comparably close to the data. We do not compare rms derivations because of doubts about certain periods, e.g., hours 36 through 50 and 96 through 110 when the air exchange data are questionable. During these periods the ARMA prediction is smooth without, jagged with the air exchange term. The zigzag pattern has a three hour cycle, wich appears correlated with the fact that the tracer gas apparatus reinjected a fresh supply of tracer gas once every three hours. During these periods the data points show the zigzag pattern only after the air exchange subtraction. If the air exchange data were true, either the total consumption or the interior temperatures would also exhibit a zigzag behavior, which is not the case. The computer simulation for these periods is smooth because it did not use measured hourly air exchange data – and for the ARMA fit that would probably have been the better choice, too. In any case, a simple ARMA model seems capable of reproducing the data as well as a massive computer simulation.

7 Conclusions

We have provided an overview of the different methods for parameter estimation in buildings. While all the methods are basically equivalent to the problem of identifying the coefficients of a linear differential equation, there are many different ways of implementing the identification process. Differences arise, amongst others, from the choice of dependent variable, the number of parameters, the choice of the finite differencing scheme, and the treatment of solar gains. Furthermore, there are linear and nonlinear algorithms; the latter promise better accuracy, but at far greater computational effort (and not always delivering better results). As basic starting point we recommend either the differential equation (equation (2.10)) or the ARMA (autoregressive moving average) model. Both are simple, and parameters can be added in a systematic way to improve the fit. In either case we urge that the results be given a standardized physical interpretation in terms of time constants and admittances. While the ARMA model is well suited for numerical work, the differential equation is more powerful because it can be integrated analytically. With the differential equation we have developed a methodology (Section 3.5) that, we believe, can yield better results than the ARMA model with linear algorithm.

Whereas previous investigators have usually obtained "good" fits (with rms errors below 1°C for T_{int}), our results are puzzling. There are two major differences from previous work: we use a larger number of different methods, and the data come from a building that is more complex. The complexity arises from energy-related activities in a large building that escape measurement (e.g., outdoor lighting, air conditioned computer rooms, kitchen exhaust), and from the window geometry. In the Enerplex Buildings with skylights, large windows on all sides, and extensive shading devices, the solar gains are far more complicated than in the test cells and passive solar houses that have been the subject of previous studies.

The basic difficulty of the identification process stems from the possibility of compensating errors. The information content of data from real buildings is very limited, and one can obtain a good fit even with wrong parameters. Such a fit may be adequate for the purpose of controlling HVAC equipment under standard, repetitive operating conditions, but it cannot be trusted to evaluate changes.

In view of the risk of compensating errors between heat loss coefficient L and solar aperture A, one may wonder if it may be safer to neglect solar gains altogether, when the window geometry is complicated. But then one pays with a systematic

underestimate of L, easily more than 10 percent even for non-solar buildings in cold weather. In mild or hot weather solar gains dominate conduction. For the understanding of cooling loads solar gains are crucial. Hence further refinement and testing of the identification process is needed. Perhaps much of the scatter in the results can be blamed on two faults of the current approach: the crudeness of the solar aperture model, and the correlations between solar radiation and exterior temperature. That, as well as the systematic exploration of the approach in Section 3.5, and the modeling of air exchange, remains subject for future research.

Acknowledgments

Supported in part by the New Jersey Energy Conservation Laboratory and by the Prudential Insurance Co. Some of this work has been carried out at the Ecole des Mines de Paris whose hospitality is gratefully acknowledged. A. Persily of the National Bureau of Standards has kindly provided air exchange data for the Enerplex Buildings, and J. Spadaro has been most helpful in preparing the data in a convenient format. I have enjoyed stimulating discussions with J. Burch, J. Gordon, D. Marchio, A. Neveu, L. K. Norford, and K. Subbarao.

References

ASHRAE, *Handbook of Fundamentals*, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, 1981.

Bacot, P., "Identification de Modeles de Comportement des Systems Thermiques, *Revue Generale de Thermique*, No. 277, 1985, p. 15.

Bacot, P., Neveu, A., and Sicard, J., "Analyse Modale des Phenomenes Thermiques en Regime Variable dans le Batiment," *Revue Generale de Thermique*, No. 267, 1984, p. 189.

Bloch, A., *Murphy's Law and Other Reasons Why Things Go Wrong*, Price, Stern, Sloan, Publishers, Los Angeles, 1980.

Box, G. E. P., and Jenkins, G. M., *Time Series Analysis: Forecasting and Control*, Holden-Day, Inc., Oakland, CA, 1976.

Burch, J., Christensen, C., and Subbarao, K., 1986, "Solar in Federal Buildings Program: Initial Results From Short-Term Monitoring in Enerplex South," Report 254/171/1-7-86, Solar Energy Research Institute, Golden, CO.

Carter, C., "A Validation of the Modal Expansion Method of Modeling Heat Conduction in Passive Solar Buildings," *Solar Energy*, Vol. 23, 1979, p. 537.

Carslaw, H. S., and Jaeger, J. C., *Conduction of Heat in Solids*. 2nd Ed., Oxford University Press, London, 1959.

Crawford, R. R., and Woods, J. E., "A Method for Deriving a Dynamic System Model from Actual Building Performance Data," Paper HI-85-37 No. 1 presented at ASHRAE Meeting Honolulu, June 1985; to be published in ASHRAE Transactions 1985, V. 91, Pt. 2.

Curtiss, P., "The Salt Gradient Solar Pond as an Alternative Energy Source," Senior thesis. Dept. of Civil Engineering, Princeton University, Princeton, NJ, Dept. of Civil Engineering, Princeton University, Princeton, NJ, Apr. 1986.

de Larminat, P., and Thomas, Y., *Automatique des Systemes Lineaires*, Flammarion Sciences, Paris, 1975.

Diamond, S. C., Hunn, B. D., and Cappiello, C. C., "The DOE-2 Validation," *ASHRAE Journal*, Nov. 1985, p. 25.

Dietz, R. N., D'Ottavio, T. W., and Goodrich, R. W., "Multizone Infiltration Measurements in Homes and Buildings Using Passive Perfluorocarbon Tracer Method," *ASHRAE Transactions*, Vol. 91, Part 2, 1985.

DOE-2 User News, PUB-439, Vol. 7, Published by Simulation Research Group, Lawrence Berkely Laboratory, Berkeley, CA, Winter 1986.

Draper, N. R., and Smith, H., *Applied Regression Analysis*, 2nd Ed., Wiley, New York, 1981.

Fels, M. F., Ed., *Measuring Energy Savings: The Scorekeeping Approach," Energy and Buildings*, Vol. 9, 1986.

Forrester, J. R., and Wepfer, W. J., "Formulation of a Load Prediction Algorithm for a Large Commercial Building," *ASHRAE Transactions*, Vol. 90, pt.2B, 1984.

Gelb, A., Ed., *Applied Optimal Estimation*, MIT Press, Cambridge, MA 1974.

Hammarsten, S., "A Critical Appraisal of Energy Signature Models," Swedish Institute for Building Research, Box 785, S-801 29 Gavle Sweden, 1986.

Hogg, R. V., 1979, "An Introduction to Robust Estimation," *Robustness in Statistics*, R. L. Launer and G. N. Wilkinson, eds., Academic Press, New York, 1979, Chap. 1.

Hsieh, E., MSE thesis, Dept. of Mechanical Engineering, Princeton University, Princeton, NJ, 1988.

ICBEM, *Control and Regulation*, Proceedings of 3rd International Congress on Building Energy Management Lausanne, 28 Sept.-2Oct. 1987. Vol. 4, A. Faist, E. Fernandes, and R. Sagelsdorf, eds., Presses Polytechniques Romandes, Switzerland, 1987.

IOP, *Inverse Problems*, Institute of Physics Publications, Ltd., Bristol, United Kingdom, 1985.

Jansen, J. E., "Application of Building Thermal-Resistance Measurement," *ASHRAE Transactions*, Vol. 88, Part 2, 1980, p. 122.

Letherman, K. M., *Automatic Controls for Heating and Airconditioning: Principles and Applications*, Pergmanon Press, Oxford, 1981.

Marquardt, D. W., "An Algorithm for Least Squares Estimation of Nonlinear Parameters," *J. Soc. Industr. Appl. Math.*, June 1963.

Marchio, D., "Idenification de Modeles Thermiques par la Methode des Moindres Carres," Report, Centre d'Energetique, Ecole des Mines, Paris, Sept. 1986.

Marchio, E., "Exploitation du Logiciel d'Identification DIATHERM sur les Donness du Batiment Enerplex North," Report, Centre d'Energetique, Ecole des Mines, Paris, Mar. 1987.

NJECL, New Jersey Energy Conservation Laboratory, 2nd Progress Report, Princeton University Center for Energy and Environmental Studies, Princeton, NJ, Oct. 1986.

Norford, L. K., and Persily, A. K., "Simultaneous Measurements of Infiltration and Intake in an Office Building, ASHRAE Transactions, to be published.

Norford, L. K., Rabl, A., and Socolow, R. H., "Measurement of Performance of Solar Heated Office Buildings," Princeton University Center for Energy and Environmental Studies Report PU/CEES 159, Jan. 1984.

Norford, L. K., Rabl, A., Socolow, R. H., and Spadaro, G. V., "Monitoring the Energy Performance of the Enerplex Office Buildings: Results for the first Year of Occupancy," Princeton University Center for Energy and Environmental Studies Report PU/CEEs 203, Dec. 1985.

Norlen, U., von Hattem, D., Hammarsten, S., and Conti, F., "Evaluating the Performance of Passive Solar Components Using Kalman Filtering," Report 1986, Swedish Building Research Institute, Box 785, S-801, 29 Gavle, Sweden.

Persily, A. K., and Grot, R. A., "Ventilation Measurements in Large Office Buildings," *ASHRAE Transactions*, Vol. 91, part 2, 1985.

Protopapas, A., "Solar Gains: A Study of the Enerplex South Building," Senior thesis, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, May 1985.

Pryor, D. V., and Winn, C. B., "A Sequential Filter Used for Parameter Estimation in a Passive Solar System," *Solar Energy*, Vol. 28, 1982, p. 65.

Rabl, A., *Active Solar Collectors and Their Applications*. Oxford University Press, New York, 1985.

Rabl, A., Norford, L. K., and Spadaro, G. V., "Steady State Models for Analysis of Commercial Building Energy Data," ACEEE Summer Study, Santa Cruz, CA, Aug. 1986.

Ryan, L., "A Comparison of Two Parameter Estimation and Control Methods," Informal Report, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, Apr. 1986.

Shurcliff, W. A., *Frequency Method of Analyzing a Building's Dynamic Thermal Performance*, W. A. Schurcliff, 19 Appleton St. Cambridge, MA, 1984.

Sonderegger, R. C., "Dynamic Models of House Heating Based on Equivalent Thermal Parameters," Ph.D. thesis, Report PU/CEES 57, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, 1977.

Sonderegger, R. C., "Diagnostic Tests Determining the Thermal Response of a House," *ASHRAE Transactions*, Part 1, 1978, pp. 691

Stephenson, D. G., and Mitalas, G. P., "Cooling Load Calculations by Thermal Response Factor Method," *ASHRAE Transactions*, Vol. 73, 1967.

Subbarao, K., "BEVA (Building Element Vector Analysis) – A New Hour-by-Hour Building Energy Simulation With System Parameters as Inputs," Solar Energy Research Institute report SERI/TR-254-2195, Mar. 1984.

Subbarao, K., "Thermal Parameters for Single and Multizone Buildings and Their Determination from Performance Data," Solar Energy Research Institute report SERI/TR-253-2617, Jan. 1985.

Subbarao, K., and Anderson, J., "A Graphical Method for Passive Building Energy Analysis," ASME JOURNAL OF SOLAR ENERGY ENGINEERING, Vol. 105, 1983, p. 134.

Wilson, N. W., Wagner, B. S., and Colbourne, W. G., "Equivalent Thermal Parameters for an Occupied Gas-Heated House," *ASHRAE Transactions*, Vol. 91, Part 2, 1985.