Model Identification
with Application to Building Control and Fault Detection

by

Peter Ross Armstrong

B.S. Engineering Design and Economic Evaluation
University of Colorado (1974)

M.S. Mechanical Engineering
Colorado State University (1980)

Submitted to the Department of Architecture
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Architecture: Building Technology

at the
Massachusetts Institute of Technology
September 2004

Signature of Author.....
Department of Architecture
June 1, 2004

Certified by.....
Leslie K. Norford
Professor of Building Technology, Department of Architecture
Thesis Supervisor

Certified by.....
Steven B. Leeb
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by.....
Stanford Anderson
Chairman, Committee on Graduate Students
Head, Department of Architecture

THESIS COMMITTEE

Leslie K. Norford, Professor of Building Technology, Department of Architecture

Steven B. Leeb, Associate Professor of Electrical Engineering and Computer Science

James L. Kirtley, Professor of Electrical Engineering and Computer Science

Model Identification with Application to Building Control and Fault Detection

by

Peter Ross Armstrong

Submitted to the Department of Architecture on June 1, 2004, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Architecture: Building Technology

Abstract:

Motivated by the high speed of real-time data acquisition, computational power, and low cost of generic PCs and embedded-PCs running Linux, this thesis addresses new methods and approaches to fault detection, model identification, and control.

Fault detection. A series of faults was introduced into a 3-Ton roof-top air-conditioning unit (RTU). Supply and condenser fan imbalance were detected by changes in amplitude spectrum of real power resulting from the interaction of impeller rotation and the dominant chassis vibration mode. Ingestion of liquid refrigerant by the compressor was identified by detecting power and reactive power transients during compressor starts. An adaptive ARX(5) model was used to detect ingestion during steady compressor operation. Compressor valve or seal leakage were detected by a change in the leakage parameter of a simple evaporator-compressor-condenser model that explains the rise in compressor load from 0.25 to .5 seconds after compressor start, i.e. as shaft speed rises from about 50% to 90% of synchronous speed. Refrigerant undercharge was also detected by changes in start transient shape. Overcharge was detected by steady state compressor power and reduced evaporator and condenser air flow were detected by steady state power draw of the respective fan motors.

Model Identification. On-line models are useful for control as well as fault detection. Model-based control of building loads requires a valid plant model and identification of such a model for a specific building or zone is a non-trivial inverse problem. The thesis develops three advances in the thermal diffusion inverse problem. Two involve thermodynamic constraints. The problem is first reformulated in such a way that the constraint on temperature coefficients becomes implicit and the problem may still be solved as an unconstrained linear least squares problem. To enforce the constraint on system eigenvalues, the problem is formulated as an unconstrained mixed (linear and non-linear) least-squares problem, which is easier to solve than the corresponding problem with linear objective function and non-linear constraints. The (usually unfounded) assumption on which the normal equations are based—that observations of the independent variables are error free—is relaxed at the cost of one more non-linear term. The resulting model coefficients are valid for predicting heat rate given zone temperature as well as for predicting zone temperature given heat rate.

Control. Three important control applications involving transient zone thermal response are HVAC curtailment, optimal start, and night precooling. A general framework for model-based control of zone and whole-building operation is developed. Optimal precooling under time-of-use rates is formulated to solve the optimal fan operation sequence using a one-day control horizon with hourly time steps. Energy and demand cost savings are presented.

Thesis Supervisors:

Leslie K. Norford, Professor of Building Technology, Department of Architecture

Steven B. Leeb, Associate Professor of Electrical Engineering and Computer Science

Acknowledgements

This research was sponsored by the California Energy Commission under contract number 400-99-011 with additional support from The Grainger Foundation, the U.S. Navy's ONR Control Challenge, National Science Foundation, and Talking Lights, LLC.

I am ever indebted for the unwavering support and confidence of my co-advisors, Les Norford and Steve Leeb. Professor James Kirtley patiently provided useful input as reader in this last year.

Without site support a field monitoring project can't even get off the ground. This was provided for the RTU fault detection testing by Jim Braun, Haoron Li, and Frank Lee of Purdue Ray Herrick Laboratory, and in the East Los Angeles County test buildings, with incredible patience and dedication, by Ron Mohr of Internal Services Division, Energy Group. It is a pleasure to thank again my Russian collaborator, Boris Nekrasov, and his staff Denis Tishchenko and Rashid Muxharlamov.

MIT project colleagues, Kwanduk Lee, Rogelio Palomera, and Rob Cox were always available to talk shop; Chris Laughmen packed in (and unpacked...and repacked) tools, equipment and computers for countably infinite field trips with gay abandon...and introduced me to glass lab.

My fellow Building Technology students have provided exceptional camaraderie; among the many who have acted as sounding boards I especially thank Henry Spindler.

For stress relief, Dwight, John, Bill, Bob, Cory, Sean, Bashar, Mike and Fran of the sailing program.

For encouraging this mid-career venture I thank PNNL colleagues Michael Kintner-Meyer, Mike Brambley, Srinivas Katipamula, Rob Pratt, John Schmelzer and Sriram Somasundaram. Steve Shankle has, in addition, provided the invaluable conduit to computing resources and administrative support and maintained continuity with the mothership throughout my educational leave from PNNL.

For being mentors: Robert Bushnell, Byron Winn, Frank deWinter, John Seem and Bill Currie.

For financial and moral support Joan Sandgren Bridges and Art Smith, Mother and Dad.

For additional moral support, Liz, Martha, Rob and Marcia, and my far-flung nieces and nephews.

For their patience and love, Abigail, Ford and Jack

To all those I have forgotten to thank in these last minutes of writing...and

To all those who didn't interfere.

Contents

Appendices

Chapter 1 Introduction

Building costs are generally broken down into the broad areas of operation and maintenance (O&M), capital cost (amortized principal plus interest), and taxes. Operational costs include utilities and operations labor. Maintenance costs include labor, parts and materials. Operation and maintenance costs are usually lumped together because the building owner has some measure of control such as investing in efficiency or replacing old equipment. For most buildings, O&M represents the largest share of ownership costs.

1.1 Industry Trends/Motivation

Because of the owner's special interest in O&M costs and associated investment opportunities, the industry-wide trends in O&M bear further scrutiny.

First, the costs of gas and electricity, despite the optimistic predictions of deregulation proponents, have not on average and in terms of constant dollars, come down. Meanwhile volatility has definitely gone up. Moreover, there is an increasing recognition that one's energy bill is not the complete picture and that there are significant external costs of waste disposal, resource depletion, land use, pollution and climate change associated with energy use.

Second, in the maintenance arena the trend to outsource maintenance services seems to be following an inexorable asymptotic path to 100% outsourcing. This means that the maintenance provider is motivated to minimize downtime and complaints and keep maintenance costs low but has little or no incentive to maintain or improve energy efficiency. In addition, there is often little or no staff continuity in HVAC maintenance for a given building. A given mechanic's responsibility is a constantly changing pool of buildings. It is ironic that efforts to allocate mechanics optimally among buildings may actually increase costs in the long-term because needed maintenance is forgotten, faults are misdiagnosed, the knowledge of each building's quirks is lost, and unnecessary repairs are made while necessary repairs are overlooked.

Third, there is a trend toward ubiquitous computing and digital communication in both the operations and maintenance arenas that are associated, again, with peculiar ironies. Operators and building mechanics now have computers to "help" with their work but often the computer applications simply translate the former pneumatic, analog, or paper and pencil systems to software without any real improvement in efficiency. Too alarmingly often there are substantial reductions in understanding and sense of ownership. Fortunately, the recognition that "computers should do the things that computers do well" is slowly emerging. In the context of this thesis, these things are: fast communication, continuous monitoring, sophisticated analysis, optimization, acquisition and use of vast data records and, when necessary, convenient access by humans to that data.

1.2 Technical Background

The first two trends are clearly adverse. The third seems to be a mixed blessing which might, if properly applied, give some relief from the first two. Our goal then is to address two O&M problems that really can benefit from sophisticated on-line analysis. These are Fault Detection and Diagnosis (FDD) and control of HVAC peak loads.

There are many O&M tasks that can be made more efficient and less tedious by application of one form or another of computer technology. The transition generally benefits from standardization and thoughtful software engineering but does not require great sophistication (such concepts as artificial intelligence or self-configuration) in most cases. Two exceptions that have held much of the building research community's interest in recent years are optimal control of transient thermal conditions and on-line fault detection and diagnosis (FDD).

Control and FDD have more in common than first (computer-centricity) meets the eye. In each activity there are important roles to be played by plant-specific models.

Model-based FDD and model-based control can take many forms. For example, fault detection can start by comparing the response given by the detailed simulation of a system (driven by the measured excitations that also drive the real system) with the measured response of the real system. Various responses may be closely or remotely related to various faults. Detection may involve little or much uncertainty. The control application raises similar questions of model type. The detailed, physics-based simulation is one extreme. At the other extreme, the model may be completely non-physical, e.g. ANN or statistics-based, or somewhere in between.

To be plant-specific often means to be adaptive and autonomous. And it often means a model that is identified or updated on-line. There may or may not be human initiation or intervention, e.g. as in the building commissioning process.

Thus our two applications of interest, FDD and control of HVAC peak loads, are bound by two common requirements: the need for powerful, standardized, low cost computing resources, the need for large amounts of data, and the need for reliable, plant-specific models.

The computing resources are at hand. The presumed platform is the ubiquitous PC because it has the right hardware (powerful CPU, mass storage, DMA, Ethernet and various other IO) , software (Linux and Octave), and is incredibly cheap ($200 single-board computer).

The need for large amounts of data is partly answered by the computer/network resources. However, obtaining sufficiently detailed data in a useful form can be a daunting task. One recent and highly motivating advance is the non-intrusive load monitor (NILM).

A non-intrusive load monitor (NILM) identifies individual loads by sampling voltage and current at high rates and reducing the resulting start transients or harmonic content to concise "signatures" (Leeb 1992, Shaw *et al* 1998, Laughman *et al* 2003). The NILM's traditional role of end-use metering in buildings is needed by the HVAC peak control application because room temperatures respond to specific electrical loads (i.e. the light and plug loads of the room in question) as well as to weather and heating and cooling stream inputs. But, almost incidentally, the NILM is also extremely well suited to the tasks of measuring and analyzing electrical signals with high information content that may be useful in fault detection. Use of the NILM for fault detection and diagnosis (FDD) is important because 1) it complements other FDD schemes that are based on thermo-fluid sensors and analyses and 2) it is minimally intrusive (one measuring point in the relatively protected confines of the control panel) and therefore inherently reliable.

1.3 Organization

Given the computing and data collection resources of the existing NILM platform, there are three pieces left to the puzzle and they are addressed in Chapters 2, 3 and 4 of this thesis.

Chapter 2 addresses *fault detection.* A series of faults was introduced into a 3-Ton roof-top air-conditioning unit (RTU). The NILM was used to record start transients and the power signals during steady operation. Chapter 2 presents the results analyzing these signals to detect supply and condenser fan imbalance, ingestion of liquid refrigerant by the compressor, compressor valve or seal leakage, refrigerant overcharge, and reduced evaporator and condenser air flow.

The compressor valve or seal leakage fault is detected by a change in the leakage parameter of a simple evaporator-compressor-condenser model that explains the rise in compressor load. The detection of liquid ingestion during steady operation is accomplished by a weighted-mean of k-step ahead innovations generated at each time step. The k-step-ahead forecasts are generated by a new transient event detector (TED) based on a recursively updated, exponentially-weighted ARX model.

Model Identification is the subject of Chapter 3 and parts of Chapter 2. The aggregate load process model, described in Chapter 2, is continuosly updated on line for optimal transient detection sensitivity. A refrigeration loop transient model, also developed in Chapter 2, explains start transient innovations for the vapor bybass fault. The need for such a plant-specific refrigeration loop model in the RTU FDD application is not especially onerous because RTU's are mass produced and the appropriate model would be loaded into flash-RAM at the factory and whenever a field upgrade is made. For control of HVAC peak loads, however, a room or building thermal response model is needed. Because very few of our buildings are mass produced, every model is different, if not in form, at least in the parmeter values.

Identification of such a model for a specific building or zone is a non-trivial inverse problem. Chapter 3 develops three advances in the thermal diffusion inverse problem. Two involve thermodynamic constraints. The inverse problem is formulated in such a way that the constraint on temperature coefficients becomes implicit and the problem may still be solved as an unconstrained linear least squares problem. Solution of this linear problem provides good initial guesses for the subsequent non-linear formulations. To enforce the constraint on system eigenvalues, the problem is formulated as an unconstrained mixed (linear and non-linear) least-squares problem, which is easier to solve than the corresponding linear least squares problem with non-linear constraints. The (usually unfounded) assumption on which the normal equations are based—that observations of the independent variables are error free—is relaxed at the cost of one more non-linear term. The resulting model coefficients are valid for predicting heat rate given zone temperature as well as for predicting zone temperature given heat rate.

Control is addressed in Chapter 4. Three important control applications involving transient zone thermal response are HVAC curtailment, optimal start, and night precooling. A general framework for model-based control of zone and whole-building operation is developed. Curtailment and optimal start problems are cast as one- or two-variable problems with a final state constraint. Optimal precooling under time-of-use rates with allowance for zone temperature swings is formulated to solve for the optimal discrete-time zone temperature sequence. This generally results in 24 variables for a one-day control horizon with hourly time steps. A suboptimal control in which fan power for night free cooling has a single gain parameter for each 24-hour horizon is also developed and evaluated.

A literature review is presented at the beginning, and a summary of results at the end, of each chapter. Overarching conclusions and suggested future research are presented in Chapter 5.

Chapter 2 Fault Detection in HVAC Plants

Hardware and control failures result in aggregate repair costs that are a large fraction of total U.S. HVAC operational costs (Katipamula 2003). Comfort and consequent productivity loss may represent significant additional costs. Unitary equipment appears to be just as susceptible to performance degradation as built-up systems (Breuker 1998a). Many roof-top unit (RTU) faults go undetected by occupants and maintenance staff. A "wait for complaints" strategy is often taken in spite of its many well known negative side effects. There are two generally accepted paths to correcting this situation:

1) test-and-measurement-based preventive maintenance[1] and
2) automated on-line fault detection.

The key to cost-effective on-line fault detection and diagnosis (FDD) is finding the right mix of sensors and automated analysis for a given target system. Much attention has focused on temperature-measurement-based fault detection because of low sensor cost (Rossi 1997). However, temperature sensor location is critical; sensors must be installed in specific locations regardless of the resulting susceptibility to harsh environments or damage during inspection and service activities. Consequently, more expensive but less intrusive sensor types, such as pressure and power, may be justified by improved reliability.

The non-intrusive load monitor (NILM) was developed to identify individual loads by sampling voltage and current at high rates and reducing the resulting start transients or harmonic content to concise "signatures" (Leeb 1992, Shaw *et al* 1998, Laughman *et al* 2003). In this thesis, load identification plays only a small role[2]. However, because the NILM is extremely well suited to the tasks of measuring and analyzing electrical signals with high information content it provides a very useful platform for fault detection. Changes in these signatures can be used to detect, and possibly diagnose, equipment and component faults associated with roof-top cooling units. Use of the NILM's power signature analysis (PSA) capabilities for fault detection and diagnosis (FDD) is important because 1) it complements other FDD schemes that are based on thermo-fluid sensors and analyses and 2) it is minimally intrusive (one measuring point in the relatively protected confines of the control panel) and therefore inherently reliable.

Results of field tests reported in this chapter show that some RTU thermo-fluid faults can be detected by PSA alone. Certain mechanical (compressor and fan) faults and most electrical (supply, controls, motor) faults can also be detected. In this chapter, previous research is reviewed and the ways in which electrical loads respond to faults (artificially introduced in an operating unit) are characterized. New methods of detecting bypass leakage, fan imbalance, and liquid ingestion faults are described and fault detection implementation issues are discussed. The methods developed in this chapter can be extended to other HVAC equipment configurations. In central plant applications a PSA platform will be installed at each motor control center as a standalone device or, possibly in conjunction with the model-based control implementations developed in chapters 3 and 4, as part of the control system.

---

[1] Including such concepts as re-commissioning or continuous commissioning

[2] In Chapter 3, model identification requires disaggregation of total building power into exterior, interior, and HVAC loads and, in some cases, further disaggregation of HVAC loads into fan, punp and compressor loads.

2.1 Previous Research

The frequency of faults in unitary cooling equipment has been assessed from substantial insurance claim and service record data bases by Stouppe and Lau (1989) and Breuker and Braun (1998a). Stouppe and Lau (1989) compiled 15,716 failure records and attribute 76% of faults to electrical components, 19% to mechanical and 5% to refrigerant circuit components. Of the electrical failures 87% were motor winding failures from deterioration of insulation (vibration, overheating) unbalanced operation and short cycling (overheating). The reported numbers refer to events; costs of repair were not published. Rossi and Braun (1993) reviewed various detection and diagnostic methods and later (1997) developed effective rule-based methods.

Rossi (1997) and Breuker and Braun (1997, 1998a, 1998b) applied detection, diagnosis and evaluation methods and performed statistical analyses of detection sensitivity at specific fault levels. Two prototype systems, designated the low-cost system (2 temperature inputs, 5 outputs) and the high-performance system (3 inputs, 7 outputs), were developed and fault detection (without excessive false alarm rates) was generally found to be effective when capacity and COP dropped by 10 to 20% for five common mechanical faults:

Refrigerant leakage, undercharge or overcharge;
Liquid line restriction;
Loss of volumetric efficiency (leaky valves, seals);
Fouled condenser coil;
Dirty supply air filter.

2.2 Fault Detection Tests

Of the five faults examined by Rossi and Breuker, NILM detection was attempted for all but the liquid line restriction fault. Purely electrical faults were not tested. Contact erosion and bounce have already been successfully tested (Shaw 2000). Detection of an unbalanced three-phase voltage condition (failure of any phase to maintain a preset threshold voltage for a preset number of cycles while the other two hold up) is completely straightforward. Three new mechanical faults, rotor imbalance and flooded start and liquid ingestion, were therefore selected for testing. The test unit is pictured in Figure 2.1

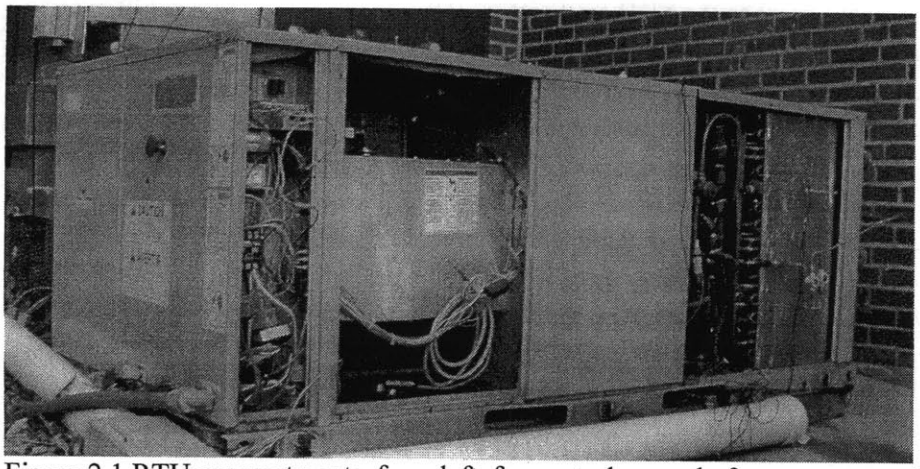

Figure 2.1 RTU compartments, from left, for controls, supply fan, compressor/condenser.

The RTU has two fans and a compressor as shown in Figure 2.2. The fans are connected across lines A and C; the compressor is wired as a 3-phase delta load. Nameplate data and stator resistance measured in the field are summarized in **Figure 2.2**Appendix A. Observation of start transients under fault conditions requires a special procedure for each type of fault. Baseline start transients were collected by starting each motor, letting it run for ~5 seconds, and waiting for it to stop spinning after shutdown. This process was repeated 5 times while the NILM wrote the data to a single file.

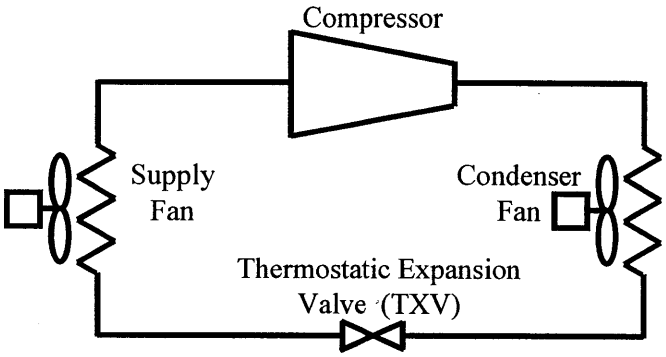

Figure 2.2 Schematic of direct expansion air conditioner.

Evaporator and condenser blockage were artificially introduced by inserting strips of cardboard against the heat exchanger inlet face. This achieves a quasi-uniform flow distribution of known blockage factor. Refrigerant over- and undercharge conditions were achieved by evacuating, then adding a known mass of refrigerant to obtain the desired level of charge. Leaky valves or seals were mimicked by opening the hot bypass valve a specific amount and making repeated compressor starts. Fan imbalance was artificially introduced by adding small weights of known mass near the rotor tip at one point on the circumference. For the supply fan, binder clips were fastened to a vane from inside the squirrel cage. For the condenser, a known mass was affixed at the tip of one blade.
To detect small changes in an electrical load signal it is essential to suppress the 120Hz component and consider only the "dc" components of power (P1) and reactive power (Q1). This is accomplished by a special digital filter implemented in the NILM software (Leeb 1992, Shaw 2004).

In subsequent sections we show plots of the start trajectories of P1 or Q1 for each motor load. Repeat-ability of these transients is important to the FDD application. The trajectories are time-shifted and superimposed to highlight similarities or differences between repeated transient shapes. Additional start transients, recorded after fault introduction, are plotted to show any change. Appendix B provides a summary of all the tests performed on the unit.

2.2.1 Short Cycling

Compressor start transients with no pressure relief between starts are shown in Figure 2.3. Starts with pressure relief between are shown in Figure 2.4. Fans were off in both these tests. The compressor cycling frequency of about three starts per minute used during the test is much higher than the shortest cycling frequency of normal operation and is therefore representative of a short-cycling fault. Note that each unit on the time (horizontal) axis is a 60-Hz half cycle, that is, 1/120 s = 8.333 ms. This time scale, which corresponds to the sampling rate most used in NILM and PSA signal

processing, appears throughout the thesis and consistently denoted by "time index (sampled at 120Hz)."

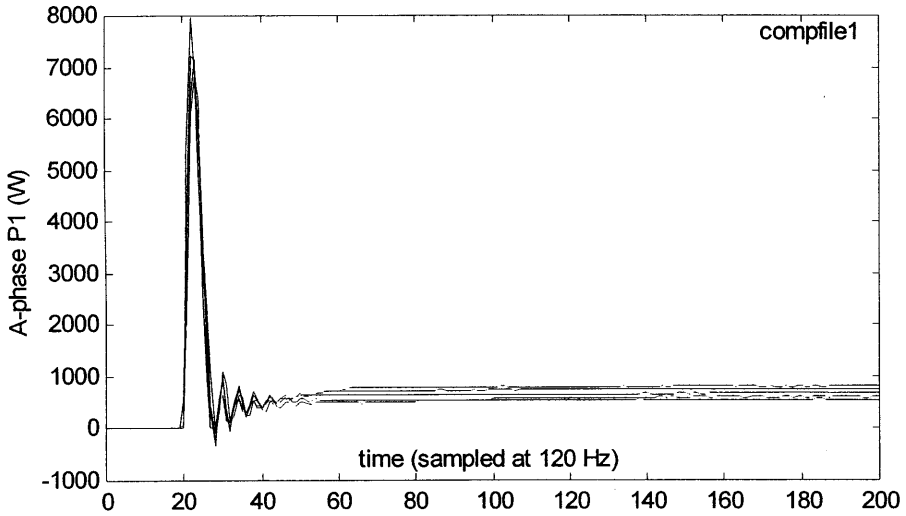

Figure 2.3 Five compressor starts with no pressure relief. The main difference between start traces is that the final value of P1 (time index >50) increases with successive starts.

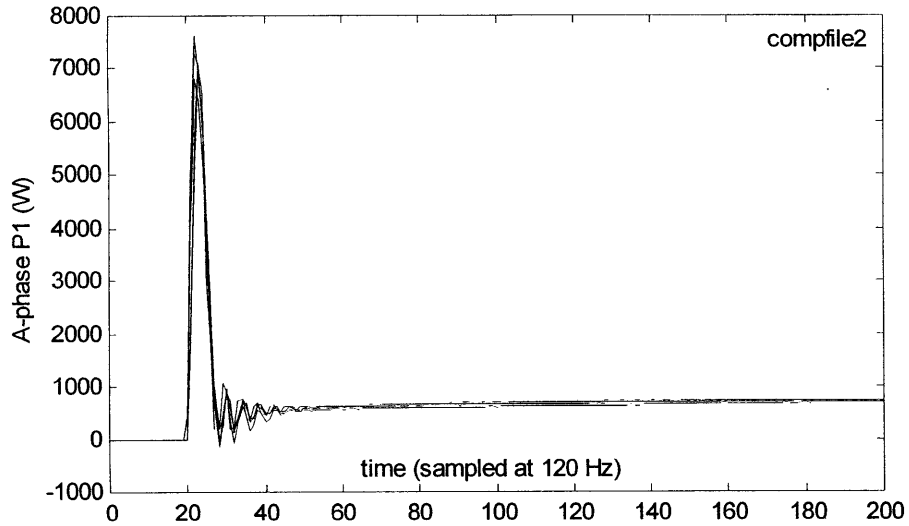

Figure 2.4. Five compressor starts with pressure relief (zero initial head).

2.2.2 Refrigerant Charge

Evaporator and condenser capacity are adversely affected by incorrect refrigerant charge. The system can be overcharged or undercharged. In either case, excessive portions of the evaporator or condenser are relegated to sensible, rather than boiling or condensing, modes of heat transfer. With undercharge there is too much vapor being superheated in the evaporator and desuperheated in the condenser. With overcharge too much liquid collects in the condenser[3]. The TXV can maintain capacity only at the expense of increased volumetric flow, in the case of undercharge, and increased pressure rise in both cases. Thus increased compressor power is a symptom of both under- and over-

[3] There is also increased danger of liquid ingestion.

charge. In some cases suction pressure will be too high for the TXV to maintain capacity and loss of capacity will be an additional symptom, albeit not observable by the PSA.

The increase in compressor power is illustrated in Figure 2.5 for overcharge. With undercharge, however, power does not quickly settle to a steady value. Two start transients representative of undercharge are compared to the normal start transients in Figure 2.6. Figure 2.7 shows the mean P transient for each fault level on one plot and Figure 2.8 shows the phase angle, A, for each of the three conditions.

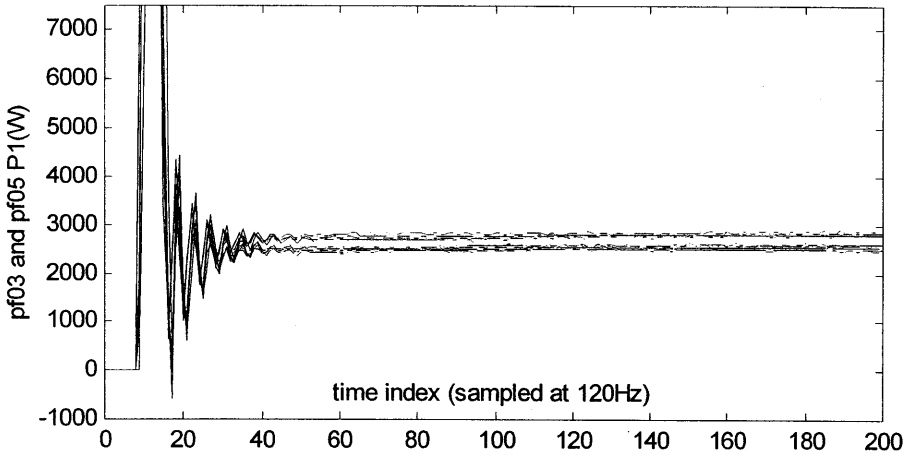

Figure 2.5 Compressor start transients: lower traces normal, upper with 20% overcharge.

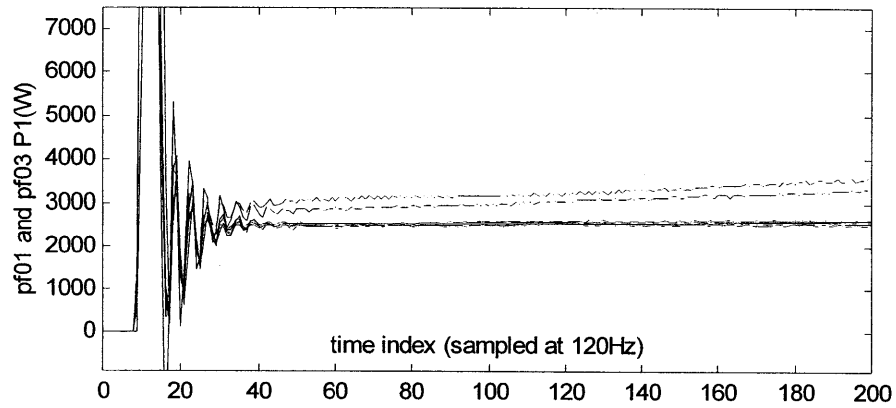

Figure 2.6 Start transients: lower seven traces normal, upper two traces 20% undercharge. For clarity, only two of the four undercharge traces are shown.

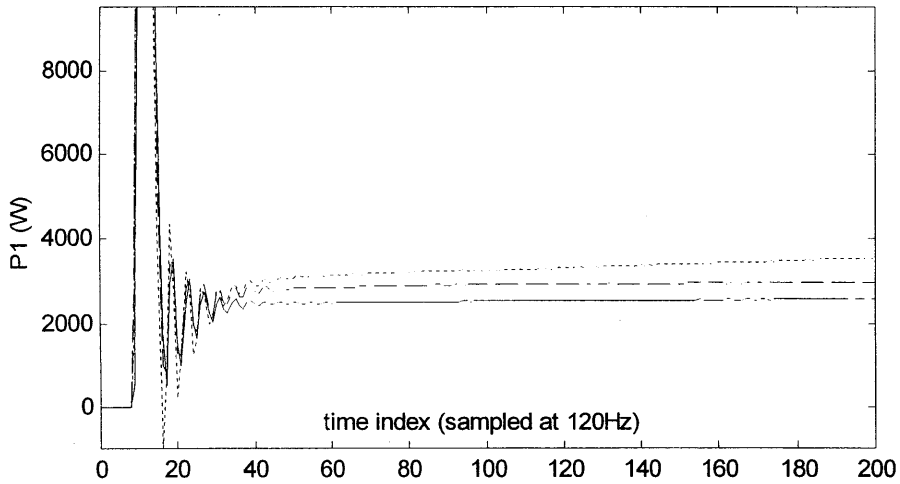

Figure 2.7 Real power start transients, mean of repetitions for each fault level: normal charge (solid), 20% undercharge (dotted), and 20% overcharge (dashed).

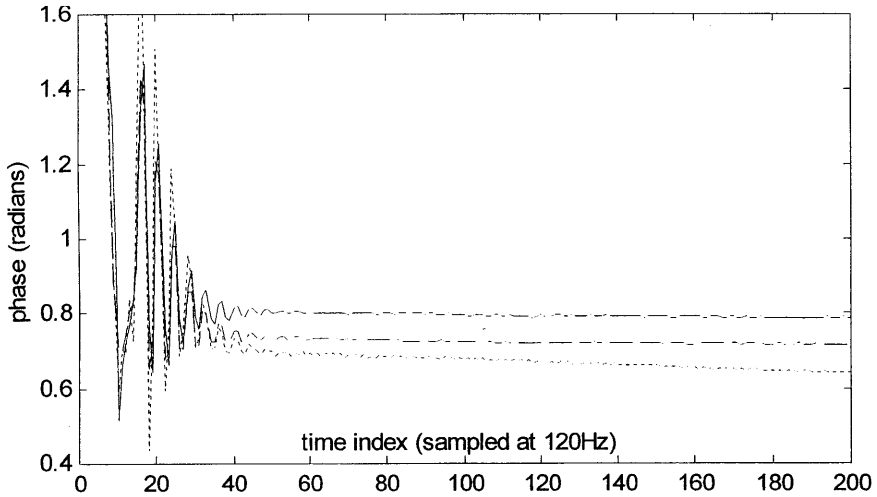

Figure 2.8 Phase angle start transients, mean of repetitions for each fault level: normal charge (solid), 20% undercharge (dotted), and 20% overcharge (dashed).

Detection sensitivity and false alarm rate can be estimated from the small field test samples of mean P1 and Q1. Detection can be based on P, Q, or one of the derived parameters, $R = (P^2 + Q^2)^{½}$, or phase angle, $A = \tan^{-1}(P,Q)$. The sample mean and variance are listed for each parameter (P, Q, R, A) and each condition (80%, 100% and 120% of normal charge) in Table 2.1. Two of the parameters, power (P) and phase (A), are useful for detection. The estimated population distributions for P and A, both assumed to be Gaussian, are plotted in Figure 2.9 and Figure 2.10.

Sample size is an important factor in detection. One advantage of on-line fault detection is that the parameters for the no-fault condition can be established with relatively high confidence because, in a new RTU, there will usually be a large number of observations, n, before a fault occurs and uncertainty is proportional to $n^{-½}$. The vertical dashed line in each plot indicates the sample mean for the no-fault condition. If this mean value were confirmed after many fault-free observations, the probability that the new mean indicated for one of the fault conditions (based on the indicated number of post-fault observations) is truly different from the fault-free mean is given by the area under the

distribution minus the area of the tail on the other side of the vertical line. The NILM can be programmed to implement standard tests (Wild 2000) for change of sample mean given actual sample sizes for the fault-free and post-fault observations that have accumulated at any given time.

We did not assess in these tests the variation of steady-state compressor power that occurs with operating conditions. The observed changes in compressor power alone cannot, therefore, be attributed to a particular fault (overcharge versus undercharge) or distinguished from normal variations in capacity. Discharge-suction pressure difference or condensing and evaporating temperatures could be used to normalize for conditions.

Table 2.1 Detection sensitivity for refrigerant undercharge (80%) and overcharge (120%).

Charge=	80%	100%	120%
N=	4	7	7
Avg(Pss)=	3891	2542	3009
sdev(Pss)=	203	97.5	219
Avg(Qss)=	2675	2585	2583
sdev(Qss)=	24	16.3	22.8
Avg(Rss)=	4723	3626	3967
sdev(Rss)=	177	74.1	177
Avg(Phss)=	0.6033	0.7941	0.7107
sdev(Phs)=	0.0219	0.0181	0.0327

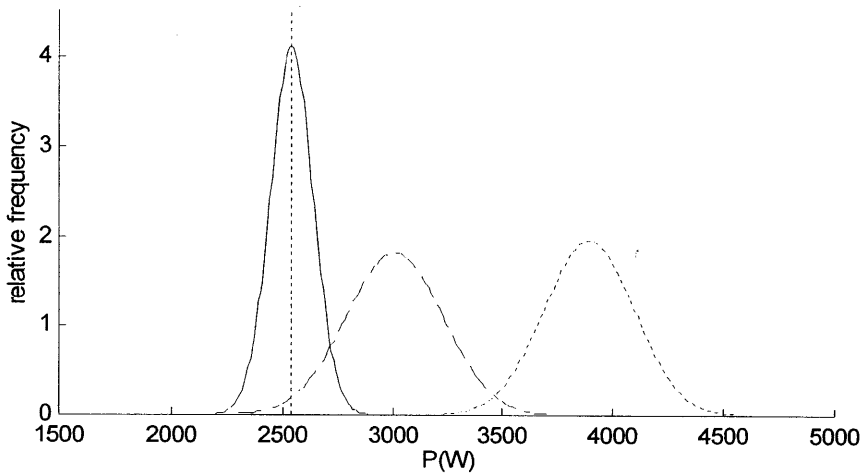

Figure 2.9 Estimated distributions of steady-state power. Because only a small portion of the tail of the 20% overcharge envelope (centered at 3009W) crosses the no-blockage mean (2542W), detection at 20% overcharge will be almost immediate (~10 starts) once the normal-charge mean has been well established.

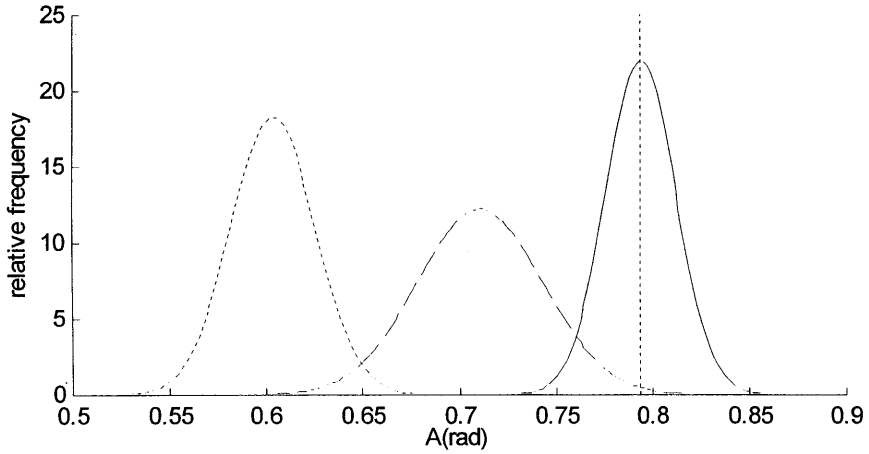

Figure 2.10 Estimated distributions of steady-state phase. Because only a small portion of the tail of the 20% overcharge envelope (centered at 0.711 radians) crosses the no-blockage mean (0.794 radians), detection at 20% overcharge will be almost immediate (~10 starts) once the normal-charge mean has been well established.

2.2.3 Vapor Bypass (Back-Leakage) Fault

Start transients were recorded (4 repetitions) for the compressor/condenser fan with the vapor bypass valve closed and two more start transients were recorded with the bypass valve ¼ turn open as shown in Figure 2.11. The transients with no bypass are the top two. The amplitude difference when the motor is developing peak torque (0.2-0.5s from initial contact) is much greater than the steady-state amplitude difference and, moreover, the shapes of the faulty and fault-free start transients are distinctly different. In contrast, the reactive power shows no such distinct change in response to the fault. The bypass valve characteristic is documented in Appendix C.

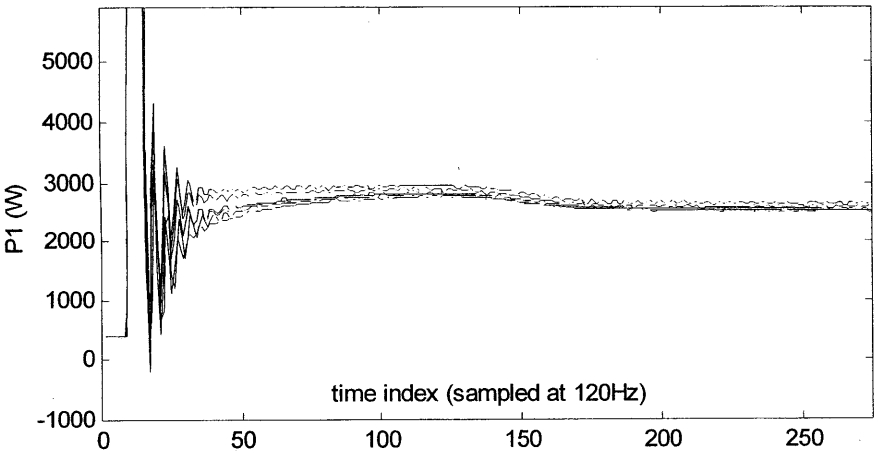

Figure 2.11 Compressor/condenser active power with bypass leakage (bottom four traces) and without (top two traces).

The mechanism for a behavior that is distinctive during the start transient, yet results in little (or no) change in steady state compressor power, is not obvious. However, it turns out that a simple model, in which evaporator and condenser vapor masses are the only state variables, explains the observed behavior.

We assume that evaporator and condenser vapor volumes are constant, that leak rate follows a power law[4], and that the compressor flow rate is smooth, not pulsing, with zero clearance volume. Under these assumptions, the rate of change in evaporator vapor mass is the rate of evaporation plus the back leakage minus the compressor suction rate, hence the following mass balance equation:

$$\frac{dM_e}{dt} = \frac{\Delta T_e u_e}{i_{fg}(T_{sat}(P_e))} + \sqrt{\rho_e \rho_c} C (P_c - P_e)^n - \rho_e V \frac{\omega}{2\pi}$$

where

V is the effective compressor displacement, m^3,
ω is the compressor shaft speed, rad/s,
$\rho_e = M_e/V_e$ is the evaporator (suction) vapor density, kg/m^3,
$\rho_c = M_c/V_c$ is the condenser (discharge) vapor density, kg/m^3,
$P_e = P_{sat}(\rho_e)$, $T_e = T_{sat}(P_e)$ are the evaporator pressure and temperature, Pa, °K,
$P_c = P_{sat}(\rho_c)$, $T_c = T_{sat}(P_c)$ are the condenser pressure and temperature, Pa, °K,
$\Delta T_e = T_{sa} - T_e$ is the supply air – refrigerant temperature difference, °K,
u_e is the effective evaporator conductance, W/m^2-°K,
C,n are the leak parameters with leakage flow modeled as a power law, $m^3 s^{-1} Pa^{-n}$,
$i_{fg}(T)$ = heat of vaporization as a function of temperature, kJ/kg, and.
$P_{sat}(\rho)$, $T_{sat}(P)$ are the other saturation properties, Pa, °K.

The corresponding condenser mass balance is

$$\frac{dM_c}{dt} = \rho_e V \frac{\omega}{2\pi} - \frac{\Delta T_c u_c}{i_{fg}(T_{sat}(P_c))} - \sqrt{\rho_e \rho_c} C (P_c - P_e)^n$$

where

$\Delta T_c = T_{amb} - T_c$ is the ambient air – refrigerant temperature difference, and
u_c is the effective condenser conductance.

Compressor power (averaged over one shaft turn) is given by the energy balance[5]

$$\dot{W} = \left(\frac{dM_c}{dt} + \sqrt{\rho_e \rho_c} C (P_c - P_e)^n \right) \frac{P_c}{\rho_c} + \frac{\Delta T_c u_c}{i_{fg}(T_{sat}(P_c))} - \frac{dM_e}{dt} \frac{P_e}{\rho_e} - \frac{\Delta T_e u_e}{i_{fg}(T_{sat}(P_e))}$$

Simulated power transients with two different leak levels are shown in Figure 2.12[6]. The MatLab code for the simulation is listed inFigure 1. **Flow-pressure data for bypass valve of type used in back-leakage tests.**

---

[4]Several flow modes are possible; use of the simple power law is based on Appendix K empirical results.

[5]Energy dissipated by the leak is assumed to add to discharge superheat, which is recognized, but not explicitly modeled; the ideal gas work is given by: $\dot{W} = \frac{\omega}{2\pi} V P_e \frac{k}{k-1} \left(\left(\frac{P_c}{P_e} \right)^{(k-1)/k} - 1 \right) = \frac{\omega}{2\pi} V \rho_e c_p (T_c - T_e)$

[6] The simulated power in both cases is much lower than the measured power because superheating, flow losses, and mechanical losses are not modeled.

Appendix D. With a leak fault, condenser pressure builds more slowly until the vapor reaches the saturated condition. Evaporator vapor is already saturated so pressure is initially almost constant, thus initial suction density and flow rate are the same for both cases. With similar mass flow rates, less head translates to lower power initially. At steady state, suction flow rate is higher for the faulty case due to higher evaporator pressure. Despite loss of capacity, with higher flow rate and lower head there is little change in steady state compressor input power.

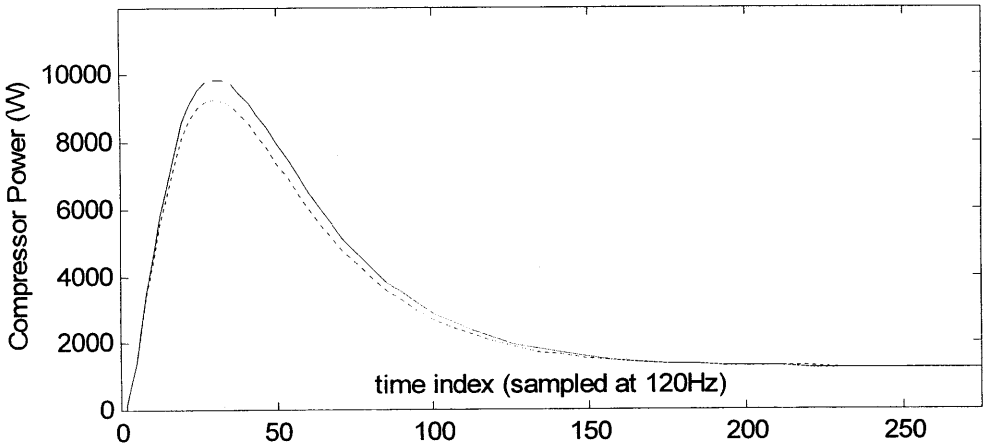

Figure 2.12 Simulated transient compressor power with leak fault (dashed) and no fault.

2.2.4 Flow Blockage

On the air side, reduced flow (blockage or restriction) is one of the main faults of interest. Condenser fan start transients were measured at three levels of blockage: 0%, 14% and 39% of coil face area. The compressor was locked out and six to eight repetitions were made at each fault level. Figure 2.13 shows the P1,Q1 transients for the no-fault case. Figure 2.14 shows the mean P transient for each fault level on one plot and Figure 2.15 shows the phase angle, A, for each fault level. Figure 2.16 gives the mean static pressure transient for each fault level.

Condenser fan power is a good fault detection criterion because it is normally quite constant. Use of temperature-rise-based detection, on the other hand, is problematic for the condenser fan blockage fault. Both inlet and outlet sensors are exposed to the elements. The inlet (ambient) temperature sensor in particular is susceptible to radiation, icing, and wet-bulb bias type errors.

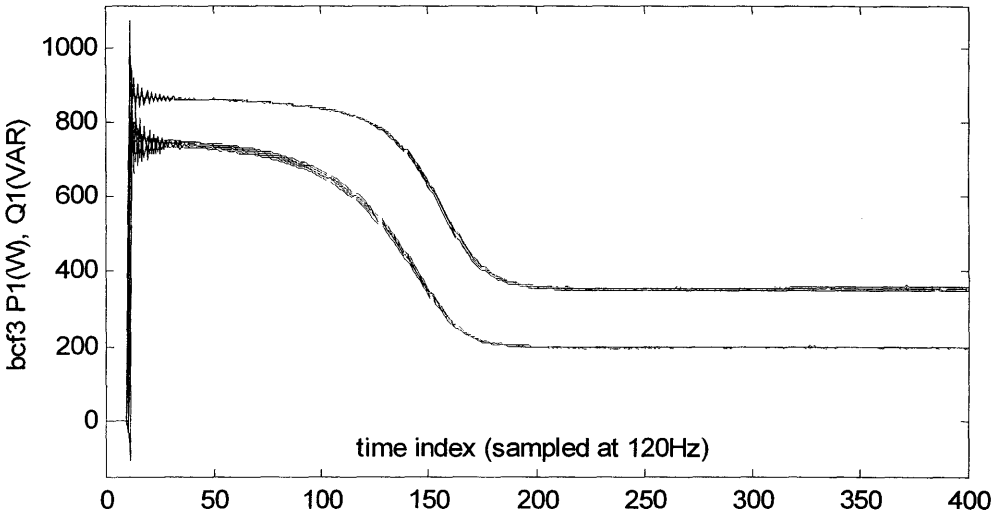

Figure 2.13 Condenser fan start transients with no blockage, P is larger than Q, 6 reps.

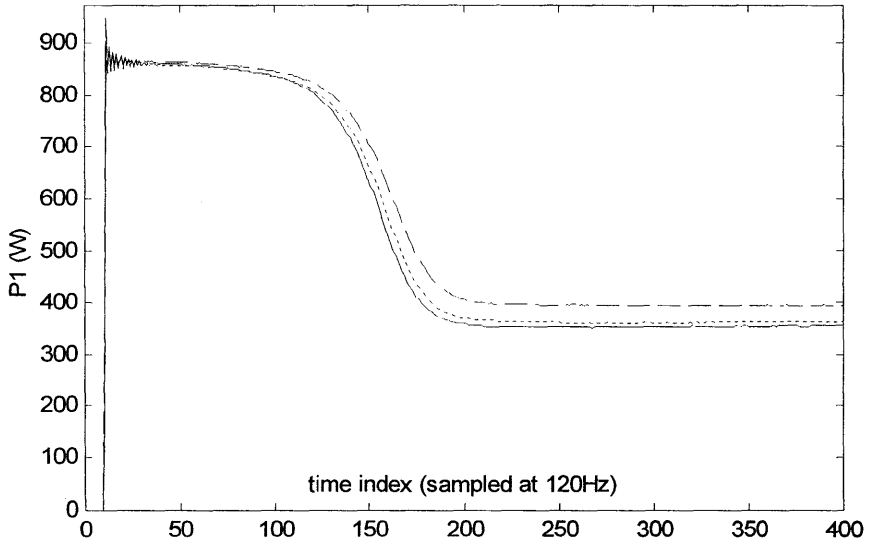

Figure 2.14 Condenser fan real power start transients, mean of repetitions for each fault level: 0% blockage (solid), 14% blockage (dotted), and 39% blockage (dashed).

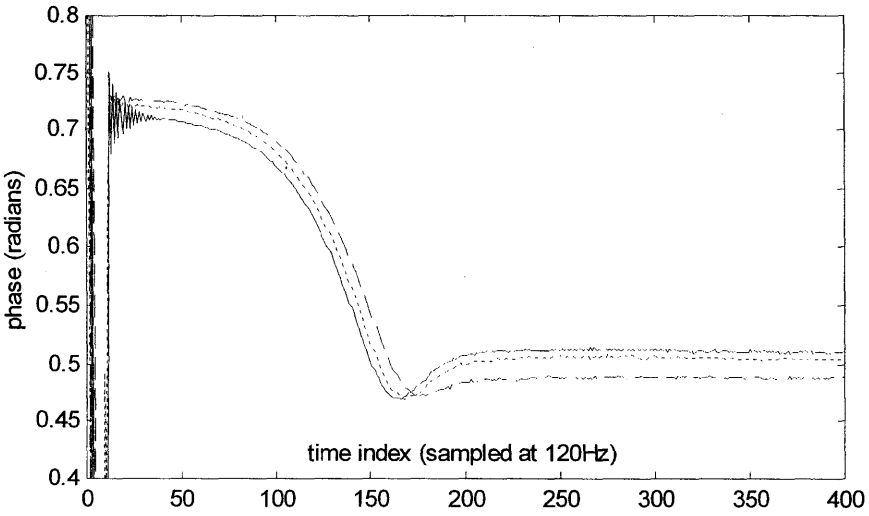

Figure 2.15 Condenser fan phase during start transients, mean of repetitions for each fault level: 0% blockage (solid), 14% blockage (dotted), and 39% blockage (dashed).

The sample mean and variance are listed for each parameter (P, Q, R, A) and each condition (0%, 14%, and 39% blockage) in Table 2.2. Only one of the parameters, P, is useful for detecting blockage because the other variances are large relative to their change in mean with fault level. The estimated population distributions for P, assumed to be Gaussian, are plotted in Figure 2.17.

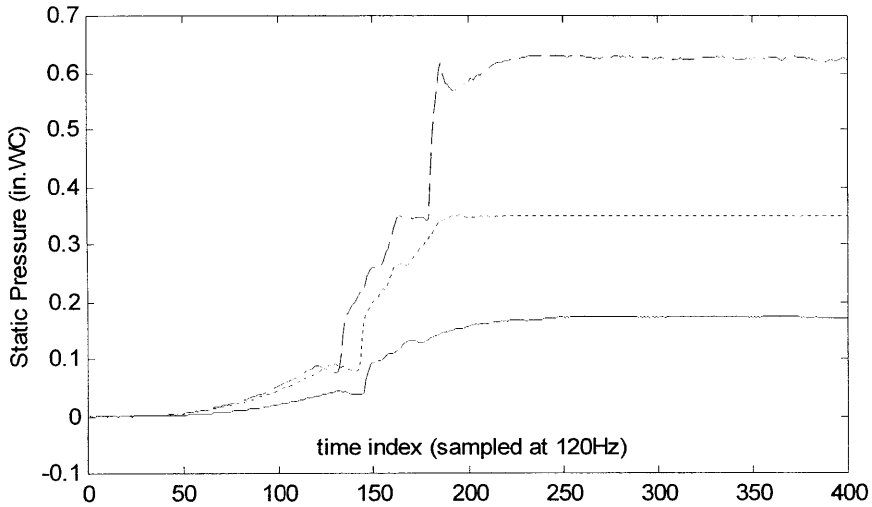

Figure 2.16 Condenser fan pressure start transients, mean of repetitions for each fault level: 0% blockage (solid), 14% blockage (dotted), and 39% blockage (dashed).

Table 2.2 Detection sensitivity for condenser fan blockage.

Blockage	0%	14%	39%
N=	6	8	7
avg(Pss)	353	361.4	393.7
sdev(Pss)	3.17	2.69	2.24
avg(Qss)	198.2	200	209.2
sdev(Qss)	1.39	1.62	1.88
avg(Rss)	404.8	413.1	445.8
sdev(Rss)	3.15	2.64	1.99
avg(Phss)	0.5117	0.5055	0.4885
sdev(Phs)	0.00361	0.00414	0.00481

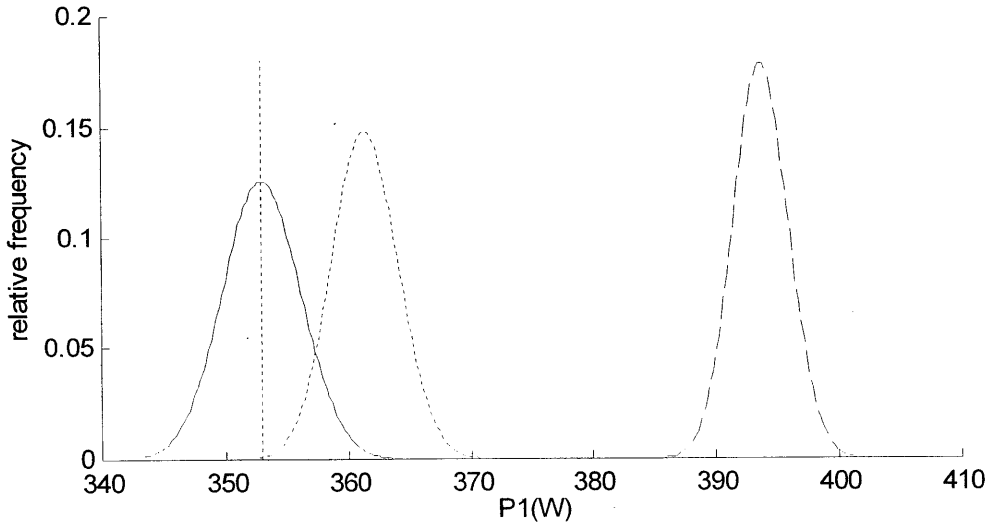

Figure 2.17 Estimated distributions of steady-state power. Because only a small portion of the tail of the 14% blockage envelope (centered at 361W) crosses the no-blockage mean, 353W, detection at 14% blockage will be almost immediate (~10 starts) once the no-blockage mean has been established.

Supply fan start transients were measured with four levels of blockage: 0%, 10%, 50% and 100% of coil face area. Six to eight repetitions were made at each fault level. Figure 2.18 and Figure 2.19 show one the P1 and Q1 transient responses for all three fault levels.

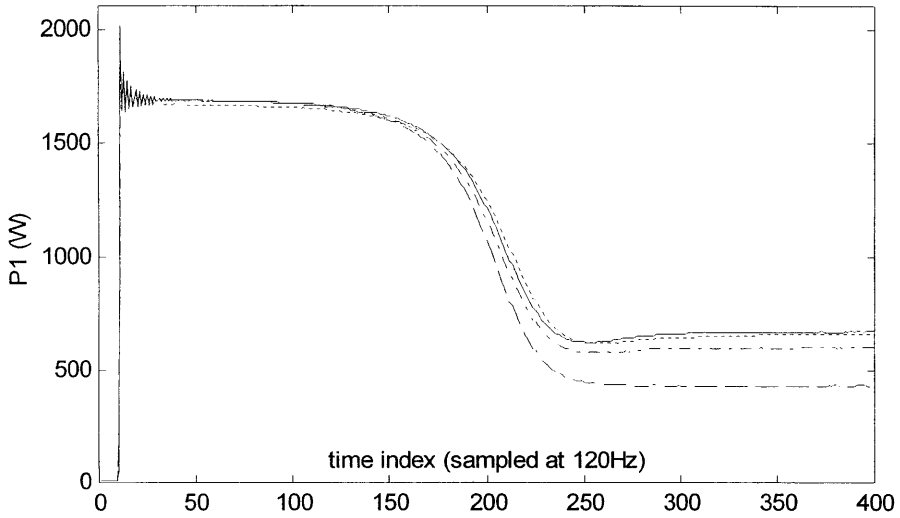

Figure 2.18 Supply fan real power start transients, mean of repetitions for each blockage level: 0% (solid), 10% (···), 50% (·–), and 100% blockage (– –).

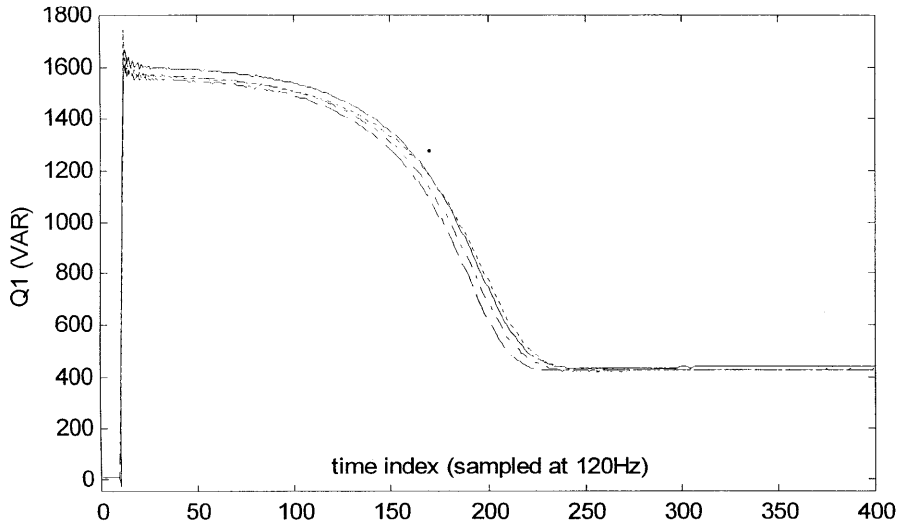

Figure 2.19 Supply fan reactive power start transients, mean of repetitions for each blockage level: 0% (solid), 10% (···), 50% (·–), and 100% blockage (– –).

The steady-state sample mean and variance are listed for each parameter (P, Q, R, A) and each condition (0%, 10%, 50%, and 100% blockage) in Table 2.3. Two of the parameters, R and A, are not useful for detecting improper charge because their variances are large relative to the change in mean of R and A with fault level. The other two parameters, P and Q, are useful. The estimated population distributions for P and Q are plotted in Figure 2.20 and Figure 2.21.

Note that, in contrast to the condenser fan response, supply fan power and reactive power both *decrease* with increasing blockage. This is a consequence of differences in fan curves and the points on those curves that correspond to the respective no-blockage conditions. Moreover, in the case of

the supply fan, detection of a blockage is best accomplished by testing the hypothesis that *both* the reactive and real components of steady-state electrical load have deviated from the established no-fault mean values.

Table 2.3 Detection sensitivity for supply fan blockage.

Blockage	0%	10%	50%	100%
N=	8	6	6	8
avg(Pss)=	662	654	600	444
sdev(Pss)=	6.1	5.6	6.4	7.5
avg(Qss)=	437	428	425	423
sdev(Qss)=	4.4	3.2	2.1	2.1
avg(Rss)=	793	782	735	613
sdev(Rss)=	6.7	5.6	6.2	4.5
avg(Phss)=	0.584	0.581	0.617	0.763
sdev(Phs)=	0.0045	0.0037	0.0035	0.0104

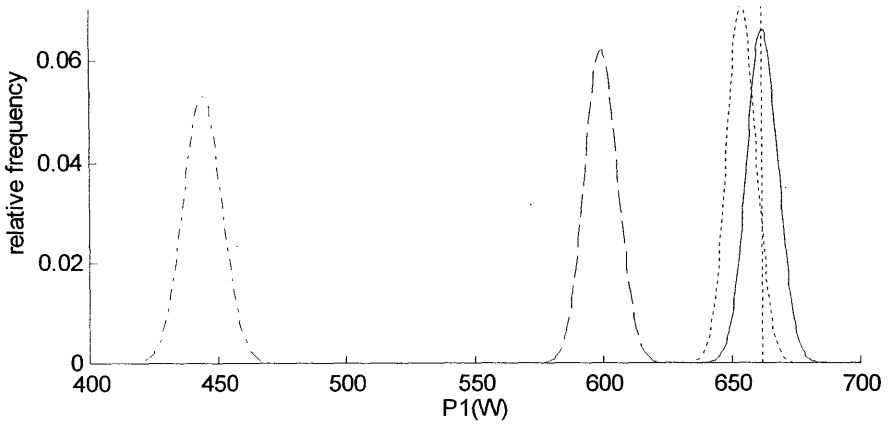

Figure 2.20 Estimated distributions of steady-state power. Because the 10% blockage envelope, centered at 654W, overlaps the no-blockage mean (662W), it would take several hundred start transient observations before a 10% blockage could be reliably detected based on P1 alone.

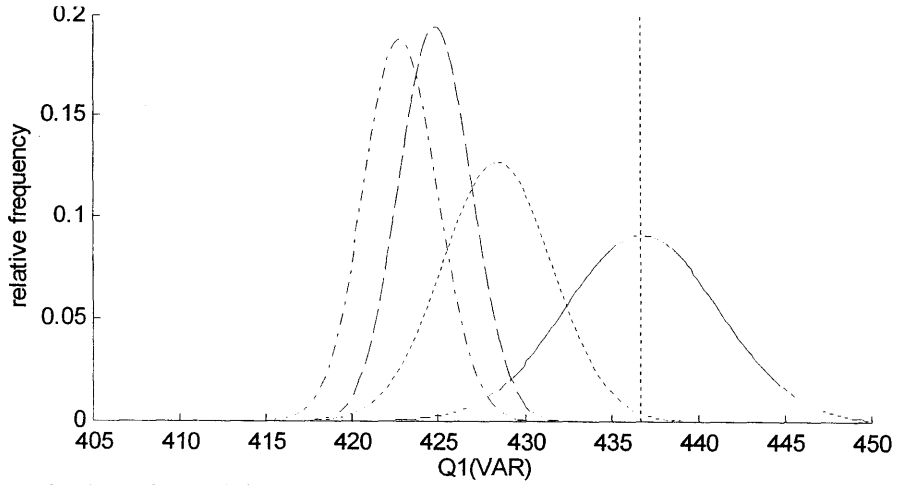

Figure 2.21 Estimated distributions of steady-state reactive power. Because only a bit of the 10% blockage envelope (centered at 428VAR) crosses the no-blockage mean, 437VAR, detection at 10% blockage will be almost immediate (~10 starts) once the no-blockage mean has been established.

From the test results it appears that a 10% blockage can be readily detected. However, variations in system resistance resulting from air density changes (Armstrong 1983) upstream or downstream of the fan and changing damper settings may result in substantial uncertainty in the no-fault means. Further work is needed to understand if the supply fan blockage fault is truly amenable to detection by power signature analysis alone in the face of such ubiquitous system disturbances.

2.2.5 Load Spectra With Artificially Introduced Imbalance Faults

The condenser fan and supply fan are susceptible to faults that result in impeller or rotor imbalance, vibrations, and associated periodic variations in motor load. Such load variations can be detected by power signature analysis. Supply fans are typically of the squirrel-cage type with direct or belt drive. Imbalance can result from a bent impeller air foil, dirt accumulation or from a foreign object, e.g. piece of filter material, becoming lodged inside the impeller cage. Although not explored here, a bent shaft or damaged sheave will result in periodic motor load variations that can be easily detected and it is likely that serious belt wear can also be detected. Condenser fans are typically of the direct drive axial type with a three- or four-bladed impeller. A cracked blade or blade root is the most common serious fault. A slightly bent blade or shaft or the accretion of dirt will also lead to imbalance and, eventually, to mechanical failure. Early detection is important because the cost to repair collateral damage to a motor or condenser is much higher than the cost to replace a fan impeller.

Periodic load variations may be detected by Fourier decomposition of the power signal. To test this fault detection method, we set the NILM to record spectral envelope data for ten minutes of steady operation. The magnitudes and frequencies of peaks in the amplitude spectrum provide a useful description of shaft or impeller imbalance and associated vibrations. For the purpose of demonstrating basic detection feasibility, we have subtracted the mean value of P1 and collected very long time-series sequences so that the Fourier transform gives good estimates of P1 amplitude spectra[7]. Discrete changes in shaft imbalance were induced by attaching small masses to the impellers.

Condenser fan. Shaft harmonics data for the condenser fan are shown in Figure 2.22. The peaks at 59.5Hz (Figure 2.23) and 29.7Hz coincide with the fundamental frequency and half-frequency of shaft speed (nominally 3550 rpm); their magnitudes do not change much with fault level. However, the interaction of unbalanced impeller rotation with one of the lower frequency modes in the RTU structure results in a peak at 18.4 to 18.7Hz that is very sensitive to fault level as shown in Figure 2.24. Note that the frequency of this peak drops moderately as imbalance is increased.

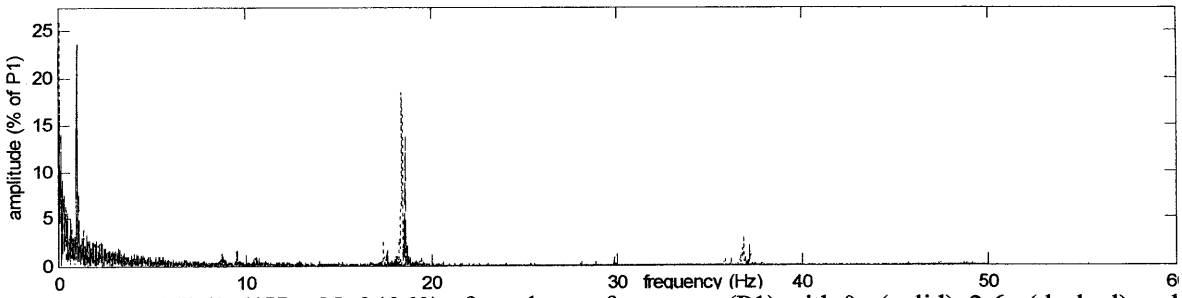

Figure 2.22 FFT (0-60Hz , N=34260) of condenser fan power (P1) with 0g (solid), 2.6g (dashed) and 5.3g (dotted) imbalance fault levels.

[7]The long sequences are more than adequate but a moving average of P1 will useful for short sequences. An appropriate windowing or multi-taper technique may be more suitable on low-cost embedded platforms.

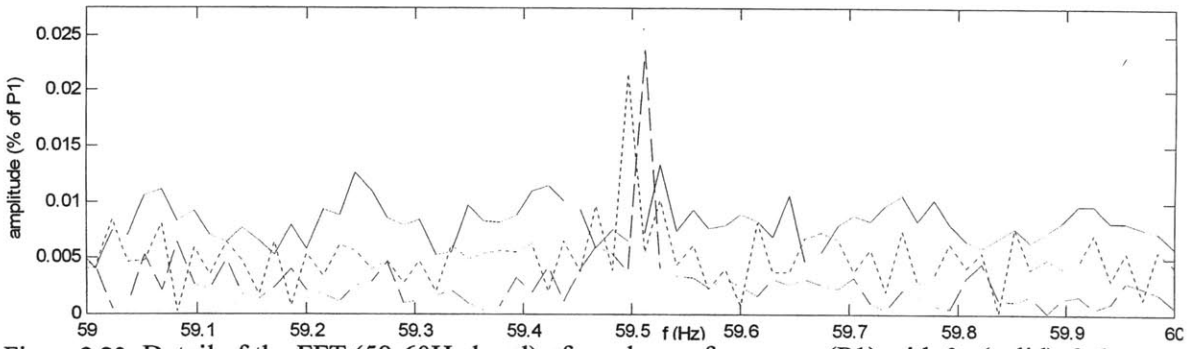

Figure 2.23 Detail of the FFT (59-60Hz band) of condenser fan power (P1) with 0g (solid), 2.6g (dashed) and 5.3g (dotted) imbalance fault levels.

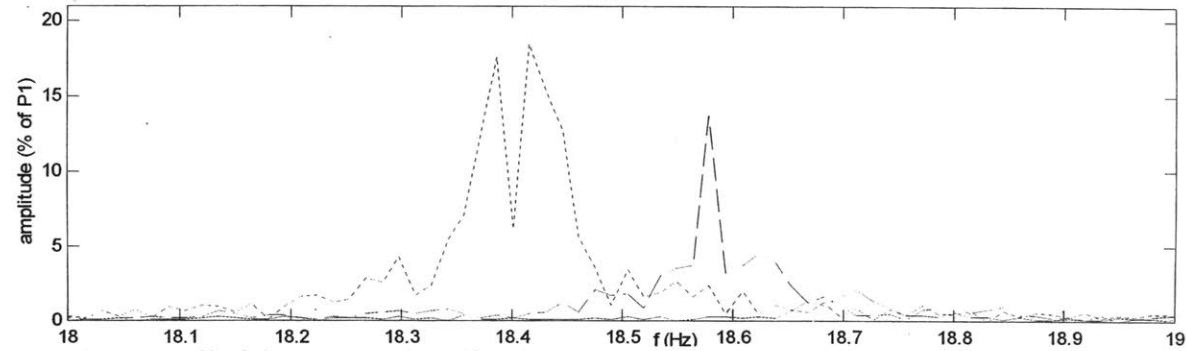

Figure 2.24 Detail of the FFT (18-19Hz band) of condenser fan power (P1) with 0g (solid), 2.6g (dashed) and 5.3g (dotted) imbalance fault levels.

Supply fan. The supply fan motor and impeller are pictured in Figure 2.25.

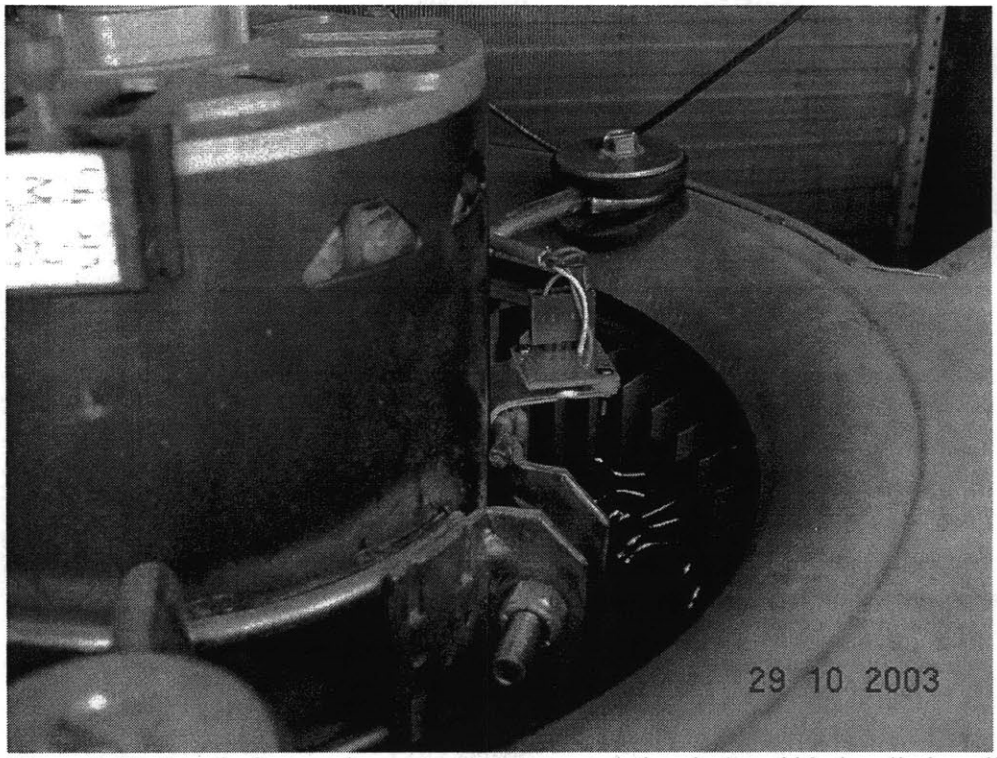

Figure 2.25 Supply fan accelerometer (PCB on angle bracket) and binder clip impeller weights.

Shaft harmonics data for the supply fan are shown in Figure 2.26. The peaks at 59.6 Hz and 29.8 Hz coincide with the fundamental frequency and second harmonic[8] of shaft speed (nominally 1770 rpm). As with the condenser fan, the interaction of unbalanced impeller rotation with one of the lower frequency modes in the RTU structure results in a peak that is very sensitive to fault level. Figure 2.27 shows the peak for fault-free operation at around 19Hz; the frequency of the peak drops to below 18.9Hz, and its magnitude increases many times, as fault level is increased.

The results for both fans show that the existence and relative level of imbalance faults can be inferred from signals measured by the standard NILM platform. Appropriate scaling factors for reporting actual fault levels can be determined for a given RTU design by the simple tests performed here.

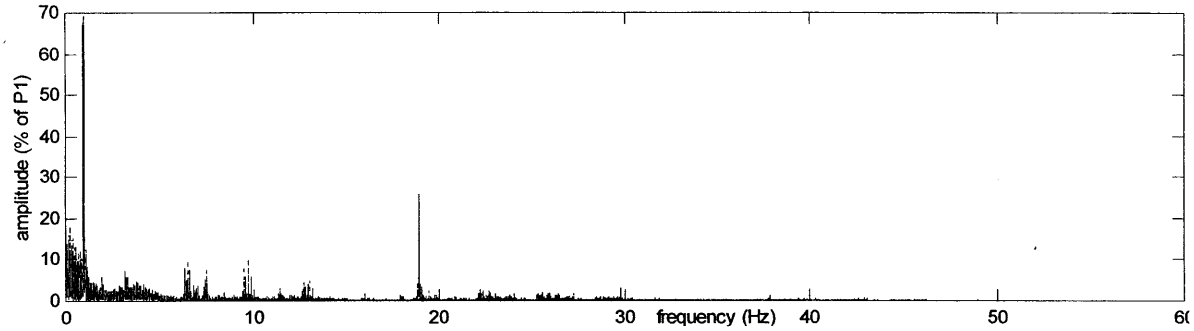

Figure 2.26 FFT (0-60Hz, N=37300) of supply fan power (P1) with 0g (solid), 8g (dashed) and 16g (dotted) imbalance fault levels.

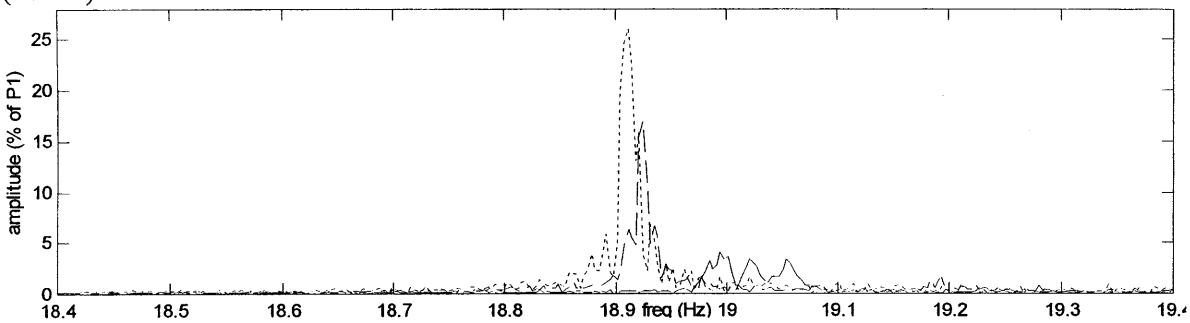

Figure 2.27 Detail of the FFT (18.4-19.4Hz band) of supply fan power (P1) with 0g (solid), 8g (dashed) and 16g (dotted) imbalance fault levels.

2.2.6 Compressor Liquid Ingestion

Compressor damage by liquid ingestion is likely to be preceded by one or more incidents in which small, relatively inconsequential, amounts of liquid[9] enter the machine. The ability to detect flooded starts and liquid slugging by power signature analysis was tested. Initial tests were performed in the field by condensing high-side refrigerant in reservoirs (the two sight glass fixtures pictured in Figure 2.28) and releasing the liquid to the suction line as the compressor was started. Because the liquid is not conveyed directly to the scroll inlet, it falls to the bottom of the can and the only noticeable effect is increased mass flow rate during the first few seconds of operation as the liquid rapidly evaporates in the can. Figure 2.29 shows the corresponding compressor power increases, followed by small abrupt drops in power when evaporation is complete.

[8] "a component...twice the rotational speed is generally linked with rotor anisotropy" Genta (1998) p.410.
[9] An amount of liquid that fills the clearance volume, about 2% of displacement, can cause damage.

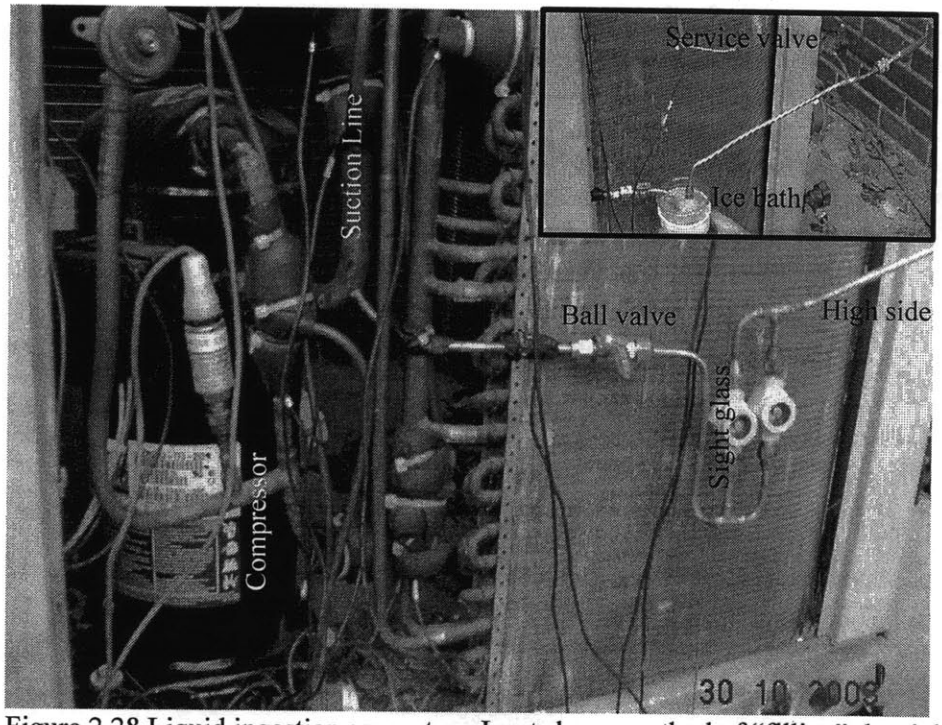

Figure 2.28 Liquid ingestion apparatus. Inset shows method of "filling" the sight glass.

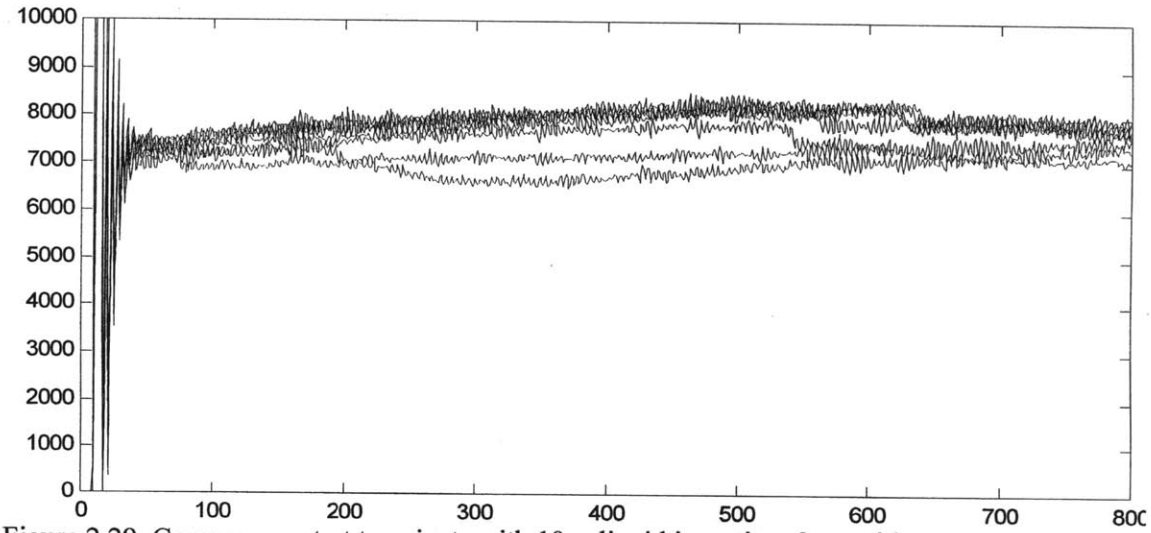

Figure 2.29 Compressor start transients with 10cc liquid ingestion, 8 repetitions.

To introduce liquid in a manner that properly simulates the flooded start condition, it was necessary to perform additional tests in the lab using a small semi-hermetic compressor. The nameplate data for this unit are documented in Table 2.4.

For the flooded start fault, pairs of successive starts were monitored. Liquid was injected in the first start of each pair and not in the second. By alternating between wet and dry starts certain confounding effects, such as change in stator temperature and build-up of liquid residue on the valves, are minimized. A series of five pairs of wet-dry starts was observed with 1.0 cc oil injected above the intake valve ports immediately before each wet start. Dichloromethane (R-30) was injected after about 4s and the compressor was run another 10-20s to clear oil out of the head and valve assemblies.

Table 2.4 Parameters of the compressor used for laboratory liquid ingestion tests.

Make, Model	Copeland KAMA-007A-TAC	
Refrigerants	R-12, R-22, R502	
Power	50/60 Hz, 3 – Phase, 208/230 V	
Resistance:	4.55 ohm line-to-line,+-7%	
Amps @208V:	MCC: 4.5, RLA: 3.2, LRA: 19.9	
Bore, Stroke (in):	1.375	0.938
Cylinders:	2	
CIR (in^3/Rev):	2.786	(21.23 cc/Rev)
CFH (ft^3/hr):	169.3	
Port Diameters (in):	Suction: 5/8	Discharge: 3/8
Condition (F, R-22):	-25, 105	-40, 105
Power (Watts):	696	515
Capacity (Btuh):	3250	1670
Mass Flow (lbm/hr):	43	22

The dry start transients are shown in Figure 2.30 and the wet start transients are shown in Figure 2.31. The mean wet and dry start transients are shown in Figure 2.32 where the difference in wet and dry transients is seen to be small relative to the variation among repetitions in either condition.

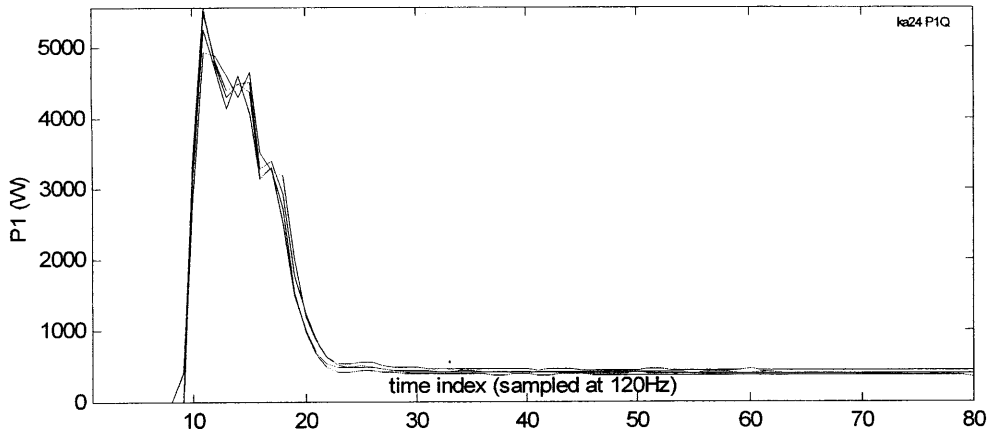

Figure 2.30 Reciprocating compressor start transients with no liquid ingestion, five repetitions.

However, the shape of the difference, shown in Figure 2.33, restores one's confidence that reliable detection will be possible. The three peaks indicate that wet start power increases only during the compression stroke, resulting in a distinct pattern that can be detected, even if the amplitude of the pattern is small relative to process noise.

A second series of alternating wet and dry starts was observed with 3.0 cc R-30 injected above the intake valve ports immediately before each wet start. The mean wet and dry start transients are shown in Figure 2.34. The shape of the difference, Figure 2.35, is exactly that obtained from the wet starts involving injection of 1.0 cc of oil.

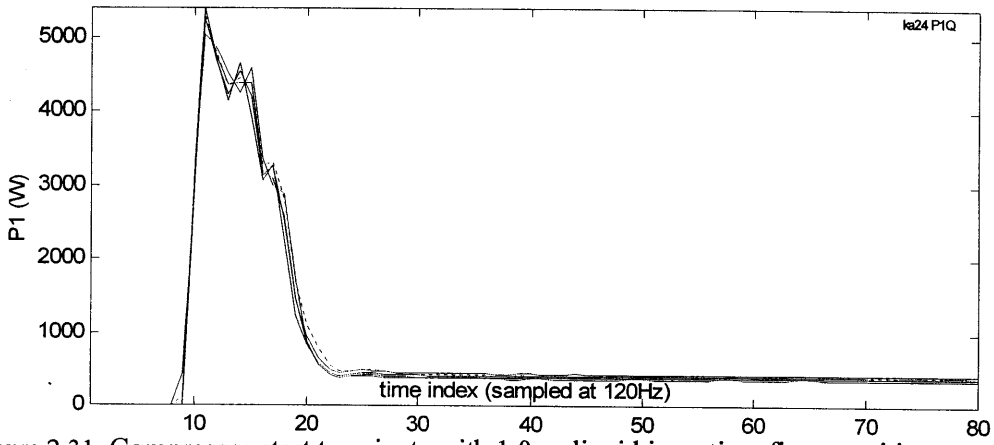

Figure 2.31 Compressor start transients with 1.0 cc liquid ingestion, five repetitions.

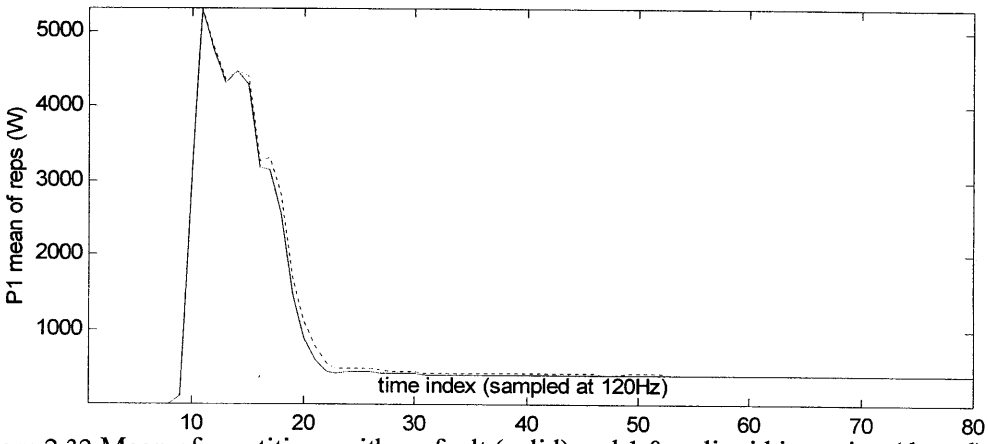

Figure 2.32 Mean of repetitions with no fault (solid) and 1.0 cc liquid ingestion (dotted).

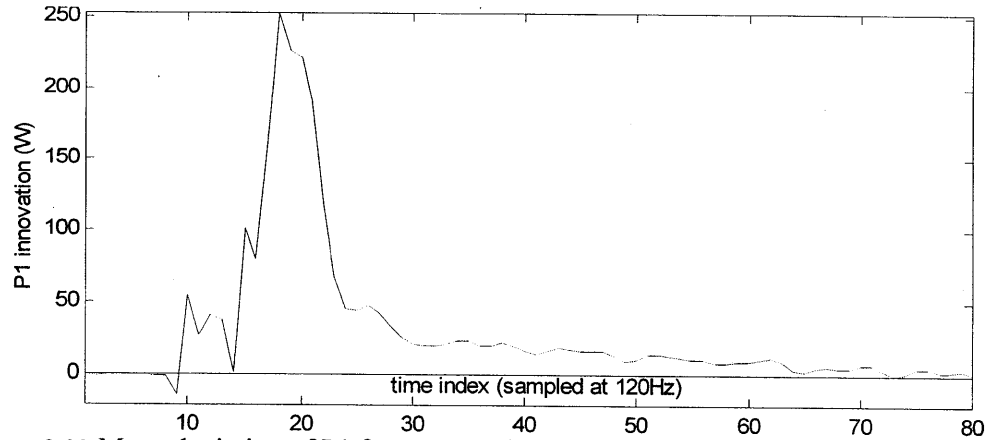

Figure 2.33 Mean deviation of P1 for starts with 1.0 cc liquid from P1 for fault free starts.

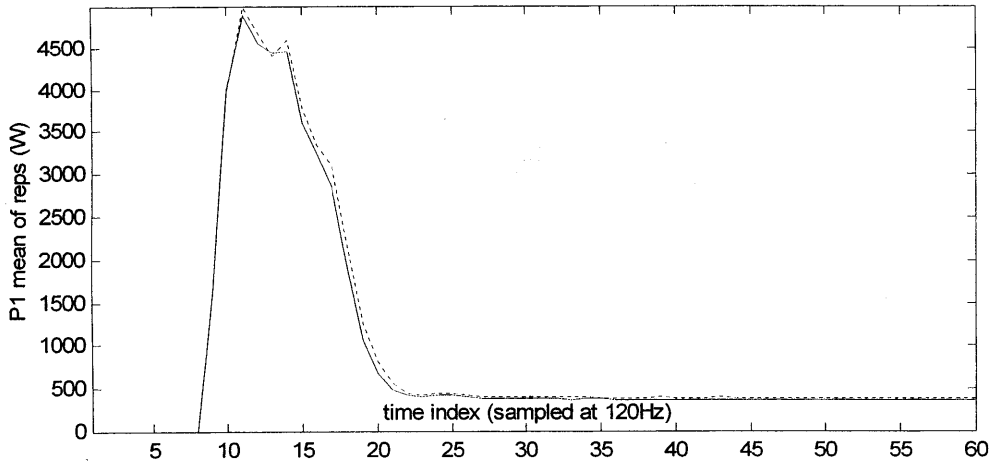

Figure 2.34 Mean of repetitions with no fault (solid) and 3.0 cc liquid ingestion (dotted).

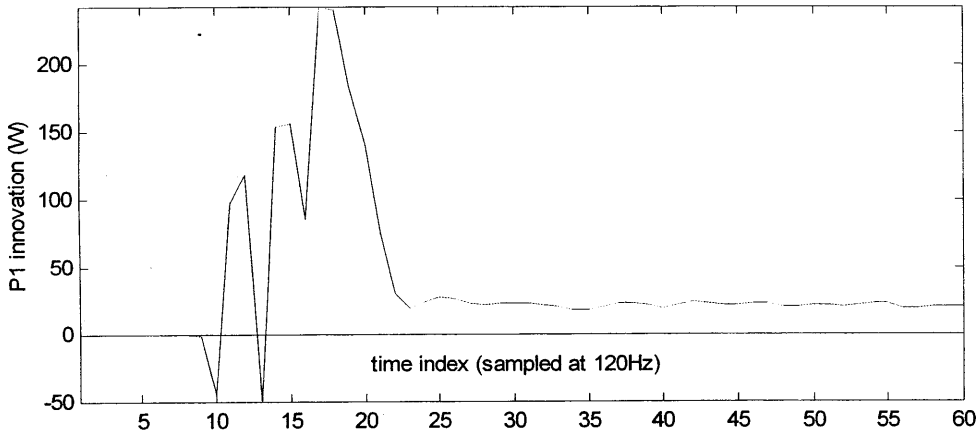

Figure 2.35 Mean deviation of P1 for starts with 3.0 cc liquid from P1 for fault-free starts.

Another liquid ingestion fault, called *liquid slugging,* involves liquid entering the compressor during steady operation. This might be caused by an intermittent TXV fault. To observe this effect, equal doses of oil were injected, at the rate of about one injection every 20s, during steady operation. A series of five events was observed with 1.0 cc oil injected at the suction port for each event. The P1 transients observed in the test are plotted (superimposed) in Figure 2.36. Compressor power increases briefly (by about 15W (2.5%) for 300ms) and then returns gradually to normal. Note that a reciprocating compressor presents a cyclic load component corresponding to shaft speed because more torque is required during suction and compression than when the crank is taking the pistons through their top and bottom positions. The shaft speed cycle interacts with line frequency to produce strong beats, evident in Figure 2.36 at 3.6Hz. At higher time resolution, Figure 2.37, the amplitude of the shaft cycle power (P1) fluctuations is seen to increase substantially when liquid is injected. The corresponding Q1 traces show little or no sign of the disturbance.

As in the flooded start case, the effect on P1 of ingesting non-damaging amounts (<5cc) of liquid is small, making reliable detection difficult. There are three aspects of the difficulty: 1) the small

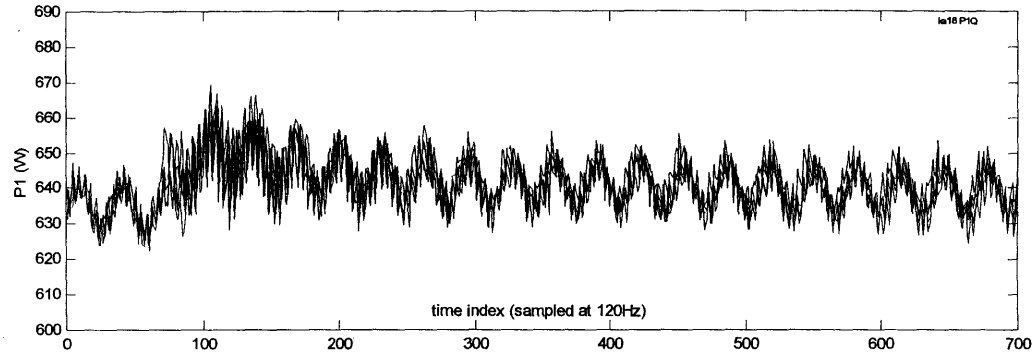

Figure 2.36 Real power during steady run liquid ingestion, five repetitions.

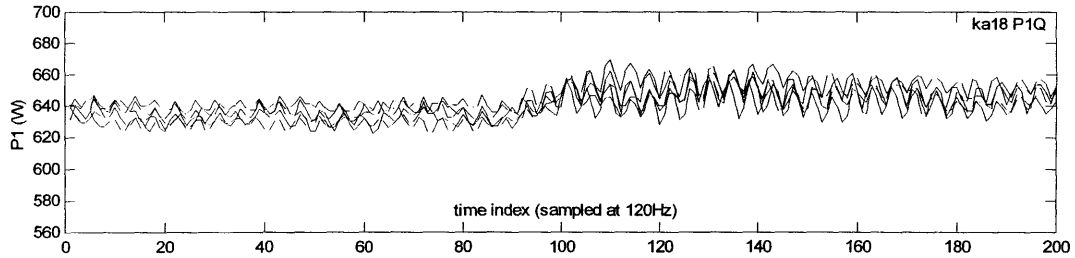

Figure 2.37 Detail showing increase in amplitude of shaft-speed cycle.

magnitude of the change relative to normal variations in P1, 2) the lack of abrupt transient features, and 3) lack of a distinct repeatable transient shape. There is the further difficulty, not apparent in the 1-10s time frames plotted above, that the dc component of P1 varies gradually with refrigeration load, ambient temperature, and stator temperature.

To address these difficulties, we have modeled the time-series of P1 produced in steady operation as a slowly varying autoregressive process (Seem 1991, Ljung 1999). That is, the process that gives rise to y(t)=P1 is modeled as:

$y(t) = \Phi(q)y(t) + \Delta + n(t)$

where

$\Phi(q)$ is an n^{th}-order polynomial in the back-shift operator, q, $\Phi(q) = 1 + \varphi_1 q^{-1} + \varphi_2 q^{-2} + \bullet\bullet\bullet + \varphi_n q^{-n}$

Δ is the detrending parameter, and

n(t) is observation noise

The model coefficients are assumed to vary slowly with time. In vector notation, let

$\mathbf{x}(t) = [y(t-1)\ y(t-2)\ \bullet\bullet\bullet\ y(t-n)\ 1]$ and

$\mathbf{b} = [\varphi_1\ \varphi_2\ \bullet\bullet\bullet\ \varphi_n q\ \Delta]^T$, thus

$y(t) = \Phi(q)y(t) + \Delta + n(t) = \mathbf{x}(t)\mathbf{b} + n(t)$

The model coefficients are estimated recursively by exponentially weighted least squares:

$\mathbf{b} = \mathbf{b}_{t-1} + \mathbf{A}^{-1}\mathbf{x}^T(t)(y(t) - \mathbf{x}(t)\mathbf{b}_{t-1})$, where

$\mathbf{A} = \lambda\mathbf{A}_{t-1} + \mathbf{x}^T(t)\mathbf{x}(t)$

The efficient recursive least-squares algorithm that avoids matrix inversion ($\mathbf{A}^{-1}$) at every time step, based on the work of Kalman and Bierman, is given accessible presentation in, e.g., Ljung (1999) and

Wunsch (1993), and the various methods for obtaining initial values of **b** and **A** (e.g. **A=I**) are also reviewed. An ARX(5,1) model was used with forgetting factor $\lambda = 0.995$ and the exogenous input was set to unity so that the corresponding model coefficient is a slowly varying (i.e., exponentially-weighted moving average) of the "dc" component of P1. The first 14 *k*-step-ahead forecast errors (*P1(k)-P1*) are plotted for the first ingestion event in Figure 2.38.

The adaptive model provides not only k-step-ahead predictions, but also estimates of the variances of prediction errors. Together these numbers may be used to reduce detection to an adaptive scalar threshold test based on the norm of 1/variance-weighted squared innovations:

$$J(t_m, k_H) = \mathbf{n}^T \mathbf{R}^{-1} \mathbf{n},$$

where **n** is the vector of innovations for the time-t_m model over a prediction horizon $(t_m+1: t_m+ k_H)$, **R** is the diagonal matrix of mean-square-error (Figure 2.39) for recent k-step-ahead predictions, $k=1:k_H$, and k_H is the prediction horizon.

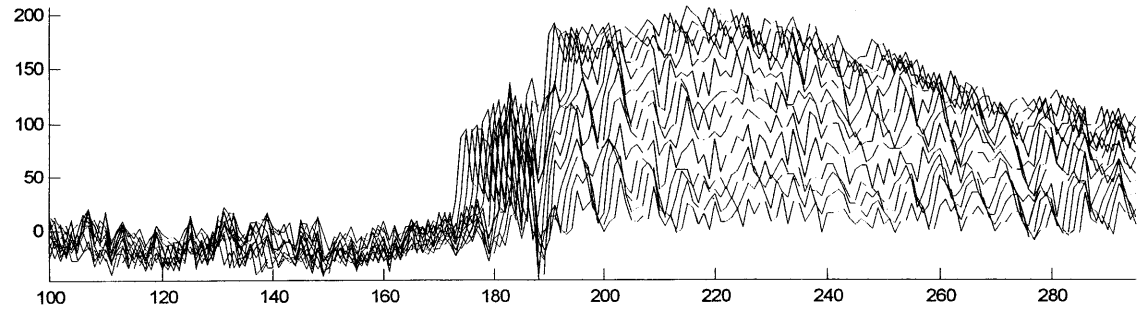

Figure 2.38 k-step-ahead prediction errors from recursive ARX(5,1), λ=0.995, k=1:14.

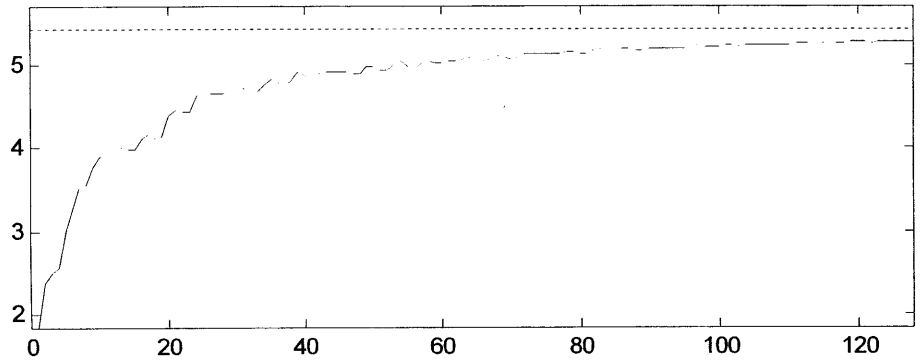

Figure 2.39 Sqrt(variance) for k-step-ahead prediction errors of the ARX(5,1) model during 10 seconds (n=1200) of steady operation. The dashed line is the variance about the mean.

The time evolution of this norm, obtained from the *k*-step-ahead predictions of Figure 2.38 and additional predictions out to k_H=128, are shown in Figure 2.40. The smoothness of *J* is the result of ensemble averaging. Because transients vary in duration, and may or may not be associated with a change of mean, it will be useful[10] to evaluate *J* over several prediction horizons, e.g. k_H=*32, 64, 128, 256* half cycles. In this case, the duration of the transient vastly exceeds the useful prediction horizon and the predictor reduces to the most recent trend line after about 25 time steps. In cases of small,

[10] The duration of the horizon that gives maximal *J* is an indication of transient duration that can be of further use in identifying the load associated with the transient. The recursive ARX-based TED is ideal (and sorely needed) for detection of off-transients because only very short prediction horizons are needed.

short transient events however, this detector is about three times more sensitive than a simple adaptive trend line and it is the optimal detector for transients of any magnitude, duration, and shape, provided the background process is reasonably linear and quasi-stationary. A transient event detector (TED) based on the given vector norm has, moreover, the desirable statistical properties of properly weighted ensemble sampling while retaining the simplicity of a scalar test. The adaptive process model, in addition to detecting a new transient event, has the further important function of isolating the detected transient from the background signal that is the sum of all loads that are active when the new load is switched on. This improves the confidence interval for finding a match with a start transient in the database of known loads.

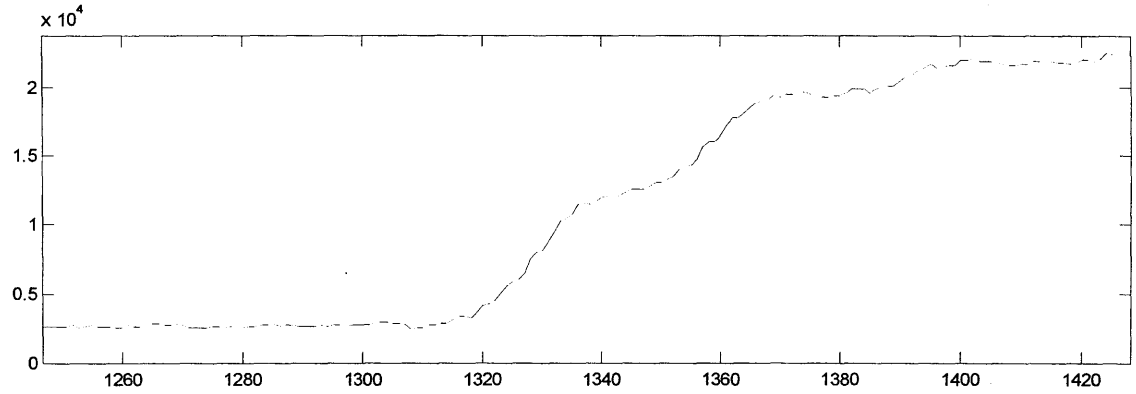

Figure 2.40 Transient event detection by norm of ARX(5,1) k-step-ahead prediction errors, k=1:128.

A diagnosis of liquid slugging could be made with utmost reliability by applying the foregoing filter to suction and discharge pressure signals of similar or moderately lower time resolution. The contemporaneous detection of anomalous power and pressure transients during otherwise steady operation is a sure indication of liquid slugging.

2.3 Implementation and Application

2.3.1 Hardware Implementation.

The NILM or other PSA platform requires only current and voltage sensors at the RTU feed, all located safely in the electrical box. A single-phase platform requires one current sensor and one voltage sensor, while a three-phase platform requires two of each sensor. There may be a niche for a PSA-based FDD device with one current and three voltage sensors so that phase imbalance or phase loss can be immediately detected. The single-phase PSA device can be implemented in a low cost 150 MHz PC-on-a-chip with a two channel A/D converter. The incremental RTU manufacturing cost could be as low as $200. Integration with RTU controls is a natural extension which reduces the incremental cost of FDD capability and may be especially attractive for high SEER designs with variable-speed fans or compressors.

2.3.2 Fault Detection Architecture.

PSA-based fault detection uses five generic detection methods.

A *change in mean* power (P1) or reactive power (Q1) was the only method of load identification implemented in the original NILM (Hart 1992). A change of mean is measured in four steps. First the current mean aggregate load is established recursively by a running, moving, or exponentially weighted moving average. Second, when an appliance switches on or off, a transient event detector

(TED) freezes the current means, logs the time, and tracks aggregate load until the transient condition has settled. Third, after the transient has settled, accumulation of the new mean aggregate load begins. After a new mean aggregate load has been established (typically 10-100s) the change of mean is computed. Once the change in mean has been estimated it is compared to the values established for known loads. Each known load is characterized by its first four odd power harmonics (Pn,Qn) and a confidence ellipsoid in (Pn,Qn) space. If a match is found the nominal as well as measured load parameters (Pn,Qn), transition type (on or off) and timestamp are logged to a tag file. If a match is not found the closest (in a weighted Euclidean norm sense) can be identified and additional tests (e.g. start transient shape) may be applied in succession until the appliance most likely to be the source of the start or stop transient and associated change of mean is identified. If the appliance so identified is one for which simple faults have been established, the type and level of fault are determined based on deviations of the measured change of mean from the expected values.

The *amplitude spectrum* of P1 is computed based on a fixed sample size. The calculation is a fast-Fourier transform on 2^n points of P1 where n is on the order of 12 to 16 and P1 is sampled at 120Hz. When the newly computed spectrum falls outside the confidence region of the baseline spectrum, the PSA enters test mode. New spectra are accumulated until a rigorous statistical test is satisfied or the predetermined maximum test sample size is exceeded. The baseline can include amplitude spectra for two fans operating together as well as spectra for the individual fans. When imbalance is detected with two fans running, the spectrum must be decomposed to determine which fan is at fault. An alternative approach is to track gradual shifting in amplitude and frequency of the peaks identified with imbalance faults by Kalman filter.

Start transients are identified by pattern matching. The process is initiated by a transient event detector (TED) which responds to deviations from *k*-step-ahead predictions generated by adaptive (Pn,Qn) ARX forecasts. The first step is to identify which motor has started or stopped. The next step is to determine if the pattern is within the confidence limits of the established fault-free pattern and if not, what type and level of fault is present. If a motor start transient cannot be identified the event is considered an anomalous transient.

The TED will detect fault-induced transients as well as those associated with individual loads starting or stopping. These transients will, in general, not match any of the established transients. For the RTU, two such *anomalous transients* are liquid slugging and intermittent contacts. The estimated change of mean will be close to zero in both cases. For the intermittent contact fault the time average of (P1,Q1) during the transient will not differ significantly from the mean before or after the transient. Also note that although the contact fault may occur on only one phase of a delta-connected load, the transient will appear on all three phases in such a way that the faulty phase contact can be identified. For liquid slugging the time average of some elements (P1,Q1) will be significantly greater than the mean before or after the transient.

Phase (current or voltage) imbalance results in elevated motor temperature which increases the likelihood of early stator insulation failure. Loss of phase causes very rapid overheating and quickly leads to catastrophic failure of three-phase induction motors. A 3-phase PSA-FDD device or a device that monitors one phase of current but all three phase voltages can respond instantly to loss of phase. This eliminates the need for a dedicated loss-of-phase detector and provides the necessary remote diagnosis to reduce service cost while expediting repair.

2.3.3 Application to larger equipment.

The observations reported here are from a small package rooftop unit. Larger units typically have a belt-driven supply fan, multiple condenser fans, and multiple compressors. A number of changes will be needed for application to larger units. The maximum order of the TED AR(n) process model would be expected to grow (n=3+2m where m is the number of motors). A larger number, m, of start exemplars must be maintained. The possibility of simultaneous motor starts (although normally not allowed by the control sequence) and simultaneous motor stops (*not* disallowed) should be addressed. The amplitude spectra generated during steady run periods must be disaggregated before a growth in power (integrated under selected peaks) can be reliably attributed to a particular motor.

Large package DX split systems are much like large rooftop packages except that supply fans, condenser fans and compressors may be on separate circuits. The PSA-FDD device would be installed at the appropriate HVAC subpanel or motor control center. Refrigeration racks present a similar application scenario.

In small central plants, one typically finds a rack of semi-hermetic reciprocating compressors, chilled water and condenser water pumps, and one or more (often 2-speed) cooling tower fans. Large central plant chillers often have dedicated controls that also provide FDD capability based on compressor current and several temperature measurements that may include oil and stator winding temperatures. However a NILM or PSA-FDD device installed at the HVAC subpanel to provide balance-of-system (pumps, fans) diagnostics would collect additional useful information about root causes of chiller trip events as well as diagnosing the auxiliary equipment itself.

2.4 Summary and Discussion

We have demonstrated that PSA can detect a variety of important faults in a rooftop air conditioner. The methods for detecting these faults are summarized below and a preliminary assessment of the value of fault detection follows.

2.4.1 Detection Method by Fault Type

Air-side blockage faults can be detected by a shift in steady state (P1,Q1) and, to some extent, by changes in start transient shape. However the change magnitudes and directions are related to blockage magnitude via the interaction of fan curve, air-side flow-pressure curve, motor torque curve and no-fault operating point. Thus diagnosis requires more information than that needed for detection, e.g. addition of static pressure. Flow transitions clearly evident in the pressure start transient are also evident in P and Q start transients. Detectable flow transitions may be useful in acquiring very reliable and complete diagnoses.

Short cycling can be detected in the obvious way, by logging time between start transients. The advantage over detection by a temperature-sensor-based FDD system is that very short—even incomplete—start transients are detected. This is important because just a few inrush current events in quick succession will raise stator winding temperature to the point of insulation damage. The second method of short cycling detection relies on the shape of individual start transients being sensitive to residual head. The two methods should be implemented together because they are complementary. A time-stamp and the thermal conditions at the time of each short-cycle occurrence can be saved to aid in diagnosing a root cause.

Refrigerant charge faults result in elevated head pressures which translate to higher steady state compressor loads. Detection of overcharge requires normalizing compressor load over the range of thermal operating conditions. Detection of undercharge is possible based only on the distinct shape of its start transient. The refrigerant vapor bypass fault has a distinct effect on the start transient shape: the change of magnitude in the latter half of the transient is two to five times the change in steady state power.

Liquid ingestion results in transients that PSA can detect in two ways. The change in start transient shape is small in magnitude but has distinct compression stroke features. The shapes of liquid ingestion transients during steady operation are not repeatable in the usual sense but they always result increased power of 0.1-1s duration that appears clearly as positive deviations from k-step-ahead forecasts. Thus any positive power transient that fails to match a known start transient, and that exceeds the adaptive model innovation norm threshold, probably indicates liquid ingestion. Conditions recorded at the time of occurrence can be saved to aid in diagnosing a root cause.

Shaft harmonics observed with no fault indicate that direct drive fan harmonics are not as obvious as belt driven shaft harmonics (Lee 2004) but still potentially useful for load identification. The results of tests using small weights to unbalance the impeller indicate that PSA can identify degradation of impeller balance well before the added repair costs of a catastrophic failure are incurred.

PSA-based fault detection uses five generic detection methods. *Current and voltage* imbalance (or complete loss of phase) are detected by simple ratio and threshold tests. A *change-of-mean* results from a change in steady state mechanical load that is detected as a change in electrical power, P1, with corresponding changes in reactive and apparent power and a corresponding change in power factor. The *amplitude spectrum* of power, P1, is used to detect shaft, coupling, rotor, impeller, and belt-drive faults. *Start transients* are identified by pattern matching. The motor associated with a measured transient is first identified using a generous confidence interval and a tighter confidence interval is then applied to detect a possible fault. *Anomalous transients* are detected by an adaptive linear filter and may arise from faults that do not produce repeatable patterns; the list of such known faults is necessarily incomplete.

Event sequence analysis is a PSA capability *per se* (it can use load state sequences detected in any number of means) but is a fault detection method that is important to implement in the PSA-based device. An event sequence analysis compares the timing and sequence of motor state (on, off) changes to the sequences expected in response to given controller commands. The PSA platform's very high time resolution of motor start and stop detection and ability to log the sequences for playback by a technician are put to good use here.

The suite of faults that are best detected by PSA complements, to large extent, the faults that are best detected by temperature-based FDD systems. Faults that are primarily electrical or electro-mechanical are difficult to detect by thermal measurements but can be detected by PSA using the methods indicated in Table 2.5. PSA can also play a role in detecting many of the more interesting non-electrical faults as summarized in Table 2.6. A comprehensive FDD system could employ an extension of the models developed by Rossi (1997) and Breuker (1998), which currently use return air and ambient air conditions as inputs, by adding compressor power as an input.

Table 2.5 Summary of electrical and electro-mechanical RTU faults detectable by PSA alone.

Fault	NILM Method
Loss of phase	Current and Voltage
Locked rotor	Start transient
Slow starting motor	Start transient
Unbalanced voltage	Voltage
Short cycling	Event sequence
Motor disconnect/failure to start	Event sequence
Incorrect control sequence	Event sequence
Contactor (improper contact closure)	Phase current interruption transient
Fan rotor faults that result in imbalance	Amplitude spectrum in steady operation
Compressor mechanical faults (not tested)	Amplitude spectrum in steady operation

Table 2.6 Summary of non-electrical faults. Some of these can only be detected by the thermal measurements and some can only be detected by PSA, but detection of most can benefit by having both types of measurements.

	Fault	PSA Method	7T[1]	10T[1]
1	Refrigerant leakage, undercharge, or overcharge	Change of mean	√	√
2	Liquid line restriction		√	√
3	Loss of volumetric efficiency (leaky valves, seals)	Start transient	[2]	√
4	Fouled condenser coil	Change of mean	√	√
5	Dirty supply air filter	Change of mean [3]	√	√
6	Liquid ingestion	Anomalous transient		
7	Vapor line restriction	[4]		
8	Non-condensable gas	[4]		
9	COP degradation (root cause: 1, 2, 3, 4, 5, 7 or 8)	[5]	[6]	[6]

[1] 7T FDD measures Twb, Tamb, Tevp, Tsh, Thg, Tcnd, and Tsc; 10T FDD also measures Tra, ΔTcndair, and ΔTevpair (Breuker 1998).
[2] low sensitivity
[3] system flow-pressure-temperature model may be needed for reliable detection
[4] not tested
[5] with addition of refrigerant flow (or comnpressor map) and head pressure
[6] with addition of air-side flow and RH

2.4.2 Value Assessment

We can now make a preliminary assessment of the value of fault detection. The very useful results of Breuker's large sample[11] analysis are summarized in Figure 2.41 and Table 2.7. Value is realized in several areas such as reduced down time, better efficiency, avoidance of consequential damage, and the efficient dispatch of repair staff and replacement parts. The data needed to assess current costs in each of these areas apparently do not exist. However, it is safe to say that these costs are, in aggregate, at least equal to the current aggregate cost of repair. Indeed, it is often the case that only consequential damage is repaired in the first service call and the failure recurs one or more times before the root cause is actually identified and corrected. The information recorded at the first sign of a fault, and additional data recorded about any consequential failures, is therefore of great value in reducing callbacks.

[11] "a statistically representative subset of the data" drawn from "over 6000 separate fault cases"

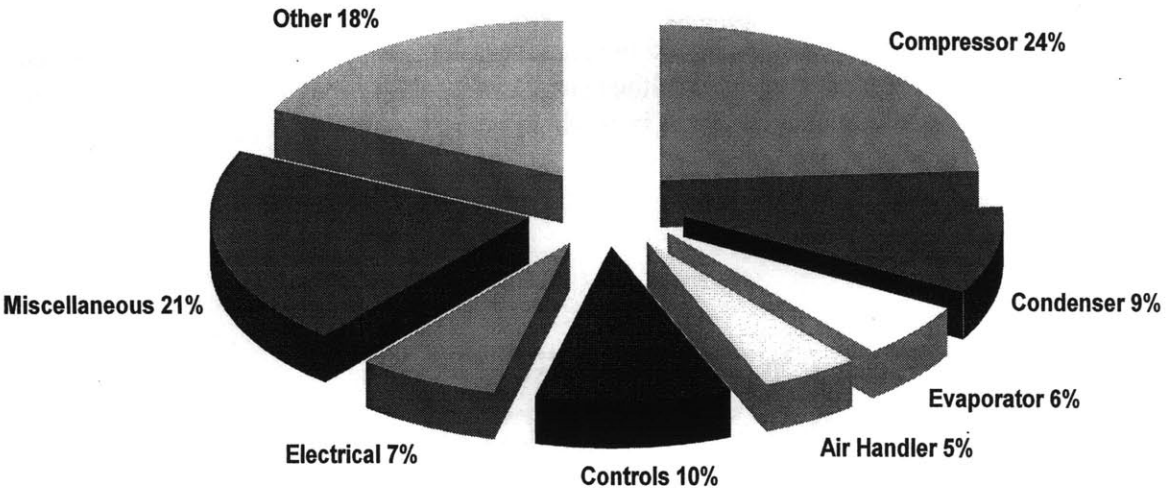

Figure 2.41 Cost of Repair by affected component or subsystem (Breuker 1998a).

Let's consider the fraction of total repair costs associated with faults that PSA can detect assuming, based on the foregoing discussion, that the value of detection is "on the order" of the cost to repair. Breuker's compilation helps us to do just that. The first column of Table 2.7 lists repair cost as percent of total based on Breuker's Table 2; the second and third give cost fraction as a percentage of column 1 based on Breuker's Tables 3-8 (1998a). Assignment of faults to columns 2 and 3 represent the author's assessment, based on the test results reported in Section 2.3, of PSA-based FDD potential for the fault subcategories as tabulated by Breuker.

Table 2.7 Cost breakdown for repair of faults.

		Detection by PSA Certain		Detection by PSA Probable	
Compressor	24	short, open, locked rotor	66	broken internals, vapor bypass	11
Condenser	9	motor, belt failure, cycle control	51	fouling, coil leak	36
Evaporator	6		0	fouling, leak, loss of capacity	75
Air Handler	5	motor, impeller, belt, electrical[12]	67	damper motor, duct leaks	20
Controls	10	compressor interlocks, fan sequence	37		0
Electrical	7	Contactor, loose connection	44	fuse/breaker, short/open	27
Miscellany	21	water loop/pump, fan belt	29	refrigerant leak, installation/startup	42
Other	18	n.a.		n.a.	
Total %	100	Repair of all PSA detectable faults (sum of products with column 1)	37	Faults likely to be detected by PSA (sum of products with column 1)	22

The PSA-detectable faults represent at least 37% of all repair costs and an additional 22% of repair costs are attributed to faults PSA is likely to detect. Together this indicates a possibility of detecting faults associated with 59% of all repair costs or 72% (59/(1-.18)) of attributed repair costs. The fault category that represents the largest repair cost, 16%, is compressor short, open, or locked rotor. (Compressor replacement is typically the single most expensive repair to an RTU.) The database analyzed by Breuker does not provide root cause analysis of these or any other of the reported failures, but likely root causes include

[12] Note that Breuker (1998a) associates *motor* and *belt failure* faults with the air handler rather than with the evaporator, presumably because the air handler and evaporator are treated as separate components whereas the condenser (motor, fan, belt (if present), and heat exchanger) is treated as a single component.

internal damage (e.g. from liquid ingestion), overheating from condenser or evaporator fouling, fan failure, or bypass leakage, and the standard electrical faults such as loss of phase, contactor faults, and short cycling. All of these root causes would be detected by PSA before serious compressor damage ensued for an immediate reduction in total repair costs of at least 10%.

Chapter 3 Identification of Building Thermal Response

This chapter derives a general transient thermal response model for on-line identification that can be applied to walls, zones, and whole building envelopes. Several important aspects of building HVAC operation—FDD, curtailment, verification, model-based control and commissioning—can be improved by use of accurate building-specific plant[13] models. The aggregate potential savings, as will be shown in Chapter 4, is substantial. But such control applications will be economically useful only if the process of identification can be reliably automated. The models and identification methods developed in this chapter are tested using time series data obtained in a test room, at the Internal Services Division (ISD) office building in East Los Angeles County, and at the 60-unit Zubkova apartment building in Ryazan, Russia.

The chapter has seven main sections: *Motivation, Previous Work, Model Synthesis, Test Room Results, ISD Test Results, Zubkova Test Results and Summary.* The materials and geometries are documented for the test room in Appendix E, for the ISD Building in Appendix F, and for the Zubkova Building in Appendix G. describes a general canonical form for the discrete-time transfer function model of an arbitrary thermal system. Appendix I documents the model identification code with constrained time constants. Appendix J documents the model identification code by a new method related to *total least squares*. A representative fragment of the data file input to the identification codes appears in Appendix K.

3.1 Motivation

Identification of building thermal response models is motivated by three broad areas of application:

Peak shifting (includes curtailment and verification of demand savings),
Model-based control, and
Fault detection and diagnosis (FDD).

Additional narrow areas of application include commissioning, planning (e.g. for efficiency retrofits) and verification of energy and demand savings[14]. Load shedding and peak shifting strategies are constrained by building and HVAC plant characteristics as well as climate, utility rate structure and building occupancy requirements. The analytical framework for dealing with different climates, rate structures and occupancies is well established (Armstrong 2002, Braun 1990, Kintner-Meyer 1995, Morris 1994, Norford 1985, Xing 2004). Reliable one-day-ahead weather forecasts can be downloaded daily or hourly. Internal gains, predominantly light and plug loads, can be disaggregated by a NILM from the building total (Leeb 1992, Laughman 2003) and future loads can be similarly extrapolated from the accumulated history (Seem and Braun 1991).

---

[13] Here the "plant" refers to the building envelope, a distributed R-C system with some mildly non-linear R's.
[14] Note that demand and energy savings, while related, are often treated differently because transient response usually has little effect on energy savings but is a dominant factor in demand savings.

However, thermal response to weather, internal loads and HVAC inputs is unique to each building and difficult to characterize empirically (Braun and Chaturvedi 2002, Norford 1985, Pryor 1982, Rabl and Norford 1991). The development of a general model and model identification procedure sufficiently robust to be automated is therefore a central and challenging prerequisite to the implementation of useful curtailment or peak-shifting (CEC, Stoecker 1981), precooling (Keeney 1997), and optimal start (Seem 1989) control strategies. Verification of curtailment for allocating incentives would, ideally, be based on the same model that is used for control.

Model-based control, particularly of zone temperature control loops, can be used to improve comfort and reduce the frequency and amplitude of control variable excursions. The plant model provides a rational way to handle some or all of the easily measured system excitations (sunshine, outside temperature or sol-air temperature, light and plug loads) that are normally treated as "disturbances" in traditional PID control of a plant. Model-based control also addresses problems that often occur when control loops are cascaded. Even if O&M providers resist giving up SISO control loops, a properly identified model can be very useful in the accurate tuning and cascading of such loops. Tuning of control loops is a significant (costly and error prone) element of the HVAC commissioning and re-commissioning processes.

Finally, on-line identification of zone and building thermal response models is essential to detection and diagnosis of HVAC faults that affect comfort and overall HVAC system efficiency (Rossi and Braun 1993). Fault detection can be as simple as checking that zone temperatures respond to HVAC control modes (economizer operation, including control of openings for natural cooling, changes in flow rate or supply temperature, etc.) as predicted by a properly identified model.

We expand on certain aspects of the control application in Chapter 4. However, the key element common to all of the foregoing control, verification, FDD, and commissioning applications is an accurate building-specific plant model. In some cases the required level of accuracy may only be achievable by an adaptive model. Such models, whether adaptive or not, will be economically useful only if the process of identification can be reliably automated. Models and identification methods that have been proposed for on-line identification of building thermal response are reviewed in the *Previous Work* section. The model structure found to be most generally useful and the new identification methods developed for this thesis are presented in the *Technical Approach* section.

3.2 Previous Work

In the area of thermal response modeling we have seen one-node models used by a number of investigators (Dexter 1981, Klein 1994, Richardson 1981, Saunders 1987, Horn 1982, Subbarao 1985) to approximate the dynamics. However, most general purpose building simulation programs use a response factor (DOE 2.1) or transfer function (TRNSYS, BLAST, TARP, EnergyPlus) model of each envelope component. Walton (1983), Klein (1996), and Clark (1985) describe simulation algorithms used in all aspects envelope modeling.

Simple design methods based on the results of building simulations have been developed. Transfer functions did not become part of a recommended design procedure (ASHRAE 1997, Rudoy 1975, Sowell 1985) until recently (ASHRAE 2001). The RTS method is based on conduction transfer functions and can be implemented in a spreadsheet but is too tedious for hand calculation.

The approximations inherent in transfer function models are satisfactory for most building characteristics and for most model fidelity requirements (Mitalas 1968). Indeed, lower-order models are desirable in many situations. Techniques have been developed for deriving reduced-order models given a detailed thermo-physical description of the building (Balcomb 1983, Richardson 1981, Seem 1989). These techniques invoke certain criteria for minimizing model errors such as reproducing exact response to a limited set of driving functions. Balcomb's criterion is response at 0, 0.2 and 1.0 cycles per day. Seem has examined two different criteria: reduced-order models based on the system's dominant roots and reduced-order models based on the Pade´ approximation (Sinha 1983).

Analysis of building test data is another topic germane to the question of reference model-based control. Most of the work in this area has been applied to the "generalization problem." One such question is how to estimate general building parameters, such as UA product or "effective thermal capacitance," from measurements obtained under transient conditions (Armstrong 2000). Another generalization involves estimation of annual energy use in response to a standard weather series such as one of the widely used TMY tapes. Model validation (Judkoff 1999, Spitler 2002) is yet another area for which generalization of short-term test data may be useful. Work in the area of short-term testing and analysis has been reported by Barakat (1987), Bushnell (1978), Kahwaji (1986), McCabe (1983), Duffy (1987), Pryor (1982), Sonderegger (1978), Subbarao (1985) and Wilson (1984). An extensive review of these techniques has been presented by Rabl (1988).

There is some pertinent literature on optimal start/stop control, however Park points out that the details of the algorithms used in commercial products are not generally made public. The schemes that have been reported vary with respect to complexity of the underlying building model that is assumed (Hittle 1981, Seem 1989).

The foregoing is an overview of the literature. Specific topics germane to model identification, model order reduction and other research germane to the applications of interest are reported below.

Recovery Response. E.G. Smith developed a method for determining recovery time in 1940 (Smith, 1940a, 1940b, 1942). He neglects the thermal capacitance of room air and assumes that all envelope walls and contents are initially in a steady-state condition. He further assumes that the heat flux at the surface of each wall and at the surface of each object contained in the conditioned space is constant during the recovery period and that all walls and contents are of homogeneous composition. With these assumptions Smith was able to derive an exact but tedious solution for recovery time. The solution is tedious in that tables of thermal "absorbance," a function of recovery time, film resistances and material properties and thickness, had to be computed by hand in 1940 for each wall construction and for each type of furnishing contained within the conditioned space.

Peak Cooling Load Models. The thermal storage effect has long been considered an essential factor in the calculation of peak cooling loads due in part to the relatively high incremental cost of installed capacity and in part to the highly transient nature of cooling loads. The temperature recovery problem raises similar concerns. It is therefore not surprising that the literature on cooling load analysis is rich in techniques that are potentially applicable to the model identification and peak shifting problems.

The ASHRAE hand-calculation procedures (Rudoy,1975) are based on the results of hour-by-hour simulations using transfer function models to account for thermal storage. It has been found that the storage effect for many different kinds of constructions can be adequately accounted for by categorizing these constructions as light, medium or heavy. (Barakat, 1987, has added an extra-light category for framed residences.) Recent efforts to more accurately account for the storage effects (Sowell, 1985a,b and Chiles 1985) have still used the transfer function as the underlying model.

The ASHRAE thermal response factors method was developed for computing cooling loads and room temperature trajectories under transient conditions (Stephenson 1967). The room temperature was originally assumed to be constant for the cooling load calculations. When the cooling load exceeded plant capacity a separate set of response factors called "response factors for variation of room air temperature" was required. These factors were used to compute the deviation from the cooling setpoint and the method was later extended to include continuously varying room temperatures found in intermittently occupied rooms and rooms in which capacity is proportional to temperature over a throttling range (Mitalas 1969). The companion to Stephenson's paper (Mitalas 1967) derives the homogeneous slab response factors (Appendix A) and the room thermal response factors (Appendix B), but the response factors for variation in room air temperature are never derived. Ceylan and Meyers (1980,1985) developed a general transfer function method that can be used to obtain a relationship between the room temperature and net sensible heating or cooling of the room in a form similar to that envisioned by Stephenson and Mitalas. The method of Pawelski (1979) is similar to that of Ceylan and Meyers in starting with a finite element model; Pawelski's use of statistical techniques results in a parsimonious but not necessarily robust model. More recently Seem (1987,1988) has developed a method specifically for the whole room problem that yields particularly robust and parsimonious whole room transfer functions.

One-node Models. One-node models have been used extensively to account for thermal storage effects in the dynamic response of buildings. Sonderegger (1978) and Duffy and Saunders (1987) for example, have used one-node models to generalize the results of short-term building tests. One-node models have been used in hour-by-hour simulation programs, e.g. TRNSYS (Klein 1983).

Richardson and Berman (1981) used Fourier decomposition to determine optimal one-node model parameters. Several alternate methods of estimating one-node wall model parameters were developed by Horn (1982). Horn also reviews the DHC method (Balcomb,1980) and the thermal admittance method (Davies,1973 and Milbank,1974) and presents a very thorough analysis of the simulation errors that result when the various one-node wall models are used. Horn recommends further investigation of two-mode models.

Model Order Reduction. Pawelski et al (1984) used correlation to find the ARMA coefficients for a complex roof-attic structure. He used a finite difference model of the structure driven by two weeks of hour-by-hour weather data and imposes the constraint sum(**b**)/sum(**d**) = UA, where **b** is the coefficient vector associated with outdoor temperature, **d** is the coefficient vector associated with attic envelope heat load, and UA is from a steady-state solution of the finite difference model. Balcomb (1983) and Seem (1987,1988) have developed the more sophisticated model-order reduction techniques described previously.

Models from short-term testing. Quentzel (1974) and Bushnell (1978) have used linear correlation to estimate parameters affecting daily energy use in buildings. Quentzel correlates daily furnace run time with daily degree days. Bushnell correlates energy use, inferred from daily utility meter readings, with daily mean solar irradiation, outside temperature and wind speed. Neither investigator chose to model thermal storage effects even though the monitored building used night setback schedules. As a result standard errors for the correlations were on the order of the daily heat loss that would result from a 5F temperature difference across the heated envelope.

Sonderegger (1978) used a one-capacitor, two-resistor model to obtain "equivalent thermal parameters" of a light frame townhouse. The parameters were evaluated by linear regression using four days of measured data during which the room temperatures were allowed to float and internal gains were relatively constant. The principal time-varying heat input was due to solar irradiation of the interior envelope surfaces.

Pryor et al (1980) used a two-capacitor R-C network to model a passive solar residence with water-wall storage. The one- and two- capacitance models are simple enough to be expressed as finite difference equations with unknown parameters. Kahwaji develops a method of finding the thermal parameters of homogeneous layered walls from transient measurements based on the exact solution of the conduction equation. To model higher order systems that are more complex than a homogeneous wall, e.g. a room or zone, it is desirable to use transfer functions or variations (e.g. ARMAX models) thereof. These models are especially appropriate for distributed parameter systems because the state of a finite order model used to approximate such a system is not physically meaningful. The input/output relationship is the thing of interest.

McCabe et al (1982) and Seem and Hancock (1985) estimated transfer function parameters by linear least squares to characterize the performance of a solar mass wall component. Forrester and Wepfer (1984) estimated DARMA time series model parameters to identify a load predictor for a large office building. The goal in the forgoing investigations was to find models that were good predictors, not necessarily to find physically meaningful parameters. Subbarao (1985) has estimated transfer function parameters and analyzed the resulting models in an effort to obtain physically meaningful thermal parameters of monitored buildings. Taylor and Pratt (1988) use a finite difference formulation of a fixed-topology, two-node R-C model to estimate equivalent thermal parameters of ELCAP buildings.

Hittle and Bishop (1983) show that a thermal delay line (1-dimensional conduction through a multi-layer slab) has time constants of alternating magnitude and that the largest temperature response time constant is always greater than the largest heat flux response time constant. Harrison, *et al* (1981) develop a superimposable room transfer function to model short-term transient response.

Peavy, *et al* (1973) use data from a simple masonry test building to validate NBSLD and perform similar NBSLD validation work using data from a carefully monitored 4-bedroom townhouse (1975). Bushnell (1978) fits a model with solar, wind, auxiliary heat and indoor and outdoor temperature terms to data from four buildings. An adequate 0^{th} –order model is obtained using one-day timesteps.

Summary. The three most common models for envelope model identification are steady-state models, difference models derived for a specific topology, and transfer function models. Frequency domain methods have been used mainly for network reduction.

Steady state models have been successfully used to estimate wall U-values, and to estimate effective solar aperture, UA, and a linear infiltration coefficient at the zone- or building-level. No information can be obtained about thermal capacitance or transient response.

Topology-specific difference equations can usually be constrained to return thermodynamically feasible models in a straightforward manner, e.g., by constraining the coefficients to be positive. However, the method is limited in practice to 2^{nd}-order models and the 2^{nd}-order topologies that can be handled are too simple for most buildings.

Time-domain transfer function models are the most general linear models that have been used and their coefficients can be easily estimated by ordinary least squares. However the relations between the model coefficients and the buildings physical parameters is complex. If these relationships are ignored (as is usual when taking a purely statistical approach, e.g., applying a multi-input ARMAX model) the resulting model is likely to violate conservation of energy, to exhibit resonant behavior, or to be unstable or non-causal.

In this chapter, constraints that ensure physically realistic behavior are described and identification methods that impose those constraints are developed. These developments allow the practitioner to take advantage of the generality and model-order flexibility of transfer function models without getting into trouble.

3.3 Model Synthesis

Room air temperature, T_{zone}, responds to weather, internal gains and thermal energy delivered by the HVAC system as shown schematically in Figure 3.1. "Weather" includes the effects of sun, wind and outdoor temperature on all exterior surfaces, as well as direct gains through windows. Internal loads, $Q_{L\&P}$, are usually dominated by office equipment and lighting, i.e. "light and plug" loads. Heating or cooling effects, Q_{HVAC}, may be delivered by forced air, natural ventilation, or various types of hydronic terminal equipment.

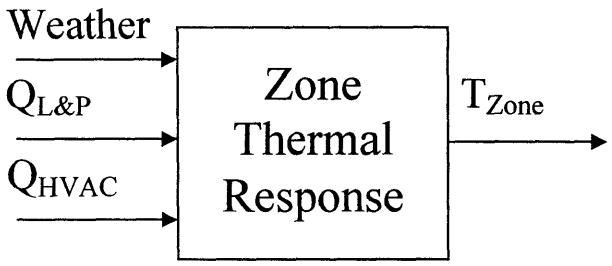

Figure 3.1 Thermal response model.

The response to a given input is determined by the physical parameters and geometry of the building envelope, internal partitions, furniture, and other contents. The thermal capacitance or "mass" as well as the thermal transmission properties of the building fabric determine these response relationships. The engineering data needed to predict how a given building will respond are complex and subject to substantial uncertainty and variation over time. However the response function illustrated in Figure 3.1, if it can be properly identified, is just the sort of crystal ball that is needed for improved control, commissioning and fault detection.

For each application discussed in the *Problem Statement* section, there is an objective function or control law that has a *thermal response model embedded in it.* It is critical that the thermal response model represent not the intended design but, rather, the building as built and currently used. The problem of producing such a model, with minimal fuss, has been addressed in this research project by new robust identification methods. These methods enable one to reliably obtain models based on measured excitations and the corresponding response. The required thermal excitation and response time-series data may include data obtained by state-of-the art devices (wireless, roving, NILMs installed at the service entrance and HVAC subpanel) as well as time-series data routinely monitored by existing HVAC control systems.

The technical approaches for characterizing envelope thermal response and for assessing the application of such models to control, commissioning, and FDD are developed followed by brief summary of test zone instrumentation. Results of application are presented in the *Results* section which follows *Model Synthesis.*

3.3.1 Thermal Response Model

The previously described control and fault detection applications have one very basic thing in common: the need to predict building- or zone-specific thermal response. The ideal properties of a thermal response model are that it should be simple yet give accurate predictions over the expected

range of operating conditions, including extremes and abnormal conditions. A model based on the physics of the problem, if sufficiently simple, is often best. A model that is generally applicable to linear and weakly nonlinear thermal systems and some of its key properties are derived below.

Heat transfer and storage in the building structure and contents is governed (Carslaw and Jeager 1959) by the diffusion (Fourier's) equation:

$$\rho c_p \frac{\partial T}{\partial t} = k\nabla^2 T \quad (3\text{–}D)$$
$$\rho c_p \frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2} \quad (1\text{–}D) \tag{1}$$

where k=conductivity (assumed isotropic) and ρc_p=thermal capacitance are the local material properties of interest. In steady-state, (1) reduces to Laplace's equation:

$$k\nabla^2 T = 0 \quad (3\text{–}D)$$
$$k\frac{\partial^2 T}{\partial x^2} = 0 \quad (1\text{–}D) \tag{2a}$$

to which boundary conditions are applied after integration to obtain heat flux (W/m^2 or Btuh/ft^2):

$$\vec{q} = -k\nabla T \quad (3\text{–}D)$$
$$q_x = -k\frac{dT}{dx} \quad (1\text{–}D) \tag{2b}$$

Applied to a wall (or an enveloping surface) in steady state:

$$Q_z = u(T_w - T_z) \tag{3}$$

where T_w is a uniform exogenous temperature[15], T_z, is zone[16] temperature (Seem 1987), and Q_z is net heat input (flux integrated over area, W or Btuh) to the zone and $u = \Sigma UA$ (W/K or Btuh/°F) is the parallel conductance of all sub-areas defined on the envelope.

With multiple walls:

$$Q_z = \sum_{w=1}^{W} u_w(T_w - T_z) \tag{4}$$

An important constraint implicit in (4) will become apparent momentarily. Consider the discrete-time, linear, time-invariant dynamic system with one wall. We have used the symbol t for continuous time (1) but t is now, and hereafter, defined as the discrete time index. It can be shown via a state-space formulation and the state-transition matrix of the resulting diagonalized system that the 1-D case of (1) corresponds to the following discrete time expression:

$$B^{0:n}(\phi_z, Q_z) + B^{0:n}(\theta_w, T_w) - B^{0:n}(\theta_z, T_z) = 0 \tag{5}$$

where the coefficient vectors each have $n+1$ elements:

$\phi_z = [\phi_{z,0}\ \phi_{z,1} \cdots \phi_{z,n}]$
$\theta_z = [\theta_{z,0}\ \theta_{z,1} \cdots \theta_{z,n}]$
$\theta_w = [\theta_{w,0}\ \theta_{w,1} \cdots \theta_{w,n}]$,

the time-varying condition and response vectors have corresponding elements:

---

[15] Several exterior surface or sol-air temperatures may be included as forcing functions
[16] Seem's linearized star model approximates the weighted mean of air and mean-radiant temperatures

$Q_z = [Q_z(t)\ Q_z(t\text{-}1) \cdots Q_z(t\text{-}n)]$
$T_z = [T_z(t)\ T_z(t\text{-}1) \cdots T_z(t\text{-}n)]$
$T_w = [T_w(t)\ T_w(t\text{-}1) \cdots T_w(t\text{-}n)]$,

and each backshift polynomial term in (5) expands to a sum, as follows:

$$B^{0:n}(x,y) = \sum_{k=0}^{n} x_k y(t-k) \cdot$$

It is convenient, when disassociating the 0-lag terms of (5), to use a more general form of the backshift polynomial notation in which the index of summation starts not at 0 but at i:

$$B^{i:n}(x,y) = \sum_{k=i}^{n} x_k y(t-k) \cdot$$

Note that one coefficient in (5) can be assigned arbitrarily; by convention (Stephenson 1971, Walton 1983) $\phi_z = -1$ so that (5) becomes:

$$Q_z = B^{1:n}(\phi_z, Q_z) + B^{0:n}(\theta_w, T_w) - B^{0:n}(\theta_z, T_z) \qquad (6a)$$

for one wall and, for multiple walls (Seem 1987) we have:

$$Q_z = B^{1:n}(\phi_z, Q_z) + \sum_{w=1}^{W} B^{0:n}(\theta_w, T_w) - B^{0:n}(\theta_z, T_z) \qquad (6b)$$

This form is the sought-after (Figure 3.1) response function that predicts heating or sensible cooling load, Q_z, given zone temperature, T_z, and several exogenous sol-air temperatures, T_w.

To satisfy the steady-state response model (4) the coefficients must satisfy the following constraint:

$$\frac{\sum_{w=1}^{W}\sum_{k=0}^{n}\theta_{w,k}}{\sum_{k=0}^{n}\phi_{z,k}} = \frac{\sum_{k=0}^{n}\theta_{z,k}}{\sum_{k=0}^{n}\phi_{z,k}} = \sum_{w=1}^{W} u_w \qquad (7a)$$

equivalently:

$$\sum_{w=1}^{W}\sum_{k=0}^{n}\theta_{w,k} + \sum_{k=0}^{n}\theta_{z,k} = 0 \qquad (7b)$$

In the steady-state formulations (3, 4) the sum of temperature coefficients equals zero by definition but in the discrete time (DT) formulation (5, 6) constraint (7b) must be consciously enforced. This constraint appears to have been overlooked in the thermo-fluid identification literature where models in the form of (6,8) have been used.

The DT model (5) can be evaluated recursively for Q_z as in (6), or for T_z as follows:

$$\theta_{z,0} T_z = B_0^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_1^n(\theta_z, T_z) \qquad (8)$$

or, equivalently, T_z = RHS expression divided by $\theta_{z,0}$.

An expression for Q_z like (6) is known as a comprehensive room transfer function, or CRTF (Seem 1987). We will call the corresponding expression for T_z (8) the *inverted CRTF* or iCRTF. Expressions (6) and (8) are the two basic forms of desired "crystal ball" postulated in Figure 3.1.

There are additional conditions (constraints) the coefficients must satisfy for the CRTF (6,8) models to be physically (thermodynamically) plausible. To see this one must start with the fact that the

system can be represented by $1+W$ transfer functions for zone thermal load response, Q, and $1+W$ transfer functions for zone temperature response, T_z, where W is the number of walls exposed to different driving temperatures or sol-air conditions. In the time (z-transform) domain we have:

$$G_{Q/Tz}(z) = \frac{Q_z}{T_z} = \frac{B_0^n(\theta_z)}{B_0^n(\phi_z)} \qquad (9) \qquad G_{Tz/Q}(z) = \frac{T_z}{Q_z} = \frac{B_0^n(\phi_z)}{B_0^n(\theta_z)} \qquad (10)$$

$$G_{Q/Tw}(z) = \frac{Q_z}{T_w} = \frac{B_0^n(\theta_w)}{B_0^n(\phi_z)} \qquad (11) \qquad G_{Tz/Tw}(z) = \frac{T_z}{T_w} = \frac{B_0^n(\theta_w)}{B_0^n(\theta_z)} \qquad (12)$$

Note that the CRTF is given by

$$Q_z(z) = G_{Q/Tz}(z)T_Z(z) + G_{Q/Tw}(z)T_W(z)$$

and the iCRTF is

$$T_z(z) = G_{Tz/Q}(z)Q_Z(z) + G_{Tz/Tw}(z)T_W(z).$$

The roots of the denominators correspond to the eigenvalues, λ_k, and thus to the time constants, $\tau_k = -1/\lambda_k$, of any system (CRTF or ICRTF) of interest as follows:

$$\ell n\left(\frac{1}{r_k}\right) = \lambda_k \Delta t \qquad (13)$$

equivalently

$$\ell n(r_k) = \frac{\Delta t}{\tau_k} \qquad (14)$$

The polynomial form may be reconstructed from its roots by Pascal's triangle to satisfy:

$$\prod_{k=1}^{n}(r_k - B) = \frac{1}{x_n}\sum_{k=0}^{n} x_k B^k \qquad (15)^{17}$$

where

$[x_k] = \varphi_z$ for the CRTF, and

$[x_k] = \theta_z$ for the iCRTF.

For diffusion processes, the characteristic frequencies, or poles, λ_k, must be real and negative, corresponding to time constants, τ_k, that are real and positive. Note that the roots of x (θ_z or ϕ_z) are monotonic in both τ and λ. It has been shown (Hittle and Bishop 1983) that zone flux and zone temperature response function pole locations must alternate along the real axis. That is, if the roots of θ_z, $[\lambda_{\theta k}]$, and the roots of ϕ_z, $[\lambda_{\phi k}]$, are each presented as ordered sets, then together they satisfy:

$$\lambda_{\theta 1} < \lambda_{\phi 1} < \lambda_{\theta 2} < \lambda_{\phi 2} < \cdots < \lambda_{\theta n} < \lambda_{\phi n} \qquad (16)$$

Unless the coefficients of the transfer function denominators are somehow constrained, the parameter estimation process is likely to return one or more thermodynamically infeasible poles. With some additional manipulation of (16), as will be shown in 3.3.2, a bounded nonlinear search algorithm, such as Matlab's fminbnd(), can be made to handle this constraint set.

---

[17] Multiplying through by x_n gives $x_n\prod_{k=1}^{n}(r_k - B) = \sum_{k=0}^{n} x_k B^k$, which has cross terms but is more useful in practice; note that the vector (6) and scalar backshift operators are related by $B^{k:k}(x,y) = x_k B^k(y) = x_k y(t-k)$

3.3.2 Model Identification and Application.

There are several ways to estimate the parameters of (6) and (8) given noisy data. There are a number of different ways, in terms of problem formulation, to account for observational noise and model constraints (Golub 1980, Wunsch 2004). Three formulations relevant to zone thermal response (CRTF and iCRTF) modeling are developed below.

Model Identification with U-value Constraints. The discrete-time, linear, time-invariant simulation model (5) can, in principle, be obtained from observations of thermal response under a range of (zone heat rate and sol-air temperature) excitations by applying linear least squares.

However, unconstrained least-squares analysis minimizes the model-observation deviations without regard for the thermodynamic constraints previously noted. Moreover, the inverted model typically is very inaccurate under certain transients and it would be desirable not to have to maintain two incompatible models.

We therefore normalize temperatures by subtracting current zone temperature or one of the sol-air temperatures from all other current and lagged temperatures. This eliminates the current zone (or sol-air) temperature term, reduces the order of the least squares problem by one, and results in a solution that satisfies the U-value constraint (Armstrong 2000).

Model Identification with Constrained Poles. The constraints represented by eqns (14-16) cannot be implemented in such a simple manner. However, a formulation that fits a standard bounded non-linear search model is possible. From eqns 14 and 16 one can write the constraint as

$$r_{\theta 1} > r_{\phi 1} > r_{\theta 2} > r_{\phi 2} > \cdots > r_{\theta n} > r_{\phi n} \tag{17}$$

We can now define a new search vector, x, each element of which has simple bounds, $0 < x < 1$ from which the roots can be evaluated recursively:

$$\begin{aligned}
r_{\theta 1} &= x_{\theta 1}\, r_{\theta max} & 0 < x_{\theta 1} < 1 \\
r_{\phi 1} &= x_{\phi 1}\, r_{\theta 1} & 0 < x_{\phi 2} < 1 \\
r_{\theta 2} &= x_{\theta 2}\, r_{\phi 1} & 0 < x_{\theta 2} < 1 \\
r_{\phi 2} &= x_{\phi 2}\, r_{\theta 2} & 0 < x_{\phi 2} < 1 \\
&\cdot & \cdot \\
&\cdot & \cdot \\
&\cdot & \cdot \\
r_{\theta n} &= x_{\theta n}\, r_{\phi n-1} & 0 < x_{\theta n} < 1 \\
r_{\phi n} &= x_{\phi n}\, r_{\theta n} & 0 < x_{\phi n} < 1
\end{aligned} \tag{18}$$

where

$$r_{\theta max} = exp(\Delta t / \tau_{\theta max})$$

There are two useful properties—besides reducing the multi-variable constraints (16) to a set of simple bounds on individual variables (18)—that arise from this formulation. First, note that $x_{\theta 1}$ corresponds to the largest time constant and a reasonable upper limit, τ_{max}, on this time constant can almost always be easily estimated. Second, minimum spacing of time constants can be imposed by using a lower bound that is greater than zero or an upper bound that is less than one.

Implementation of the foregoing model identification formulation, coded as a Matlab script and a function called by Matlab's built-in bounded, nonlinear least squares routine, *lsqnonlin*, are presented in Appendix I.

Without pole constraints, least squares identification will typically return parameters that correspond to an unstable model (simulation grows without bound), complex or imaginary eigenvalues (resonance) or negative time constants (noncausality). None of these three properties is possible in a diffusion system.

Model Identification with Uncertainty. Evaluation of thermal response is needed in both of the two modes, (6) and (8). Ordinary least squares (OLS) analysis minimizes the Euclidean norm of observation residuals. Thus if OLS is applied to heat rate observations in order to identify a CRTF model (6), temperatures forecast by the corresponding ICRTF model (8) are likely to have a relatively large error norm. Conversely, if OLS is applied to zone temperature observations in order to identify an ICRTF model, heat rates forecast by the corresponding CRTF model are likely to have a relatively large error norm. Total least squares (Golub 1980) may provide the better model parameter estimates if reasonable estimates of the uncertainties are available for all the variables.

The exogenous temperature errors are of less concern than the zone temperature and heat rate errors because we are not using the models to estimate or forecast the exogenous temperatures. This suggests an alternate formulation:

$$Q_z + \theta_{z,0} T_z = B^{1:n}(\phi_z, Q_z) + B^{0:n}(\theta_w, T_w) - B^{1:n}(\theta_z, T_z) \tag{19}$$

in which the objective function for m observations $(t = 1{:}m)$ is

$$J_g = \|n[1{:}m]\| = \|Q[1{:}m] + \theta_{z,0} T_z[1{:}m] - x_g[1{:}m] b_g\| \tag{20}$$

where

$$x_g(t) = [Q(t\text{-}k,\ k=1{:}n)\ T_z(t\text{-}k,\ k=1{:}n)\ T_1(t\text{-}k,\ k=1{:}n)\ T_2(t\text{-}k,\ k=0{:}n)\ \cdots\ T_W(t\text{-}k,\ k=0{:}n)]$$

$$b_g = [\phi(1{:}n)\ \theta_z(1{:}n)\ \theta_1(1{:}n)\ \theta_2(0{:}n), \cdots, \theta_w(0{:}n)].$$

In this formulation the zero-lag zone temperature coefficient, $\theta_{z,0}$, is estimated by non-linear least squares. The LHS observation vector is then generated and the remaining coefficients are estimated at each iteration by OLS. Matlab code is listed in Appendix J. The foregoing objective function (20), in combination with the non-linear formulation that constrains poles, is another approach to be tested.

3.3.3 Model Application.

Now we are in a position to fill in some of the Figure 3.1 details. Evaluation of thermal response is needed in both of the two modes described by eqns 9-12. For example, model-based control typically treats T_z as a given trajectory (Kintner-Meyer 1994) and simulated Q is added to the command given by a local loop controller. Conversely, for optimal start, a Q trajectory is given and the response of interest is temperature as start of occupancy approaches. For curtailment, both modes are important. In deciding whether a given level of curtailment is feasible a Q trajectory (corresponding to the reduction in chiller power called for by the ISO) is given and the response of interest is comfort, i.e., the zone temperature trajectory. Once committed to curtailment the verification of demand reduction

can be based on the original CRTF model which tells us how much heating or cooling *would have been* required if room temperature had been maintained.

The heat rate response mode of model application is illustrated in Figure 3.2. Inputs include all of the exogenous conditions of weather (temperature, solar gains on the main opaque and window surfaces of the envelope) and building operation (light and plug and certain other—mostly non-HVAC—electrical gains, metabolic gains). Zone temperature is an input and heat rate is the output of interest. The transfer functions of eqns 9 and 11 give this response by superposition.

The temperature response mode of model application is illustrated in Figure 3.3. Inputs include the same exogenous conditions of weather and building operation. Net heat rate is an input and MRT or zone air temperature[18] is the output of interest. The transfer functions of eqns 10 and 12 give this response by superposition.

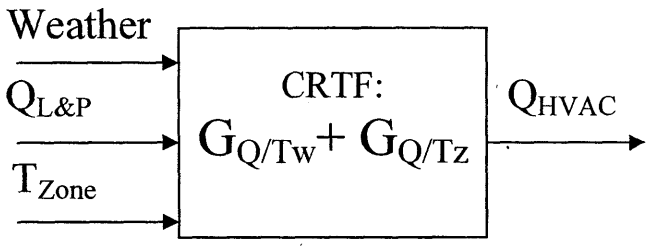

Figure 3.2 Building thermal response modeled as a comprehensive room transfer function (CRTF)

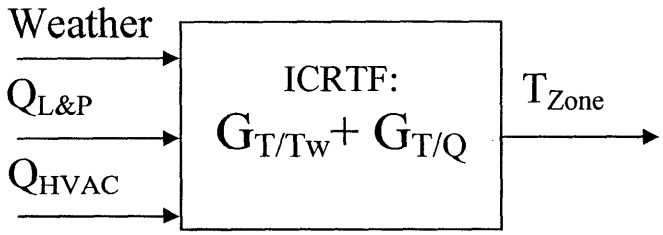

Figure 3.3 Building thermal response modeled as an inverted CRTF

Training and cross validation data collected from three very different buildings are used in the next three sections to demonstrate the identification methods. The building characteristics, instrumentation and test conditions are documented. The resulting thermal response models are described and characterized in terms of model order, equivalent thermal parameters, parameter confidence intervals, and residual analysis.

[18] Any of several response temperatures may be defined, including star temperature [Davies, Seem 1987], operational temperature, $T_{op}=(MRT+T_{air})/2$, or other measures of perceived temperature.

3.4 Test Room Results

Two large rooms, shown in Figure 3.4, were instrumented to provide training and testing data with which the model and identification methods can be evaluated.

3.4.1 Room Description.

The thermal regime in a room is complex. Surface temperatures are constantly changing under the influence of conduction (most strongly in windows and exterior walls) convection and radiation. Convection may at times be dominated by buoyant forces, fan-induced, or virtually nonexistent. Boundary layer flows develop in response to surface-to-bulk-air temperature differences. The bulk air may be stratified (stagnant or in a simple, stable mode, such as displacement ventilation), turbulent (mixed), or even divided into two regime-layers, one mixed and one stratified.

Temperature sensors generally respond to the radiant field as well as to local air temperature and the split depends strongly on local air velocity and on sensor size and emissivity. The most troublesome measurements for advanced control are therefore often the temperatures of room air and the MRT fields. For control applications, we need a single estimate of zone air temperature and/or of occupant-

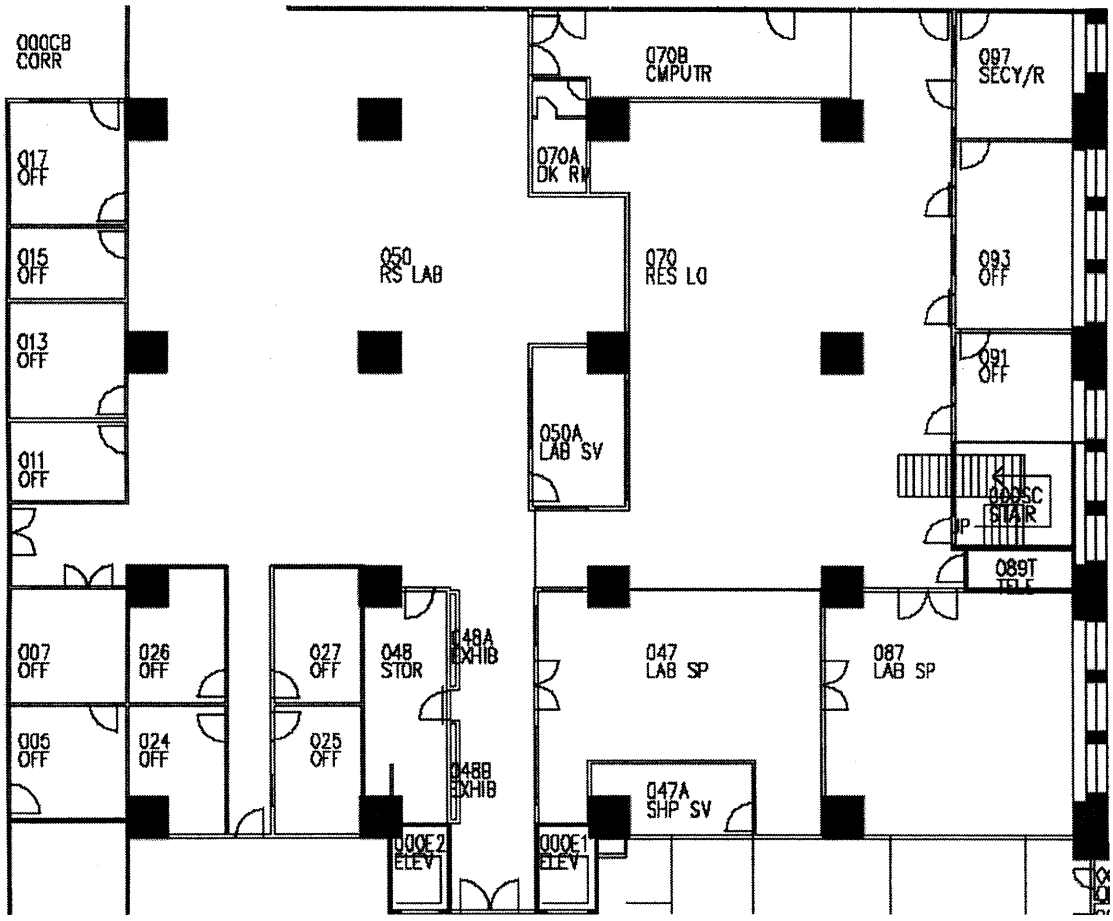

Figure 3.4 Main lab floor plan showing monitored zones in lower right. Room 047 is the shop and room 087 is the prototyping lab. Columns are 46"x 46" concrete.

perceived temperature. Also note that the air stream heat rate is directly proportional to the difference between supply and exhaust temperature. The latter is often difficult to measure in the case of plenum air returns where interior and perimeter air returning at different temperatures will mix.

To obtain reliable spatial average zone temperatures, as well as indications of local variations, seven to eight temperature loggers were deployed in each of the two test zones.

A two-foot diameter axial fan draws air from the main lab areas (4850 ft^2 to north and west) via a two-foot diameter supply duct, and distributes supply air to the shop (740 ft^2), to the proto lab (650 ft^2), or to both simultaneously. Zone air is allowed to return to the main lab through one or both sets of double doors.

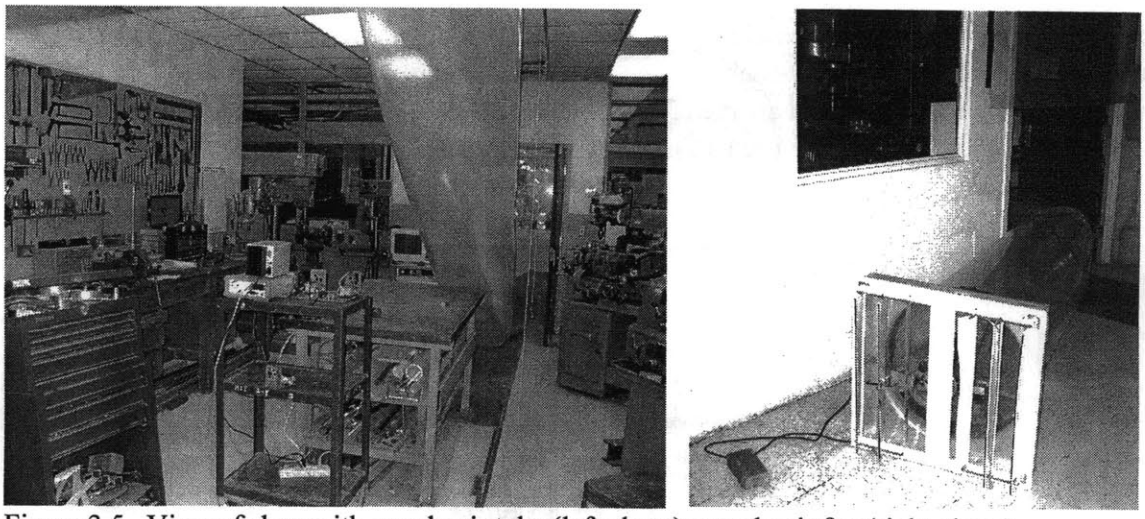

Figure 3.5. View of shop with supply air tube (left photo); supply air fan (right photo)

The lighting load (ten 60W fixtures) presents a connected load, measured by the Fluke 39, of 590W = 2013 Btuh. It is the completely dominant source of internal gain in the shop. The proto-lab has computers and bench meters that add about 900 W to the 530 W connected load of lighting. The supply air fan flow rate was measured and the corresponding thermal capacitance rate calculated to be 1620 Btuh/°F (855 W/K). Wall, floor, and ceiling construction are documented in Appendix E.

Temperature loggers were deployed as indicated in Table 3.1 to assess spatial variations. Additional loggers monitored supply and exit air stream temperatures. The loggers sample at 2Hz, record sample averages at 5-minute intervals, and are accurate to +/-0.1°F in the room temperature (65-80°F) range.

3.4.2 Test Protocol.

The existing HVAC fan coil unit serving the zones was turned off at the start of monitoring. Internal gain variations were effected by turning the lights on and off. Envelope loads resulted primarily from diurnal temperature excursions in the main lab and outside. Selected temperature trajectories logged 6 September through 10 October, 2003, are plotted in Figure 3.6. Figure 3.7 shows the same trajectories for the identification training and testing periods.

Table 3.1. Temperature logger locations

	Logger Serial#	Location Code	Room	planLoc	elevation	Moved 20030909	Group Average
C	678777	Sh-E-1	shop	East	Floor		Shop
D	678778	Sh-N-3	shop	North	Hat		Shop
E	678779	Pl-C-2	protoLab	center	Table		protoLab
F	678780	Sh-N-1	shop	North	Floor		Shop
G	678781	Fan-In-1	fanInlet		Floor		Shop
H	678782	Pl-E-3	protoLab	East	Hat		protoLab
I	678783	Fan-In-2	fanInlet		26"AFF		
J	678784	Sh-N-2	shop	North	Table	shopOut	
K	678785	Sh-E-3	shop	East	Hat		Shop
L	678786	Pl-W-3	protoLab	West	Hat		protoLab
M	678787	Sh-S-3	shop	south	Hat		Shop
N	678788	8788	shop				Shop
O	678789	Pl-W-1	protoLab	West	Floor		protoLab
P	678790	Sh-S-2	shop	south	Table	protOut	
Q	678791	Sh-E-2	shop	East	Table	protIn	
R	678792	Sh-S-1	shop	south	Floor		Shop
S	678793	Pl-S-1	protoLab	south	Floor		protoLab
T	678794	Pl-N-1	protoLab	North	Floor		protoLab
U	678795	Pl-S-3	protoLab	south	Hat		protoLab
V	678796	Fan-out	duct				

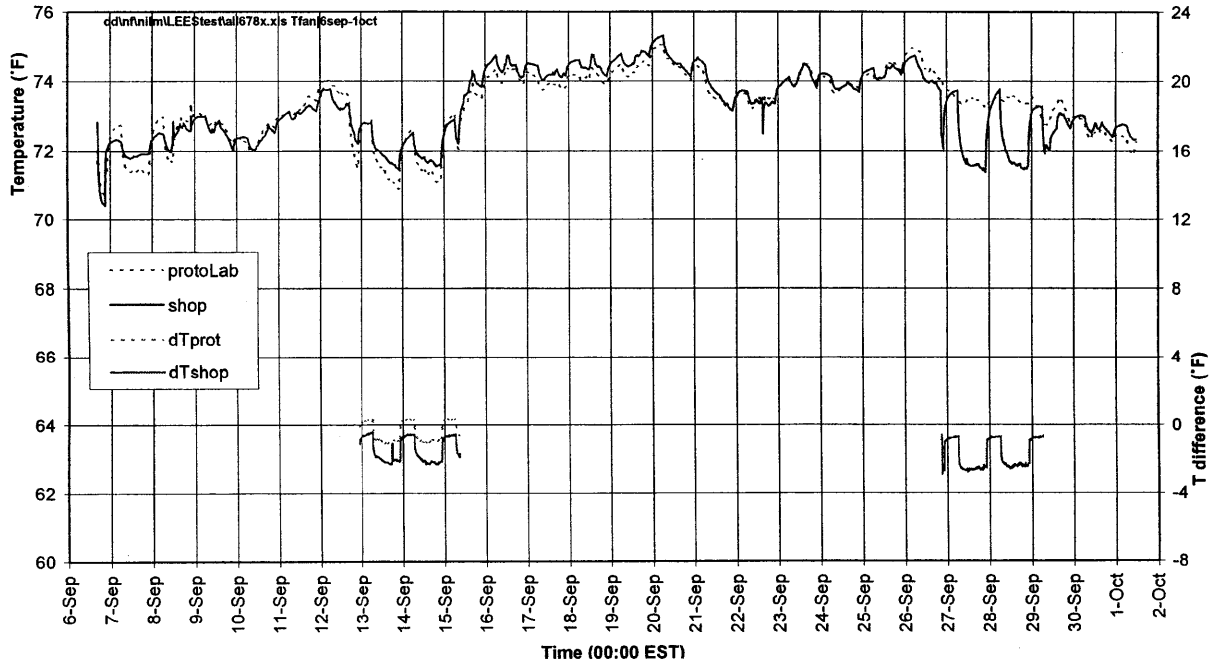

Figure 3.6 Temperatures measured in September and early October, 2003. Temperature differences are shown only during periods of fan operation. In the 12-15 September test air is delivered to the proto-lab and exits to the shop by the connecting door (dTprot). Air exits the shop by its west door (dTshop). In the 26-29 September test the connecting door is shut and air is delivered directly to the shop, leaving again by the west door.

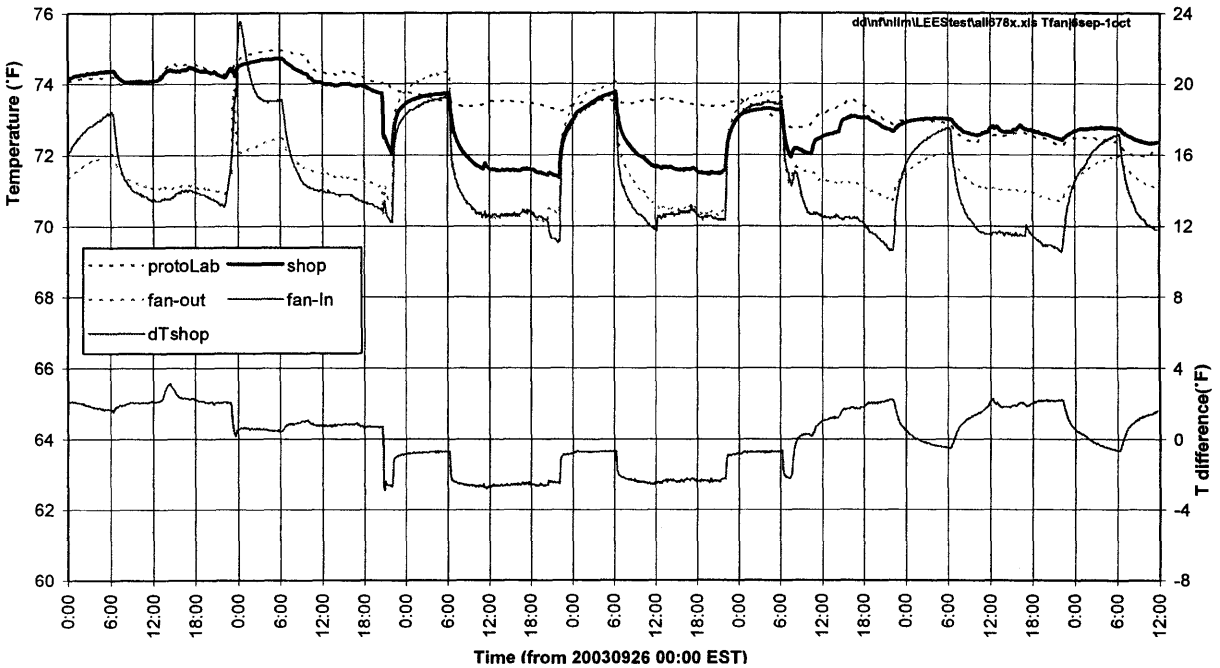

Figure 3.7. Temperatures measured in the test zones from 25 September to 1 October 2003.

3.4.3 Zone Thermal Conditions.

One of the main challenges to practical model identification is the accurate measurement of excitations and responses. For building thermal response, the spatial variation of temperature is a particular concern. Consider the conditions monitored during the period from which training and testing data sets have been drawn, specifically the six days beginning 10:00 am, 25 September.

The test zone temperature trajectories can be explained, at least qualitatively, in terms of three exogenous conditions: ambient temperature, fan operation, and the internal gains of lighting and other electrical loads.

Ambient temperature. The shop is strongly driven by temperatures in the main lab because the thermal resistance of the large windowed partitions is low. The prototyping lab, with its large east window exposure to the outside, is more strongly driven by outside temperature and is driven, additionally, by diffuse solar gains. The partition between the two zones means that each is driven to some extent by the temperature in the other. Thus the proto-lab has four significant exogenous inputs (three temperature and solar radiation) while the shop has only two temperature inputs.

Internal gains. Proto-lab internal gains consist of lighting and computers. The heat from computers generally rises in a plume and diffuses across the ceiling. This, and the fact of additional heat from lighting fixtures in the ceiling, promotes thermal stratification. Weak lateral circulation occurs, even in the absence of fan operation, to place a limit on the degree of thermal stratification that can develop. The shop is dominated by gains from lighting in the suspended ceiling and behaves similarly.

Fan operation. Zone air is generally mixed when the fans are operating.

Against this background of generally expected behavior one can begin to understand the observations from individual temperature loggers deployed in the shop.

Individual logger temperatures are plotted against their average in Figure 3.8. Local temperature is seen to vary in a somewhat systematic way from the average. However, the biases are quite large (right hand band of points with average temperature in the 73.7-74.8˚F range) and more variable when the fan is off.

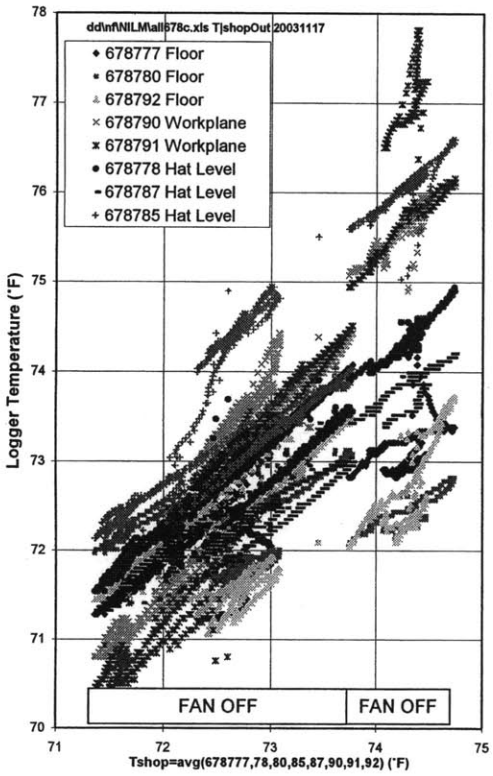

Figure 3.8. Temperatures recorded by eight spatially distributed loggers are plotted against temperature averaged across the loggers. With the fan on, local temperatures deviate in a fairly systematic way from the average. However, the biases are quite large (right hand band of points with average temperature in the 73.7-74.8˚F range) and more variable when the fan is off.

The exhaust temperature is plotted against the average of logger temperatures in Figure 3.9. Exhaust temperature provides a well correlated, but biased, estimate of average zone temperature during fan operation. The bias in a real building with plenum return will generally be a function of lighting power. For the particular exhaust path arrangement present during these tests, the exhaust temperature is very poorly correlated with average room temperature when the fan is off.

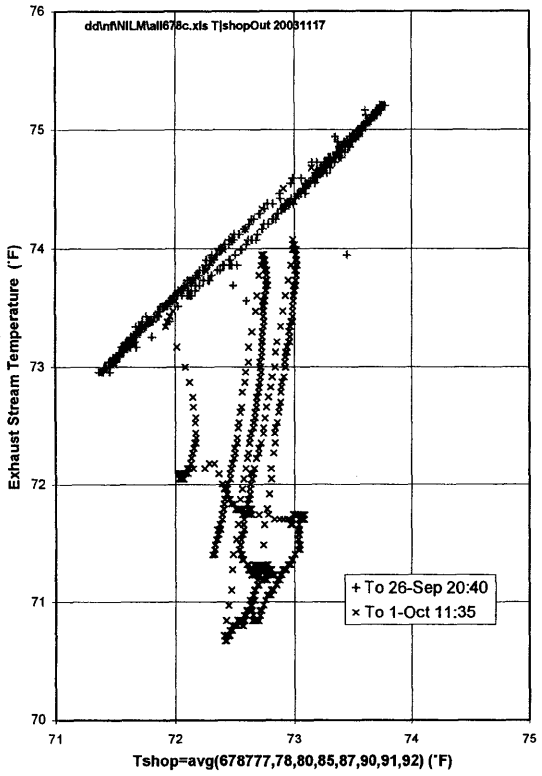

Figure 3.9. Exhaust stream temperature plotted against the average of eight spatially distributed logger temperatures. The fan is on (+) from 26-Sep 20:40 to 29-Sep 06:15 and then remains off (x) until 1-Oct 11:35.

One can imagine an ideal air temperature sensor that would return a spatial average of the bulk air temperature in the occupied region, i.e. to about 2 m above the floor. An ideal MRT sensor would return the temperature of a black, spherical receiver that is unaffected by air temperature and has unobstructed view from a central location[19] of floor, ceiling, and perimeter wall surfaces. In occupied buildings one is generally constrained to sensors mounted on the ceiling (i.e., at the highest stratum) or along perimeter walls (i.e. in the boundary layer flows) and further constrained by current practical sensor technologies to a temperature that is some weighted average of the local air and MRT. The costly practice of averaging over multiple temperature sensors remains, for the foreseeable future, the only practical approach to acquiring representative continuous time-series data. One can only seek to minimize the number of sensors by judicious selection of sensor locations.

[19] i.e., a typical occupant location

3.4.4 Model Identification Results.

Here the discrete-time, linear models developed in the *Model Synthesis* section are identified and used to simulate the dynamic thermal responses of one of the two test rooms. The parameters for each model are determined by least-squares fit using each of the three response models: CRTF, ICRTF and hybrid. The models have identical terms but different objective functions; the identified coefficients can therefore be quite different. Each model is used to simulate, in turn, the two responses of interest: zone temperature and zone heat rate. The results are assessed by residual analysis and cross-validation. In all time series plots of Q and T_z, the forecast response is solid and the observed is dashed.

In all three model identification cases, the exogenous temperatures and time-shifted zone temperature are converted to temperature differences by subtracting any unlagged exogenous temperature series from the other (current and lagged) temperature series. This eliminates the selected unlagged reference series term, reduces the order of the least squares problem by one, and results in a solution that satisfies the constraint on sums of temperature coefficients (7).

The first model identified for the shop is in the form of a CRTF (6) with two thermal capacitances and 15-minute time steps:

$$Q_z = B_1^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_0^n(\theta_z, T_z) \qquad (6)$$

With $\phi_{z,0}$ equal minus one, the least squares solution is given in Table 2 where T is in °F and Q is in Btuh. The building load coefficient (UA) corresponding to the estimated model coefficients is 15634 Btuh/°F. The normalized width of each coefficient's confidence interval (CI) is given in Table 3.2.

Table 3.2. Coefficients, confidence intervals, and time constants from CRTF identification.

Term	Coefficient	|coef|/CI	Root	Tau(hr)
$Q_{,k-0}$	-1.0000	9.6		
$Q_{,k-1}$	1.2024	11.3	1.0205	n.a.
$Q_{,k-2}$	-0.1856	1.4	0.1819	0.34
$T_{x,k-1}$	-1383	1.2		
$T_{x,k-2}$	210	0.3		
$T_{z,k-1}$	-1351	0.7	1.3351	n.a.
$T_{z,k-2}$	2371	0.7	0.4196	0.66
$T_{x2,k-0}$	-757	0.4		
$T_{x2,k-1}$	1163	0.2		
$T_{x2,k-2}$	-2227	0.3		

Shop temperature responses produced by the second-order models are compared to the measured responses in Figure 10. Cross-validation was performed by dividing the data set into a training set (27 Sep 03:30 to 28 Sep 06:30) and a testing set (28 Sep 06:30 to 29 Sep 09:30), shown on the left and right sides of Figure 3.10. Simulated heat input trajectories are compared to the measured trajectory in Figure 3.11. The training and testing sets are again shown on, respectively, the left and right sides. Residual norms are $se(Q,T_z)_{Train}$=*186, 0.138* and $se(Q,T_z)_{Test}$=*445, 0.330.* The residuals corresponding to Figure 3.10 are plotted in Figure 3.12 where large errors are apparent during transient operation (the residuals corresponding to Figure 3.11 differ only by a scale factor). Further evidence that the transient errors are excessive is provided by the tails of the cumulative distributions, Figure 3.13. Note also that the model predictions over the test set (dashed) are significantly biased.

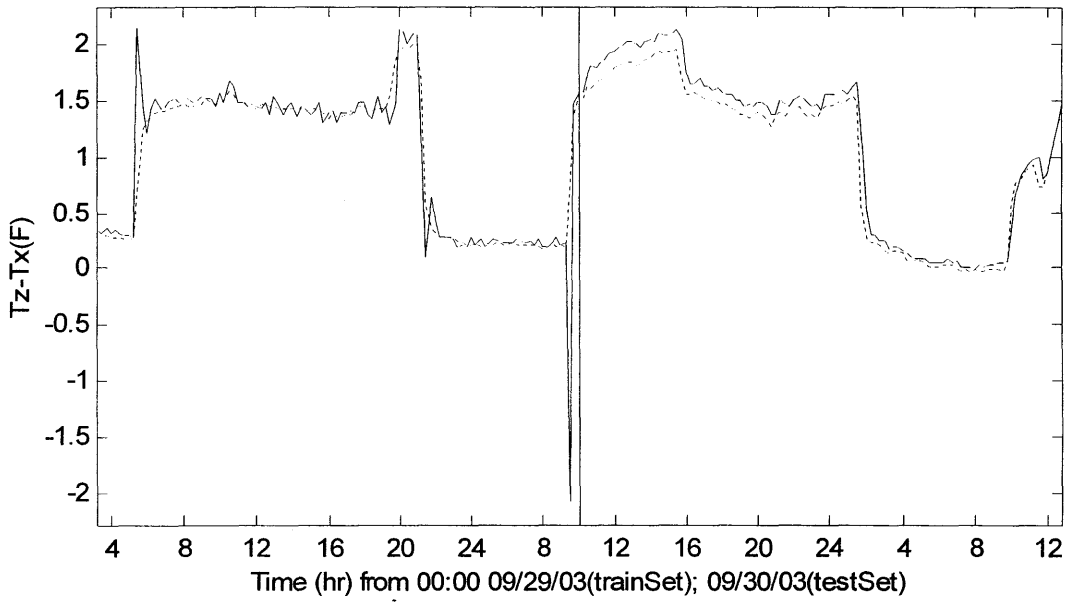

Figure 3.10. Forecast (solid) of temperature by inversion of identified CRTF model

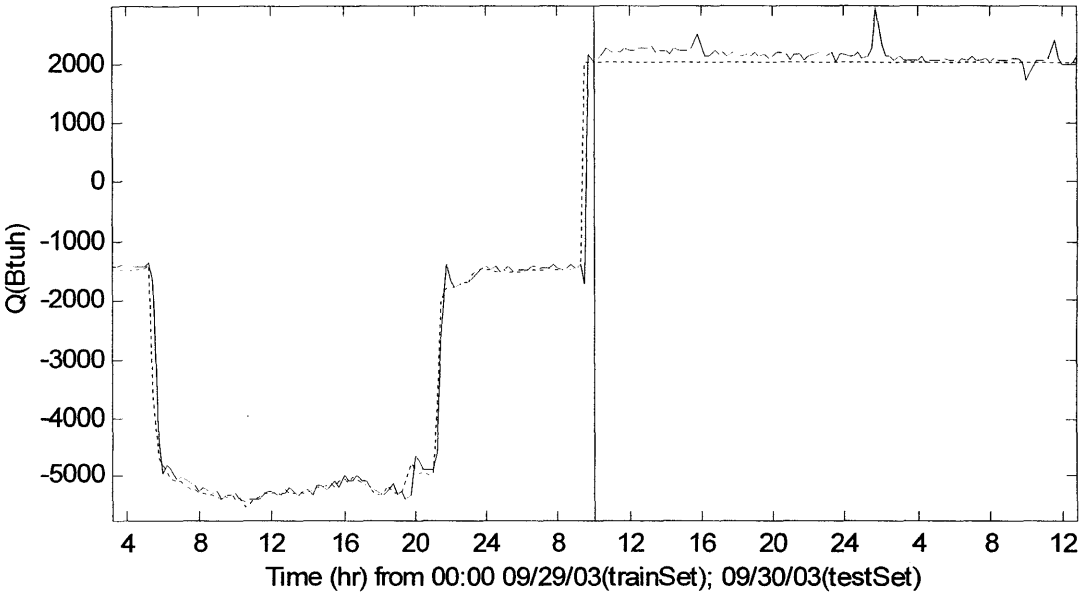

Figure 3.11. Forecast of heat rate by identified CRTF model

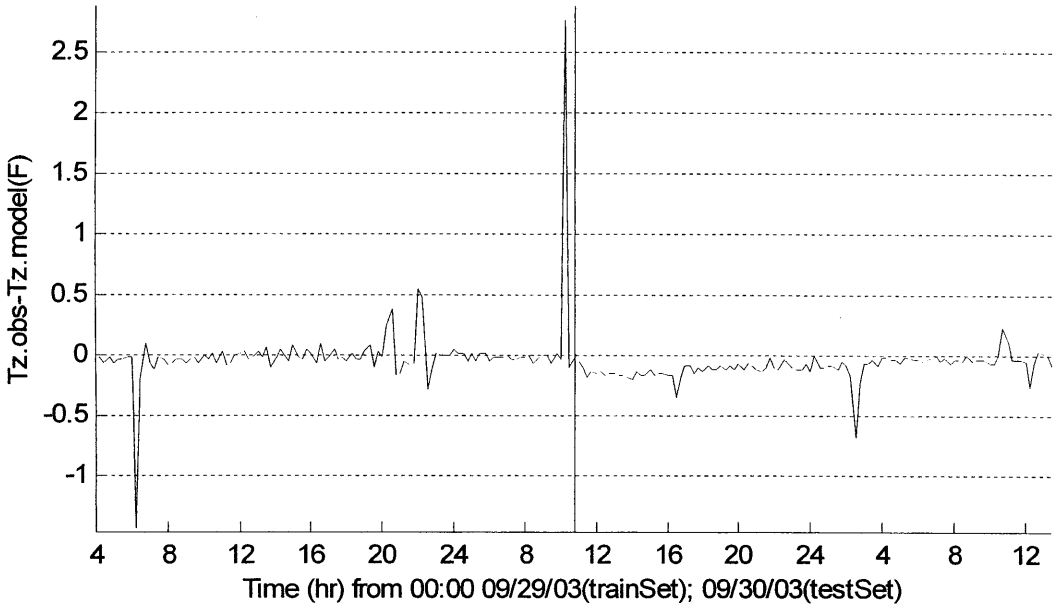

Figure 3.12. Deviation of predicted response from observed over training set (left) and test set (right).

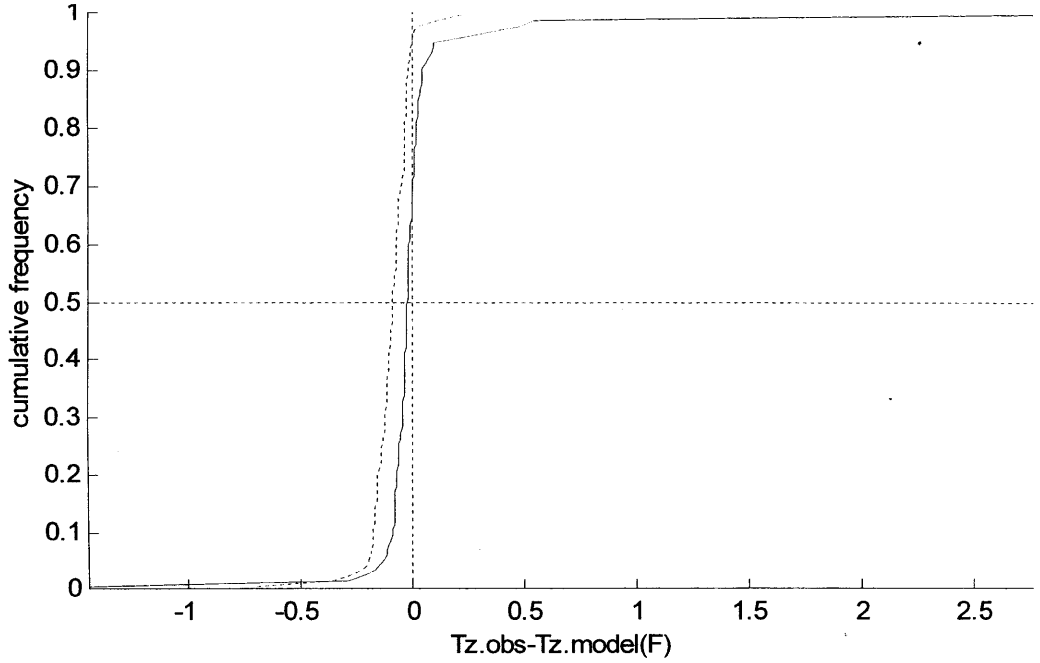

Figure 3.13. Cumulative distributions of prediction error over training set (solid) and test set (dashed).

The second model identified for the shop is in the form of an ICRTF (8) with two thermal capacitances and 15-minute time steps:

$$\theta_{z,0}T_z = B_0^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_1^n(\theta_z, T_z) \tag{8}$$

With $\phi_{z,0}$ equal minus one, the least squares solution is given in Table 3.3 where T is in °F and Q is in Btuh. The zone load coefficient (UA) corresponding to these estimated model coefficients is 8317 Btuh/°F.

Simulated shop temperature responses produced by the second-order models are compared to the measured responses in Figure 3.14. Simulated heat input trajectories are compared to the measured trajectories in Figure 3.15. The training and testing sets are shown on, respectively, the left and right sides; $se(Q,T_z)_{Train}$=*11650, 0.043*; $se(Q,T_z)_{Train}$=*6328, 0.023*. The residuals corresponding to Figure 3.14 are plotted in Figure 3.16 where large errors are apparent during transient operation (the residuals corresponding to Figure 3.15 differ only by a scale factor). Further evidence that the transient errors are excessive is provided by the tails of the cumulative distributions, Figure 3.17. Note also that the model predictions over the test set (dashed) are again significantly biased.

Table 3.3. Coefficients, confidence intervals, and time constants; iCRTF identification.

Term	Coefficient	|coef|/CI	Root	Tau(hr)
$Q_{,k-0}$	-1.0000	0.7		
$Q_{,k-1}$	5.9271	2.7	4.5104	n.a.
$Q_{,k-2}$	-6.3899	3.5	1.4167	n.a.
$T_{x,k-1}$	-139580	14.7		
$T_{x,k-2}$	46230	5.2		
$T_{z,k-1}$	-272720	52.5	0.8887	4.88
$T_{z,k-2}$	405790	20.6	0.5992	1.12
$T_{x2,k-0}$	-145230	7.5		
$T_{x2,k-1}$	-9180	0.1		
$T_{x2,k-2}$	-4580	0.0		

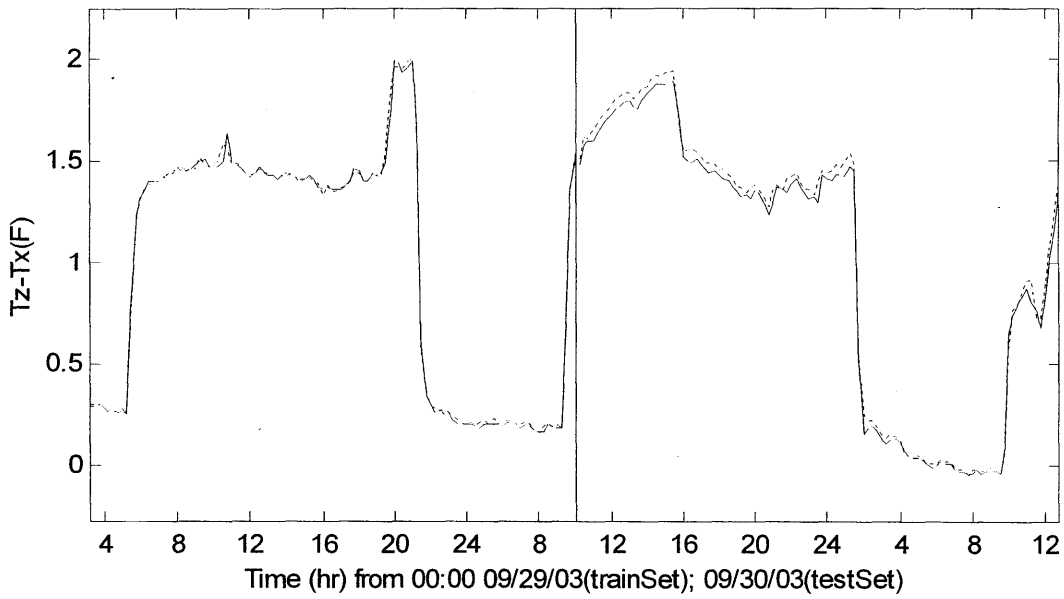

Figure 3.14. Forecast of temperature by identified ICRTF model

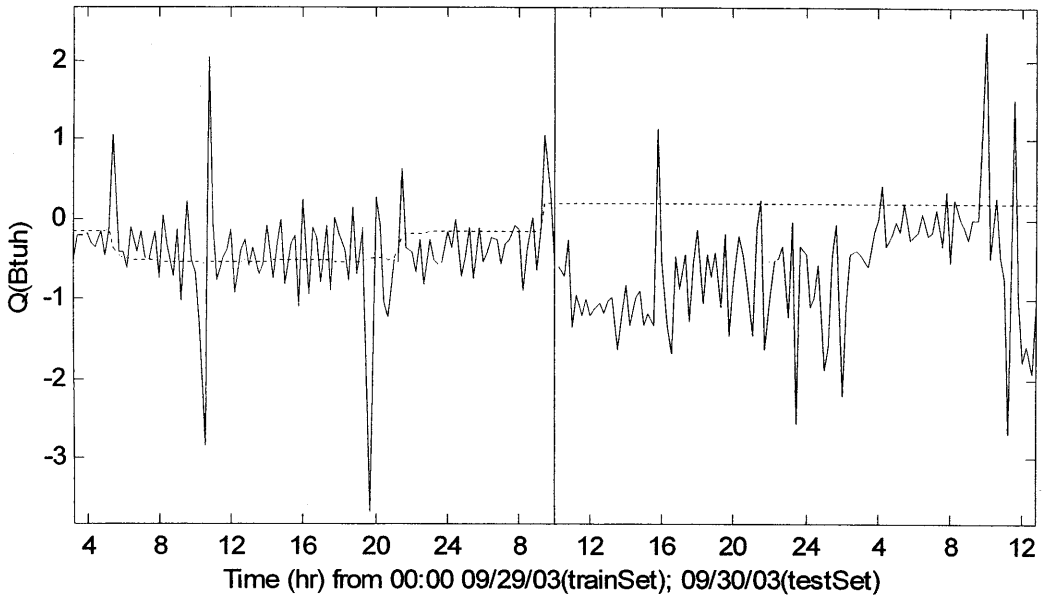

Figure 3.15. Forecast of heat rate by inversion of identified ICRTF model (multiply by 1000)

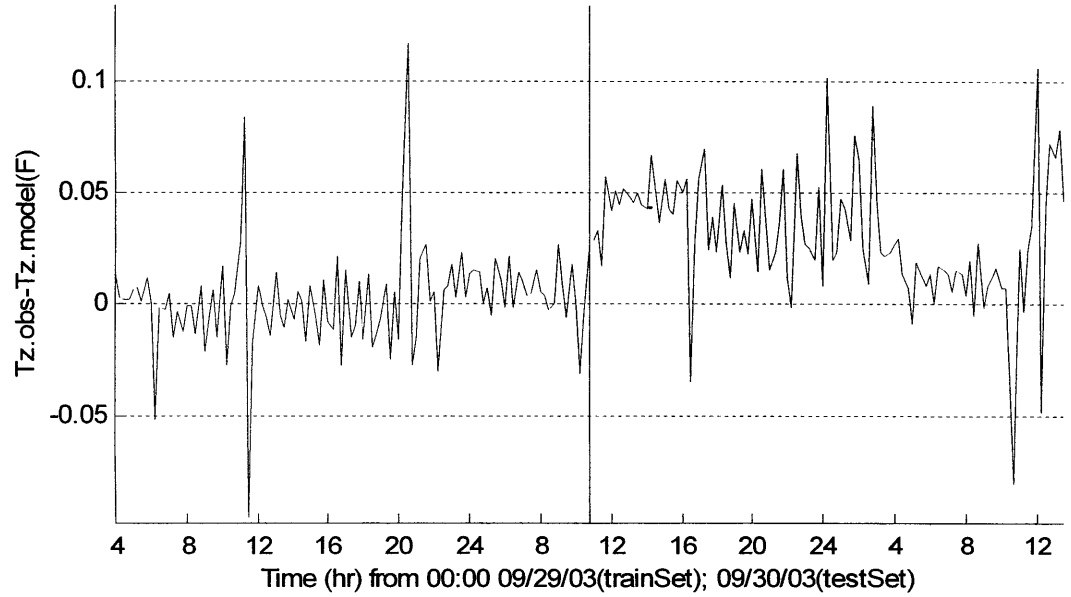

Figure 3.16. Deviation of ICRTF-predicted response from observed over training (left) and test set (right).

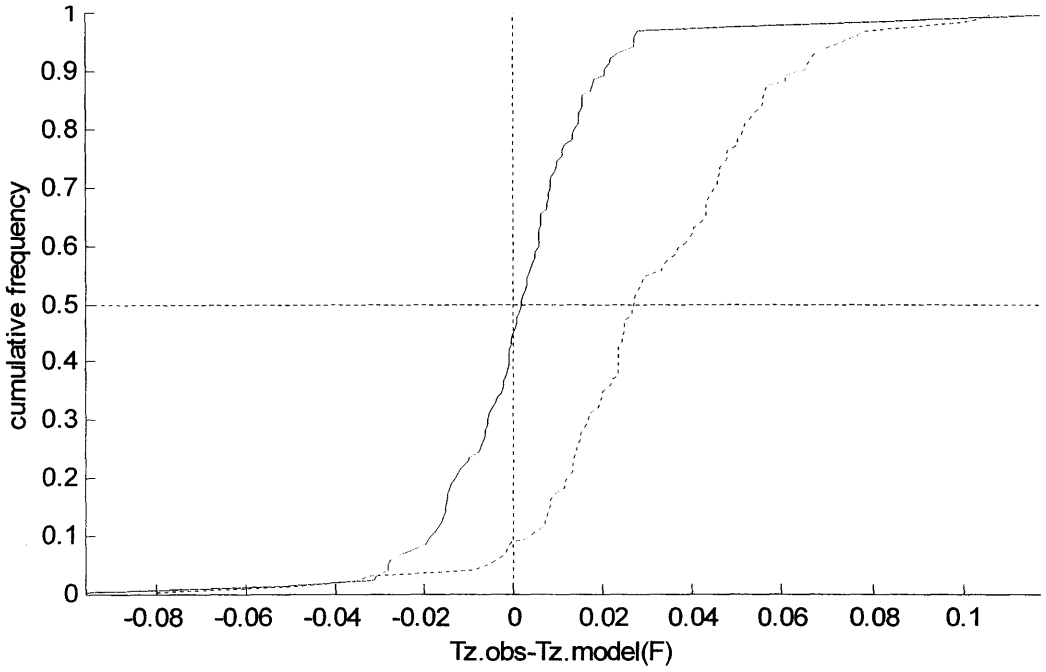

Figure 3.17. Cumulative distributions of ICRTF prediction errors over training (solid) and test set (dashed).

The third model identified for the shop is in the hybrid form (17) with two thermal capacitances and 15-minute time steps:

$$Q_z + \theta_{z,0} T_z = B_1^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_1^n(\theta_z, T_z) \qquad (17)$$

With $\phi_{z,0}$ equal minus one, the least squares solution is given in Table 3.4 where T is in °F and Q is in Btuh. The zone load coefficient (UA) corresponding to these estimated model coefficients is *4980* Btuh/°F.

Table 3.4. Coefficients, confidence intervals, and time constants from hybrid identification.

Term	Coefficient	|coef|/CI	Root	Tau(hr)
$Q_{,k-0}$	-1	8.5	0.9374	8.90
$Q_{,k-1}$	1.3697	12.2	0.4323	0.69
$Q_{,k-2}$	-0.4052	2.9		
$T_{x,k-1}$	-6274	8.5		
$T_{x,k-2}$	1839	2.7	0.9636	15.52
$T_{z,k-1}$	-10956	n.a.	0.556	0.98
$T_{z,k-2}$	16650	10.8		
$T_{x2,k-0}$	-5870	3.9		
$T_{x2,k-1}$	797	0.1		
$T_{x2,k-2}$	-2310	0.2		

Simulated shop temperature responses produced by the 2nd-order models are compared to the measured responses in Figure 3.18. Simulated heat input trajectories are compared to the measured trajectories in Figure 3.19. The training and testing sets are shown on, respectively, the left and right sides. Training residuals norms are *se(Q,Tz)= 346, 0.0316;* testing residual norms are *se(Q,Tz)= 498, 0.0455*. The residuals corresponding to Figure 3.18 are plotted in Figure 3.20 where large errors are apparent during transient operation (the residuals corresponding to Figure 3.19 differ only by a scale factor). Further evidence that the transient errors are excessive is provided by the tails of the cumulative distributions, Figure 3.21. Note that the model predictions over the test set (dashed) are less biased than the predictions by the CRTF and ICRTF models..

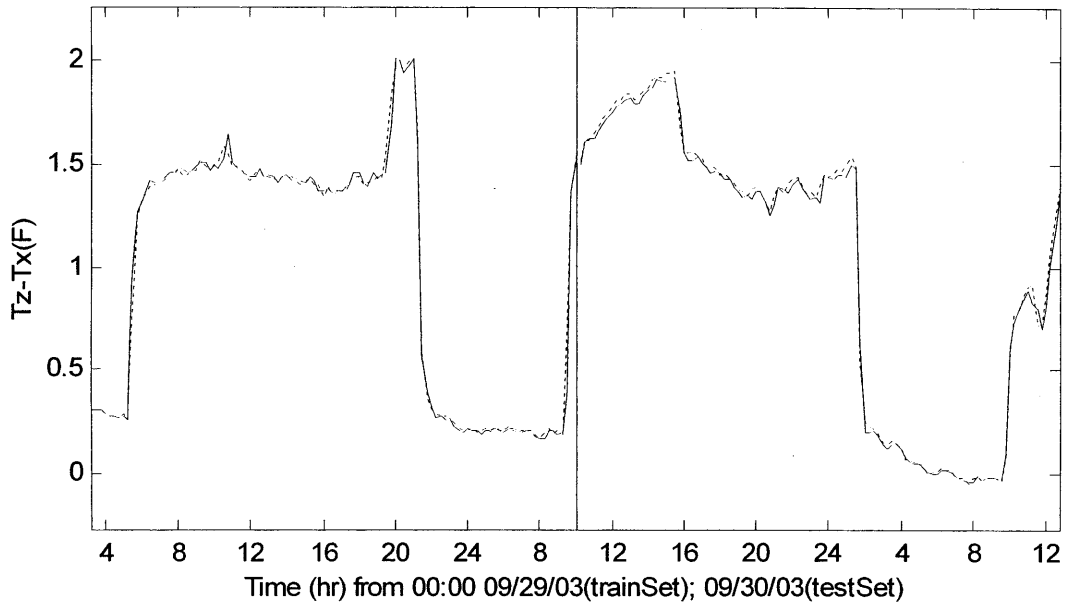

Figure 3.18. Temperature by model identified using weighted sum of zero-lag observations.

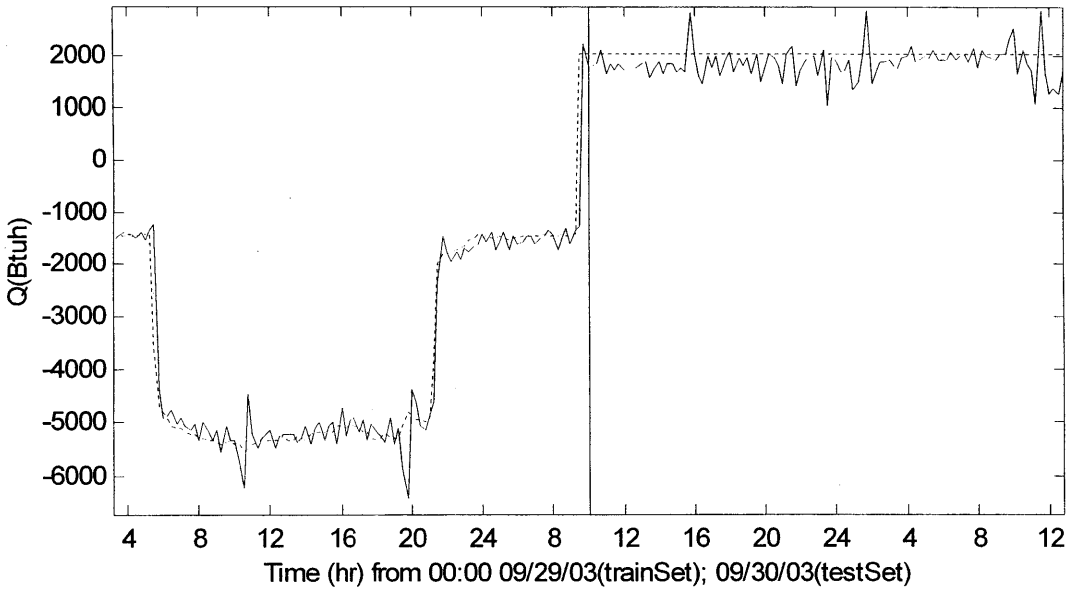

Figure 3.19. Heat rate by model identified using weighted sum of zero-lag response observations.

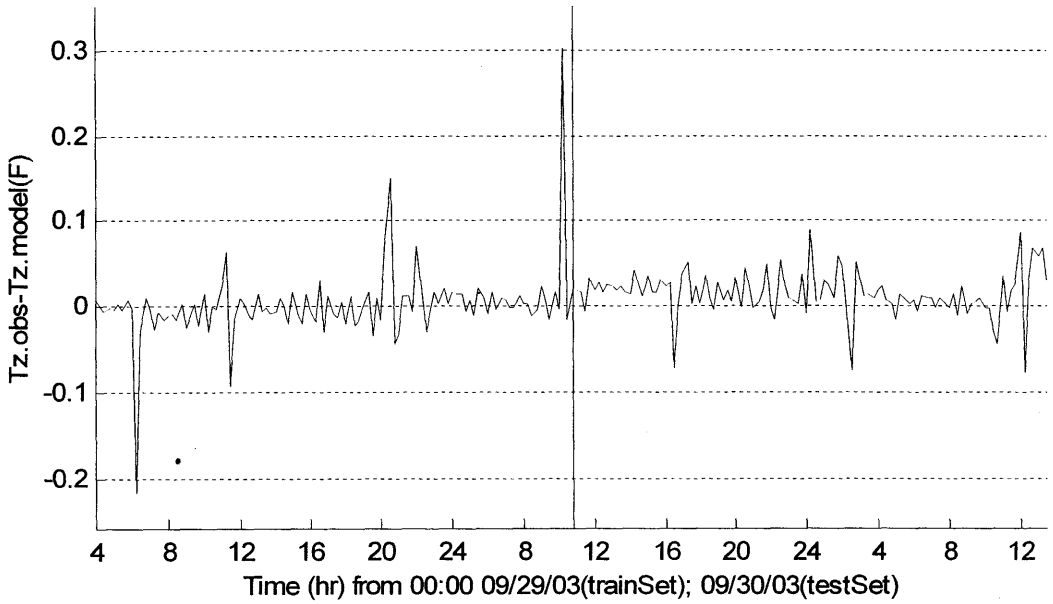

Figure 3.20. Deviation of hybrid-predicted response from observed over training (left) and test set (right).

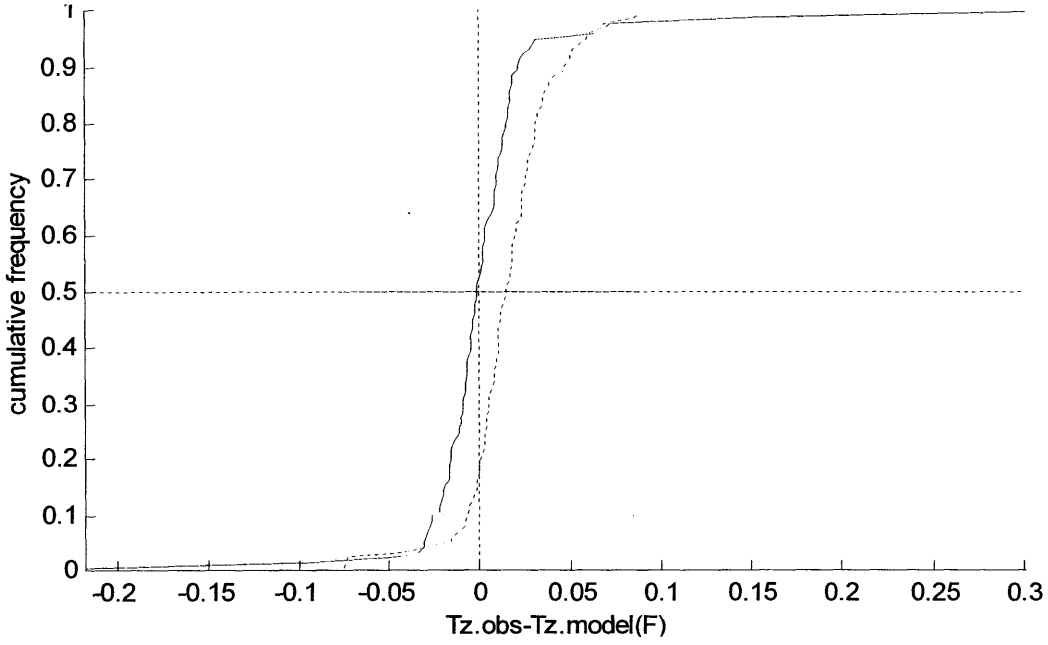

Figure 3.21. Cumulative distributions of prediction errors over training (solid) and test set (dashed).

3.4.5 Results Summary.

Three ways of estimating the parameters of a general model were assessed. The residual analysis indicates that the hybrid model identification procedure reduces bad behavior of inverted models identified with heat rate on the one hand, or zone temperature, on the other, as the response whose prediction error is to be minimized. The hybrid approach uses a linear combination of heat rate and zone temperature which results in a model that is a good compromise for both prediction jobs. The bias of predictions over the testing set was also somewhat smaller with the hybrid procedure. The

behavior of the three fitting procedures can be viewed in terms of the equivalent thermal parameters as well as in terms of observation-prediction residuals.

Residual norms. The deviation of predicted from observed heat rate with CRTF model identification is on the order of 10% of rms heat rates measured during training and testing periods. The norms are 186 Btuh for training and 445 Btuh for testing. When zone temperature is treated as the response the norms are 0.14°F for testing and 0.33°F for training. These error norms are quite reasonable but the large deviations seen in Figure 3.10 and Figure 3.12 during transients are unacceptable.

The deviation of predicted from observed zone temperature with iCRTF model identification is on the order of 5% of the rms temperature variations measured during training and testing periods. The norms are 0.02°F for training and 0.04°F for testing. When heat rate is simulated by the inverted model, the norms are unacceptably large: 6328 Btuh for testing and 11650 Btuh for training.

Hybrid identification gives the best of both worlds in terms of overall error norms (0.03°F and 346 Btuh for training; 0.05°F and 498 Btuh for testing) and deviations during transients (Figures 14-15).

UA. The zone load coefficients inferred from the three identification procedures are 15634 Btuh/°F for the CRTF, 8317 Btuh/°F for the iCRTF, and 4980 Btuh/°F for the hybrid formulation. The calculated UA is only about 2000 Btuh/°F. This is partly explained by the large return air opening that provides a bi-directional free-convection path when the fan is off.

Time constants. Two of the four time constants obtained by CRTF identification violate diffusion system constraints. A similar result is obtained with iCRTF identification. In both cases the resulting models are unstable. The hybrid identification returns time constants that are all positive real and, furthermore, obey the alternating magnitude constraint derived by Hittle and Bishop (1983).

3.5 ISD and ECC Test Results

The results obtained in the test zones are promising, but application to real buildings is essential to demonstrate the feasibility of model-based control and similar applications. Here we consider two test buildings, ISD and ECC, that were monitored under the CEC PIER project. The two buildings are of particular interest because night precooling tests performed in 2003 showed, at least qualitatively, that precooling is feasible in terms of existing thermal storage capacity. These buildings were not monitored continuously for long periods but fairly complete (electrical internal gains, heating and cooling capacities, weather data, and room temperatures) data sets of about one continuous month for each building were collected. The building footprints and orientations are indicated in the aerial view of Figure 3.22.

Figure 3.22. Aerial view of monitored buildings on the LA County Eastern Avenue Campus: ISD (left at 34.051N, 118,174W), ECC (top at 34.058N, 118.166W), and the building hosting a weather station (bottom at 34.049N, 118.170W). I-10 and I-710 intersect at the upper right and Eastern Ave. snakes north-south just west of ISD. The top edge of the photo faces north exactly.

Dimensions and aggregate areas of the ISD and ECC envelope elements are described in Appendix F. First-order models were obtained for both buildings. The ECC Building did not yield a very useful model because some of the important inputs—internal gains and incident solar radiation—could not be properly measured or estimated[20] and the room temperature sample was small; the statistical significance of several coefficients are marginal in the first-order model. In Chapter 4's demonstration of precooling potential, therefore, we will use ISD and Zubkova models, but no ECC model. Instrumentation used to obtain training data and model identification results are described below.

3.5.1 Instrumentation.

Two NILM installations were made in each building, one to track HVAC (chillers, fans, pumps), the other to track whole building (lighting, plugs, other non-HVAC) loads. Each NILM, except at the ECC service entrance, is shadowed by a K20-6 (traditional end-use metering) logger with one-minute integration intervals.

The K20-6 loggers serve to monitor the non-power variables needed to properly control load shedding activity and analyze the effectiveness and comfort impacts of such actions[21] as well as to verify NILM end-use disaggregation. One K20-6 logger can monitor 16 end-use circuits, 16 status or pulse channels, and 15 analog channels. The electrical end-use channels are documented in Appendices A and B.

Return air temperature and humidity are measured in each building to provide an estimate of average zone conditions. The return temperature is typically a biased estimate[22] of average zone temperature when zone air returns through ceiling plenums because most of the heat from lights is added to the air as or after it enters the ceiling plenum.

Room temperatures in selected zones (Table 3.5 and Table 3.6) are monitored by unobtrusive battery-powered microloggers. The loggers sample temperature at 2 Hz and the sample averages are recorded at 5-minute intervals. Data storage capacity is 24 days; data are retrieved as needed to characterize envelope response and assess test results.

Coil air-side temperature rise is measured in the ISD hot and cold decks and net air-handler temperature rise is measured in ECC. (Installation at the ECC chilled water coil was impractical because of the very large size (~12' x 36') of the coil face).

Air-side temperature differences, measured by a thermopile with multiple upstream and downstream junctions distributed over the projected coil face area, determine sensible cooling capacity . The thermopile puts out a very low level signal (about 50 uV per Kelvin per pair of junctions) that is amplified by an auto-zeroing op-amp configured for the appropriate gain (200 to 2000 depending on the number of junction pairs in the thermopile and the maximum temperature difference expected). The op-amp output is connected to a K20-6 analog input channel. (K20-6 input range is fixed at 0 to 5V).

---

[20] The east wing of ECC casts a time-varying shadow on the large east-facing window area and a parking garage casts a time-varying shadow on the large west-facing window area of the north wing. The parking garage and exterior lighting represent an unknown, but probably substantial, part of the measured electrical load

[21] The variables needed for control, including end-use loads output by a NILM, would normally be monitored by the control system. Prototype NILMs may be interfaced to control systems in one or more of the project sites after the effectiveness of load shedding has been demonstrated. However, for the initial tests control and monitoring of load shedding actions and responses will be implemented independently.

[22] The ideal sensor measures what the occupant perceives, T_{occ}, which is a weighted average from head to toe at whatever location the occupant chooses to occupy at a given time. A single fixed sensor can only provide an estimate, T_{sense}, of occupant-perceived temperature. In the most general linear model $T_{occ} = \mathbf{C_o x} + \mathbf{D_o u}$ and $T_{sense} = \mathbf{C_s x} + \mathbf{D_s u}$. These relations show that there is at least the possibility of obtaining a better estimate of T_{occ} from T_{sense} via a state observer.

Table 3.5. ISD zone temperature micro-logger locations:

Floor	Wing	Side	Dept.	Room	Serial No	Comment
0	S	Core			522639	Mid open office; lost after 2003.01.24
1	N	Core			332632	Mid open office
1	S	Core			495219	Mid open office, Ron's cubicle
2	N	E			332635	Private office
2	N	Core			495190	On T-stat outside west private office
2	S	W			495213	Private office, Susan Lopez
2	S	Core			522646	Mid open office
3	S	Core			522644	Mid open office, replaced 2002.06.26
					500862	

Table 3.6. ECC zone temperature micro-logger locations:

Floor	Wing	Side	Dept.	Room	Serial No	Comment
G	N	core		0101	332630	sheriff (reception and open office)
L	N	core			495205	large open office
2	N	W		2700	495199	large open office near phone closet T2N4
2	E	N			332268	large private office (Jo Schiff)
3	E	core	406		332633	court room converted to arts & crafts
3	N	core		3511	522461	
4	E	S		400C	522645	four-person office (Angela Smith)
4	N	core	417		522637	courtroom
5	E	core	421		522640	courtroom (Judge John L.Henning)
5	C	S		Hall	522638	on T-stat near copy machine & stairwell
5	N	core		424	332631	courtroom
6	N				522643	three-person office and reception (Lisa Romero/Ann Fragraso)

Boiler run times at each firing stage are monitored by installing relays in parallel with the boiler gas valves. The relays provide contact closure inputs to the K20-6 loggers that monitor thermal loads in each building. Note that boiler run time is the primary measure of heat supply in buildings with reheat coils (ECC) and is a redundant measure in buildings with central heating coils (ISD).

The temperature and relative humidity (RH) of return, mixed and supply air are monitored by RTDs and HyCal RH sensors in the ECC.

Weather is monitored at the Communications building. Outdoor temperature is measured by an RTD mounted in a radiation shield and solar radiation is measured by a LiCor pyranometer mounted on the wind cross-arm. Wind is measured by a 3-cup anemometer and an active (op-amp-based) rectifier to convert the AM signal to dc. Barometric pressure is measured by a Setra 207. Direct and diffuse components of solar radiation are determined by a shadowband pyranometer employing a LiCor model PY-200 detector and controlled by a Campbell Scientific CR10 logger.

3.5.2 Model Identification.

The parameters for each model are again determined by least-squares fit to the monitored conditions and thermal responses of each building. The data sets for these buildings are somewhat limited (3 weeks) and the zone temperature observations (2 loggers per floor) were not sufficient to obtain a very accurate estimate of mean zone temperature. The analysis for these buildings is therefore of limited value in validating the identification process and only basic parameters and error bands are reported here.

As in the case of the test zones, sol-air temperatures and time-shifted zone temperatures are converted to temperature differences by subtracting current zone temperature from the other (current and lagged) temperatures. This eliminates the current zone temperature term and reduces the order of the least squares problem by one, and results in a solution that satisfies the constraint on sums of temperature coefficients.

The model identified for ISD is in the form of (8) with one thermal capacitance and an hour timestep:

$$-\phi_{z,0}Q_z = B_1^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_0^n(\theta_z, T_z) \qquad (8)$$

With $\phi_{z,0}$ equal minus one, the least squares solution is

$$Q_{z,k} = Q_{z,k-1} + 194.4T_{z,k} - 198.1T_{z,k-1} + 0.2249T_{x,k} + 3.456T_{x,k-1}$$

where T is in °F and Q is in kW. The building load coefficient (UA) corresponding to these estimated model coefficients is 16367 Btuh/°F. The time constant for Q is 3.93h and the time constant for T_z is 53.0h. The normalized width of each coefficient's confidence interval (CI) is given in Table 3.7.

Table 3.7. ISD model coefficients and confidence intervals

Term	Coefficient	CI/|coef|
$Q_{,k-1}$	0.7751	212.5
$T_{x,k-0}$	-0.2249	0.745
$T_{x,k-1}$	-3.456	11.3
$T_{z,k-0}$	194.4	109.3
$T_{z,k-1}$	-198.1	109.3

The response produced by the model for ISD is compared to the measured response in Figure 3.23.

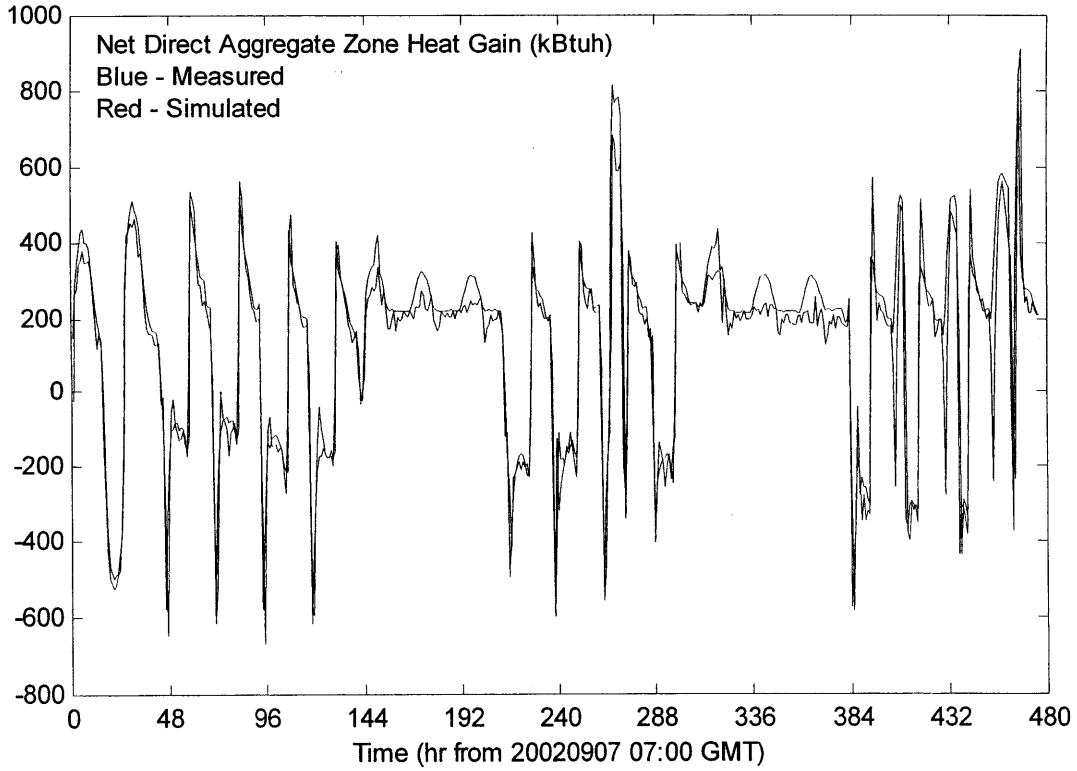

Figure 3.23. Measured and simulated trajectories of ISD net zone heat gain from HVAC air streams, solar radiation and electrical loads.

The model identified for the ECC is in the form of (7) with two thermal capacitances and hour time steps:

$$\theta_{z,0}T_z = B_0^n(\phi_z, Q_z) + B_0^n(\theta_w, T_w) - B_1^n(\theta_z, T_z)$$

With $\theta_{z,0}$ equal one, the least squares solution is

$$T_{z,k} = .8873T_{z,k-1} + .0991T_{z,k-2} + .2580T_{x,k} - .0323T_{x,k-1} + .0201T_{x,k-2} + .00160Q$$

where T is in °F and Q is in kW. The building load coefficient (UA) corresponding to these estimated model coefficients is 29004 Btuh/°F. This is about twice the UA of ISD, which is reasonable for a much larger, albeit better insulated, building. The normalized width of each coefficient's confidence interval (CI) is given in Table 3.8.

Table 3.8. ECC model coefficients and confidence intervals

Term	Coefficient	CI/|coef|
$T_{z,k-1}$	0.8873	10.8
$T_{z,k-2}$	0.0991	1.2
$T_{x,k-1}$	-0.0323	1.0
$T_{x,k-2}$	0.0201	1.0
Q	0.0160	3.4

The response produced by the model for ECC is compared to measured response in Figure 3.24. The magnitude of the root-mean-square of the residual, 0.33°F, is comparable to the observation uncertainty but it is visually obvious the the residual is not random. One of the weaknesses of least squares is that it does not penalize structure within the residual time series, only its 2-norm.

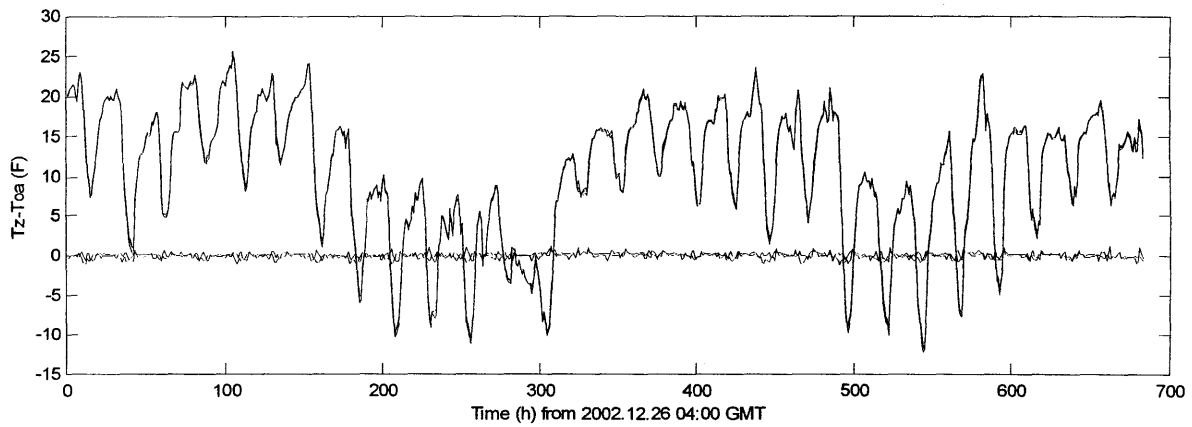

Figure 3.24. Measured and simulated trajectories of ECC indoor-outdoor temperature difference (note $T_{oa} \equiv T_x$).

3.5.3 Results Summary.

Attempts to identify physically plausible second-order models of ISD and ECC buildings failed. This failure can be at least partly attributed to the lack of a good spatial average zone temperature. The possibility that the system parameters change significantly with fan operation may also have a role. Finally, in the case of ECC, the diurnal variation in shading patterns. While the objective of identifying useful second- or third-order models was not achieved, the first-order models are reasonably good predictors of response.

3.6 Zubkova Test Results

A 60-unit, 5-story apartment building located at 22-1 Zubkova St., Ryazan, Russia serves as the fourth test case for model identification. This building was instrumented with 24 high-resolution temperature loggers carefully and uniformly distributed among the five floors and four cardinal exposures. All building utilities, including cooking gas, domestics hot water, cold service water, and waste water enthalpy flows, as well as space heating water and electricity were monitored. A weather station on the roof measured wind speed, solar radiation, and outdoor temperature. Thus we have a very complete and accurate set of data for model identification. Another reason for selecting this building is its construction type of precast panels for partitions as well as exterior walls and floor decks (see Appendix G). The concrete floor slabs of each intermediate story are relatively well coupled to occupied spaces both above and below. The resulting large surface-to-floor area ratio for directly exposed concrete slab sections of 10-15cm thickness is ideal for the performance of chiller curtailment strategies as well as the precooling strategies assessed in the next chapter. The south wall, which actually faces 17.5° west of south, and east gable end are pictured in Figure 3.25.

Figure 3.25 South facade of test building at 22-1 Zubkova, Ryazan (54.62N, 39.72E).

3.6.1 Test Conditions.

In Ryazan apartment buildings, the cold water enthalpy stream is typically larger in magnitude than the heat input by electric power. The reason for this surprising fact is that cold water, entering at 3 to 5°C in winter, is mainly used in the toilets and thus has a long residence time in the flush tanks. Waste water typically leaves at temperatures close to room temperature. The seasonal average values of all the heat inputs are tabulated in Table 3.9. The average, maximum and minimum values of the daily mean outdoor temperature time series for 4 December 1996 to 16 April 1997 were -4.9, 10.9, and -27.5°C. Horizontal solar radiation averaged 63.2 W/m2, and wind speed averaged 3.8 m/s (BNW 1997). Daily average values of zone temperature and utility enthalpy rates appear in Figure 3.26.

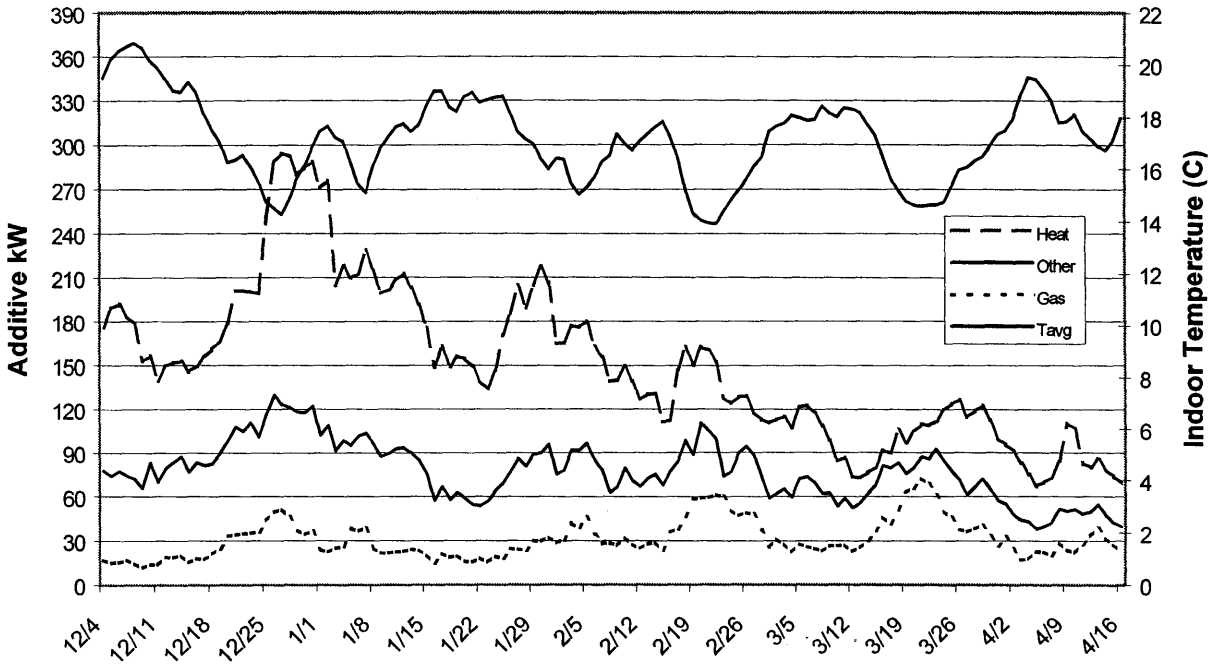

Figure 3.26. Average daily enthalpy rates for 221 Zubkova utilities from 1996.12.04 to 1997.04.16

Table 3.9. 1996.12.04 to 1997.04.16 average of zone temperature and utility thermal inputs

	Units	Zubkova 22-1
Indoor Temperature	°C	17.1
Gas Stoves	kW	31.34
Electrical	kW	9.93
Heating System	kW	71.88
Hot Tapwater	kW	33.78
Cold Tapwater	kW	-10.07
DHW Recirculation	kW	11.97
Net Thermal Power	kW	148.83

When district heat supply temperatures were below the values established by the district heating outdoor temperature reset schedule, the room temperatures dropped to as low as 14°C. Tenants coped as best they could by drawing less cold tap water, drawing more hot water, and using their kitchen stoves to stay warm. Gas use more than doubles during periods of heat deficit—that is, at times when indoor temperatures are much below 18°C.

3.6.2 Model Identification.

The parameters of 2nd- and 3rd-order models were determined by least squares. Results are presented in Table 3.10 for the 3rd-order model and in
Table **3.11** for the 2nd-order model. The Q time constants and UA parameters of the building inferred from the two models are similar. The T_z time constants for the 2nd-order model are more in line with our expectations than the corresponding time constants of the 3rd-order model and, furthermore, the interleaving constraint on Q,T_z time constants is satisfied by the 2nd-order model but not by the 3rd. The auxiliary terms are xx1=constant, xx2=solar on horizontal, and xx3=windspeed*dT product.

Table 3.10 Model coefficients and confidence intervals (CI) for 3rd-order model with 2-hour timestep

2 hour time step; merge221c(,,,2,2)			
Term	coeff	CI	|coeff|/CI
Q1	1.0413	0.054	19.2764
Q2	-0.2657	0.0733	3.6228
Q3	0.004	0.0507	0.0791
Tz0	-50.8952	8.5922	5.9234
Tz1	69.3284	2.8871	5.3797
Tz2	-7.6737	3.1546	0.5833
Tz3	-12.2181	8.4002	1.4545
Tx1	-0.2401	1.0225	0.2348
Tx2	3.8853	1.0091	3.8504
Tx3	-3.2555	0.5602	5.8112
Xx1	-3.355	3.1904	1.0516
Xx2	0.016	0.0122	1.3175
Xx3	-0.0016	0.0191	0.0828
	----------residual trend f(k_t)----------		
	coeff	CI	|coeff|/CI
T	20.1908	14.2465	1.4173
t^2	-64.3678	44.1274	1.4587
t^3	44.8002	32.6226	1.3733
	---------time constants (tau/Δt)-------		
Tz tau/dt:	8.386268	4.716192	0.098969
Q tau/dt:	2.158929	1.07939	0.242199
UA(w)=6.619817			
Se(Q,Tz)= 16.978106 8.592161			

Table 3.11. Model coefficients and confidence intervals for 2nd-order model with 2-hour timestep

Term	Coeff	CI	|coeff|/CI
Q1	1.0676	0.0553	19.3107
Q2	-0.2773	0.0487	5.6914
Tz0	-48.4847	8.6212	5.6239
Tz1	72.474	12.7778	5.6719
Tz2	-25.4102	8.4852	2.9947
Tx1	2.5783	0.9387	2.7467
Tx2	-0.9994	0.5864	1.7044
Xx1	-1.2471	3.2696	0.3814
Xx2	-0.0059	0.0116	0.5102
Xx3	-0.0011	0.0199	0.0529
	----------residual trend f(k_t)----------		
	coeff	CI	|coeff|/CI
T	26.2203	14.8004	1.7716
T^2	-78.7943	45.8432	1.7188
T^3	51.828	33.8911	1.5293
	---------time constants (tau/Δt)-------		
Tz tau/dt:	14.4539	1.733363	
Q tau/dt:	2.100208	1.239904	
UA(w)=6.774821			
Se(Q,Tz)= 17.671588 8.621222			

The tabulated results nevertheless show the distinct possibility of a meaningful 3rd order model. The auxiliary heat rate terms, when statistically significant, satisfy physical expectations and the first two (xx1 and xx2) are at least statistically marginal. To further explore the 3rd-order model, one might try fitting 1-hour timestep data with a model that has assumed time constants of 16.37, 9.43, 1.73 for T and 4.32, 2.16, 1.24 for Q

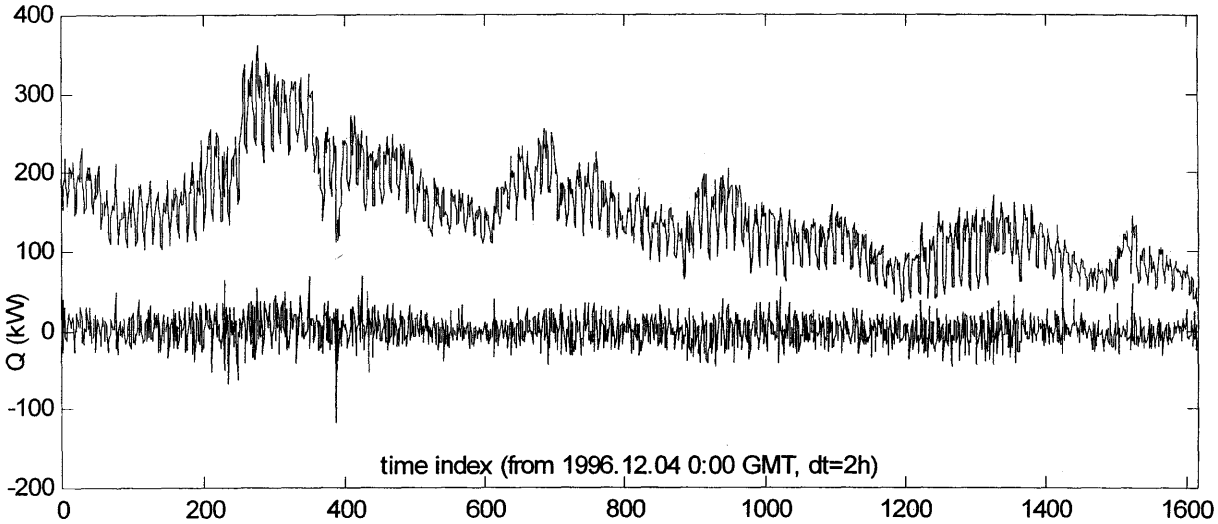

Figure 3.27. Observed (dots) and predicted (solid) net enthalpy rate of utilities plus direct solar gains. Residual is the solid trace centered on zero.

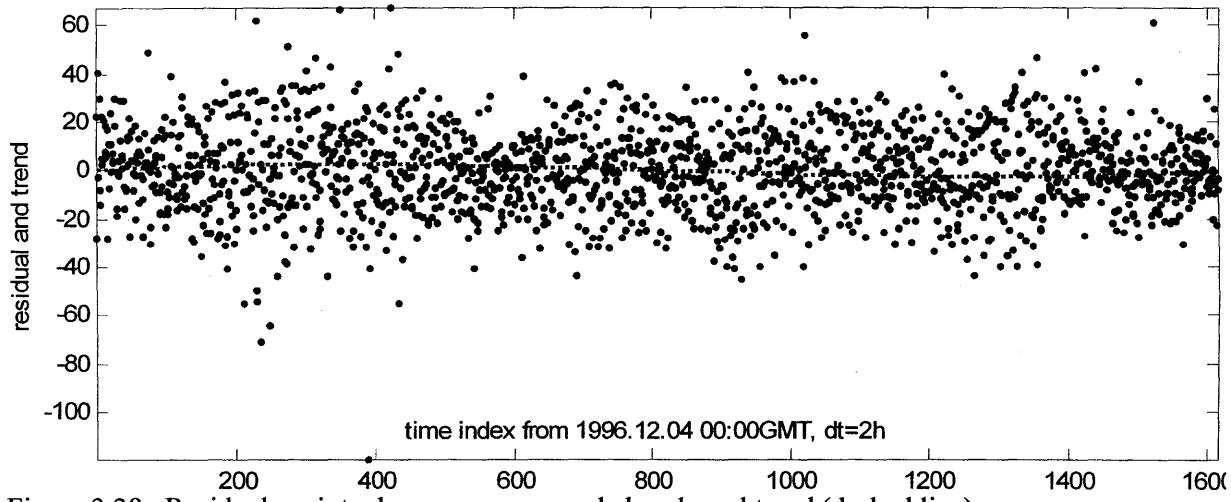

Figure 3.28. Residuals points shown on an expanded scale and trend (dashed line).

The trend in the residuals is small but significant and probably results from the model's use of total solar radiation on the horizontal. The ratios of 1) radiation incident on vertical surfaces and 2) direct gains through windows to radiation incident on the horizontal vary systematically and substantially over the diurnal cycle and through the monitored part (December to April) of the heating season.

The distribution of residuals (observations - predictions) is shown in the cumulative frequency plot, Figure 3.29, where the deviation from a normal distribution (dotted) is seen to be small except in the tails, Figure 3.30. One of the extreme values coincides with interruption of district heat service on 5

January 1997; others may be bad data, non-linear physics, or other model error. The autocorrelation, Figure 3.31, shows a strong diurnal peak. This can be partly (perhaps mostly) attributed to the model using total solar radiation on the horizontal for both direct gain (windows) and sol-air temperature (exterior envelope) terms. The manual operation of windows (exhaust rates were monitored but data was collected from the loggers only sporadically) and treatment of cold service water heating load as instantaneous may also contribute to the problem.

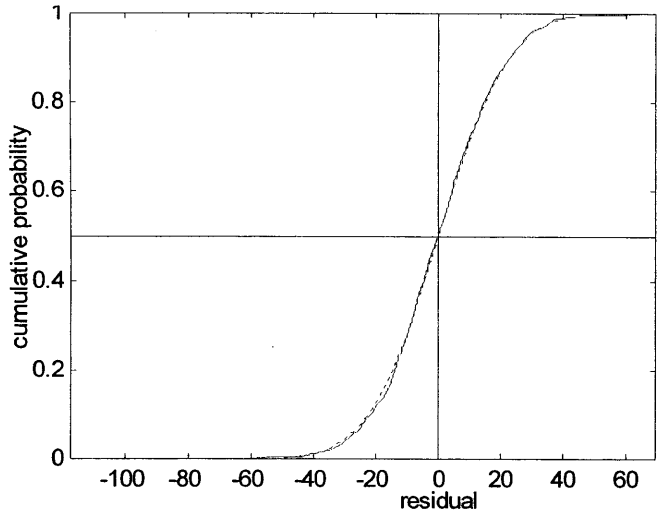

Figure 3.29. An almost perfect match of residuals (solid) with the normal distribution (dash)

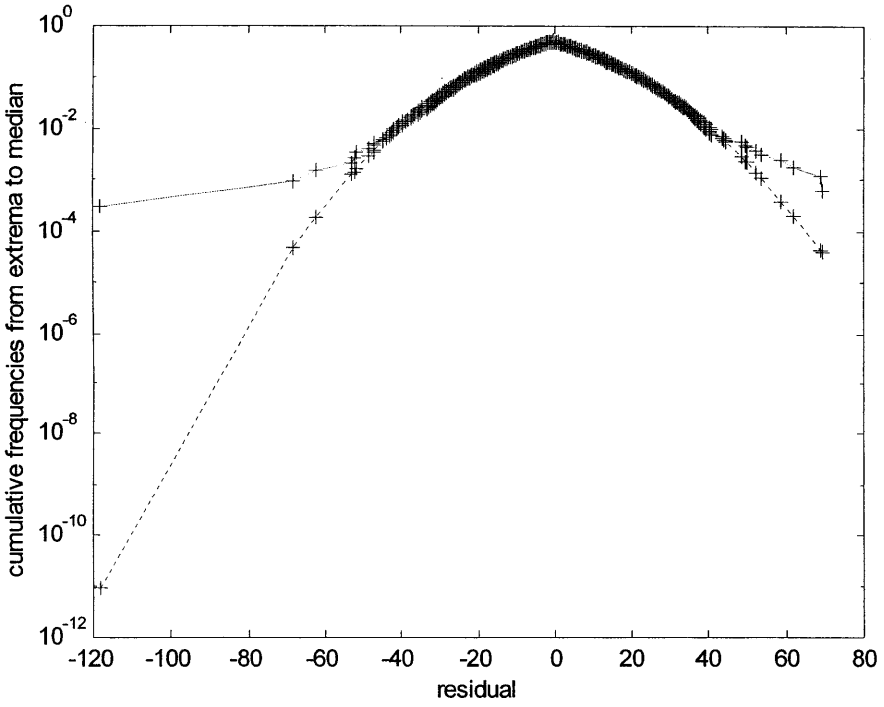

Figure 3.30. Semi-log plot showing large deviations from the fitted normal distribution for the ~15 of the 1620 residuals; these are, of course, the residuals of greatest magnitude.

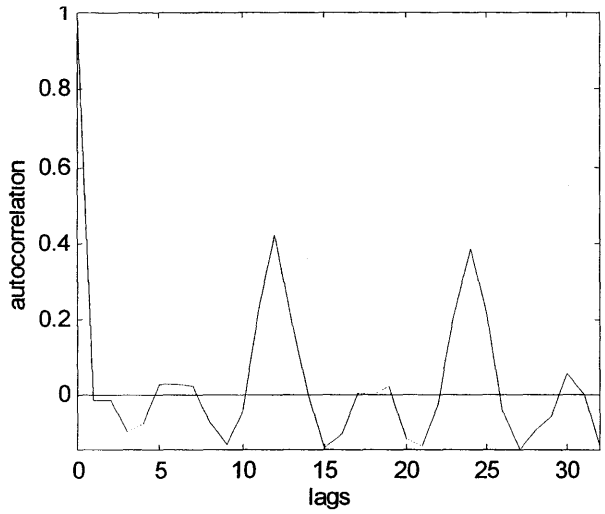

Figure 3.31. Excessive correlation of residuals separated by 24 hours (12 lags)

3.6.3 Results Summary

The Zubkova data set consistently produces physically valid 2nd-order models with only the sum-of-T-coefficients constraint imposed. The distributions of residuals are Gaussian except for about 1% which may be considered outliers. The UA inferred by the models ranges within 10% of the calculated UA and the value does not change much with model order. The inferred dominant time constants are shorter than the calculated value by factors of 4 to 8 but the interleaving of zone temperature and heat rate time constants is generally correct. The autocorrelation indicates a diurnal cycle in the residuals. This may be attributed to the use of solar radiation on the horizontal as an input when in fact the system is driven mainly by beam and diffuse radiation incident on the north and south walls. Other data limitations are the lack of an albedo measurement and lack of kitchen temperature measurements. Kitchen temperatures are important because the kitchen generally gets warmer than other rooms during cooking and cooking generally occurs on a diurnal schedule. Other occupant behaviors, such as window opening, may also be significant.

3.7 Summary

A general building thermal response model and three thermodynamic constraints have been identified. The constraints are crucial to on-line identification for the model-based control application because the observations used for model identification are inevitably noisy and often incomplete or biased as well. Without the constraints, a physically implausible model that behaves badly under certain conditions (e.g., fast transients or temperature extremes) will often be accepted as the "best" model implied by the data.

Formulations were developed for enforcing the constraints using the simple optimization tools of ordinary least squares and bounded-search nonlinear least squares. The steady-state conductance constraint is handled algebraically. This approach eliminates one unknown but, more importantly, has the invaluable advantage of leaving the problem in unconstrained ordinary least squares form. The constraint on time constants is a non-linear constraint. To avoid the difficulties traditionally associated with non-linear constraints, a way was found to present the problem as a nonlinear least squares problem. A further modification was devised to enforce the constraint of alternating roots by imposing simple bounds on the transformed unknowns. However there are still some unresolved numerical problems that prevented application of the time constant constraints to our test buildings.

The problem of identifying coefficients for predicting both (Q and T_z) responses of interest using the CRTF form for heat load, Q, and its inverse, the ICRTF, for zone temperature, T_z, has been addressed. The "errors in variables" (aka *total least squares*) approach was tried with poor results. An alternative approach was developed by defining a new response that is the sum of the zero lag terms for Q and T_z. This approach has the advantage of great simplicity but the disadvantage of making the estimation problem nonlinear in one unknown, the ratio of Q and T_z zero-lag coefficients. In the case of the test room, which exhibited bad behavior when its conventionally identified models were inverted, the new procedure gave very satisfactory results.

Second-order models were identified for the test cases that had spatially averaged zone temperature observations. For the two LA buildings where the number of zone temperature loggers was relatively small for the magnitude of expected spatial variations it was possible to identify only first-order models. The massive building came close to giving a third-order model and it is likely that better observations of excitations, particularly the solar gains, would result in reliable identification of valid third-order models.

There are only a few really useful evaluations of reduced-order models in the literature (Horn 1982, Seem 1989), and they do not fully answer the question of the minimum model order needed for model-based control and other applications. The question should probably be more thoroughly explored using simulation before further effort is expended on real data.

Chapter 4 Optimal Use of Building Thermal Capacitance

This chapter puts the general transient thermal response model of Chapter 3 to use as the kernel of an on-line controller. Several strategies are used to shift peak cooling loads by reducing zone temperatures before the time of peak load when outside air and lower electric rates can be used to advantage.

4.1 Control Function Descriptions

The three main control functions that require a plant model are precooling and curtailment (both forms of peak shifting), optimal start, and set-point control under large disturbances (model-based control). Precooling is the most challenging of these with its variety of possible implementations involving various approaches to optimal or heuristic control as well as variations predicated by the cooling air distribution system (CV, VAV, dual deck, etc.) Here we describe some important details of the more important variations and define useful nomenclature along the way.

4.1.1 . Curtailment and Peak-shifting Strategies

Peak electrical demand may be reduced by

1) reducing non-HVAC loads (lighting and plug loads), and/or
2) reducing HVAC cooling capacity (fans, chillers, pumps and cooling tower).

Reduction of base load is the first measure that should be considered because energy savings accrue for all hours of operation, are not subject to occupant, operator, or controller behaviors, and generally result in additional cooling plant and distribution system load reductions. Lighting and office equipment are significant base loads in many buildings. Improvements in heating and cooling plant and air distribution efficiency are also classified as base-load reduction measures.

Curtailment. Short-term reductions in plug and lighting loads are generally the second most desirable measure because they produce additional cooling capacity reductions without occupant discomfort. The bonus, of course, is that curtailment of lighting and plug loads yields an immediate reduction in cooling load which the control system will sense and automatically respond to by effecting just the capacity reduction needed to maintain room conditions near setpoint. Dispatchable loads may include such equipment as copiers and printers, two-level lighting, general lighting (if task lighting or centrally controlled dimmable ballasts (LBNL 2003) are available) or even task lighting (if sufficient daylighting is available), and may be extended to workstations of users who are able to indulge in non-computer-related work without undue disruption of their productivity. Many of these additional reductions would be controlled manually. Occupants could be notified by email, phone broadcast, or over a public address system. Some advance training and periodic "fire drills" will be necessary in most situations to implement this sort of curtailment strategy effectively. An ability to predict and confirm responses to such occupant-implemented load curtailment is key to its effectiveness.

A more aggressive curtailment measure that can be implemented involves further reduction in cooling capacity, either by direct control or by raising room temperature (and humidity, if implemented) set points. The setpoint changes can be abrupt or gradual, depending on the curtailment response desired. The occupants will experience loss of comfort in either case, and the amount of additional

capacity reduction, and its duration, will be limited by occupants' tolerances for elevated temperature and humidity levels.

The last, most difficult, but potentially most efficient and least disruptive, curtailment measure requires control functions that anticipate curtailment by anywhere from one to sixteen hours and increase cooling capacity modestly during this pre-curtailment period so that zone conditions are as cold and dry as can be tolerated immediately before the curtailment period and the thermal masses of building structure and contents are at or even below this minimum tolerable temperature. With such a favorable initial state, the reduction in cooling capacity and its duration can be made significantly larger—i.e., the amount of cooling capacity that can be shed during the curtailment event is significantly increased.

There are at least three distinct schemes for implementing precooling as a retrofit:

1) Short period of reduced setpoint before the anticipated load shedding time.
2) Extended period of reduced setpoint (starting the day with a lower setpoint or ramping it down gradually during the pre-curtailment period)
3) Night cooling (outside air with or without chiller operation)

The last two schemes are generally referred to as peak-shifting rather than curtailment.

Peak Shifting. Curtailment is a short-term mechanism to prevent a utility's aggregate load from exceeding its available generation capacity. Curtailment programs usually provide no direct incentive for efficient operation. Utilities have used (extensively in the last two decades) demand- and time-of-use- (TOU) based rate structures, on the other hand, to encourage longer-term reduction of peak loads[23]. Peak-shifting is a customer response that typically benefits both the utility and the customer under a time-of-use- or simple demand-based rate structure[24]. The incentive to improve energy efficiency, although often ignored by cream-skimming measures, does exist under TOU rates.

Night cooling is a peak shifting strategy that is well suited to office buildings and similar air-conditioned buildings in climates where cool nights and significant daytime cooling loads coincide through much of the year. Night cooling can save energy by using ambient conditions (cool night air) to remove heat from the building structure and contents or by using forms of mechanical cooling that are particularly efficient under night cooling (low lift, small load) conditions. The maximum usable cooling capacity stored on a given day is proportional to the normal zone cooling setpoint minus the thermal capacitance weighted average of contents and structure temperatures at the start of occupancy. The actual capacity is roughly (but not exactly, because the envelope temperature field is affected by outdoor, as well as indoor, conditions) proportional to the change in the thermal-capacitance-weighted average of contents and structure temperatures between start and end of occupancy.

---

[23] for which a utility typically incurs higher generation cost, T&D loss, and risk of service interruption.

[24] "Real-time" pricing with 24 hour advance rate notification is another effective rate structure that has been tested, and appears to have an assured future, in electric utility markets.

Stored capacity is limited in four ways: total thermal capacitance, rate of charge/discharge, night temperature, and comfort constraints. Within the framework of these basic limitations, however, there are a number of control strategies that can be used to implement night cooling.

Constant volume(CV) precooling. The simplest strategy is to cool the building at night whenever outdoor temperature is below return air temperature and to modulate the cooling rate to just maintain the minimum comfortable temperature (MCT) until occupancy. On warm nights there may be no cooling, on mild nights the minimum comfortable temperature may not be reached.

This strategy is suboptimal for two reasons: no part of the mass can ever be cooled below the MCT and the cost of fan energy has not been considered.

CV delayed start. The simple strategy can be improved by running the fan fewer hours at night. The objective in this strategy is to reach the MCT not too long before occupancy and, on days with little cooling load, to not reach MCT. Implementation is relatively simple because there is only one variable, start time, to be determined each night.

CV subcooling/tempering. Delayed start is still not optimal because the potential to cool part of the mass below MCT is not exploited. A third CV strategy, therefore, add another variable, the tempering time. With this strategy the optimal night cooling start and stop times are determined each day.

Constant volume objective function. The foregoing strategies have been described in terms of mass state at start of occupancy. However this state is not observable and, in any case, is only indirectly related to the objectives of minimizing plant demand and energy costs. An objective function that represents daily electricity cost will therefore be used. For the most elaborate CV strategy, daily cost can be minimized by enumerating all possible start and tempering times and comparing total daily cost for each[25].

Variable volume precooling. The fan energy required on any given day can be further reduced by modulating fan speed during the night cooling phase. This is a much more difficult optimization problem because any sense in which daily cost is monotonic with fan speed in a given hour is lost. Two approaches will be used to make the problem tractable: discretization of fan speed and application of a fan modulating function in which gain is determined daily by optimizations analogous to the CV-delayed-start and CV-tempering strategies.

4.1.2 . Recovery/Optimal Start

Temperature recovery refers to the restoration of comfort conditions just prior to the start of each occupied period. In heating mode, maximum energy savings are usually achieved by starting recovery at the last possible moment. Controls that approach or accomplish this are often referred to as *optimal start.*

---

[25] In practice, a one hour time step is convenient and reasonable in terms of control resolution. Also, complete enumeration is found to be unnecessary. One starts with the earliest start time and shortest (zero) tempering time. Tempering time is increased until daily cost stops decreasing. The process is repeated with progressively later start times until the daily cost stops decreasing with lateness of start time.

Equipment (plant, distribution or terminal unit) capacity, envelope thermal characteristics, and conditions (initial zone and outside temperatures, internal gains, daytime setpoint) determine how long it will take for controlled zones to recover from night setback. It is generally desirable, for energy efficiency reasons, to set zone temperatures back during unoccupied periods. Yet, because first-cost and part-load-efficiency penalize over-sizing, equipment specified on the basis of steady-state design load may not have sufficient excess capacity to effect timely recovery of zone temperature on the coldest or hottest mornings of the year. Additionally, the potential for heating and cooling recovery problems has been compounded in recent decades as buildings have become tighter, better insulated, and better designed with respect to undesirable solar gains.

The engineer must reconcile these conflicting design and operation objectives by providing sufficient excess capacity under design load conditions to recover from a given setback temperature within a reasonable recovery time.

The dimensionless response of room temperature to a step change in zone heat input rate, Q, is defined as:

$$\Psi(t) = \frac{T_r(t) - T_r(0)}{T_r(\infty) - T_r(0)}, \text{ where}$$

$T_r(t)$ = room temperature at t, time from start of recovery,
$T_r(0)$ = initial room temperature at $t = 0$, and
$T_r(\infty)$= limit to which room temperature tends as $t \rightarrow \infty$ under given boundary conditions.

Recovery time is the time, t, that satisfies
$T(t)=T_d$ (day setpoint) when
$T(0)=T_n$ (night setpoint).
The design load is given by
$Q_{des} = UA(T_n - T_{des})$, where
T_{des} is the design ambient temperature, and the limiting temperature is
$T_r(\infty) = T_{des} + Q_p/(UA)$, where
Q_p is the total plant capacity, and
$Q_p - Q_{des}$ (excess plant capacity) is the step excitation that results in the response $\Psi(t)$.

The room temperature trajectory during recovery, $T(t)$, depends on conditions (initial room temperature, outside temperature, heating capacity and internal gains). The dimensionless response, $\Psi(t)$, on the other hand, is determined completely and uniquely by the thermal parameters of the building envelope and its contents (Armstrong 1992). Dimensionless response has a form consistent with (8):

$$\psi(t) = 1 - (\gamma_1 exp(\lambda_1 t) + \gamma_2 exp(\lambda_2 t) + \cdots + \gamma_n exp(\lambda_n t))$$

where λ_i are the eigenvalues (13-16) of the system or a reduced-order approximation of the system.

The step response is a reasonable basis (always conservative, albeit moderately so in most practical cases) for sizing plant and distribution equipment. However, it is too conservative, in most cases, as a basis for optimal start control (Seem 1989). An on-line model is ideal because it safely allows the latest possible plant start times on mornings when temperature recovery is needed.

4.1.3 Model-Based Control

Room temperature is usually controlled by a thermostat and terminal unit which we may represent together as a compensator. However the room temperature is strongly influenced by highly variable conditions—weather, adjacent zone temperatures, solar and internal gains—which are generally treated as disturbances. A long-ago recognized way to improve control is to measure the important disturbances and estimate their effect on the response. One such arrangement is shown in Figure 4.1. The level of potential improvement depends on the relative importance of the measurable disturbances. But the effectiveness of this approach is also directly related to the accuracy of the plant model which, in many cases, must adapt to changing plant parameters.

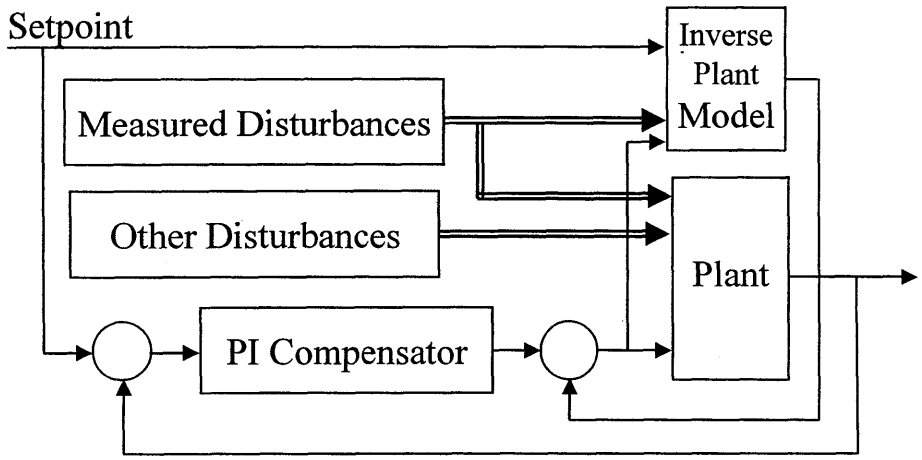

Figure 4.1 A typical scheme for model-based control; the plant model is inverted in the sense that the setpoint is input while the control signal is output. For room temperature control, the control signal is heating or cooling capacity and the setpoint is room temperature. The left hand summing node returns the deviation of room temperature from the desired setpoint and the right hand summing node returns the open loop capacity estimated by the plant model plus a correction signal comprising terms proportional to the deviation from setpoint and its time integral.

Variations of model-based control may include additional features such a predictive control. Ways may be provided to deal with loss of sensor(s) or loss of internet weather forecast inputs. Also note that the traditional compensator can be tuned for model-based control or stand-alone control. (In the latter case, near-optimal compensator gains are obtained with knowledge of the plant model.) Both sets of gains can be stored so that reasonable control can be maintained if the plant model loses key sensor inputs.

4.1.4 Commissioning and FDD

The purpose of commissioning is to find and correct faults and to properly adjust various flow component and control loop settings[26]. The close relationship of commissioning to FDD has two main facets:

[26]documentation and confirmation of control I/O points, confirmation of HVAC functionality by component, and calibration of individual sensors and actuators comprise the other main tasks, albeit usually straightforward, of a commissioning process.

1) fault detection tasks that are feasible *without a priori knowledge of as-built behaviors*, can and should be automated and made part of the commissioning process;
2) properly identified models of HVAC and other building subsystems[27] are essential to serve as a baseline for continuous fault detection and continuous commissioning. It is most reasonable to identify such baseline models during the commissioning process when there is a high confidence that each subsystem modeled is operating correctly.

In this chapter the precooling application will be explored.

4.2 Previous Work

The analytical framework for dealing with different climates, rate structures and occupancies is well established (Braun 1990, Kintner-Meyer 1995, Morris1994, Norford 1985, Rabl 1991, Xing 2004). Reliable one-day-ahead weather forecasts can be downloaded daily or hourly. Internal gains, predominantly light and plug loads, can be disaggregated by a NILM from the building total (Leeb 1992, Laughman 2003) and future loads can be similarly extrapolated from the accumulated history (Seem and Braun 1991, Forrester 1984).

However, thermal response to weather, internal loads and HVAC inputs is unique to each building and difficult to characterize empirically (Braun and Chaturvedi 2002, Norford 1985, Pryor 1982, Rabl and Norford 1991). The development of a general model and model identification procedure sufficiently robust to be automated is therefore a central and challenging prerequisite to the implementation of useful curtailment or peak-shifting (Kintner-Meyer 2003, Stoecker 1981), precooling (Keeney 1997), and optimal start (Seem 1989, Armstrong 1992) control strategies. Verification of curtailment for allocating incentives would, ideally, be based on the same model that is used for control (Goldman 2002).

The literature dealing with optimal control of discrete storage (e.g. a tank of ice or cold water) is dominated by the very thorough work of Henze and Krarti (1995, 1997) and Kintner-Meyer (1994, 1995). Of particular interest to both discrete storage control and building mass control is an assessment of the impact of forecast uncertainties (Henze 1999).

The literature dealing with precooling of building mass is richer and more varied in part because of the uncertainties in building thermal response and in part because of the severe capacity and storage limitation constraints (Andresen 1992, Brandemuehl 1990, Braun 1990, Braun and Chaturvedi 2002, Coniff 1991, Eto 1984, Keeney 1996 and 1997, Morris 1994, Rabl 1991, Ruud 1990). Night precooling studies have used simplified analyses, as well as simulation, to show annual savings in mechanical cooling energy input of up to 50%. Unfortunately there has been only one successful demonstration of significant precooling benefits in a real building (Keeney 1997). Control of night precooling involves the same building-specific thermal response model that is used for curtailment.

Reducing peak cooling loads, the usual basis for curtailment, is addressed in a number of papers Questions of incentives and dispatch have dominated the discussion; building transient thermal response has not been handled very rigorously except in the simulation studies (Haves 2001, Goldman 2002, Kintner-Meyer 2003, Stoecker 1981).

[27]Possibly including whole-building performance models

Optimal start requires some kind of thermal response model. HVAC controls have traditionally used very simplistic models (Hittle) More reliable control can be obtained with more realistic models but the economic benefit with respect to the best simplified models appears not to be very great (Seem 1989, Armstrong 1992).

4.3 Experimental Demonstration of Peak Shifting Potential

The transient thermal responses of the LA test buildings (ECC and ISD) to manually introduced control changes were observed in a series of field tests. Field tests included summer afternoon load shedding (curtailment) sequences and summer and winter pre-cooling sequences in both buildings. The test protocols and observed responses are documented in this section.

4.3.1 Curtailment Test Protocol

A summer load shedding protocol proposed by Haves (2001) provides a number of load shedding scenarios. All scenarios require heat to be completely shut off to eliminate the inadvertent heat gains (poorly insulated pipes, stuck valves, etc.) that inevitably exist in real world distribution systems. The ideal method of cooling plant modulation is to raise all zone setpoints gradually by the same amount. The practical approximation given by Haves is to raise all setpoints abruptly by the same amount. This will cause chillers to shut down for 30-60 minutes in most buildings. Since the L.A. test building control systems do not provide a way to raise zone setpoints simultaneously, we simply shut down the chillers. This has the same effect up until the time when one or more zones reaches its new setpoint. An alternative approach for ISD is to shut down one or more compressors or, for ECC, to reduce chiller capacity by modulating the inlet vanes. Such capacity modulation actions let us trade off load reduction against duration.

The most extreme action given by the Haves protocol is by no means the most extreme electric load shedding scenario possible nor does it represent the most extreme thermal excitation possible. Rather, the protocol aims to provide significant load reduction with modest comfort impact over short (1- to 4-hour) time frames. In the case of ECC, the test was implemented by disabling HVAC equipment in the sequence indicated below:

1) supply fan speed frozen at current setting (32Hz)
2) return fan speed frozen at current setting (25Hz)
3) hot water pumps to OFF
4) chilled water pumps to OFF
5) chiller and tower pumps and fans turn off automatically after step 4.

ISD cooling load was shed, as at ECC, by turning off hot and chilled water pumps. The chillers and cooling towers go down automatically upon loss of chilled water flow.

A log of load shedding activities is correlated with NILM measurements to quantitatively assess the impacts in terms of both load reduction and loss of comfort for occupants.

The NILMs monitor electric load changes effected by occupants and plant operators during curtailment events and the K20-6s and microloggers monitor weather and zone comfort conditions as well as coil loads.

Data obtained during load shedding are analyzed to determine the following:

Reductions, from baseload, achieved by curtailment of lighting and plug loads;

Reduction, from model-predicted HVAC load, achieved by the combination of lighting and plug load curtailment, and

further reduction in cooling capacity;

Disaggregation of HVAC load reduction by the two above-mentioned actions;

4.3.2 Night Precooling Test Protocol.

A constant volume precooling strategy was tested in ISD and a variable volume strategiy was tested in ECC. Since model-based optimal control could not be implemented on the existing control systems the tests were performed manually by an MIT research assistant and site contact as operators.

For ISD, the *CV Subcooling/Tempering* strategy was implemented in the conventional way, with both supply and return fans operating, and in a low fan energy mode with windows opened and only the return fan operating.

For ECC, the *Variable Volume* cooling strategy is not practical as a manually implemented strategy. A suboptimal variant was tested instead in which the fans are operated at a reduced but fixed static pressure and the terminal boxes are allowed to modulate air flow to each zone.

In both tests, the economizer dampers were held open manually until outside air exceeded return air enthalpy on the constant comfort line and the chillers were held off until average or least comfortable zone conditions reached a corresponding predetermined comfort limit.

4.3.3 ECC Load-Shedding Test Results

Load curtailment tests followed the main elements of the Haves protocol, described earlier. The 25 June 2002 test was effected by simply turning off the chilled-water and hot-water pumps at the motor control panel (MCP). The chiller (immediately) and cooling tower (within one minute) then shut down automatically. The averages of pre-curtailment temperatures shown in Figure 4.2 were 72.05˚F for zones and 72.16˚F for return air; both were very steady (standard deviation <0.1˚F) over the 100 minutes preceding the test.

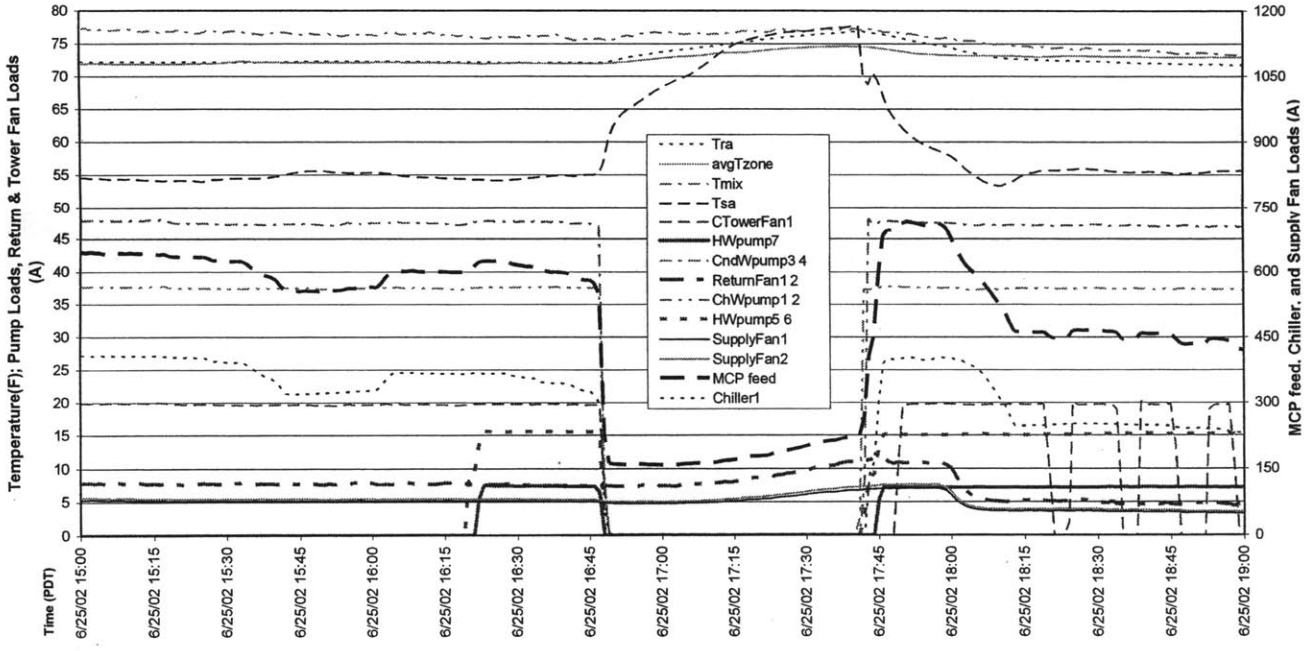

Figure 4.2. Temperatures and HVAC motor loads during ECC load shedding test.

It is not surprising that the return air is warmer than the average temperature across zones because return air is taken from the ceiling level and usually picks up some heat from ceiling lights as it leaves a room. Both temperatures begin to rise almost from the instant—indicated in Figure 4.2 by the 400A (330kVA) drop in MCP load at 16:48—that the plant was shut down. Both rise along similarly shaped trajectories after the chiller stops but the return air temperature rises more rapidly. This, too, is not surprising because the zone temperature sensors were placed on file or desktop cabinets or on wall thermostats where the proximate slower responding surface temperatures affect the measurements by radiant coupling. The zone sensors are indicating something close to "operative temperature" defined [ASHRAE 2001] as the average of air temperature and mean radiant temperature (MRT). The zone average and return air temperatures after 53 minutes without cooling were 74.65 and 76.89°F. When the chiller is turned back on, the return air temperature again responds more quickly than the zone sensors and we see that it approaches within 0.1°F of the pre-test temperature 50 minutes later. At this point the return temperature is 0.8°F below the sluggishly responding average zone temperature. The return air temperature drops below the pre-test value (overshoots in the direction of initial response) after the chiller is restarted--a direct result of slow zone thermostat response. Fan power rises gradually during curtailment as air flow increases to satisfy terminal units demand for more cooling and static pressure is maintained. The pressure setpoint could be reduced to prevent this.

Zone temperatures (transient behavior and dispersion among zones) observed during the June 2002 site visit were analyzed to assess the consistency of zone thermal conditions. The zone temperatures, their average, and the standard deviation across zones, are plotted in Figure 4.3 and Figure 4.4. Note that loggers resided together in a bag prior to deployment (2002.06.21 15:00-16:30 PDT), thus the data taken prior to deployment is only useful for showing that the loggers track each other quite well.

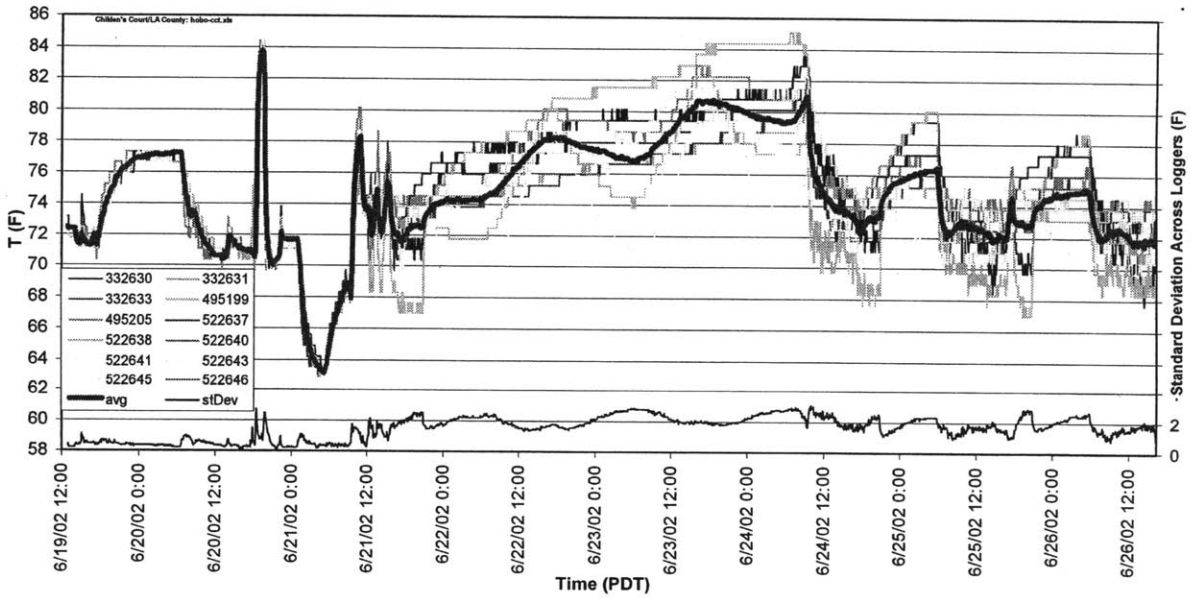

Figure 4.3. Children's Court micro-logger data from launch time (2002.06.19 13:00 PDT) through first download/relaunch (2002.06.26 16:00-16:45 PDT).

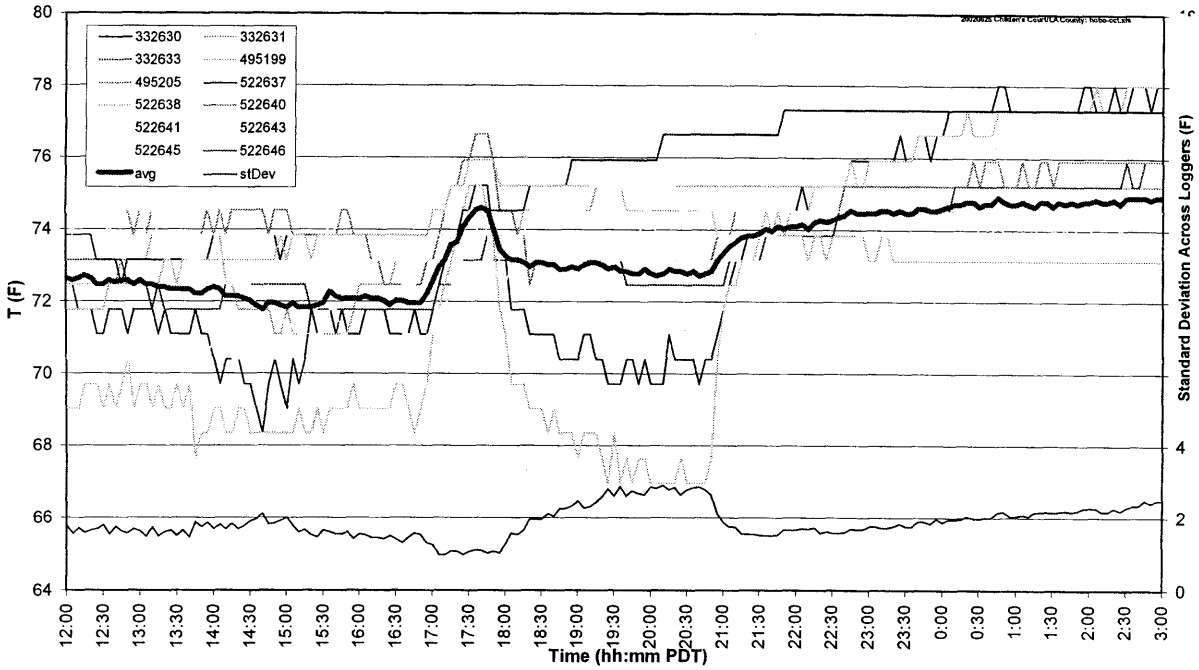

Figure 4.4. Children's Court zones before, during, after the chiller OFF step test of 25 June

4.3.4 ISD Load-Shedding Test Results

ISD cooling equipment was shut down, as at ECC, by turning off hot and chilled water pumps. The chiller and cooling tower shut down automatically. Prior to shut down, return air and average zone temperatures were 74.8 and 72.0 ±0.1°F, as shown in Figure 4.5. The higher return-air temperature may be partly caused by air leaking from the hot deck. Also note in Figure 4.5 that the return air temperature is less responsive to the step change than the average zone temperature. Supply duct leaks alone could only account for this if they were on the order of 50% of supplied air; thermal coupling to the underside of the floor decks also contributes to attenuation of rapid changes in return air temperature. Over half the total building load (600 – 350 = 250kVA) was cut during the test. A further 12% was cut when the fans were shut off at 5:30 pm. The average zone temperature rose 5.5°F in the first 40 minutes (fans on) and 0.5°F in the next 40 minutes (fans off). The zone temperature responded quickly after the test, in part because the hot deck was not restored until much later; return air temperature responds slowly. It is not clear why, with the large zone temperature rise, the chillers did not return to high capacity at the end of the test.

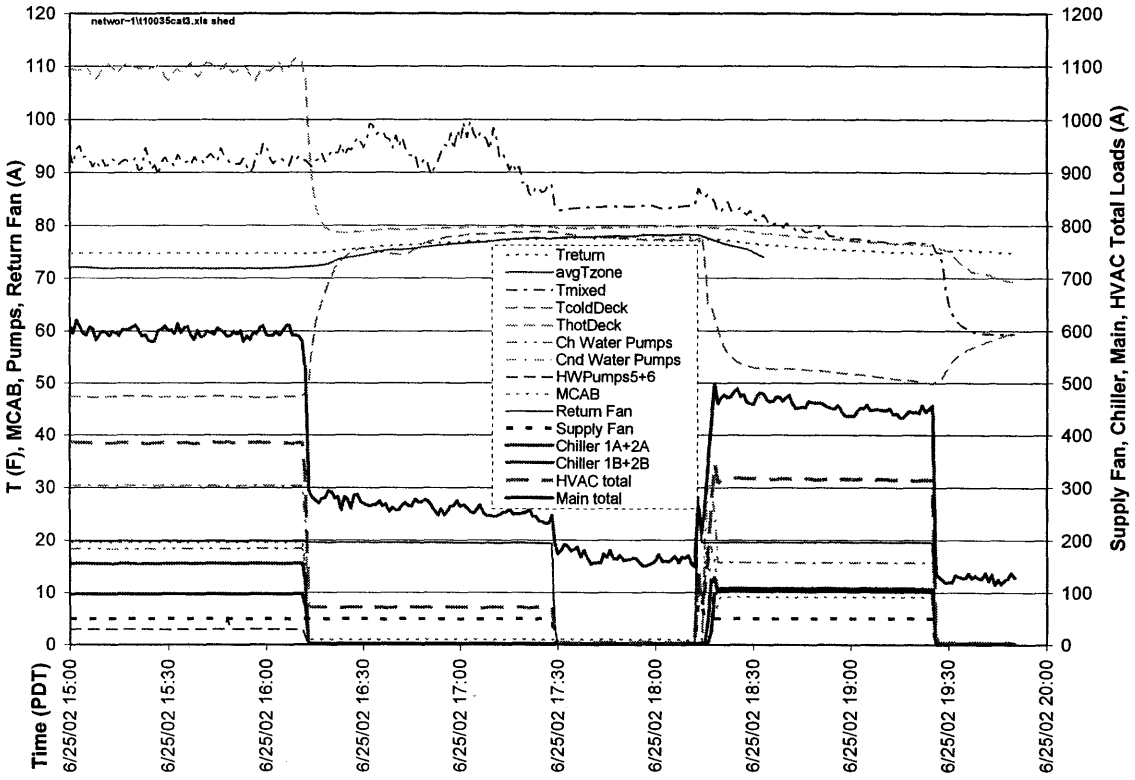

Figure 4.5. Temperatures and HVAC motor loads during ISD load shedding test.

4.3.5 Children's Court Precooling Test Results

A second ECC precooling test was completed 23 January 2003 in a week of mixed weather and building occupancy. Monday, 20 January, was sunny, Tuesday hazy, Wednesday and Thursday moderately overcast. Temperatures were typical for January except Wednesday morning was about 5°F below normal. The building was closed to the public on Monday for Martin Luther King Day.

The chiller was turned off 17:00 PST on Wednesday afternoon and the supply fan static pressure (SP) setpoint was reduced from 2.2 to 0.5 inches (water gauge) at 01:00 early Thursday morning. The setpoint was restored to 2.2 inches at 07:00 and the chiller was restored at 10:15. The HVAC and building electrical loads are shown in Figure 4.6 and the temperature trajectories are plotted in Figure 4.7.

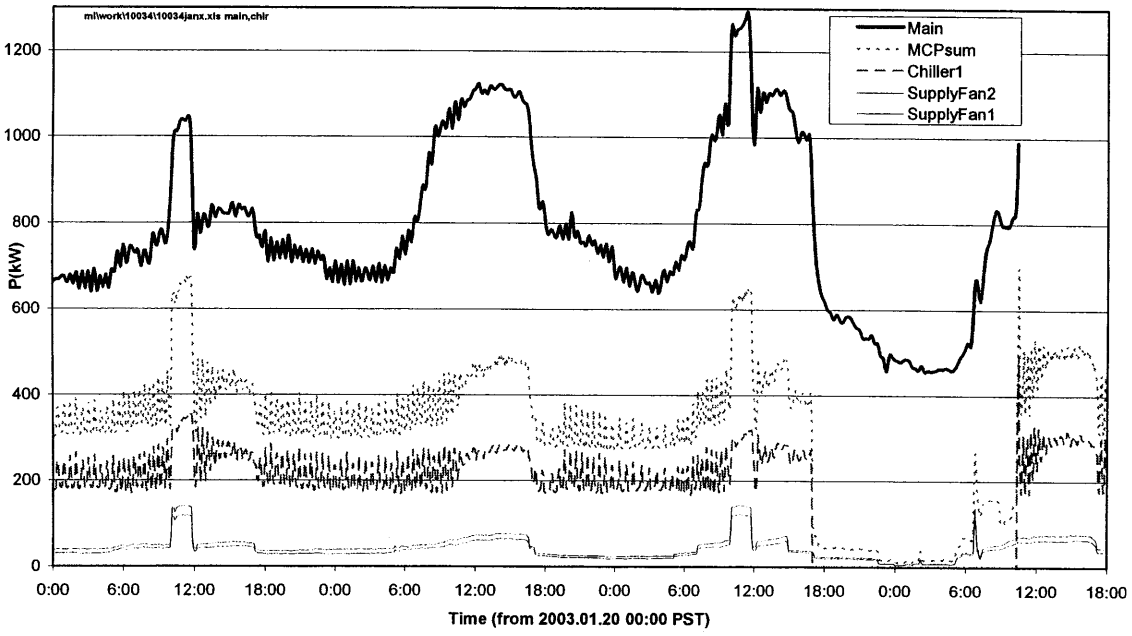

Figure 4.6 ECC electrical loads for 20-23 January 2003.

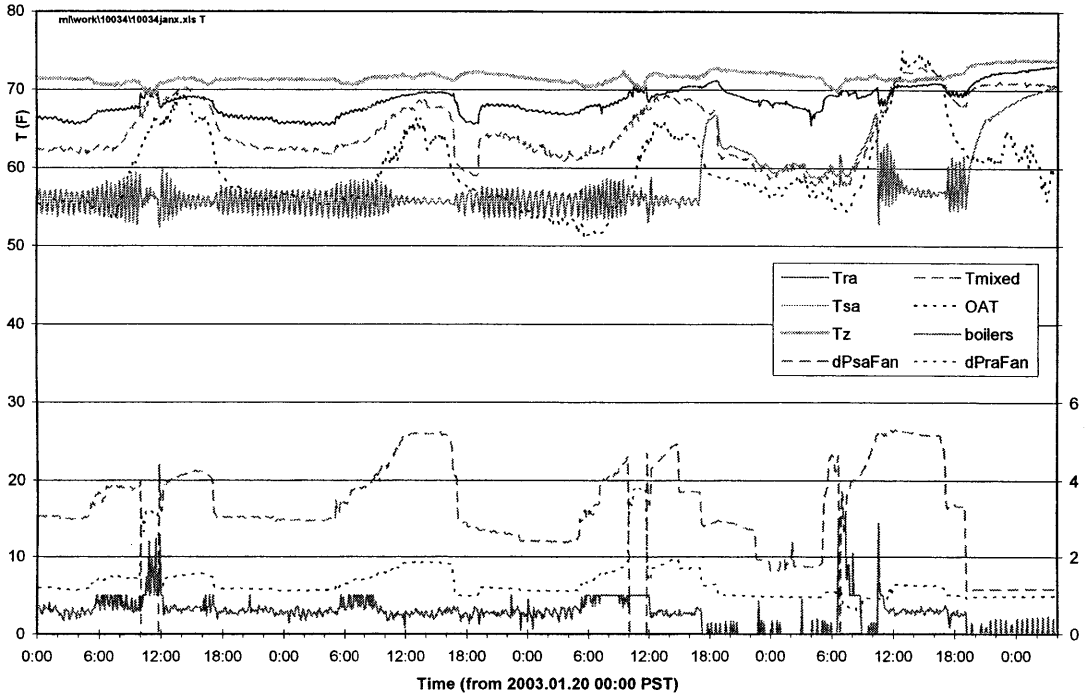

Figure 4.7 ECC temperatures; fan inlet velocity pressure and boiler duty cycle signals.

The reason for large supply fan loads from 10:00 to noon Monday and Wednesday is unknown; some possibilities are a change in zone setpoints, a SP sensor fault, or some unexplained meddling with SP setpoint schedule. Note that the mean supply air temperature doesn't change significantly even though its fluctuations are greatly diminished. Chiller power increases to maintain supply air temperature with the increased supply air flow rate. Chiller pump loads (not shown) are steady. Zone temperatures drop as would be expected in response to a step change in cooling capacity.

Also note that the supply air temperature control loop is poorly tuned for part-load operation. The result is oscillation of the coil control valve, reflected in chiller power, at about 2.2 cycles per hour. The amplitude of these oscillations decreases during times of high cooling load (high supply air flow rate). The daily mean conditions, loads, and room temperatures averaged across 12 zones are summarized in Table 4.1.

Table 4.1 24-Hour average conditions at Children's Court for 20-23 January 2003

		Monday	Tuesday	Wednesday	Thursday
CtowerFan (CT1)	kW	1.8	1.3	1.3	1.5
CndWpump (P3)	kW	28.2	28.3	25.6	10.3
ReturnFan (R1,R2)	kW	0.3	0.3	0.3	5.3
SupplyFans (S1,S2)	kW	100.4	90.6	81.5	74.9
HWpumps (P5,P6)	kW	3.8	3.8	3.5	1.9
ChWpump (P1)	kW	23.6	23.7	22.0	9.1
Chiller1	kW	236.7	222.1	203.5	99.6
MCP feed	kW	394.8	370.0	337.6	202.7
C-H main	kW	415.2	465.7	462.4	306.5
Tra	°F	67.5	67.3	68.6	69.2
Tmixed	°F	65.2	63.8	64.6	63.9
Tsa	°F	55.8	55.8	56.6	59.7
Rhra	%	49.9	52.8	51.6	50.3
Rhmixed	%	44.8	51.4	45.0	40.8
Rhsa	%	70.8	74.3	73.2	65.2
Tz (Hobos)	°F	71.0	71.4	71.5	71.7
Tamb (outside)	°F	60.1	58.5	57.2	61.9
S (plain LiCor)	W/m2	154.1	114.2	75.9	77.1
Total (unshaded)	W/m2	154.1	114.4	75.9	77.7
Diffuse1 (shaded)	W/m2	27.7	55.9	55.3	62.3
Prt'l (partly shaded)	W/m2	136.1	101.8	69.8	73.1

Because the chiller operates inefficiently at part load, there is little variation in average chiller power except on the day of the precooling test, when it was completely shut off for 16 hours. The chiller average power Thursday was thus reduced from over 200 kW to 100 kW. The associated average cooling tower and chilled water pump loads were also reduced substantially from 49 to 21 kW. Supply fan power is about 16kW (4 kW daily average) during the precooling phase and about 112 kW during the day. Baseline chiller and pump power are very high Monday through Wednesday because no unoccupied periods were scheduled on the control system. The baseline daily average would have been about 150 kW for the chiller and 30 kW for pumps under a normal night lockout schedule. Estimated savings from this reduced baseline are, nonetheless, a considerable 45% which represents about $50k/year. The extra fan energy (compared to a fan-off-at-night baseline) is only about 10% of chiller and pump savings thanks to the reduced static pressure during precooling.

4.3.6 ISD Building Precooling Test Results

The test was completed Wednesday, 22 January. In contrast to Children's Court, the ISD Building was partly occupied on the preceding Monday, 20 January 2003. Temperature trajectories are plotted in Figure 4.8 and HVAC and building electrical loads are plotted in Figure 4.9.

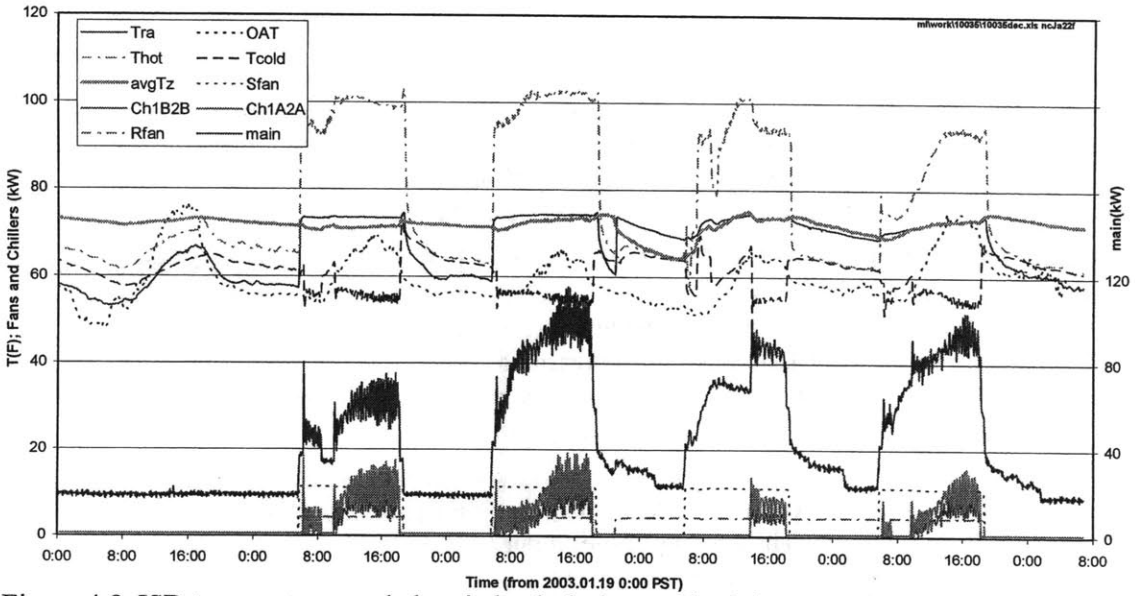

Figure 4.8 ISD temperatures and electric loads during week of the precooling test.

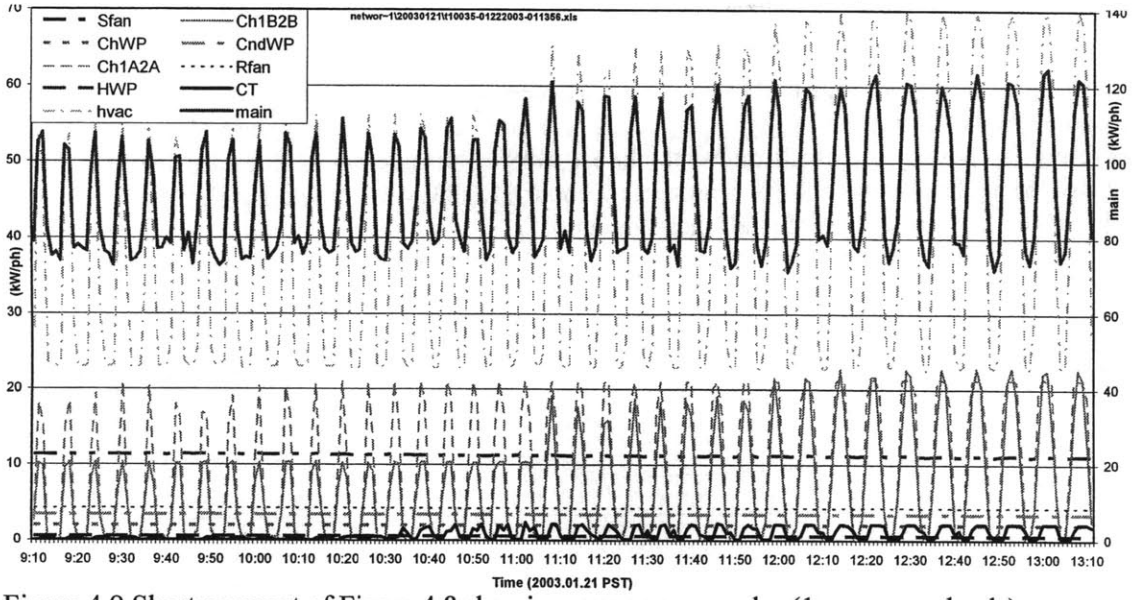

Figure 4.9 Short segment of Figure 4.8 showing compressor cycles (1m-average loads)

For this precooling test the chiller was turned off 18:00 PST on Tuesday afternoon and restored at 13:50 the following day. The windows were opened at 21:30 Tuesday and the return fan was started. Return temperature dropped from 73.2 to 69.0°F and average room temperature dropped from 73.5 to 64.3 F during the precooling period, which ended at 05:30 Wednesday morning with closing of the windows. The supply fan started at 05:40 per its normal schedule. The exhaust damper was closed

and the return damper opened from 07:00 to 8:45 allowing the occupied space temperature to rise from 67.9 to 71.7°F. The space was thus effectively in a tempering mode from 05:30 to 08:45. This amount of tempering is more than was necessary; tempering would have ended at 07:20 when the zone reached 70°F, had the controls been properly automated.

Average zone temperature reached 74°F at 11:45am. The chiller remained locked out until 13:50, by which time the zone temperature had reached 75.1°F. The return fan remained on through 24 January but the windows were not opened that night. This resulted in limited precooling Thursday morning by outside air being drawn through the hot and cold deck systems even though the supply fan was off. Note that return temperature approaches room temperature progressively more closely from Monday (after the previous week and weekend in which there was no precooling) to Thursday. Several days of precooling are needed to cool the building's floor deck structures from their under (ceiling plenum) sides. It is fair to conclude that 1) a one-day test is not sufficient to demonstrate the full potential and 2) precooling on Saturday, as well as Sunday, night may be cost effective.

Chiller cycling is apparent in Figure 4.8. Details of this cyclic behavior are shown to better advantage on the expanded time scale of Figure 4.9. One may conclude that the cycling could be largely eliminated by modifying the compressor sequencing logic so that compressors come on one by one, rather than in pairs. The response to precooling is presented on an expanded time scale in Figure 4.10 where the cycling tendency, even under the heavy zone temperature pull-down load that existed when the chillers were finally started at ~14:00, can be clearly seen.

The daily mean conditions, loads, and room temperatures averaged across six zones are summarized in Table 4.2. The ISD chillers, in contrast to the Children's Court chiller, operate more efficiently at part load (in spite of the cycling). The average chiller input is therefore significantly lower on cooler and partial occupancy days. Average chiller power on the day of the precooling test is, nevertheless, less than half of the average power used on the other three days. Fan power, on the other hand, changes little from day to day except that average return fan power is higher Tuesday through Thursday because it was run during certain unoccupied hours on all three of those days.

On the day of precooling, average chiller power was reduced from ~22 to 8.3 kW. Associated average cooling tower and chilled water pump loads also dropped substantially from ~8 to 3.7 kW. Average return fan power increases from ~7 to almost 13 kW. Net average savings are therefore about 12 kW from a baseline of 58 kW representing about $15,000/year. Precooling by the supply fan would reduce savings by roughly half these amounts.

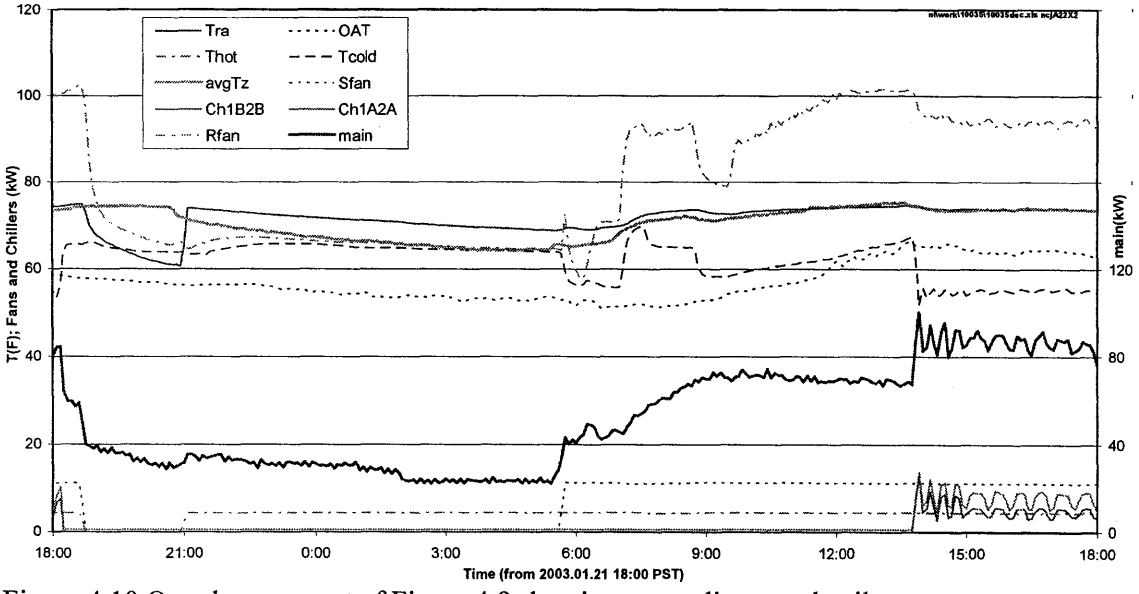

Figure 4.10 One-day segment of Figure 4.8 showing precooling test details.

Table 4.2 Daily Summary ISD Building Data for 20-23 January 2003

		Monday	Tuesday	Wednesday	Thursday
Chiller1	kW	12.45	13.88	5.58	9.87
Chiller2	kW	8.97	10.18	2.74	6.32
ChWpumps	kW	2.57	2.98	1.16	2.37
CtowerFans	kW	4.45	5.16	1.99	4.10
CndWpumps	kW	1.15	1.14	0.54	0.81
ReturnFan	kW	6.94	8.56	12.80	9.96
SupplyFans	kW	18.19	18.27	18.19	17.93
Hwpump	kW	0.89	0.89	0.87	0.88
HVAC feed	kW	58.14	63.94	46.02	54.75
ISD main	kW	111.61	173.73	157.20	163.46
Tra	F	67.3	69.6	72.0	69.8
Rhra	%	54.7	52.2	49.1	49.7
Thot	°F	84.0	84.5	79.1	76.3
Tcold	°F	60.2	59.6	61.8	59.9
Rhcold	%	60.5	63.7	61.5	60.3
Rhmix	%	37.5	41.4	53.8	53.3
B1lo	s/min	21.1	19.6	13.8	0.0
B2lo	s/min	35.3	34.9	33.7	35.3
B2hi	s/min	5.5	4.6	18.6	27.4
Tz (Hobos)	°F	71.6	72.0	70.8	72.2
Tamb (outside)	°F	59.4	58.5	57.2	61.9
S (plain LiCor)	W/m2	154.1	114.3	76.0	77.1
Total (unshaded)	W/m2	154.1	114.4	75.9	77.7
Diffuse1 (shaded)	W/m2	27.7	55.9	55.3	62.3
Prt'l (partly shaded)	W/m2	136.1	101.8	69.8	73.1
Barometer	PSIA	14.59	14.61	14.64	14.59

4.4 Simulation

To compare cooling control strategies we must estimate annual energy use and cost for each strategy while all other aspects of building operation are held constant. These "constants" include fan and chiller performance curves; fan, chiller, and auxiliary equipment control sequences; weekly and holiday schedules that describe internal gains, minimum outside air and fan static pressure; zone temperature setpoint schedule and related control parameters; utility rate schedule, and weather.

The information flow and main functional elements of the simulation are shown in Figure 4.11. The thermal response model of chapter 3 is integrated with economizer operation as described in 4.4.1. This model is of of general form and is made building-specific by plugging in the model order, number of walls, and coefficients. The chiller and internal gain models are described in sections 4.4.2 and 4.4.3.

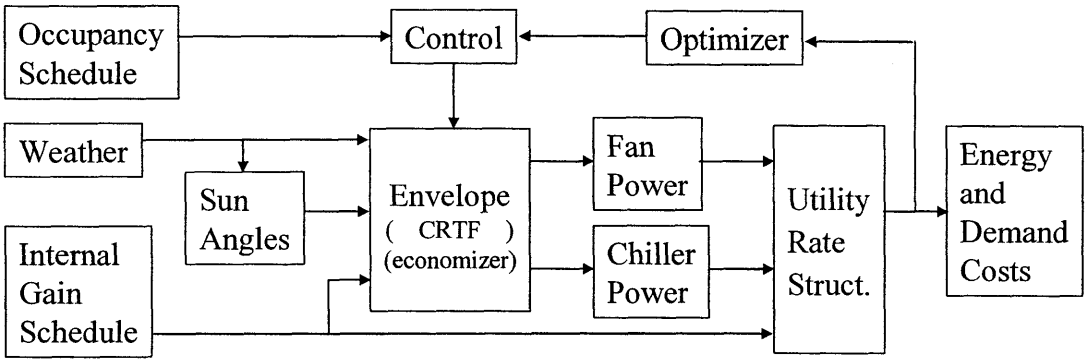

Figure 4.11 Simulation block diagram showing main functions, structures and information flow

4.4.1 CRTF for Simulation

Recall from the *Model Synthesis* subsection (3.3) that the discrete-time model (5) can be evaluated recursively for Q_z or T_z:

$$\text{for } Q_z\text{:}\quad -\phi_{z,0}Q_z = B_1^n(\phi_z,Q_z) + B_0^n(\theta_w,T_w) - B_0^n(\theta_z,T_z) \qquad (6a,b)$$

$$\text{for } T_z\text{:}\quad \theta_{z,0}T_z = B_0^n(\phi_z,Q_z) + B_0^n(\theta_w,T_w) - B_1^n(\theta_z,T_z) \qquad (8a)$$

With $\varphi_z = -1$ eqn (6) becomes:

$$Q_z = B_1^n(\phi_z,Q_z) + B_0^n(\theta_w,T_w) - B_0^n(\theta_z,T_z)$$

When the zone heat balance involves exchange of outside air with mass flow rate F, the foregoing expression for Q_z becomes:

$Q_z = Fc_p(T_z - T_{oa}) - Q_{star}$, where $Q_{star} = Q_{h/c} + Q_{solrad} + Q_{ig}$, the expression for T_z becomes:

$$T_z = \frac{Fc_p + Q_{other} + B_1^n(\phi_z,Q_z) + B_0^n(\theta_w,T_w) - B_1^n(\theta_z,T_z)}{\theta_{z,0} + Fc_p} \qquad (8b)$$

and it is easy to see that with $F = 0$, (8b) reduces to (8a).

4.4.2 Chiller Performance

ISD coil loads were monitored for two weeks, 29 August – 9 September, 2002, a period which included very hot weather. The total coil load is computed from its sensible, $Fc_p(T_{cold} - T_{mix})$ and latent, $Fh_{fg}(w_{mix} - w_{cold})$, air-side components. The measured 5-minute average sensible and total coil load are plotted in Figure E-1 against chiller power. Note that the latent fraction increases moderately with load, as expected for an essentially constant air-side flow rate. Chiller-specific power, expressed as chiller power per unit coil load (kW/Ton) increases with load. This is also expected. The condenser and evaporator areas are fixed so the approach temperatures increase with load resulting in the lift temperature increasing even faster than the condenser-chilled water temperature difference. Only one chiller was operated during this period. There are four compressors per chiller and the resulting performance in each of four operational stages, clearly evident in the plot, is summarized in Table 4.3.

Table 4.3 ISD chiller performance summary for 29 August – 9 September 2002

Qtot (kBtuh)	Qtot (RefTon)	Qlat/Qtot (-)	Pch/Qtot (kW/RT)
770	64	0.28	0.52
1270	106	0.31	0.65
1500	125	0.30	0.81
1700	142	0.31	0.93

For discrete-time simulation it is inconvenient and unnecessary to emulate the cycling between stages that occurs with periods less than the simulation timestep. A power law, shown in Figure 4.12, is therefore used to represent chiller performance in the simulation. Control parameters determine maximum capacity and at what load the second chiller starts.

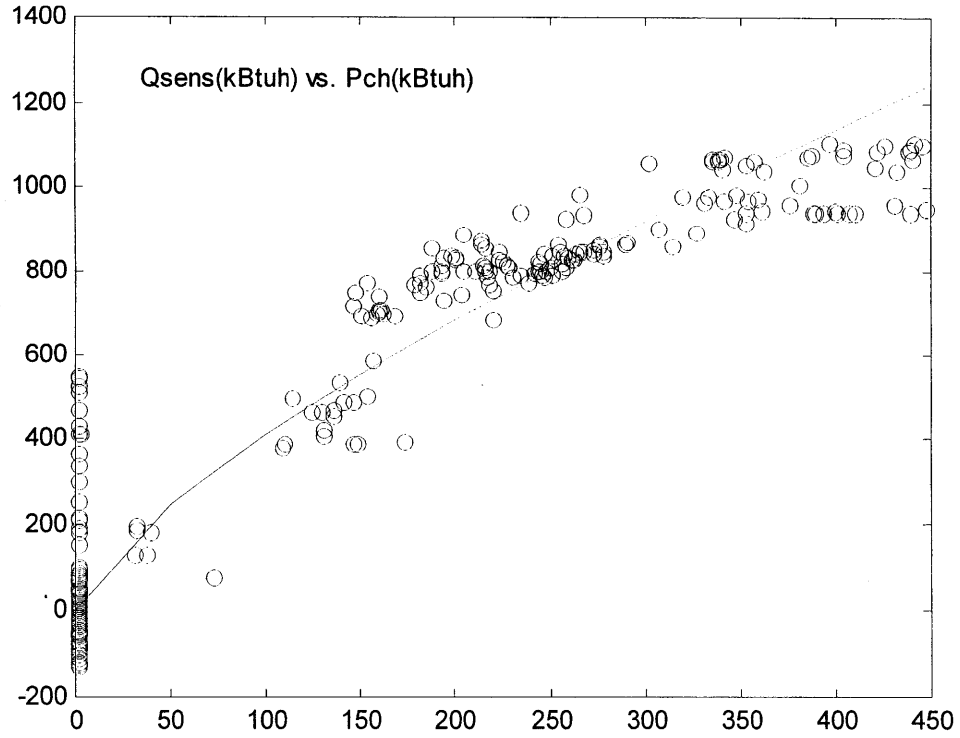

Figure 4.12 Power law used to simulate ISD part-load chiller performance.

4.4.3 Internal Gain Schedule

The measured internal gains are shown in Appendix P???. Weekday and weekend profiles are distinct and repeatable. The ISD building is supposed to be operated in weekend mode on Fridays but some people do work and the resulting lighting and plug loads are variable. The 24-hour weekday and weekend profiles shown in Figure 4.13 were obtained by averaging the observed kW numbers.

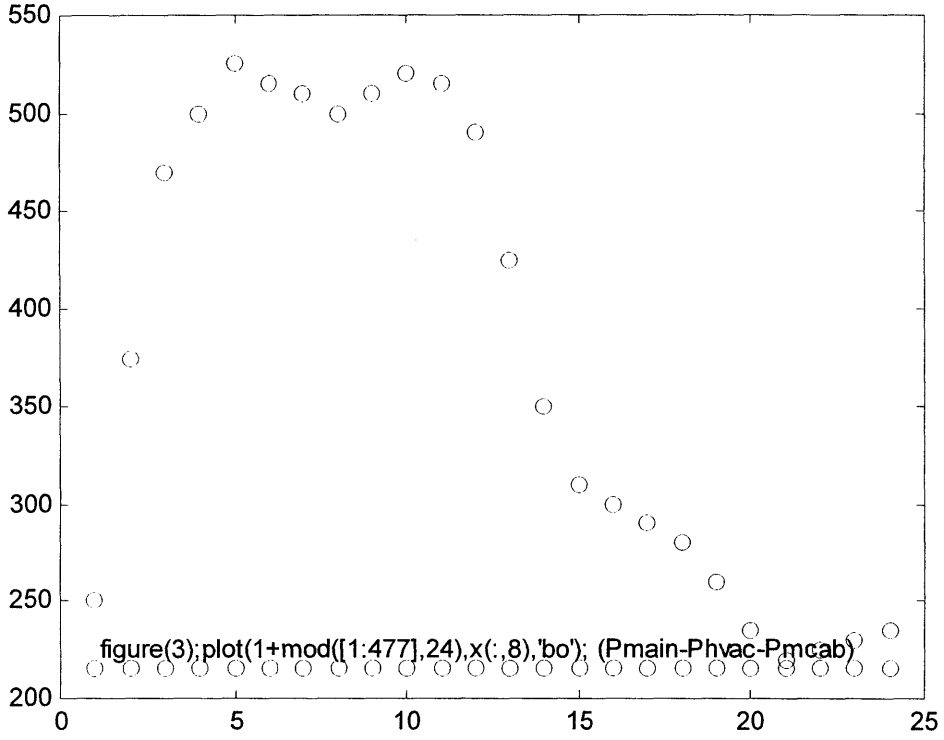

Figure 4.13 Occupied and unoccupied internal gain schedules used for ISD simulations (kW).

4.4.4 Control of Fan and Outside Air

The following truth table summarizes the control structures used in the simulation in a form that can be readily ported to an EMCS for a demonstration. The logical input columns (1-6) have entries of 1=true, 0=false and *=either. The zone temperature setpoint relation $T_z|T_e$ refers to the economizer set point, T_e, which may be different in occupied and unoccupied periods. For example $T_z<T_e$ generally turns the fan off in unoccupied periods and reduces outside airflow to the minimum required for ventilation in occupied periods. The output columns (7-8) refer to fan and damper operation. F_{nom} is the nominal (maximum) supply fan mass flow rate and $\%F/F_{nom}$ is the fraction of F_{nom} expressed as a percent.

1	2	3	4	5	6	7	8	9	10	11
Occupied	Night fan	Fan gain	$T_z > T_{oa}$	$T_z \mid T_e$ (1)	No econ	$\%F/F_{nom}$	%OA	H/C SP	TQ (2)	Notes
1	*	*	*	*	1	100	15	1		(3)
1	*	*	0	*	0	100	15	1		
1	*	*	1	<	0	100	15	1		
1	*	*	1	≥	0	100	100	1		(4)
0	*	*	*	*	1	0	0	2	T	(3)
0	0	0	*	*	0	0	0	2	T	
0	1	0	0	*	0	0	0	2	T	
0	1	0	1	<	0	0	0	2	T	
0	1	0	1	≥	0	100	100	2		(4)
0	1	>0	*	*	0	Illegal		2		
0	0	>0	0	*	0	0	0	2	T	
0	0	>0	1	<	0	0	0	2	T	(5)
0	0	>0	1	≥	0	(6)	(6)	2	T	(4,6)

(1) T_z, the economizer cooling setpoint, may take different values in occupied and unoccupied periods
(2) evaluate T_z when H/C setpoints are satisfied: $T_H < T_z < T_C$; else evaluate Q_z unless noted otherwise in this column
(3) when "NoEcon" is set we always zero "NightFan" and "FanGain"
(4) when result is $T_z < T_e$, set $T_z = T_e$, evaluate CRTF, and reduce fan power so that airflow, F, corresponds to Q_z
(5) usually set T_e = -99C
(6) $F=g*(T_z - T_{oa})$ where g is daily fan gain, the optimization variable.

4.4.5 Setpoint Control

Because discrete thermostat setpoints are used, the rate of HVAC heating and cooling does not vary continuously with zone temperature. The operating point at each timestep (Q, T_z) cannot therefore be solved by a derivative method. We instead use a sequential algorithm to determine at which setpoint, or within which thermostat band, the zone temperature must be. Once this is determined, the heat rate and exact zone temperature can be evaluated.

There are three zone setpoints for heating, T_{H1}, economizer cooling, T_{C1}, and vapor-compression cooling, T_{C2}. The setpoints must satisfy the inequality

$$T_{H1} < T_{C1} < T_{C2},$$

and the three corresponding maximum plant capacities must satisfy

$$-Q_{H1} > 0 > Q_{C1} > Q_{C2}.$$

Note that economizer cooling capacity at 100% outside air is given by:

$Q_{C1} = -F_{nom}c_p(T_{C1} - T_{OA})$ with $T_Z > T_{C1}$, or zero, whichever is less,

and that vapor-compression cooling capacity is:

$Q_{C2} = -F_{nom}c_p(T_{C1} - T_{OA})^+$ plus vapor-compression capacity at the highest stage of cooling.

The CRTF heat rate and new zone temperature (Q_z, T_z) are evaluated at each timestep by applying the plant capacities and corresponding setpoints progressively starting with[28] Q_{C2}:

$$T_z = CRTF(Q_{C2}, F_{nom});$$

if $T_z < T_{C2}$ $\quad Q_C = CRTF(T_{C2}); \; T_z = T_{C2}$

if $Q_C > Q_{C2}$ $\quad T_z = CRTF(Q_{C2}, F_{nom})$

[28] One might want to start at the other end in a heating-dominated climate.

if $T_z < T_{C1}$ $\quad Q_C = CRTF(T_{C1});\ F=F_{nom}*Q_C/Q_{C2}$,

if $Q_C > Q_{C1}$ $\quad T_z = CRTF(Q_{C1},\ F_{nom})$

if $T_z < T_{C1}$ $\quad Q_C = CRTF(T_{C1});\ F=F_{nom}*Q_C/Q_{C1}$,

if $Q_C > 0$ $\quad T_z = CRTF(0,\ F_{min})$

if $T_z < T_{H1}$ $\quad Q_H = -CRTF(T_{H1});\ F=F_{nom}*Q_H/Q_{H1}$,

if $Q_H > Q_{H1}$ $\quad T_z = CRTF(-Q_{H1},\ F_{min})$

As soon as the bracketing setpoints have been found, the progression can be terminated and the heat balance solved for $Q_{H/C}$:

$$Q_{CRTF} + Q_{H/C} + Q_{IG} + Q_{DSG} = 0$$

where

$Q_{CRTF} = \theta_{xo}T_x - \theta_{zo}T_z + Q_P$., and

Q_{DSG} = rate of direct solar gain,

$Q_H = Q_{H/C}$, $Q_C = 0$, when $Q_{H/C} > 0$, and

$Q_H = 0$, $Q_C = Q_{H/C}$, when $Q_{H/C} < 0$.

The final evaluation of zone temperature is then given by:

$$T_z = ((Fc_p + \theta_{xo})T_x + Q_P + Q_H + Q_C + Q_{IG} + Q_{SR})/(Fc_p + \theta_{zo})$$

4.5 Results for ISD Building Driven by LA TMY

Here the annual performance of night precooling control strategies is evaluated by using TMY2 Los Angeles weather data to drive the ISD thermal response model. The cooling capacities provided by the chiller plant and by outside air are separately integrated over the year for each case simulated. These numbers and the annual electricity (energy and demand) costs together provide a complete picture of the alternative control strategies described in Section 4.1.1.

The base case, no night cooling with the economizer setpoint equal to the mechanical cooling setpoint (74.5°F), is shown in Figure 4.14. Mechanical cooling represents almost 60% of the total annual load. As the economizer setpoint is reduced, the total cooling load is increased but the annual mechanical cooling share is reduced to about 50% of the total with the most extreme economizer setpoint of 68°F. The cooling shares are not very sensitive to economizer setpoint because there is little daytime economizer cooling potential on days when there would be significant chiller load without economizer cooling.

With economizer cooling enabled at night (fan runs whenever zone temperature is above the economizer setpoint and outdoor temperature is below zone temperature) mechanical cooling is immediately reduced by 30% for case 6 (economizer setpoint = mechanical cooling setpoint = 74.5°F) simply by eliminating morning pull down loads. Total annual cooling is immediately increased by 20% in response to lower zone temperatures on many nights. Note, furthermore, that the mechanical share drops significantly (an additional 30%) and total load increases significantly (another 17%) as the economizer setpoint is lowered. These results are shown in Figure 4.15.

With optimal nightly delayed-start times, the annual fan energy (hours of fan operation) is greatly reduced, the amount of economizer cooling is moderately reduced, and the amount of mechanical cooling is increased slightly as shown in Figure 4.16. Total annual cooling numbers are 10-15% lower. However, as shown in Figure 4.17, there is very little cost savings because off-peak energy rates are relatively low.

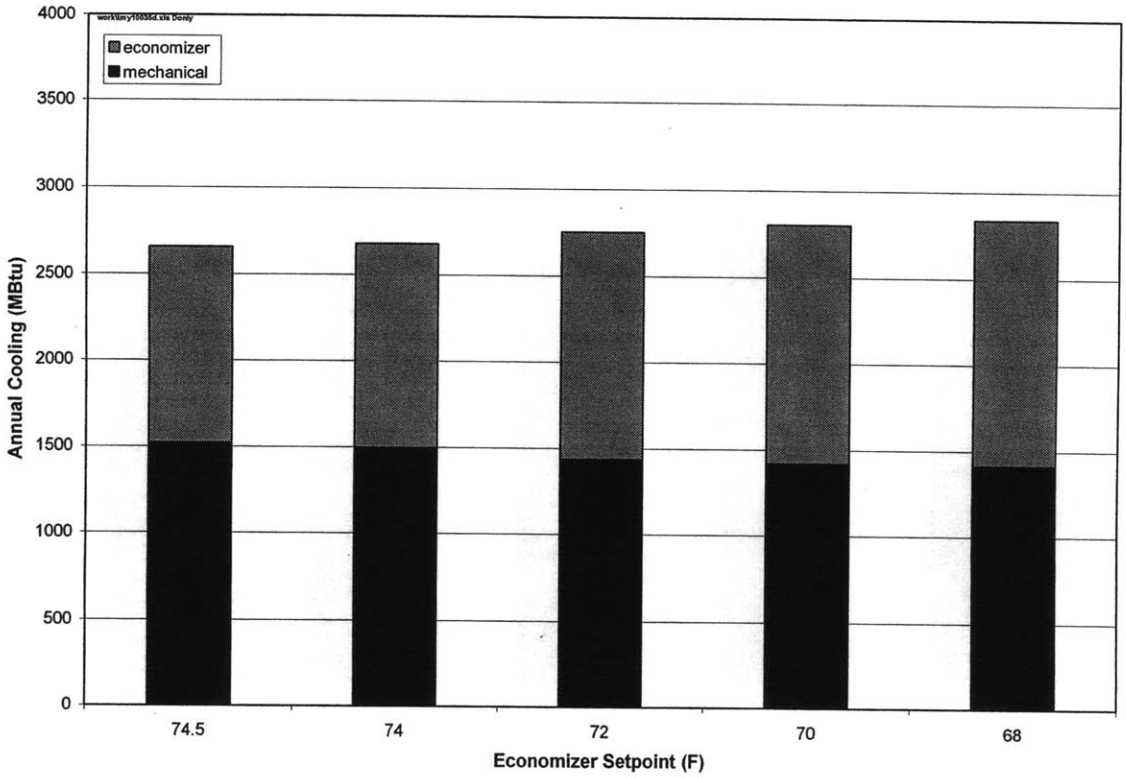

Figure 4.14 No night cooling, cases 1-5.

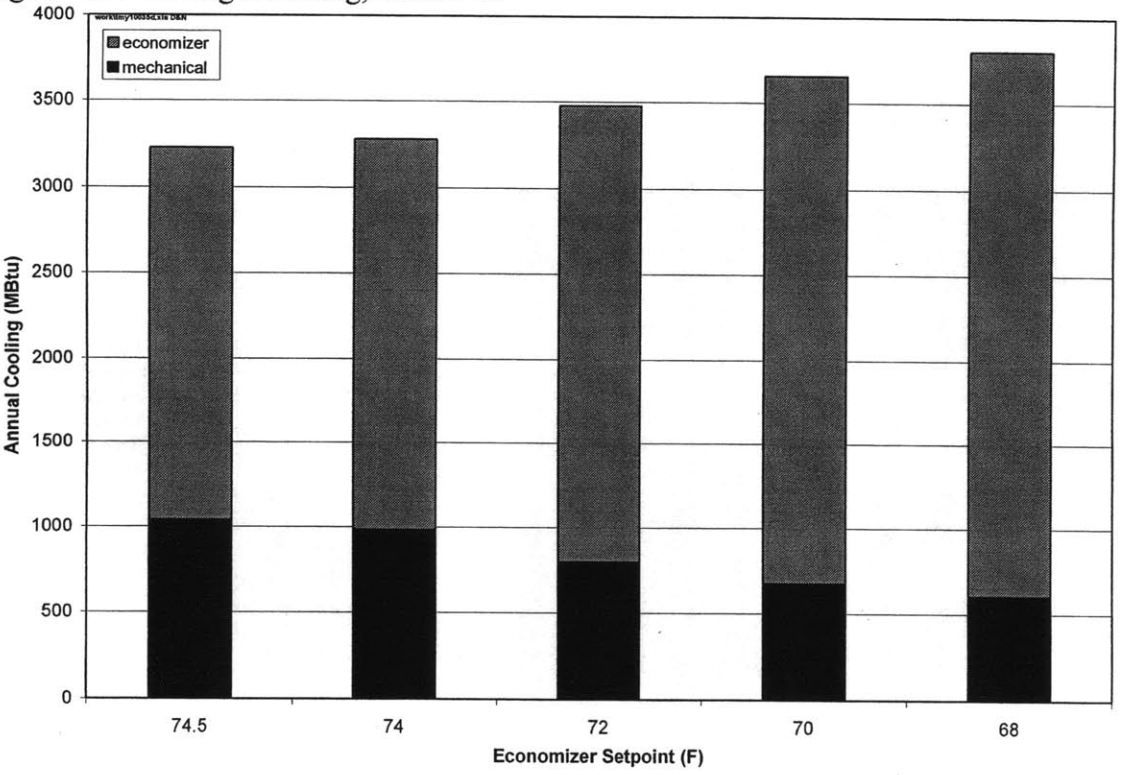

Figure 4.15 All night cooling, cases 6-10.

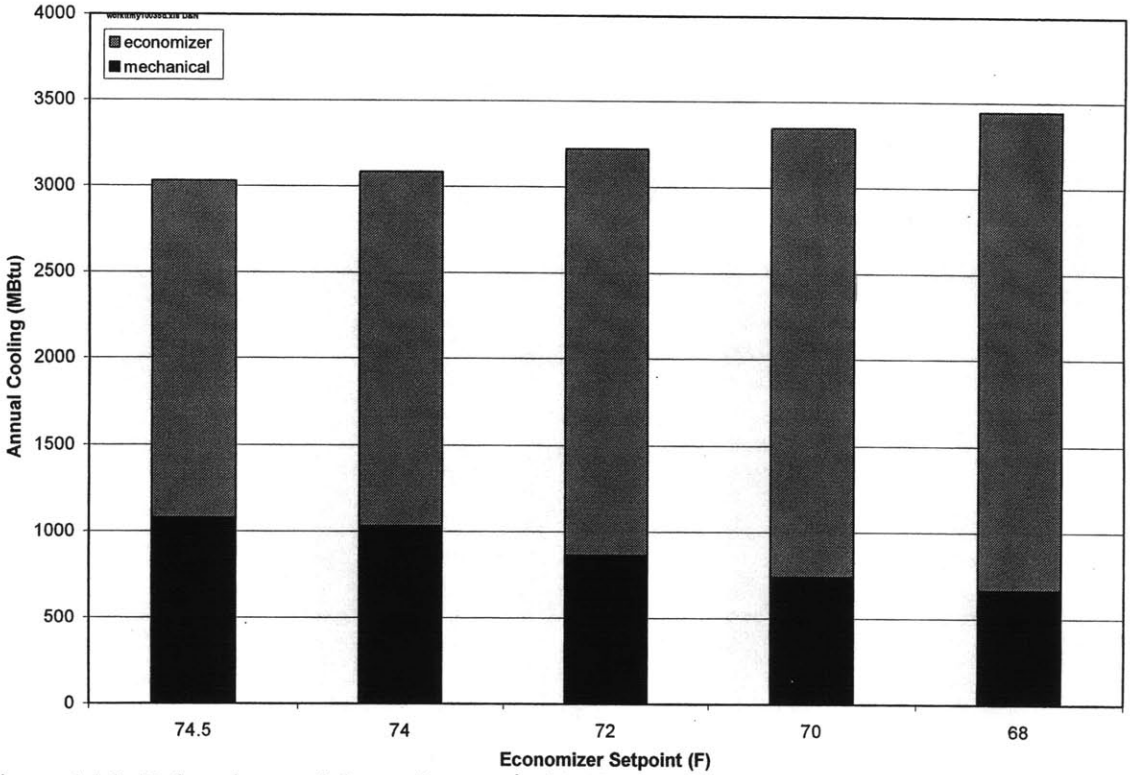

Figure 4.16. Delayed-start night cooling, cases 11-15.

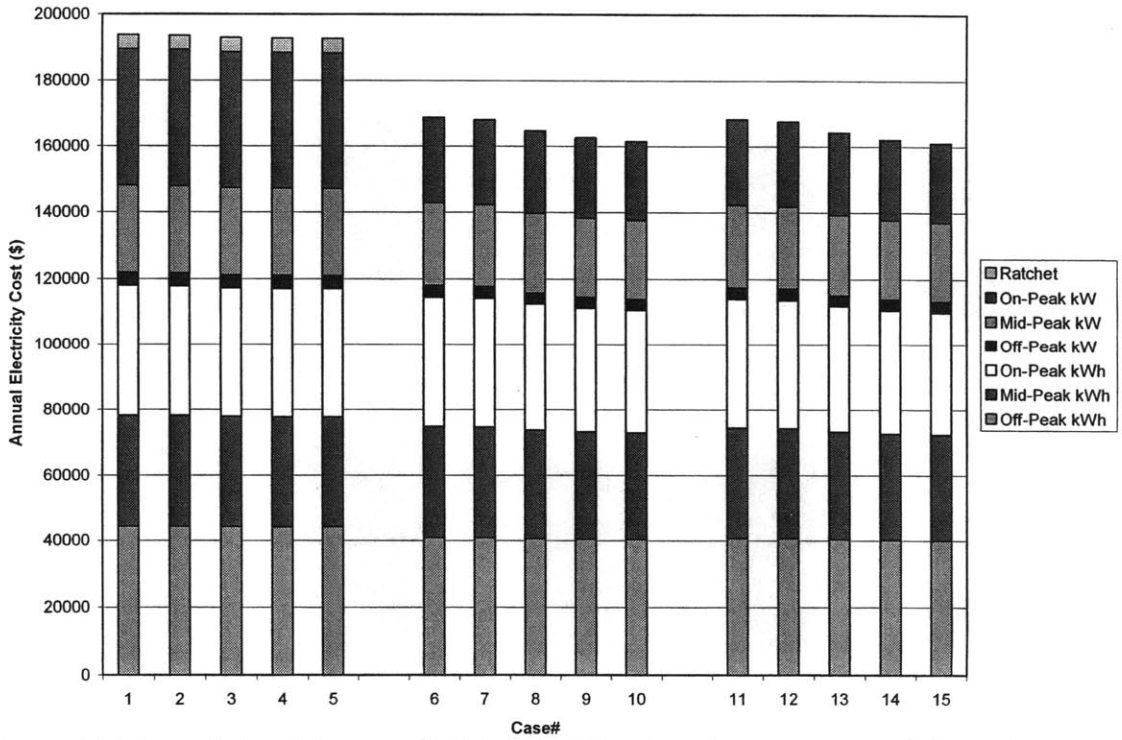

Figure 4.17. Annual electricity costs (building total) by time of use, energy, and demand

4.6 Summary

In Chapter 4 a simulation environment was developed for the study of optimal and suboptimal control strategies involving building thermal transient response. Chiller performance curves, utility rate structures, and building thermal response models can be plugged in to enable building-specific analysis.

A precooling control strategy with variable subcooling time and variable tempering time was developed. The strategy results in significant reduction of mechanical cooling using typical Los Angeles weather, utility rate structure and office occupancy schedules. Sensitivity of annual savings to building mass and mean annual outdoor temperature was characterized.

In the base case, no night cooling with the economizer setpoint equal to the mechanical cooling setpoint, mechanical cooling represents 60% of the total annual load. This was reduced to 50% with the economizer setpoint reduced by 6°F.

With economizer cooling enabled at night mechanical cooling is reduced another 30% by eliminating morning pull down loads and yet another 30% when the economizer setpoint is reduced by 6°F. With optimal nightly delayed-start times, total annual cooling numbers are 10-15% lower and there is a small cost savings from reduced fan hours.

Because it is primarily a controls measure, and because it does not involve utility coordination[29], the implementation cost for night cooling is potentially quite low. However, packaging is again a key issue. The control package must provide reliable, autonomous identification of the plant's thermal response.

[29] Verification becomes a matter of internal evaluation, rather than being a contractual issue.

Chapter 5 Conclusion

The related themes of fault detection, diagnostics, and control are bound by the important role that a plant-specific model plays in each application.

Model-based FDD and model-based control can take many forms. For example, fault detection can start by comparing the response given by the detailed simulation of a system (driven by the measured excitations that also drive the real system) with the measured response of the real system. Various responses may be closely or remotely related to various faults. Detection may involve little or much uncertainty. The control application raises similar questions of model type. The detailed, physics-based simulation is one extreme. At the other extreme, the model may be completely non-physical, e.g. ANN- or statistics-based, or somewhere in between.

To be plant-specific often means to be adaptive and autonomous. And it often means a model that is identified or updated on-line. There may or may not be human initiation or intervention, e.g. as in the building commissioning process.

Consequently, the type of model is a key issue. There exists a white-box to black-box continuum in which the white-box model is based entirely on process physics and the black box model is based entirely on any sort of generic pattern matching that has no physical interpretation. It is rarely desirable to settle on either extreme. Even a model that *can* have a purely physical interpretation, e.g. the CRTF, can often also have a statistical interpretation, e.g., in the case of the CRTF, an ARMAX interpretation, and it would be rare for a plant to behave as a perfect CRTF so we must always be prepared to interpret model identification results statistically as well as physically.

White-box models typically present parameter estimation difficulties. If overcome the white- or grey-box model is typically better suited, for extrapolation, than a black-box model identified from the same training set. The discrepancy between the testing set prediction error norm and the training set prediction error norm is likely to be less for a light-grey-box than for a dark-grey-box model. These concerns are especially germane to building applications where identification data often lack richness or reflect a low signal-to-noise ratio. Thus this thesis has been largely concerned with identification and application of grey box models of the lighter sort and finding a good balance between reliability of identification and accuracy in application.

5.1 Thesis Accomplishments

Significant advances have been made in both direct and model-based fault detection, in zone- and building-level model identification, and in precooling control and design.

Chapter 2 shows how the NILM's signal acquisition and PSA capability can be used to detect faults, prefault behavior, and performance degradation. A recursive ARX-based filter for detection of normal and anomalous transients was developed. The consistent association of anomalous transients with liquid slugging, a dangerous condition that leads to compressor failure, was demonstrated by compressor bench tests. Liquid ingestion during steady operation results a transient power increase of 0.1-1s duration that appears clearly as positive deviations from k-step-ahead forecasts made by an adaptive ARX(5) model. Any positive power transient that fails to match a known start transient and that exceeds the adaptive model innovation norm threshold indicates liquid ingestion.

The following additional fault detection capabilities were demonstrated.

- Air-side blockage faults were detected by a shift in steady state (P1,Q1).
- Short cycling can be detected by PSA in the obvious way, by logging time between start transients. The advantage over detection by a refrigerant-temperature-based FDD system is that very short—even incomplete—start transients that promote severe overheating are detected. A second method of short cycling detection relies on the shape of individual start transients being sensitive to residual head.
- Refrigerant charge faults result in elevated head pressures which translate to higher steady state compressor loads. Undercharge can also be detected from the start transient shape.
- The refrigerant vapor bypass fault had a distinct effect on the start transient shape: the change of magnitude in the latter half of the transient was two to five times the change in steady state power.
- Flooded starts were detected by a distinctive change in start transient shape.
- PSA's ability to identify degradation of impeller balance well before the added repair costs of a catastrophic failure are incurred was demonstrated by spectral analysis.

The value of FDD is realized in several areas such as reduced down time, better efficiency, avoidance of consequential damage, and the efficient dispatch of repair staff and replacement parts. The available data, summarized in Table 5.1, are not sufficient to make precise statements about FDD savings without making some assumptions about costs avoided through detection of incipient faults or about the logistical savings achieved by remote FDD. For example, if we assume that early detection and repair of faults *leading to* compressor short, open, or locked rotor failures is 100% effective at preventing ultimate compressor failure, the savings are (.24)*(.66)=10% of all repair costs for this one class of faults alone.

Table 5.1 Breakdown of repair cost by fault from Breuker; columns 2 and 3 are percents of column 1.

		Detection by PSA Certain		Detection by PSA Probable	
Compressor	24	short, open, locked rotor	66	broken internals, vapor bypass	11
Condenser	9	Motor, belt failure, cycle control	51	fouling, coil leak	36
Evaporator	6		0	fouling, leak, loss of capacity	75
Air Handler	5	motor, impeller, belt, electrical	67	damper motor, duct leaks	20
Controls	10	compressor interlocks, fan sequence	37		0
Electrical	7	Contactor, loose connection	44	fuse/breaker, short/open	27
Miscellany	21	water loop/pump, fan belt	29	refrigerant leak, installation/startup	42
Other	18	n.a.		n.a.	
Total %	100	Repair of all PSA-detectable faults (sum of products with column 1)	37	Faults likely to be detected by PSA (sum of products with column 1)	22

A very general model of wall, zone and whole-building thermal response was developed in Chapter 3. Three thermodynamic constraints were also identified. The constraints are crucial to on-line identification for the model-based control application because the observations used for model identification are inevitably noisy and often incomplete or biased as well. Without the constraints, a physically implausible model that behaves badly under certain conditions (e.g., fast transients or temperature extremes) will often be accepted as the "best" model implied by the data.

Formulations were developed for enforcing the constraints using the simple optimization tools of ordinary least squares and bounded-search nonlinear least squares. The steady-state conductance constraint is handled algebraically. This approach eliminates one unknown but, more importantly, has the invaluable advantage of leaving the problem in unconstrained ordinary least squares form. The constraint on time constants is a non-linear constraint. To avoid the difficulties traditionally associated with non-linear constraints, a way was found to reformulate the problem as an unconstrained nonlinear least squares problem. A further modification was devised to enforce the constraint of alternating roots by imposing simple bounds on the transformed unknowns.

The problem of identifying coefficients for predicting both (Q and T_z) responses of interest using the CRTF form for heat load, Q, and its inverse, the ICRTF, for zone temperature, T_z, has been addressed. The "errors in variables" (aka *total least squares*) approach was tried with poor results. An alternative approach was developed by defining a new response that is the sum of the zero lag terms for Q and T_z. This approach has the advantage of great simplicity but the disadvantage of making the estimation problem nonlinear in one unknown, the ratio of Q and T_z zero-lag coefficients. In the case of the test room, which exhibited bad behavior when its conventionally identified models were inverted, the new procedure gave very satisfactory results.

Second-order models were identified for the test cases (test room and Zubkova building) that had spatially averaged zone temperature observations. For the two LA buildings where the number of zone temperature loggers was relatively small for the magnitude of expected spatial variations, it was possible to identify only first-order models. The massive Zubkova building came close to giving a third-order model and it is likely that better observations of excitations, particularly the solar gains, would result in reliable identification of valid third-order models for such buildings.

In Chapter 4 a simulation environment was developed for the study of optimal and suboptimal control strategies involving building thermal transient response. Chiller performance curves, utility rate structures, and building thermal response models can be plugged in to enable building-specific analysis.

A precooling control strategy with variable subcooling time and variable tempering time was developed. The strategy results in significant reduction of mechanical cooling using typical Los Angeles weather, utility rate structure and office occupancy schedules. Sensitivity of annual savings to building mass and mean annual outdoor temperature was characterized.

Past research has focused on the savings potential of various precooling strategies in buildings of typical design. The question of how best to distribute building mass and how to estimate savings potential in early design phases based on mass distribution and the important parameters of lighting and equipment gains (over which designer and owner may have some discretion) is addressed for the first time in this thesis. A simple bin method based on step response is proposed for designers.

5.2 Future Research

A number of directions and tasks of various sizes could be beneficially pursued further.

All of the detection schemes discussed in Chapter 2 operate on 60Hz spectral envelopes (Leeb 1992). In single-phase applications this greatly simplifies the process of matching start transients. However,

in 3-phase applications a preprocessor that computes 3-phase power and reactive power does not really need to implement the spectral envelope filter, In fact, eliminating the filter will allow possibly useful high frequency signals pass. An immediately apparent benefit is a much higher resolution determination of contact closure. This in turn gives higher resolution, and more repeatable transient shapes, more reliable identification of start transients, and better detection of faults that result in changes to those shapes.

Detection of the flooded-start fault, in particular, would benefit. This benefit should be quantified and the impact on detection of other faults explored. Fitting the adaptive ARX directly to the unfiltered, high-sampling-rate P,Q signals provides an alternate transient event detector. This would help detect very short transients and small fault-induced innovations with significant signal power above 60Hz based on transients in admittance (Z), P, and Q generated without spectral envelopes

Availability raw 3-phase P and Q signals opens the possibility of estimating aggregate winding temperature (a function of locked-rotor[30] 3-phase input-terminal admittance) from inrush power, P, and the simultaneous V and Q signals. Temperature is a concern during heavy load conditions, when the off cycles are short. Although the temperature will only be estimated during the first few ms of excitation (when shaft power is insignificant) and the exact source of overheating (stator or rotor) will be indeterminate, any indication of overheating is useful because so many failures ultimately are the result of overheating. Moreover, if the rate of temperature rise can be estimated during the first few milliseconds, the presence of partly shorted windings or hot spots might be immediately detected.

The success of fan imbalance fault detection raises the possibility of detecting other similar mechanical faults. One likely candidate in this class of faults is motor support failures such as disintegration of a rubber bushing or fatigue failure of a support strut. This capability could be quickly tested at the Purdue RTU.

A vapor-compression testbed with small interchangeable reciprocating and scroll compressors should be established in the lab. One important liquid slugging detection question involves use of compressor suction and discharge pressures sampled at high (>=60Hz) rates to explore the lower limit of liquid slug mass ingestion detectable in otherwise steady operation. Other PSA-based fault detection areas that might benefit from pressure data should also be further tested. These include vapor bypass and over- and under-charge faults. In another testbed activity, the feasibility of a COP diagnostician based on compressor power and suction and discharge pressures should be explored using either a compressor map or a liquid line flow meter to track refrigeration capacity.

Finally, after accumulating sufficient laboratory and attended field testing experience, and refining and automating the detection algorithms, a large-scale deployment of PSA-based RTU FDD prototypes should be made. (A central plant demonstration would also be useful.) This would serve the dual purpose of a very thorough testing of PSA hardware, software, and application details and of collecting data about the frequency and nature of real faults that occur during the deployment.

There are still some unresolved numerical problems that prevented application in Chapter 3 of the time-constant-constrained models to our test buildings. Alternate ways to implement the constraint

[30] Locked rotor here denotes any negligible shaft load condition, e.g. near-zero torque or near-zero shaft speed.

should be explored. One possible method is the superposition of first order models fit to the original signal and the successively generated residual vectors. The successive models could have the same time step or start with a time step appropriate for the longest expected time constant and move progressively to a timestep corresponding to the raw sampling rate the latter approach addresses numerical problems typically associated with stiff systems, of which the Zubkova Building is a good example. Synthetic data sets and simple data sets, such as the time series of interior and exterior surface heat fluxes and temperatures taken during U-value testing may be useful test cases.

Time invariant model is almost always preferred from both the identification and simulation points of view. However fan operation makes the assumption untenable in some cases. One possible approach in these cases is to define auxiliary inputs related to fan operation that allow a more elaborate, but time invariant model to predict response over a complete data set as well as two simpler models, one for fan on and one for fan off operation, can perform on the disaggregated data.

It is still difficult to obtain data of acceptable quality (accuracy and completeness) for model identification. Data sets that include zone pressures and possibly intake and exhaust rates are needed when infiltration loads are significant. Separate observations of beam, diffuse, and ground-reflected radiation are needed when solar gains, incident-angle dependent transmission, snow cover, and seasonally or diurnally varying shading are significant. Finally, accurate and spatially well-resolved observations of room and plenum temperatures are critical to the model applications addressed here.

The Chapter 3 models pertain only to *sensible* heating and cooling loads. The transient latent cooling load response of the building envelope and its contents to changes in zone humidity conditions may be important for the curtailment application (Fairey 1983). The mass diffusion process may be too nonlinear for the successful application of approximate linear models of transient response. Future research should assess the importance of latent load to the curtailment control application and, if found to be important, should seek ways to implement the model-based control autonomously and reliably.

Chapter 4 results are based on thermal response of the occupied space. An alternative form of precooling could move cool air through a building's ceiling plenums. To explore this it is necessary to use separate plenum and conditioned-space thermal response models. Further research should determine the benefits of :

1) subcooling the plenum at night
2) maintaining an optimal daytime plenum setpoint schedule (Coniff 1991).

Chapter 4 focused on precooling. Curtailment and optimal start are both simpler, but equally important, applications of model-based control. It will be important to develop, simulate, and field test optimal start and optimal curtailment control strategies based on the identified building thermal response models. It is critical for practitioner acceptance to test the precooling, optimal start, and curtailment strategies in real buildings

Once optimal mass placement and optimal control of plenum, as well as occupied space temperature control has been developed to the point of routine application it will be possible to assess the national potential for savings by precooling and optimal start. The assessment should look at retrofit in

existing buildings as well as the potential in properly designed new buildings. The potential curtailment resource should be assessed in a similar manner.

The potential benefit of night precooling raises some policy questions. Standards (ASHRAE 90.1) are based on equipment and envelope specifications that result in least-life-cycle-cost building performance. Advanced control strategies have generally been avoided because they are so easily (and often are) abandoned soon after occupancy. However it may be argued that, in the case of control strategies that make use of building thermal mass, such standards effectively discourage adoption of advanced controls. A buyer of an existing building who wants to implement such controls is faced with the prospect of poor performance or of high initial cost to add mass that isn't already there. The policy research question then is "should buildings be required to be *precooling ready*?" and a corollary is "(how) should a performance-based standard trade mass for other envelope properties?"

References

Andresen I., M J. Brandemuehl, 1992. Heat storage in building thermal mass - a parametric study, *ASHRAE Transactions* 98(1) AN-92-8-3, 910-918

Armstrong, P.R., 1983. Instrumentation of solar heating and cooling systems at CSU, invited paper, Proc. *2nd Workshop on Performance Monitoring*, FSEC, Cape Canaveral.

Armstrong, P.R., C.E. Hancock and J.E. Seem,1992. Commercial building temperature recovery: design procedure based on step response model, *ASHRAE Transactions*, 93(1)

Armstrong, P.R., et al, 2004. Fault detection based on motor start transients and shaft harmonics measured at the RTU electrical service, *Int'l Refrigeration and Air Conditioning Conf. at Purdue.*

Armstrong, P.R., J.A.Dirks, R.W.Reilly, J.W.Currie, R.J.Nesse, B. Nekrasov, O.V. Komarov., 2000. Russian apartment building thermal response models for retrofit selection and verification, *ACEEE Summer Study on Energy Efficiency in Buildings.*

ASHRAE. 2001. *Handbook of Fundamentals*, Atlanta

Balcomb, J.D. 1981. Heat storage duration, *Proc. 6th National Passive Solar Conf.*, Portland, OR. LA-UR-81-2186.

Balcomb, J.D. 1983. Thermal network reduction, *Proc. Annual ASES Conf.*, Minneapolis, MN. LA-UR-83-869; also in LA-9694-MS

Barakat, S.A., 1987. Experimental determination of the z-transfer function coefficients for houses, *ASHRAE Transactions*, 93 (1) 3022.

Brambley, M.R., R.G. Pratt, D.P. Chassin, S. Katipamula, 1998. Automated diagnostics for outdoor air ventilation and economizers. *ASHRAE J.* 40(10).

Brandemuehl M.J., M.J. Lepoer, J.F. Kreider, 1990. Modeling and testing the interaction of conditioned air with building thermal mass, *ASHRAE Transactions* 96(2) SL-90-16-1, 871-875.

Braun J E., 1990. Reducing energy costs and peak electrical demand through optimal control of building thermal storage, *ASHRAE Transactions*, 96(2) SL-90-16-2, 876-888.

Braun J E., Klein S A., Mitchell J W., Beckman W A., 1989. Applications of optimal control to chilled water systems without storage, *ASHRAE Transactions*, 95(1), 663-675.

Braun, J.E. and Nitin Chaturvedi, 2002. An inverse grey-box model for transient building load prediction, *Int'l J. HVAC&R Research*, 8(1) pp.73-99.

Breuker, M.S. and J.E. Braun, 1998a. Common faults and their impacts for rooftop air conditioners, *Int'l J. HVAC&R Research*, 4(3) pp.303-318.

Breuker, M.S. and J.E. Braun. 1998b. Evaluating the performance of a fault detection and diagnostic system for vapor compression equipment. *Int'l J. HVAC&R Research*, 4(4) 401-425.

Bushnell, R.H., 1978. The measured heat loss coefficient of solar heated buildings, *ASES, Proc. Annual Conf.*, Denver.

Carslaw, H.S. and J.C. Jeager, 1959. Conduction of Heat in Solids, Oxford University Press.

Ceylan, H.T., and G.E. Myers. 1980. Long-time solutions to heat conduction transients with time-dependent inputs. ASME J. Heat Transfer 102(1) pp.115-120.

Ceylan, H.T., and G.E. Myers. 1985. Application of response-coefficient method to heat-conduction transients. ASHRAE Transactions 91(1A).

Coniff, J.P., 1991. , ASHRAE Trans., 96(2) 910-9.

Davies, M.G. 1973. The thermal admittance of layered walls. Building Science 8, 207-220

Dexter, A.L. Self-tuning Optimum Start Control of Heating Plant, *Automatica*, V.17, No.3 (1981)

Ehrich, F, and D. Childs, 1984. Self-excited vibration in high performance turbomachinery, *Mechanical Engineering*, May.

Eto, J.H., et al, 2002. Innovative Developments in Load as a Reliability Resource, LBNL-50968

Eto, J.H., 1984. Coolung strategies based on indicators of thermal storage in commercial building mass, *Annual Symp. On Improving Building Energy Efficiency in Hot and Humid Climates*, ESL, Texas A&M.

Forrester, J.R. and W.J. Wepfer, Formulation of a load prediction algorithm for a large commercial building, *ASHRAE Transactions*, V.90, pt.2B (1984).

Genta, G., 1998. *Vibration of Structures and Machines—Practical Aspects*, 3rd ed. Springer-Verlag.

Golub, G.H. and C.F. Van Loan, 1980. An analysis of the total least squares problem, *SIAM J. Numerical Analysis*, 17(6).

Gordon, J.M, and K.C. Ng, 2000. Cool Thermodynamics, Cambridge International Science Publishing ISBN 1898326 908.

Harrison, H.L., W.S. Hansen, and R.E. Zelenski, 1981. Development of a room transfer function model for use in the study of short-term transient response. ASHRAE Trans. 87(l): 461-472.

Hart, G.W., 1992. Nonintrusive appliance load monitoring, *Proc. IEEE*, 80(12) pp.1870-1891.

Haves P, Norford L K, DeSimone M., 1998. A standard simulation test bed for the evaluation of control algorithms and strategies, *ASHRAE Transactions*, 104(1) #4140 (RP-825), 460-473

Haves, P., and F. Smothers, 2002. Guidelines for Emergency Energy Reduction in Commercial Office Buildings (LBNL Report)

Haves, P., and L. Gu, 2001. Guidelines for Operation of Demand Responsive HVAC Systems (LBNL Report)

Henze G P, Dodier R H, Krarti M., 1997. Development of a predictive optimal controller for thermal energy storage systems. International Journal of HVAC&R Research, vol.3, no. 3.

Henze G P, Krarti M., 1999. The impact of forecasting uncertainty on performance of a predictive optimal controller for thermal energy storage systems. ASHRAE Trans., 105(1) SE-99-02-02, 553-561.

Hittle, D.C. and R. Bishop, 1983. An improved root-finding procedure for use in calculating transient heat flow through multilayered slabs, *Int'l J Heat Mass Transfer*, 26(11) 1685-1693.

Horn J.C., 1982. Investigation of various thermal capacitance models; M.S. Thesis, M.E. Dept, University of Wisconsin, Madison, WI.

Kahwaji, G., P. Burns and C.B. Winn, Flux, Heat Storage, and Effective Thermal Properties From in Situ Temperature Measurements in Passive Solar Components, *ASME J. Solar Energy Engineering* (Jan. 1986) New York, NY.

Katipamula S., R.G. Pratt, D.P. Chassin, Z.T. Taylor, K. Gowri, M.R.Brambley, 1998. Automated fault detection and diagnostics for outdoor air ventilation systems and economizers: methodology and results from field testing, *ASHRAE Transactions*. 105(1).

Katipamula, S. and S. Gaines, 2003. Characterization of Building Controls and Energy Efficiency Options Using Commercial Building Energy Consumption Survey, PNWD-3247.

Katipamula, S., 1989. A Study of the Transient Behavior During Start-up of Residential Heat Pumps, Ph.D. Thesis, Mechanical Engineering, Texas A&M

Katipamula, S., O'Neal, D. L., 1991. Performance Degradation During On-Off Cycling of Single-Speed Air Conditioners and Heat Pumps: Model Development and Analysis, ASHRAE Trans, 97(2) 316 – 323.

Keeney K R., Braun J E., 1997. Application of building precooling to reduce peak cooling requirements, ASHRAE Trans., 103(1) PH-97-4-1, 463-469

Keeney K., Braun J., 1996. A simplified method for determining optimal cooling control strategies for thermal storage in building mass, IJ HVAC&R Research, 2(1) 59-78.

Kintner-Meyer M. and A.F.Emery, 1995. Cost optimal analysis and load shifting potentials of cold storage equipment, *ASHRAE Trans.*, Vol.101, Part 2, #3929, 539-548

Kintner-Meyer, M., 1994. An Investigation of Optimal Sizing and Control of Air-Conditioning Systems in Commercial Buildings, PhD Thesis, M.E., U. Washington.

Kintner-Meyer, M., C. Goldman, O. Sezgen, and D. Pratt, 2003. Dividends with Demand Response, LBNL-52980

Kintner-Meyer, M., et al, 2003. How and Why Customers Respond to Electricity Price Variability: A Study of NYISO and NYSERDA 2002 PRL Program Performance, LBNL, PNNL, Neenan Associates.

Kusuda, T., 1969. Thermal response factors for multilayer structures of various heat conduction system, *ASHRAE Transactions*, Vol 75, p. 246.

Laughman, C.R., et al, 2003. Power Signature Analysis, IEEE Power and Energy, March 2003.

Lawton, L., M. Sullivan, K. Van Liere, A. Katz, and J. Eto, 2003. A Framework and Review of Customer Outage Costs: Integration and Analysis of Electric Utility Outage Cost Surveys, LBNL-54365

LBNL. http://eetd.lbl.gov/ea/EMS/EMS_pubs.html

LBNL. http://www.lbl.gov/Science-Articles/Archive/sb-EETD-internet-controls.html, November 2003.

Lee, D.K. 2003. Electric Load Information System Based on Non-Intrusive Power Monitoring, Ph.D. thesis, Massachusetts Institute of Technology.

Leeb, S. B. 1992. A Conjoint Pattern Recognition Approach to Non-Intrusive Load Monitoring. Ph.D. thesis, Massachusetts Institute of Technology.

Leeb, S. B. and J. L. Kirtley, A Transient Event Detector for Nonintrusive Load Monitoring. U.S.Patent Number 5,483,153, Issued January 1996.

Leeb, S. B., S. R. Shaw, and J. L. Kirtley, Transient event detection in spectral envelope estimates for nonintrusive load monitoring, *IEEE Transactions on Power Delivery*, vol. 7, pp. 1200-1210, July 1995.

Leeb, S.B. and S.R. Shaw, 1994. Harmonic estimates for transient event detection, IEEE UPEC.

Li, Haorong, and J.E. Braun, 2002. On-line models for use in automated fault detection and diagnosis of HVAC&R equipment, Proc. ACEEE Buildings Conference, Monterey.

Liu, Zheji and W Soedel, 1994. An investigation of compressor slugging problems, Int'l Compressor Engineering Conf. at Purdue.

Ljung, L., 1999. System Identification: Theory for the User, Prentice Hall.

Luo D., L. K. Norford, S. R. Shaw, and S. B. Leeb. 2001. Monitoring HVAC loads from a centralized location--method and field test results. ASHRAE Trans. Vol. 107.

Luo, Dong. 2001. Detection and Diagnosis of Faults and Energy Monitoring of HVAC Systems with Least-Intrusive Power Analysis. MIT PhD Thesis.

Mannion, Sean, 2004, private communication, United Refrigeration, 1 March 2004, Somerville, MA.

Mitalas, G.P. and D.G. Stephenson, Room Thermal Response Factors, *ASHRAE Transactions*, V.73, Pt.I (Jan. 1967).

Mitalas, G.P., An Assessment of Common Assumptions in Estimating Cooling Loads and Space Temperatures, *ASHRAE Transactions*, 71, Pt.II (1965).

Mitalas, G.P., Calculation of Transient Heat Flow Through Walls and Roofs, *ASHRAE Transactions*, 74, 182-188 (1968).

Morris F B., Braun J E., Treado S J., 1994. Experimental and simulated performance of optimal control of building thermal performance, ASHRAE Trans., 100(1) 3776, 402-414.

Mouen, E.D., Equivalent Thermal Properties of Hollow Elements Found via the Thermal Response Factor Method, *ASHRAE Transactions*, V.81, Pt.1 (1975).

Norford L K., Rabl A., Socolow R H., Persily A K., 1985. Measurement of thermal characteristics of office buildings, Proc. BTECC III Conf., ASHRAE SP 49, 272-288.

Norford, L. K. and R. D. Little. 1993. Fault detection and load monitoring in ventilation systems. ASHRAE Transactions, Vol. 99, Part 1, pp. 590-602.

Norford, L. K. and S. B. Leeb. 1996. Nonintrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. Energy and Buildings 24, pp.51-64.

Norford, L. K., J. A. Wright, R. A. Buswell, and D. Luo. 2000. Demonstration of fault detection and diagnosis methods in a real building. Final Report, ASHRAE 1020-RP.

Pawelski, M.J., Mitchell, J.W., and Beckman, W.A., Transfer Functions for Combined Walls and Pitched Roofs, *ASHRAE Transactions*, V.85, Part 2 (1979).

Peavy, B.A., D.M.Burch, F.J.Powell and C.M.Hunt, 1975. Comparison of Measured and Computer-Predicted Thermal Performance of a Four-Bedroom Wood-Frame Townhouse, NBS BSS 57.

Peavy, B.A., F.J.Powell and D.M.Burch, 1973. Dynamic Thermal Performance of an Experimental Masonry Building, NBS BSS 45.

Pedersen, C.O. and G.H. Youker An Experimental Study of Equivalent Thermal Properties of a Wall Section With Air Cavities, Paper AT-76-6 #5, RP102, *ASHRAE Transactions*.

Percival, D.B. and A.T. Walden, 1993. Spectral Analysis for Physical Applications, Cambridge U. Press.

Pipes, L.A., 1957. Matrix analysis of heat transfer problems, *Franklin Inst. J.*, 263(3).

Press, W.H., et al, 1992. Numerical Recipes in Fortran 77: The Art of Scientific Computing, Cambridge U. Press, Cambridge.

Pryor, D.V. and C. Byron Winn, A Sequential Filter Used for Parameter Estimation in a Passive Solar System, *Solar Energy*, V.28, No. 1, pp. 65-73 (1982).

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

Rabl, A. and L.K. Norford, 1991. Peak load reduction by preconditioning buildings at night, *Int'l J Energy Res.*, 15(781-798).

Reddy T.A., 1989. Application of dynamic building inverse models to three occupied residences monitored non-intrusively; BTECC-IV, Proceedings of the ASHRAE/DOE/BTECC/CIBSE Conference, Orlando, Florida. 654-663. ISBN 0-910110-66-2.

Reilly, R.W., P.R. Armstrong, and L.A. Klevgard, 1997. Post-Retrofit Energy Consumption—Ryazan Demonstration, Enterprise Housing Divestiture Project Report, Battelle Northwest Lab.

Richardson, R.W. and S.M.Berman, Optimum lumped parameters for modeling building thermal performance, *Construction Press* (1981).

Rossi T.M. and Braun J.E., 1993. Classification of fault detection and diagnostic methods; in IEA Annex 25, edited by Hyvarinen and Kohonen.

Rossi, T.M. and J.E. Braun, 1993. Classification of fault detection and diagnostic methods, *Building Optimization and Fault Diagnosis Concepts*, Int'l Energy Agency Annex 25

Rossi, T.M. and J.E. Braun. 1997. A statistical, rule-based fault detection and diagnostic method for vapor compression air conditioners. HVAC&R Research 3(1), pp.19-37.

Ruud, M.D., J.W. Mitchell, S.A. Klein, 1990. Use of building thermal mass to offset cooling loads, *ASHRAE Transactions* 96(2) SL-90-14-2, 820-829.

Seem J.E., Klein S.A., Beckman W.A., Mitchell J.W., 1989. Model reduction of transfer functions by the dominant root method; ASME J. Heat Transfer.

Seem, J.E., J. M. House, and R. H. Monroe. 1999. On-line monitoring and fault detection. *ASHRAE Journal*, July 1999, Vol.41, pp. 21-26.

Seem, J.E. and C.E. Hancock, A Method for Characterizing the Performance of a Thermal Storage Wall from Measured Data, Thermal Performance of the Exterior Envelopes of Buildings III, ASHRAE/DOE/BTECC, (1985) Clearwater Beach, FL.

Seem, J.E. and J.E. Braun, 1991. Adaptive methods for real-time forecasting of building electrical demand, ASHRAE Trans. 1991, vol.97, Part 1, paper# NY-910-10-3, 710-721.

Seem, J.E., 1987. Modeling of Heat Transfer in Buildings, PhD Thesis, UW, Madison.

Seem, J.E., P.R. Armstrong and C.E. Hancock, 1989. Algorithms for Predicting Recovery Time from Night Setback, ASHRAE Transactions 90(1) Atlanta, GA.

Shaw, S. R., C. B. Abler, R. F. Lepard, D. Luo, S. B. Leeb, and L. K. Norford, 1998. Instrumentation for high performance nonintrusive electrical load monitoring, *ASME J Solar Energy Engineering*, vol. 120, pp. 224-229.

Shaw, S.R. and C.R. Laughman, 2004. A Kalman-filter spectral envelope preprocessor, *IEEE Trans. Instrumentation and Measurement* (in review).

Shaw, S.R., 2000. System Identification Techniques and Modeling for Nonintrusive Load Diagnostics, Ph.D. thesis, Massachusetts Institute of Technology.

Shaw, S.R., and S.B. Leeb, 1999. Identification of Induction Motor Parameters from Transient Stator Current Measurements, IEEE Trans. Industrial Electronics, 46(1) pp.139-149.

Smith, E.G., A Method of Compiling Tables For Intermittent Heating, *Heating, Piping and Air Conditioning*, V.14, No.6 (June 1942).

Smith, E.G., A Simple and Rigorous Method for the Determination of the Heat Requirements of Simple Intermittently Heated Exterior Walls, *Journal of Applied Physics*, V.12 (August 1941).

Smith, E.G., The Heat Requirements of Simple Intermittently Heated Interior Walls and Furniture, *Journal of Applied Physics*, V.12 (August 1941).

Sonderegger, Robert C., Diagnostic Tests for Determining The Thermal Response of a House, *ASHRAE Transactions*, AT-78-4 No. 4.

Stephenson, D.G. and Mitalas, G.P., Calculation of Heat Conduction Transfer Functions for Multi-layer Slabs, *ASHRAE Transactions* V.77, Part II, (1971).

Stephenson, D.G. and Mitalas, G.P., Cooling Load Calculations by Thermal Response Factor Method, *ASHRAE Transactions*, V.73, Part I, (1967).

Stoecker, W.F., R.R. Crawford, S. Ikeda, W.H. Dolan, D.J. Leverenz, 1981. Reducing the peak of internal air-conditioning loads by use of thermal swings, *ASHRAE Transactions* 87(2) 599-608.

Stouppe, D.E. and T.Y Lau, 1989. Air conditioning and refrigeration equipment failures, *National Engineer*, 9/89.

Subbarao, K. and J.V. Anderson, Graphical Method for Passive Building Energy Analysis, *ASME JSEE* V. 105, pp. 134-141 (May 1983).

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).

Taylor Z T, Pratt R G. The effect of model simplifications on equivalent thermal parameters calculated form hourly building performance data; Proc 1988 ACEEE SS Energy Efficiency in Buildings, Vol 10, 268-278.

Tomczyk, J.,1995. Troubleshooting and Servicing Modern Air Conditioning and Refrigeration Systems, BNP.

Van Huffel, S. and J. Vandewalle, 1991. The Total Least Squares Problem: Computational Aspects and Analysis, SIAM

Walton G N. Thermal analysis research program (TARP). NBSIR 83-2655, March 1983.

Wild, C.J. and G.A.F.Seber, 2000. Chance Encounters: A First Course in Data Analysis and Inference, Wiley.

Wilson,N.W., W.G.Colborne and R.Ganesh, "Determination of Thermal Parameters for an Occupied House," KC-84-01, No.3, ASHRAE Transactions V.90, Pt.2 (1984)

Wunsch, C, 2004. Discrete Inverse and State Estimation Problems With Geophysical Fluid Applications, draft revision of The Ocean Circulation Inverse Problem, 1996, Cambridge U. Press.

Wunsch, C.I., 2003. Discrete Inverse and State Estimation Problems with Fluid Flow Applications. 3 April 2003 draft, revised chapters 1 and 2; originally published 1996 as The Ocean Circulation Inverse Problem, Cambridge U. Press, Cambridge.

Appendix A: Motor Data

RTU model No: D1EE060A25C, s/n NEHM0J9328; Manufacturer: York..

Compressor: compliant scroll type; hermetic. The suction line ID is 7/8" and the discharge is ½". Note that stator resistances are for the given winding in parallel with the series combination of the other two. Thus R ≅ 1.170Ω for individual windings.

Supply Fan: 10"W x 12"dia; 2-sided inlet; inlets are 10" diameter at throat; direct drive.

Condenser Fan: 22" diameter axial, 3-bladed.

Note that all three motors are mounted with their shafts vertical. This may be the cause of beats observed in the shaft harmonics data obtained with two or more motors running at the same time.

	compressor	Supply fan	condenser fan
Brand	Copeland		
Model	ZR57K3-TF5-250		
HP	4.75	1	0.5
LRA	123		
RLA	18	6.6	2.3
RPM	3500	1750	3550
R_{AB}	0.775		
R_{BC}	0.784		
R_{CA}	0.781	4.44	11.64
R_{START}		16.63	45.2

Appendix B. Test Summary

The parameters for each raw data file collected in April and October, 2003, are tabulated below in Tables B-1 and B-2. All files were created by the following NILM/Linux command line:

```
prep /dev/pcl1710 |chansift |prune >fnam
```

Each plot in the body of the thesis was produced from data contained in one of these files. The first column contains the start time (hh:mm) to the nearest minute and the filename is listed in the second column. The potential connection used during a given test is listed in column 3; potentials were measured from phase A to neutral (AN) phase A to phase B (AB) or phase A to phase C (AC). The type of test—"V" for transient start training or "S" for shaft harmonics sampling—is listed in column 4. The number of repetitions is listed, for "V" tests, in column 5 and the duration, for "S" tests, in column 6. The load that is active during a given test is listed in column 7.

Table B-1. April 22, 2003 Test Summary

CDT	file	PT	Test	reps	duration	Load
10:20	fantest1	AN	V	5		supply fan
	fantest1	AN	V	5		condenser fan
10:31	fantest2	AB	V	5		supply fan
	fantest2	AB	V	5		condenser fan
10:52	hotplate1	AN		1		reference load: hotplate
10:56	hotplate2	AC		1		reference load: hotplate
11:08	fantest3	AC	V	5		supply fan
	fantest3	AC	V	5		condenser fan
11:16	shaftfile1	AC	S		180	supply fan
	shaftfile1	AC	S		180	condenser fan
11:24	shaftfile2	AN	S		180	supply fan
	shaftfile2	AN	S		180	condenser fan
11:29	compfile1	AN	V	5		compressor, no pressure relief between
11:35	compfile2	AN	V	5		pressure relief between starts fans off.
11:46	compfile3	AN	V	5		compressor/condenser fan, relief between
12:00	shaftfile3	AN	S		20	condenser fan with 13g imbalance.
12:19	shaftfile4	AN	S		40	condenser fan with 22x2" tape.
12:27	shaftfile5	AN	S		100	supply fan with 2.45g imbalance
12:36	shaftfile6	AN	S		60	supply fan with 6.05g imbalance
13:55	cfanblockfile1	AN	V	5		condenser fan with 3/8 of coil face area
14:06	sfanblockfile1	AN	V	5		supply fan with return blocked and OA 70%
15:03	hotbypass1	AN	V	4		compressor/condenser fan with hot bypass
	hotbypass1	AN	V	2		compressor/condenser fan with bypass

Table B-2. October 29-31, 2003 Test Summary

File Name	Date	time	Duration /reps	V	condition	notes
parftial-fill-1	10/29	16:48	5reps	A-N	20% undercharged	
pf01a	10/29	16:48	5reps	A-N	20% undercharged	
parftial-fill-2	10/29	17:33	5reps	A-N	20% undercharged	
parftial-fill-3	10/30	11:16	5reps	A-N	normal charge	
parftial-fill-4	10/30	11:16	5reps	A-N	normal charge	
parftial-fill-5	10/30	12:06	5reps	A-N	20% overcharged	
parftial-fill-6	10/30	12:17	5reps	A-N	20% overcharged	
parftial-fill-7	10/30	13:25	5reps	A-N	1/8 turn bypass	20% overchrg.
parftial-fill-8	10/30	13:35	5reps	A-N	1/2 turn bypass	20% overchrg.
liquid-slugging-1A	10/30	13:53	5reps	A-N	5cc (20%oc)	practice session
liquid-slugging-2A	10/30	13:56	4r + 5r	A-N	5cc (20%oc)	4r w/high valve open; 5r w/high valve closed
liquid-slugging-3A	10/30	14:20	4r + 5r	A-N	5cc (20%oc)	
liquid-slugging-4A	10/30	16:48	8reps	A-N	10cc (20%oc)	
blocked-supply-fan-1	10/30	18:04		A-N		1-5v==0-2"wc
blocked-supply-fan-2	10/30	18:12		A-N		1-5v==0-2"wc
blocked-supply-fan-3	10/30	18:15		A-N		1-5v==0-2"wc
blocked-supply-fan-4	10/30	18:27		A-C	0% blockage	1-5v==0-2"wc
blocked-supply-fan-5	10/30	18:30		A-C	50% blockage	1-5v==0-2"wc
blocked-supply-fan-6	10/30	18:33		A-C	100% blockage	1-5v==0-2"wc
blocked-supply-fan-7	10/31	19:25		A-C	10% blockage; no dP	1-5v==0-2"wc
blocked-supply-fan-8	10/31	19:30		A-C	10% blockage; with dP	1-5v==0-2"wc
blocked-condenser-fan-1	10/31	20:05		A-C	35/82=39% blockage	1-5v==0-1"wc
blocked-condenser-fan-2	10/31	20:09		A-C	13.7% blockage	1-5v==0-1"wc
blocked-condenser-fan-3	10/31	20:11		A-C	0% blockage	1-5v==0-1"wc
liquid-slugging-5A	10/31	8:54	3reps	A-N	10cc w/cold cmpr	20% overchrg.
supply-fan-test-9	10/31	9:30	30s	A-C	0g	
supply-fan-test-10	10/31	9:40	10m	A-C	0g	
supply-fan-test-11	10/31	9:52	5reps	A-C	0g	
supply-fan-test-12	10/31	9:54	5reps	A-C	8g	
supply-fan-test-13	10/31	9:56	30s	A-C	8g	
supply-fan-test-14	10/31	9:57	10m	A-C	8g	
supply-fan-test-15	10/31	10:09	5reps	A-C	16g	
supply-fan-test-16	10/31	10:12	30s	A-C	16g	
supply-fan-test-17	10/31	10:13	10m	A-C	16g	
liquid-slugging-6A	10/31	10:52		A-N	10cc w/cold cmpr	20% overchrg.
Condenser-fan-test-8	10/31	11:29	5reps	A-C	0g	
Condenser-fan-test-9	10/31	11:32	30s	A-C	0g	
Condenser-fan-test-10	10/31	11:32	10m	A-C	0g	
Condenser-fan-test-11	10/31	11:50	5reps	A-C	7.78g	
Condenser-fan-test-12	10/31	11:52	30s	A-C	7.78g	
Condenser-fan-test-13	10/31	11:54	10m	A-C	7.78g	
Condenser-fan-test-14	10/31	12:02	5reps	A-C	15.56g	
Condenser-fan-test-15	10/31	12:05	30s	A-C	15.56g	
Condenser-fan-test-16	10/31	12:10	3m	A-C	15.56g	

Appendix C. Bypass Valve Flow-Pressure Test

Compressor back leakage faults were simulated by a bypass valve connected between the discharge and suction lines. A rough estimate of refrigerant flow rate through the valve was obtained by testing an identical valve. The test apparatus consisted of a nitrogen cylinder, pressure regulator, the valve under test, and a rotameter, connected in series, with the rotameter outlet discharging to the atmosphere. Flow-pressure data were obtained at each of several valve positions: $^1/_{16}$ turn open, $^1/_8$ turn open, and $^1/_4$ turn open.

At each position, the valve inlet pressure was varied from 0 to 200 psig in 40 psi increments and the flow rate indicated by the rotameter was recorded at each pressure.

The results are plotted in Figure 1.

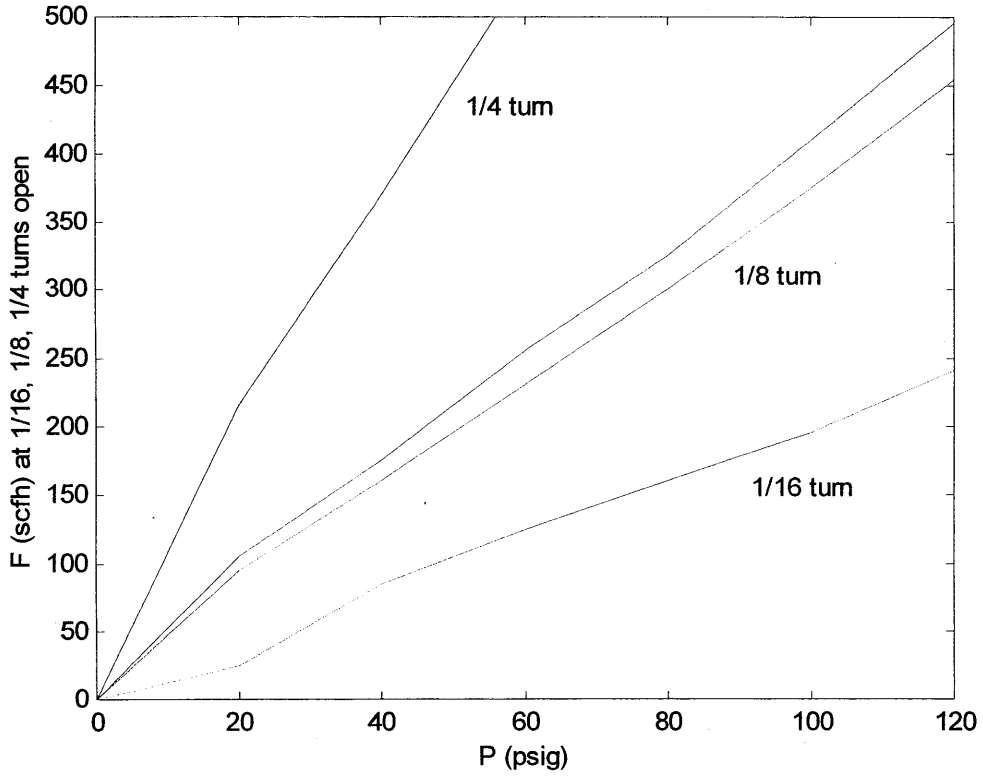

Figure 1. Flow-pressure data for bypass valve of type used in back-leakage tests.

Appendix D. Transient Evaporator-Compressor-Condenser Model

```
function simCmpr3(tf)
%2004.01.23pra simulate A/C start transient: x(1:2)=condenser,evaporator vapor mass
%future:x(3)=rotor speed (inertia); x(4:5)=cnd,evp wetted fractions; x(5:7)=Tcwall,Tewall
if 1;%modified US engineering units
   b.T2Psat =[0.0049165 0.32174  87.582  2963.2]';%se= 503, rmsy=1.63e+005, cv=0.00308
   b.rho2Psat =[2.2967 -233.46    7788 -730.25]';%se= 457, rmsy=1.63e+005, cv=0.0028
   b.T2ifgsat =[-1.3517e-6 -0.00053832 0.015297  214.01]';%se=0.562,rmsy= 495,cv=0.00113
else;%SI (kPa) units
   b.T2Psat =[0.0014 0.1232 10.6315 299.7518]';%se=24.1, rmsy=7.81e+003, cv=0.00308
   b.rho2Psat =[0.0000 -0.0434 23.2536 -34.9874]';%se=21.9, rmsy=7.81e+003, cv=0.0028
   b.T2ifgsat =[-0.0000 -0.0013 0.0177 101.1145]';%se=0.312, rmsy= 222, cv=0.00141
end; b.T2ifgsat
p.R = 0.08;     %R134 J/kgK
p.gam=0.191;  %R/cp = (cp-cv)/cp
p.k = 1.25;     %cp/cv
%p.Psuc =.4;    %R134 0.5kPa at 5C, 0.3kPa at 1C
p.Tsuc = 86;    %F NOT USED
p.Psuc = polyval(b.T2Psat,p.Tsuc)%psf NOT USED
p.rhosuc=poly2x(b.rho2Psat,p.Psuc,150);%NOT USED
p.Tdschg=212;  %F NOT USED
%CONDENSER & EVAPORATOR STUFF
p.Tca  = 70;    %F
p.Tea  = 70;    %F
p.PsaTca=polyval(b.T2Psat,p.Tca);
p.PsaTea=polyval(b.T2Psat,p.Tea);
p.rhosaTca=poly2x(b.rho2Psat,p.PsaTca,300);
p.rhosaTea=poly2x(b.rho2Psat,p.PsaTea,300);
p.Vc = 0.016;%.066; %ft3 volume of condenser & upstream to cmpr discharge port
p.Ve = 0.016;%.066; %ft3 volume of evaporator & downstream to cmpr suction port
p.UAc= 0.25;%.25; %Btu/s/F (1750Btuh/F)
p.UAe= 0.25;%.25; %Btu/s/F (1750Btuh/F)
%COMPRESSOR STUFF
p.cratio=20; %compression ratio
p.Vs= 0.0020;%.0017;   %ft3 cmprVolume/revolution
p.w = 60;       %rev/s cmpr shaft speed
p.Vdot=p.Vs*p.w/(2*pi);%NOT USED
p.Cleak=0.00001;%.0001;%cfs/psf power law coefficient
p.xleak=0.5; %power law exponent
M0=[p.Vc*p.rhosaTca p.Ve*p.rhosaTea]'
fprintf(1,'    Mc       rho       Pc       Tsat    dMdt dMcompr dMcondns dMleak  Qrej(W)
Leak(W) dMass(W)\n');% TEST
   opts=odeset('reltol',.0001);
   [t,x]=ode45('cmprODE2',[0 tf],M0,opts,b,p);
   figure(1); plot(t,x,'r:'); xlabel('time (s)'); ylabel('Me,Mc(lbm)');
   for i=1:length(t); y=cmprResim2(t(i),x(i,:),[],b,p); z(i,:)=y;end
   figure(2); plot(120*t,z(:,6),'r:');
   xlabel('time index (half-cycles)'); ylabel('Compressor Power (W)');
   figure(3); plot(120*t,z(:,7),'r:');
   xlabel('time index (half-cycles)'); ylabel('Compressor I.G. Power (W)');
   figure(4); hold on; plot(120*t,z(:,8),'r:');plot(120*t,z(:,9),'r:');
   xlabel('time index (half-cycles)'); ylabel('Pc,Pe (psia)');
   fprintf(1,'    Mc       rho       Pc       Tsat    dMdt dMcompr dMcondns dMleak
        Qrej(W)  Leak(W) dMass(W)\n');
   p.Cleak=0.000001;%.00005
   [t,x]=ode45('cmprODE2',t,M0,opts,b,p);figure(1); hold on; plot(t,x,'b-')
   for i=1:length(t); y=cmprResim2(t(i),x(i,:),[],b,p);z(i,:)=y;end
   figure(2); hold on; plot(120*t,z(:,6),'b-');
   figure(3); hold on; plot(120*t,z(:,7),'b-');
   figure(4); hold on; plot(120*t,z(:,8),'b-');plot(120*t,z(:,9),'b-');
return

function [x]=poly2x(b,y,x)
bd=polyder(b); e=y-polyval(b,x);
while abs(e)/y>0.000001
   e=y-polyval(b,x);
   dy=polyval(bd,x);
   x=x+e/dy;
end
```

```
function [y]=cmprResim2(t,x,flag,b,p)
revs=p.w/(2*pi);%compressor shaft rps = rpm/60
Mc=x(1);          %vapor resident in condenser & upstream to cmpr discharge port
Me=x(2);          %vapor resident in evaporator & downstream to cmpr suction port
rhoc=Mc/p.Vc;     %condenser density=mass*volume
rhoe=Me/p.Ve;     %condenser density=mass*volume
Pc=polyval(b.rho2Psat,rhoc);%condenser pressure
Pe=polyval(b.rho2Psat,rhoe);%condenser pressure
if Pc<p.PsaTca;Pc=rhoc*p.R*p.Tca; if Pc>p.PsaTca; 'ERROR';end; end
if Pe>p.PsaTea;Pc=rhoc*p.R*p.Tea; if Pc<p.PsaTea; 'ERROR';end; end
Tcsat = poly2x(b.T2Psat,Pc,80); %saturation temperature
Tesat = poly2x(b.T2Psat,Pe,80); %saturation temperature
dTc= max(Tcsat-p.Tca,0);  %dT<0 has same effect as dT=0: no heat rejection
dTe= min(Tesat-p.Tea,0);  %dT<0 has same effect as dT=0: no heat rejection
ifgc  = polyval(b.T2ifgsat,Tcsat);%enthalpy of vaporization, Btu/lbm
ifge  = polyval(b.T2ifgsat,Tesat);%enthalpy of vaporization, Btu/lbm
rhoms= sqrt(rhoe*rhoc);  %mean density
dp = max(Pc - Pe,0);
m1=rhoe*p.Vs*revs;
m2=dTc*p.UAc/ifgc;
m3=rhoms*p.Cleak*dp^p.xleak;
%mdot = compressorFlow - condnstnRate - leakRate
y(1)=rhoe*p.Vs*revs - dTc*p.UAc/ifgc - rhoms*p.Cleak*dp^p.xleak;
y(2)=-rhoe*p.Vs*revs - dTe*p.UAe/ifge + rhoms*p.Cleak*dp^p.xleak;
y(3) = dTc*p.UAc*1055;% HEAT REJECTION RATE (W)
y(4) = dTe*p.UAe*1055;% HEAT REJECTION RATE (W)
%y(5) = ((Pc/p.Psuc)^p.gam - 1)*p.Psuc*p.Vs*revs/p.gam*1.355;%underest P(W)745/550J/ft-lb
y(5) = -y(3)-y(4)-(Pc*m3/rhoc/p.gam)*((Pe/Pc)^p.gam - 1)*1.355; %add LEAK PWR DISSIPATION
y(6) = y(5)+(y(1)*Pc/rhoc+y(2)*Pe/rhoe)*1.355;%(W) add WORK associated w/net chg in mass
kk=(p.k-1)/p.k;
y(7) = p.Vs*revs*Pe/kk*((Pc/Pe)^kk-1)*1.355;%(W)
y(7)=max(y(7),0);
y(8)=Pc/144; y(9)=Pe/144;
fprintf(1,'%7.3f %8.2f %9.0f %9.1f %8.3f%8.3f%8.0f%8.0f',Mc,rhoc,Pc,Tcsat,y(1),m1,m2,m3)
fprintf(1,'\n')
return

function [f]=cmprODE2(t,x,flag,b,p)
f=x;
revs=p.w/(2*pi);%compressor shaft rps = rpm/60
Mc=x(1);          %vapor resident in condenser & upstream to cmpr discharge port
Me=x(2);          %vapor resident in evaporator & downstream to cmpr suction port
rhoc=abs(Mc)/p.Vc;     %condenser density=mass*volume
rhoe=abs(Me)/p.Ve;     %condenser density=mass*volume
Pc=polyval(b.rho2Psat,rhoc);%condenser pressure
Pe=polyval(b.rho2Psat,rhoe);%condenser pressure
if Pc<p.PsaTca;Pc=rhoc*p.R*p.Tca; if Pc>p.PsaTca; 'ERROR';end; end
if Pe>p.PsaTea;Pc=rhoc*p.R*p.Tea; if Pc<p.PsaTea; 'ERROR';end; end
Tcsat = poly2x(b.T2Psat,Pc,80); %saturation temperature
Tesat = poly2x(b.T2Psat,Pe,80); %saturation temperature
dTc= max(Tcsat-p.Tca,0);  %dT<0 has same effect as dT=0: no heat rejection
dTe= min(Tesat-p.Tea,0);  %dT<0 has same effect as dT=0: no heat rejection
ifgc  = polyval(b.T2ifgsat,Tcsat);%enthalpy of vaporization, Btu/lbm
ifge  = polyval(b.T2ifgsat,Tesat);%enthalpy of vaporization, Btu/lbm
rhoms= sqrt(rhoe*rhoc);  %mean density
dp = max(Pc - Pe,0);
m1=rhoe*p.Vs*revs;
m2=dTc*p.UAc/ifgc;
m3=rhoms*p.Cleak*dp^p.xleak;
%mdot = compressorFlow - condnstnRate - leakRate
f(1)=rhoe*p.Vs*revs - dTc*p.UAc/ifgc - rhoms*p.Cleak*dp^p.xleak;
f(2)=-rhoe*p.Vs*revs - dTe*p.UAe/ifge + rhoms*p.Cleak*dp^p.xleak;
return
```

```
function PVcurves
%R134 T(C), P(kPa)
T=[-100 -90 -80 -70 -60 -30  -20   -10    0     10    20    30    40     50   60   70   80
90 100 101.06]';
P=[.56 1.52 3.67 7.98 15.91 84.38 132.73 200.6 292.8 414.6 571.7 770.2 1016.6 1318 1682 2117
2633 3244 3972 4059]';
T3 = -103.3; P3=.39; Tc=101.06;
title('R134');
%plot(log(T-T3),log(P)); ylabel('log(P[kPa])'); xlabel('log(T-T3[K])');
plot((T-T3),log(P)); ylabel('log(P[kPa])'); xlabel('T-T3[K]');
%b=fitPv(log(P-P3),log(T-T3));
b=fitPv(log(P-P3),(T-T3));

function b=fitPv(P,T)
x(:,1)=0*T+1;
x(:,2)=T;
%x(:,2)=T.*T;
x(:,3)=sqrt(T);
[b,bi,r,ri,stats]=regress(P,x);
hold on; plot(T,x*b,'r')

return
```

Appendix E. LEES Test Zones: Materials, Construction.

The floor is concrete cast in place on grade, thickness unknown, finished in asphalt tile (12" squares).

The shop has a floor area of 740 ft^2. The north and west walls are light partitions of standard hollow-core concrete/masonry units (CMU; 4" nominal, 3.5" actual thickness) to 56" above finished floor (AFF) and glass to 104" AFF. The shop is entered by a 60 x 84" double door on its west wall.

The dividing wall between the two zones is CMU (4" nominal, 3.5" actual thickness) extending from floor to top of ceiling plenum (124"AFF). A 60 x 84" double door provides access between the two rooms. The 46" square columns at either end of the east wall extend 42" into the proto-lab space and are flush with the partition on the shop side.

The proto-lab has a floor area of 650 ft^2. The north wall is CMU (4" nominal, 3.5" actual thickness) extending from floor to ceiling (104"AFF) with a72 x 90" double door leading to the main lab. The east wall is 8" CMU to 56" AFF and metal frame windows, facing a parking lot, to the ceiling. The window units are 64 x 88" in 16" deep concrete mullions. The top mullion is 12" high and the columns between windows are 16" wide. The 46" square columns at either end of the east wall extend 29" into the space and 16" to the outside.

The south wall of both rooms appears to be of 8" CMU; same for the shop pantry wall. The pantry is entered by a 36x 90" door.

The shop ceiling is suspended acoustic tile (SAT) in a standard 24" grid, suspended 104" AFF. The proto-lab ceiling is suspended acoustic tile (SAT) in a standard 24" grid, suspended 94.5" AFF.

The UA of the dividing wall and double doors, assuming no airflow through gaps and other leaks, can be estimated. The controlling uncertainty in this estimate is film resistance. We have used a combined convection radiation coefficient for vertical interior walls of xxx. The R-value of a 4" hollow-core CMU wall is taken to be 1.51 $ft^2F/Btuh$. The R-value of a solid-core wooden door is taken to be 1.41 $ft^2F/Btuh$.

The UA of the portion of the proto-lab wall exposed to the main lab can be estimated as follows. We have used a combined convection radiation coefficient for vertical exterior walls of yyy. The R-value for the window units is taken to be $0.02 + 1/xxx + 1/yyy$ $ft^2F/Btuh$.

The UA of the portion of shop wall exposed to the main lab can be estimated based on the R-value of the fixed glass being zzz $ft^2F/Btuh$.

Appendix F. ISD Building Description

The ISD Building is a 70,000 ft^2 structure built in 1973 with a constant volume dual duct HVAC system. The plant consists of two boilers and two four-stage reciprocating chillers. Two NILMs have been installed, one at the service entrance and one at the central fan/chiller motor control panel. Thermal instrumentation includes temperature and humidity of return-, mixed-, hot-deck and cold-deck air. Fan inlet pressure taps measure flow rates at the supply and return fans while thermal anemometers measure "mass velocity" (labeled "rhoV" in the plots) to determine the damper-controlled division of supply air between hot- and cold-decks.

Table F-1. ISD Building HVAC Electrical Loads.

		Circuit & Motor Ratings			
Name	Function	Ckt	Amps	HP	rpm
P1	Chilled water pump 1	30		7.5	
P2	Chilled water pump 2	30		7.5	
P3	Condenser water pump 1	30		10	
P4	Condenser water pump 2	30		10	
P5	Hot water pump 1	15		1.5	
P6	Hot water pump 2	15		1.5	
CP1	Circulation pump 1			0.5	
SF	Supply fan	200		60	
RF	Return fan	60		20	
C1.1	Chiller stage 1 compressors				
C1.2	Chiller stage 2 compressors				
C1.3	Chiller stage 3 compressors				
C1.4	Chiller stage 4 compressors	300			
TF1A	Toilet fan 1A				
TF1B	Toilet fan 1B				
TF2	Toilet fan 2				
EF1	Exhaust fan 1			0.5	
EF2	Exhaust fan 2			0.5	
EF3	Exhaust fan 3				
EF4	Exhaust fan 4			0.33	
EF5	Exhaust fan 5			0.25	
EF6	Exhaust fan 6			0.75	
CAC1	Control air compressor 1			1	
CAC2	Control air compressor 2			1	
	Strip heater, 13kW				

Internal Services Department (ISD) Building
BIS#7022; LACO#5863
1100 N Eastern Avenue
Los Angeles, CA 90063
Function: Offices
Completion date: 1 May 1973
Lighting retrofit: 29 November 1996
Floor area net/gross (ft2): 58,826/45,646
Operating hours: 11 (Monday-Thursday)
Heat: two 1.1 Mbtuh (input) gas-fired hot water boilers
Cooling: two 2-stage reciprocating chillers
Distribution: Dual-duct constant volume system; built-up AHU
Electric Utility: SCE, I-6 (interruptible) rate
Contact: Ron Mohr

Table F-2. ISD Building K20 Channel Assignments (all are A-phase currents).

K20 Channel	Name	Function	Circuit & Motor Ratings			CT FS amps
			Ckt	Amps	HP	
2	P1	Chilled water pump 1	30		7.5	25
2	P2	Chilled water pump 2	30		7.5	
3	P3	Condenser water pump 1	30		10	50
3	P4	Condenser water pump 2	30		10	
7	P5	Hot water pump 1	15		1.5	10
7	P6	Hot water pump 2	15		1.5	
6	CP1	Circulation pump 1			0.5	
0	SF	Supply fan	200		60	150
5	RF	Return fan	60		20	50
1	C1B+2B	Chiller stages 3&4 (both chillers)				300
4	C1A+2A	Chiller stages 1&2 (both chillers)				
6	TF1A	Toilet fan 1A				10
6	TF1B	Toilet fan 1B				
6	TF2	Toilet fan 2				
6	EF1	Exhaust fan 1			0.5	
6	EF2	Exhaust fan 2			0.5	
6	EF3	Exhaust fan 3				
6	EF4	Exhaust fan 4			0.33	
6	EF5	Exhaust fan 5			0.25	
6	EF6	Exhaust fan 6			0.75	
6	CAC1	Control air compressor 1			1	
6	CAC2	Control air compressor 2			1	
6		Strip heater, 13kW				
8	LAC	G,1,2 floor lighting (w/retrofit)				150
9	DSA	General plug loads				300
10	MCAB	Penthouse HVAC loads (CT)				100
11	Lift1+2	Hydraulic elevators				150
12	DEA	Emergency loads				150
13	LAG	3rd floor lighting (no retrofit)				50
14	Main	Service entrance (5A piggy back CT)				1500
15	MCA	Main HVAC loads				600

NILM Load Data.

HVAC loads were cycled on and off, one load at a time, while recording at 120Hz the power, P, and reactive power, Q, fed to the MCC. Second and higher harmonic spectral envelopes were not recorded during these training tests. When interpreting the recorded signals it is important to keep in mind that the spectral envelope is not an instantaneous signal but rather a one-cycle moving average of the selected harmonic (in this case the 60Hz fundamental) of the power signal.

Each pump was turned on, allowed to run for about 10 seconds, turned off, and left off for about 10 seconds. This standard *on/off* sequence was repeated five times for each pump. The supply and return fans were operated longer because they have longer start transients. The chillers also produced long *on* cycles because they have a programmed multi-stage sequence control that requires operation at a given stage for a minimum time before switching to the next higher stage.

V-sections were fit to all start and stop transients by the interactive Matlab program *vtrain*. The program locates events and allows the user to define the section boundaries and averaging weights used to aggregate points taken from all (or a subset) of the observed transients.

The mean values of P and Q and their standard deviations, computed for data points recorded after each load had reached steady state, are presented in Table F-3. The steady-state means and standard deviations serve two purposes. The real (P) and reactive (Q) components are used to identify off transitions and the real component is used to estimate energy used by a given load during each of its operating periods. Standard deviations are generally 1% of average load. However, the standard deviations for the two small pumps (P5 and P6) are about 3% of average load. These numbers suggest a simple "base-value-plus-percent-of-signal" noise model. Note that two sets of lead-lag pumps have significantly different steady-state loads: P1,P2 (chilled water) and P5,P6 (heating water).

Table F-3. Loads recorded during ISD NILM training (multiply P and Q values by 20).

		Circuit & Motor Ratings				P (W)		Q (VAR)			Change		
Name	Function	Ckt	Amps	HP	Rpm	Average	s.d.	Average	s.d.	PF	P(W)	Q(VAR)	PF
P1	ChW	30		7.5		287.7	2.9	-344.9	2.5	0.64			
P2	ChW	30		7.5		268.1	2.7	-202.5	2.3	0.80			
P3	CndW	30		10		488.7	4.3	-336.7	2.6	0.82			
P4	CndW	30		10		465.3	5.0	-332.8	2.8	0.81			
P5	HW	15		1.5		69.7	2.3	-87.3	2.2	0.62			
P6	HW	15		1.5		77.8	2.1	-98.8	2.2	0.62			
S+R	Fans					1869	42	-1456	13	0.79			
SF	fan	200		60		1460	35	-1092	7	0.80	409	-364	0.75
RF	Fan	60		20		476	28	-391	10	0.77			
	Fans					2002	40	-1420	26	0.82			
C1.1	Chiller					2775	45	-2115	26	0.80	773	-695	0.74
C1.2	Chiller					4162	47	-3470	23	0.77	1387	-1354	0.72
C1.3	Chiller					5581	55	-5023	22	0.74	1419	-1553	0.67
C1.4	Chiller	300				6762	48	-6475	17	0.72	1181	-1453	0.63

ISD Floor, Wall and Window Areas

The ISD is built on a 24'x24' structural grid 10 bays wide (the long side running almost east to west; the "south" wall normal actually points ~30 degrees east of south) and 4 bays deep (the short side). The main footprint is thus a bit over 240'x 96'≈ 23,000 ft^2 . The basement is only 3 bays deep (floors 1 and 2 extend 1 bay further north) giving it a ~240'x 72'≈ 17,300 ft^2 footprint. The basement is essentially windowless. Windows on the main floors (1 and 2) are placed every 4 feet to fit the 24-foot grid and the floor-to-floor height is 14 feet. The windows are operable but usually remain closed. The third floor (penthouse) has no windows except on the 50-foot east wall, looking on to a rooftop terrace, which is essentially all glass with a 20" (clear) transom strip of fixed glazings above an 80" (clear) main span of sliding doors and fixed glazings.

There are roof doors at the tops of both stairwells, one at the north-east corner (north-most point) and one near the center of the west end of the building (just west of and adjacent to the elevators).

Window types are described in Table F-3 and their aggregate surface areas are summarized by location in Tables F-4 and F-5.

Table F-4. ISD Window Types

Type	Dimensions (H x W, in.) Finished	Clear	Clear Area (ft^2)	Mullion Area (ft^2)
A (single casement)	66.0 x 36.5	61.5 x 31.25	13.35	3.38
B (narrow fixed)	108.0 x 17.0	64.7 x 13.5	9.44	3.31
		36.0 x 13.5		
C (wide fixed)	108.0 x 99.0	64.7 x 95.5	66.78	7.47
		36.0 x 95.5		
D (door)	83.5 x 36.0	66.7 x 24.8	11.35	9.53
E (narrow transom)	24.0 x 72.0	20.0 x 70.5	9.79	2.21
G (double casement)	66.0 x 22.5	61.5x18.3	7.82	2.50
I (sliding door)	83 x 36	80 x 33	18.33	2.42
J (door-height, fixed)	83 x 40	80 x 37	20.56	2.50
K (transom)	23 x 40	20 x 37	5.14	1.25

Table F-5. ISD Window Inventory

Floor	Wall	Type	Qty	Clear (ft^2)	Mullion (ft^2)
1	N	A (single casement)	54	720.7	182.5
1	N	B (narrow fixed)	2	18.9	6.6
1	N	C (wide fixed)	1	66.8	7.5
1	N	D (door)	4	45.4	38.1
1	N	E (narrow transom)	2	19.8	4.4
1	S	A (single casement)	60	800.8	202.8
1	W	H (double casement)	5	78.2	25.0
2	N	A (single casement)	60	800.8	202.8
2	S	A (single casement)	60	800.8	202.8
2	W	H (double casement)	5	78.2	25.0
3	E	I (sliding door)	2	36.66	4.8
3	E	J (door-height, fixed)	12	246.7	30.0
3	E	K (transom)	14	72.0	17.5

Table F-6. ISD Window and Wall Areas

		Wall Dimensions (ft)	Gross(ft^2)	Clear(ft^2)	Mullion(ft^2)
	N	2x240x14 + 110x11	7,930	1672.3	441.9
	E	2x96x14+72x14+50x11	4,246	355.4	52.3
	S	3x240x14 + 110x11	11,290	1601.6	405.6
	W	2x96x14+72x14+50x11	4,246	156.4	50.0
	Roof	240x96	23,040		
	ground	240x(96+12)	25,920		

Condition of HVAC Equipment. The air-side instrumentation was designed to measure air-side zone sensible heat balance within 10% (3% repeatability) for a properly functioning CV system[31]. However the CV system is not in great condition as outlined below.

1) air distribution. Some access doors have been found open and there are probably leaks in hot and cold decks as well. The condition of zone control dampers is not known; damper leakage at full shut position should be less than 2% of design flow to ensure reasonable efficiency. The cooling coil condensate pan has rusted through so that water puddles on, and re-evaporates from, the floor of the mixed air plenum changing the latent/sensible load ratio at the coil. This affects system performance and performance measurement but the effects are difficult to quantify.

2) Zone thermostats and controls dampers do not communicate with the EMCS. There is therefore no way to monitor or change zone temperatures and damper positions remotely. In particular, supervisor

[31] CV system serving ISD is inherently inefficient but this in itself does not affect model identification

must be able to such that Controls that could reduce simultaneous heating and cooling by resetting hot and cold deck temperatures are therefore not possible with the existing control hardware. The chilled water temperature and hot water temperature control loops have a similar configuration that will not support more efficient control and coordination of supply air and water temperatures.

3) Upgrade central plant equipment. Install 2-speed chilled water pumps or replace with VSD pumps. Install 2-speed condenser water pumps or replace with VSD pumps. Install 2-speed cooling tower fans or replace with VSD fans.

4) Upgrade central plant controls. Replace zone temperature control loops so that setpoints can be centrally controlled and local loop temperature and damper positioning commands are available to the supervisory control system. Replace pneumatic hot and cold control loops with DDC. Replace boiler temperature controls with DDC so that boiler temperature can be reset. Replace chilled water temperature control loop with DDC so that it can be reset by the supervisor. Reprogram (or replace with appropriately programmable) compressor sequence controls so that supervisor can select local loop control on chilled water or cold deck temperature or over-ride for modes such as optimal start or curtailment; compressors should come on one at a time alternating between the north and south chiller units and commissioning should check that when one chiller is off line for service, the appropriate number of compressors come on in the other chiller. Install DDC local loop controls to make chilled water, condenser water, and tower air flow rates approximately proportional to chiller load when in fail-safe mode but program supervisor to control operation for maximum whole plant efficiency in normal operating mode. Cooling tower air and water-side discharge temperatures and cooling coil air- and water-side inlet and outlet temperatures must be available to the supervisor.

5) Economizer control is currently based on a mixed air temperature control loop and a minimum outdoor temperature set point. The latter is set up to command minimum outside air when outside air is above 65F regardless of return air temperature. There is no provision to go to 0% outside air during morning warmup in winter. It appears that an enthalpy control loop based pneumatic sensors was in operation when the building was new.

6) There is no provision to shut off the hot deck coil in warm weather or to shut off the cold deck coil in very cold weather.

7) The air distribution system should be re-balanced to provide uniform supply air flow over conditioned floor area, correct building pressure, and correct ventilation air to each zone.

If a VAV retrofit is pursued, the supervisor must be able to reset hot and cold deck temperatures such that simultaneous heating and cooling are minimized or, alternatively, total cost of central plant operation is minimized. In addition, supply air static pressure reset control and demand ventilation control should be provided so that fan power can be controlled by the supervisor in conjunction with optimal operation of other central plant.

Appendix G. Zubkova Building: Materials, Construction, and Thermal Properties.

Russian apartment buildings of precast panel construction typically have 35-cm-thick exterior walls. A standard 1.5-m-high by 2-m-wide window unit contains two double-glazed operable sash panels of nominal 1.0 x 0.5 and 0.5 x 0.5m size and a 1.5m square fixed sash. Small bedrooms and kitchens sometimes use a smaller (1.5 x 1.5 m) standard window. Panels are attached by welding metal tabs that protrude at standard locations along their edges. After welding, the panel joints are filled with grout, which is sometimes covered with mastic or a similar sealant. A floor deck typically comprises an array of 150-mm precast concrete panels. Over this is an air gap formed by 15-25mm furring strips covered with ~20mm wood planks or particle board and linoleum.

Each apartment typically has a balcony accessed from the main room by a partly or fully double-glazed door. The main room has an additional window in rare cases. All other rooms, including the kitchen have a single standard window unit. The bathroom is an exception (bathrooms are never situated on an exterior wall). There is a hot water radiator under each window. Attics are of two types: "cold" and "warm"—the latter type is heated by apartment exhaust air passing through on its way to a roof stack.

Except for the double-glazing of windows and the reported use of light aggregates in exterior panels, most Russian building envelopes are uninsulated. A one-room apartment has a kitchen, bathroom, and one main room and a floor area of 30 to 35 m2; a two-room apartment has one additional room and a floor area of 48 to 54 m2; and a three-room apartment adds yet another room for a floor area of 67 to 73 m2. Heat, service, and wastewater risers are exposed in apartments as they traverse the vertical extent of a building. Electrical risers, on the other hand, pass through a chase and the treatment of floor penetrations varies. Doors and windows, passive ventilation channels, smoke control channels, electrical chases, and stair and elevator shafts are the main paths available for air movement into, out of, and through the building.

Thermal parameters of the 60-unit apartment at 221 Zubkova, Ryazan, are summarized below.

The front wall faces 17.5° E of North (NNE) and is 74m long.
The gable ends are 11m long and are mostly opaque with just one balcony window/door unit per floor
The balance of window area is about equally distributed between front and back walls.

Envelope areas and U-values by component are from building plans, a field audit, and U-value tests:
kitchen windows $1.8m^2 * 50 = 90\ m^2$
large windows $2.8m^2\ 100 = 280\ m^2$
balcony window/door units of $3.72m^2\ 40 = 148.8\ m^2$
common space windows of $1.85m^2\ 32 = 59.2\ m^2$
window U-value (double glazed 1.5cm air gap) assumed to be 3.06 $Wm^{-2}K^{-1}$.
Two-thirds of balconies assumed to be covered U-value of 2.16 $Wm^{-2}K^{-1}$.
Common space windows are assigned a higher U-value of 3.69 $Wm^{-2}K^{-1}$.
Window Total $578m^2$ $UA=1716.7WK^{-1}$.

Perimeter: 176.5m; height: 2.55*5; net wall area: $2250 - 578 = 1672\ m^2$.
Median wall U-value (Armstrong 2002): 2.424 $Wm^{-2}K^{-1}$, giving a wall UA of 4052 WK^{-1}.

Attic floor/basement ceiling median U-value (Armstrong 2002): 2.651 $Wm^{-2}K^{-1}$; area: $811m^2$. Attic

and basement buffer spaces are heated by ventilation exhaust and hot water pipe losses. Assumed buffer-space temperature differences as fraction of total temperature difference are 0.1 for interior to attic and 0.5 for interior to basement. Based on these assumptions:

effective attic UA: 301 WK^{-1}

effective basement UA: 1075 WK^{-1}.

Total envelope conductance is the sum as recapitulated below:

Window area = 578m^2	UA = 1717WK^{-1}
Wall area: 2250 – 578 = 1672 m^2;	UA = 4052 WK^{-1}
Attic (effective UA)	UA = 301 WK^{-1}
Basement (effective UA)	UA = 1075 WK^{-1}
Total envelope conductance:	UA = 7145 WK^{-1}

Partitions are precast concrete, thickness 10cm (4in). The wall between apartment blocks is assumed equivalent to a gable end wall. Each block has 4 transverse partitions and 2.5 longitudinal partitions. The volume per floor per block is therefore about 86m^3. With 4 blocks and 5 floors the total volume is 1725 m^3 and the corresponding participating mass is about 3,320,000 kg. The building UA is 7-12kW/K giving a maximum time constant of 60-110hr.

Ryazan, Russia. Elevation: 558 feet Latitude: 54° 37′N Longitude: 039° 43′E

Temperature Stats (F)

Average (19 years)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	De
42	16	18	27	44	57	64	66	63	53	41	29	21

Highest Mean Daily (19 years)

95	38	45	60	77	88	92	95	90	81	72	54	46

Lowest Mean Daily (19 years)

35	11	12	22	37	48	56	58	55	46	36	26	16

Extreme High

95	38	45	60	77	88	92	95	90	81	72	54	46

Extreme Low

-36	-31	-20	-9	18	32	38	43	37	27	10	-10	-36

Average Number of Days With Precipitation Years on Record: 19

195	23	17	16	13	12	13	13	12	14	16	22	24

Average Number of Days Above 80F/26C Years on Record: 19

16	---	---	---	---	2	5	6	3	---	---	---	---

Average Number of Days Above 65F/18C Years on Record: 19

99	---	---	---	3	15	23	27	22	8	1	---	---

Average Number of Days Below 32F/0C Years on Record: 19

160	30	27	27	10	---	---	---	---	1	11	24	30

Average Number of Days Below 0F/-17C Years on Record: 19

19	8	6	1	---	---	---	---	---	---	---	---	4

Appendix H. Model Canonical Forms.

Inversion, i.e., transformation between the CRTF (6) and iCRTF (8) models is a frequent and sometimes confusing operation (see simulation code listings in Appendices I and J). A set of four canonical forms have been found to be useful. The coefficient index notation (i:n) indicates the range of a coefficient vector's association with past samples. Thus, for example:

$\phi_z(0{:}n) = [\phi_{z,0}\ \phi_{z,1} \cdots \phi_{z,n}]$ $\quad$ $\phi_z(1{:}n) = [\phi_{z,1}\ \phi_{z,2} \cdots \phi_{z,n}]$

$\theta_z(0{:}n) = [\theta_{z,0}\ \theta_{z,1} \cdots \theta_{z,n}]$ $\quad$ $\theta_z(1{:}n) = [\theta_{z,1}\ \theta_{z,2} \cdots \theta_{z,n}]$

$\theta_w(0{:}n) = [\theta_{w,0}\ \theta_{w,1} \cdots \theta_{w,n}]$ $\quad$ $\theta_w(1{:}n) = [\theta_{w,1}\ \theta_{w,2} \cdots \theta_{w,n}]$

The range of lags may also be expressed in operator notation by the backshift operator with the range expression appended as a superscript:

$B^{i:n}(\phi) = [\phi_i\ \phi_{i+1} \cdots \phi_n]$

$B^{i:n}(\theta) = [\theta_i\ \theta_{i+1} \cdots \theta_n]$

$B^{i:n}(q) = [\ q(t\text{-}i)\ q(t\text{-}i\text{-}1) \cdots q(t\text{-}n)]$

$B^{i:n}(T) = [T(t\text{-}i)\ T(t\text{-}i\text{-}1) \cdots T(t\text{-}n)]$

Equations (6) and (8) were developed for the simplest case involving just one exogenous temperature, e.g. a wall model. Here we extend the subscript used with temperature coefficients to indicate which temperature is meant: z refers to the inside wall surface or zone temperature and $w = 1{:}W$ refers to exogenous temperatures or sol-air temperatures 1 through W.

Note that the zero-lag term of the first exogenous temperature is absent from all of the forms. Its value is given by the u-value constraint (7).

CRTF form

$x_c = [Q(t\text{-}k,\ k{=}1{:}n)\ T_z(t\text{-}k,\ k{=}0{:}n)\ T_1(t\text{-}k,\ k{=}1{:}n)\ T_2(t\text{-}k,\ k{=}0{:}n) \cdots T_W(t\text{-}k,\ k{=}0{:}n)]$

$b_c = [\phi(1{:}n)\ \theta_z(0{:}n)\ \theta_1(\ 1{:}n)\ \theta_2(0{:}n) \cdots \theta_w(0{:}n)]$ with $\phi(0) = -1$ implicit

$Q = x_c\ b_c$

iCRTF form

$x_i = [Q(t\text{-}k,\ k{=}0{:}n)\ T_z(t\text{-}k,\ k{=}1{:}n)\ T_1(t\text{-}k,\ k{=}1{:}n)\ T_2(t\text{-}k,\ k{=}0{:}n) \cdots T_W(t\text{-}k,\ k{=}0{:}n)]$

$b_i = [\phi(0{:}n)\ \theta_z(1{:}n)\ \theta_1(\ 1{:}n)\ \theta_2(0{:}n),\cdots,\ \theta_w(0{:}n)]$ with $\theta_z(0) = -1$ implicit

$T_z = x_i\ b_i$

Un-normalized complete form

$x_u = [Q(t\text{-}k,\ k{=}0{:}n)\ T_z(t\text{-}k,\ k{=}0{:}n)\ T_1(t\text{-}k,\ k{=}1{:}n)\ T_2(t\text{-}k,\ k{=}0{:}n) \cdots T_W(t\text{-}k,\ k{=}0{:}n)]$

$b_u = [\phi(0{:}n)\ \theta_z(0{:}n)\ \theta_1(\ 1{:}n)\ \theta_2(0{:}n),\cdots,\ \theta_w(0{:}n)]$ with no particular scaling imposed

$0 = x_u\ b_u$

Q-normalized complete form

$x_n = [Q(t\text{-}k,\ k{=}0{:}n)\ T_z(t\text{-}k,\ k{=}0{:}n)\ T_1(t\text{-}k,\ k{=}1{:}n)\ T_2(t\text{-}k,\ k{=}0{:}n) \cdots T_W(t\text{-}k,\ k{=}0{:}n)]$

$b_n = [\phi(0{:}n)\ \theta_z(0{:}n)\ \theta_1(\ 1{:}n)\ \theta_2(0{:}n),\cdots,\ \theta_w(0{:}n)]$ with $\phi(0) = -1$ explicit.

$0 = x_n\ b_n$

It is apparent that a T_z-normalized form can also be immediately defined. However, one is more likely, in practice, to use an iCRTF form obtained, on the fly, from the CRTF form or from either one of the complete forms.

Note: we insert the t-ratio from the residual error norm when converting from CRTF or iCRTF canonical form to either of the complete forms (see function rootstu(bs,n)).

Appendix I. Script for Model Identification with Constrained Roots

```
%testctf7x.m 2003.06.07 Hittle-Bishop constraint: via 0<x<1; xbx
%is search vector, rbr becomes an intermediate vector in resimctfxbx()
%testctf7.m 2003.06.06 2-step search at each increment of model order
%testctf6.m 2003.01.12 fit progressively higher order models
if 1;%---generate sine responses by qctf() and tctf() top p.14-----
 bb=[.53355063E-4 .013730926 .064549573 .037668124 .33696494E-2 .38707349E-4];
 cc=[5.9708847 -10.761687 5.967169 -1.1142596  .057851433 -.54859184E-3];
 dd=[1.0 -1.2165741 .38310083 -.029577257 .37410267E-3 -.43038736E-6];
 crrx=sum(bb)/sum(cc); cc=crrx*cc;
 cof=[bb;-cc;dd];
 np=length(bb);n=np-1; nnp=n+np;
 m=4*12;%30;
 tz=[1:m]'; tz=0*tz; tz(1:m/2)=1+tz(1:m/2);
 tx=[tz(1+m/4:m);tz(1:m/4)];
 qstart=0*tz(1:n);tx=[qstart;tx];tz=[qstart;tz];
 q=qctf(cof,qstart,tx,tz);
 figure(1); plot(tx);hold on;
 plot(tz,'g');plot(q,'r');
 plot([0;length(tx)],[0;0],'k');
 [be,ce,de,se]=lsqctf(5,q,tx,tz);
 fprintf(1,'norm(bb-be)=%11.4g\n',norm(bb-be'))
 disp(bb);disp(be')
 fprintf(1,'norm(cc-ce)=%11.4e\n',norm(cc-ce'))
 disp(cc);disp(ce')
 fprintf(1,'norm(dd-de)=%11.4g\n',norm(dd-[1,-de']))
 disp(dd);disp([1,-de'])
end
if 1;%---test manipulation of CTF coefficients and roots bot p.14-----
 np=length(bb);n=np-1;
 b=[-cof(3,2:np),cof(1,:),-cof(2,:)]';
 dbcof=[roots(cof(3,:));cof(1,:)';roots(cof(2,:))];
 disp(1./log(1./roots(cof(3,:))'));%tau/timestep q  (dd)
 disp(1./log(1./roots(cof(2,:))'));%tau/timestep Tz (cc)
 disprbr(dbcof)
 phi=poly(dbcof(1:n));
 bb=dbcof(np:n+np)';
 thz=poly(dbcof(np+np:n+n+np));
 thz=(sum(bb)/sum(thz))*thz;
 newb=[-phi(2:np),bb,thz]';
 disp('|newb-b|=');disp(norm(newb-b))
end
x=zeros(m,3*np);
for i=1:np
   x(:,i)= q(1+np-i:m+np-i);
   x(:,np+i)=tx(1+np-i:m+np-i);
   x(:,np+np+i)=tz(1+np-i:m+np-i);
end;
y=x(:,1);
hold off
xx=x(:,2:3*np); fig=1
%check error norm given by the correct model----------
figure(2);
rbr=dbcof;
xbx=rbr;%xbx(1)=rbr(1)
xbx(nnp+1)=rbr(nnp+1);
xbx(1)=rbr(1)/rbr(nnp+1);
for i=2:n;
   xbx(nnp+i)=rbr(nnp+i)/rbr(i-1);
   xbx(i)=rbr(i)/rbr(nnp+i);
end
fig=.99
if 1
   f=resimctfxbx(xbx,xbx,xx,y);  plot(f); hold on; disp('xbxnorm');norm(f)
   opti=optimset('maxfune',8000*length(dbcof),'maxiter',4000,'tolfun',1e-11,'tolx',1e-
11,'diffmin',1e-9);
   xbx=fig*xbx; xbx(nnp+1)=xbx(nnp+1)*1.005/fig
   b = 1. + 0*xbx; lb=.1*b; ub=.9*b;
   lb(np:nnp)=deal(0);ub(nnp+1)=xbx(nnp+1);
```

```
   [xbx,sse,e]=lsqnonlin('resimctfxbx',xbx,lb,ub,opti,xbx,xx,y+.2*randn(length(y),1));
   figure(3);plot(e)
   disp([norm(rbr-dbcof),sse,e'*e,sqrt(sse/length(e))])
   rbk=xbx;%rbk(1)=xbx(1) DELETE THIS STUFF
    rbk(1)=xbx(1)*rbk(nnp+1);
    for i=2:n;rbk(nnp+i)=xbx(nnp+i)*rbk(i-1);rbk(i)=xbx(i)*rbk(nnp+i);end
   disprbr(rbk)
   e=resimctfxbx(xbx,xbx,xx,y);
   disp('rbr-dbcof|,sse,sse,rms=')
   disp([norm(rbr-dbcof),sse,e'*e,sqrt(sse/length(e))])
   alle=e;figure(4);plot(e)
end
return

function f=resimctfxbx(xbxr,xbx,x,y)
%this uses past predicted y=ye; also try with past values of y data
% f=resCTFrbr(rbr,x,y) returns m residuals given parm estimate,
% b=[roots(phi);thx;roots(thz)], and data x(mxlen(b)),y(mx1).
% Columns of x correspond to [phi',thx',thz'] elements. See testctf3
np=(length(x(1,:))+1)/3; n=np-1; nnp=n+np;
if length(xbx)==length(xbxr);%for 2-step t-consts search when incrementing model order
   xbx=xbxr;
else
   xbx(n-1:n)=xbxr(1:2);
   xbx(np:n+np)=xbxr(3:2+np);
   nr=length(xbxr);
   xbx(n+np+n-1:n+np+n)=xbxr(nr-1:nr);
end
rbr=xbx;%rbr(1)=xbx(1)
rbr(1)=xbx(1)*rbr(nnp+1);
for i=2:n;
   rbr(nnp+i)=xbx(nnp+i)*rbr(i-1);
   rbr(i)=xbx(i)*rbr(nnp+i);
end
dbc=[rbr;0];
p=poly(rbr(1:n));
dbc(1:n)=-p(2:np)';
p=poly(rbr(1+np+n:n+np+n));
dbc(1+np+n:np+np+n)=p'*sum(rbr(np:np+n))/sum(p);
ye=0*y; ye(1)=x(1,:)*dbc;
for k=2:length(y);
   x(k,2:n)=x(k-1,1:n-1);
   x(k,1)=ye(k-1);
   %x(k,1)=y(k-1);
   ye(k)=x(k,:)*dbc;
end
f=y-ye;
```

Appendix J. Model Identification by Ordinary and Hybrid Least Squares

```
>>agglees9('lees1oct',1,2000,890,919,1210,600,[2],[3]);
function b=aggLEES9(fnam,fLtg,mcp,jhead,jtail,khead,ktail,nlist,malist);
%train/test sets from fnam between head and tail line#s;
%fit model orders in nlist, e.g. [0 1 2 3] and
%time-aggregations in malist, e.g. [1 2 3 6]
%fLtg, mcp for problem specific heat rate calcs
%aggLEES9.m 20031218pra use canonical order within x,b
data=load([fnam '.txt']); size(data);
%1:8 lotusdate TfanIn1 TfanIn2 TfanOut Tprot Tshop dTube dTshop (degF)
time.nXTicks=15; time=timecalc(data,jhead,khead,time);
for ma=malist;
   %---------problem specific heat rate calcs---------
   q=data(:,1)>=37890.86806 & data(:,1)<37892.26389; %ltgOff & fanOn
   %q=2390*(1-data(:,1))+2581*data(:,1).*data(:,8); %Qltg+G*dTshop
   Qltg=fLtg*590*3600/1055;%fLtg*[W]*[W2Btuh]=fLtg*2013Btuh
   q=Qltg*(1-q)+mcp*q.*data(:,8); %Qltg+G*dTshop
   %---------end problem-specific heat rate calcs---------

   [jq,jz,jx,jx2,jn]=laggedTs([q data(:,6) data(:,3) data(:,5)],ma,jhead,jtail,nlist);
   [kq,kz,kx,kx2,kn]=laggedTs([q data(:,6) data(:,3) data(:,5)],ma,khead,ktail,nlist);
   time=timacalc(ma,jn,kn,time);
   plotqtzi([1,2],jq,jz,jn,kq,kz,kn,time,'iCRTFid')
   plotqtzi([3,4],jq,jz,jn,kq,kz,kn,time,'CRTFid')
   plotqtzi([5,6],jq,jz,jn,kq,kz,kn,time,'QTzid')
   uscale=1.
   for n=nlist;%model order (number of poles)
      np=n+1
      xj=[jq(:,1:np) jz(:,1:np) jx(:,2:np) jx2(:,1:np)];
      xk=[kq(:,1:np) kz(:,1:np) kx(:,2:np) kx2(:,1:np)];

      y=jz(:,1);
      x=[uscale*jq(:,1:np) jz(:,2:np) jx(:,2:np) jx2(:,1:np)];%Tz-response

      disp('iCRTF LScovWgt')
      alph2=0;
      Rbb=alph2*eye(length(x(1,:)));Rnn=eye(length(x(:,1)));
      [b,bse]=lscovwgt(x,Rbb,y,Rnn);
      bs=bi2bns(b,bse,n,x,y)
      bs2str(bs,n);
      u=rootstu(bs(:,1),bs(:,2),n)
      seqtz=plotcrtf([1,2],xj,n,bs,0);
      seqtz=plotcrtf([1,2],xk,n,bs,jn);

      if 0
      disp('iCRTF lscov')
      [b,bse]=lscov(x,jz(:,1),eye(length(x(:,1))));
      bs=bi2bns(b,bse,n,x,jz(:,1))
      %bs=[b bse abs(b./bse)];
      bs2str(bs,n);
      u=rootstu(bs(:,1),bs(:,2),n)

      disp('iCRTF LSVDcov')
      [b,bse]=lsvdcov(x,jz(:,1),eye(length(x(:,1))));
      bs=bi2bns(b,bse,n,x,jz(:,1))
      bs2str(bs,n);
      u=rootstu(bs(:,1),bs(:,2),n)
      end

      y=uscale*jq(:,1);
      x=[uscale*jq(:,2:np) jz(:,1:np) jx(:,2:np) jx2(:,1:np)];%q-response
      disp('CRTF LScovWgt')
      alph2=0;
      Rbb=alph2*eye(length(x(1,:)));Rnn=eye(length(x(:,1)));
      [b,bse]=lscovwgt(x,Rbb,y,Rnn);
      bs=bc2bns(b,bse,n,x,y)
      bs2str(bs,n);
      u=rootstu(bs(:,1),bs(:,2),n)
      seqtz=plotcrtf([3,4],xj,n,bs,0);
```

```
      seqtz=plotcrtf([3,4],xk,n,bs,jn);
        bz=-bs(np+1)
       if bz<200; bz=1000;    end
      xx=[jq(:,2:np) jz(:,2:np) jx(:,2:np) jx2(:,1:np)];
      bz=fminbnd('lsqz',bz,50*bz,[],xx,y,jz(:,1))
      [b,bse]=lscov(xx,y+bz*jz(:,1),eye(length(y)));%b s/b unchngd
      b=[b(1:n);-bz;b(np:length(b))];
      bse=[bse(1:n);0;bse(np:length(bse))];
      bs=bc2bns(b,bse,n,x,y)
      bs2str(bs,n);
      u=rootstu(bs(:,1),bs(:,2),n)
      seqtz=plotcrtf([5,6],xj,n,bs,0);
      seqtz=plotcrtf([5,6],xk,n,bs,jn);
   end;%n
end;%ma
for i=1:6;figure(i);v=axy(.5); plot([jn jn],[v(3),v(4)]);hold off;end
return;%-----------------------------------------------------------------

function v=axy(f)
vi=axis; axis('tight');
v=axis;
vi=(1-f)*v+f*vi; %tighter y limits
v=[v(1) v(2) vi(3) vi(4)]; %x-fullyTight; y-partlyTight
axis(v); %apply
return;%-----------------------------------------------------------------
function e=lsqz(bz,x,yq,yz)
%cost fcn name passed to fminbnd('lsqz',bzl,bzu,options,x,y,kz(:,1))
%x is the CRTF data matrix with Tzone column removed
%yq=qzone is the vector observations; u=theta0 associated with yz=Tzone
y=yq+bz*yz; cov=eye(length(y)); one=diag(cov);
rmsq=norm(yq-mean(yq)*one);
rmsz=norm(yz-mean(yz)*one);
b=lscov(x,y,cov);
ye=x*b;
e=norm(yq-ye+bz*yz)/rmsq +norm(yz-(ye-yq)/bz)/rmsz;
return;%-----------------------------------------------------------------
function [kq,kz,kx,kx2,kn]=laggedTs(data,ma,khead,ktail,nlist)
%function [kq,kx,kz,kx2]=laggedTs(data,clist,ma,khead,ktail,nlist)
nn=max(nlist);
m=fix((length(data(:,1))-khead-ktail-ma*nn)/ma)
mm=m*ma
k=khead+1; kf=khead+mm+ma*nn;
q=decima(data(k:kf,1),ma); %Qltg+G*dTshop
tz=decima(data(khead+1:khead+mm+ma*nn, 2),ma);
fprintf(1,'%21s Tshop staircase\n',datestr(now));
wp=length(data(1,:)); %future multiple exogenous temperatures
tx=decima(data(khead+1:khead+mm+ma*nn, 3),ma); %Tmain=TfanIn2
tx2=decima(data(khead+1:khead+mm+ma*nn, 4),ma); %Tprot
%NOTE len(tx,tz,q) = len(kx,kz,kq)+nn
kq= q(1+nn:m+nn);
kx=tx(1+nn:m+nn); kx0=kx;
kz=tz(1+nn:m+nn)-kx0;
kx2=tx2(1+nn:m+nn)-kx0;
kn=length(kq);
for n=1:nn
   k=nn-n;
   kq=[kq q(1+k:m+k)];
   kz=[kz tz(1+k:m+k)-kx0];
   kx=[kx tx(1+k:m+k)-kx0];
   kx2=[kx2 tx2(1+k:m+k)-kx0];
end
return;%-----------------------------------------------------------------
function t=timecalc(data,jhead,khead,t);
t.StepHr=[1 2 3 4 6 8 12 18 24 36 48];%allow these axis increments
%following round-off correction works only if 1/t.StepHr is an integer
t.StepHr=1.0/round(1/(24*(data(2,1)-data(1,1))));
t.Startj=data(jhead,1)+1900*365.25-13;%convert from lotusdate to a.d.
t.StartjRem=t.Startj-floor(t.Startj);
t.Startk=data(khead,1)+1900*365.25-13;%convert from lotusdate to a.d.
t.StartkRem=t.Startk-floor(t.Startk);
%tickStartHr=ceil(24*t.StartRem)
```

```
t.StartjStr=datestr(t.Startj,2);%convert to datestring
t.StartkStr=datestr(t.Startk,2);%convert to datestring
return;%-----------------------------------------------------------------
function t=timacalc(ma,jn,kn,t);
   t.StepAggHr=ma*t.StepHr;
   %following works only if 1/tAggHr is an integer
   t.StepXTickHr=ceil((jn+kn)*t.StepAggHr/t.nXTicks);
   t.StepsPerTick=t.StepXTickHr/t.StepAggHr;
   t.StartjTickHr=t.StepXTickHr*ceil(24*t.StartjRem/t.StepXTickHr);
   t.Startjmndex=round(rem(t.StartjTickHr,24)/t.StepAggHr);
   t.Startjindex=round(t.StartjTickHr-24*(t.StartjRem))/t.StepAggHr;
   t.StartkTickHr=t.StepXTickHr*ceil(24*t.StartkRem/t.StepXTickHr);
   t.Startkmndex=round(rem(t.StartkTickHr,24)/t.StepAggHr);
   t.Startkindex=round(t.StartkTickHr-24*(t.StartkRem))/t.StepAggHr;
   t.BaseXLabel=['Time (hr) from 00:00 ' t.StartjStr '(trainSet); ' t.StartkStr '(testSet)'];
return;%-----------------------------------------------------------------
function plotqtzi(fig,jq,jz,jn,kq,kz,kn,t,ti)
%could pass jn,kn as t.jn,t.jn
   figure(fig(1)); plot(jq(:,1),'k:'); title(ti);
   s=t.StepsPerTick; set(gca,'XTick',[[t.Startjindex:s:jn] [jn+t.Startkindex:s:jn+kn]]);
   set(gca,'XTickLabel',round(t.StepAggHr*[[t.Startjmndex:s:jn] [t.Startkmndex:s:kn]]));
   xlabel(t.BaseXLabel); ylabel('Q(Btuh)');hold on
   plot([jn+1:jn+kn],kq(:,1),'b:');
   figure(fig(2)); plot(jz(:,1),'k:'); title(ti);
   set(gca,'XTick',[[t.Startjindex:s:jn] [jn+t.Startkindex:s:jn+kn]]);
   set(gca,'XTickLabel',round(t.StepAggHr*[[t.Startjmndex:s:jn] [t.Startkmndex:s:kn]]));
   %set(gca,'XTick',[[1:s:jn] [jn+1:s:jn+kn]]);
   xlabel(t.BaseXLabel); ylabel('Tz-Tx(F)');hold on
   plot([jn+1:jn+kn],kz(:,1),'b:');
return;%-----------------------------------------------------------------
function seqtz=plotcrtf(fig,x,n,bs,tb)
%usu figure(fig(1:2)) already exist w/axes, titles, plot of observed response
c='bgrycmk';%tb=base time index; for aggLEES.m tb(train)=0, tb(test)=jn
np=n+1;
b=bs(:,1);
   %CRTF (Qz response)
  y=x(:,1)-(x*b)/b(1);
  figure(fig(1));plot([tb+1:tb+length(y)],y,c(np))
  y=x(:,1)-y; sse=y'*y;%reuse y for r
  seqtz(1)=sqrt(sse/(length(y)+1-3*np));

  %ICRTF (Tz as response)
  iz=np+1; %bz=b(iz);
  y=x(:,iz)-x*b/b(iz);
  figure(fig(2));plot([tb+1:tb+length(y)],y,c(np))
  y=x(:,iz)-y; sse=y'*y;
  seqtz(2)=sqrt(sse/(length(y)+1-3*np));
  sprintf('se(q,tz)= %0.5g %0.5g',seqtz)
return;%-----------------------------------------------------------------
function b=rootstu(b,bse,n)
%given un-normalized complete b, print useful stuff; later: return log obj
%assumes b in un-normalized complete form: q[0:n] Tz[0:n] Tx[1:n] Tx2[0:n]
np=n+1; W=(length(b)+1)/np - 2
%in either complete canonical form: u=sum(ce)/sum(de)
utot=sum(b(np+1:np+np))/(sum(b(1:np)))
for i=2:W;j=np*(i+1); u(i)=sum(b(j:j+n))/(sum(b(1:np))); end
u(1)=utot-sum(u(2:W))
seqtz=[bse(1) bse(np+np)]
%be=b(np:np+n-1); %insert be(0)?
ce=b(np+1:2*np);
rc=roots(ce)
;nrc=sum(imag(rc)==0);%abs(rc)==-rc
de=b(1:np);
rd=roots([de])
;nrd=sum(imag(rd)==0);%isreal(rd)
return;%-----------------------------------------------------------------
function b=bi2bn(bi,n)
np=n+1;
b=[bi(1:np,:) 0*bi(1,:) bi(np+1:length(bi),:)];%insert null row
b=-b/bi(1);
return;%-----------------------------------------------------------------
```

```
function bs=bc2bns(b,be,n,x,y)
np=n+1;%[b,bse] convert CRTF model id result to normalized canonical form
r=y-x*b; %compute std error of fit
se=sqrt((r'*r)/(length(y)-length(b)));

%move observation [coeff,se] to RHS
bs(:,1)=[-1; b(1:np+n); b(np+np:length(b))];
bs(:,2)=[sqrt((r'*r)/(y'*y)); be(1:np+n); be(np+np:length(b))];
 %be(1:np+n); se; be(np+np:length(b))];

%normalize
%bs=-bs/b(1);
%compute t-ratios
%bs(:,3)=abs(bs(:,1)./bs(:,2))
return;%-----------------------------------------------------------------

function bs=bi2bns(b,be,n,x,y)
np=n+1;%[b,bse] convert iCRTF model id result to normalized canonical form
r=y-x*b; %compute std error of fit
se=sqrt((r'*r)/(length(y)-length(b)));
%THIS IS se(data); NEED TO CONVERT TO se(coeff) using expression for CV?

%move observation [coeff,se] to RHS
bs(:,1)=[b(1:np); -1; b(np+1:length(b))];
bs(:,2)=[be(1:np); sqrt((r'*r)/(y'*y)); be(np+1:length(b))]; %s/b se(coeff)! ! !
 %be(1:np); se; be(np+1:length(b))];

%normalize
bs=-bs/b(1);
%compute t-ratios
%bs(:,3)=abs(bs(:,1)./bs(:,2))
return;%-----------------------------------------------------------------

function bstr=bs2str(bs,n)
np=n+1;%pretty-print the (complete) model [coeff,se] vectors and t-ratios
%scale by u-value?
t=abs(bs(:,1)./bs(:,2))
xt=[np+1:length(t)];

if 0;
   blanco=char([0*t+32]);
   bstr=[num2str(bs(:,1)),blanco,num2str(bs(:,2)),blanco,num2str(t)];
else
   disp([bs(1:np,:),t(1:np)])
   disp([bs(xt,:), t(xt)])
end
return
blanco=char([0*t(1:np)+32]);
bstr=[num2str(bs(1:np,1)),blanco,num2str(bs(1:np,2)),blanco,num2str(t(1:np))]
xt=[np+1:length(t)];
blanco=char([0*t(xt)+32]);
bstr=[[num2str(bs(xt,1)),blanco,num2str(bs(xt,2)),blanco,num2str(t(xt))]];
return;%-----------------------------------------------------------------

function [b,be]=lsvdcov(x,y,cov)
%Numercal Recipes Ftn (1992) 15.4.3-5 and 15.4.18-19
%corrected per Wunsch,C.I., Ch 2.5, 2003 rev of The Ocean Circulation Inverse Problem (1996)
m=length(x(1,:))
yb=cov\y;%yb=y/diag(cov);%for diagonal cov
%[u,w,v]=svd(cov\x,0);%svd([x(i,j)/sig(i)]);for diag
[u,w,v]=svd(x,0);%svd([x(i,j)/sig(i)]);for diag
s=diag(w);
b=0*x(1,:)';be=b; k=0; tol=sqrt(m)*s(1)*4e-15
while s(k+1)>tol;
   k=k+1;
   if k==m;
      break;
   end
end
%below tol tests c/b avoided by using k as upper summation limit
for i=1:m;
```

```
   if s(i)>tol; b=b+(u(:,i)'*yb)*v(:,i)/s(i); end
   for j=1:m;
      if s(j)>tol; be(i)=be(i)+(v(i,j)/s(j))^2;end
   end
   be(i)=sqrt(be(i));
end
return;%-------------------------------------------------------------------
function [b,be]=lsvdw(x,y)
%Wunsch,C.I., Ch 2.5, 2003 rev of The Ocean Circulation Inverse Problem (1996)
m=length(x(1,:))
yb=y;
[u,w,v]=svd(x,0);%svd([x(i,j)/sig(i)]);for diag
s=diag(w);overs=1./s
b=0*x(1,:)';be=b; ey=0*yb; p=0*eye(m); k=0; tol=sqrt(m)*s(1)*4e-15
while s(k+1)>tol; k=k+1; if k==m; break; end; end
for i=1:k;
   b=b+overs(i)*(u(:,i)'*yb*v(:,i));%(5.77) w/o bias part
   ey=ey+(u(:,i)'*yb)*u(:,i);;      %(5.78)
   p =p +overs(i)*overs(i)*(v(:,i)*v(:,i)');%(5.83) w/o bias part
end
ey=ey-yb;
p=diag(p)
length(yb)-m
be=sqrt(((ey'*ey)/(length(yb)-m))*p);
return;%-------------------------------------------------------------------
function [b,be]=lscovwgt(E,Rbb,y,Rnn)
%Wunsch,C.I., Ch 2.5, 2003 rev of The Ocean Circulation Inverse Problem (1996)
%eqns 2.7.19-21 with provision for Rbb=0
Rnn=inv(Rnn);
m=length(E(1,:))
if rank(Rbb)==m; p=inv(inv(Rbb)+E'*Rnn*E);
else; p=inv(E'*Rnn*E);
end
b=p*E'*Rnn*y;
ey=y-E*b;
p=diag(p);
length(y)-m
be=sqrt(((ey'*ey)/(length(y)-m))*p);
return;%-------------------------------------------------------------------

function [b,be,k]=lsvdwgt(E,Rbb,y,Rnn)
%Wunsch,C.I., Ch 2.5, 2003 rev of The Ocean Circulation Inverse Problem
%(1996);  %COULD USE Rbb,RnnCholeski factors and work in xfmd space
Rbb=inv(Rbb);Rnn=inv(Rnn);
m=length(E(1,:))
yb=y;
[u,w,v]=svd(x,0);%svd([x(i,j)/sig(i)]);for diag
s=diag(w);overs=1./s
b=0*x(1,:)';be=b; ey=0*yb; p=0*eye(m); k=0; tol=sqrt(m)*s(1)*4e-15
while s(k+1)>tol; k=k+1; if k==m; break; end; end
for i=1:k;
   b=b+overs(i)*(u(:,i)'*yb*v(:,i));%(5.77) w/o bias part
   ey=ey+(u(:,i)'*yb)*u(:,i);;      %(5.78)
   p =p +overs(i)*overs(i)*(v(:,i)*v(:,i)');%(5.83) w/o bias part
end
ey=ey-yb;
be=sqrt(((ey'*ey)/(length(yb)-m))*diag(p));
return;%-------------------------------------------------------------------

function b=tlsq(scaleq,nn,x,y)
%Golub & VanLoan, 1989, Matrix Computations, Algorithm 5.1.1
np=nn+1; n=length(x(1,:)); m=length(x(:,1));%n data; m parms
sm=[scaleq*ones(np:1);ones(n-np)];
%use nn for CRTF canonical form; np for complete form
sm=diag(sm);%x-->x*sm [3 PL]
[u,w,v]=svd([x y],0);
[diag(w) log10(diag(w))]
er=-[x y]*v*v';
b=[x+er(:,1:n)]\[y+er(:,n+1)];
return;%-------------------------------------------------------------------
```

Appendix K. Fragment of data LEES1oct.txt for input to aggLEES9.m

lotusdate	TfanIn1	TfanIn2	TfanOut	Tprot	Tshop	dTube	dTshop
37891.14583	73.69	73.52	74.21	73.61	73.84	-0.576	-0.737
37891.14931	73.73	73.52	74.21	73.62	73.85	-0.562	-0.737
37891.15278	73.69	73.52	74.25	73.61	73.85	-0.562	-0.737
37891.15625	73.73	73.52	74.25	73.61	73.85	-0.547	-0.737
37891.15972	73.73	73.52	74.25	73.61	73.85	-0.547	-0.737
37891.16319	73.69	73.56	74.25	73.60	73.84	-0.548	-0.737
37891.16667	73.73	73.52	74.25	73.59	73.85	-0.547	-0.737
37891.17014	73.73	73.52	74.25	73.59	73.85	-0.547	-0.737
37891.17361	73.73	73.56	74.29	73.59	73.86	-0.562	-0.693
37891.17708	73.73	73.56	74.25	73.58	73.86	-0.533	-0.780
37891.18056	73.78	73.56	74.29	73.58	73.87	-0.548	-0.736
37891.18403	73.78	73.56	74.29	73.56	73.86	-0.548	-0.736
37891.1875	73.78	73.56	74.29	73.56	73.86	-0.548	-0.693
37891.19097	73.78	73.56	74.25	73.55	73.86	-0.519	-0.780
37891.19444	73.73	73.56	74.25	73.55	73.86	-0.577	-0.693
37891.19792	73.78	73.56	74.29	73.55	73.87	-0.548	-0.736
37891.20139	73.78	73.56	74.29	73.54	73.86	-0.548	-0.693
37891.20486	73.78	73.60	74.29	73.53	73.88	-0.534	-0.736
37891.20833	73.78	73.60	74.29	73.53	73.87	-0.534	-0.736
37891.21181	73.78	73.60	74.29	73.52	73.88	-0.534	-0.736
37891.21528	73.78	73.60	74.29	73.52	73.88	-0.534	-0.736
37891.21875	73.73	73.60	74.29	73.53	73.87	-0.548	-0.736
37891.22222	73.78	73.60	74.29	73.52	73.88	-0.534	-0.736
37891.22569	73.78	73.60	74.29	73.52	73.88	-0.534	-0.736
37891.22917	73.78	73.60	74.29	73.52	73.88	-0.534	-0.736
37891.23264	73.78	73.60	74.34	73.52	73.88	-0.519	-0.736
37891.23611	73.78	73.60	74.34	73.52	73.89	-0.562	-0.693
37891.23958	73.78	73.65	74.34	73.52	73.88	-0.505	-0.736
37891.24306	73.78	73.60	74.34	73.52	73.88	-0.519	-0.736
37891.24653	73.78	73.60	74.34	73.50	73.88	-0.519	-0.736
37891.25	73.78	73.60	74.34	73.50	73.89	-0.562	-0.693
37891.25347	73.78	73.65	74.34	73.50	73.89	-0.548	-0.693
37891.25694	73.78	73.65	74.34	73.50	73.89	-0.548	-0.693
37891.26042	73.78	73.60	74.34	73.50	73.90	-0.562	-0.693
37891.26389	73.56	73.39	73.82	73.49	73.78	-0.230	-1.299
37891.26736	73.00	72.78	72.91	73.49	73.53	-0.101	-1.945
37891.27083	72.57	72.31	72.39	73.48	73.32	-0.143	-2.161
37891.27431	72.22	72.01	72.09	73.47	73.17	-0.158	-2.332
37891.27778	72.01	71.79	71.92	73.47	73.06	-0.230	-2.331
37891.28125	71.88	71.66	71.79	73.45	72.96	-0.229	-2.374
37891.28472	71.75	71.58	71.71	73.44	72.89	-0.244	-2.374
37891.28819	71.66	71.53	71.58	73.44	72.82	-0.244	-2.416
37891.29167	71.53	71.40	71.49	73.44	72.76	-0.272	-2.416
37891.29514	71.45	71.32	71.40	73.43	72.71	-0.272	-2.459
37891.29861	71.36	71.32	71.36	73.42	72.67	-0.272	-2.458
37891.30208	71.32	71.28	71.32	73.41	72.63	-0.272	-2.458
37891.30556	71.23	71.19	71.19	73.41	72.58	-0.244	-2.544
37891.30903	71.19	71.15	71.19	73.41	72.53	-0.229	-2.544
37891.3125	71.15	71.10	71.15	73.41	72.50	-0.272	-2.501
37891.31597	71.06	71.06	71.10	73.40	72.47	-0.287	-2.500
37891.31944	71.02	71.02	71.02	73.39	72.43	-0.258	-2.586
37891.32292	70.97	71.02	70.97	73.38	72.41	-0.287	-2.543
37891.32639	70.93	70.97	70.97	73.39	72.37	-0.272	-2.543
37891.32986	70.89	70.93	70.97	73.38	72.36	-0.301	-2.543
37891.33333	70.85	70.93	70.93	73.37	72.32	-0.287	-2.543
37891.33681	70.85	70.89	70.85	73.36	72.30	-0.287	-2.586
37891.34028	70.80	70.85	70.85	73.36	72.28	-0.272	-2.585
37891.34375	70.76	70.85	70.85	73.37	72.27	-0.287	-2.585
37891.34722	70.72	70.80	70.80	73.37	72.24	-0.244	-2.628
37891.35069	70.72	70.76	70.76	73.38	72.23	-0.315	-2.585
37891.35417	70.67	70.76	70.72	73.37	72.19	-0.258	-2.628
37891.35764	70.63	70.72	70.72	73.38	72.17	-0.287	-2.628
37891.36111	70.59	70.72	70.67	73.39	72.16	-0.315	-2.585
37891.36458	70.59	70.72	70.63	73.38	72.15	-0.287	-2.628
37891.36806	70.54	70.63	70.63	73.37	72.11	-0.244	-2.671
37891.37153	70.54	70.63	70.63	73.37	72.11	-0.287	-2.628
37891.375	70.50	70.63	70.63	73.37	72.10	-0.301	-2.628
37891.37847	70.50	70.63	70.54	73.38	72.08	-0.287	-2.671
37891.38194	70.50	70.59	70.54	73.37	72.05	-0.301	-2.627
37891.38542	70.46	70.59	70.54	73.37	72.04	-0.272	-2.627
37891.38889	70.42	70.59	70.54	73.36	72.03	-0.287	-2.627
37891.39236	70.42	70.54	70.50	73.36	72.00	-0.272	-2.670
37891.39583	70.37	70.54	70.46	73.36	71.99	-0.258	-2.670
37891.39931	70.37	70.50	70.46	73.37	71.98	-0.315	-2.627
37891.40278	70.37	70.50	70.46	73.37	71.97	-0.272	-2.670
37891.40625	70.33	70.50	70.42	73.37	71.96	-0.301	-2.670
37891.40972	70.33	70.46	70.42	73.37	71.94	-0.272	-2.670
37891.41319	70.29	70.46	70.42	73.37	71.93	-0.287	-2.670
37891.41667	70.29	70.46	70.33	73.39	71.91	-0.272	-2.713
37891.42014	70.29	70.37	70.37	73.38	71.90	-0.287	-2.713
37891.42361	70.24	70.37	70.37	73.39	71.89	-0.301	-2.670
37891.42708	70.24	70.37	70.33	73.39	71.88	-0.315	-2.713
37891.43056	70.24	70.37	70.33	73.39	71.87	-0.315	-2.670
37891.43403	70.24	70.37	70.33	73.39	71.85	-0.272	-2.670
37891.4375	70.20	70.37	70.29	73.39	71.84	-0.258	-2.713
37891.44097	70.16	70.33	70.29	73.39	71.84	-0.329	-2.626
37891.44444	**70.20**	**70.37**	**70.29**	**73.39**	**71.83**	**-0.301**	**-2.670**

Appendix L. "Why It Works" and Relation of Hybrid Method to TLS

The hybrid method appears to, but actually does not, correspond exactly to an implementation of total least squares (Golub, Van Huffel) in which measurement error is attributed to one of the data columns, $T_z(t)$, as well as the vector of observations, $Q_z(t)$, the ratio of T_z to Q_z error is θ_{z0}, and the exogenous temperature(s) are without measurement error.

In most thermal systems of interest, the ratio of T_z to Q_z error is related to the quick mass coupling parameters, $h_q A_q$ and $M_q c_p/\Delta t$. The heat rate of interest is a sum of electric and solar gains and a flow-temperature product of the form $(T_{s/a} - T_{r/a})mc_p$, where s/a and r/a refer to supply and return air streams. Although solar gain is a function of effective aperture area, which is not known exactly (may, indeed, be one of the unknown model parameters) the electric and solar gain measurements are usually made with very low *random* error. The random error in Q is dominated by entering and leaving fluid stream temperature errors.

However, the foregoing LS-TLS formulation gives a different, and not nearly as reliable a result as the hybrid method. The reason is that the LS-TLS objective function is minimized by the LS-TLS solution when the T_z and Q_z errors are uncorrelated. This is not the case. In fact, the errors are highly correlated because a rise in zone air temperature corresponds to heating of the air. Thus the term, $Q_z - \theta_{z0} T_z$, is very insensitive to short-term disturbances.

The errors in exogenous temperatures, on the other hand, are likely to be dominated by systematic (often diurnal or correlated with other weather variables involving sun, wind, and rain) rather than random errors. TLS weights (variances and covariances) are not much help. Thus one are motivated to measure the exogenous temperatures as best we can, and having done so, one are just as well off to treat them as perfectly accurate.

3180-6