ELECTRICAL POWER DISAGGREGATION IN COMMERCIAL BUILDINGS WITH APPLICATIONS TO A NON-INTRUSIVE LOAD MONITOR

by

Ruel Davenport Little

B.A. Physics, The Johns Hopkins University
1986

SUBMITTED TO THE DEPARTMENT OF
MECHANICAL ENGINEERING IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1991

Signature of author...
Department of Mechanical Engineering
May 10, 1991

Certified by...
Leslie K. Norford
Assistant Professor of Building Technology
Thesis Supervisor

Certified by...
Leon R. Glicksman
Director, Building Technology Program
Reader

Accepted by...
Ain A. Sonin
Chairman, Departmental Committee on Graduate Studies

Electrical Power Disaggregation in Commercial Buildings with Applications to a Non-Intrusive Load Monitor

by

Ruel Davenport Little

ABSTRACT

Methodologies and design considerations are presented for a non-intrusive load monitor for commercial buildings. Techniques for extracting information from a building's aggregate electrical signals, both current and voltage, are presented. This information is disaggregated in software and used for equipment fault diagnosis. Methods of augmenting this information using equipment start-up and shut-down transient analysis and building audits are considered. The monitor design can assume different levels of intrusiveness. These are discussed in terms of providing the most efficient implementation given the application and the amount of information desired.

Thesis Supervisor: Dr. Leslie K. Norford

       Title: Assistant Professor of Building Technology

ACKNOWLEDGMENTS

TABLE OF CONTENTS

FIGURES

TABLES

CHAPTER 1 - INTRODUCTION

Buildings use energy for heating, ventilation and air conditioning (HVAC), and to power occupant equipment. The trend in building technology has been to reduce these elements of building energy use by many different methods. There has been difficulty however, in verifying the effect of these methods on the overall energy consumption of the building, which has produced uncertainty as to how effective these methods are. From a maintenance point of view, too, the overall consumption of energy by the building is of interest. Failed equipment clearly impacts building energy use, but partly running systems also have characteristic patterns of energy use which can inform a building operator of problems.

This thesis investigates considerations of monitoring issues of these energy characteristics using building electrical power signals. The bulk of the energy consumed in a commercial building comes from electrical equipment which can be monitored individually or as an aggregate signal at the building electrical power lines. This energy tie to electricity, plus the simplicity of measuring this signal, provides an informative view of the building's energy use. While the individual equipment signal contains information about that piece of equipment, the aggregate building signal contains information from equipment throughout the building. A non-intrusive load monitor would use this aggregate signal to gather information about individual equipment and systems.

We present a preliminary design of a non-intrusive load monitor that takes advantage of the information contained in an aggregate signal to address the characteristics of the individual equipment and systems of equipment of which the aggregate is composed. We are interested in utilizing the power signal in such a way as to extract as much information about the building equipment as possible.

Initially, our project began as the design of a non-intrusive load monitoring device that would be a low cost method to replace sub-metering. The monitor, placed on a building's utility feed would have access to the total building voltage and current information. It would then be able to determine the workings of the building's electrical equipment based on general knowledge of buildings, including their thermal and electrical design.

This point of view changed after the investigation of several buildings showed that the aggregate electrical signal to a building depends on a vast range of characteristics of the particular building. These characteristics include dependencies as varied as its thermal envelope, the location of the its HVAC equipment, HVAC control strategies, type of building, and more. In contrast to this, if we were to sub-meter the building it would implicitly divide up the building information by monitoring each piece of equipment. For example, when sub-metering a fan, one naturally asks what kind of fan it is, what kind of control strategy it has, where it supplies air to, and other questions that involve its disposition in the building system; whereas with a whole building, one asks what kinds of fans are in the building and how they interact with it. We concluded that keeping strictly to a non-intrusive definition caused us to miss a considerable amount of valuable building information.

We have therefore changed the focus of the design problem from dealing with a large amount of aggregate information (with the advantages of cost and non-intrusiveness) to the design a monitor with respect to how much information can we extract from a building using the fewest intrusive techniques such as monitoring locations and individual equipment characterization. We present a set of intrusive levels that enable a load monitor to be applied to a a given building in the informationally most efficient way. We do not advocate any particular level of intrusiveness. Rather, our goal is to illustrate how the information contained in the building power signal can be extracted and used in building monitoring applications.

Chapter 2 discusses the pros and cons of the different levels of intrusiveness, while chapters 3, 4, and 5 describe methods for extracting information using building audits, software power modelling, and equipment transient analysis, respectively. The software modelling focuses on time varying equipment, which are HVAC equipment that have varying power consumption over a day. This is in contrast to constant power equipment that maintains a constant power over time. The transient analysis considers the power signature of equipment when they turn on and off.

Finally, chapter 6 suggests some strategies for using the disaggregation methods and it also discusses their applications on a simulated commercial building.

The questions we are after therefore, are how much useful information is contained in the electrical power signal in a commercial building? can it be extracted? and how intrusive must our meter and knowledge of the building be in order to extract this information?

CHAPTER 2 - NON-INTRUSIVE VS. INTRUSIVE

2.1 Introduction

The design of a non-intrusive monitor must first balance the benefits and detriments of the different levels of intrusive and non-intrusive metering. To define intrusive as collecting any specific information from the building aside from its voltage and current inputs is to cast a two sided discussion that does not allow for gradations. On the non-intrusive side, we are addressing the information contained in a total aggregate signal supplied from a utility feed, while on the intrusive side we are monitoring numerous points for which knowledge is desired. The fundamental question is what is the right amount of intrusiveness such that a building can be monitored efficiently?

As an example of this, consider the total power supplied by the utility company to a town which could be described as an aggregate signal . Certainly it would be desirable to know how much power each house is using by employing just one meter at the utility power plant. While this scheme is too complicated without further information about each house in the town, neither is it an attractive option to monitor the whole town by fully instrumenting every power outlet, which would provide an enormous amount of interesting, though perhaps not useful, information.

Why then consider a non-intrusive load monitor? Because there is a lot of common information that can be generalized about buildings. With less hardware and intrusive information, a non-intrusive monitor can be more flexible to a building and its changing needs in software. The monitor can then provide the building operators with information and diagnostics about the building instead of the operators spending their time maintaining the monitoring system to fit the building's needs.

Another dimension to the intrusive discussion is the matter of cost, which argues for non-intrusive monitoring and its fewer meters. One might be willing to pay a little more for a meter that extracts a lot of useful information. Therefore the real question surrounding cost is how much one pays for access to the varying degrees of information.

This chapter steers the question of non-intrusiveness from a strict definition to a series of levels. The levels range from the definition of non-intrusiveness to complete intrusiveness in 4 steps. Each step indicates the advantages of each intrusive undertaking in terms of a load monitor application.

First, load monitoring applications are described as a motivation for monitoring. Next, existing load monitoring techniques are presented as a background to the different possible intrusive levels of monitoring. Finally, we discuss the advantages and disadvantages of the different levels of intrusiveness that a load monitor can use.

2.2 Applications

The application for a commercial building load monitor of the type discussed here fall into two categories: understanding the normal building behavior, and detecting the abnormal building behavior. An added advantage of using the electrical power as a monitor signal for buildings is that savings in energy relate directly to savings in operating costs. In other words, electrical power is money. The applications discussed here are potential areas of building operations management that relate normal and abnormal behavior to cost savings.

2.2.1 Building load management

Utility energy rates, like telephone rates, vary based on the time of use. Utilities change the rates in an effort to reflect the varying cost of generation, transmission and distribution of electricity. Load management of a building is the practice of using electrical equipment when these energy rates are low. To manage the loads, a monitor must be able to extract from the power signal each piece of equipment's turn-on time, turn-off time, and load profile. With this information about the use of equipment, the monitor can track the building systems, thus giving the operator insight as to where, when, and how energy is being used throughout a building. Optimal start-up and shut-down times can be determined plus methods for peak shaving can be experimented with.

Another possibility of reducing operating costs with a load monitor is by using real time pricing (RTP) to reduce costs by altering the building consumption pattern. Real time pricing give continuous feedback to the building as to the current electric energy cost the utility is charging. Control schemes to optimize the cost of the energy used can be employed using load monitor data. Therefore, with a load monitor, this type of energy cost savings can be recognized, tested, and verified as being effective. [Daryanian, 1991]

2.2.2 Unintended equipment operation

The physical plant at the Massachusetts Institute of Technology (MIT) has to supervise and maintain a vast number of equipment across the campus and personnel admit that the existence let alone the condition of all the equipment eludes them. One benefit of a whole building load monitor would be to verify that the intended equipment are running within a building by using equipment starting and stopping transients. The manpower needed to check that control systems are running the equipment on proper time schedules

15

would be very large for a facility like MIT, especially equipment that is hard to access. Therefore as a first and simple check as to whether equipment and systems are running properly, the electrical operating schedules could be monitored.

In some systems a centralized computer controls the equipment throughout the building. This computer turns equipment on or off using a control signal. In order to know that the equipment actually obeyed the signal, some systems use sensors located at the equipment. With an aggregate load monitor, detecting equipment transients could replace these sensors.

2.2.3 Inefficient equipment behavior

Erratic or inefficient equipment behavior can be tracked by a monitor. While the equipment transients indicate equipment operation schedules, the time varying equipment profiles indicate operational behavior. This knowledge is useful when trying to understand how equipment are interacting within systems. If a feedback control loop causes an oscillating response one might expect to see the behavior on the electrical load side of the equipment. Certainly if the equipment exhibits erratic behavior because of an internal or external fault, the load behavior should be detected and the operator notified.

2.2.4 Predictive maintenance

Predictive maintenance can be as subtle as detecting a cracked rotor bar in a inductive motor, or as large as an increase in fan power due to a dirty filter. In either case the job of the monitor would be to detect abnormal behavior within the context of varying equipment. This is difficult to implement due to the complexities of characterizing

equipment in its "normal" state. However this is a viable application and its success depends on how accurately the monitor can characterize equipment behavior and then map these characteristics to detect long term trends.

2.2.5 Preventive maintenance schedules

Many buildings, especially hospitals, have preventive maintenance programs in place to extend the life of the equipment. These programs are usually set up based on manufacturers specified maintenance and building operator experience. There is no feedback to the operator as to whether the frequency of maintenance is correct. If the equipment is being under-maintained then it tends to fail. However, if it is being over-maintained then the operators are needlessly spending money. Although this application is similar to predictive maintenance its function would be to aid the operator in tuning the maintenance schedule to more accurately account for the equipment's environment.

2.2.6 System problems

The application for monitoring system problems is to attempt to identify problems before the tenants do. As with the predictive maintenance, the method for detecting system problems is to characterize equipment performance and then to extrapolate their characteristics to system characteristics, again with the goal of determining normal operation so that anomalous behavior can be investigated.

For example, if a cooling coil is malfunctioning because of a bad valve, then both the chiller, fan, and the pump that serve the coil can be affected. Therefore if each equipment malfunction can be tied together in a systems view of the building then these

characterizations can then be used to detect long term system trends or to pinpoint immediate problems.

2.2.7 EMCS and optimizing building energy use

All of the above applications fall under the category of optimizing the building energy use which should be the main application of a load monitor. The building load monitor as a whole should be able to provide the information to characterize, maintain, and adjust the electrical equipment as a system within a building. An existing method in the area of building optimization that could use the monitors information is energy management and control systems (EMCSs).

An EMCS, described further in section 2.2.4, is a computer run supervisory command center that monitors and controls building equipment. EMCSs monitor buildings using a variety of mechanical and electrical sensors. The load monitor discussed in this thesis is well matched with the function and environment of an EMCS, and it could incorporate the knowledge it gains from non-intrusive load monitoring into an EMCS. The monitor is designed to gather information about equipment through the electrical signal while reducing the number of electrical monitoring points for the building. This could reduce the complex network of sensors of an EMCS.

The monitor could augment an EMCS in the following ways.

1) An EMCS turns equipment on and off using a control signal. Some EMCS have sensors to determine if the equipment actually turned on or off. The monitor could replace these sensors or just provide another check on equipment state. In particular, equipment that is not controlled or instrumented by the EMCS could be observed.

18

2) Equipment performance characteristics gathered by the monitor could augment or replace the EMCS extensive sensor network.

3) Health analysis through equipment transient monitoring is another possible piece of equipment information the monitor could supply an EMCS.

The extent to which a load monitor can be made to realize these applications is not answered by this thesis. Rather, this thesis addresses the promising paths to gather information about the building through the electrical load signal which these applications can use.

2.3   Load monitoring techniques

There are many ways of monitoring a building's power consumption on the market today. The general trend is to design techniques to measure the building's normal operation and to use that information to optimize building power use, or else to monitor it for abnormal trends and events.

The following methods of monitoring are practiced:

2.3.1 Spot metering

An inexpensive technique based on estimation, spot metering consists of taking instantaneous measurements on major branch circuits. The measurements are made several times over a day or week with a hand held meter. The information gathered from the metering is then used to help determine order-of-magnitude estimates of such loads as

lighting and HVAC equipment which make up the major portion of electrical loads in commercial buildings. With spot metering, time-varying load profiles are often mis-estimated due to the low sample frequency of measurements. Spot metering is performed using a portable watt-hour meter.

2.3.2 Electro-mechanical watt-hour meters

Electro-mechanical meters are the standard "rotating disk" meters used on both homes and commercial buildings. Very accurate and reliable, they are generally used for billing purposes. Many electro-mechanical meters are tied into a central microprocessor unit, allowing multiple points to be monitored. These types of set-ups can be used to separate the billing in an apartment building.

2.3.3 Current transformers

Equipment and buildings are commonly monitored by a direct connection to the voltage source using voltage taps, and a connection to the current using current transformers (CTs). The voltage and current measurements typically are connected to a microprocessor that can provide such useful data such as real and reactive power and power factor from any load signal. Typically CTs are installed on all equipment that require monitoring. Therefore each piece of equipment will have a set of CTs and voltage taps leading to the central microprocessor. For inductive three phase motors (i.e. most HVAC equipment) this requires two CTs and three voltage connections. Aggregate building connections require three CTs and four voltage connections. Finally all single phase connections need just one CT and two voltage taps. Thus the number of monitor

connections increases by about five for each piece of new equipment. Therefore, we note that from a cost perspective, it is advantageous to reduce the number of monitoring points.

2.3.4 Energy management and control system

Although not explicitly a load monitoring technique, EMCSs are systems that use load monitoring as part of their overall function. EMCSs provide a centralized command center for automatic operation of HVAC equipment. The automation combines energy saving schemes with the informational database of the system. The system can use past, current, and predicted monitoring values to optimize the equipment energy use while monitoring occupant comfort.

One monitoring function of an EMCS that has been tested is the optimization of HVAC systems. An optimal control strategy is applied to building equipment based on predetermined equipment characteristics. The information can then be used for fault detection, which uses power as an index to compare optimal building performance with real building activity. Deviations are measured and statistical analysis is employed to determine the possibility of a fault . [Pape, 1991]

With a potentially large amount of information to be extracted from a building by a monitor, there should be some type of centralized intelligent component of a monitoring system that can organize and asses the information. EMCSs are designed for just this function. In addition, some EMCS have expert systems built into their architecture, which are discussed next.

2.3.5 Expert systems

There is no known use of expert systems for load monitoring in the literature, however, there are studies of expert system used to monitor buildings. One study has developed a rule-based systems in the field of building fault analysis. This system gathers information from instruments or visual inspection that can be applied to the HVAC systems and their problems. Application-specific knowledge is gathered about building maintenance problems and their solutions through interviews of building personnel and literature searches. These data, combined with an inference engine that can be run on personal computers, have been assimilated to aid building operators with building maintenance diagnosis. This type of system could incorporate the data from a load monitor to enhance its diagnosis. [Hildebrand, 1986]

2.3.6 Spectrum analysis of motors

A device for fault detection of inductive motors has been developed by the Entek company in England and introduced in 1988. The device, called a motormonitor, monitors the current of a motor that is operating under normal conditions. The company claims that the device can detect such motor faults as broken rotor bars, scored rotors, and rotor eccentricity. It also claims that electrical and mechanical faults can be identified separately. The identification of the faults is accomplished by computer inspection of a spectrum analysis of the motor current draw. Apparently the company has determined the frequency characteristics of each of the operating faults and has implemented a computer identification scheme in the frequency domain of the motor current. The monitor is designed for motors 20 hp or larger. [Entek, 1988]

2.3.7 Residential load monitor

In 1986, a load monitor was developed for residential houses by MIT to non-intrusively identify equipment operation schedules. Attached to the utility billing monitoring device, the monitor is non-intrusive in that it looks at the total power used by the house. With only the total power signal, it attempts to sort out the equipment that is operating based on the equipment start-up and shut-down changes in power level. The key to the residential monitor's success is its ability to monitor both the total current and voltage, thereby enabling the device to calculate real and reactive power. The device uses a sample speed of one Hertz.

Since the appliances found in a household typically have constant power profiles, the device can match the characteristic start-up and shut-down level changes in real and reactive power. In the study, it was noted that some knowledge of normal household use schedules were very helpful in matching level changes. [Hart, 1985]

2.4   Non-intrusive discussion

Without limiting a non-intrusive load monitor to total building power information, we define a *non-intrusive* load monitor as a device that extracts as much information from the total building power as possible in software, using the least amount of hardware and *a priori* building knowledge. This is to say that hardware and building knowledge should be considered, and that they should be minimized in such a way as to be used in the most efficient possible way. Given the previous load monitoring techniques as a background, we now describe the different levels of intrusiveness in load monitoring with the goal of

23

leading the reader through the reasoning behind the level of intrusiveness prescribed by this thesis. Furthermore, the benefits and detriments of each level are described.

2.4.1 Intrusive level 1

The least intrusive monitoring method relies exclusively on the knowledge of power, voltage, and current, for the whole building. This method, used in the residential monitor, represents the quintessential non-intrusive load monitoring technique: to be able to understand the dynamics of building equipment, and thus the building, using one meter at one placement point. At first inspection, the advantages of this method are that it would be cheap in terms of hardware, neat in terms of installation, universal in terms of building types, and it would not require a researcher to enter the building except for installation. The disadvantages are that a commercial building can be very complex, and the advantages just stated may not hold up well when put into application. Therefore the level 1 intrusiveness becomes less appealing when applied to commercial buildings because of the complexity of the equipment, which are also very co-dependent.

The cost of a level 1 load monitor should be lower than an equivalent monitor with many monitoring points due to hardware and installation costs. However, the limited information that the monitor is able to extract from the signal must be compared to the cost of the extra hardware and labor. In addition, if the numerical processing is too complex for a simple computer then cost may be inflated. One argument against this reasoning is that technology is moving fast enough to make cost assessment of a computer outdated in a very short time.

With a level 1 monitor, the installation of a single monitoring point at the utility feed can be problematic because it requires a building shut-down in order to avoid dangerous installation of the CTs. Even though the split core CTs do not require the utility cables to

24

be disassembled in order to install them, the voltages involved at that point in the utility feed are high enough to make any work too hazardous to perform while the cables (which are not sheathed) are live. Therefore the power must be cut off to the building in order to install the monitor at that point. With some buildings this is not a problem, but with others, like hospitals and computer facilities that run around the clock, it can be an obstacle.

A major difficulty with the level 1 monitor is the time varying nature of many commercial building equipment. This equipment varies such that the information supplied by the aggregate building power signal cannot provide equipment information past the knowledge supplied by the start-up and shut-down transients. In the diurnal cycle of a building, the equipment start in some power state and finish in another. In order to track the states of the time varying equipment during the cycle, the monitor must have access to some independent variable which the equipment power can be shown to depend on. This would require another level of intrusiveness.

The appeal of the universality of a level 1 monitor may be overshadowed by the non-universality and constantly changing load characteristics of equipment. Without any knowledge of equipment in the building, the monitor would need a library of equipment transients in order to recognize the multitude of brand and design differences in equipment transients. Given that the library can be completed for different equipment types (fans, chillers, pumps, lights, etc.), then the monitor must consider all possible cases. That is to say, if the monitor checks a transient to see if it is a fluorescent light, it must check it against an average or characteristic fluorescent light transient, which by the nature of statistical averages, will have a region of uncertainty. This uncertainty is much greater for a monitor that has no *a priori* knowledge of the types of equipment within a building than a monitor that has equipment information. Therefore the solution to this last difficulty is to obtain a knowledge base of the major equipment and their transients within each specific building. This method is part of the next level of intrusiveness.

25

2.4.2 Intrusive level 2

This next level of intrusiveness deals with extracting the time varying information contained in the daily cycles of the equipment, and cataloging the types of transients in a specific building. This information can be used not only to determine how a normal building uses its energy, but also to detect when the normal building is operating improperly.

The following "extra" intrusive information is supplied to the monitor in this level. First, independent variables are monitored for time varying equipment. For example, flows for variable volume air systems, and chiller evaporator temperature difference can be used to model the time varying equipment. Next, a building audit of the major equipment is procured which supplies the monitor with nameplate information (size, maximum current, voltage, brand, etc.), equipment use schedules, and the type of equipment (constant volume pump, variable inlet vane fan). Finally a description of how each piece of equipment fits into the building as a system is supplied.

This information on the elements of the building system allow the monitor to associate start-up and shut-down transients and time varying loads with equipment. The independent variables also give the monitor the ability to model the time varying equipment so that equipment load profiles can be established. This information could then be utilized by the monitor for equipment optimization and fault detection. The linking of equipment by system can prove analytically useful by placing equipment into groups that affect each other. Since some equipment track other equipment very closely, making them seem like one piece of time varying equipment, the information can also resolve the load profile redundancy. In other words this intrusive aspect of level 2 provides a road map for fault diagnosis by revealing equipment interdependencies and characteristics.

The intrusiveness inflicted on the building would consist of sensors wired into some equipment and a survey of the building's electrical plans. For a large commercial

building, the number of time varying equipment needing sensors is on the order of 1-3 for chillers and 2-8 for air handlers (note that return fans typically track supply fans as will be discussed later, so only one sensor is necessary). In some cases the sensors already exist in buildings with EMCS already in place. Therefore the information could be acquired with little inconvenience and cost.

Finally, the equipment information would also allow for the matching of transients with actual building equipment rather than making the monitor guess the equipment make-up of the building. Not only would the library of transients be reduced and possibly simplified, the monitor would have a closed set of equipment with which to associate the transients.

2.4.3 Intrusive level 3

At the third level of intrusiveness we divide up the building into different electrical equipment groups by installing several monitoring points in close proximity to the building power supply. With different types of equipment branching from the main utility feed at different points, the load can be disaggregated by hardware. This technique provides the monitor with a method of partially disaggregating a large building in hardware before tackling the smaller pieces in software.

Based on the wiring design of the building, it may or may not be possible to separate the HVAC equipment, which represents the main time varying loads, from the other building loads. If possible, this disaggregation would have the benefits of dividing the aggregate signal into smaller groups of known equipment, which would allow the monitor to efficiently use its observational resources to attack each monitoring point. The following list contains examples of this aspect of level 3.

1) The noise from small appliances can be separated from other monitoring points.

2) The different metering spots can use smaller ranged analogue-to-digital converters. With the increased signal-to-noise from (1), the monitor can detect smaller changes in the current signal at a reduced cost.

3) The different metering spots can have different front end hardware to more closely match the needs of the signal being monitored. For example the HVAC monitoring points could concentrate on the time varying loads instead of the transients since the independent variables will indicate what equipment turned on or off, and when. The other monitoring points could then look for specific loads, such as lighting, using transient detection, thus further reducing the transient identification problem.

4) Separating noisy constant loads from time varying loads increases the accuracy of the modelling due to the reduction in the complexity of the model, and in the noise it has to account for.

If the time varying loads in a building are too numerous to model using one monitoring point, then more points could be used further to reduce the system to a manageable size. How many pieces of time varying equipment can be modelled off one point depends on the nature of the equipment and the accuracy of the model desired.

The disadvantage of level 3 monitoring is that there is an increased number of measurement points in the building. As the number of points increases to a value close to the number of major equipment in the building, the intrusive level can then be considered level 4.

2.4.4 Intrusive level 4

The final level of intrusiveness depends on the brute force method of monitoring all equipment or at least all major equipment. This allows for complete knowledge of equipment load profiles and equipment schedules. However, the non-HVAC equipment would not be realistic to monitor individually since it is scattered throughout the building, thus requiring extensive monitor lines.

The only real added advantage that level 4 has over level 3 is that accuracy of knowledge has been increased. If it is determined that the monitor applications do not need the added accuracy supplied by the individual monitoring locations, then it should certainly be avoided. Considering the case of the motor diagnostics performed by the Entek instrument, an accurate signal is needed in order to determine the finer motor characteristics. If several motors are tested from one signal, the accuracy of the diagnostics would suffer. In other words, as the required diagnostics on the building and equipment become more subtle, the monitoring of the equipment, in general, will become more intrusive to the building.

The diagnostics required for a building depends on the application of the monitor. The next section revisits the applications of section 2.2 to consider them from an intrusive levels viewpoint.

2.4.5 Applications revisited

We now list the previous applications and the levels of intrusiveness that suit them

Building load management

For load management, the monitor characterizes equipment profiles so that power use schemes can be implemented. This requires at least a level 2 monitor for the characterization, but a level 3 and 4 will further increase the characterization accuracy. However, for simple, order-of-magnitude building management, level 2 is sufficient.

Unintended equipment operation

This application bases its knowledge on equipment transients which can be detected by a level 1 monitor. At level 2, building knowledge allows transient detection to tax software less by providing the types of equipment in the building. For level 3, the accurately, and therefore the probability, of the detection increases. Level 4 detection is implicitly exact.

Inefficient equipment behavior

Here again the level 2 monitor is necessary to characterize the equipment load profiles. In this application, however, the profiles must be accurate so the monitor can detect small changes in equipment characteristics. Therefore we suggest a level 2 or 3 as a non-intrusive method, while level 4 would be the most accurate.

Predictive maintenance

Predictive problems can show up as very small perturbations in start-up transients or large changes in system characteristics. A level 2 monitor could detect larger problems which we address in chapter 4, while level 3 would probably be significantly more efficient

due to the reduction of noise. For very the small transient changes, level 4 would be the best.

Preventive maintenance schedules

Preventive maintenance deal with the same types of equipment problems as predictive maintenance, so the intrusive level should be similar, depending on the desired accuracy.

System problems

System problems deals with equipment failures after they happen, which can be handled by a level 2 monitor. Subtle power problems that occur on non-electrical equipment but affect electrical equipment would need a higher level of intrusiveness, perhaps 3.

EMCS

An EMCS could use any level of intrusiveness to augment the information it gets from other sensors. As the levels get more intrusive, an EMCS would have more information by which to make decisions and, therefore the best level would be 4. We do not advocate this level because of the complexity that the monitoring system would then have. We believe that the most efficient level to use would be level 3.

Table 2.1 shows a list of the applications with their least intrusive and suggested levels.

Table 2.1: Applications of load monitoring with intrusive levels

Application	Least intrusive level	Suggested level	Notes
Building load management	2	2	
Unintended equipment operation	1	2	Increasing level improves detection probability
Inefficient equipment behavior	2	2 or 3	Increasing level improves equipment characterization
Predictive maintenance	2	3 or 4	Level 4 may be needed for subtle detections
Preventive maintenance schedules	2	3 or 4	Increasing level improves equipment characterization
System problems	2	3	Increasing level improves system characterization
EMCS	1	3	Use varying levels depending on desired information

In summary, the important decision criteria of a non-intrusive load monitoring are accuracy versus uncertainty, informational content versus hardware flexibility, and cost tradeoffs. As the level of intrusiveness increases, the noise inherent in the measurements will decrease and the accuracy of the monitor will increase. In addition, the computational complexity of disaggregating the signal becomes simpler and the uncertainty of the time varying equipment characterization and the transient detection will decrease.

Since the informational content of the building signal is may be separated with more accuracy and less uncertainty with increasing level of intrusiveness, the informational content of the equipment data will increase. That is to say, finer equipment characteristics

can be detected. In contrast to this, the flexibility of the monitor to changing building

equipment will decrease. A level 4 monitor must update its hardware as equipment are

changed or replaced in a building. A level 3 monitor is more flexible, however, its

hardware separation of loads can be disrupted if new equipment are added to the building.

A level 2 monitor requires only an input to a database as equipment change. Finally, a level

1 is completely flexible.

The issue of cost has not been quantitatively analyzed, however, it will increase as

the intrusive level and the number of monitoring points increase. We have not developed a

price for information so we can not numerically compare the levels of intrusiveness using a

cost analysis. The level of intrusiveness depends on the information desired and the cost

constraint that one is operating within. Figure 2.1 summarizes the qualities of

intrusiveness that we have discussed.

| Intrusiveness | Intrusiveness | Intrusiveness |
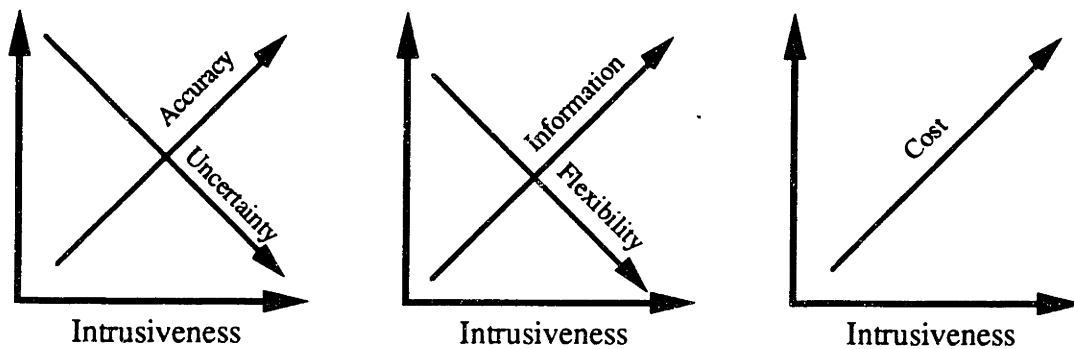
Figure 2.1: A graphical representation of the effect of intrusiveness on,
from left to right: a) Accuracy and uncertainty of collected information.
b) Collectable information and monitor flexibility c) Cost of hardware
and installation

Each building has different attributes that a load monitor must consider, whether it

is the systems within the building, the function of the building, the climate that it is located

in, or the architecture that was used to design it. With this diversity, a level 1 load monitor

faces an enormous task to extract the information for the applications discussed. Conversely, because of the standards by which buildings are constructed and the similar elements of equipment that a building uses, a level 4 monitor does not make use of the knowledge of building design and the abundance of information contained in an aggregate power signal. On the other hand, depending on the application, level 2 or level 3 maintains a balance between knowledge and intrusiveness. These two levels extract information without imposing unnecessary intrusiveness, while also allowing the monitor to be flexible to the building's changing needs.

CHAPTER 3 - BUILDINGS AND EQUIPMENT

3.1   Introduction

Taking an informational point of view, this chapter focuses on the types of
equipment one would find in a commercial building. Initially, we set out to characterize all
buildings from a level 1 viewpoint. In sections 3.2 - 3.5 we generalize the size of
equipment per square feet of space in the building to give typical equipment scales versus
building size. We then apply these generalizations by examining the equipment and
systems of two large buildings, one tall and one wide. The dimensions of these two
building are also contrasted in terms of specialized equipment needs.

The complexities of the building systems we encountered from our building studies
lead us to consider a level 2 informational description of building systems, given in section
3.8. This system description reduces the plethora of possible system types into a set of
common equipment types that can be flexibly linked together to form any system. This
method of describing the building system has applications to a monitor database and system
fault detection.

The intent of the informational survey of the commercial buildings in this chapter is
to pinpoint the level of intrusiveness that is necessary to extract information from an
aggregate electrical signal. This building information can then be used to augment a
monitor's signal processing for transient detection and time varying load disaggregation.

3.1.1 What is a building?

In terms of a commercial building and its equipment, there are general characteristics that can be quantified. A building, in general, provides people with an environment that is different from its surrounding environment. A more specific function of a commercial building is to provide an environment for a number of people performing a specific function, or many different functions. For instance, a bottling factory provides a large conditioned space to run automated equipment around the clock. In contrast, an office building must maintain environments in many different offices (zones) nominally from 9 am to 5 pm. A hospital is yet another example of a building with a specific function, in which there are many zones with the constraint of sanitary air flow throughout the building. Furthermore, the hospital may have a laboratory facility with explicit temperature and humidity needs.

Listed and described here are the basic equipment elements and sizes of a typical commercial building. The data were compiled from a combination of the *ASHRAE Handbooks* and interviews with HVAC suppliers [ASHRAE,1983,1984,1985].

3.2 Fans

3.2.1 Air volume

When selecting equipment for buildings, there are general rules that HVAC engineers follow in sizing the equipment. We spoke to representatives at The Trane and Carrier HVAC companies as to the general sizing of their equipment in commercial buildings. In general, they sized fans by a need of 400 cfm of air for every ton of cooling

from a chiller, which is derived from the thermal heat capacity of dry air and a generalized temperature for chilled water. This corresponds to a need of one ton of cooling load from a chiller for every 400 square feet of conditioned space. The air volume need for a commercial office building is thus:

$$\text{Air volume need} = \frac{400 \frac{\text{cfm}}{\text{ton}}}{400 \frac{\text{ft}^2}{\text{ton}}} = 1 \frac{\text{cfm}}{\text{ft}^2}$$

Although fans systems vary in size, a rough estimate power versus flow for the fans we have encountered is 1 Watt/cfm. This implies that a generalized fan power would be 1 Watt/ft^2.

3.2.2   Supply fans

The building air volume requirement is created by one or two supply fans. Typical supply fan air volumes range from 15,000 to 50,000 cubic feet per minute (cfm) depending on the size of the building. For buildings that are large enough to need more air volume, separate groups of supply fan systems supply different zones in a building. This configuration can be seen in tall buildings that have the upper and lower floors conditioned by different fan sets. The supply fans are distinguished from other fans in the building by being the largest, with 0.5 hp/ton of cooling as an upper limit for their size.

Mechanical Room Fans
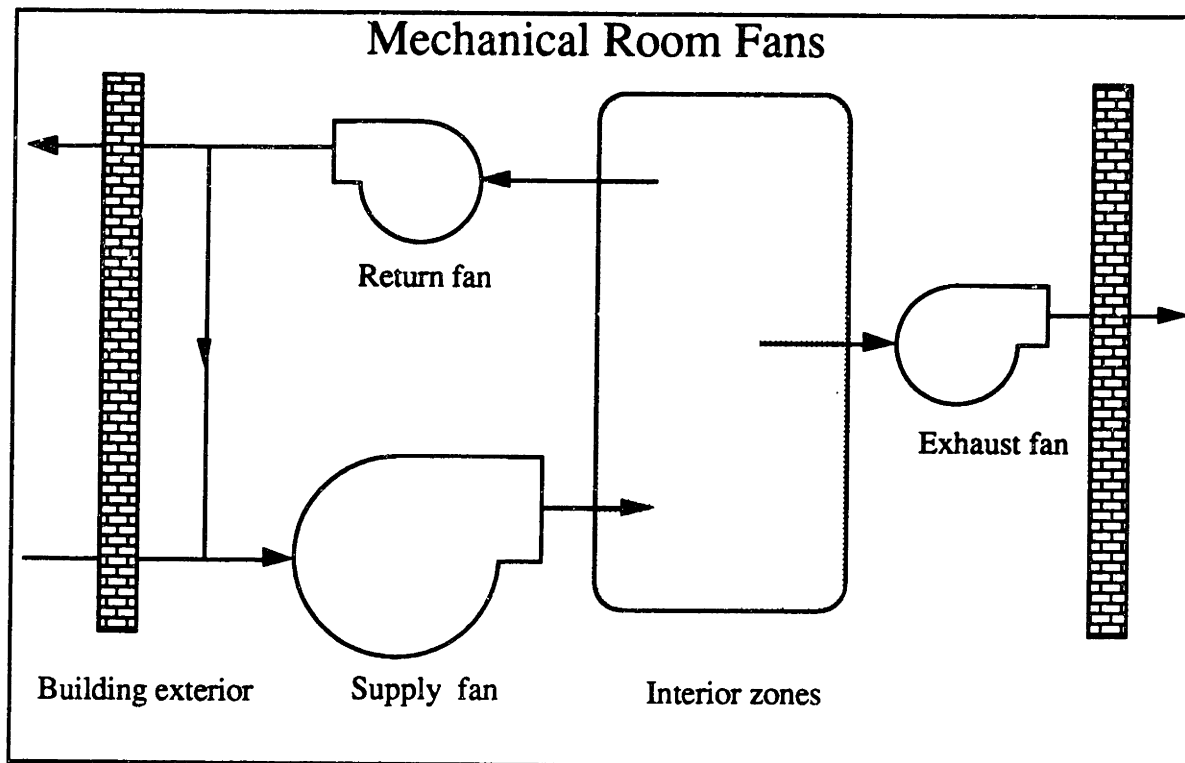
Figure 3.1: Schematic of general fan types in a commercial building. In small buildings the return fans are not needed.

Control

Control schemes for building fans depend on the building and the HVAC system chosen. However, three types of control can be generalized for centrifugal fans: constant volume, variable speed drive (VSD), and variable inlet vane (VIV) controlled fans. The constant volume fans may not be considered as much a control scheme as an open loop system where the fan is sized for the peak cooling air quantity need for the building. The heating requirement is then implemented with reheat coils. Control to the space is adjusted by cooling or heating to the air. Constant volume fans are therefore considered constant power.

The VSD and VIV control schemes attempt to reduce power usage by delivering a volume of cooled or heated air to the space to be conditioned as needed. The advantage is

that the temperature of the cooled or heated air is controlled to a fixed temperature while the amount of air to the space is varied. This allows for optimization of heat delivered to the air and air delivered to the room.

The VIV control varies the air volume using inlet vanes to throttle the flow to a constant speed fan. This reduces the amount of air that the fan must push, while also wasting energy due to the inefficiencies of the vanes as they reduce the air volume. The VIV control eliminates this inefficiency by varying the speed of the fan. Both these methods produce a variable volume that matches the need of the space to be conditioned.

The control sensor for both the VSD and the VIV systems is typically a static pressure measurement located two thirds of the way downstream in the duct. Due to this control method, the VIV and VSD systems are inherently time varying in their power usage. The use of power is dictated by the complex needs of the conditioned spaces that the fans supply. Consequently, the control function of the supply fan is to produce a positive static pressure in the ducts while the air volume that it pushes affects the electrical energy used by the fan. Supply fans are also required to deliver to a building a minimum outdoor air requirement as specified by the ASHRAE standard 62-90 . This minimum air requirement maintains a dc level in the supply fan power and must be considered in any monitoring application.

3.2.3 Return fans

Return fans exhaust and recirculate controlled proportions of the building air. They consist of possibly many fans amounting to a total horsepower of a fraction (< 50%) of the supply fan total horsepower. In smaller buildings it is possible that return fans are not needed.

Control

Return fans are controlled using the same methods as the supply fan, and a return fan will generally have a control scheme to match the supply fan. Therefore a VIV supply fan will have a VIV return fan.

The control measurement for the control fan should be designed to maintain the supply fan flow rate minus a minimum outside air quantity to account for exhaust flows and to slightly pressurize the zone. One technique is to measure both the flow of the supply and return fans and control the return fan flow by a fixed amount lower than that of the supply fan. Some fans measure a fixed pressure differential between the return air plenum and the zone being conditioned. Other fans use the supply duct static pressure measurement minus a field calculated constant, which is essentially a open loop control method that doesn't take into account possible fluctuations in the air supply system. The flow differential and the return air plenum methods, in contrast to the static pressure method, are active closed loop control techniques that take into account the air supply system, however, they also require a return fan sensor.

Any control method for the return fan will tend to track the supply fan with varying degrees of success. This roughly translates to a similar time varying power usage pattern.

3.2.4 Exhaust fans

Exhaust fans consist of many fans scattered about the building used to expel unwanted and non-recirculable air out of the building. These fans are typically less than 10 hp except for such special purposes as kitchen range hoods in cafeterias. Most exhaust fans service toilets, mechanical rooms, and elevators. The exhaust fans are constant

volume fans that affect the building by slightly reducing the dynamic zone pressure produced by the supply fan and return fan pairs.

3.2.5 Miscellaneous fans

There are many other possible fan systems within a building's HVAC system. However these fans tend to be of fractional horsepower and are used for local distribution of air. One example is fan powered zone boxes used in VSD systems. Another possibility is laboratory exhaust hoods, which are sometimes several horsepower in size. Since all these fans are small and depend on building specific factors we will not consider them as general fans. However there are two cases where significant fans show up. First, in larger cooling towers there are often fans that are used to increase evaporative cooling. These fans sometimes have more that one speed setting. Secondly, in some chillers the condenser is cooled by air instead of water. This is done using one fan or a set of fans that amount to a substantial fan size.

3.3 Chillers and pumps

Chiller systems in commercial buildings function as centralized heat removal equipment. There are many types of chillers to match the general needs of different buildings. In describing a general chiller system, we include all possible pumps that generally are used with larger buildings. In some cases, chillers don't use pumps and liquid to distribute the heat, rather they use fans and air. Therefore, these general statements about chillers systems describe what one could expect to find in a chiller system.

Due to building differences, the number and the sizes of the pumps and chillers may vary, but generally the make-up of the components of the system will not.
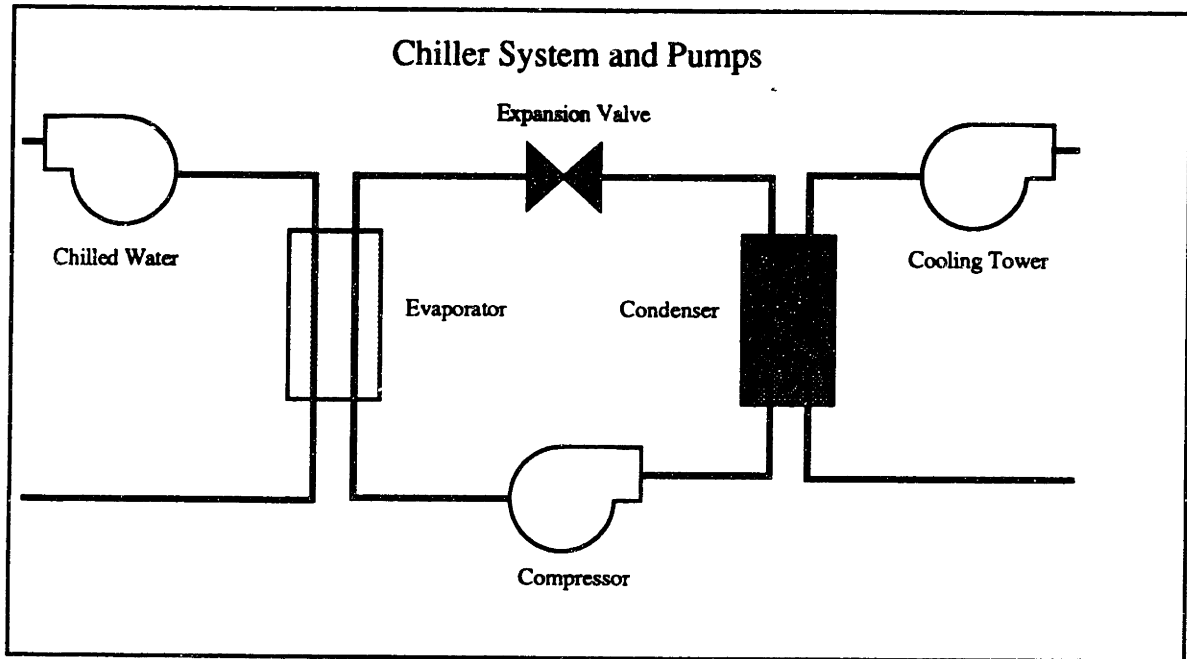
Figure 3.2: A schematic of a chiller system and pumps.

3.3.1 Compressors

The various types of compressors typically found in chillers and their range in sizes are in Table 3.1. The thermal power of the compressor can be converted to an electrical power using the coefficient of performance of the chiller (COP). A typical COP is about 3.5.

Table 3.1: Chiller types

	Capacity	Thermal power
Reciprocating	1/16 - 150 tons	(.046 - 112 kW)
Helical Rotary	100 - 750 tons	(350 - 2600 kW)
Centrifugal	100 - 10000 tons	(350 - 35000 kW)

Compiled from ASHRAE [1983]

The centrifugal compressor chiller is the most commonly used chiller because it has a long life expectancy and it has a large range of cooling loads that it can be designed for. This makes it a low maintenance and high versatility chiller. It has typically one compressor per chiller in contrast to the reciprocating chillers that can have multiple compressors. Chiller power usage depends mainly on cooling load within the building which is naturally time varying.

3.3.2 Chilled water pumps

Pumps in buildings are usually centrifugal pumps. For chilled water circulation there are generally several pumps of identical or very similar sizing plus one for standby (emergency) use. An additional smaller pump is sometimes used for cooling in the winter for such areas as the core of the building that always have a cooling load. These pumps are about 20% the size of the summer pumps and therefore run at a reduced chiller water capacity to circulate water through the system. The chiller does not run at this time.

3.3.3   Cooling tower pumps

The cooling tower pumps circulate the warmed water from the chiller condenser to the cooling tower which then expels the heat through evaporation. The cooling tower pumps are set up in the same way as the chilled water pumps.

3.3.4   Domestic water pumps

In buildings that require water above 90 feet (approximately six stories) pumps are employed to boost the municipal water pressure which ranges from 40 to 60 psi. In cases of very tall buildings, there are water reservoirs spaced at various heights in the building. The water is then pumped up to these reservoirs by a series of medium sized pumps. This has the added cost savings and safety benefit of not requiring high pressure piping.

3.3.5   Hot water pumps

For HVAC systems that require heating coils, water can be used as the thermal transport medium and pumps are necessary. Pumping sizes should be on the same order as the chilled water pump sizes in the building.

3.3.6   Air compressors

Many buildings use compressed air control lines to control the HVAC equipment. These control lines need air compressors to maintain a high pressure. The air compressors

are typically under 10 hp in size and frequently cycle on and off. The cycling frequency is on the order of minutes, and the compressor is considered a constant power device.

3.4 Tenant loads

Loads in the building that do not correspond to HVAC functions can be considered tenant loads. Although there are an innumerable number of obscure devices that can be plugged into a building we have noted three common loads.

3.4.1 Lights

Just as HVAC equipment is part of building design, so too is lighting. A common design estimate for lighting is 1-1.5 watts per square foot of building space. A very rough estimate of total electrical lighting load for a building is that it is comparable in magnitude to the ventilation load. We also note that lights are constant loads with predictable use patterns.

3.4.2 Office equipment

Common modern office equipment includes personal computers, computer printers, and copiers with the caveat that the term modern for office equipment is changing very fast as is exemplified in the last ten years with the rise of personal computers. Our current estimate is one personal computer per office in a professional office building. We have

also been given estimates of 1 to 5 watts per square foot for office equipment by building operators.

3.4.3 Personal air conditioners

Room air conditioners, which can be seen sticking out of windows on some buildings, generally are installed in circumstances where the building's HVAC system is faulty or not adequately handling the cooling load. This is especially true for older buildings, particularly buildings that have gone through changing needs. A typical unit designed to condition a 220 square foot office requires about 750 Watts. Most units have a power consumption of about 4 Watts/ft^2.

3.5 Special problems associated with dimensions

As buildings get larger than approximately 500,000 square feet, their size and shape begin to play a role in the types of equipment that are in their mechanical rooms. More specifically, a tall building has a large gravitational potential against which to move liquids throughout the building, causing the building to need additional pumps of larger than normal size. The ventilation systems, however, have easy access to floors and space within the building because of the floor stacking.

In contrast to these special needs of tall buildings, a low story wide building of similar size will have less need for large pumps due to the lack of potential to work against. Also the ventilation distribution system will be larger in order to deliver air over a sprawling space. Plus, the increase in duct length results in more air friction and air leaks

requiring larger fans. In summary tall buildings need added pumping capacity, while wide buildings need added fan capability. An example of this relationship is illustrated in section 3.7.2.

3.6   Electrical wiring configuration

The attachment of a monitor in a building requires some knowledge of the electrical wiring configuration in commercial buildings. We studied the transformer vault layout in building E17 on the MIT campus, where the members of the physical plant staff removed the high voltage panel enabling us to map the electrical branching of the utility electrical supply which is shown in figure 3.3. The transformation of utility feed voltages down to building voltages followed a linear set-up. The utility line is fed through a large transformer in the mechanical room that steps the voltage down to 480 volts. A 480 volt bus bar line is then run horizontally. Bolted to the bus bar are vertical rising cables called *risers* that either pass upwards through to each floor via housing or are connected to motor control breakers that run HVAC equipment. The tenant risers can be accessed by a breaker box on each floor of the building where a breaker box again transforms the voltage down to an appliance level of 120 volts. The HVAC equipment that runs off 480 volts is tapped directly off the bus bar in the mechanical room where the HVAC equipment is often located.

In terms of attaching the monitor, the whole building can be monitored at a spot on the bus bar just after the utility transformer. However, it could be advantageous to attach the monitor in two or more places such that the tenant loads and the HVAC loads can be separated in an intrusive level 3 hardware sense. It is common to have a linear configuration for the bus bar. On the other hand, the risers seem to be different for each
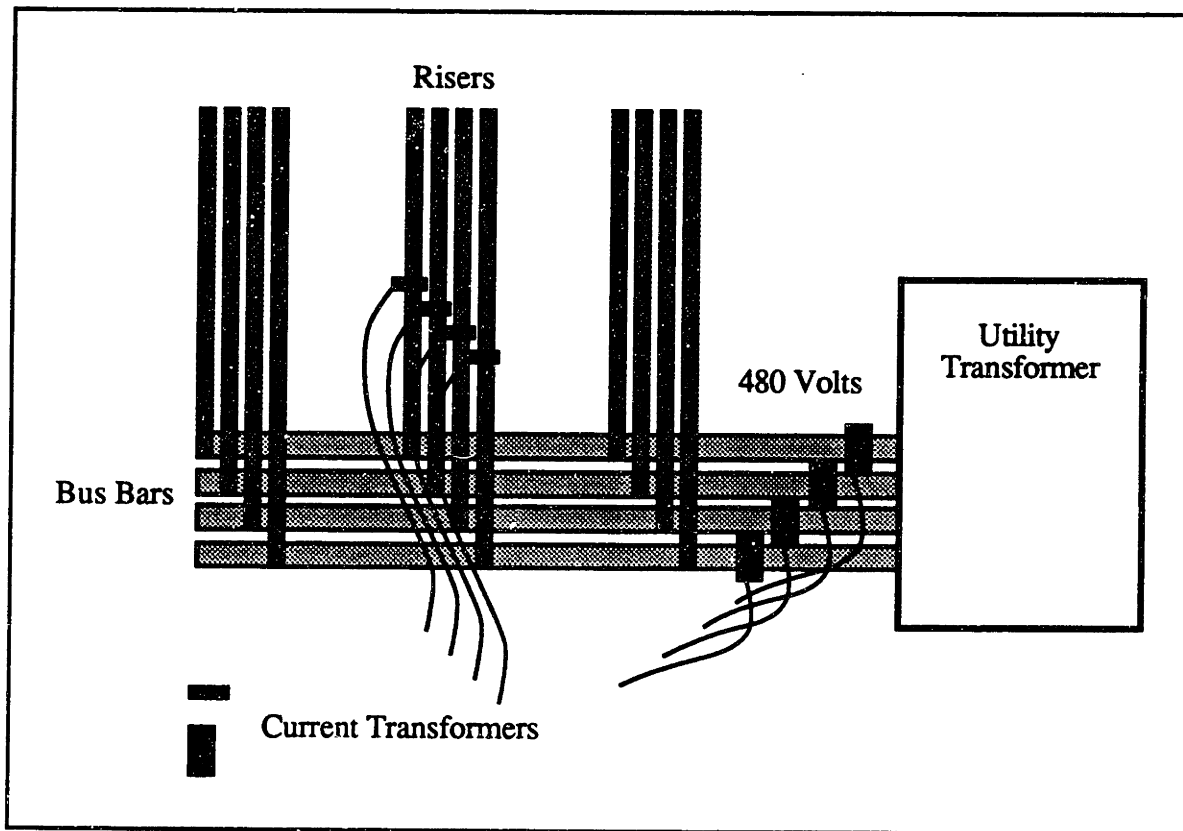
Figure 3.3: A schematic showing how the utility transformer supplies electricity to a four-bar bus. The various building loads get their supply from the risers that tap into the bus bars in a linear configuration . A monitor could tap into the whole building using CT's just after the transformer or on the risers.

building such that the bus bar taps are configured for greatest locational convenience. That is to say, the tenant load taps and the HVAC taps on the bus bar can be inter-dispersed. In order to separate the loads there would have to be a set of CTs located at each tap. One way to achieve the separation and possibly reduce the number of CT's used, is to monitor the whole building at the transformer, and then monitor the set of equipment with the fewest number of bus bar taps, either tenant or HVAC. The load not measured could then be reproduced by subtracting the monitored equipment from the total. The subtraction could possibly be done in hardware.

We have not encountered any buildings that have a branched set-up where the tenant risers and the HVAC risers split at one junction. This would be advantageous to a level 3 monitor because it would only require two sets of CT's to separate the loads.

3.7   Building examples

3.7.1   The John Hancock Tower

The John Hancock Tower in Boston, Massachusetts is an example of both a large building and a tall building. About twenty years old, the Tower is a modern building with fairly up-to-date equipment and computer control. We chose to study the Tower because of its overall size and vertical dimension. Because of these attributes the Tower has both large general needs and very specific building needs.

The Tower is a sixty-two story building with approximately one million square feet of floor space. From the equipment list in appendix B, it is obvious that this large scale building has an extensive amount of space to condition and that the Tower uses large-scale systems with some miscellaneous equipment to satisfy specific local needs. However, dividing the building into areas conditioned, figure 3.4 shows that the tower is really three smaller sized buildings stacked one on top of another. There is a transformer vault for each of these areas that turns the utility service into 277/480 volts, located on floors BFR, 7, and 62. This partitioning of transformers would require three load monitors, thus further creating the effect of three separate buildings. The transformers supply power to the mechanical rooms and the areas indicated in figure 3.4. The mechanical rooms in turn supply the areas with HVAC. As well as being divided up into smaller areas, each area is divided up into north and south zones.

GENERAL LAYOUT OF HVAC IN THE JOHN HANCOCK TOWER

North   South   Transformer vaults

62nd
& 61st   Mechanical Room
Floor   Mechanical Room

Zone 1 (61st)

                                                vault
                                                #3
47th
Floor

Zone 2 (61st)

35th
Floor

Zone 1 (7th)

22nd
Floor

Zone 2 (7th)   vault
                                                #2

7th
Floor   Mechanical Room

                Zone 3 (7th)

3rd
Floor

                Lobby and cafeteria

                                                vault
                                                #1
                Basement Fan Room
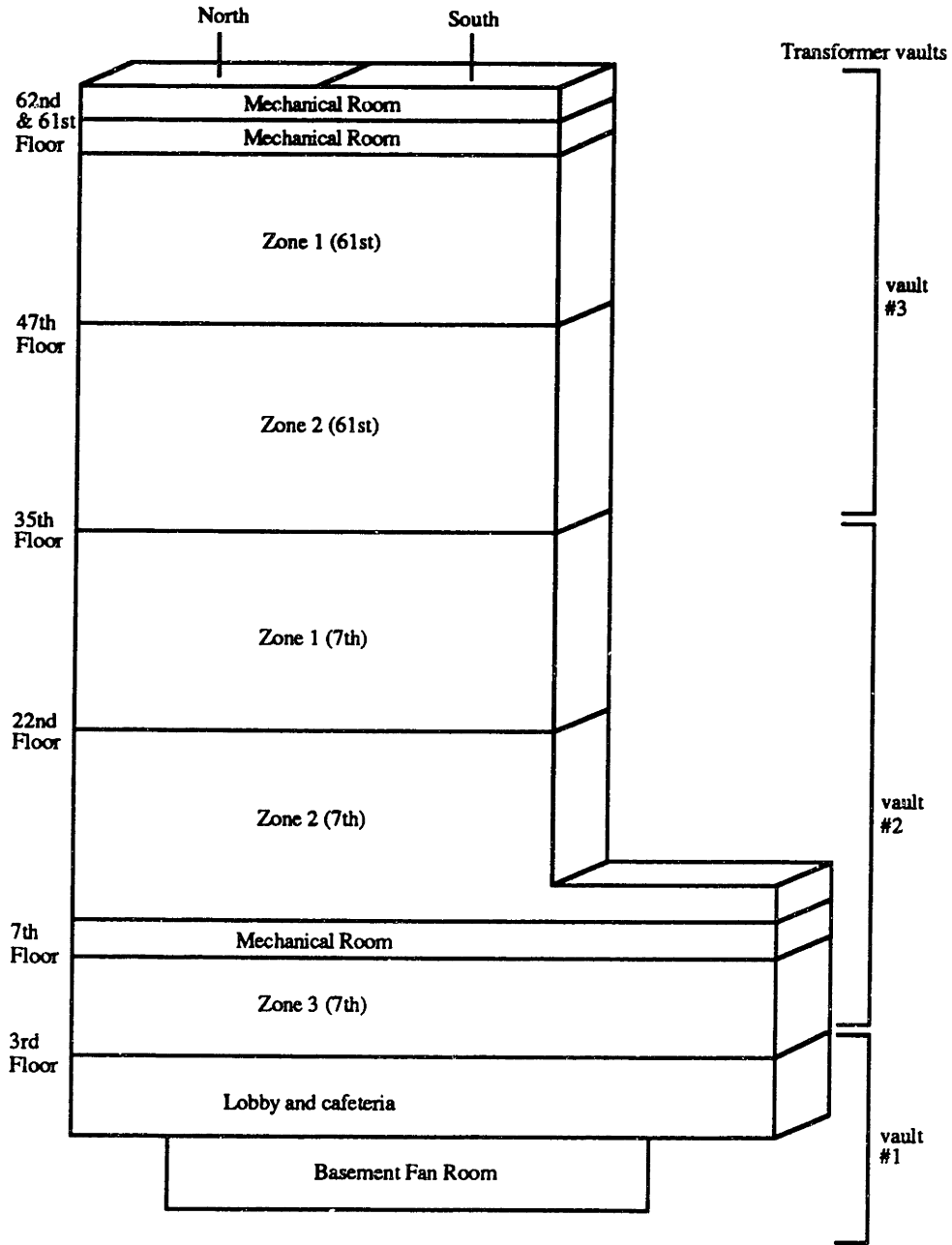
Figure 3.4: The John Hancock building layout with the different zones and transformer services indicated.

50

The Tower, being a tall building, must push domestic water up to the higher floors exceeding the water pressure supplied by the utilities. Therefore extra pumps are needed for the higher floors where the pumps fill reservoirs at various levels in the building. The pumps cycle as water levels get low.

General building equipment

Based on the Tower equipment lists in appendix B, tables 3.3 and 3.4 show the general sizing of the fans and pumps, respectively. Table 3.2 shows the chillers that the building used and the floors that house them.

Table 3.2: Centrifugal chillers of the John Hancock Tower

#	Size	kW electric	Location
1	1500 tons	1500	7th floor
1	1000 tons	1000	7th floor
1	550 tons	550	7th floor
2	1325 tons	1325	61st floor
1	400 tons	400	61st floor

Table 3.3: Fans of the John Hancock Tower

Fans	Exhaust		Supply		Return	
General	(5)	10 hp	(4)	250 hp	(8)	20-25 hp
Interior		---	(10)	60 hp	(11)	30 hp
Mech. room	(15)	40 hp		---		---
Kitchen	(2)	40-75 hp	(2)	15 hp		---
Toilets	(4)	15-20 hp		---		---
Misc.	(10)	0.8-5 hp	(15)	30-60 hp	(5)	7.5-15 hp

Note: The general supply fan is for the high pressure perimeter induction units.

Table 3.4: Pumps of the John Hancock Tower

Pumps	Chilled water	Condenser	Domestic water	Cooling tower make up	Misc
7th floor	(4) 200 hp	(6) 400 hp	(6) 100-150 hp	(3) 150 hp	(6) 0.75-5 hp
61st floor	(3) 100 hp	(5) 250 hp	(3) 100-150 hp	none	(5) 2-15 hp

It is noted here that the general equipment listed in tables 3.2 - 3.4 indicate that there are overlapping equipment sizes between equipment types, listed in table 3.5. This means that absolute power consumption alone cannot distinguish between equipment. There must be a method to decipher between equally sized equipment other than power consumption. A suggested method is described in chapter 4.

Table 3.5: Relative equipment sizes of the John Hancock Tower

Chillers	400 - 1500 kW
Supply fans	10 - 200 kW
Return fans	5 - 20 kW
Pumps	75 - 300 kW

Finally we note some specific information of the complex building system encountered.

HVAC specific information for the Tower

- Equipment floors are Basement Fan Room (BFR), #7, and #61
- There is a heat recovery chiller serving the computer room that runs 24 hours a day and is used to help heat the building.
- Normal heat is supplied by utility steam.
- There are perimeter induction units: 4 pipe system with no unit fans.
- The ventilation system is VAV with inlet vanes for the interior.

• For the three transformer vaults located on floors BFR, #7 and #62, the
configuration of the wiring is: transformer, to riser, to breaker that supplies three
floors north and south. Each transformer converts to 277/480V.

• Equipment start up time depends on outside temperature.

• Equipment shut down time is scheduled for 6 pm.

The building has an aggregate load monitoring system supplied by the utility

company. The data are delivered in watts at a 15 minute sampling interval. Figure 3.5

shows an example of a three-dimensional load profile of a data set over the month time

interval of 13 September 89 to 14 October 89, where the valleys in the plot indicate the

reduced weekend loads. The building operators use this profile for peak shaving strategies

in summer months. For our load monitoring applications, it would be descriptively

advantageous to have three-dimensional plots for individual equipment profiles.
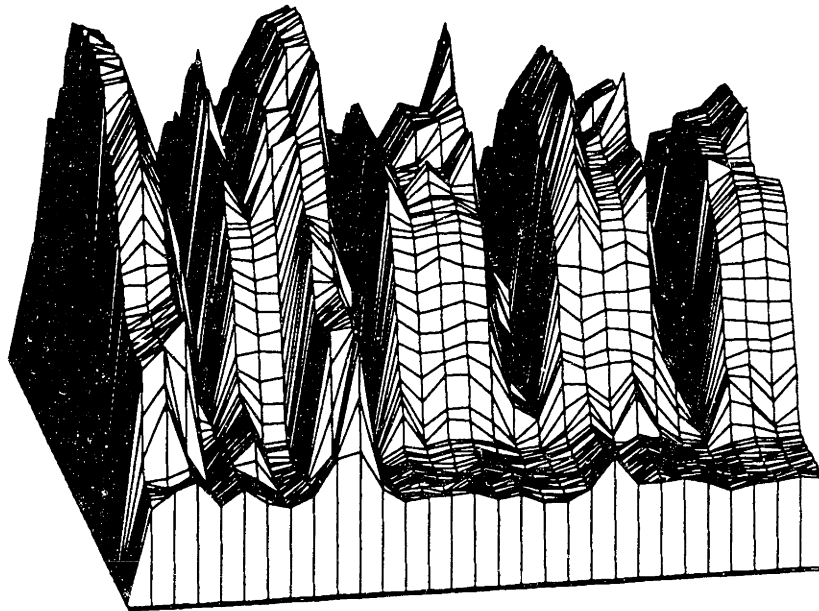
*Figure 3.5: A graphical representation of the John Hancock Tower's electrical energy
consumption from 13 September to 14 October 1989. The bottom axes are minutes for
each day going into the page, and days of the month running left to right.*

3.7.2 Bull building

The Bull building is a new three story office building in Billerica Massachusetts. It consists of 440,000 square feet rented to a single tenant, Bull HN Information Systems Inc. The building's HVAC equipment are described in tables 3.6 - 3.8.

Table 3.6: Refrigeration Equipment of the Bull building

Equipment	Capacity	Compressor	Condenser fan	Total kW
3 Rooftop chillers	167 ton 194 kW	4x50 hp	14x1 hp	194
	Ice making	Chilled water	Compressor	
3 Ice maker	160 ton ice	250 ton	363 hp	

Table 3.7: Pumps of the Bull building - one operating, one back-up

Function	type	hp
secondary chilled water	split case	30
primary chilled water	vertical turbine	40
recirc ice tank	vertical turbine	30

Table 3.8: Air Handling Units (VAV) of the Bull building

AHU	Supply hp	Return hp
1	(2) 60	(2) 20
2	(2) 75	(2) 30
3	(2) 75	(2) 25
4	(2) 50	(2) 20

The ice making equipment are typically operated during the night when electrical rates are down thus providing a cooling storage system for conditioning the next day. The rooftop chillers are used as backup for the ice makers and for cooling the computer room at night while the ice is being made. The ventilation system is VAV and time clocks with mechanical set points are used to start and stop the major equipment.

The electrical wiring configuration has the ice makers and the air handlers wired from four main banks of switchgear which also have tenant loads associated with them. The chillers and some pumps are wired from separate switchgear.

Also noted were the relative sizing of the equipment in the building. As noted with the John Hancock building, the different equipment types, as described in table 3.9, overlap in power. [Tabors, 1991]

Table 3.9: Relative equipment sizes of the Bull building

Chillers	190 - 325 kW
Supply fans	45 - 67 kW
Return fans	18 - 36 kW
Pumps	13 - 36 kW

The Bull building falls more into the wide building type in contrast to the Hancock Tower's tall dimension. Tables 3.5 - 3.6 reveal a marked difference in pump sizes with respect to the Tower. The fans are difficult to compare since the Tower has two large (250 hp) fans for perimeter induction units. However, using the interior VAV fans for comparison, there is a slightly larger fan size for the Bull building. Also a comparison of the Tower's 10 supply fans and 11 return fans to the Bull building's 8 supply fans and 8 return fans gives the Bull building a higher total power per square foot than the Tower by 1 $hp/600$ ft^2.

3.8   System descriptions

The next step in describing HVAC in a building is to obtain a level 2 list the types of systems found in buildings. In any type of load monitoring, knowledge of the HVAC system configuration is necessary. Such information as what equipment is time varying, how each piece of equipment is controlled, and what are the inter-dependencies between the equipment, all help understand the dynamics of the equipment load profiles. The knowledge can also be helpful in location equipment faults.

A dirty cooling coil provides an example of this kind of equipment knowledge and linking. In order to maintain the air temperature past the coils, the pump responsible for chilled water will respond to the reduced heat transfer by pumping more water through the coils. If the water through the coils is at full flow, and the air temperature is still high, then the chiller, which maintains the chilled water temperature, will experience a reduction in load. In addition, the temperature of the air being supplied to the zones will increase, causing the zones to request more air for cooling, and increasing the fan load. Finally, if the fan cannot supply enough flow to the zones (on a hot day) then the zones will heat up and affect the tenants. Because the system is automatically controlled, there are no indications that the system is running inefficiently until the tenants are affected.

Although this thesis makes no claims that the events just described can be detected in order to narrow down a dirty cooling coil, we do suggest that if the events can be measured, a rule-based system to ask IF...THEN questions could be devised to utilize the links of the HVAC systems [Huffer, 1989]. This method could be useful for detecting some HVAC faults.

Examples of different HVAC systems are listed in appendix A. For a more complete description the *ASHRAE Systems Handbook* gives finer details [ASHRAE, 1984]. The list also describes some typical uses of each system. However, what is most notable about these different systems is their diversity considering the elements of each of

the systems are pumps, fans, chillers and some non-electrical equipment. When first trying to categorize these systems, we felt a need to describe hybrid systems using common elements. Instead of listing and characterizing each possible system, we now suggest that a simpler and more efficient method is to link together equipment elements to form specific systems as they are encountered. What is meant by links is the physical-electrical or physical-physical connections between equipment elements. This method of elements and links is used by building simulation programs such as TRNSYS and HVACSIM+ with the intent of numerically modelling each element and simulating the total building response to specified initial conditions. The method we present here is based on similar reasoning, but we are not stressing the modelling and simulation which is very computationally intensive. Rather we are focusing on the links between the elements. These links provide a road map into the building and its cause and effect characteristics so that a human or a computer can ask IF...THEN questions.

For a element/link map, the following equipment element knowledge is noted:

Element
> A common type of equipment found in a building.

Specifications
> • varying/constant power equipment
> • setpoint data
> • type of actuator for control (if any)
> • size (pertinent units)

Load
> Variables that affect electrical load. If possible the theoretical physical model that determines the electrical load.

Control - setpoint
> The physical variable that controls the equipment and type of control if possible (e.g. PID).

Links
> Other equipment or physical variables the equipment is connected to in the building as a system.

Schedules

The times the equipment is turned on or off.

Element-links example

An example of a system with an element and link list compiled follows. The system is a typical variable air volume system.
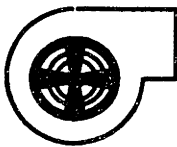
Supply fan with inlet vane dampers:

Specifications:
* Varying
* Set point:
* Actuator: inlet vanes
* Size: 100 hp

Load:

Determined by volume of air moved by fan; pressure drop across fan; and the efficiency of the fan and motor, which is a function of load. This can be expressed by a fan curve which depicts pressure versus air volume.

Model: Power = A + B(flow)+ C (flow)2 for a fixed pressure setpoint

Control - set point:

Static pressure downstream in duct.

Links:

a. Room air dampers.

b. Resistance to airflow.

Schedule:

weekdays	on: 7 am	off: 6 pm
weekends	on: 10 am	off: 4 pm
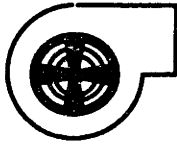
Return fan with inlet vane dampers:

Specifications:
> * Varying
> * Set point: Air flow rate
> * Actuator: inlet vanes
> * Size: 50 hp

Load:
> Determined by volume of air moved by fan; pressure drop across fan; and the efficiency of the fan and motor, which is a function of load. This can be expressed by a fan curve which depicts pressure versus air volume.
>
> Model: Power = A + B(flow)+ C (flow)2

Control - set point:
> Air flow rate = supply fan air flow rate - minimum outside air flow rate.
>
> (one form of return fan control)

Links:
> a. Supply fan flow rate.
> b. Room air dampers.
> c. Resistance to airflow through return air plena and ducting.

Schedule:
> | weekdays | on: 7 am | off: 6 pm |
> | weekends | on: 10 am | off: 4 pm |
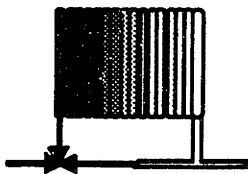
Cooling coils

Specifications:
> * Varying
> * Set point: Temperature of air past coils.
> * Actuator: Three-way bypass valve controlling flow.
> * Size: NA

Load:

> Heat removed from air = heat gained by coils.

> Model : $M_{air} C_p (T_{air,in} - T_{air,out}) = M_{water} C_{pwater} (T_{water,out} - T_{water, in})$

Control - set point:

> Temperature of air past coil.

Links:

> a. Supply fan flow rate (air flow over cooling coils)

> b. Inlet air dampers (temperature of air supplied to ventilation system)

> c. Chiller set point (temperature of water supplied to cooling coils)

> (NOTE: flow of water through coils is controlled by a valve)
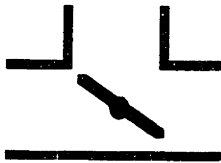
Inlet/return/relief damper

Specifications:

> Not applicable.

Load:

> Not applicable.

> (Note that if the dampers are driven by electrical motors the load is negligible.)

Control - set point:

> Minimum outside air: if $T > T_{high}$ or $T < T_{low}$

> Temperature: if $T_{high} > T > T_{low}$

Links:

> Outside temperature.

Zone dampers

Specifications:

> Not applicable.

Load:

     a. Not applicable.

     (Note that if the dampers are driven by electrical motors the load is negligible.)

     b. If dampers include fan powered boxes the fans could be listed here as fractional horsepower loads, or listed as separate elements.

Control - set point:

     a. Temperature of zone

     b. Minimum and maximum airflow

Links:

     a. Set point of cooling/heating coils (Supply temperature)
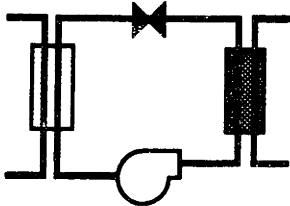
Chiller

Specifications:

     * Varying

     * Set point: Leaving chilled water temperature.

     * Actuator: Contained within chiller.

     * Size: 100 tons

Load:

     Determined by the amount of heat to be removed from the incoming
          evaporator water and the efficiency of the chiller at that evaporator load.

     Model: Biquadratic using chiller load and the temperature difference
          between the leaving condenser and chilled water flows. [Braun, 1987]

Control set point:

     Leaving chilled water temperature.

Links:

     a. Cooling tower (Incoming condenser water temperature)

     b. Cooling coils (evaporator water inlet temperature).

     c. % of full load at which the chiller is running.

Schedule:

weekdays	on: 8 am	off: 5:30 pm
weekends	on: 11 am	off: 3 pm
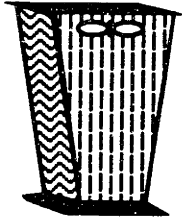
Cooling tower

Specifications:

* Varying - stepwise for fan

* Set point: Condenser water temperature.

* Actuator:

a. The temperature is controlled by a three-way bypass valve that controls the amount of condenser water that passes through the cooling tower.

b. There is often a fan included in the cooling tower unit in order to increase the evaporative cooling. This is a case where the HVAC element can be split into two linked elements or left as one.

* Size: NA

Load:

Determined by the amount of heat to be removed from the incoming condenser water. This is a function of the flow rate and the temperature of the water.

Control - set point:

Condenser water temperature.

Links:

a. outside environment - outside air temperature; outside air humidity.

b. chiller demand (temperature of water to be cooled).

c. tower fan (depending on how it is grouped).

Schedule:

weekdays	on: 8 am	off: 5:30 pm
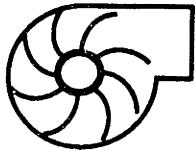
Pumps

Cooling tower pump

Specifications:
> * Constant
> * Size: 50 hp

Load:
> Pressure differential caused by path of water through the cooling tower or the bypass. The three-way bypass valve determines the path. The load also depends on the efficiency of the pump and motor. This can be expressed by a pump curve which depicts pressure versus flow.

Links:
> a. Cooling tower bypass valve.

Schedule:
> weekdays   on: 8 am   off: 5:30 pm

Chilled water pump

Specifications:
> * Constant
> * Size: 50 hp

Load:
> Pressure differential caused by path of water through cooling coils or bypass. The three-way bypass valve determines the path. The load also depends on the efficiency of the pump and motor. This can be expressed by a pump curve which depicts pressure versus flow.

Links:
> a. Cooling coil (bypass valve).

Schedule:
> weekdays   on: 8 am   off: 5:30 pm
> weekends   on: 11 am   off: 3 pm

The element/link database would provide a monitor with level 2 knowledge. For transient detection this information can help the a level 2 monitor over a level 1 in many ways. The element list provides a verification as to what equipment exist in the building, reducing the need for the monitor to reconstruct the equipment from the transient signature (which we do not know for sure is possible), while the size of the equipment in specifications can be correlated with the magnitude of the transient. We have also mentioned previously that there can be a difficulty with time varying loads in which they can have misleading transient magnitudes because of the time varying nature of the load. This can cause difficulty when matching start-ups with shut-downs. However, with sensors recording the independent modelling variable for the equipment, a transient will be indicated by a sensed change in the variable. This information will allow the time varying equipment to be identified and leave the equipment detection algorithms free to detect constant load equipment. In the case where there are no sensors on time varying equipment the schedules information can be used as a template which the transients should follow.

The load database provides information for disaggregation of time varying loads. Because the specifications supply the monitor with the information as to whether the equipment load will vary between start-up and shut-down, the constant load equipment can be subtracted off the total signal to leave only the time varying signal. The monitor, knowing what equipment are causing the variations, will the be able to use the monitor sensors and the load model to disaggregate the equipment load in software.

Finally, with the disaggregated information, the links could be used to track down any system abnormalities. Using an expert system with a knowledge base of how the equipment is controlled and linked in the physical system could provide answers to system faults where many equipment are involved.

In summary, we began by looking at buildings from an intrusive level 1 viewpoint, but we found that commercial buildings are difficult to generalize due to their diverse sets of equipment. The two buildings investigated have shown how buildings can have be a

complexity of systems and a simplicity of equipment. This apparent dichotomy serves as a motivation for our method of describing a building system which views the building at a level 2 perspective. This method shows that buildings can be described more simply. Furthermore, the information obtained by this viewpoint could be very useful for a load monitor trying to disaggregate the total power signal. With a background of general building equipment and their functions, we have shown that the level 2 informational method has applications of transient detection and system characterization. Finally, we have indicated that a level 3 hardware disaggregation can be used to increase the accuracy and reduce the complexity of the level 2 intrusiveness.

CHAPTER 4 - SOFTWARE DISAGGREGATION

4.1   Introduction

One application of a load monitor is to track the various loads in a commercial

building. In hardware, this is done by monitoring the separate pieces of equipment. For

an aggregate power signal, however, the disaggregation is accomplished in software. This

disaggregation requires a level 2 intrusiveness which states that the aggregate load signal

must be decomposed into pieces using knowledge gathered about the building using

independent data.

In this chapter we present a method to disaggregate a power signal into equipment

component profiles using software. The method is then tested using fan data taken from a

commercial building.

In determining the equipment models, we first distinguish between two classes of

equipment: time varying and constant power. For constant power equipment, the device is

meant to run at constant operating conditions:

$$Power = f(constant)$$

Power can vary due to system fluctuations and in these situations the constant

power model breaks down. We assume that the fluctuations in time are negligible and the

constant power equipment is stable in time. The constant power equipment is therefore

modeled by determining the value of the constant which can be accomplished using device

start-up transients handled in chapter 5. This discussion, then, focuses on time varying

equipment. For these devices, the power varies with some set of independent variables:

Power = $f(\xi_1, \xi_2, ..., \xi_n)$

To isolate this equipment, the constant power equipment will have to be subtracted off the aggregate signal containing the multi-variable time varying equipment.

Once a model of equipment has been suggested, and the power and independent variable data for the equipment taken, then an estimation of the parameters can be made using a parameter modelling technique such as least squares. Errors, confidence intervals, and fit indexes may be then obtained to judge the noise and adequacies of the model after a model has been fit. In the software disaggregation case, however, we have one more added "dimension." The power measurement is the sum of all the equipment powers. Consequently the complexity of our models increases as each piece of equipment is added to the power:

$$\text{Total Power} = f_1(\xi_1, \xi_2, ..., \xi_n) + f_2(\xi_1, \xi_2, ..., \xi_n) + ... + f_m(\xi_1, \xi_2, ..., \xi_n)$$

The problem of understanding a model using multiple inputs (independent variables), discussed here, is present in many fields of science. One example is the untangling of flows in a building using tracer gas measurements in the field of building air flow analysis where it is known as the inverse problem. In statistical books the problem is known a response surface analysis [Box, 1986].

Given this problem set-up, we are faced with two tasks. First, we must determine how the solution of the model can be used for identifying equipment characteristics and detecting faults. Second, we must understand how to calculate a solution to the problem. The solution should consist of a model fit that can reproduce the time varying signal given set of independent variables. With that solution, we must understand how the complexity of the problem causes our results to break down and how to gauge how much it has broken down.

4.1.1 Equipment characterization

With respect to the first task, parameter modelling is used to characterize the equipment or the system that we are monitoring. With some metric of characterization, we can devise a method with which to test whether the building is following its nominal character, or has deviated to some other characteristic. As dictated by imperfect measurements and noise, comparisons between two different characteristics lends itself to statistics and confidence intervals of models. A technique for detecting faults, or abnormal characteristics, is described by Pape and Mitchell [Pape, 1991]. A general description of their technique follows:

- Using a quadratic model for equipment, determine the independent variables to be used for a piece of equipment. Research by Braun has shown that fans, pumps, and chillers can be modelled as quadratics [Braun, 1987].

- Take measurements over the full range of the *healthy* piece of equipment.

- Fit the data to the model using a regression, recording the sum of the squared error.

- Three methods are used to detect faults in the equipment being monitored:

- **Method 1:**

  A statistical confidence region is defined using a t-distribution. The confidence interval is based on the error distribution of the equipment power as calculated using the scalar value of the sum of the squared errors.

  As real time data are taken for a system being monitored, the predicted power is then calculated using the model fit for normal operation. This value is compared to the real power and the difference of the two measurements is compared to the confidence region. If it lies outside the region, a fault is noted.

69

The confidence region calculated by this method has a fixed magnitude and therefore does not vary along the range of the fit.

• **Method 2:**

The cumulative sum of the errors between new values and the predicted values is calculated over time. Given a symmetric error distribution and no system faults present, the sum should maintain a near zero value. With a fault present, the errors will be biased towards a new model and the cumulative sum will grow in the positive or negative direction. Thus the absolute value of the cumulative sum can be checked against a threshold value for fault detection.

• **Method 3:**

A number of real time data points are taken for a piece of equipment being monitored. A fit to the new data is made and the models are compared. The comparison uses a t-distribution to test whether the two sample sets are significantly different.

In the context of a load monitor, methods 2 and 3 focus on longer time frames than method 1. Running over one day, the load monitor could run method 1 to continuously check for faults. Method two could be used for intermediate time intervals on the order of hours to check trends. Finally method three could be used at the end of the day over the whole data set.

We will focus here on method 1 for "real time" applications. A deeper understanding of the statistical model comparisons must therefore be clarified.

4.2   Models

The equipment models considered here are for the two largest, time varying electrical energy users commonly found in commercial buildings: fans and chillers. Work by Braun has shown that the power requirement for chillers, fans, and pumps can be adequately represented as quadratics [Braun, 1987]. Our analysis uses this body of work as a basis for which to characterize equipment.

We also assume that these models are correct in a general sense. If we were to properly form the equipment models for a building, we would have to consider each piece of equipment separately. For each test there would be an iterative process of analysis and conjecture between the model and experiment [Box, 1986]. In this way, the characteristics of each system-equipment combination could be accounted for. That is to say that the monitor uses general equipment models that cannot account for any individual system dynamics. With our monitoring design, the tuning of the sub-characteristics of the equipment models will rest with the building operator or experimenter, not on the monitor itself. We suggest that the design of the monitor allow for this type of adjustment.

4.2.1   Fan

The response variables for a fan are the pressure drop across the fan and the volume flow through the fan. Using Braun's empirical model:

$$\text{Electric power} = a_o + a_1\text{Flow} + a_2\text{Flow}^2 + a_3\Delta P + a_4\Delta P^2 + a_5\text{Flow }\Delta P$$

Where,

Flow   - volumetric flow rate

ΔP - change of pressure across fan

$a_0 - a_5$ - coefficients

For fan systems that are controlled using a fixed static pressure differential, it is found that the fan efficiency varies with flow, and the ΔP collapses into the coefficients:

Fan power $= a + b(\text{Flow}) + c(\text{Flow})^2$

If the static setpoint changes, then the system curve fit will change. Most variable volume fan system that we have encountered are controlled with a fixed static pressure, so we use this model for fans.

4.2.2 Chiller

Similar research has been applied to empirical models of chillers in which a quadratic model has been fit. The model was tested on both variable speed chillers and fixed speed chillers with variable inlet vane control.

The model requires a measurement of the chiller load and the temperature difference between the leaving condenser and chilled water flows for its two independent variables.

$$\frac{P_{chiller}}{P_{design}} = a_0 + a_1 X + a_2 X^2 + a_3 Y + a_4 Y^2 + a_5 XY$$

where

X - Ratio of the load on chiller to a design load

Y - Leaving water temperature difference divided by a design value

$P_{chiller}$ - Electrical chiller power

P_{design} - Electrical chiller power associated with the design condition

$a_0 - a_5$ - coefficients

[Braun, 1987]

4.3 Least squares

The method of least squares for parameter estimation is a well documented technique. Since our analysis involves potentially many models and configurations, we use the matrix form of the least squares analysis. The solution involves a number of aspects including model set-up, numerical methods, and pathological cases. We therefore cover these aspects rather than the theory of least squares which can be found in many textbooks [Draper, 1981].

4.3.1 Model set-up

The matrix form of a quadratic model such as a fan is formulated, using the fan model discussed:

$$\begin{Bmatrix} \text{Power}_1 \\ \text{Power}_2 \\ \text{Power}_3 \\ \text{Power}_4 \\ \cdot \\ \downarrow \end{Bmatrix} = \begin{bmatrix} 1 & \text{Flow}_1 & \text{Flow}_1^2 \\ 1 & \text{Flow}_2 & \text{Flow}_2^2 \\ 1 & \text{Flow}_3 & \text{Flow}_3^2 \\ 1 & \text{Flow}_4 & \text{Flow}_4^2 \\ \cdot & \cdot & \cdot \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{Bmatrix} a \\ b \\ c \end{Bmatrix}$$

where each row represents an observation. In compressed form this is:

$$\{\text{Power}\}_{n\times 1} = [\text{Flow}]_{n\times 3} \times \{\text{Coefficients}\}_{3\times 1}$$

which is in the general form of:

$$\{b\} = [A] \{x\}$$

with n equations and three unknowns.

Since there are more equations than unknowns, we have an overdetermined set of linear equations. The quadratic model can be considered a linear equation in terms of the coefficient matrix with the squared terms representing another independent variable, $\xi = \text{Flow}^2$. Solving the set of linear equations means solving the matrix equation:

$$\{x\} = [A]^{-1} \{b\}$$

The problem solution, disregarding model considerations such as noise and fit, depends on inverting the [A] matrix. Mathematically, inversion of a matrix requires that the matrix be square. However, the matrix equation can be solved in a least squares sense by

74

using the *normal equations*. These equations are formed by multiplying the matrix equation by $[A]^T$.

$$[A]^T \{b\} = [A]^T [A] \{x\}$$

This has the same form of the original equation $\{b\} = [A] \{x\}$, only the new power matrix is a 3x1 matrix, and the new [A] matrix is a 3x3 square matrix which is invertible. Thus the solution to our problem using the normal equations is:

$$\{x\} = \left([A]^T [A] \right)^{-1} [A]^T \{b\}$$

For a set of overdetermined equations, the normal equations give a solution for $\{x\}$ that is not necessarily exact, rather among all possible vectors, the solution is the closest in a least squares sense such that:

$$r = | [A]\{x\} - \{b\} |$$

Where r, called the residual of the solution, is minimized.

A common problem with the normal equations is that they suffer from roundoff errors because of the number of matrix operations performed. We now present a technique that reduces this problem.

4.3.2 Singular value decomposition

To get a numerical inverse of [A], a technique called singular value decomposition (SVD) is used. SVD is well suited for least squares analysis because it has two advantages

to the previously mentioned normal equations. SVD avoids numerically produced round-off errors by reducing the number of matrix operations performed. It also allows for "solutions" of normally pathological, singular, cases such that the singularity is pinpointed and can be removed. This provides a method for solving the system of equations as completely as the data permit without the computational errors and singularities that hinder the normal equations.

SVD is based on a theorem of linear algebra that states: "Any MxN matrix [A] whose number of rows M is greater than or equal to its number of columns N, can be written as the product of an MxN column-orthogonal matrix [U] , an NxN diagonal matrix [W] with positive or zero elements, and the transpose of an NxN orthogonal matrix [V]" [Press, 1988]

$$[A] = \begin{bmatrix} & & \\ & U & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix} \begin{bmatrix} & & \\ & V^T & \\ & & \end{bmatrix}$$

Since an orthogonal matrix has the property $[U]^T = [U]^{-1}$, and a diagonal matrix, $[W]^{-1} = \text{diag}(\frac{1}{w_i})$, the the solution to the general equation $\{x\} = [A]^{-1}\{b\}$ is:

$$[A]^{-1} = [V]\left(\text{diag}(\frac{1}{w_i})\right)[U]^T ;$$

$$\{x\} = [V]\left(\text{diag}(\frac{1}{w_i})\right)[U]^T\{b\}$$

By theorem, we are guaranteed the SVD, therefore the only problem that can arise is if any of the w_i's are zero or very close to zero so that its size is dominated by round off

error or computer precision. In this case the [A] matrix is considered singular, and an index of the degree of singularity is defined as:

$$\text{Condition number} = \frac{\text{Max}(w_i)}{\text{Min}(w_i)}$$

This number is referred to as the *condition number* of the matrix [A], where a matrix is defined as singular if the condition number is infinite. If the condition number is large then the matrix is called *ill-conditioned*. In the case of numerical computation, too large means if the the reciprocal of the condition number approaches the machine's precision (10^{-12} for double precision) [Press, 1988].

Figure 4.1 graphically shows two cases of conditioning. A solution to a set of equations which depends on the precision to which the equations are known, and the data space they represent. The ill-conditioned case clearly shows a large region of uncertainty while the well conditioned case reflects the uncertainty of the data. The condition number resulting in the least possible uncertainty for any set of data is 1.

When a matrix is singular or ill-conditioned, it implies that a linear combination of columns in the matrix equals or nearly equals another column in the matrix. Again the closeness is measured by the condition number and is referred to as *rank deficiency* of a singular matrix. In the case of the flow model given, a linear combination can form in the columns of the [A] matrix if the flow does not vary over a large range and flow $\approx$ (flow)^2. It will be shown later that the problem of rank deficiency causes more difficulty when modelling multiple equipment.

The way the specter of the rank deficiency is avoided or at least controlled is to reduce the complexity of the model by collapsing the deficient column that provides no new information into the columns it is dependent on. This implicitly reduces the coefficient
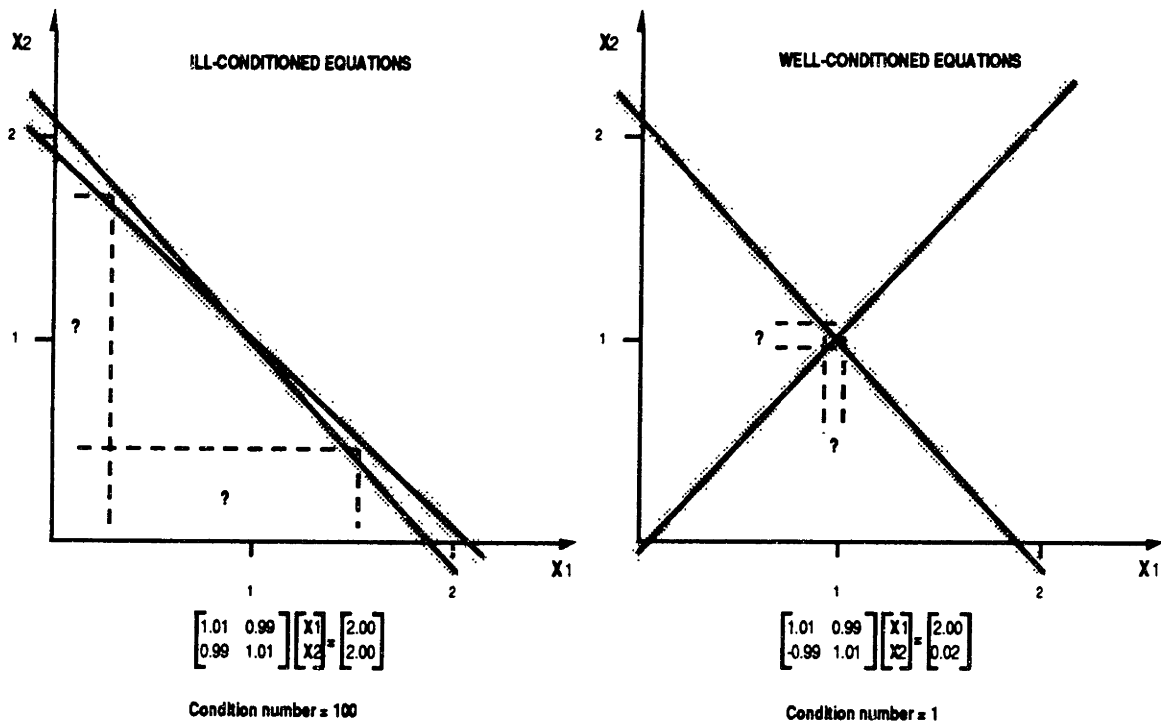
77

Figure 4.1: A graphical representation of well- and ill-conditioned sets of equations. Each equation has the same solution, but different data sets. Figure a) shows how uncertainties in knowledge of the linear equations, represented as the grey area around the lines, is magnified when the lines describe a similar space. Figure b) shows two equations covering an orthogonal data space, where the uncertainties of the data reflect the uncertainties of the solution. [Gerald, 1989]

matrix by one and is achieved by making the small w_i's to be zero after they are inverted, thus zeroing the corresponding answer for that coefficient. The remaining coefficients then solve the "new" reduced model. Using the fan model example, this would mean collapsing the quadratic term into the linear term and the model would become a linear model. This makes sense since over a small interval quadratic functions can be approximated as linear.

78

4.3.3 Errors based on the covariance matrix

In solving the equation $\{b\} = [A]\ \{x\}$ in a least squares sense, a model fit and errors are calculated for the equipment. The fit represents the equipment's characteristic response to its independent variables. As new data are collected from the building, they can be compared to the model fit as in method 1. The result of the comparison should statistically indicate that the new data lie either within the bounds of the characteristic fit, or in some other deviated model fit. A common method for indicating the bounds of a model fit is by confidence intervals that are constructed using the covariance matrix of [A]. Inspecting a covariance matrix generated from simulated data shows that there is valuable information contained in data and model descriptions that can help augment the scalar confidence region defined by Pape as the data vary with the independent variable. As will be shown, the confidence interval is not a constant but varies with the independent variable. Furthermore, the nature of the [A] matrix itself has a large bearing on how the confidence intervals are shaped.

Consider the covariance matrix of the set of equations $\{b\} = [A]\ \{x\}$:

$$Cov = s^2(A^TA)^{-1}$$

Now, let $\{A_0\}$ be a vector of predictor values. For instance, for our fan model, $\{A_0\}$ would be a 3x1 matrix for one data observation of flow such that:

$$P_T = \{A_0\}^T\{x\},$$

where P_T is a scalar estimate of power.

The variance of the power estimate is then:

$$\text{Var}(P_T) = \{A_0\}^T (\text{Cov}) \{A_0\}$$

$$= s^2 \{A_0\}^T (A^T A)^{-1} \{A_0\}$$

Where s is the estimated variance of the least squares fit.

Let $s_{P_T} = \sqrt{\text{Var}(P_T)}$, therefore the confidence interval for P_T is:

$$P_T \pm t_{n-p}(\alpha/2)\, s_{P_T}$$

$t_{n-p}(\alpha/2)$ is a two sided t-distribution with n-p degrees of freedom. Here n is the number of data points and p is the number of parameters being fitted. [Rice, 1988]

It should be emphasized that P_T describes the confidence interval of the model fit and not the confidence interval of the data about the fit.

The role of the variance of P_T on different models is illustrated in figure 4.2 where variance is plotted against flow for both a linear and quadratic model fit. The data points, and the noise are simulated, with $\sqrt{s^2}$ being 20, or about 10% of the Y value.

The parabolic shape of the P_T estimate of figure 4.2b indicates that the uncertainty of the model fit increases at the ends of the linear fit. This effect can be understood if one thinks of many different data sets taken independently for that same equipment. Considering the linear model, a fit to ten sets of data would produce ten lines with slight perturbation varying around the true line. The different lines would agree most often in the center of the data and disagree most markedly at the ends, which is illustrated by figure 4.2b. Figure 4.2c and 4.2d show the same simulation for a quadratic fit.

Now, in considering the variance of P_T the most notable effect is that the dependency of the model variance on the noise of the data is represented by a scalar, s^2. Therefore, changing the noise of the simulated data only scales the magnitude of the model variance, not the shape. The [A] matrix and the $\{A_0\}$ vector, that depend on the independent variables, determine the shape.
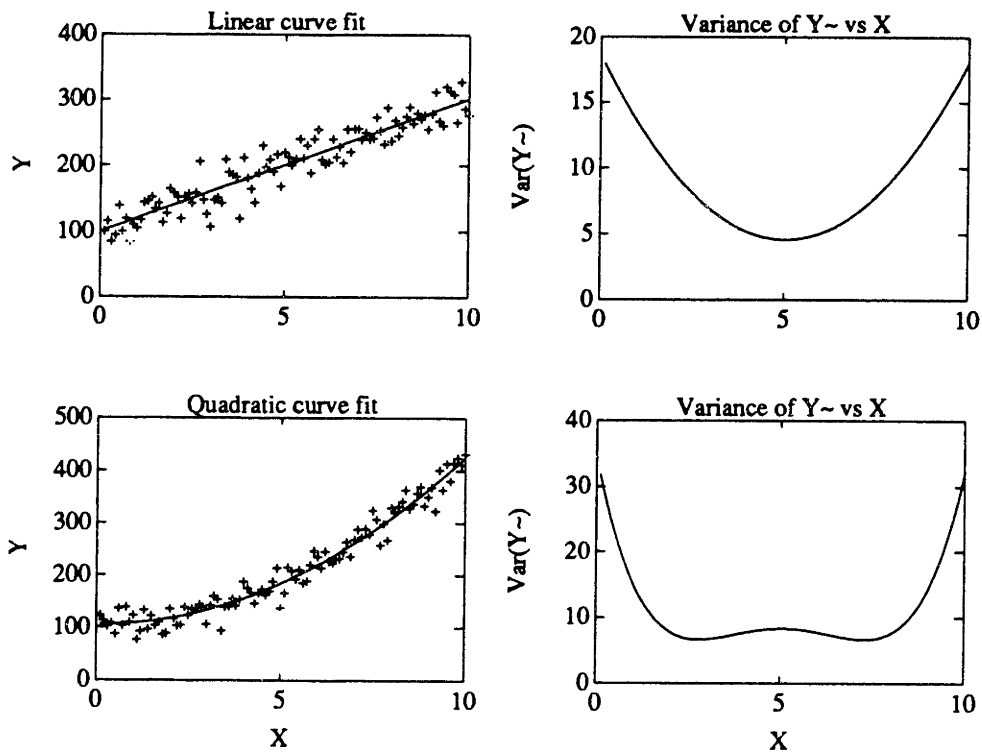
Figure 4.2: *Examples of model variance. From top to bottom, left to right: a) A linear curve fit to simulated data. The fitted line is Y~. b) A plot of the variance of the estimated linear curve Y~ versus X. c) A quadratic curve fit to simulated data. The fitted curve is Y~. d) A plot of the variance of the estimated quadratic curve Y~ versus X.*

An example of this dependency on the [A] matrix is shown in figure 4.3a. This set of data simulates more realistic flow and power data, where data are clustered into areas of frequent fan operation. The resulting model variance plot, figure 4.3b, shows the region with the lowest variance is at the top end of the middle and densest cluster.

As previously emphasized, the meaning of Var(P_T) and the confidence interval associated with it refers to the model fit and not the points about the model. As described above, small differences in the data set cause perturbations in the model fit. Therefore the 90% confidence interval described by Var(P_T) and the t-distribution is the band that contains the *true* model 90% of the time the experiment is run. With this confidence

81

interval for our simulated model, we now find the confidence interval for a predicted value around the calculated model fit.
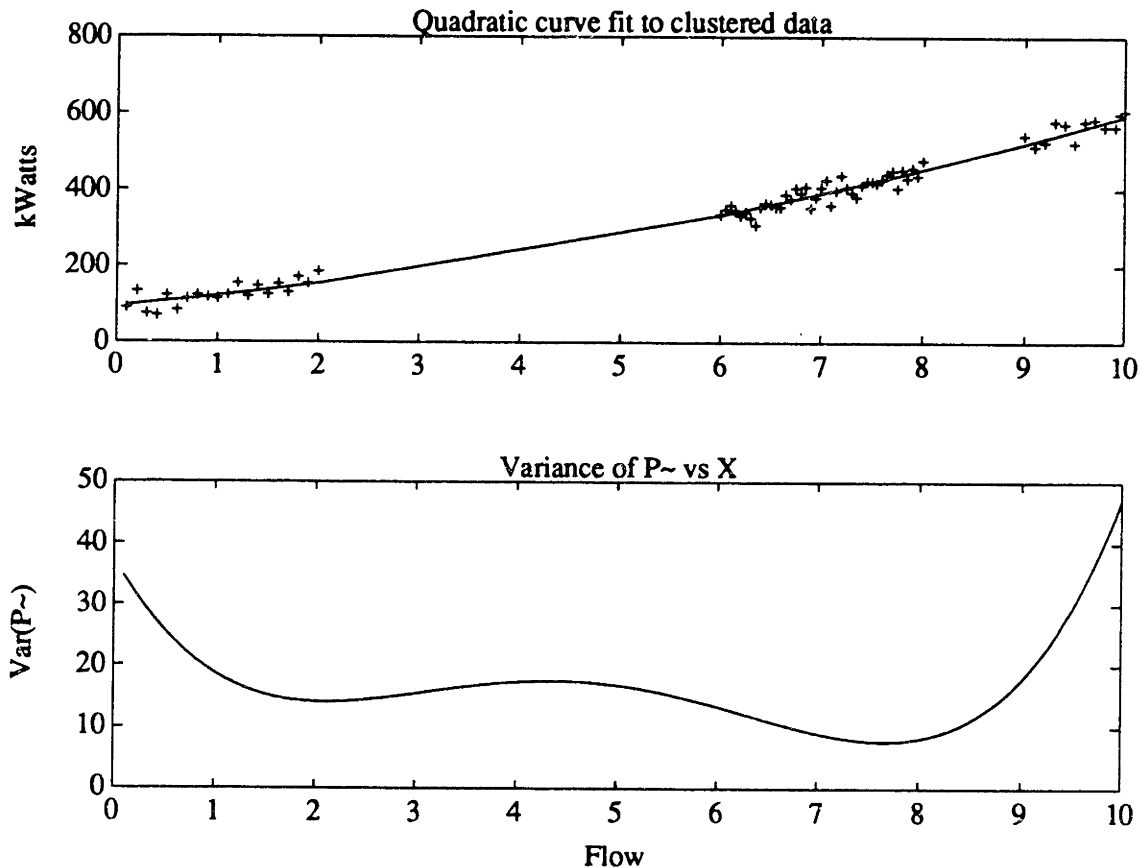
Figure 4.3: An example of possible fan power data versus flow. a) The line through the data points is a quadratic curve fit P~. b) A plot of the variance of the estimated P~ versus flow. Note how the lowest variance is near the upper part of the middle and most dense cluster.

A predicted value 90% confidence interval is a region that will contain 90% of future observations from the same model distribution. In simpler terms, with a predicted value region we can be 90% sure that a new data point will fall into the region given the same system characteristics, model, and error distribution. This region represents the information necessary to predict faults as described by Pape.

82

For any distribution, the confidence interval for future observations depends on the distribution of errors from the mean. If we combine the variance of the distribution of errors about the model with the variance of the model fit, the result is a confidence interval for a future value that considers the uncertainty of the model fit.

With normally distributed errors and a confidence interval for the model fit, the predicted value confidence interval is therefore the sum of the error confidence and model confidence interval or:

$$Var(P_{Future}) = Var(P_T) + s^2$$
$$= s^2 \left[\{A_0\}^T (A^T A)^{-1} \{A_0\} + 1 \right]$$

and,

$$\hat{P}_{Future} = P_T \pm t_{n-p}(\alpha/2) \, s_{P_{Future}}$$

An example of the shape of the model fit and the $\hat{P}_{Future}$ regions are shown in figure 4.4 using the simulated data model of figure 4.2c and a 90% confidence interval. The dashed lines close to the model fit show the model confidence intervals, while the outer dotted lines enclose the predicted value region. Of course these data have been simulated with even spacing along the range of values. Any number of confidence region shapes are possible as figure 4.3 indicates.

The 90% predictive confidence intervals for a linear curve fit have increasing errors at the ends of the fit with the narrowest interval in the middle, which indicates that the varying nature of the confidence interval depends on the model used [Draper, 1981]. These shapes have not been accounted for in previous work on HVAC fault detection.
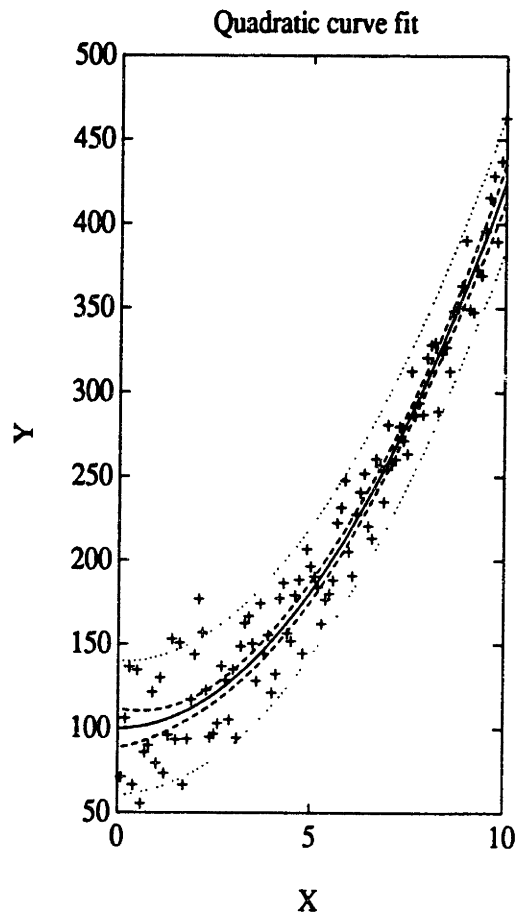
Figure 4.4: 90% confidence intervals for a simulated quadratic model. The dashed lines are the confidence intervals for the quadratic model fit represented by a solid line. The dotted lines enclose the confidence region for predicted future data values.

In a real modeling situation, errors will occur in both the dependent (power) and independent (flow) variables. The errors may or may not be normal. It has been stated that regardless of the error types and or dependencies the shape of the variance plot will depend on the distribution of the independent variable. This is true only if the assumptions we make when applying an unbiased least squares estimate hold. These assumptions are that the errors in the variables are normally distributed and have a mean of zero. Any deviation from these assumptions means that the variance plot does not necessarily describe the situation.

In summary, the errors considered in this discussion concentrate on the natural errors that occur from fitting a model to a set of data. These errors vary with the set of independent data collected, and the type of models used. However, this analysis demonstrates how the covariance matrix can aid in describing the errors as they vary with flow. Furthermore, we have augmented the confidence intervals that have been used to predict future observations. With a method for creating confidence intervals for equipment has been presented, we now look at difficulties in achieving this calculation.

4.4   Multiple Fans

We now consider an exercise in software disaggregation of multiple equipment from one aggregate power signal using independent variables. Considered here is the case of two fans with total power and the flows of both fans measured. The matrix representation is:

$$
\begin{pmatrix} P_T \\ \downarrow \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{bmatrix} 1 & F_1 & F_1^2 & F_2 & F_2^2 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{pmatrix} a_1+a_2 \\ b_1 \\ c_1 \\ b_2 \\ c_2 \end{pmatrix}
$$

where,

P_T - Total power

F_1 - Fan 1 flow

F_2 - Fan 2 flow

a_1, b_1, c_1 - Fan 1 coefficients

a_2, b_2, c_2 - Fan 2 coefficients

This has the dimensions:

$$\{\text{Total Power}\}_{n \times 1} = [\text{Flows}]_{n \times 5} \times \{\text{Coefficients}\}_{5 \times 1}$$

which is in the general form of:

$$\{b\} = [A] \{x\}$$

Notice that the constant parameters for each fan are lumped into one coefficient shown as $a_1 + a_2$. This is necessary because otherwise the corresponding data in the [A] matrix would form a linearly dependent columns and the matrix would be singular as discussed.

Visually, the two fan model can be expressed as a surface in the positive three-dimensional coordinate system. The power for any combination of flows fall on the surface, which when fitting the model is approximated in a least squares sense (see figure 4.5). In terms of experimental design, with the condition number and the covariance matrix as our indicators to how well the data spans the models space, the following question is posed. What set of data gives a good fit, and what set of data give a bad or insufficient fit? The motivation here is to have the monitor be responsive to when good fit data is being acquired.

In a building, the response of the fans depends on the excitation of the building thermal load. This is uncontrolled from the monitor's perspective. Therefore the data set by which the monitor processes data is at the mercy of the building and its environment.

86

This environmental determinacy can give rise to two modeling problems, collinearity and lack of variance.

4.4.1 Collinearity

When modeling two devices using a least squares method, a difficulty arises when the independent variables for the devices are collinear or linearly dependent. As previously stated, the [A] matrix in the expression $[A]\{x\} = \{b\}$ becomes rank deficient and thus singular. The problem is then unsolvable in its configuration. However, by collapsing the independent measurements into one set of independent variable data, the problem can be solved with the resulting coefficient vector, $\{x\}$, containing the sum of the coefficients from the two models.

This problem arises when modeling a typical ventilation system, comprised of two fans, a supply fan, F_1, and a return fan, F_2.

Now let the return fan be collinear with the supply fan, but lag it in magnitude by an amount J, where J ranges within the bounds of the supply fan. In other words, the maximum value J can have is the minimum value of the supply fan flow, P_1 .

$$F_2 = F_1 - J$$

This is often the case in practice. The airflow J represents air that slightly pressurizes a building and leaves via leakage and exhausts in electrical closets and toilets. If we know the value J, the fan model fits can be separated as follows.

Let F_1 be our independent modeling variable F_T,

$$P_T = a_1 + a_2 + b_1F_1 + b_2F_2 + c_1F_1^2 + c_2F_2^2$$

$$= a_1 + a_2 + b_1F_1 + b_2(F_1 - J) + c_1F_1^2 + c_2(F_1 - J)^2$$

$$= (a_1 + a_2 - b_2J + c_2J^2) + (b_1 + b_2 - 2c_2J)F_1 + (c_1 + c_2)F_1^2$$

A least squares fit to total power P_T using F_1 as the independent variable will determine our modeling parameters in the three combinations:

$$A = a_1 + a_2 - b_2J + c_2J^2$$

$$B = b_1 + b_2 - 2c_2J$$

$$C = c_1 + c_2$$

We therefore have three equations and six unknowns, and need three more equations.

Three equations can be obtained by using power and flow information from the daily start-up and shut-down transients.

$$\left.\begin{array}{l} P_{1i} = a_1 + b_1F_{1i} + c_1F_{1i}^2 \\ P_{2i} = a_2 + b_2F_{2i} + c_2F_{2i}^2 \end{array}\right\} \quad \text{start-up transients at i}$$

$$\left.\begin{array}{l} P_{1f} = a_1 + b_1F_{1f} + c_1F_{1f}^2 \\ P_{2f} = a_2 + b_2F_{2f} + c_2F_{2f}^2 \end{array}\right\} \quad \text{shut-down transients at f}$$

In matrix notation:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & -J & J^2 \\ 0 & 1 & 0 & 0 & 1 & -2J \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & F_{1i} & F_{1i}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & F_{2i} & F_{2i}^2 \\ 1 & F_{1f} & F_{1f}^2 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \end{Bmatrix} = \begin{Bmatrix} A \\ B \\ C \\ P_{1i} \\ P_{2i} \\ P_{1f} \end{Bmatrix}$$

This is in the form of $[A]\{x\} = \{b\}$.

Initially, when this set of equations was worked out on a computer, the $[A]$ matrix turned out to be rank deficient. On further analysis, it can be shown that the reason for this rank deficiency is the fifth row. It was found that in the system of equations, the condition of $F_2 = F_1 - J$ was used for the transients so that $F_{2i} = F_{1i} - J$. The result of this relation is the first, second, and third rows in the $[A]$ matrix become a linear combination of the fourth and fifth rows (also the fifth and sixth rows). This results in the matrix being rank deficient.

In order to maintain six linearly independent equations so as to solve for the coefficients of the fans, either the start-up or shut-down transients must not have the same collinear relation as the two tracking fans. If we assume that the fans start-up into a collinear relationship and then increase their flow to compensate for building pull-down, acceptable start-up measurements could be achieved by staggering the start-up times of the fans. Alternatively, if the warm-up period of the each fan does not follow the collinear relationship then start-up staggering would not be necessary.

At shut-down the fans would be tracking each other. Therefore, in order for these point to be used in the $[A]$ matrix, the shut-down events must be staggered *and* the flow must vary over the staggered interval. Without the flow varying the $F_{2f} = F_{1f} - J$ relationship would hold and the matrix would be rank deficient. We note that depending on

the flow to vary at shut-down is a less appealing algorithm that staggering the start-up events.
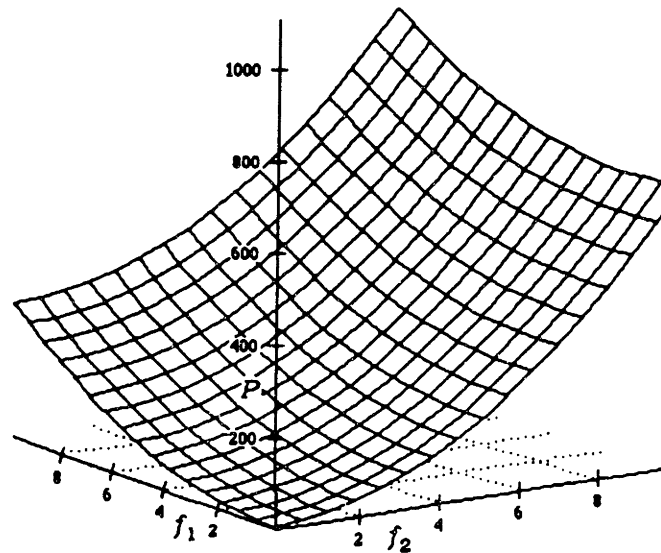
Figure 4.5: An example of power depending on the quadratic of two flows. The resulting least squares estimate must fit the surface.

It has previously been reported that the issue of collinearity between equipment such as fans is a cause for difficulty. However, the process of disaggregating the power of two fans with independently related flow is a three dimensional surface problem (see figure 4.5), whereas the same two fans with collinear flow patterns reduce the problem to a two dimensional curve problem described in figure 4.6. From a curve fitting perspective, the errors associated with a three dimensional problem as compared to a two dimensional problem are much more severe. Therefore collinearity can help the accuracy of model fits when one piece of equipment accurately tracks another piece of equipment.

In conclusion, if we are given two fans with collinear flow measurements to be modeled as discussed, then the total power measurement alone is adequate to determine the individual fan coefficients provided their is no collinearity in the start-up measurements or shut-down measurements.
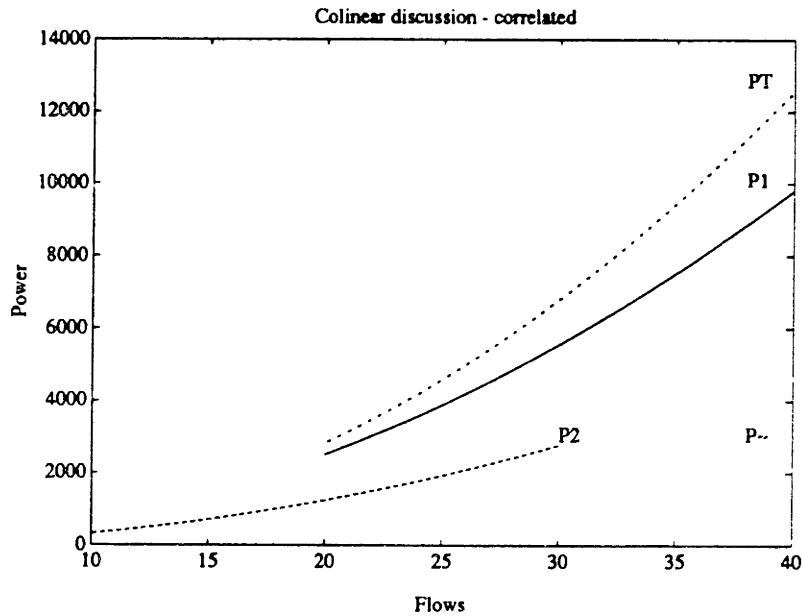
Figure 4.6: An example of how a two-dimensional least squares estimate becomes one-dimensional when the fans have collinear flows. PT is the total power, while P1 and P2 are the powers of the respective fans. P2 is plotted with the F1 scale to show that it has different flow magnitudes. When the three-dimensional surface plot of PT versus F1 and F2 is collapsed into a two dimensional curve versus F1, P~ is the new P2 in the F1 frame of reference.

4.4.2 Lack of data variance over the total fan range

To judge a model fit, the statistical application of the model must be considered. Because the model fit will be used to predict equipment operation, it is desired that the predictive capabilities cover the range of equipment operation. Considering the region of

frequently operation is also important. Since the equipment spends the most time there, the model should be accurate there. These two criteria, however, have conflicting attributes which are model bias and model spread. Model bias is introduced in the fit due to the model's, in this case a polynomial, inability to fit the real response of the equipment. As the range over which the equipment is measured increases, the lack of fit of the polynomial causes the confidence in the fit to decreases. If the range of measurement is decreased, the polynomial will tend to fit the model better. However, the confidence in the ends of the model will decrease due to the lack of range in which to fit the model.

It is desired for the monitor to track the equipment characteristics change as they vary with time, so we must consider how we are mapping the characteristics. If, as suggested, the equipment is run over its full range in some test/set-up run, the equipment could be well characterized and the fit can be used to test the equipment in real time. However, a problem arises if a change in the equipment characteristic is detected. In order to compare the new characteristic against even newer data there must be a significant number of new data points to form a new model fit. But the luxury of running the fan though its range is not practical. Therefore the building would have to randomly allow the monitor to collect the necessary data for a well-conditioned fit.

This difficulty is further complicated by the results from the fan data in the next section, where there are two distinct characteristics that depend on certain daytime system conditions that may not be possible to recreate when gathering the data for the test/set-up run. The data may therefore be entirely up to the buildings fluctuations to collect. This makes problems like collinearity and lack of variance significant.

The situation of a lack of variance in the flow data set can both be a difficulty or a plus depending on one's intent. One positive result of a lack of variance is that the model becomes simpler. Equipment with a normally quadratic response can be considered constant with no variations, linear with small variations, and quadratic with significant variations. The SVD procedure should help in determining which model should be used

based on the rank of the data matrix. A problem can be encountered with a lack of variation if normally varying equipment remains constant for a period of time. This situation essentially connotes a period in which little predictive model information can be gathered.

With the two-fan model where the model fit is the surface described by figure 4.5, an interesting advantage can be gained when one fan is constant in power while the other varies. This is a case where the model collapses into a single fan model (the fan that is varying). Intuitively, if one fan varies and the other is constant, the response present is seen to be due to the the varying fan. Visually, on the two dimensional surface plot, this represents a line running up the curved surface in one fan dimension. Here again the model simplifies, from two quadratics to one. And again the SVD method can help determine when a fan is constant enough to eliminate it from the model fit. This decision can be set by how small a w_i has to be in order to set its reciprocal to zero in the SVD least squares equation. In general, any situation where the data fall on a line in the response surface representation of the model, the model can be simplified to a lesser dimension.

4.5   Fan tests

Application of the methods discussed in this chapter were performed on two supply and return fan pairs. The fans tested are part of a variable air volume (VAV) system in an office building (E23) on the MIT campus. The VAV fans are controlled by Allen-Bradley variable speed drives. Control variables for the speed are static pressure at a sensor located approximately two thirds down stream in the duct. The specific fan sizes are:

Fan pair 1:   supply: 40 hp return: 10 hp
Fan pair 2:   supply: 40 hp return: 7.5 hp

In term of the building information discussed in chapter 3, we have:
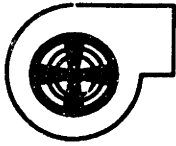
Supply fans with variable speed drives:

Specifications:
* Varying
* Set point: static pressure
* Actuator: digital
* Sizes: 40 hp and 40 hp

Load:
  Model: Power =
    A + B(flow)+ C (flow)2
Control - set point:
  Static pressure downstream in duct.
Links:
  a. Room air dampers.
  b. Resistance to airflow.
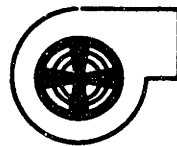
Return fans with variable speed drives:

Specifications:
* Varying
* Set point: static pressure
* Actuator: digital
* Sizes: 10 and 7.5 hp

Load:
  Model: Power =
    A + B(flow)+ C (flow)2
Control - set point:
  Supply fan setpoint minus a
  constant
Links:
  a. Supply fan flow rate.
  b. Resistance to airflow through
  return air plena and ducting.

Both fan pairs have the return fan track the supply fan based on static pressure bringing up the question of how closely the two fans track. It was found that on a minute to minute time frame there are variations in the power ratio of the fans, however the correlation between the fans over a day is very strong. The calculated correlation for the day shown in figure 4.7 is 0.98. Such a strong correlation leaves the task of separating the supply and return fan powers to transient detection, which must provide a start-up and shut-down power value for the fans to be disaggregated.

94
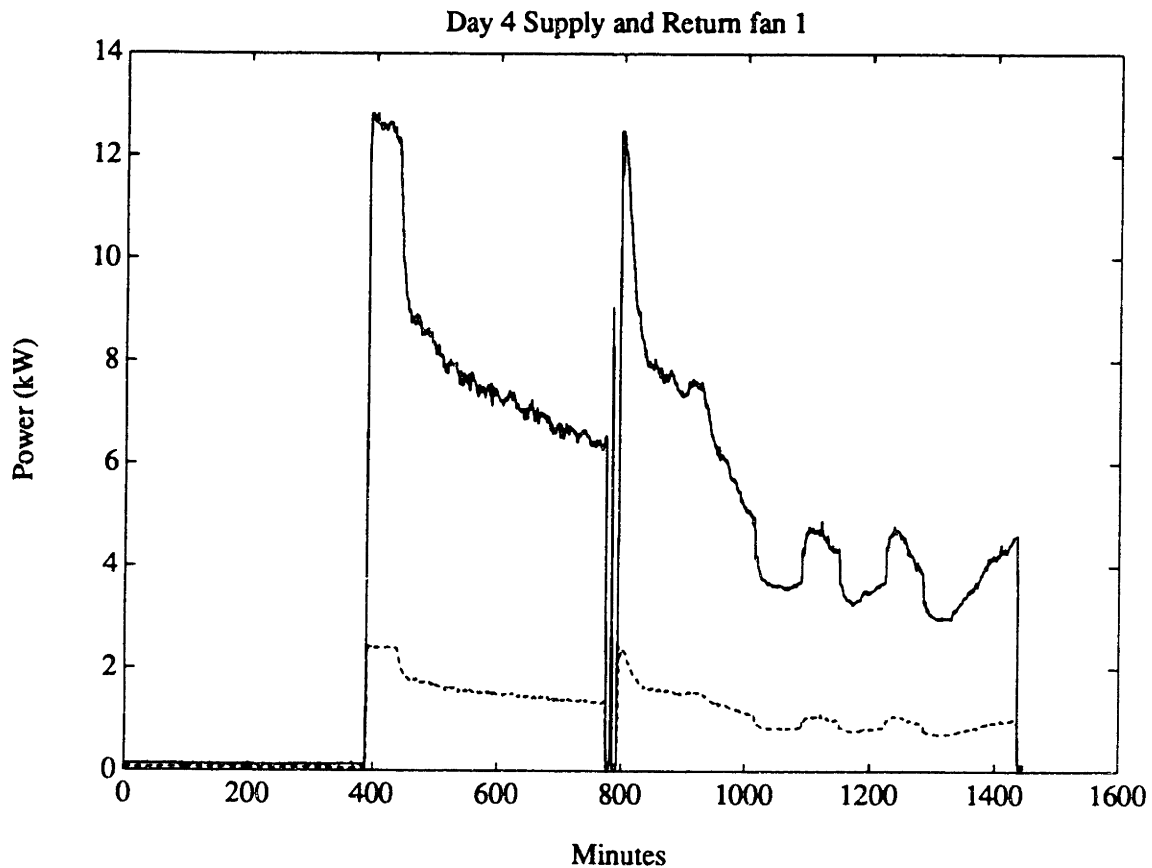
Figure 4.7: Electrical power use for fan pair 1 over one day. The
supply fan profile is the solid line, while the return fan is the dashed
line. One data point was take every minute.

The flow measurement of the supply fans consists of an array of pitot tubes

measuring average velocity pressure across the duct just after the fan. The pitot array

output is connected to a Setra 264 differential pressure transducer with a 0-5 volt output.

Data were collected using a permanently installed Synergistics load monitoring system.

This system collected power and flow data from the two fan pairs at a rate of 0.25 Hz,

which the monitor averaged and recorded every minute.

The tests were designed to:

• Characterize a system
• Apply a flow model to a fan

- Look at the errors of the fits
- Apply the software disaggregation method

4.5.1 Characterizing a system

Figure 4.8 shows a plot of power versus flow for supply fan 1 over a three weekday period. What is obvious from the plot is that the fan has three characteristic modes: one lower curve, one upper curve, and one maximum power plateau.

Figure 4.8: Power versus flow plot of supply fan 1 over a three weekday period. The upper curve corresponds to intervals of decreasing fan power and steady state, while the lower curve corresponds to intervals of increasing power. The plateau is at maximum fan power.

On further investigation it was discovered that the lower curve corresponds to regions where the fan power and flow were increasing, whereas regions of decreasing and steady state values fall on the upper curve. Also, the plateau at the top end of the plot represents times when the power has reached a maximum. Th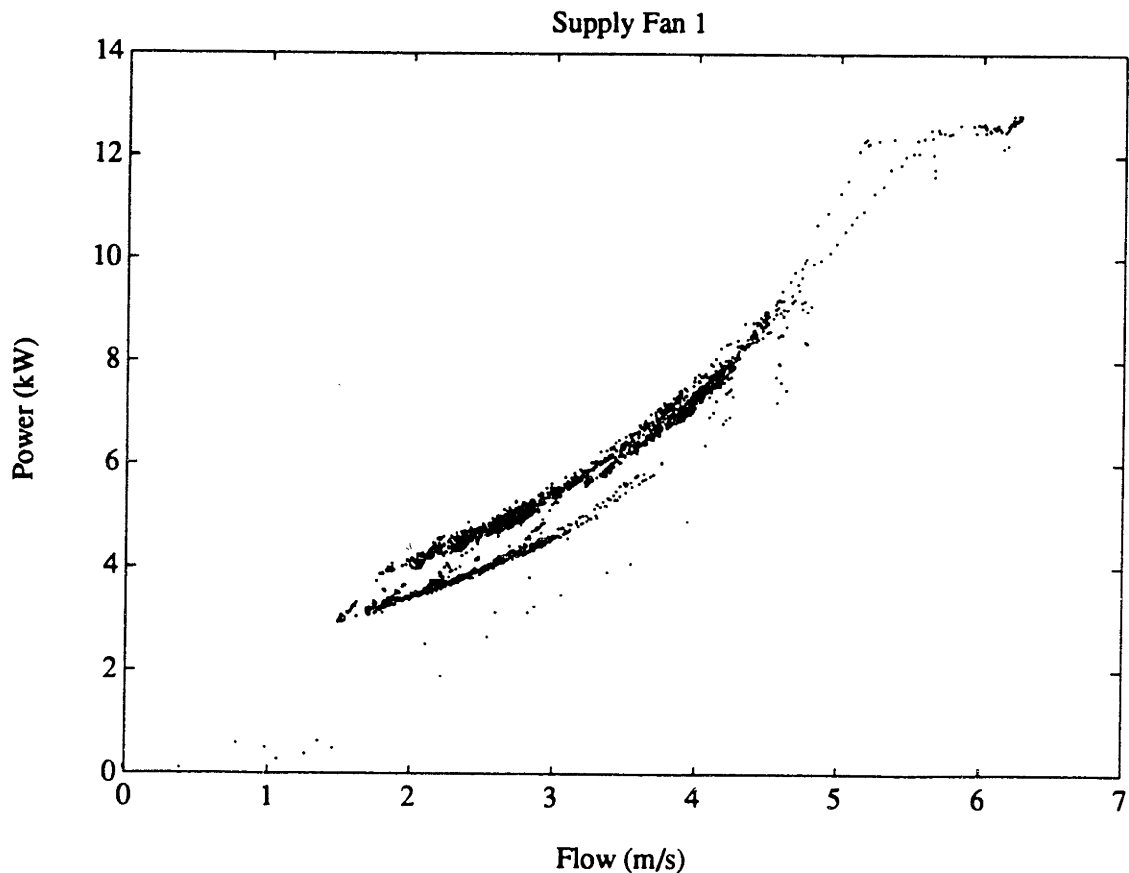e data scattered below the curves are due to data collected during transients and inactive periods. We will exclude these points as well as the plateau for our model fit analysis.

The explanations for the two curved regions stem for the nature of the VAV system and its control. We believe that the two system curves form a control region hysteresis where an offset from the setpoint is being observed. For the lower curve, this offset results from a static pressure that is slightly lower that the setpoint as the demand for flow is increasing. The increase in flow is due to the zone boxes in the areas being supplied with air. As the louvres of these boxes open for more air flow, the duct static pressure drops, translating to a system curve of lower power. In contrast, as a demand for flow is throttled by the zone boxes, a positive pressure offset causes the fan system to ride on the higher curve. Since most of the steady state regions come after a decrease in flow (see figure 4.7), the steady stated powers fall on the upper curve. One way to eliminate this hysteresis loop would be to add integral control to the controller. The resulting control loop should then eliminate the offset over time.

We suspect that the plateau region also has to do with the fan control. If the control signal at that power level is requesting higher fan speed, but the fan cannot deliver it, the zone boxes will be open. This causes the fan flow to increase and the static pressure to decrease. With the control signal demanding full power, the fan power stays constant, causing a plateau effect.

The existence of the hysteresis curves emphasizes that statistics cannot be applied blindly to determine system characteristics in a load monitor application. Without tuning

the models to fit the data and keeping an eye on the condition number, the model variance, and the model fit, the monitor can fail to characterize the system.

The next task is to fit a quadratic through a set of data over a day for supply fan 1.

4.5.2 Flow model

Figure 4.9 shows the curve fits for both system curves of supply fan 1. Based on the confidence interval discussion, the upper data estimated will have narrow predictive confidence intervals in the dense cluster in the middles of the curve. The accuracy near the sparsely grouped data points will be larger. The bottom curve seems very uniform, however, it varies over a small range. Overall the fits are very reasonable with standard deviation and R^2 statistic for the curves being 0.1 kW / 0.99 for the upper curve and 0.04 kW / 0.98 around the lower curve.

Although the data were modelled separately using data from increasing and decreasing power intervals, it appears that the top data set could again be modelled by two curves. We did not investigate the cause of this "third" curve.

A single curve was also fit through the whole data set with a standard deviation of 0.21 kW and a R^2 value of 0.98. The high R^2 value is probably due to the small effect that the few points associated with the lower flows have on the total fit. However, the standard deviation increased for the single fit, so the confidence interval will be larger.

Table 4.1 summarizes the various statistical measurements for the three model fits. First, the condition number is about the same for all three fits which indicates that each independent variable matrix has similar conditioning for the model fit. Second, the sum of the squared error shows the bottom curve with the lowest value and the single fit with the largest, as one would expect. Third, the maximum variance of the model fit over the range of independent variables is listed. As previously discussed, the maximum is at the ends of the

curves. Fourth, the s^2 and maximum variance of the fit are combined with a 95% t-distribution value to give the maximum confidence value. In the final column this value is calculated as a maximum interval. For the single curve fit the confidence interval is about 25% of the lowest fan power, which corresponds to the worst case scenario. For the upper and lower curves the worst case value reduces to 10% and 5% respectively.
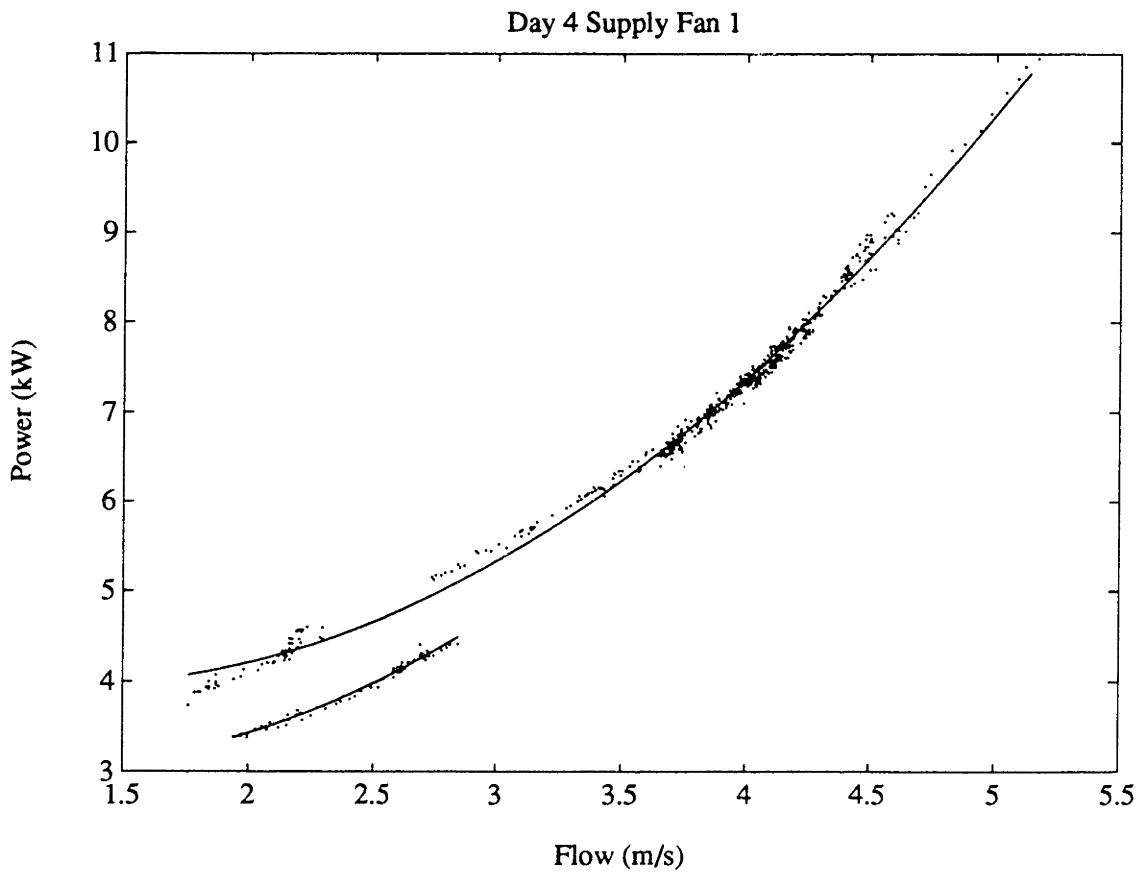

Figure 4.9: Supply fan 1 curve fits for upper and lower curves.

The variance of the model fit is very small due to the success of the fit and distribution of the independent variables. However we cannot neglect this effect when modelling equipment because we have no control over the independent data that the equipment produces.

99

The equipment must supply its own well-conditioned data that may or may not provide variances as small as these.

Table 4.1: Curve fit confidence intervals

	condition number	s^2 (kW^2)	maximum variance for the model fit (kW^2)	maximum 95% confidence value (kW)	maximum confidence interval (ΔkW)
upper curve	4.3	0.010	5.0×10^{-4}	0.20	0.40
bottom curve	3.5	0.002	2.0×10^{-4}	0.08	0.16
single curve	3.7	0.044	2.0×10^{-3}	0.42	0.84

4.5.3 Residual plots

The residuals of the top system curve are presented in figure 4.10 as a check against systematic errors. The plot shows random excursions about the zero level for the first 400 data points. However, for the last 200 points there seems to be some systematic trend in the data. This trend is explained by the curve in figure 4.9, where between 2.5 and 3.5 m/s the curve under-predicts the data. We further note that the data presented were taken at times of the day that the fan load is decreasing or at steady state. Therefore the time is fragmented along the x-axis.

An area we have not considered in this thesis is the error of the flow measurements. The pitot tube arrays that are typically used for flow measurements can become dirty which affects their accuracy. This error in measurement can also vary with the flow in the duct, which can cause the least squares assumptions to break down in our models. In section 4.5.5 we suggest an the control signal as an alternative independent variable for time varying equipment which would not be as subject to these errors.
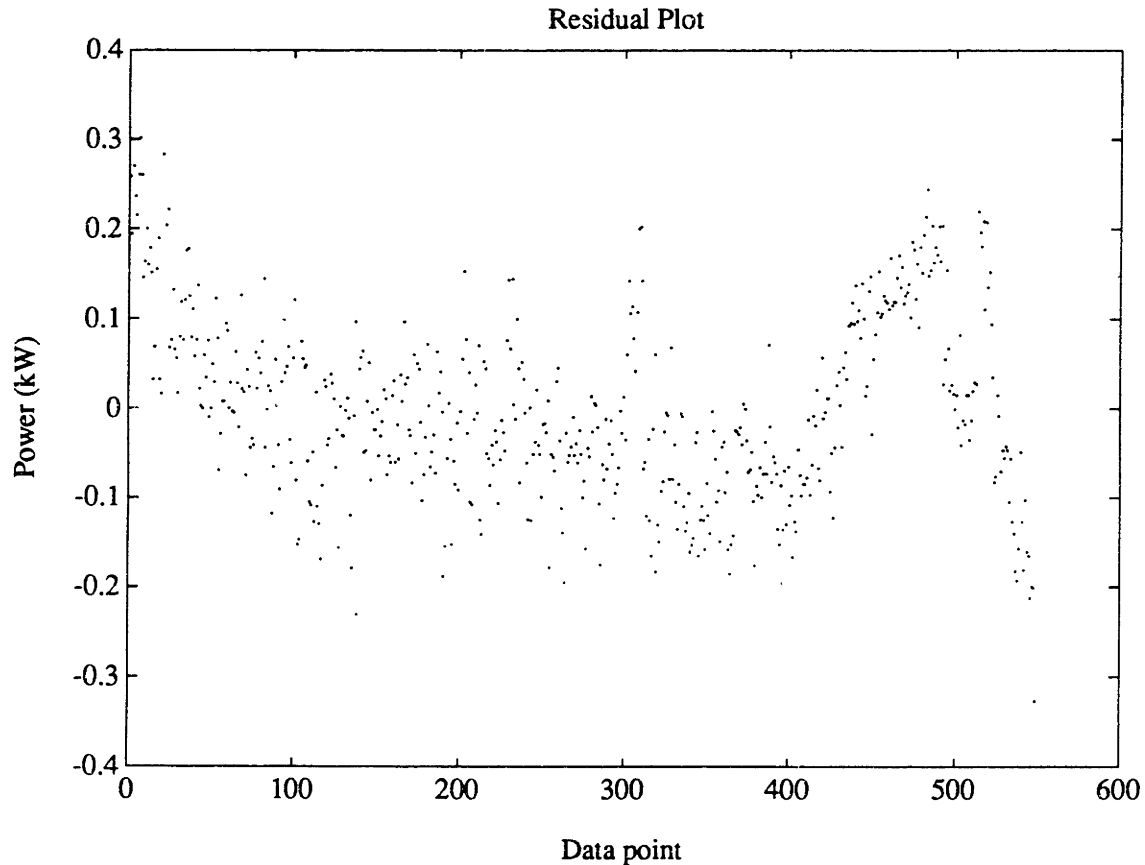
Figure 4.10: Residual plot of the upper curve fit of figure 4.9. The errors appear to be random until data point 400 where there is some systematic fluctuation. This is probably due to a poor model fit in that interval. Note that the x-axis is not continuous in time due to the fragmented nature of the data.

4.5.4 Disaggregation problem

The disaggregation problem stated is to estimate equipment power from an aggregate power signal using independent variables. With supply and return fan pairs 1 and 2, there are four fan powers that can be summed to create an aggregate signal, and two flow measurements, one for each supply fan. We are therefore able to compare the curve fits for each fan pair separately and then compare them with the curve fits from the disaggregation problem. Figure 4.11 shows the results of this comparison for the upper

system curve data of figure 4.9. The top data cluster of figure 4.11 is the fan 1 supply and return aggregate power signal versus flow, and the bottom data cluster is the same for the fan 2 supply and return aggregate signal. The fitted curves through the data are the curves calculated for the data as separate fan pairs, and as an aggregate load, with arrows pointing to the aggregate fit. The error between the two curve pairs causes the aggregate model fit to have a larger confidence interval.

We cannot tease apart the supply and return fans of the two aggregate curves because of the control strategies implemented by the fan pairs. The supply and return fans are programmed to track each other so the relation $F_2 = F_1 - J$ will always hold. One drawback to our disaggregation solution is that although the start-up transients do not have the $F_2 = F_1 - J$ relationship as desired, they also do not fall on the fitted curves. Therefore they can not provide the necessary information to pull apart the supply and return fans as suggested in section 4.4.1. In this case the supply and return fans pairs must be treated as one piece of time varying equipment, which will only allow fault detection to be applied to the fans as a pair.

A list similar to table 4.1 is compiled in table 4.2 for each fan pair individually and for the fans as an aggregate load. Table 4.2 describes some interesting aspects of the disaggregation method. First, the condition number increased from the individual fan models to the aggregate fan model. This illustrates the tendency of condition number to increase as the complexity of the model, in this case a surface, increases. This increase is due to the flow matrix as discussed in section 4.3.3.

A second point is that the worst case error consisting of the maximum confidence interval divided by the lowest power value did not increase more that 5% between the fan pair 2 curve fit and the aggregate fit. This slight increase is due to the distribution of the flow data.

102
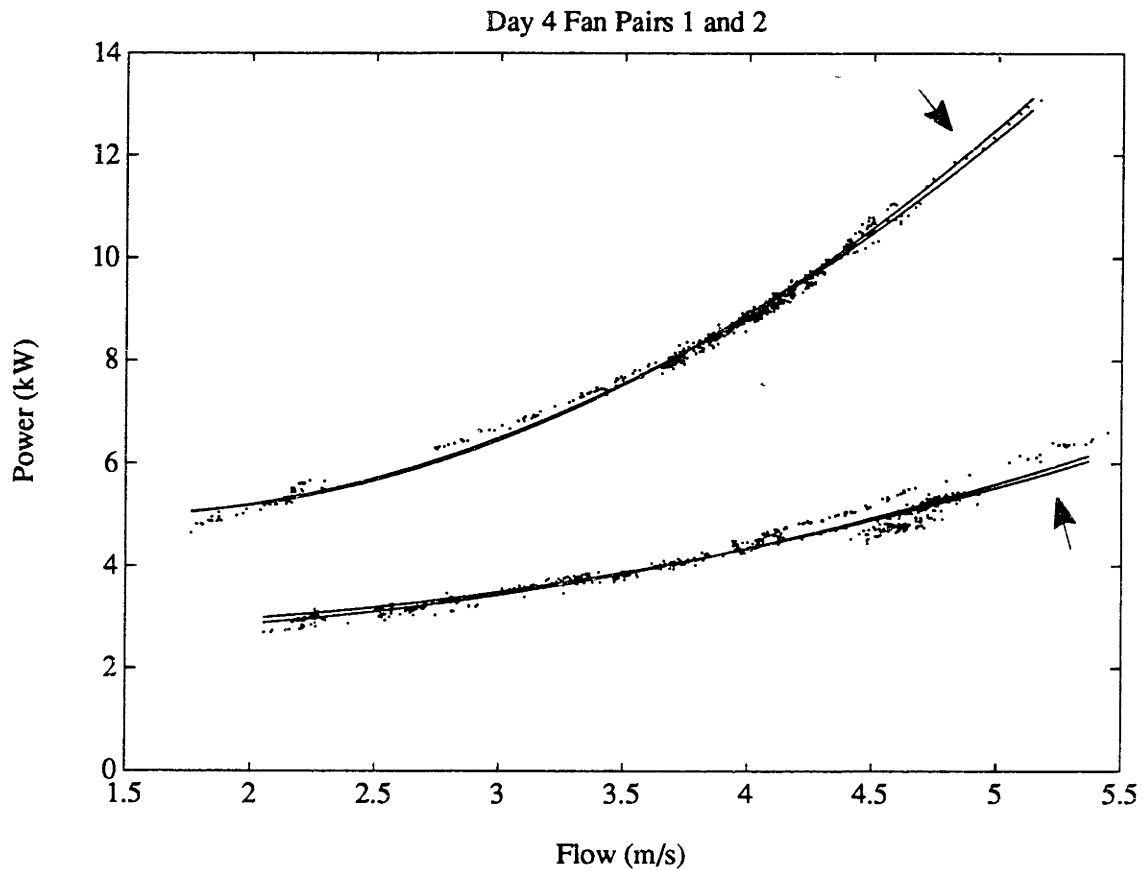
Figure 4.11: The disaggregation technique applied to fan pairs 1 and 2 for periods of increasing and steady state power. The top data cluster is the total power from fan pair 1, while the bottom data cluster is the total power data from fan pair 2. The two curve fits for each data cluster are fits for the cluster alone and fits for the data as an aggregate signal. The aggregate signal model lines are indicated by the arrows.

Table 4.2: Curve fit confidence intervals for disaggregation problem

	condition number	s^2 (kW^2)	maximum variance for the model fit (kW^2)	maximum 95% confidence value (kW)	maximum confidence interval (ΔkW)
fan pair 1	4.3	0.015	7.4×10^{-4}	0.25	0.50
fan pair 2	3.3	0.029	6.5×10^{-4}	0.34	0.68
fan pair 1+2	5.8	0.042	4.0×10^{-3}	0.42	0.84

The maximum confidence value column in table 4.2 represents the maximum size in kW that a deviation from the model fit can assume before a fault is detected. In reference to the method 1 presented in section 4.1.1, this defines the smallest fault in the equipment power signal that our model fit can detect.

Figure 4.12 illustrates the complexity of the model fit variance as it varies over the flows of the two fan sets. The axes for the plot are the flow for the fan pair 1 on the near axis and the flow of fan pair 2 on the left axis. The z-axis is the variance of the model fit. As illustrated in table 4.2, the maximum variance is very small compared to the s^2 value. For the confidence interval calculation it is, for practical purposes, negligible.

We emphasize that even though the model fit variance is not a factor in the data sets presented, it is still an important effect when characterizing equipment. We note three cases where it can have a significant contribution to the confidence interval.

1) Sparse data set

If a monitor has detected a change in the system characteristics of a piece of equipment, it may then try to characterize the new system. Initially, the monitor will have few data points in which to calculate the fit. The model fit variance will then have a significant bearing on the predictive confidence intervals calculated from the model. Consequently, the confidence region will be large and the likelihood of a future data point falling outside the interval is low. However, as each new data point is added to the data (provided it has fallen in the confidence region), the confidence region will shrink and the model fit variance will again become insignificant.
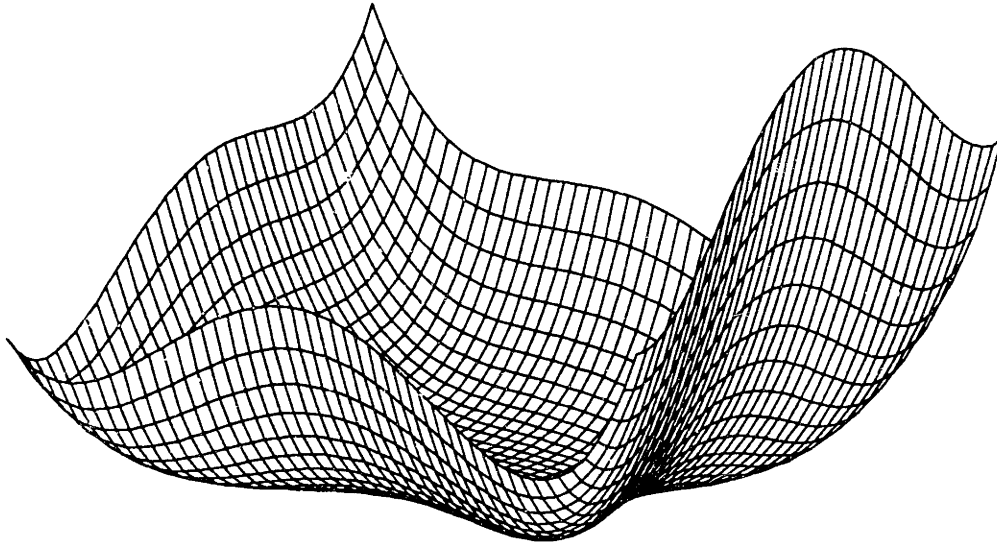
Figure 4.12: The three-dimensional surface of the model fit variance of the disaggregation problem. The near axis is the flow of fan pair 1 and the left axis is the flow of fan pair 2.

2) Poor data spread

If the collected equipment data do not vary over the equipment operating range then the model fit variance can be significant. An example of this would be a straight line fit through a round cluster of data. A monitor could encounter this situation if a piece of equipment does not undergo any building stimulus.

3) Unmodelled areas

When equipment operates in a new range where there is no model fit, a monitor can only base its observations on previous data. At the edge of the operating range, the model fit variance turns sharply up as illustrated in figure

4.12. If the out-of-range data is to be tested, then the model fit variance must be considered.

4.5.5   Control signal

As an alternate independent variable, we measured the control signal of supply fan 1. The control signal for this system varies the speed of the fan based on the offset of a static pressure measurement from a setpoint. The Allen-Bradley fans use a pneumatic control signal. We installed a Setra 208E pressure transducer on this line to produce an output of 0-5 volts.

The power has a nearly linear dependency on control signal, however, we used a quadratic fit to stay consistent with our models. The data for one day are shown in figure 4.13, which also indicates two model fits. The longer fit, as with previous fits, represents the periods of decreasing and steady power, while the shorter fit represents the periods of increasing power.

A comparison of control signal to the flow as an independent variable to the power model can be seen in table 4.3, while flow versus power data are shown in figure 4.14. Note that the plateau of the flow data model was omitted from the fit, while the control signal data did not require this procedure. Three fits are compared: the decreasing power (long), increasing power (short), and the total data set (single).

Table 4.3 clearly shows, for the case of supply fan 1, that the control signal produces a better fit overall. The single curve for the control signal is markedly better than the single curve for the flow data because of the smaller hysteresis loop in the control signal data. We suspect that this may be the case in general.

13
12
11
10
9
Power (kW)
8
7
6
5
4
3
2.8   3   3.2   3.4   3.6   3.8   4   4.2   4.4
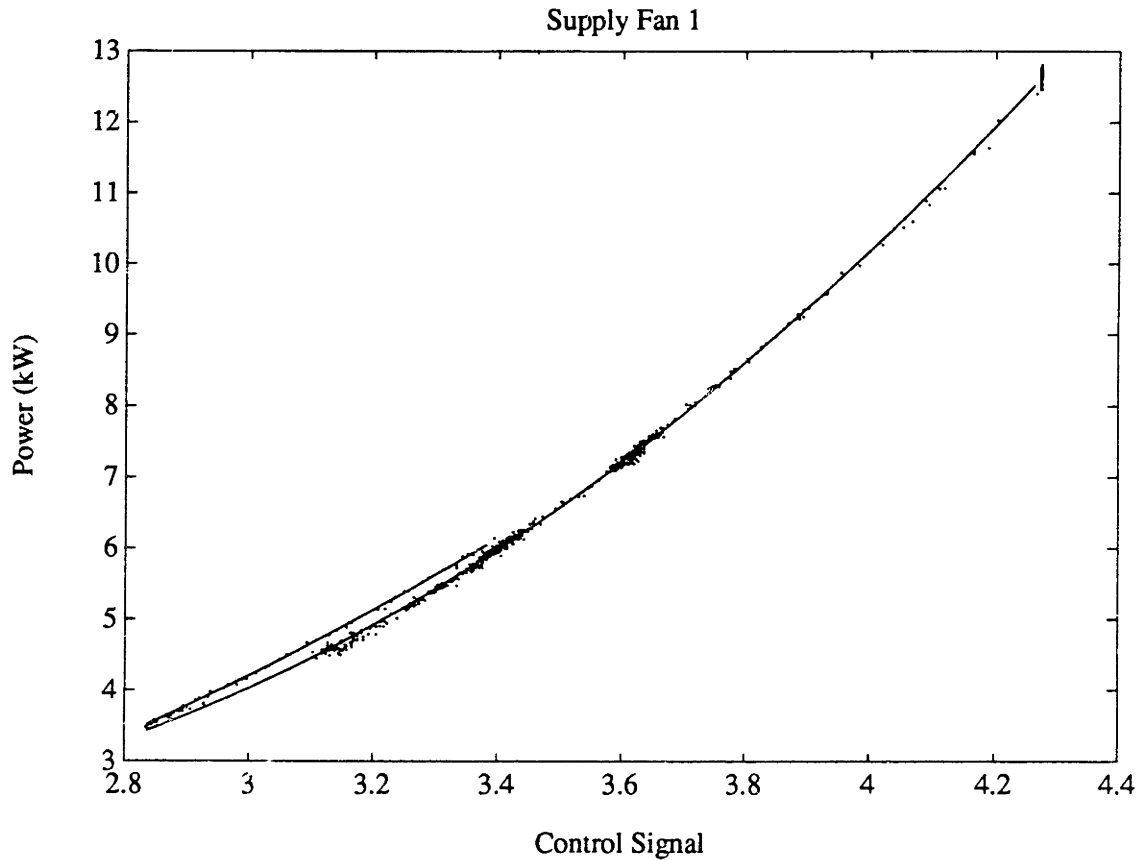
Control Signal

Figure 4.13: Supply fan 1 curve fits using control signal as an independent variable. The long curve is a fit to decreasing and steady power periods. The short curve is a fit to increasing power periods.

Table 4.3: Comparison between control signal and flow as an independent variable

	condition number	s^2 (kW^2)	maximum variance for the model fit (kW^2)	maximum 95% confidence value (kW)	maximum confidence interval (ΔkW)
Control					
long curve	3.1	33 x10^{-4}	1.2x10^{-4}	0.12	0.24
short curve	3.2	6.6x10^{-4}	1.0x10^{-4}	0.06	0.12
single curve	2.9	49 x10^{-4}	1.0x10^{-4}	0.14	0.28
Flow					
long curve	2.4	3.3 x10^{-2}	17 x10^{-4}	0.37	0.74
short curve	3.0	0.12 x10^{-2}	1.5 x10^{-4}	0.07	0.14
single curve	2.5	7.3 x10^{-2}	39 x10^{-4}	0.55	1.1
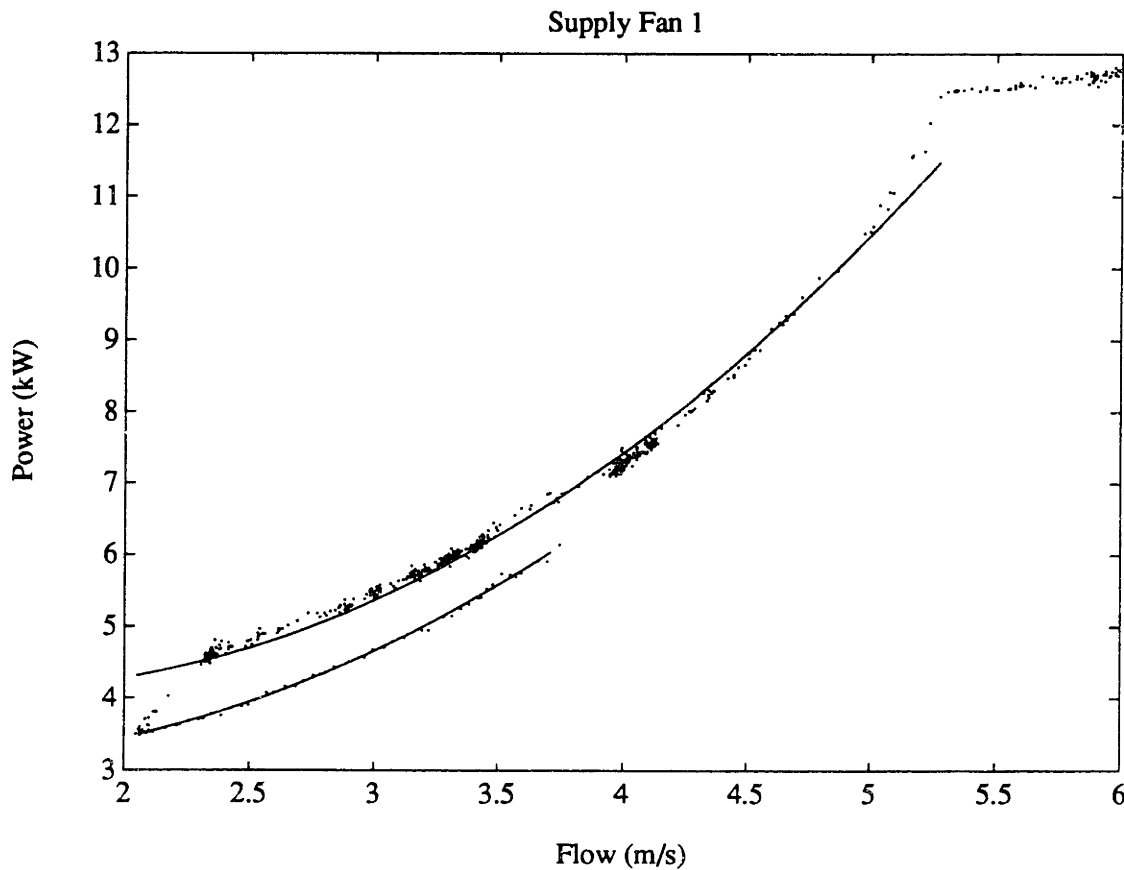
Figure 4.14: Supply fan 1 curve fits using flow as an independent variable. Data were taken simultaneously with the control signal data. The long curve is a fit to decreasing and steady power periods. The short curve is a fit to increasing power periods.

This comparison suggests that the control signal might be a better measurement than a flow measurement for a level 2 monitor to model equipment. In addition to the improved fit, the control signal has other advantages. These advantages include ease of measurement, universality in equipment , and stability of errors.

The first advantage stems for the fact that a control signal can be measured directly from the equipment. This close proximity eliminates the need to attach sensors in out-of-the-way places. For example, a flow measurement sometimes has to be installed in awkwardly placed duct-work.

The second advantage is the universality of control signals. Almost all time varying equipment have them. This would reduce the types of sensors that would have to be used to obtain the necessary independent variables for the building's HVAC system.

The third advantage comes from the flow measurement errors previously discussed. A measurement of a control signal can be very precise, with no variations of errors that can depend on the magnitude of the signal. Furthermore, the control signal would not require such maintenance as cleaning pitot tubes for flow measurements.

These advantages are based on our VSD fan experiments. To generalize them requires that more work be performed to model other equipment using their control signals.

In summary, this chapter has addressed the disaggregation of power signals using a level 2 intrusive viewpoint. We have presented and tested a method for determining the confidence intervals for quadratic model fits. In addition, the technique of singular value decomposition has been presented as a method of numerically extracting equipment characteristics from power data using ill-conditioned data. For these ill-conditioned data sets, a method of further breaking up a collinear signal using start-up values of equipment was presented.

The tests performed have indicated that software disaggregation can be very accurate provided the collected data are well-conditioned. Furthermore, equipment control signals are suggested as an accurate and economical independent variable.

CHAPTER 5 - TRANSIENTS

5.1 Introduction

Equipment transients are the electrical power fluctuations a device experiences as it is turned on or off. The information contained in these transients can be used by a level 1 to a level 4 monitor. For a level 1 monitor, the transients indicate that a device has turned on or off, while a level 4 monitor would evaluate the dynamics of the event more precisely and possibly detect changes in the equipment's normal transient behavior. Depending on the accuracy by which the transients can be measured, transients contain information that can be used by a load monitor in both equipment detection and health.

5.1.1 Information to be gained

Electrically, an equipment transient is a dynamic event that reflects the physical nature of the device. This nature can be viewed as unique to the device and a generalization of its function. For example, the start-up of an induction motor that is driving a fan will have a general shape based on the electrical and physical design of the motor. On the other hand, the size of the motor and the load that it is driving will perturb both the magnitude of the electrical power use and length of time of the event. These perturbations can make the transient different from other equipment of similar motor design but different function. There are, however, limiting factors to transient uniqueness, particularly measurement considerations such as noise, accuracy, and reproducibility. How uniquely we can separate transients of equipment of similar design and size has not been studied as of the completion of this thesis.

111

A generalization of the transients method can be applied by a load monitor to identify the different types of equipment in a building using the building's power signal. By creating a library containing generalized transients of different equipment, a monitor can look to match events with equipment types. For inductive motors, the work-horse of building HVAC equipment, this would mean distinguishing between fans, pumps, and compressors. By distinguishing equipment, we mean that a transient can be identified as being a motor, then a pump, then a pump of a certain size. The size would be reflected in the magnitude of the transient. Therefore, with a defined set of criteria, a monitor should be able to identify reasonably different types of equipment as they turn on in a building.

The identification of transients can provide two applications for a load monitor. One use would be to individually record all the equipment transients in a building and then compare the recorded transients to transients in the aggregate signal. As a monitor matches a known transient with a transient in the signal, the specified equipment can be recorded as operating. This would be a useful way of detecting when, and how often equipment operates.

Another use of transient uniqueness is as a metric for equipment health. As trends in the equipment transients are recorded, the change signifies a change in the physical state of the devices . As with the Entek spectrum monitor[1], this can be used to flag equipment that may need maintenance. Here again, noise and measurement capability will affect the ability of a monitor to detect the changes in attributes, which has not bee studied yet.

In summary, there is a relationship between the electrical and physical characteristics in a device. With proper understanding of these equipment characteristics a monitor could exploit the similarities and differences of transients of building equipment.

---

[1] See chapter 2

5.2 Detection methods

5.2.1 Residential load monitor

A non-intrusive load monitor was developed at MIT in 1986 for residential houses, which used transient detection to monitor equipment use. Several houses were tested, in combination with a survey of their electrical equipment. After the equipment turn-on levels were quantified at a sample rate of one Hertz, the meter tracked appliances by comparing level changes using different power measurements. These changes were then mapped into real and reactive plots for each of the two 120 volt line voltages from the utility. Such plots represent a four-dimensional space where the level changes of each appliance in the house are clustered. By matching the on level change with the off level change based on magnitude and sign of the change, the monitor tracked the appliance use and proved successful in determining their use patterns. [Hart, 1985]

An important factor in making the analysis successful is that residential appliances tend to run at constant power, and therefore on and off cycles create step changes in the time domain. One of the main hurdles in distinguishing between different appliances was found to be the presence of noise in the aggregate signal.

It should be mentioned that when noise inhibited matching the on and off changes, a method of multi-state analysis was tested on the data. The method is aimed at mending the patterns of on and off changes so that if an on is missed the off won't be improperly matched. Multistate analysis was also used to detect multistate machines such as dishwashers and dryers. Much work has been put into this type of analysis. [Hart, 1987]

113

5.2.2 Electrical equipment

There are four main types of equipment in commercial buildings: fans, pumps, chillers, and lights. Furthermore, because of the rising personal computer use in commercial buildings and their use of a power supply typical to office equipment, we have added them to the list of common loads. The following is a list of the common equipment types and their characteristics in reference to start-up transients.

fan - induction motor, long start-up with high inertia

pump - induction motor, short start-up with high torque

chiller - compressor, multistate with a complex and long start-up

lights - fluorescent and HID - reactive, incandescent - resistive

personal computers - fast start-up, nonsinusoidal current wave form

Some equipment have start-up sequences that reduce loading on the motor until it has approached a steady state speed. For this reason pumps may have very similar transients as compared to fans. However, we still expect pumps and large fans of similar size to have different start-up lengths.

5.2.3 Equipment signature

A research team in the Laboratory of Electromagnetic and Electronic Systems at MIT working in conjunction with us on the monitor project have been leading the work on the transient detection problem. The group has many sophisticated pattern matching methods such as neural networks and auto-regressive speech algorithms that they are considering. These methods are beyond the scope of this thesis. Instead, we will

114

introduce the method of cluster analysis and some of the characteristics of transient shapes that illustrate the feasibility of detecting them.

Cluster analysis

After a transient is recorded in both current and voltage, its features can be quantified and used as indices to determine its "signature". These features are determined by the researcher and can be as simple as a step change in power or complex as a vector containing every recorded power point in the transient. The advantage of the single point is the speed of processing and the simplicity in storage while the disadvantage is the reduction of unique information in the measurement. In contrast, a power vector contains the complete information about the transient. However, it is also computationally slow, and requires a large storage space. Furthermore, its shape can vary between similar equipment which makes pattern matching difficult. Thus, one of the priorities of the transient detection problem is to choose a set of criteria that describe transient signatures. As a result of these considerations, we suggest that the cluster analysis methods that were applied to the residential monitor can be extended to commercial buildings in order to categorize the transient as falling into a predefined equipment group, for example, a light or a fan.

It is important to note that the criteria selected for cluster analysis can be based on patterns that may not be readily associated with a physical nature of the device. It may be necessary for the equipment health analysis to have a separate set of metrics. That is to say a metric for cluster analysis that differentiates a device from another may not add any equipment health information and vice versa.

Cluster analysis is the last step in the progression of category sorting problems. When classifying objects or items, the categories of sorting are known, while for other cases of discriminate analysis only part of the category structure is known. The objective, then, for these cases is to group items into their different classifications and to classify new

115

observations. In cluster analysis little is known about the structure of the categories. What is available is a set of observations with unknown identity. The cluster analysis methods seeks to discover a category structure that fits these observations into classes which are defined by the researcher.

For our transients analysis, we define a cluster as a group of equipment transient observations that can be classified as a type of equipment.

Elements of cluster analysis

1) Choice of data units

This constitutes defining what observations are to be performed. In the case of statistical observations, a data unit can be a random sample of a larger group. This subject carries with it all the connotations associated with statistical sampling such as bias and error.

2) Choice of variables

The selection of attributes that most clearly define separate clusters. Poor selection of the variables leads to blending of clusters and the inability to classify observations. Many scientists emphasize parsimony of variables to minimize measurement.

3) Cluster criteria

The cluster criteria are used to define the boundaries of a cluster in which an observation is attached to. In other words, the cluster criteria define the cluster space within the categories that a classification is made. Many different criteria are possible.

4) Number of clusters

A number of clusters must be chosen so that the variables and criterion are designed to reflect distinct groups. Some algorithms have a cluster for each

116

different element, while others stop at lower levels of specificity, thus grouping many elements into one cluster.

5) Algorithms

Two methods of cluster analysis are hierarchical and nonhierarchical. In hierarchical methods variables are linked together to create a tree of classification. For each observation, variables are tested against each branch of the tree, thus following a path to a cluster. For example, this is equivalent to testing a playing card for color, suit, and then value. Since each decision is permanent and effects the next level of decision, the hierarchical method quickly reduces the number of possibilities it must check for, while also having a side effect of not being able to mend early mistakes.

Nonhierarchical methods define a multidimensional space using the cluster variables and then partitioning off the clusters. Various algorithms are suggested for defining partitions between the clusters with the cluster criterion. However, the basic idea is to create partitions for the clusters and then alter the membership of the clusters so as to improve the partitions. Altering the membership is accomplished by varying the items 1-4 above.

[Anderberg, 1973]

5.3   Slow sample detection

5.3.1   Level 4 monitor

With a slow sample speed on the order of one Hertz, a monitor attached to one electrical device will acquire information about the device's steady power and the level

changes of its transients. The level change can be recorded as a magnitude parameter of the transient. Therefore a single device can be monitored to record the operating history and the level change can be used as a rudimentary check on the health of the device.

5.3.2 Level 1 monitor

As discussed with the residential monitor, a multi-equipment load monitor running at a slow sample speed can use the magnitude of the level changes in real and reactive power transients to identify different equipment. With time varying loads in commercial buildings, however, the magnitude of the start-up transient does not necessarily match the magnitude of the shut-down transient. In fact, due to the similar power ranges of time varying equipment[2], smaller equipment can be accidently matched with larger equipment in this scenario.

5.4 Fast sample detection

5.4.1 Level 4 monitor

We believe that the sample rate for analyzing equipment transients should be increased to the rate of the Nyquist frequency of their major features to allow for equipment specific information to be extracted from the power signal. The fast sample rate monitoring offers a finer perspective on the character of equipment transients, which can be then used to expand the dimensions in the cluster analysis for transient separation on an aggregate

---

[2] As discussed in chapter 3

power signal. We therefore constructed a test apparatus to analyze individual transients. The goals of the tests were to:

• Verify the uniqueness of transients.

• Determine proper sampling rates for typical transients.

• Compare faulty equipment with normal equipment.

• Test different cluster variables for detection.

The bench set-up for the tests consisted of a current and a voltage probe that were measured by a Nicolet 3091 digital storage oscilloscope capable of storing two sets of 4000 points, and also a RS232 port for downloading the data onto a personal computer. The current probe used was a Tektronix P6302 clamp-on current probe with a Tektronix AM503 amplifier. An in-house voltage divider scaled down the 120 volt signal by a factor of 11. It should be noted that the voltage divider was not electrically isolated from the oscilloscope and that this type of apparatus requires care with respect to the polarity of the voltage measurements into the oscilloscope. The oscilloscope ground reference should be considered with caution as well.

With our apparatus we were able to measure 120 volt tenant loads of which lighting is the most common. One aspect of lighting that is quickly gaining popularity in commercial building is the use of efficient electronic fluorescent ballasts. We tested an Advance Mark VII ballast with two 4 ft fluorescent lamps. The current and power waveforms for this start-up are shown Figure 5.1 at a sample rate of 2 kHz. Because the power used by the device is small, the voltage waveform is not affected by the transient and is termed *stiff*.
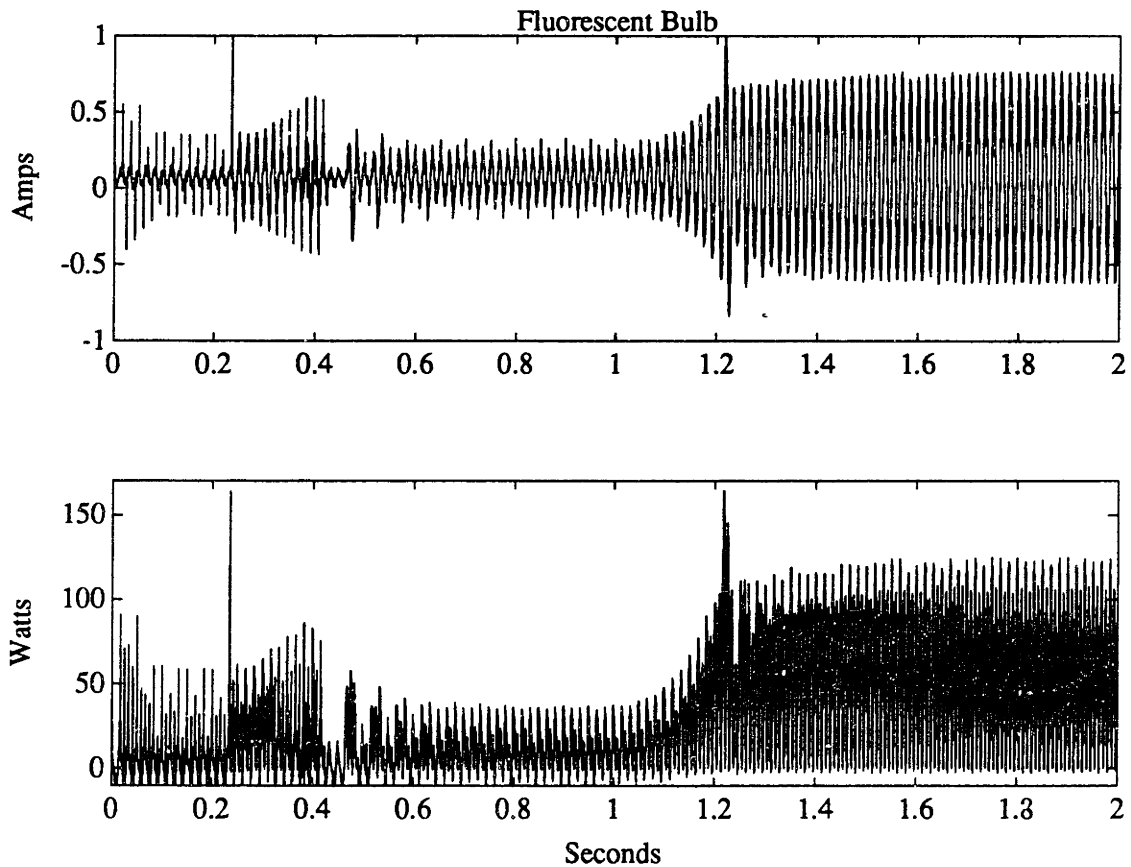
Figure 5.1: Start-up transients for light fixture consisting of an Advance Mark VII ballast and two 4 ft. fluorescent lamps. From top to bottom: a) The current waveform. b) The instantaneous power waveform.

Figure 5.1a and 5.1b clearly show very distinct features. The length of the transient is long, about 1.2 seconds, allowing for many marked features such as the spike at 0.2 seconds and the horn shape that increases in power at 1 second. These characteristics have physical meaning according to the solid state circuit in the the ballast that is designed to first start the light, and then to regulate the current flowing through the fluorescent bulb's arc.

PQ Plots

The information contained in the voltage and current signal can be combined to form real and reactive power, which contain the values of the current, the voltage, and the phase angle θ between them. Because of the phase angle, the instantaneous power calculated from product of the instantaneous voltage and the instantaneous current, can be negative. For this reason, the definition of the real power is calculated by taking the integral of instantaneous power over at least one wavelength, which will always be positive.

$$P_{ave} = \frac{1}{\Delta t} \int_{t-\Delta t}^{t} v(t)\, i(t)\, dt$$

In order to define a reactive power (Q_{ave}), we first introduce the value of apparent power (P_{app}). This value represents the power neglecting the phase angle, where the current and voltage value are calculated separately as:

$$V_{eff} = \sqrt{\frac{1}{\Delta t} \int_{t-\Delta t}^{t} v^2(t)\, dt}$$

$$I_{eff} = \sqrt{\frac{1}{\Delta t} \int_{t-\Delta t}^{t} i^2(t)\, dt}$$

and apparent power follows as:

$$P_{app} = V_{eff} \times I_{eff}$$

121

P_{ave} and Q_{ave} are then defined in terms of θ and P_{app} as:

$$P_{ave} = P_{app} \cos(\theta)$$

and

$$Q_{ave} = P_{app} \sin(\theta)$$

[Hayt, 1978]

These relations can be calculated for steady state conditions. However, during a transient the phase angle lacks definition since the signals are varying so quickly. The literature has many different methods to numerically define the θ and Q [Wyatt, 1990]. For our purposes we calculate Q_{ave} using the following relation that can be verified based on the previous definitions.

$$P^2_{app} = P^2_{ave} + Q^2_{ave}$$

Solving for Q_{ave}, we then calculate the cosine of θ which is termed the *power factor* (PF).

$$PF = \cos(\theta) = \frac{P_{ave}}{P_{app}}$$

Since P_{ave} and P_{app} can be calculated for any given voltage and current in a transient, we can then calculate Q_{ave}. For our PQ plots we have calculated the average values of the real and reactive power by employing the definition of average and apparent power over a moving window of one steady state wave length (60 Hz).

Figure 5.2 shows P_{ave}, Q_{ave}, and power factor for the Advance fluorescent ballast. Note how in figure 5.2c the power factor settles to a near unity value. This indicates that the phase angle is nearly zero and that the apparent power is nearly equal to average power, the case for resistive (incandescent) lights. This is due to the balast's circuit that is designed to decrease the reactive power of the fluorescent lights.

With a definition of real and reactive power, we can now plot the fluorescent lamp transient in a PQ plot as shown in figure 5.3a. First note that the time element has been lost in the plot, while the actual path of the transient starts at the bottom of the plot and ends at steady state at the upper dense area. Most of the transient features occur before one second in the time domain and are located in the lower area of the plot. In particular, many features link the two steady state clusters at the bottom. The P and Q time plots of figure 5.2 help clarify the time history.

In contrast to the healthy transient of figure 5.3a, two faulty bulbs were installed and tested with the same ballast. One of the bulbs was virtually dead, while the other had difficulty striking an arc. Visually, the bulb would begin to light, go out, and then begin to light again, giving a flickering effect. The PQ plot for this ballast-lamp combination is shown in figure 5.3b. The plot shows the failure of the bulb to reach the steady state value reached in figure 5.3a. In fact, the faulty bulb remains in a dynamic state that keeps trying to make the jump to a steady state P. The initial excursion into Q is part of the ballast start-up.
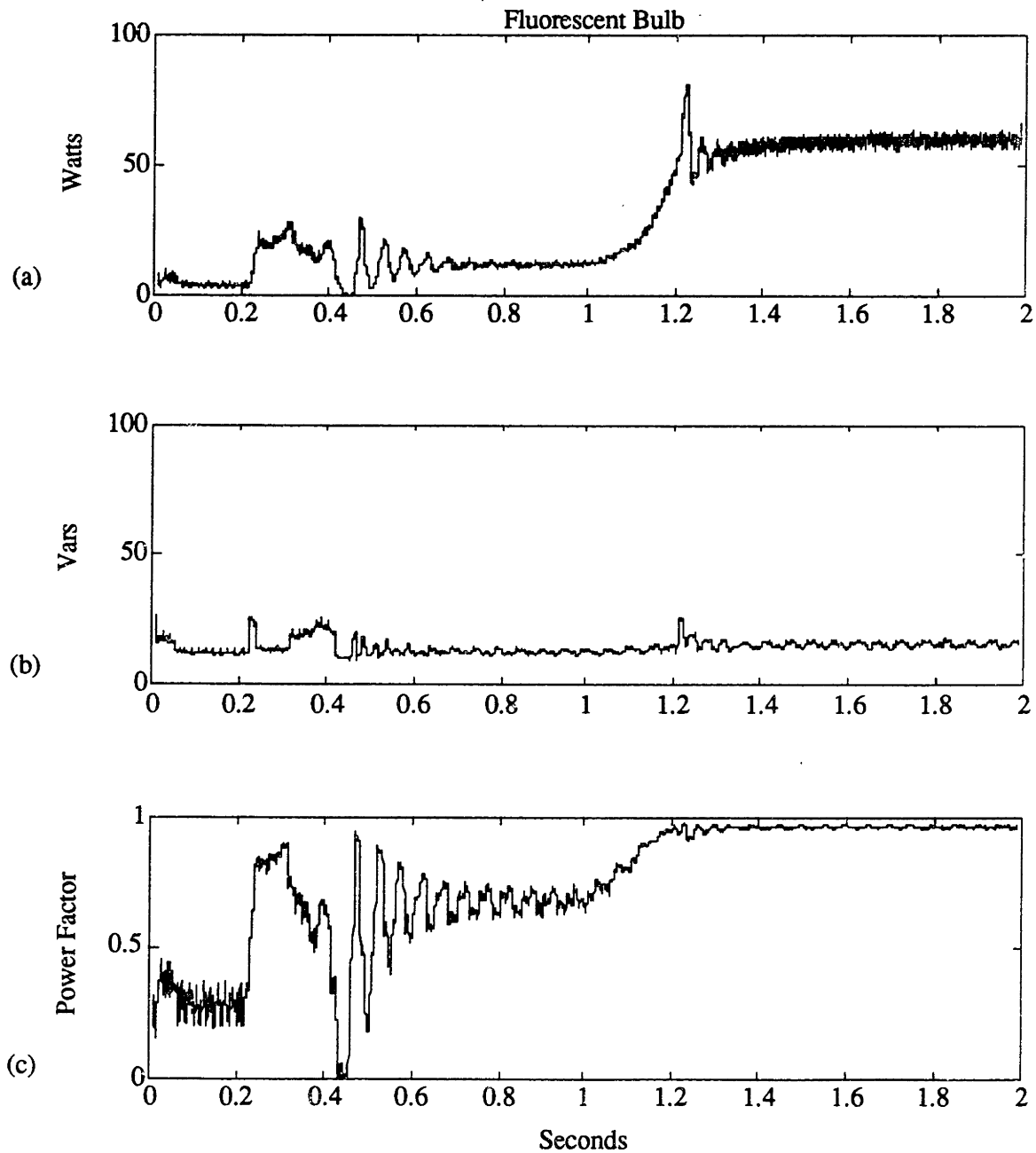
Figure 5.2: Start-up transients for light fixture consisting of an Advance Mark VII ballast and two 4 ft. fluorescent bulbs. The average values of the real and reactive power were calculated by using a moving window of one (steady state) wave length. From top to bottom: a) The average power waveform. b) The average reactive power waveform. c) The power factor.
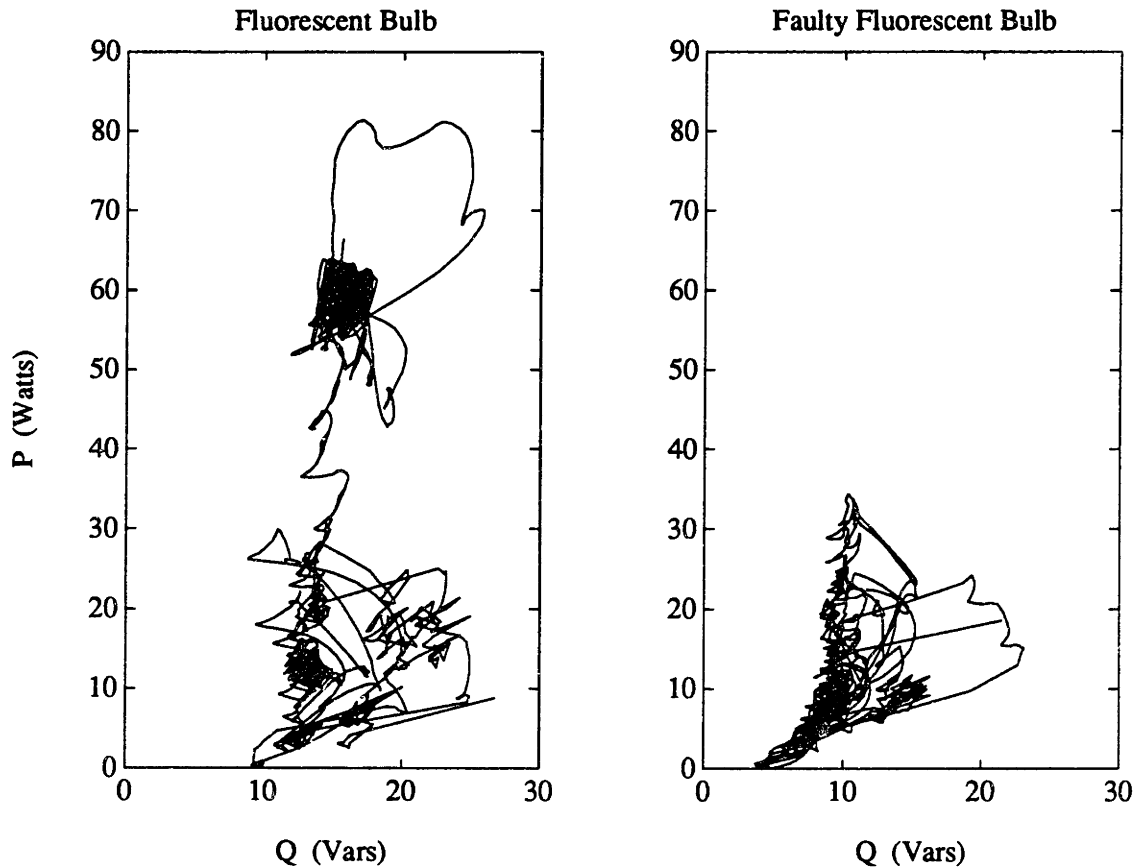
Figure 5.3: Real and reactive (PQ) plots. From left to right: a) The normal start-up transient of the fluorescent light fixture described in figure 5.2. b) The same fixture with two bad bulbs inserted in place of the normal bulbs

Other transients tested were an IBM personal computer XT and a shop vacuum shown in figure 5.4a and 5.4b. Both plots have a time history path that starts at the top end and finishes at the bottom. Although the magnitude of the plots are different, the aspect ratio of the P and Q axes is the same. One could argue that the shape of the two PQ plots is similar enough in the presence of noise that they would be indistinguishable. However, the sample rate for the PC is 20 kHz while the vacuum was sampled at 2 kHz. This is a significant difference that separates the two transient.
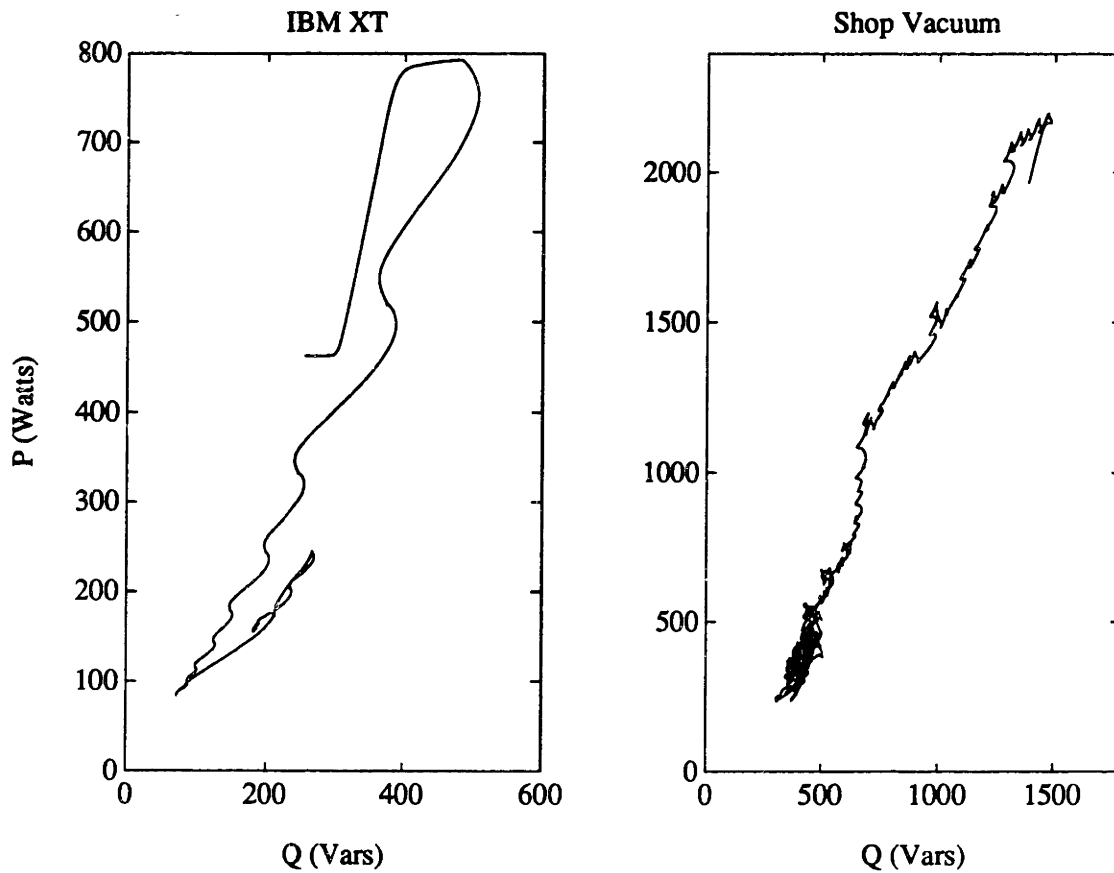
Figure 5.4: PQ plots for two tenant loads. From left to right: a) A PQ plot of an IBM personal computer XT at a sample rate of 20 kHz. b) A PQ plot of a shop vacuum cleaner at 2 kHz. Note that the vacuum has a moving average power window of 5 wavelengths to cut down noise.

Because commercial buildings have their lighting wired in banks of fluorescent lights, we tested a bank of 11 iron core ballast and 22 lights for start-up transients signature, shown in figure 5.5. The bank of lights has a a starting PF of 0.42 and an ending PF of 0.98 in contrast to the starting PF of 0.73 and an ending PF of 0.97 for the lamp in figure 5.3. We believe that this effect is due to the complex current and voltage dynamics on the grid that the lamps are connected to. Of further note, the plot shows small clusters as it ramps up to steady state. We suspect that these are lights that didn't start

immediately with the dense start-up group at the bottom. These late starts could also be due to the local current and voltage fluctuations on the lighting grid.

Bank of Fluorescent Bulbs
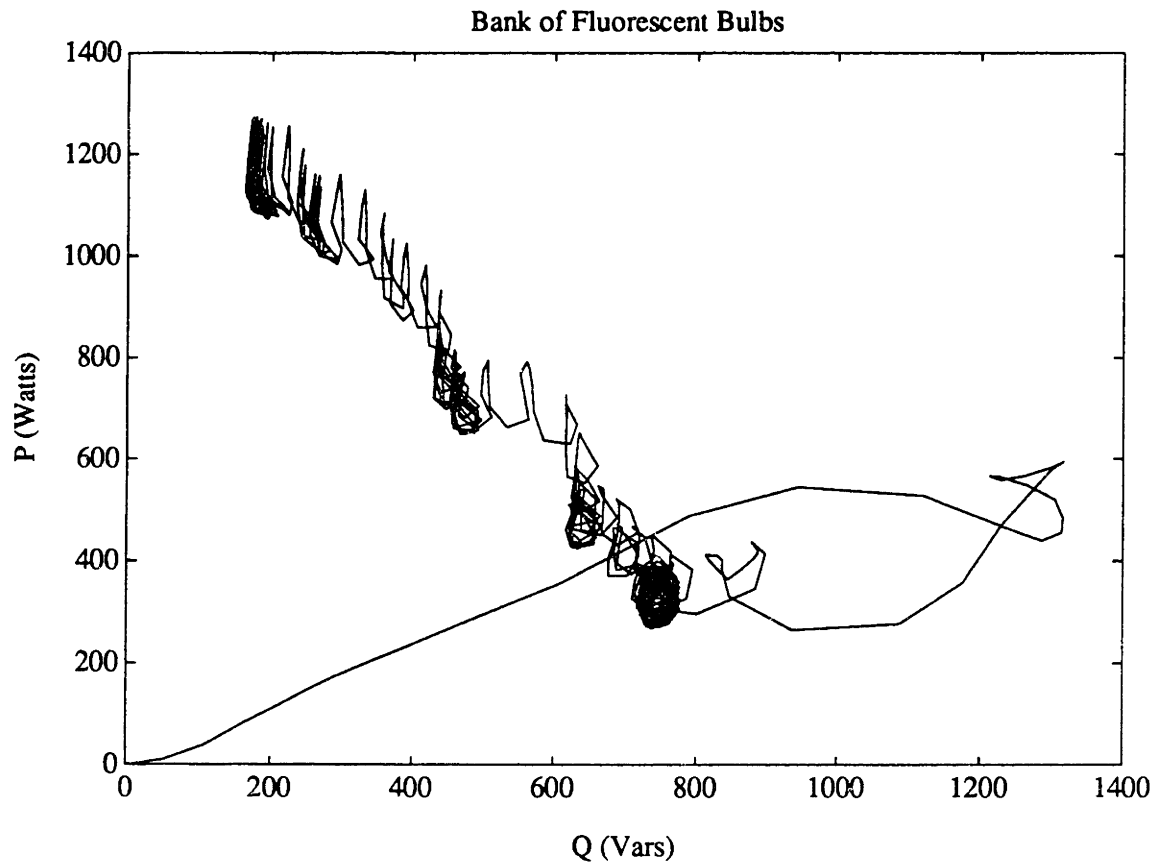
Figure 5.5: A PQ plot for a bank of 11 iron core ballasts with 22 fluorescent lights.

Shut-down transients

Up to this point we have discussed start-up transients with no mention of shut-down transients. We have found in our testing that most equipment perform a shut-down by breaking the circuit between the power and the device. This has the effect of separating any physical relationship between the equipment and the load signal. There is therefore no

information to be recorded except the magnitude of the transients. Figure 5.6 shows a shut-down transient for two Advance Mark VII fluorescent lamps sampled at 2 kHz.

Although the transient itself contains no information, it is possible that equipment can be identified at shut-down by comparing the harmonic content of the signal before and after an event. If a device contributes identifiable steady state harmonics, then there should be a difference on the aggregate signal before and after the shut-down event occurs. Difficulties will arise, however, if equipment do not have unique harmonic signatures.
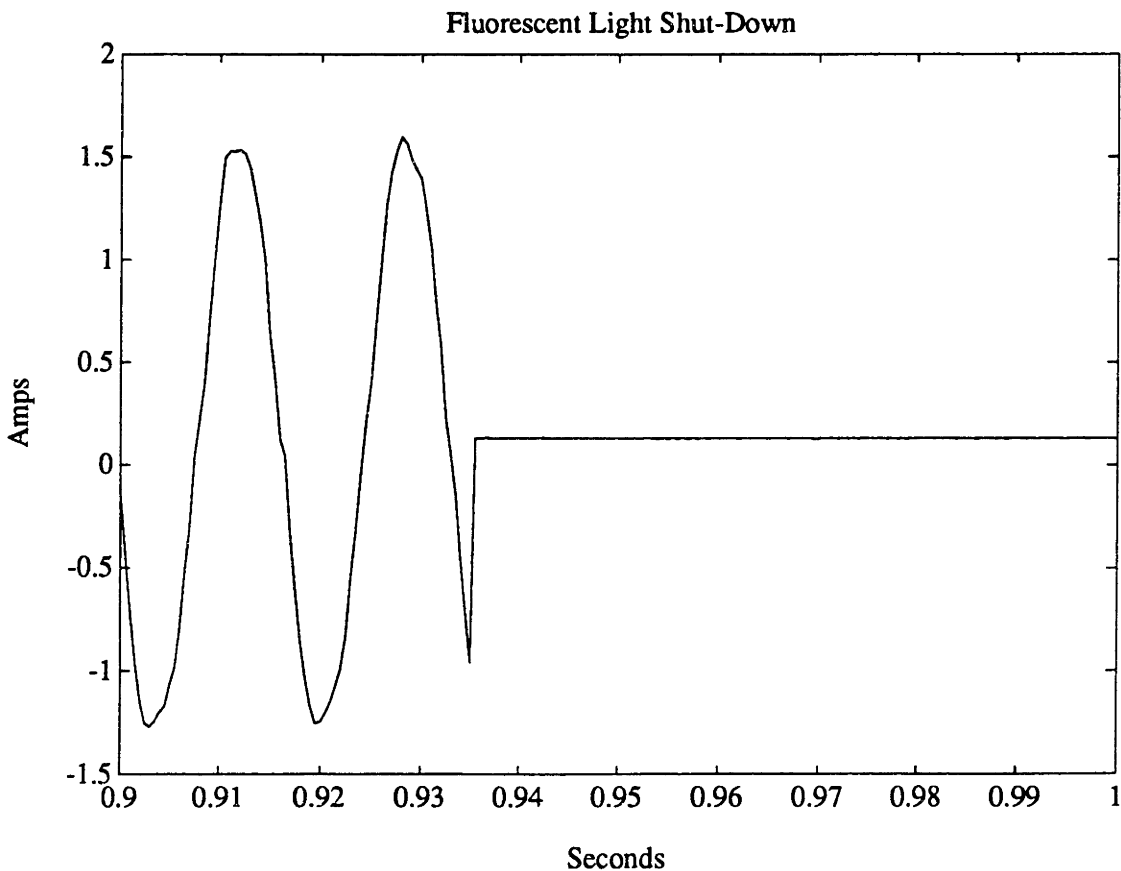
Figure 5.6: A shut-down transient in current for two fluorescent light fixtures at a sample rate of 2 kHz. Note how there is an instantaneous break in the signal.

128

The transient tests have laid a strong foundation for considering aggregate signal start-up transients as a means of detecting equipment. Sample rates between one and ten kilohertz demonstrate a transient uniqueness that has potential to be part of a cluster analysis method for detection. The faulty lamp has also shown how a transient can change when the health to the equipment is poor, while the only information the shut-down transients provide is a step change in power.

In this section we have found that the differences between transients are:

1) The magnitudes

2) The length of the transient in time

3) The shape of the PQ plot.

4) The time history in the PQ plot (how long it spends in one area)

Given these four criteria, a four-dimensional cluster space could be formed, and the transients presented would intuitively fall into considerably different areas. These attributes for cluster analysis are discussed in the following section.

5.4.2 Level 1 monitor

Multi-load transient detection in a commercial building will require a pattern recognition scheme to differentiate the many types of equipment. The single load transient tests illustrated the possibility of using cluster analysis in combination with the unique attributes of the transient. We now use the elements of cluster analysis to formulate a method for distinguishing the transients studied.

For data units, we have already chosen real power, reactive power and time. These values represent the fundamental units of the variables to be used to cluster observations.

We now suggest four the attributes that could make up the variables for the cluster analysis.

1) Time

The length of the start-up transient for each piece of equipment tested. For the fluorescent lamp the time attribute is about 1.4 seconds while the vacuum and the PC are about 1 second and 0.1 seconds respectively. Although there are significant differences in these equipment, we cannot be sure that all equipment will have such defined time attributes. Time is a good first variable that could separate many equipment.

2) Magnitude

The change in magnitude for both P and Q that characterizes the device's steady state conditions can be used as a cluster variable. The examples of P and Q given all show different start-up magnitudes. However, the PC and the vacuum have approximately the same ratio of P and Q at steady state. This could cause confusion if one was trying to use this variable to compare two PCs to a vacuum where the total magnitude change of the PQ plot would be approximately the same.

3) Power spikes

The power spikes during a transient can be viewed as excursions in the PQ plane. The maximum excursions in the P and Q directions for the transient examples in section 5.4.1 are listed in table 5.1.

Again note that the vacuum and the PC have a similar ratio that could cause problems if multiple PCs are started together. However, in this case three PC would be needed to recreate a vacuum-like transient in contrast the magnitude variable's ratio of two.

Table 5.1: Excursions in the PQ plane

	P (Watts)	Q (Vars)
fluorescent lamp	80	25
fluorescent bank	1200	1300
PC	800	500
vacuum	2100	1500

4) PQ path

The PQ path contains the time element of the transient, and essentially encompasses the previous three variables. But, the PQ path is a 4000 point vector, which contains the finer details of the path that we are not interested in at this point. Therefore, as a more compact representation of the PQ path, one could use a coarse grid that is normalized to the transient size. The path of the transient then could be recorded as it passes through the squares which would significantly reduce the vector size.

An even simpler method would be to record the location of ten points spaced equally along the length of the transient in euclidian PQ space. Each point could then be formed into a vector by including the time over which the segment occurred and the direction to the next point. This method combines the time and shape element into a small data set of a defined size.

The number of clusters that define different equipment will depend on how well the present set of observations can be distinguished from each other. In the presence of significant noise, we may be forced to make general clusters, while accurate measurements of each variable will allow each device to be classified.

We have compiled a set of equipment transients that represent various types of fluorescent lamp ballasts found in commercial buildings, which are listed in appendix C.

To illustrate the differences of these transients, table 5.2 compares all the equipment

studied, using the cluster analysis variables that we have suggested. Note how these

attributes distinguish the various equipment from each other.

Table 5.2: Cluster analysis variables of transients from Appendix C

	Time (Sec.)	Magnitude (W)	PQ Spike (W)		Path
One-fixture start-up	2	60	P - 80	Q - 25	Rising in P
Two-fixture start-up	2	120	P - 140	Q - 40	Rising in P
Three-fixture start-up	2	150	P - 160	Q - 60	Rising in P
Rapid start ballast start-up	0.4	80	P - 100	Q - 35	Rising in P
Instant start ballast start-up	0.05	75	P - 350	Q - 500	Decreasing in P and Q
Instant start ballast shut-off	0.025	75	P - 100	Q - 85	Decreasing in P and Q
Personal computer	0.13	150	P - 800	Q - 500	Decreasing in P and Q
Vacuum	1	500	P - 2100	Q -1500	Decreasing in P and Q

The cluster criterion and algorithms are beyond the analysis of this thesis. As

mentioned, there is currently research proceeding at MIT on transient pattern matching.

One of the issues that this research is addressing is transient detection during simultaneous

events. This problem stems from the necessity of a detection scheme to frame a transient in

time. In order to analyze a transient on an aggregate signal, it must have a specified

beginning and an end. One way this can be determined is if a change has occurred in either the real or reactive power signals. The disadvantage of this method is that it requires a steady power signal before and after the event, which cannot be guaranteed in the building signal.

Two cases where the power is not steady are when small tenant loads are creating noise, and during simultaneous events. For noisy tenant loads, either the loads have to be individually detected and removed, or the monitor will only detect larger transients that are not affected by the noise.

For simultaneous events, the problem becomes more complex because the transient framing can have three possible scenarios. First, one long transient can have a shorter transient occur in the middle of it. Second, two transient can begin and end at nearly the same time. Third, two transients of similar time length can overlap creating a longer total transient.

The first scenario could be solved by assuming that one transient dominates the other. The transients can then be clustered as one, and the equipment that dominates will be selected as the nearest possible cluster. This known equipment transient can then be subtracted from the total transient and the residual can then be clustered to identify the other transient. If the second transient can be successfully identified, it can also be subtracted from the total transient and the first transient identification can be verified.

The second scenario can be solved by creating clusters for all the different combinations of two equipment in the building (an argument for using building information in a level 2 monitor thereby reducing the number of possible combinations). These clusters are defined as one transient representing two equipment.

Finally, the third scenario is a hybrid of the first two scenarios. The detection using the solution of scenario two would not work because the transients are staggered and there would be too many combinations of transient to cover all possible events. The solution for the first scenario may work, however, it wouldn't be adequate for cases where the length

of the transient makes the cluster identification impossible. One possibility is that the beginning and end of the transient could be clustered as partial transients and the two transient could be inferred. We note that this is a situation that causes difficulty for transient detection.

The key question that stems from this discussion is how often do simultaneous events happen in commercial buildings, and a more practical question is how often do they occur with equipment that are of interest.

In summary, this chapter presents equipment transients as a tool for disaggregating a building's power signal. We have given an overview of cluster analysis and how it can be used with transient detection. The common attributes that can describe transients show how a cluster analysis algorithm might implement equipment detection. Furthermore, the unique attributes of the transients point to physical characteristic in the equipment. This information can be used for both equipment identification for a level 1 monitor and equipment health for a level 4 monitor.

CHAPTER 6 - MONITORING STRATEGY

6.1   Introduction

Buildings are electrically and mechanically diverse systems, so every possible contingency for load monitoring cannot be accounted for. A level 1 monitor has simplicity in hardware, which is ideal for small buildings and buildings with HVAC equipment located throughout them, however, the complexity of the problem taxes software. On the other hand, a level 4 monitor is simpler than a level 1 monitor in software and is well suited for buildings with a modest amount of centralized equipment, however, its hardware is expensive to wire throughout other buildings. Therefore a strategy for the monitoring levels is needed to properly apply a load monitor to a commercial building.

This chapter brings together the methods and discussions of the previous chapters. We take two different approaches to this task. First, we look at the methods individually, on an informational gathering level. We also restate the methods in order of knowledge needed to disaggregate a building using the tools given in this thesis: building information, hardware disaggregation, and software disaggregation.

The second approach is from an intrusive levels viewpoint. In this discussion, we simulate a medium sized commercial building and consider the various pros and cons of the intrusive levels of disaggregation.

Different levels of monitoring can be right for different buildings. These strategies are presented to take into account the advantages of the individual building to be monitored.

6.2 Building information

A level 1 monitor will at the very most supply a building operator with equipment use schedules and the total load of the building. For this reason, a building audit is a prerequisite for load monitoring. The information supplied by an audit reduces the size of the transient identification problem from generalized equipment to a known set of equipment, while supplying major equipment transient schedules to avoid errors that can be associated with simultaneous events. Also, the element and links database connects equipment behavior to system behavior, which we have suggested can be used to understand equipment characteristics on a system level.

Another aspect of the building audit is a description of the electrical wiring network. This information determines whether a level 3 monitor would be better suited for the building and the desired applications, than a level 1 or 2 monitor.

The building audit as described in chapter 3 puts equipment into a database with the following categories.

- **Specifications** of electrical equipment
- **Load** models
- **Control** method for equipment
- **Links** to other equipment
- **Schedule** of start-up and shut-down events

The database can be updated as the building changes. This allows for a flexible monitoring system as new equipment are added or replaced.

6.3 Sensor/Power Hardware

The main issue concerning intrusive hardware for a load monitor is what hardware can be most efficiently used to supply information. The installation of hardware for the load monitor ideally should be minimized both to reduce labor and instrumentation costs and to maintain the flexibility of the monitor to adapt to changing building characteristics. However, the discussion of chapter 2 suggested that adding sensors can add accuracy to transient detection and provide the opportunity to model time varying equipment. To this end we suggest the following strategy.

6.3.1 Instrument time varying equipment

Installing sensors on time varying equipment allows the monitor to address the system characteristics of daily equipment operation. The sensors for the equipment should be determined by the model for the equipment and the ease of installation. We suggest the feedback control signal of the equipment as a measurement based on our VSD fan experiments. However, models using control signals for other equipment need to be tested before the control signal can be considered a universal independent variable for power modelling. Although all equipment with control signals may not model as well as the fans in section 4.5.5, the control signal can be measured near the equipment, thus reducing the need to install sensors in out-of-the-way places.

For situations where supply and return fans are controlled by the same signal, sensors for both fans will not separate their characteristics because of collinearity. One sensor for the pair is sufficient for these cases. In a complex building HVAC system, pairing the supply and return fan is a worthwhile practice since they are usually thought of as a unit. In fact, for the case of a rooftop chiller/ventilation package where there can be a

large distance between the unit and the monitor, instrumenting the components of the system may not be desired. Again, the operator may choose to let the monitor treat the system as one unit.

6.3.2 Separate load groups using different monitoring points

By separating the tenant loads from the HVAC equipment, improved measurement accuracy may be obtained for subtler monitoring applications. The advantage of this procedure is that HVAC equipment and individual tenant equipment have different load magnitudes both in steady state and in transients. Consequently, a large number of small tenant loads in a building can produce a noisy aggregate signal. An example of this noise can be seen in figure 5.5 where a bank of 22 fluorescent lights reaches steady state with a 100 Watt average (over one 60 Hz wavelength) power band. Although periodic in nature, this steady state signal will add noise to the aggregate power signal unless it is modelled and subtracted out, or smoothed by averaging further. Unfortunately, smoothing may degrade the transient detection.

In some cases, where noise is not great enough to prevent a transient detection, it may be large enough to cause some features about the transient to be lost, and transient health information can suffer. This means that analyzing an HVAC equipment's transient on top of a steady state signal of lights and other tenant loads can corrupt HVAC transients. This argues for separating the equipment into groups that produce noise that is below the magnitude of the equipments' transients.

In steady state, the separation is not as important because the speed of sampling is slower and the noise can be averaged out. If the occurrence of transients in the building is high however, the transients can become noise and reduce the accuracy of the software disaggregation procedures.

Another strategy for hardware separation in the case of large building systems is to separate sets of equipment into systems. The net effect of this division would be to improve the conditioning of the otherwise large modelling problem. As the number of models on one monitoring point are decreased, the errors due to signal-to-noise, collinear systems, and least-squares disaggregation are reduced. In addition the separate systems can be considered to be independent when applying confidence intervals for characterization and fault detection, which can aid in the understanding of the interdependencies of the equipment. We have not determined the optimal number of systems on one monitoring point.

These hardware strategies may be hindered by the fact that the building's wiring configuration may not allow for a precise breakup of equipment. For instance, risers that supply tenant loads of upper floors in a building may also supply a rooftop chiller unit. To separate the tenant loads from the chiller unit would require a monitoring point at the top the the building which implies wiring a signal carrier to the roof or using a wireless communication, which are undesirable because of the intrusiveness of the installation. In this case, the two loads could be monitored as an aggregate signal.

6.4   Software

The software tasks can be grouped in the following way:

- Database(s)
- Output/Interface
- Transient detection
- System character/ fault detection

The databases and output/interface elements of a load monitor have not been investigated and are not discussed in this thesis. They, however, do represent an important area of research to help realize the monitors final application.

6.4.1 Transient detection

Transient detection algorithms and mathematics are not be presented in this thesis. Instead, we restate the following key attributes of transients and provide a strategy for using the various transient information with building information database.

We propose the following strategy for the monitor.

Detection

The attributes used in chapter 5 for clustering and health trend analysis are: time, magnitude, start-up spike, and PQ path. There are certainly more attributes that can be extracted from the intricate equipment transients and these attributes are presently being investigated as a part of another thesis.

Identification

Once a transient has been detected, it must be classified. For a level 1 monitor, this requires a knowledge of the vast number of different equipment transients. A level 2 or 3 monitor has knowledge of equipment schedules (for some equipment) and equipment types. These two pieces of information reduce the complexity of the transient classification. The schedule and the equipment type can be used as cluster variables of their

own. For instance, if a transient is detected at 8 am, then it would be clustered with any equipment with start-up schedules within a predetermined interval around 8 am.

Two cases can arise from this situation. First, if a device is suppose to go on at 8 am but doesn't, then there is a potential fault. Second, if an unidentified transient starts up at 8 am, then a different fault situation occurred. In both cases, there is no match found where the schedule indicated that there should have been.

These situations have interesting implications. Either the device did not start-up when it was schedule to, or it did not start-up in its usual manner. These two possibilities deal with equipment health. Note that in a level 1 situation, this type of analysis would not be possible.

The shut down events must rely heavily on the equipment schedules because they lack the wealth of information contained in the start-up transients. The shut-down events must rely on the change in magnitude and the change in harmonic content of the real and reactive power data for detection. The schedule variable in a cluster analysis adds a much needed dimension for detection. In particular, the time varying equipment must have more knowledge than real and reactive power because of their changing transient magnitudes.

Recording events

Each event is recorded in a database for operator inspection and use by the software disaggregation.

Alarms

If an event cannot be identified, the monitor should note the event for the operator. If there is an unknown event or no event at an equipment scheduled time, then the operator should be alerted for possible equipment or system problems.

Verification

If equipment schedules are available to the monitor it can use that information to verify transients even in the presence of noise. If noise prevents a transient from being detected by cluster analysis or pattern matching, the monitor can smooth the the noise of the signal by averaging. Much transient information will be lost; however, noise prevented the information from being analyzed anyway. With the smoothed signal, and the schedule information, a minimum level change should be detected to verify that the equipment did turn on. In the case of simultaneous events, the minimum level change can be added to verify both equipment.

Equipment use schedules

At the end of the day, data of equipment activity plus all unidentified events should be available to the operator or EMCS.

6.4.2 System characteristics/fault detection

The characterization of a system uses data from time varying equipment. We have proposed a strategy here that uses the data as they are collected for fault diagnosis. We then use data over a period of time to look for trends. The data that have been collected over the day are then used to augment the model data that that is used to characterize the equipment.

Therefore we have the following strategy in order of application:

142

Characterization of data

Acquire data of normal equipment operation over the entire equipment range of operation. In order to do this, one must either drive the equipment to give the proper output range, or else monitor the equipment until the building gives the necessary output. The first method is not sufficient because the equipment characteristic within the building as a system can be much different from the equipment running on its own stimulus. The second method is essentially the same as attaching the monitor and letting it collect characterization data on its own. One previously noted difficulty with this is that it is possible for some equipment never to be driven over their whole range. This is not a problem until the equipment has excursions into infrequent ranges, in which case the accuracy of the prediction will drop as discussed in section 4.5.4. There is no remedy for this except for the length of time a monitor acquires data from a building.

Data acquisition.

For system characteristics, a data rate of one sample per second to one sample per minute is sufficient. The data should be averaged over the data period to smooth out noise in the signals. If there any periodic noise exists, as with the bank of lights example, the sample rate should be at the Nyquist rate of of the noise to avoid aliasing. The sensor or power signals can then be averaged to produce a slower sample rate.

Check data for faults

Based on the predetermined model of power versus independent variable(s), the monitor should compare each new data point with the predicted model as calculated from data characterization. With the knowledge from the database and transient detection, the

143

equipment that are operating should be known at all times. A predicted load value can be calculated for each monitoring point. For aggregate loads:

1) Add up the constant loads' power using transient information. There will be some error associated with this value that will cause the following fault predictions to be less accurate.

2) Subtract constant loads from the total signal, leaving the time varying equipment. The predicted time varying loads can also be subtracted from the total signal, which should produce a constant signal. If the signal is not constant within the noise, then either the time varying models do not accurately fit the data, or there are significant time varying effects in the constant power equipment, which should be noted.

3) Using the disaggregation technique on the time varying signal, compare each new point taken to the confidence interval from the model fit. If it is outside the confidence interval then a warning should be given to the operator. Record the residual value between the predicted and actual power.

4) Keep a record of data points that are outside of the confidence interval. If there are several in a row in time, sound an alarm based on an operator selected number of events.

5) Using method 2 from section 4.1.1, add the residual from step 3 to a running sum. Test to see if the sum has reached an operator defined value indicating that a short term system trend has been detected. If a trend is detected, then alert the operator.

6) At the end of the day the monitor can apply statistics on the day's data set, which may, or may not be well-conditioned. For each monitoring point:

    a. Subtract constant loads from the day's power history using the equipment use record.

b. Test the independent variable matrix for conditioning.

c. Disaggregate the power as well as the data permit using the SVD technique.

d. Using the t-distribution and method 3 from section 4.1.1, compare the model fit for the day with the model used for power prediction. If a significant difference is detected then the operator should be alerted and the new data should be used to create a new predictive model. If there is no significant change in the model, then the data can be added to the previous data used to create the predictive model. However, this should not be done indefinitely because small fluctuations in the equipment will cause the total model to blur over time. Instead, data that are older than a predetermined value should be forgotten. In this way the predictive model will change when the characteristics of the system changes, and if no change is detected the data can be used to form a more accurate and well-conditioned model.

e. Over time the characteristics may change slowly such that day-to-day detection is not possible. To detect these long term effects, daily model data should be stored periodically and tested for significant changes against old records.

7) For any significant changes, the operator should be notified as to what problem is found and what test determined it.

8) Relevant data should be stored in a system database that links equipment.

6.5   Special procedures

6.5.1   Off-hour test

Off-hour tests have been considered as a way of isolating equipment power signals for measurement during times of building inactivity. For example, a piece of equipment could be run through a set of specific exercises during the night when few other equipment are running. The specified equipment's signal in the building would then be stronger with respect to the aggregate signal, which should not be time varying.

An anticipated use of the off-hour tests was to collect time varying data to characterize individual equipment. The aggregate load signal would have less building equipment noise and other varying equipment signals affecting it. Thus the accuracy of the model fits would be improved by a reduction of noise and a improvement in data conditioning. The drawback to the off-hour tests is that it can not recreate the system dependencies of the equipment that they experienced in the daytime. This results in different equipment model fits during the night, which would not be useful for daytime load prediction.

However, the off-hour tests can be used for transient analysis. Considering the Entek motor analyzer application discussed in chapter 2, the improved signal-to-noise provided by the off-hour tests could allow for motor transient health analysis. The monitor could test equipment on a periodic intervals suggested by a predictive maintenance schedule. Testing the motors could be performed by running them through start-up and shut-down sequences which can be controlled by a central control system like an EMCS. The transients recorded could then be analyzed in a similar manner as the Entek method. With this kind of control, a health analysis program can run the equipment through several transient cycle thus acquiring repeatability and statistical samples that a building does not naturally provide.

146

6.6 Building simulation and discussion

We have simulated a 50,000 square foot commercial office building in order to illustrate our disaggregation process. For the simulation, we chose to model the time varying loads with respect to outside temperature, which is created to represent a summer day, with a peak temperature of 90°F at two o'clock in the afternoon.

A description of the building and its equipment are as follows:

Building type:
    50,000 square foot, five story office building
Chiller:
    125 ton centrifugal chiller with a 170 hp compressor
    The chiller turns on at 7:30 am and turns off at 6 pm.
Fans:
    • Two supply and return fan pairs that distribute air to different areas of the building.
        Two 40 hp supply fans
            maximum air volume: 25,000 cfm
            control: variable speed drive
            setpoint: static pressure sensor.
            schedule: on at 7 am and off at 6 pm
        Two 10 hp return fans
            maximum air volume: 20,000 cfm
            control: variable speed drive
            setpoint: supply fan setpoint minus a constant.
            schedule: on at 7 am and off at 6 pm
    • Six 1/2 hp toilet exhaust fans running around the clock.
Pumps:
    • One 30 hp chilled water pump
    • One 30 hp cooling tower pump
        Schedule: turn on ten minutes before chiller turns on, and off ten minutes after it turns off. This is to guarantee that there is a load on the chiller before it starts its compressor. Ten minutes is an

exaggeration of the actual time which is on the order of one to two minutes.

Cooling tower:

  • One 20 hp fan with 3 speeds: off, 1/2 power, full power

    The fan turns on when more evaporative cooling is needed.

Compressor:

  • One 5 hp pneumatic control line compressor

    Compressor cycles on and off every five minutes.

Tenant loads:

  • Lighting - fluorescent lighting at 1.5 Watts/ft^2

  • Office equipment - 2 Watts/ft^2

    Schedule: on at 8:30 am off at 7 pm

  • The tenant loads are given a noise of about 10% of their total value. This addition is based on the assumption that tenant loads are noisy and that they will be turned on and off throughout the day

Table 6.1: A list of the simulated building equipment and their sizes

Equipment	kW
Chiller	125
Supply fans	(2) 30
Return fans	(2) 8
Cooling tower fan	15
Exhaust fans	(6) 0.4
Pumps	(2) 23
Compressor	4
Lighting	75
Office equipment	100

Electrical wiring configuration:

  • The utility transformer supplies 480 volts on a linear bus bar. There are three connections to the bus bar:

Riser 1: tenant loads for floors 4 and 5, and a chiller located on the roof. Note that in buildings we have investigated, the rooftop chillers have had a separate riser. However, the riser could come from the same connection on the bus bar.

Riser 2: tenant loads for floors 1, 2, and 3.

Riser 3: All other HVAC loads, which are located in a mechanical room in the basement.

Figure 6.1 shows an electrical power simulation for the building over a day. The top load profile is the aggregated signal, while the other three profiles from top to bottom are risers 1, 2 and 3. We use this figure to describe the possible building scenarios that affect the different levels of monitoring.
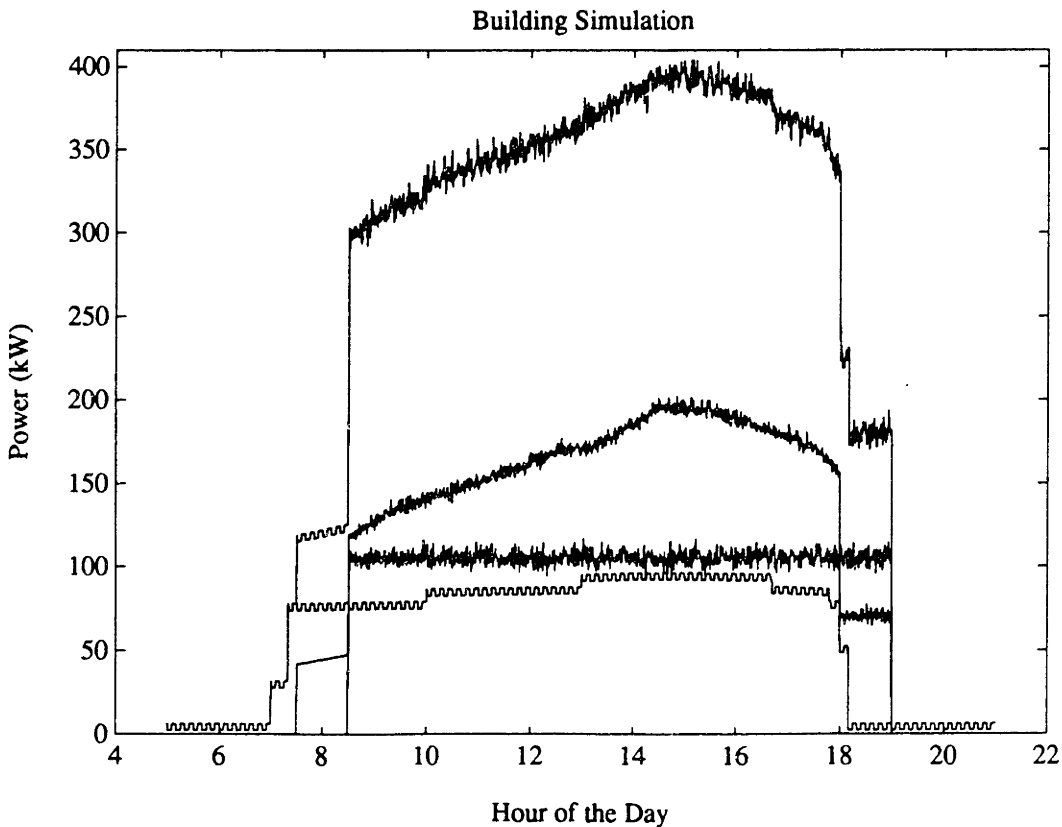

Figure 6.1: Building load profiles for a simulated building. From top to bottom: the aggregate signal, riser 1, riser 2, and riser 3.

6.6.1   Level 1:

A level 1 monitor will have only the aggregate signal in figure 6.1 as input. Inspecting the signal, the three pieces of information that are apparent are the transients, the load profile, and the noise.

The transients that can be seen by eye are the larger transients and the transients that occur at relatively quiescent building periods. The most obvious transients are the chiller and the tenant loads. The tenant loads would, in reality, take more time to start-up, since the office equipment would likely be turned on over a period of an hour or more. The transients that cannot be seen are the smaller transients during the middle of the day. For example, the cooling tower fan in riser 3, which steps through its range between 10 am and 6 pm, appears as a time varying load in the aggregate signal. Detection of these transients would require pattern matching and cluster algorithms that will be hindered by noise.

Because there are relatively few pieces of equipment running in the morning, the start-up transients for pumps and fans show up clearly. This is an advantage for transient detection (and modelling) that can be used by any monitoring level.

The shut-down transients are not as easily identified as the start-up transients. The chiller and the fans shut down at the same time and with a different magnitude than their start-up transient. This make the two transients difficult to identify and match.

The pump turns off at 6:10 pm with the noise of the tenant loads affecting the magnitude slightly, and finally, the tenant loads turn-off at 7 pm. This clearly shows how the tenant noise can affect the transient magnitude.

The total building load profile gives a basic building response to the outside temperature. However, it lacks specifics that would let a building operator identify what loads are affecting the profile.

As figure 6.1 illustrates, noise changes as equipment turns on and off. In the early morning, the only noise is from the compressor which is distinguishable by eye. When the

150

tenant loads turn on, however, the noise increases and the compressor and other transients disappear. Therefore the application of a level 1 monitor on the simulated building shows how the monitor is constrained by noise and time varying loads.

6.6.2 Level 2:

For a level 2 monitor, the equipment data and independent variables are available. In our simulated building, the time varying loads are the two fan pairs and the chiller. The cooling tower fan is considered a constant power device because it varies by a step change.

The time varying loads can then be disaggregated using the software method of chapter 4. We note that the supply and return fan of each pair are collinear, so they can only be modelled as a unit. Thus there are three time varying loads for the level 2 monitor to separate.

Because the fans pairs are the first equipment to turn on in the morning, there is a period which they can be modelled alone. This provides data for a simpler model involving only two time varying loads with reduced noise. In addition, the constant term of the models (if one is included) can be determined before the chiller is turned on, which further disaggregates the loads.

The chiller starts up an hour before the noisy tenant loads turn on, which is a period where the chiller, with the fan pairs, can be modelled with reduced noise. These early morning data for all the time varying loads should be considered more valuable than the later, more complex, and noise filled data of the middle of the day. When assessing the end-of-the-day fit as described in section 6.4.2, the lower noise data can reduce the error variance of the model by a least squares weighting method [Draper, 1981].

The equipment knowledge from a level 2 intrusiveness would point out the transients of all the HVAC equipment except the compressor and the cooling tower fan,

151

which have no schedules. As discussed in section 6.4.1, if a transient does not occur when it should, then an alarm is evoked.

In a very coarse analysis, the element and link database will tie the chiller to the two fan pairs, the cooling tower fan, and the pumps. If any equipment characteristics change or faults occur, then the equipment links database could be scanned for any physical links to other parts of the system. For example, a chilled water pump and a supply fan are connected by a cooling coil which is mapped out by the links. If the fan experiences a fault or a change in character, the the cooling coil and the chilled water pump can be considered as possible causes. Furthermore, if a history of problems develops in one system, the links may point to the fault and an expert system might either suggest that repairs be made, or an increase the preventive maintenance.

6.6.3 Level 3:

Level 3 monitoring optimizes monitoring strategies by separating the loads using hardware. The simulated building has three risers with some mixed loads. Considering the two strategies in section 6.3, one separation method is to monitor the HVAC loads apart from the tenant loads. This could partly be accomplished by monitoring the aggregate signal and riser 2 in the simulated building. With an aggregate monitoring point, the total signal of risers 1 and 3 can be calculated by subtraction. Although all the tenant loads are not separated, the noise contribution from them is reduced by more than half.

Another strategy for a level 3 separation is to split apart the time varying equipment, which can be accomplished by monitoring riser 1 and 3. The chiller and the two fan pairs can then be modelled separately, thus increasing their model fit accuracy. In addition, the magnitude of the chiller power is much bigger than the fans. Separating these two loads

separates the larger chiller model noise from the smaller fan model noise, particularly increasing the accuracy of the fan model fits.

These two schemes also benefit the transient detection problem. Each separation increases the accuracy of the monitor, while reducing the noise and the number of possible transients for that monitoring point. With fewer equipment transients there is less probability that a simultaneous transient will happen without schedule knowledge to disaggregate the event. The accuracy of the transient health analysis will also be increased due to the reduction of noise.

6.6.4 Level 4:

A Level 4 monitoring requires that monitoring points be placed on all HVAC equipment and the tenant loads be monitored by floor. This would provide the monitor with complete power information, however, it would also require 20 monitoring points. The wiring for this installation would have to connect one monitor to monitoring points from each floor, the roof, and the mechanical room. In addition, office equipment and lights would not be separated.

We have presented in this chapter a summary of the methods of disaggregation that this thesis advocates. The three information gathering methods have been addressed: building information, software disaggregation, and hardware disaggregation. Strategies for implementing these tools have produced a framework within which a monitor can work. Furthermore, we presented a discussion of a simulated building that explored the different levels of monitoring. Of these levels, 2 and 3 are the most promising in terms of extracting information from a building in a non-intrusive manner, while 1 and 4 lack efficiency in information and installation, respectively.

CHAPTER 7 - CONCLUSION

The design of a non-intrusive load monitor is a balance between its application and the level of intrusiveness desired of it. In this thesis, we have presented both a vocabulary for considering levels of intrusiveness and methods for achieving load monitor applications. With these tools, we have put forth strategies for extracting building information which consider the building to be monitored, the application for which it will be used, and the level of intrusiveness required.

These aspects of monitoring stem from knowledge of general building equipment and the diversity of the previously studied buildings. In our discussions, we have suggested level 2 and 3 for their efficient use of information and building type versus the level of intrusiveness. We have not stated that these levels are correct for all buildings, however, they contain the most promising strategies.

In addition to the levels of monitoring, we presented a software disaggregation technique that considers the conditioning of the data as well as its noise to produce confidence regions for fault prediction. This statistical approach has demonstrated that an aggregate power signal can be modelled and disaggregated using "intrusive" sensors. From the equipment tests on fans, we have also illustrated that casual application of the method can lead to poor characterization of building equipment by blindly applying generalized equipment models to a specific building We therefore see this type of load monitor application as operator interactive, so that the monitor (and operator) can learn about the building and its individual characteristics. Whether software disaggregation is used with a load monitor or not, it has applications in building fault analysis.

Transient detection analysis has been shown to have such promising applications as equipment detection and health analysis. Although not fully developed for large building

equipment, we believe that it could produce valuable information to a building monitoring system.

Finally, non-intrusive load monitoring strategies were brought together in chapter 6 to illustrate the implementation of a monitor on a simulated building. We show how our techniques are designed to extract the useful information from a power signal at the level of intrusiveness that best suits the building. This last example describes how a load monitor can take advantage of the qualities of a specific building and it also presents the need for future work.

7.1 Future work

The focus of future work should be to understand the real building applications using the strategies presented in this thesis. The one drawback to this analysis is that each building is different. Therefore to make generalization as to the effectiveness of one strategy may not commute to another building. For that reason we have tried to offer more than one strategy.

Future work with respect to building information, transient detection, and software disaggregation is as follows:

Building information

• The noise produced from different equipment types needs to be quantified for both software disaggregation and transient detection.

• The number of simultaneous transient events that occur in buildings also needs to be quantified to determine the importance of equipment detection during these transients.

• Integration of load monitoring information with a larger building system such as an EMCS or an expert systems could be investigated.

Transient detection

• Perform fast sample tests on large HVAC equipment to understand the cluster variables that differentiate these equipment

• Implement cluster algorithms in an aggregate signal situation.

• Identify framing techniques for transient detection.

Software disaggregation

• Test the control signal of other HVAC equipment as an independent variable.

• Define a reasonable number of time varying equipment that can be disaggregated from one signal with a given noise signal.

• Test the software strategy on equipment, and simulate a fault.

REFERENCES

[Anderberg, 1973]   Anderberg, M.R.,1973. "Cluster Analysis for Applications."
                     Academic Press, New York.

[ASHRAE, 1983]   ASHRAE Handbook, 1983. "1983 Equipment." American Society
                     for Heating, Refrigeration and Air Conditioning Engineers, Inc.
                     Atlanta, Georgia.

[ASHRAE, 1984]   ASHRAE Handbook, 1984. "1984 Systems." American Society
                     for Heating, Refrigeration and Air Conditioning Engineers, Inc.
                     Atlanta, Georgia.

[ASHRAE, 1985]   ASHRAE Handbook, 1985. "1985 Fundamentals." American
                     Society for Heating, Refrigeration and Air Conditioning
                     Engineers, Inc. Atlanta, Georgia.

[Box, 1986]   Box, G.E., Draper, N.R., 1986. "Empirical Model-Building and
                     Response Surfaces." John Wiley and Sons, New York.

[Braun, 1987]   Braun, J.E., Mitchell, J.W., Klein, S.A., 1987. "Performance
                     and Control Characteristics of a Large Cooling System."
                     ASHRAE Transactions, vol. 93, part 1.

[Daryanian, 1991]   Daryanian, B., Bohn, R.E., 1991. "An Experiment in Real Time
                     Pricing for Control of Electric Thermal Storage Systems."
                     submitted for presentation at the IEEE Power Systems Society
                     1991 summer meeting.

[Draper, 1981]   Draper, N.R., Smith, H., 1981. "Applied Regression Analysis,
                     Second Edition." John Wiley and Sons, New York.

[Entek, 1988]   Entek Scientific Corporation, 1988. "Motormonitor, induction
                     motor fault diagnosis." Entek Scientific Corporation. Cincinnati,
                     Ohio.

[Gerald, 1989]   Gerald, C.F., Wheatley, P.O., 1989. "Applied Numerical
                     Analysis" Addison-Wesley Publishing Company, Reading,
                     Massachusetts.

[Hart, 1985]   Hart, G.W., 1987. "Prototype Nonintrusive Appliance Load
                     Monitor." Progress Report #2, Massachusetts Institute of
                     Technology Energy Laboratory, NERES, Concord,
                     Massachusetts.

[Hart, 1987]   Hart, G.W., 1987. "Minimum Information Estimations of
                     Structure." PhD Thesis, Dept. of Electrical Engineering and
                     Computer Sciences at The Massachusetts Institute of Technology.

[Hayt, 1978]   Hayt, W.H., Kemmerly, J.E., 1978. "Engineering Circuit Analysis." McGraw-Hill Book Company, New York.

[Hildebrand, 1986] Hildebrand, A., 1986. "A Knowledge System for HVAC Diagnostics." Ms. Thesis, Graduate School of the University of Colorado.

[Huffer, 1989]   Huffer, J.E., 1989. "Diagnostics in Heating Ventilation and Air Conditioning (HVAC) Systems using a Rule-Based System." MS Thesis, Graduate College of the University of Illinois at Urbana-Champaign.

[Pape, 1991]   Pape, F.L., Mitchell, J.W., Beckman, W.A., 1991. "Optimal Control and Fault Detection in Heating, Ventilating, and Air-Conditioning Systems." *ASHRAE Transactions*, vol. 97, part 1.

[Press, 1988]   Press, H.W., Flannery B.P., Teukolsky S.A., and Vetterling W.T., 1988. "Numerical Recipes in C." Cambridge University Press, Cambridge, England.

[Rice, 1988]   Rice, J.A., 1988. "Mathematical Statistics and Data Analysis." Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

[Stanford, 1988]   Stanford, H., 1988. "Analysis and Design of Heating, Ventilating and Air Conditioning Systems." Prentice Hall, Inc., Englewood Cliffs, New Jersey.

[Tabors, 1991]   Tabors Caramanis and Associates, 1991. "Measured Performance of the Thermal Storage System at the Bull HN Information Systems, Inc. Headquarters Building Billerica, Massachusetts.", Draft Report, Tabors Caramanis and Associates, Massachusetts.

[Wyatt, 1990]   Wyatt, J.L., Ilic, M., 1990. "Time-Domain Reactive Power Concepts for Nonlinear, Nonsinusoidal or Nonperiodic Networks." Proc. IEEE Int'l Symposium on Circuits and Systems.

System	Control tolerance	Response to load variations	Maintenance requirements	sophistication
ALL-AIR SYSTEMS				
Constant volume Typical applications: Restaurants, shops, assembly areas	Fair	Good	Low	Average
Variable air volume (cooling only) Typical applications: Core zones.	Good	Very good (cooling)	Low	Above ave.
Dual duct				
Constant volume Typical applications: Office, hospital, classroom.	Good	Excellent	Moderate	Above ave.
Variable air volume Typical applications: Office, hospital, classroom.	Good	Excellent	Moderate	Well above ave.
Multizone				
Constant volume Typical applications: small office building, classroom buildings, perimeter zones of large buildings	Good	Good	Low-moderate	Average
Terminal reheat (Applied to any system) Typical applications: close control of temperature: hospitals, laboratories, offices, computer rooms	Excellent	Excellent	Moderate	Average

AIR-WATER SYSTEMS

Induction

Two pipe	Fair	Fair	High	High
Four pipe	Fair	Good	High	High
Typical applications: perimeter zones, high-rise buildings				
Fan coil units	Fair	Fair-good	Moderate	Below ave.
Typical applications: hotels, motels, apartments, perimeter offices				
Variable volume (ceiling fan box)	Fair-good	Good	Moderate	Above ave.
Typical applications: offices				

ALL-WATER SYSTEMS

Heat pumps				
air-to-air				
air-to-water	Poor-fair	Fair	High	Above ave.
Typical applications: Apartments, hotels, motels, small office buildings				

162

APPENDIX B: JOHN HANCOCK TOWER

The equipment lists of the following pages are compiled from equipment plans for the John Hancock Tower in Boston, Massachusetts. The first list is of the HVAC fans within the building, and the second list is of the pumps.

For each equipment the lists give the following information:

1) Equipment identification number

2) Description of the equipment's function

3) Type of fan or pump

4) Floors that the equipment serves

5) Area on the floors that the equipment serves (North, South, East, West)

6) Size of the equipment in horsepower

7) Floor (mechanical room) where the equipment resides

Fan#	Description	Type	Floors	Area	HP	Loc.
33	dishwasher	exhaust			5	7
77	elevator smoke	exhaust		north	3	7
76	elevator smoke	exhaust		south	2	7
35	general	exhaust	4 - 34	north	10	7
26	general	exhaust	3 - 34	south	10	7
11	general	exhaust	3 - 6	west	0.8	7
46	generator room	exhaust			40	7
31	kitchen and cafeteria	exhaust			40	7
32	kitchen range hood	exhaust			75	7
45	mechanical equip. room	exhaust	#7		40	7
34	pantry	exhaust	#3		2	7
29	switchgear room	exhaust			3	7
30	toilet	exhaust	B - 41	north	20	7
27	toilet	exhaust	B - 24	south	15	7
28	transformer vault	exhaust			7.5	7
42	air	return	22 - 34	north	30	7
41	air	return	9 - 21	north	30	7
40	air	return	22 - 34	south	30	7
39	air	return	9 - 21	south	30	7
25	air	return	4 - 6	south	25	7
36	air	return	4 - 6	north	15	7
44	primary air	return	22 - 34	north	25	7
43	primary air	return	9 - 21	north	25	7
38	primary air	return	22 - 34	south	25	7
37	primary air	return	9 - 21	south	25	7
AH10	audio distribution room	supply			?	7
AH7	dining area supply	supply	#3		30	7
17	interior	supply	4 - 6	south	40	7
22	interior	supply	22 - 34	east	60	7
21	interior	supply	9 - 21	east	60	7
20	interior	supply	22 - 34	west	60	7
19	interior	supply	9 - 21	west	60	7
AH9	interior air	supply	4 - 6		15	7
9	lobby (upper)	supply			50	7
23	primary air	supply	9 - 34		250	7
18	primary air	supply	9 - 34		250	7
AH8	primary air	supply	3 - 6		75	7
83	conference	exhaust	#58		0.3	58
78	elevator smoke	exhaust		south	3	61
79	elevator smoke	exhaust		north	3	61
55	mechanical equip. room	exhaust	#61		3	61
75	pantry	exhaust	#59		1	61
64	switchgear room	exhaust			5	61
54	toilet	exhaust	42 - 60		15	61
49	air	return	35 - 46	south	30	61
58	air	return	35 - 46	north	30	61

47	air	return	47 - 58	south	30	61
60	air	return	47 - 58	north	30	61
50	primary air	return	35 - 46	south	20	61
57	primary air	return	35 - 46	north	20	61
48	primary air	return	47 - 58	south	20	61
59	primary air	return	47 - 58	north	20	61
74	board room	supply	#59		25	61
80	building service closet	exhaust	.		1	62
67	elevator sevice room	exhaust			5	62
52	general	exhaust	35 - 60	south	7.5	62
56	general	exhaust	35 - 60	north	7.5	62
65	mechanical equip. room	exhaust	#62		3	62
66	service elevator machine room	exhaust			0.8	62
81	tank room	exhaust			0.3	62
53	toilet	exhaust	25 - 41		5	62
51	transformer vault	exhaust			10	62
72	observation deck	return	#60		25	62
73	board room	supply	#59		75	62
AH11	cooling for engineers area	supply			1.5	62
62	interior	supply	35 - 46	south	60	62
69	interior	supply	35 - 46	north	60	62
61	interior	supply	47 - 58	south	60	62
70	interior	supply	47 - 58	north	60	62
71	observation deck	supply	#60		75	62
68	primary air	supply	35 - 50	north	250	62
63	primary air	supply	35 - 58	south	250	62
14	fan room	exhaust			7.5	BFR
13	fan room	exhaust			7.5	BFR
4	switch gear room	exhaust			3	BFR
82	telephose equip.	exhaust			0.8	BFR
3	transformer vault	exhaust			7.5	BFR
12	water and steam meter rooms	exhaust			5	BFR
8	bank	return			10	BFR
5	lounge	return			15	BFR
1	maintenance area	return			10	BFR
7	retail area	return			7.5	BFR
2	securities, vault and counting	return			7.5	BFR
AH4	bank	supply			40	BFR
6	cafeteria	supply			60	BFR
AH6	fan room ventilation	supply			25	BFR
10	lobby (under window)	supply			50	BFR
AH3	lounge	supply			20	BFR
AH1	maintenance and storage lockers	supply			25	BFR
AH5	retail area	supply			10	BFR
AH2	security,vault and counting	supply			10	BFR
16	kitchen	supply			15	CFR
15	kitchen	supply			15	CFR

Pumps

pump#	Description	Type	Floors	HP	Loc.
4	chilled water	double suction		200	7
6	chilled water	double suction		200	7
15	chilled water (winter)	double suction		40	7
5	chilled water stand-by	double suction		200	7
	condensate return			1.5	7
	condensate return			0.75	7
	condensate return			0.75	7
1	condenser water	single suction		400	7
3	condenser water	single suction		400	7
16	condenser water (winter)	single suction		100	7
17	condenser water aux.	single suction		300	7
2	condenser water stand-by	single suction		400	7
17A	condenser water stand-by aux.	single suction		300	7
60	cooling tower make-up	hor. split		150	7
61	cooling tower make-up	hor. split		150	7
7	exterior cold water	double suction	9 - 34	150	7
10	exterior cold water	double suction	3 - 6	20	7
9	exterior hot water	double suction	9 - 34	100	7
12	exterior hot water	double suction	3 - 6	15	7
8	exterior water stand-by	double suction	9 - 34	150	7
11	exterior water stand-by	double suction	3 - 6	20	7
13	interior hot water reheat	double suction	C - 34	5?	7
14	interior hot water stand-by	double suction	C - 34	5?	7
62	kitchen condenser water	single suction		5	7
25	chilled water	double suction		100	61
26	chilled water	double suction		100	61
32	chilled water (winter)	double suction		25	61
22	condenser water	double suction		250	61
23	condenser water	double suction		250	61
33	condenser water (winter)	double suction		50	61
34	condenser water aux.	double suction		60	61
24	condenser water stand-by	double suction		250	61
27	exterior cold water	double suction	35 - 58	150	61
29	exterior hot water	double suction	35 - 60	100	61
28	exterior water stand-by	double suction	35 - 58	150	61
35	hot well pump (steam condesers	single suction		2	61
30	interior hot water reheat	double suction	5?		61
31	interior hot water reheat stand-	double suction	5?		61
63	radiation perimeter	single suction		15	62
64	radiation perimeter	single suction		15	62
	condensate return			2	BFR
57	fan system	single suction		3	BFR
58	fuel oil transfer	gear type		0.5	BFR
59	fuel oil transfer	gear type		0.5	BFR
39	ground floor radiation	single suction		5	BFR
40	ground floor radiation	single suction		5	BFR

APPENDIX C: TRANSIENTS

This appendix contains examples of the transients tested with the apparatus described in chapter 5. The plot on the following pages are of instantaneous power, current, and, when indicated, voltage. There is also a PQ plot for each piece of equipment. The PQ plots use average real and reactive power as defined in section 5.4.1. The averages are calculated using a window of one 60 Hz wavelength. In order, the plots are:

Two fixture start-up 2
A start-up transient for two fluorescent light fixtures (each with two lamps) with Advance Mark VII electronic serial start ballasts.
a) Power
b) Current
c) PQ plot

Three fixture start-up 1
A start-up transient for three fluorescent light fixtures (each with two lamps) with Advance Mark VII electronic serial start ballasts.
a) Power
b) Current
c) PQ plot

Rapid start, start-up 2
A start-up transient for a fluorescent light fixture with a Magnetek B240R120 electronic rapid-start parallel ballast.
a) Power
b) Current
c) PQ plot

Instant start, start-up 3
A start-up transient for a fluorescent light fixture with a Magnetek B275I120 magnetic instant-start parallel ballast.
a) Power
b) Current
c) PQ plot

Instant start turn-off 2
A turn-off transient for a fluorescent light fixture with a Magnetek B275I120 instant-start parallel ballast.
a) Power
b) *Voltage
c) PQ plot

PC start-up 1
A start-up transient for a IBM XT personal computer.
a) Power
b) Current
c) PQ plot

Vacuum start-up 2
A start-up transient for a shop vacuum.
a) Power
b) Current
c) PQ plot

Two-Fixture Start-up 2

Power(watts)

Time(seconds)

Two-Fixture Start-up 2

Amps

Time (seconds)
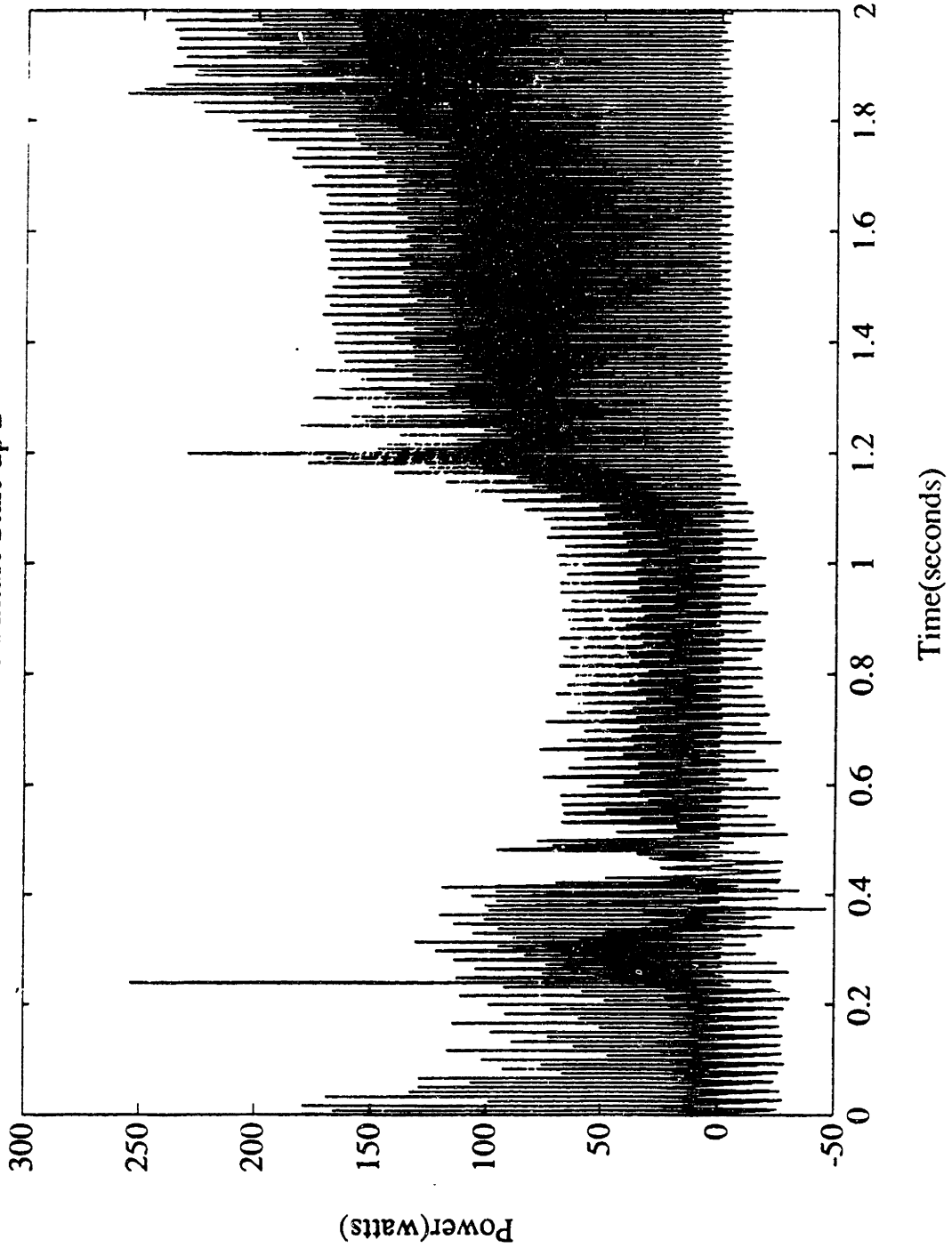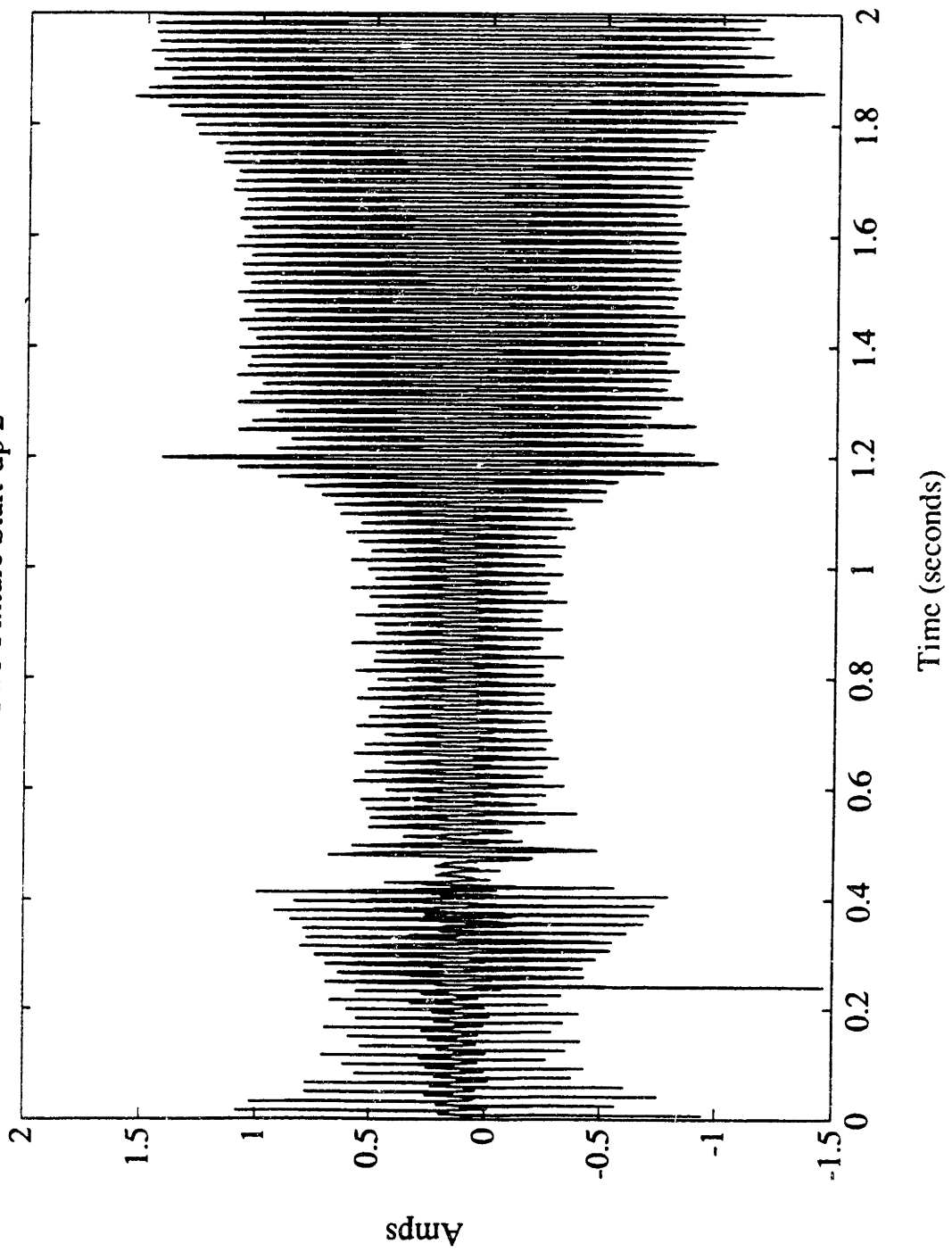
Two-Fixture Start-up 2

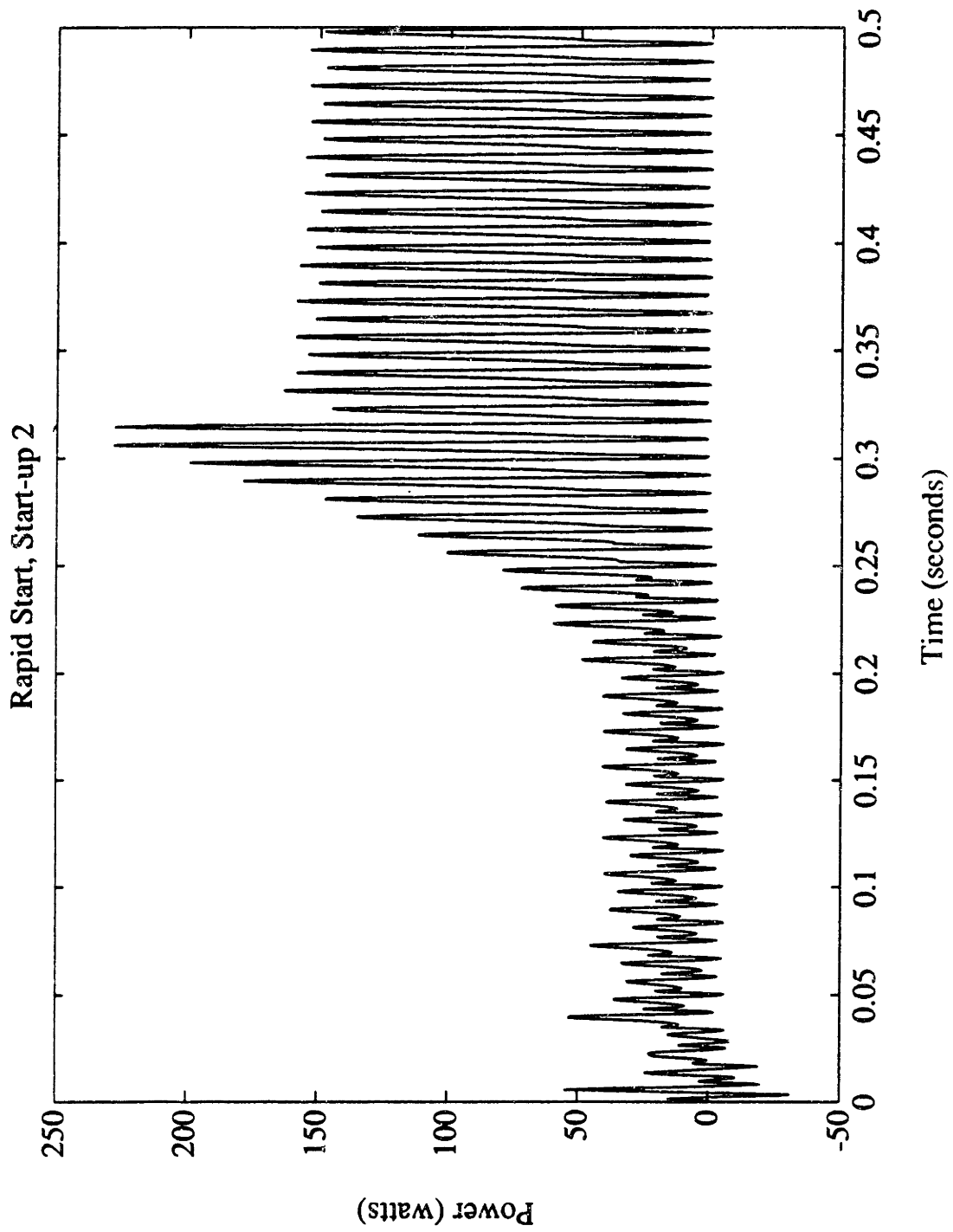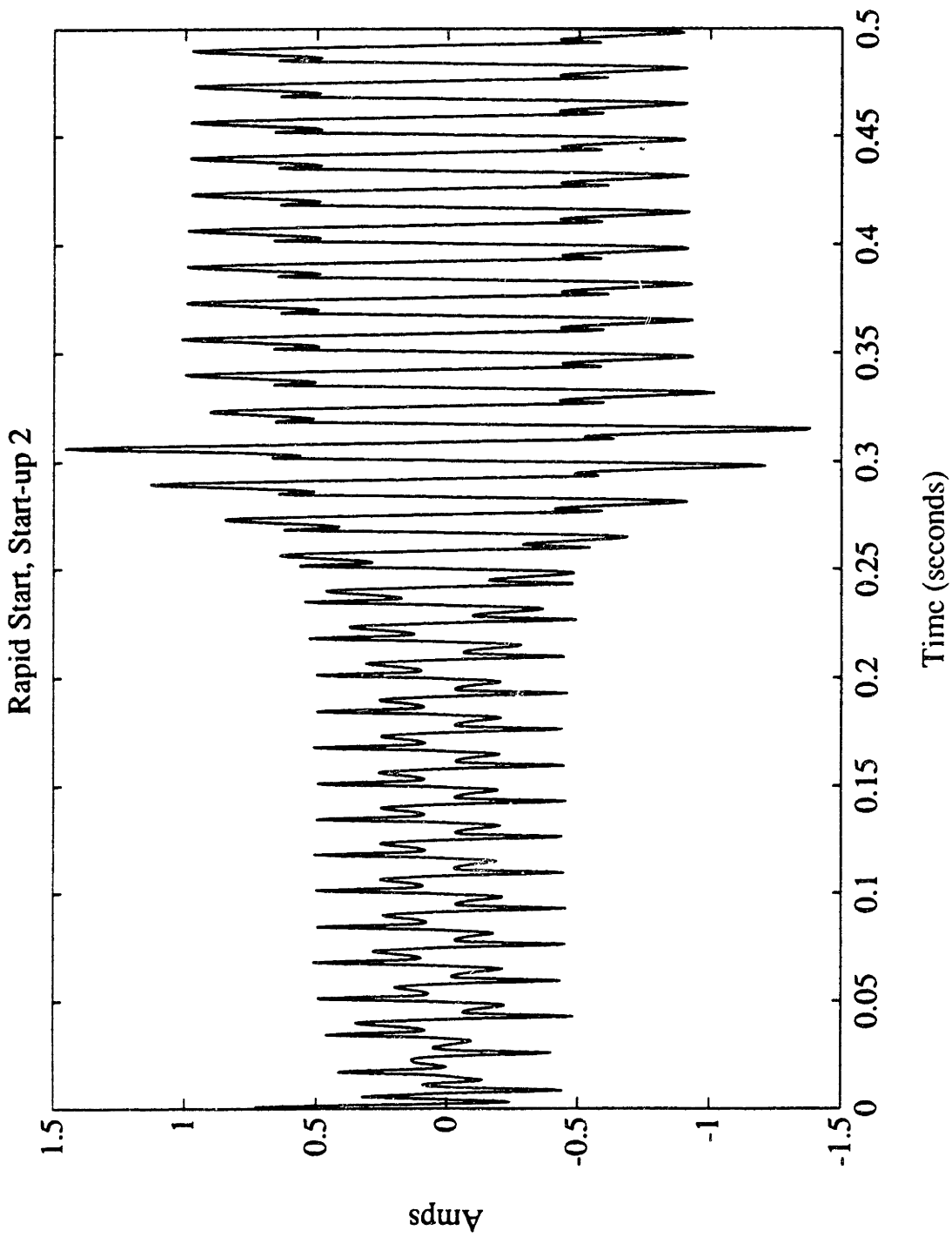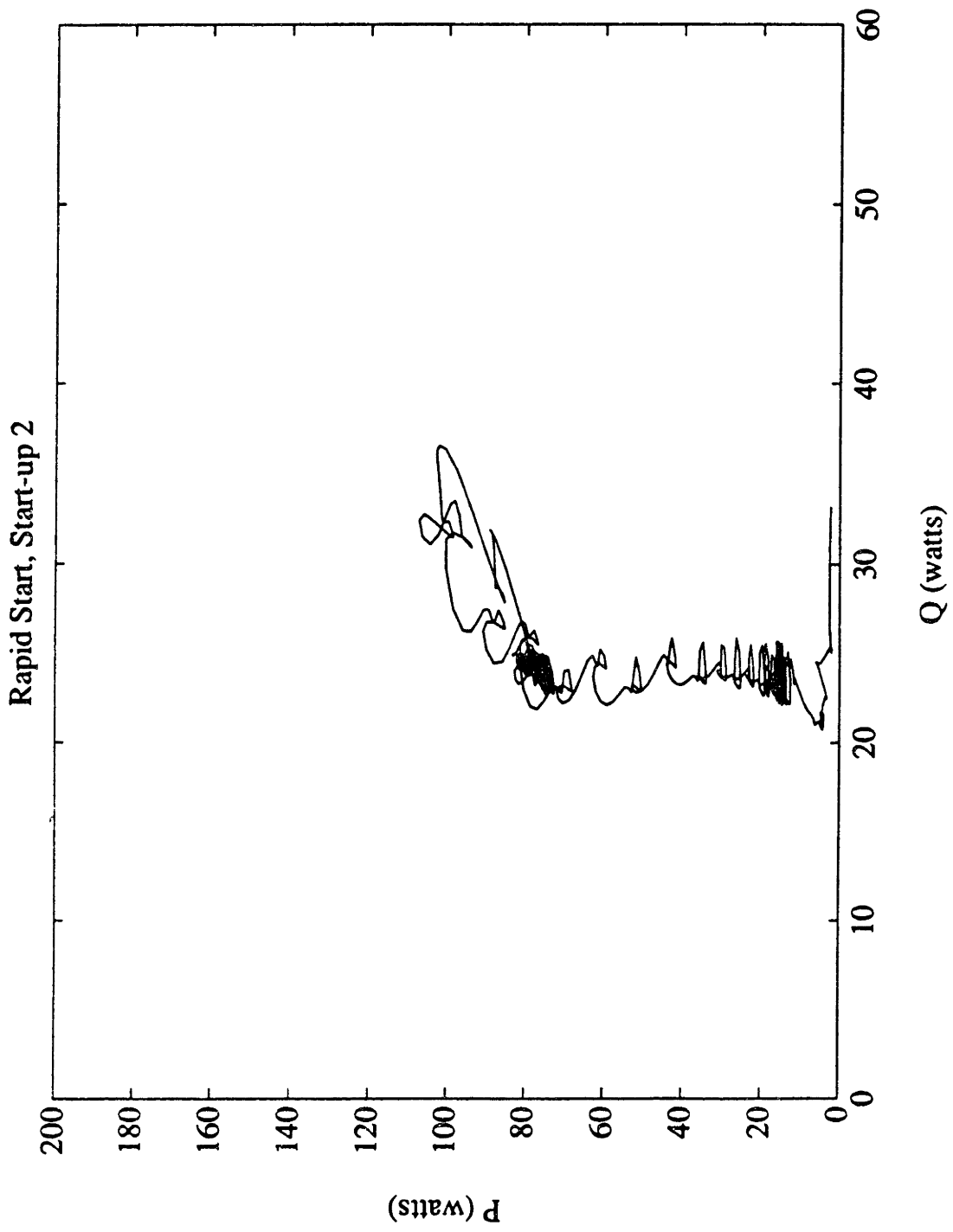Three-Fixture Start-up 1
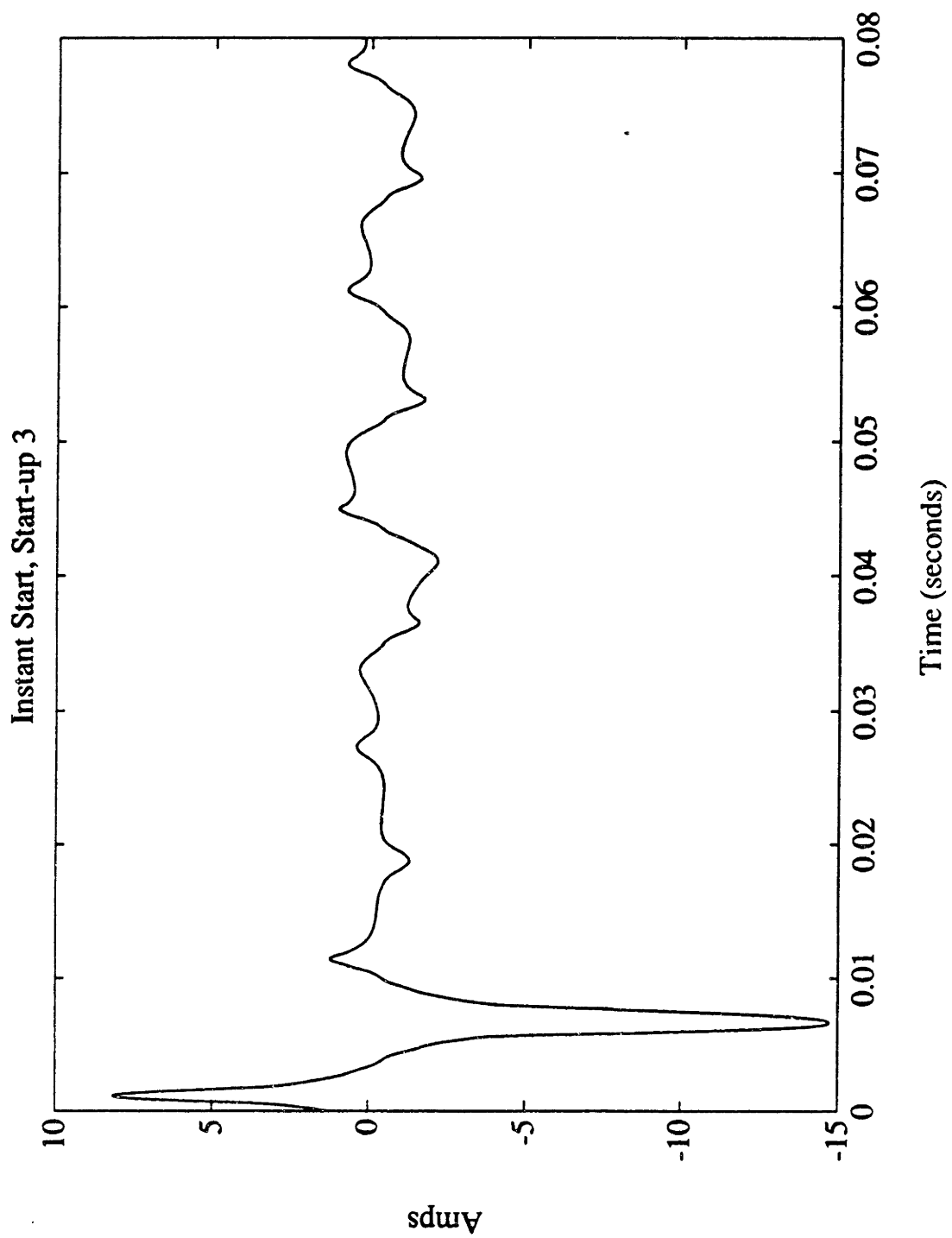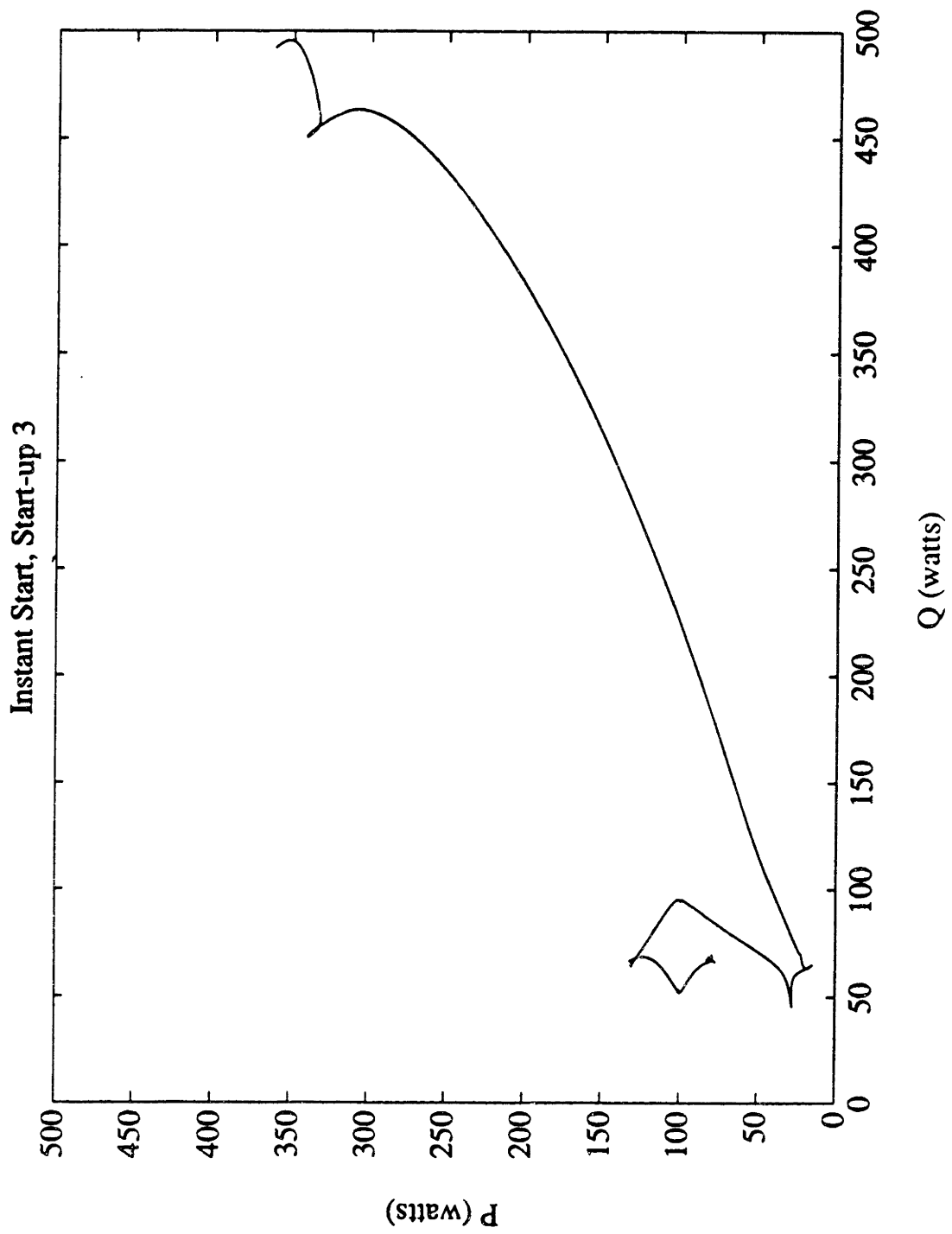
Power(watts)

Time(seconds)

171
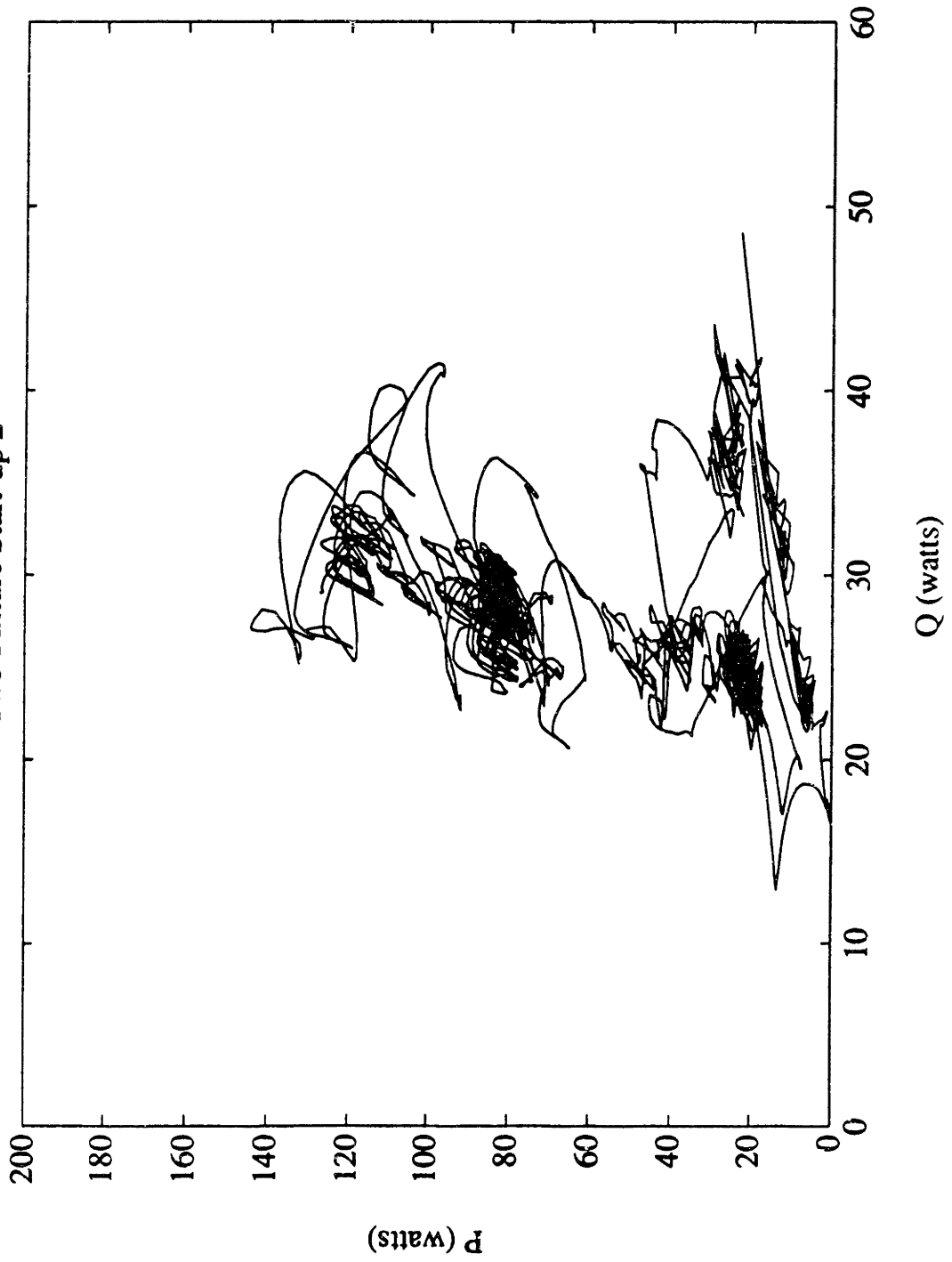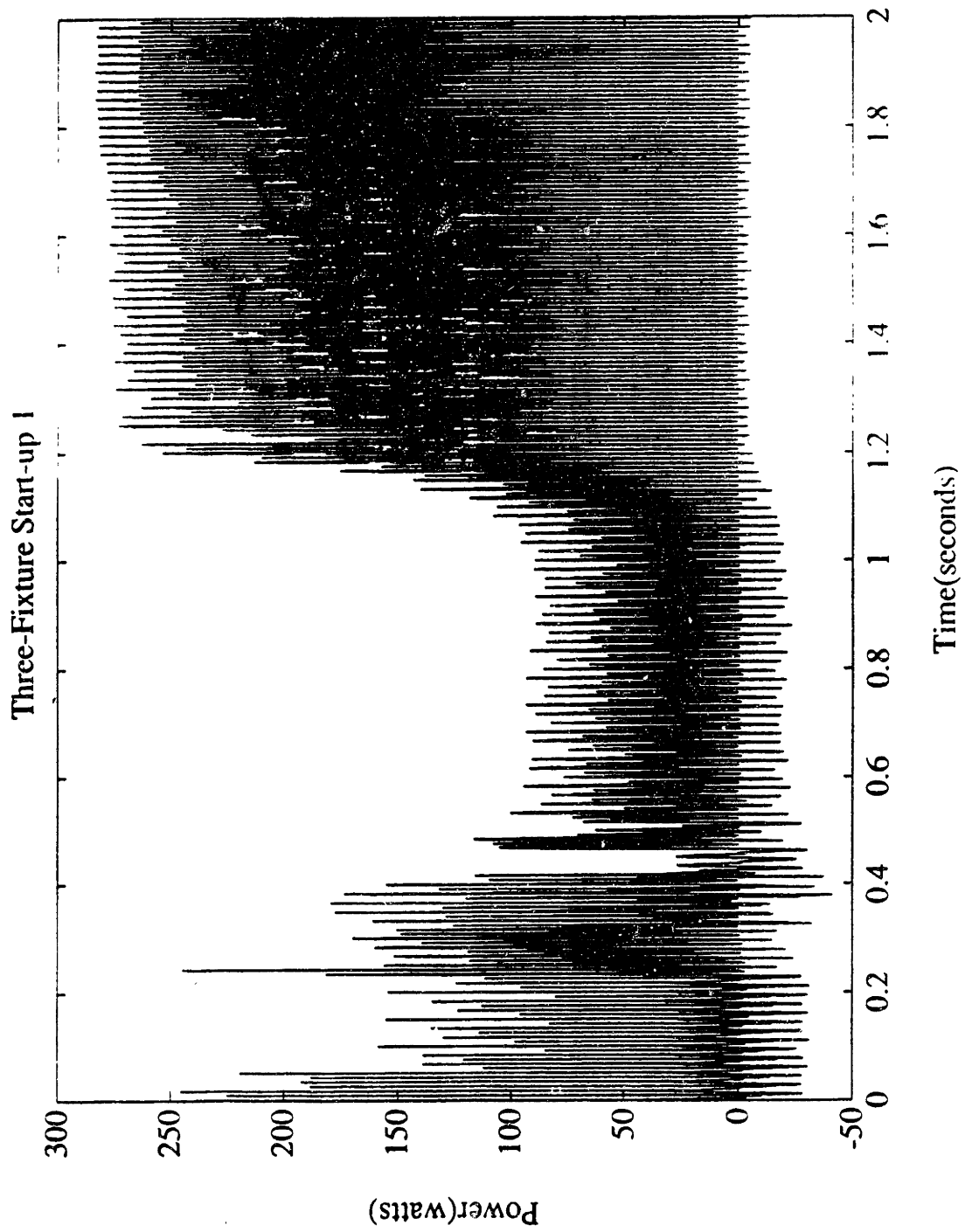
Three-Fixture Start-up 1

Three-Fixture Start-up 1

P (watts)

Q (watts)

173

Rapid Start, Start-up 2

Time (seconds)

Power (watts)

174

Rapid Start, Start-up 2

Amps

Time (seconds)

Rapid Start, Start-up 2

P (watts)

Q (watts)

176

Instant Start, Start-up 3

Time (seconds)

Power (watts)

Instant Start, Start-up 3

Amps

Time (seconds)

Instant Start, Start-up 3

P (watts)

Q (watts)

179

Instant-Start Turn-off 2

Power (watts)

Time (seconds)

Instant-Start Turn-off 2

Voltage(volts)

Time(seconds)

Instant-Start Turn-off 2

P(watts)

Q(watts)

PC Start-up 1

Power (watts) vs Time (seconds)

PC Start-up 1

Amps

Time (seconds)

PC Start-up 1

P (watts)

Q (watts)

185

Vacuum Start-up 2

Power (watts)

Time (seconds)

Vacuum Start-up 2
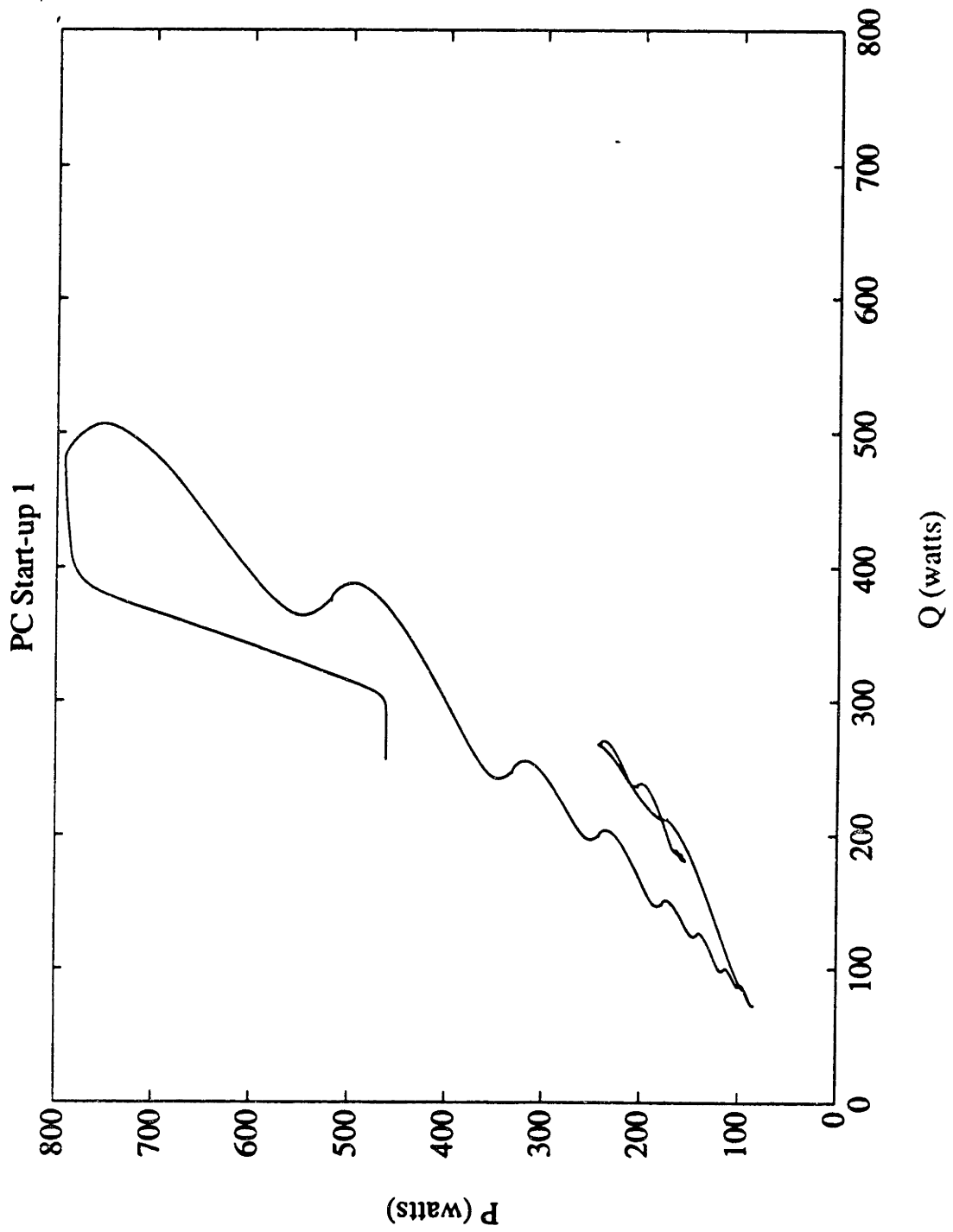
Amps

Time (seconds)

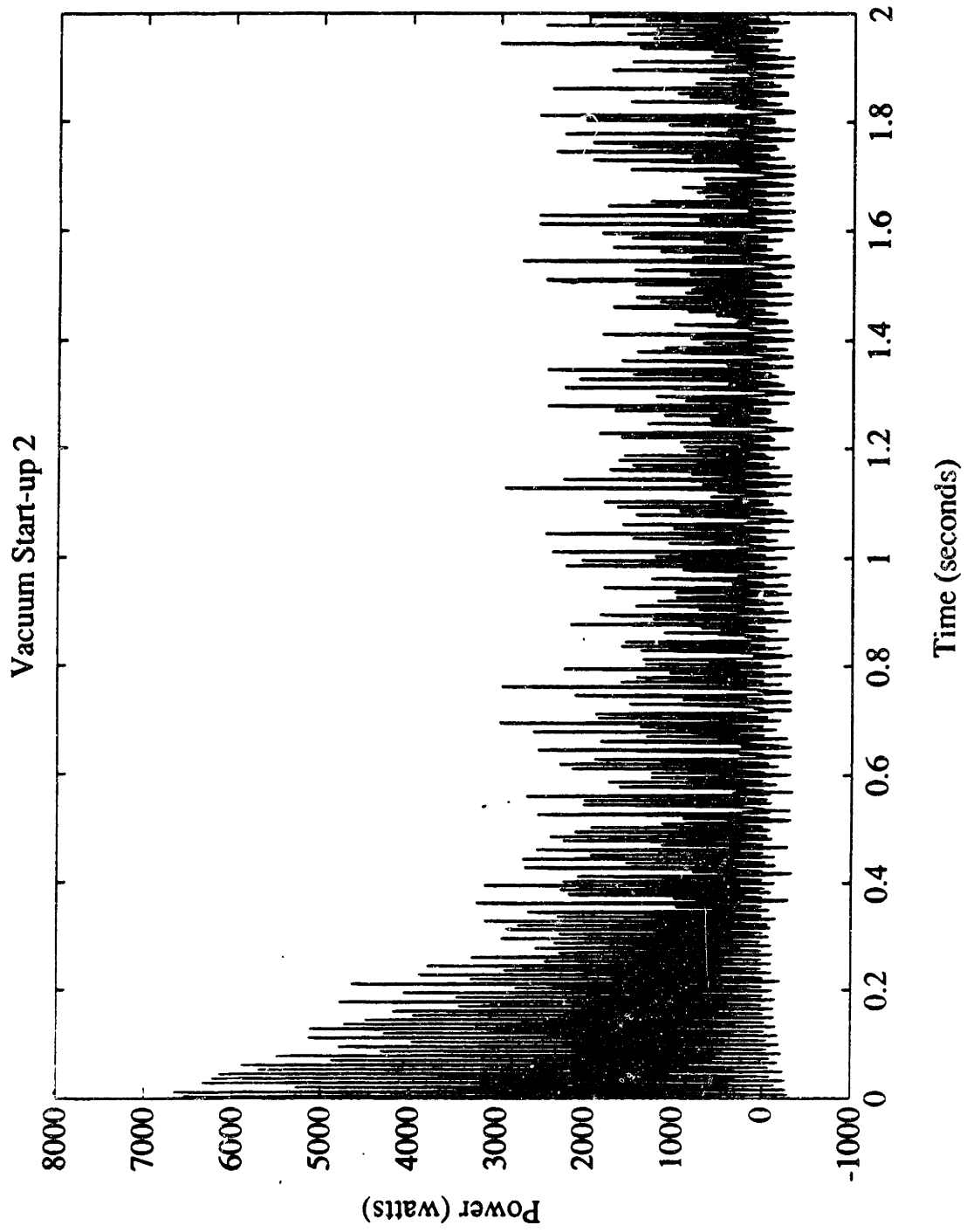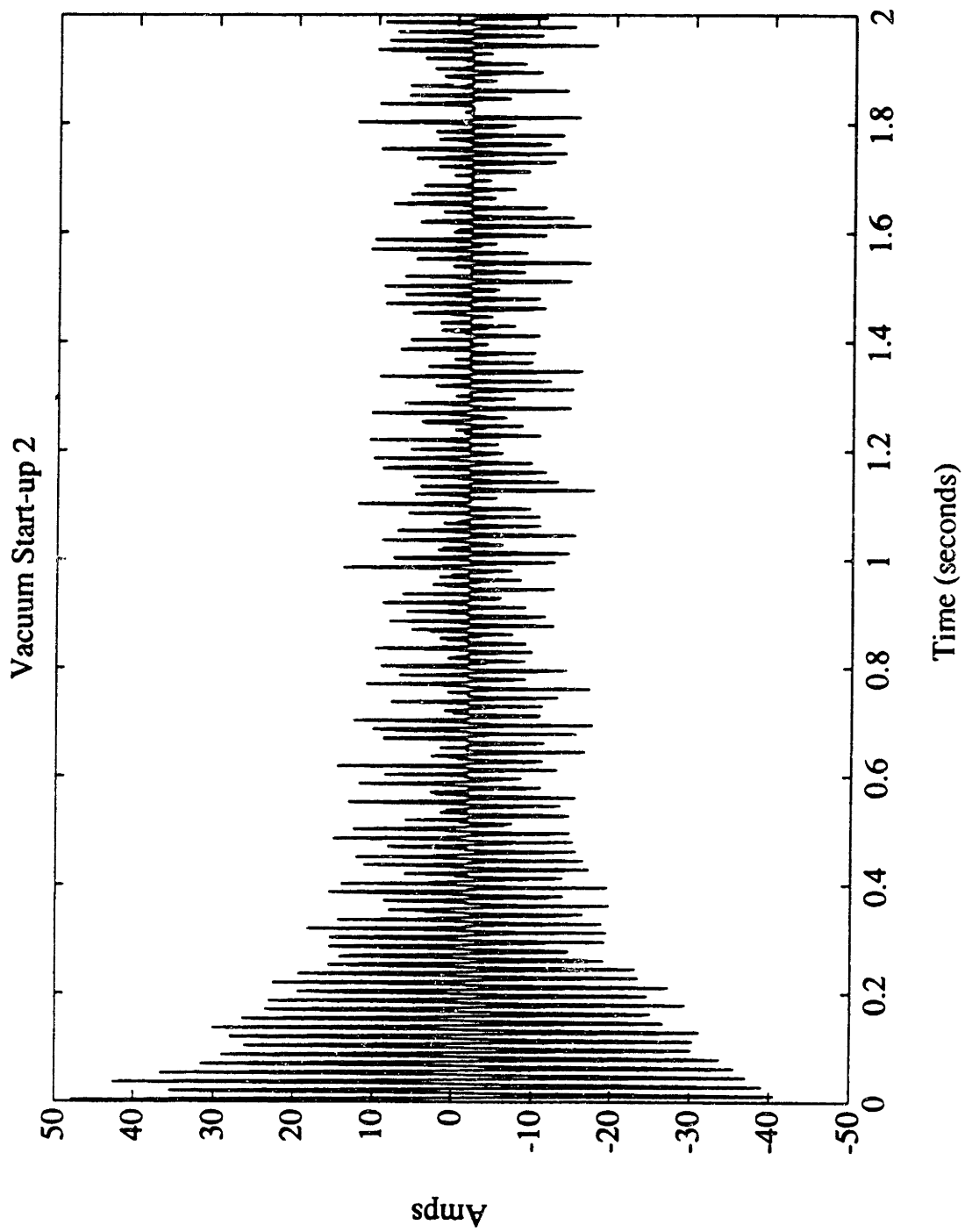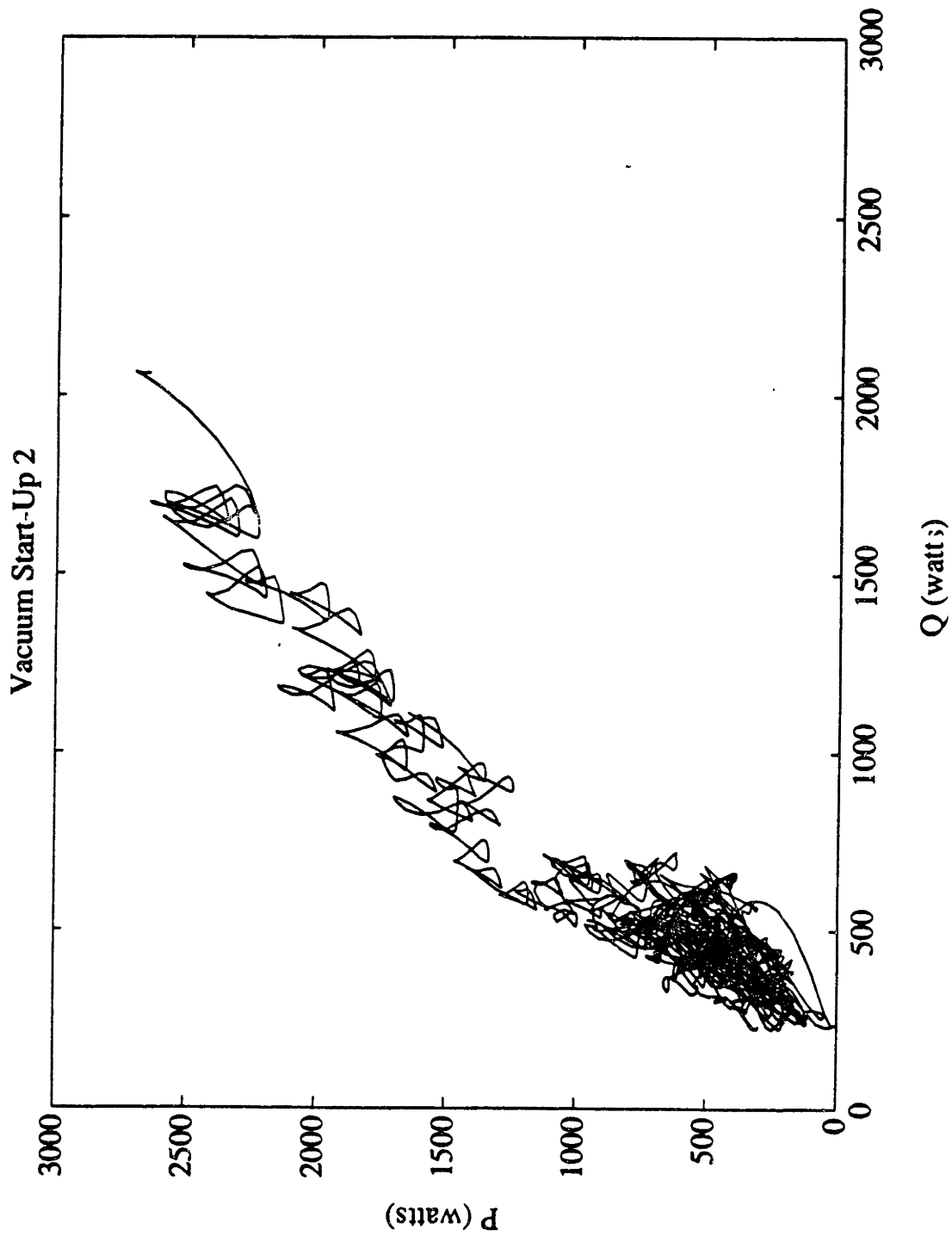Vacuum Start-Up 2

P (watts)

Q (watts)

APPENDIX D: SIMULATED ANNEALING

The method of simulated annealing is a technique that is very different from the least squares method. It is presented here as not only a valid technique, but to offer a different approach to the disaggregation of the time-varying loads.

Simulated annealing is a combinatorial minimization algorithm that is based on the Boltzman probability function: $Prob(E) \sim exp(-E/kT)$. Here, E is the energy state to be minimized; T is the "temperature" of the system; and k is the Boltzman constant. The idea is that an energy index, E, is chosen for a system to be minimized. As the system varies, the index, E1, is compared to the last system's index, E2. A new system is accepted if it is lower in energy (i.e. closer to the desired solution), or with a probability $exp(-(E2 - E1)/kT)$. The system therefore heads downhill towards minimum energy state, but it also has some probability of moving back up in energy to allow it to get out of local minima and into a more global minimum. As the algorithm searches for a minimum, the temperature T tends toward zero. Thus the system probability tends to zero for accepting an uphill move and the system "anneals" to a final state. The slower the temperature is reduced the less likely the system will get stuck in a local minimum.

To set up the algorithm for our problem on MATLAB, the following is needed:

1) Configuration: The parameters must be chosen to some initial value. We chose the parameter values of the least squares fit as a starting point.

2) Rearrangement: The parameters in our case were increased or decreased randomly by a set interval. The parameters were only allowed to vary within the error of the least squares results.

3) Index or energy function: We chose the sum of the absolute value of the difference between the total power and the estimated power.

4) Annealing schedule: The rate at which the temperature T is reduced [Press 1988].

The reason for choosing the absolute value instead of the square of the errors as with the least squares estimate is that the absolute value index is much less sensitive to outliers. This is a significant advantage especially when the sample size of the data is small and sensitive to each data point.[Rice 1988]

One advantage of simulated annealing is that one must tailor the configuration and rearrangement to the problem. This is beneficial to our problem because any knowledge that is known *a priori* can be incorporated into the algorithm. For instance, in the

MATLAB simulation we used the results from the least squares estimation as initial values of the coefficient matrix parameters. The variances of the parameters were then used as boundaries for the parameters to vary within. In this way the algorithm applied a more accurate index for minimizing the problem of the already estimated parameter set. This is one example of using *a priori* knowledge in simulated annealing.

Simulation

The implementation of simulated annealing on MATLAB used the same quadratic model as section 4.2 on the least squares simulation. The algorithm used had the following elements:

* The least squares estimates of the coefficients were used as initial values.

* The values were allowed to fluctuate within error bounds as specified by the confidence intervals of the model fit.

* The rearrangements consisted of the parameters randomly jumping up, down or staying the same by the initial amount of the confidence interval divided by ten.

* The magnitude of the jumps decreased with the temperature index.

* The choice of selecting a configuration is based on the index and the Boltzman probability: Prob= exp((stored index-new index)/(temperature index)).

* The index was the sum of the absolute value of the difference between the total power and the estimated power.

* The annealing temperature decreased with each increment of the loop counter.

Implementation

After running the simulation on MATLAB, it was first noticed that the simulation took many times longer than the least squares fit. As the temperature index was allowed to decrease at a slower rate, the annealing provided better results but the runs became restrictively long (on a 286 AT computer). This was especially true when the rearrangement methods and Boltzman constant were being tested and adjusted.

The Boltzman constant k, found in the probability function:
Prob = exp(-(E2 - E1)/kT), was adjusted to increase or decrease the probability of accepting new configurations. It was found that this constant was very temperamental to adjust. If chosen too large, the algorithm annealed too fast and returned the least squares estimate that it started with. If chosen too small the estimate that was returned was much worse that the least squares estimate that it started with.

Another difficult attribute to adjust was the method of rearrangement. Even though the rearrangement was random and looked for a minimum within the error band where it

was likely to find it, it had no way of judging whether it was headed the right direction when it found a likely configuration. In other words, it was arbitrary whether it would do better than the least squares estimate. If the bands with which the parameters were allowed to vary were small, the Boltzman constant was large, and the temperature was decreased slowly (typically a ten minute run), then the algorithm would sometimes do better. This can be seen in the following results.

The index that was used to compare the two techniques was the absolute value of the difference between the model fit and the data, which was used for the simulated annealing.

RUN	LEAST SQUARES INDEX	ANNEALING INDEX
1	9.436	9.442
2	8.978	8.982
3	127.069	126.691
4	35.143	35.565
5	376.149	376.568
6	37.96	38.03
7	703.282	704.514
8	882.330	877.353
9	139.656	140.196
10	818.013	816.290

It is not known how the least squares confidence intervals change with the improved annealing indexes.

Conclusion

After comparing the two MATLAB simulations of least squared and simulated annealing techniques, we found that the least squares method suits our problem better if the outliers are small enough for adequate models (within our needs). The simulated annealing proved to be better when data was sparse and a priori data could be used. The least squares simulation ran quickly and was easy to implement using MATLAB functions. Simulated annealing however was slow, temperamental, and required tailoring to fit our problem. The tailoring proved useful only when a priori knowledge was used.

APPENDIX E: BUILDING OPERATOR INTERVIEWS

Most of the applications for a load monitor are associated with building faults. For this reason, we proceeded with a study to interview building operators about their buildings and how the building equipment breaks down. In addition, we were interested in the frequency of equipment failures which could serve as an added area of information for a monitor's knowledge base.

We interviewed the building operators from 3 hospitals, 6 universities and 3 private buildings in the Massachusetts area. Initially we framed the interviews around the following questions:

1. *What major equipment do you maintain?*
2. *What are the most common problems associated with these equipment by type?*
3. *Which are the most frequent?*
4. *Which are the most severe?*

The interviews typically lasted about an hour and were conducted in an open-ended format. With this style, the interviewees were encouraged to discuss the building problems that they feel are important with minimal guidance from the interviewer. However, it was necessary to maintain a flow to the interview which allowed some focus on the areas the interviewer deemed important.

We had the following general impressions about this process.

1) The interviewees tended to want to answer specific questions about their facility rather that relating their experiences with equipment faults. Hence, there was a tendency to have the building operator describe the building and not the faults.

2) We felt that it was an imposing question to ask about equipment failures. The building operators seemed to be defensive about the question of what problems there are in their building, or that their equipment have any failures.

3) There was, however, a "sweet spot" at the end of the interview when the interviewee felt that there were no more questions to be asked. At that time, we found that they were more reflective on the subject, at which time it was more clear to them what it was we were after. This period in the interview produced the most useful information.

4) There is a lead time associated with knowing what questions to ask. Without knowledge of what equipment might fail, it is difficult to lead the interviewee to discuss faults. However, without talking to building operators, it is difficult to know what fails.

While providing us with insight as to how buildings are maintained and run from an electrical, and mechanical viewpoint, the interview produced the following information. Note that our questions were not designed in such a way as to allow for statistical analysis. Therefore the information compiled is based on common observations made by the building operators.

Equipment maintenance problems

Hospital have intensive preventive maintenance (PM) programs while universities have varied levels of PM. In general, major equipment did not break down more that once a year and in most cases the building operator could not think of major failures that occurred in the last 2-4 years. The buildings that did have breakdowns could be correlated to small or nonexistent maintenance programs.

Instead, the main equipment issue that was stressed by building operators was maintaining equipment by replacing parts that wore out but rarely failed. In the case of failures, the cause was usually negligence of the equipment. The main types of replaceable parts were:

Fans: belts, bearings, shims, filters
Pumps: bearings.
Chillers: No chiller faults were noted. In all cases chillers were maintained by outside contractors that routinely disassembled the units and overhauled them.

One of the reasons that our interviews did not discover many equipment failures is because maintenance facilities do not deal commonly with failed equipment, but failing equipment, which is corrected as quickly as possible. If a piece of equipment tends to failure often, then it is replaced. Hence, the question of how often equipment fail leads to an answer that does not address the issues of equipment problems.

The types of failures that the building operators found most difficult to detect and maintain were system errors. Because systems control themselves, an error in one part of

194

the system can be compensated for by another part of the system. The most noted problem by the operators was broken or stuck dampers (or valves). If a system doesn't operate for an off-season or the system does not vary much in time, the control actuators can freeze up. The system will change in efficiency and its quality of control, but otherwise there is no way of detecting the problem. Furthermore, these components of the system are generally not designed to be visually inspected.

With this building information, we see that the building operators have a need for characterizing their equipment within a system, while the individual catastrophic failure does not seem to be a priority.

One could argue that the monitor could replace the need to visually inspect HVAC equipment for problems, however, we feel this would be a mistake. The programs that we encountered were run by people who had a feel for their equipment and any intervention of that would only take valuable information away from them. The monitor can best be used as a device that provides a different perspective on the operation of building equipment. If it can be a microscope into the area of building monitoring, then it will be a valuable tool to a building operator.