Energy Analysis in New York City Multifamily Buildings: Making Good Use of Available Data

Margaret F. Fels and Cathy L. Reynolds

PU/CEES Report No. 273
October 1992

Final Report to
New York State Energy Research and Development Authority
(for Agreement #1646-EEED-BES-91)

To be published by NYSERDA

Center for Energy and Environmental Studies
Princeton University
Princeton, NJ 08544

ABSTRACT

Available energy consumption data from large multifamily buildings are assessed in terms of their quality and accuracy, and in terms of information derivable from simple analytical techniques applied to the data. Fuel consumption records from three data sets of multifamily buildings in New York City are used, and the results of applying the PRISM methodology to the data are analyzed. Data range from handwritten fuel records for one or two years taken from a survey that included data on building characteristics, to computerized fuel consumption records spanning many years for cases reporting energy conservation actions. The study emphasizes problems encountered in oil delivery data and possible data improvements that can result from careful data screening and application of simple statistical techniques for outlier detection.

New procedures for data improvement based on PRISM are derived from these analyses. Application of these procedures provides encouraging evidence that existing consumption data may be sufficient for meaningful monitoring of energy conservation in large multifamily buildings. Overall, by combining PRISM with data improvement techniques, reliable indices of weather-adjusted consumption appear generally feasible for oil-heated as well as gas-heated multifamily buildings.

ACKNOWLEDGEMENTS

This study was funded by the New York State Energy Research and Development Authority (Energy Authority), Agreement #1646-EEED-BES-91. The authors would like to thank Norine Karins at the Energy Authority for her valuable input and helpful comments on the progress of this research. We would also like to thank Fred Goldner of Energy Management Associates (formerly of the New York City Department of Housing Preservation and Development) for providing the oil consumption data from the Building Energy Use Tracking System data base and the detailed data from the Building Monitoring Data project and for offering helpful comments on our research. Thanks also to Mark Engel of Ralph Langsam Associates who provided the oil delivery and vacancy data for the Building Monitoring Data project data base, and Barbara Tillman of Grenadier Realty and George Birman for providing the consumption data for the data base of Energy Conservation Cases. We are especially grateful to Wayne Rosen and Erik Nelson of Princeton University who painstakingly pored through the data, created the PRISM-ready data files, and did a lot of the PRISM runs reported here; and to Richard DeVeaux who was very helpful in suggesting analytical techniques for our study of outlier detection. Elizabeth Steward's remarkable patience in typing and retyping drafts of this report is also appreciated.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued) **Page**

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF TABLES

EXECUTIVE SUMMARY

The overall objective of this study was to assess the validity and usefulness of available energy consumption data in New York City multifamily housing, particularly in terms of the information derivable from simple analytical techniques applied to these data. More specifically, the objective was threefold:

- to determine the quality, accuracy and usefulness of available energy consumption data (oil deliveries and utility meter readings) in multifamily buildings;
- within the constraints of available data, to determine the usefulness of PRISM (PRInceton Scorekeeping Method) for monitoring consumption and measuring energy savings in multifamily buildings in New York City, particularly those with oil heating; and
- to explore the usefulness of PRISM as a data-screening and data-cleaning tool, for improving the quality of available data and ultimately the reliability of model results obtainable from the data.

As a simple physically based model with sophisticated statistics, PRISM is well suited to data quality assessments, and to the development of procedures for improving and making better use of the data.

Three data bases from projects sponsored by the Energy Authority were used:

1) a sample of 71 multifamily buildings from the Building Energy Use Tracking System (BEUTS) for which reports of building characteristics and fuel consumption records were compiled;

2) the Building Monitoring Data (BMD) project consisting of 30 oil-heated multifamily buildings with oil delivery data spanning several years and data-recording energy management systems; and

3) Energy Conservation Cases (ECC) consisting of case studies of efficiently managed multifamily buildings.

Sample results are highlighted in this Summary.

In the study of the BEUTS sample, a technique for outlier detection, the studentized residuals test, was applied for the first time in PRISM analysis. Emphasis was on problems peculiar to oil data, such as non-fill deliveries that can show up as outliers in the consumption data. Judicious combination of outliers was shown to yield substantial, and sometimes dramatic, improvements in the quality of the PRISM fits of the data, as measured by the R^2 statistic, reflecting the goodness of fit, and CV(NAC), the relative standard error of Normalized Annual Consumption (NAC), the weather-adjusted index of consumption. In one example, data correction through outlier detection yielded an improvement in R^2 from 0.14 to 0.83 and an improvement in CV(NAC) from 0.54 to 0.07. (See Table 1 in the text.) Overall, the improvement was more pronounced for oil data than it was for gas data: the median R^2 improved from 0.74 to 0.82 for the 57 oil-heated buildings, and from 0.95 to 0.98 for the 14 gas-heated buildings (see Figures 4 and 8). The additional analysis for oil data is thus shown to be worthwhile.

In the study of the BMD data base, outliers were identified in spite of the expected high quality of the data, and their correction yielded substantial improvements in the quality of the PRISM results (Figure 9). Frequent oil deliveries in some months allowed an assessment of improvements from monthly aggregation of the deliveries, as a possible procedure for smoothing anomalous consumption data from deliveries spanning short periods (especially one to three days). The median R^2 for the 30 buildings increased from 0.70 when the original data were used to 0.84 when the data were aggregated in monthly increments; correspondingly, the median CV(NAC) decreased from 0.13 to 0.09 (Figure 11 and Table 4). However, since the fits were not improved substantially more by monthly aggregation than by outlier correction of the original data, and since more information is lost in monthly aggregation, straightforward outlier correction is recommended even in cases of frequent oil deliveries.

The availability of vacancy rates in the BMD data set led to an exploratory study of whether vacancy rates should be explicitly included in a weather adjustment of a building's energy consumption data. Several different vacancy-rate analyses were performed. The results suggest that, at least for buildings where vacancy does not vary greatly from month to month, the consumption per occupied unit can be reliably and simply estimated by dividing NAC from the PRISM run of building-level oil-delivery data by the average number of occupied units during the estimation period (Table 10).

In the study of the ECC data, many years of high-quality consumption data in conjunction with records of conservation measures provided an opportunity to test the multifamily-building PRISM approach for monitoring conservation. Even in this data set, the need for data improvements was indicated. One benefit was that trends in consumption became more clearly discernible after data improvements (Figures 17, 23 and 26). Data for one building included runtime-metered data as well as oil billing records, with more frequent summer readings in the former than the latter. Comparison of PRISM results showed the difference between the two data sets being primarily attributable to additional summer information in runtime-metered data (Figure 17 vs. Figure 22, and Table 12). This case study verifies the importance of distinct summer information for reliable estimates of weather-adjusted consumption.

Using the lessons derived from analyses of the three data sets, procedures were formulated specifically for outlier detection and data improvements and more generally for enhanced PRISM analysis. The procedures are presented as guidelines for future PRISM applications to multifamily buildings.

From the examples in this study, it is clear that painstaking analysis can transform a seemingly not useful data set into one that can yield reliable and useful consumption indices, making the extra work for data improvement worthwhile. Although the reliability of PRISM estimates for oil-heated multifamily buildings on average remains somewhat lower than has been seen in studies of houses and multifamily buildings with gas heating, as well as of houses with oil heating, the successful application of PRISM to a large fraction of the oil-heated buildings analyzed, with reliable NAC estimates and high model R^2 statistics, is an encouraging indication that readily available consumption data may be sufficient for meaningful monitoring of energy conservation in large multifamily buildings.

Energy Analysis in New York City Multifamily Buildings: Making Good Use of Available Data

Margaret F. Fels and Cathy L. Reynolds

Center for Energy and Environmental Studies
Princeton University
Princeton, NJ 08544

October 1992

1. INTRODUCTION

Research objective

Good energy information in buildings is needed both to identify energy conservation opportunities and to measure the actual savings achieved by whatever energy conservation actions are taken. For single-family houses, energy bills have provided a useful source of the needed information. Furthermore, very simple analytical tools applied to these data can produce reliable weather-adjusted and time-adjusted consumption indices needed for providing measurements of savings. One such method is PRISM (PRInceton Scorekeeping Method), whose applicability is far better understood for single-family houses than it is for multifamily buildings (Fels, ed., 1986). This study uses PRISM to explore the usefulness of available data in large multifamily buildings in New York City.

In multifamily housing, particularly low-income, acquisition of required information may be less straightforward than it is for single-family houses. Record-keeping is not at all uniform among building managers, and experience suggests that records that are obtainable may not be accurate, e.g., a building may be billed for oil delivered to another customer. Nevertheless, there is a wealth of useful energy information for multifamily buildings, just as there is for single-family houses. Monthly electric and natural gas meter readings are available from energy bills; oil delivery data are, at least in theory, recorded on bills; and in many buildings more detailed data are collected through energy management systems. Since instrumentation to monitor energy consumption directly is very costly, both in terms of equipment and person-time requirements, the best possible use of available data is warranted simply on the basis of cost savings.

An added motivation for improved energy analysis techniques in large multifamily buildings is the vast energy conservation potential represented by that sector (Hewett, 1988). For example, mechanical system retrofits are likely to be more cost-effective in multifamily buildings than in single-family buildings because of lower surface-to-volume ratios in larger buildings. Since engineering estimates are often inaccurate predictors of energy savings, better use of available data on actual consumption is needed for more accurate and widely useful performance monitoring of retrofits in multifamily buildings.

The overall objective of this study is to assess the validity and usefulness of available energy consumption data in New York City multifamily housing, particularly in terms of the information derivable from simple analytical techniques applied to these data. More specifically, the objective is threefold:

- to determine the quality, accuracy and usefulness of available energy consumption data (oil deliveries and utility meter readings) in multifamily buildings;
- within the constraints of available data, to determine the usefulness of PRISM for monitoring consumption and measuring energy savings in multifamily buildings in New York City, particularly those with oil heating; and
- to explore the usefulness of PRISM as a data-screening and data-cleaning tool, for improving the quality of available data and ultimately the reliability of model results obtainable from the data.

As a simple physically based model with sophisticated statistics, PRISM is well suited to data quality assessments. When it works well, such as when applied to a year of consumption data for a building, the PRISM results provide verification of data quality as well as useful consumption indicators. Furthermore, when PRISM does not work well, the method becomes a useful pre-processor of the data, an identifier of data problems and anomalies, and a guide for possible data improvements. Thus, the study is intended not only to test the quality of available consumption data but also to develop procedures for improving and making better use of the data.

PRISM is a simple standardized method for determining weather-adjusted estimates of energy savings. It has been extensively validated on gas-heated and electrically heated houses, and has been used effectively in evaluations of numerous residential energy conservation programs (see, for example, Hirst, 1987, and Beschen and Brown, 1990; see also Gerardi, 1991, an Energy Authority study). Over the years, the method has been subjected to more validation tests and interpretability studies than probably any other energy analysis tool based on measured data. Nevertheless, an understanding of its applicability to large multifamily buildings, particularly those with oil heating, has been a major gap in the research. Previous PRISM studies of single-family homes heated by oil (Fels et al., 1986) and of multifamily buildings heated by gas (DeCicco et al., 1986, Goldman and Ritschard, 1986) have indicated that PRISM might indeed be a valuable tool for oil-heated as well as gas-heated multifamily buildings. However, this hypothesis needs testing on real data. The availability of several rich data sets for multifamily buildings in New York City provided the impetus for this study.

Synopsis of PRISM

PRISM, the PRInceton Scorekeeping Method developed at Princeton University, uses utility bills from before and after retrofit installation, together with average daily temperatures from a nearby weather station for the same time periods, to determine a weather-adjusted index of consumption, Normalized Annual Consumption (NAC), for each house (or building).* The NAC index provides a measure of what energy consumption would be under typical weather conditions. The total energy savings are then derived as the difference between the NAC in pre- and post-weatherization periods. An energy conservation effect is thus neither masked by a cold winter nor exaggerated by a warm one, nor is it obscured if the time covered by billing periods in one "year" is longer or shorter than another.

Required input for PRISM includes:

- individual-building consumption data, for the heating fuel, including exact meter reading (or oil delivery) dates, for approximately one year in each period of interest;
- average daily temperature data, for the periods of interest, from a nearby weather station; and
- long-term degree-days computed from daily temperatures from the same weather station over about 10 years.

Using an iterative procedure, three physical parameters result from the model applied to the data for the heating fuel: reference temperature, or break-even temperature, below which the heating fuel is required; heating slope, as a measure of the heat-loss rate of the house; and base-level consumption, a measure of appliance use in the house. As shown in Figure 1, NAC is then derived from these parameters as the estimate of the house's weather-normalized consumption. In addition to these output values, statistical analyses are performed, giving the reliability of each estimate.

PRISM differs from other weather-normalization procedures in that the building's break-even temperature is treated as a variable rather than a fixed value, such as the commonly used 65°F. That value of the reference temperature

---

*A complete description of PRISM, together with a range of PRISM applications, is available in a special issue of Energy and Buildings (Fels, ed., 1986).

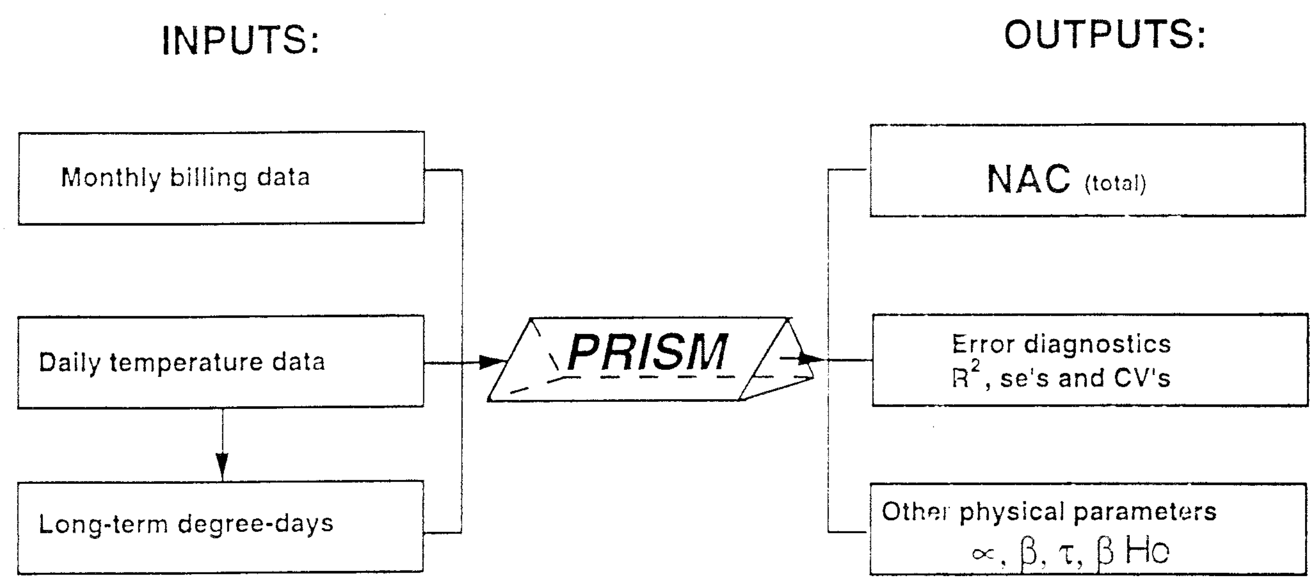


Figure 1. Schematic diagram showing the data requirements for PRISM (PRInceton Scorekeeping Method) and the estimates that result from it.

corresponding to the best linear fit of consumption data to heating degree-days (as determined by the R^2-statistic) is adopted by PRISM as the "best" reference temperature; the other parameters are then estimated with respect to that value. This produces a better statistical model of consumption. Furthermore, when the reference temperature is well determined, this procedure leads to a better decomposition of total heating fuel consumption into heating and non-heating components. The PRISM feature of variable reference temperature is particularly well suited to large buildings, for which the reference temperature is likely to be considerably below 65°F.

PRISM's physical foundation allows a physically meaningful interpretation of the results, and meaningful questioning of the results when they are physically unreasonable. Its statistical underpinnings ensure accurate error diagnostics for determining how good the fit of the data is and how much confidence to place on NAC, savings estimates, and the individual PRISM parameters. These error diagnostics are as important as the estimates themselves for determining the validity of the PRISM model.

PRISM's most important feature is the high reliability and stability of NAC. Experience from previous studies has shown consistently that NAC is generally much better determined than the individual parameters, and, in the extreme case, can be highly reliable even when the individual parameters are physically unreasonable and badly determined. For example, an overly high reference temperature, with huge standard errors, is likely to be accompanied by a lower-than-reasonable base-level estimate; when combined they often yield a well determined NAC estimate. This study presents numerous opportunities to test whether this feature of PRISM generalizes to multifamily buildings.

Variations of PRISM are being developed to model cooling as well as heating. The PRISM Heating-Only (HO) model is the original PRISM model and the one most widely used. The HO model is for application to heating fuels, such as natural gas or oil used for heating and other purposes, or electricity used for heating, lighting, etc., but not cooling. The PRISM model used throughout this report will be the three-parameter HO model.

Data bases for this study

Three sources of consumption data for multifamily buildings in New York City, all from Energy Authority-sponsored projects, were used for this study:

1) The Building Energy Use Tracking System (BEUTS) consists of 765 multifamily buildings for which reports of building characteristics and fuel consumption records were compiled. The BEUTS survey was conducted by the New York City Department of Housing Preservation and Development (HPD). Buildings with fairly complete consumption records were selected for this study.

2) The Building Monitoring Data (BMD) project consists of 30 buildings with data-recording energy management systems. The data set compiled for this study consisted of several years of oil delivery data supplied by the managing agent for the buildings, and sample detailed monitored data in computer-readable form from Fred Goldner of Energy Management Associates, who is carrying out a companion study of these buildings.

3) Energy Conservation Cases (ECC), from an HPD program in the mid-1980s that predates the BEUTS project, consists of 23 case studies of efficiently managed multifamily buildings. Data for seven ECC buildings were provided for this study.

The selection of buildings for these data sets and the results of the PRISM analyses of their consumption data are described in the next three sections (Sections 2-4). Wherever possible, lessons for future energy analysts of multifamily buildings are gleaned from the results; specific procedures are described (Section 5) and conclusions are summarized at the end of the report (Section 6).

2. BEUTS DATA BASE

Description of sample of buildings

The first data base, obtained from the New York City Department of Housing, Preservation and Development (HPD), consists of a subset of the 765 multifamily buildings that participated in the Building Energy Use Tracking System (BEUTS) project (Judd et al., 1989). Each building's BEUTS report contained data on the buildings' characteristics and reported fuel consumption.

Selected from these reports, the subsample for this study was designed to include a specifically large number of buildings to cover a wide range of data problems and possibilities. In general, buildings were selected to have energy billing or fuel oil delivery data spanning at least one year with corresponding meter reading dates or oil delivery dates specified, and preferably with some of the data in fairly short (monthly) increments. Buildings with less than adequate data were intentionally included as well. (A large fraction of the buildings had very sparse fuel consumption records, such as only a single annual total.) The subsample included gas-heated as well as oil-heated buildings.

The consumption data for the resulting subset of 71 buildings may be classified as follows:

- handwritten consumption data with meter reading or oil delivery dates (N=50), including the buildings that seem to have the most complete and accurate records, and selected cases with data apparently determined by truck limitations on amount of fuel delivered, rather than tank fill-ups;
- utility-supplied consumption data with meter reading dates (N=3);
- consumption data with no meter reading or oil delivery dates, but only the month specified (N=16); and
- data histories kept in a computerized (FASER) data base (N=2).

Our original subsample of 79 buildings was reduced to 71 because of illegible data or lack of data for the heating fuel. One case of a dual-fueled heating system, showing comparable amounts of oil and gas used over a year, could not be used: inconsistent consumption periods meant that the sum of the fuels were not available.

The consumption data compiled from the BEUTS data base for the resulting 71 buildings were then entered in computer-ready PRISM format for further analysis. In most cases, one to two years of energy data were recorded. In addition, a data base of daily temperatures for 1970-1991 from the New York City National Weather Station was prepared (NOAA, 1970-1991).

Example from the BEUTS data

A sample of the original consumption data for one oil-heated building (Building #343) and the corresponding PRISM consumption file are shown in Figures 2a and 2b, respectively. Running these data through the PRISM Heating-Only (HO) model gave the plots and results shown in Figure 3. Clearly, the heating consumption follows closely the heating degree-days for corresponding periods; the resulting R^2 = 0.833 indicates that 83% of the month-to-month variability is explained by outside temperature. The Normalized Annual Consumption, or NAC, as the estimate of the amount of oil this building would consume under average weather conditions, is well determined: NAC = 46,080 (± 3,380) gallons/year, i.e., the relative standard error of NAC, or CV(NAC), is only 7.3% of the estimate.* PRISM also indicates that 71% of the total consumption is for space heating. Apparently, this building uses oil for domestic hot water heating as well as space heating; this was confirmed in the BEUTS building description data sheet for this building.

*Definitions: NAC = Normalized Annual Consumption; CV(NAC) = [standard error(se) of NAC]/NAC, or, equivalently, the relative standard error of NAC. CV(NAC) is written alternatively as a ratio (e.g., 0.065) or as a percent (6.5%).

Question 10 (Fuel usage)

a)

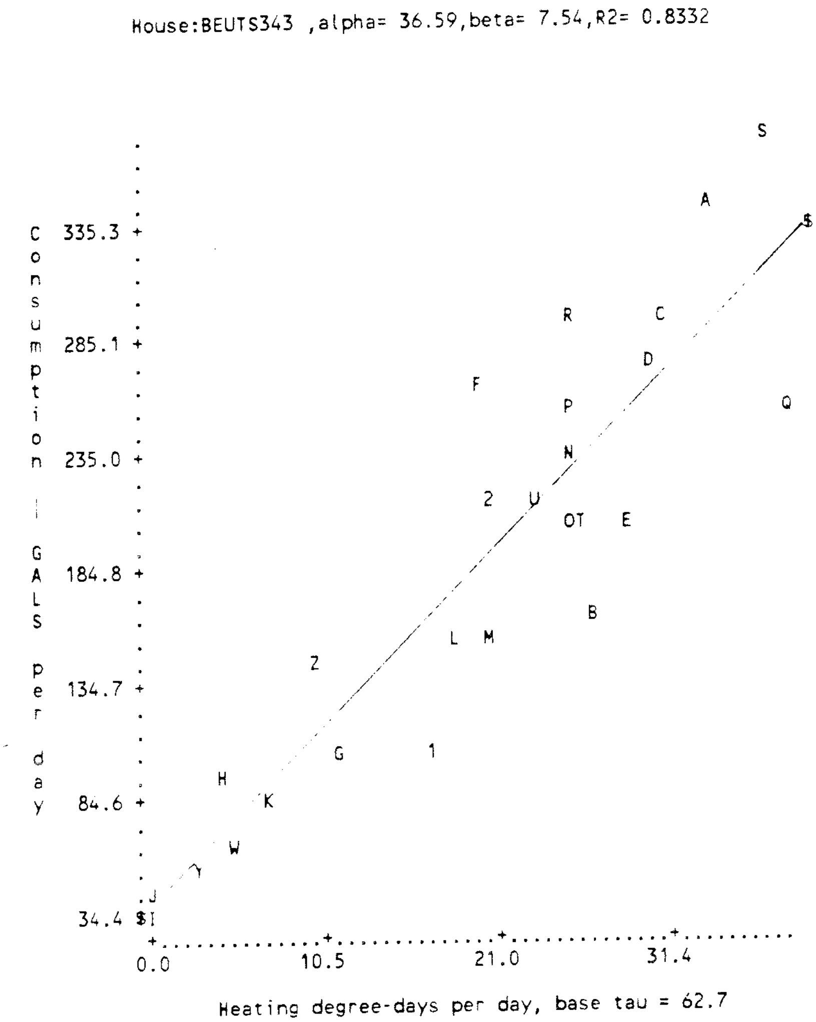
YEAR 1 DELIVERY DATE	# of Gallons/ Therms/Mlbs
1/8/87	3,038
1/20/87	3,010
1/30/87	2,508
2/9/87	2,913
2/17/87	3,003
3/4/87	3,005
3/18/87	3014
4/15/87	3000
6/4/87	3,001
8/26/87	3,000
10/19/87	3,000
11/9/87	3,000
12/2/87	2,501
12/16/87	3,001
12/31/87	3,001

YEAR 2 DELIVERY DATE	# of Gallons/ Therms/Mlbs
1/6/86	2,976
1/16/86	3,472
2/3/86	3,000
2/13/86	2,976
2/24/86	2,976
3/8/86	2,480
3/19/86	2,950
4/17/86	2,976
5/20/86	3,000
8/1/86	2,513
10/1/86	3,001
11/7/86	3,031
11/26/86	3,003
12/12/86	2,500
12/24/86	2,860

b)

```
BEUTS343        1 06 86    U GALS
*1929/1947
*PRE-WAR
*39 APTS
*6 FLOORS
*STEAM
*#6 OIL
3472            1 16 86
3000            2 03 86
2976            2 13 86
2976            2 24 86
2480            3 08 86
2950            3 19 86
2976            4 17 86
3000            5 20 86
2513            8 01 86
3001           10 01 86
3031           11 07 86
3003           11 26 86
2500           12 12 86
2860           12 24 86
3038            1 08 87
3010            1 20 87
2508            1 30 87
2913            2 09 87
3003            2 17 87
3005            3 04 87
3014            3 18 87
3000            4 15 87
3001            6 04 87
3000            8 26 87
3000           10 19 87
3000           11 09 87
2501           12 02 87
3001           12 16 87
3001           12 31 87
-9
```

Figure 2. a) Sample of original consumption data for Building #343 from the BEUTS data base; b) corresponding PRISM "Meter" file for the same data.



PERIODS:

A	JAN 6,1986 to JAN 15,1986
B	JAN 16,1986 to FEB 2,1986
C	FEB 3,1986 to FEB 12,1986
D	FEB 13,1986 to FEB 23,1986
E	FEB 24,1986 to MAR 7,1986
F	MAR 8,1986 to MAR 18,1986
G	MAR 19,1986 to APR 16,1986
H	APR 17,1986 to MAY 19,1986
I	MAY 20,1986 to JUL 31,1986
J	AUG 1,1986 to SEP 30,1986
K	OCT 1,1986 to NOV 6,1986
L	NOV 7,1986 to NOV 25,1986
M	NOV 26,1986 to DEC 11,1986
N	DEC 12,1986 to DEC 23,1986
O	DEC 24,1986 to JAN 7,1987
P	JAN 8,1987 to JAN 19,1987
Q	JAN 20,1987 to JAN 29,1987
R	JAN 30,1987 to FEB 8,1987
S	FEB 9,1987 to FEB 16,1987
T	FEB 17,1987 to MAR 3,1987
U	MAR 4,1987 to MAR 17,1987
G	MAR 18,1987 to APR 14,1987
W	APR 15,1987 to JUN 3,1987
I	JUN 4,1987 to AUG 25,1987
Y	AUG 26,1987 to OCT 18,1987
Z	OCT 19,1987 to NOV 8,1987
1	NOV 9,1987 to DEC 1,1987
2	DEC 2,1987 to DEC 15,1987
O	DEC 16,1987 to DEC 30,1987

Figure 3. PRISM plot of consumption vs. heating degree-days for BEUTS Building #343. Line drawn represents PRISM fit of the data.

Note that the oil deliveries for this building are generally close to 3,000 gallons, but they are spaced unevenly, as shown in Figure 2a. Figure 2b shows the wide variation in per-day consumption. This is the opposite of what we see for gas-heated buildings, which have evenly spaced consumption periods (e.g., monthly) and widely varying consumption quantities across periods. It is reassuring that the PRISM results are quite reliable for this oil-heated building, as is commonly the case for gas-heated buildings (Fels et al., 1986).

PRISM analysis of the BEUTS data

PRISM was run on the data for all 71 buildings for an initial assessment of the quality of the data and PRISM's reliability as applied to the data set. An objective was to determine whether "bad" model results were the consequence of inaccurate and/or inadequate data, and to find possible approaches to improvements in the data and in the model results. Thus, the usefulness of PRISM as a data screener was explored as part of this project.

Two indicators of the goodness of fit from PRISM are CV(NAC), the coefficient of variation of NAC, which ideally is very small, and the model's R^2-statistic, whose closeness to 1.0 (its maximum value) measures the extent to which consumption correlates linearly with degree-days (computed to the "best" reference temperature determined by PRISM). Figure 4 shows a plot of CV(NAC) vs. R^2 for the initial PRISM runs. Results for the 71 buildings are shown. Superimposed on the plot are the quartiles of the distribution, showing that 50% have CV(NAC) less than 0.09 and 50% have R^2 above 0.75.*

*The quartile distribution shows the interquartile range, bounded by the values below and above which 25% of the cases fall, and the median value, above and below which 50% of the cases fall. Median, rather than mean, values are emphasized in this study because of the median's relative insensitivity to unusually large or small values.

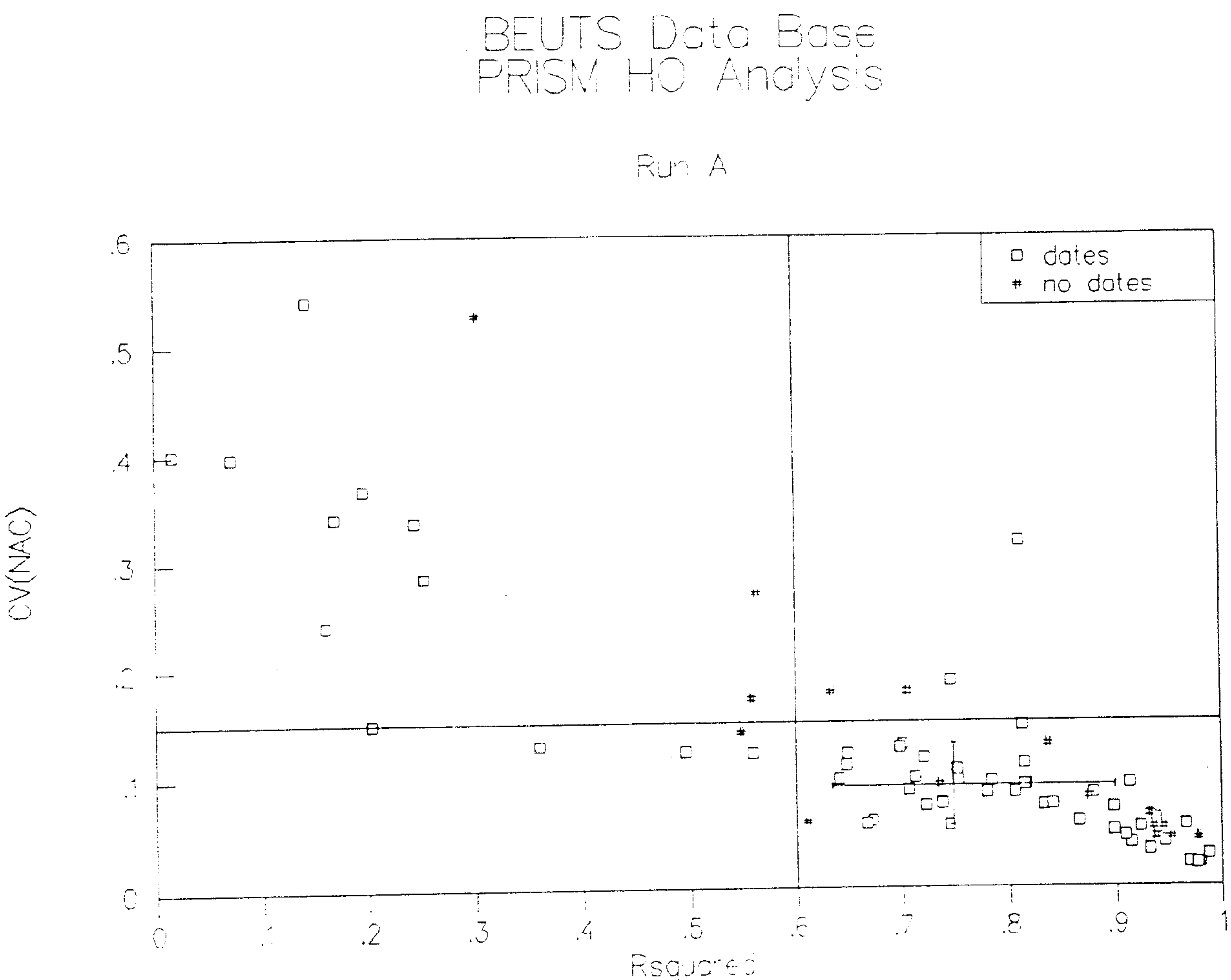


Figure 4. CV(NAC) (percent standard error of NAC) vs. R^2 for BEUTS sample of 71 buildings -- Run A using original data. Symbols distinguish those buildings that had meter reading or delivery dates from those without. For the latter, assumed dates were used. Superimposed on the plot are the quartile distributions of CV(NAC) and R^2.

The lower right-hand corner of Figure 4, representing high-R^2 and low CV(NAC) cases, is well populated, but nevertheless there are numerous cases for which the model results are not reliable. In earlier work on detached single-family houses, reliability criteria of $R^2 \geq 0.7$ and CV(NAC) ≤ 0.06 were adopted (Reynolds and Fels, 1988). By these criteria, only 18 (25%) of the 71 cases would be deemed reliable. Since it is reasonable to expect that oil-heated multifamily buildings will not model as well as single-family gas-heated buildings, these criteria may be overly stringent for our BEUTS subsample. From the way the results cluster in the plot, we decided on criteria of $R^2 \geq 0.6$ and CV(NAC) ≤ 0.15 to determine those buildings warranting more detailed analysis; these reliability cutoffs are indicated in the plot. The first-run PRISM results (called "Run A") may thus be summarized as follows:

- 51 (72%) of the 71 buildings have reliable or fairly reliable PRISM results ($R^2 \geq 0.6$ and CV(NAC) ≤ 0.15);
- 20 (18%) of the buildings have "unreliable" PRISM results ($R^2 < 0.6$ and/or CV(NAC) > 0.15).

The next step was to take a closer look at the original consumption data for the "problem" cases failing to meet the criteria defined by these cutoffs. Discrete types of data problems became evident for this data set:

(a) buildings with no meter reading (or, more commonly, oil delivery) dates specified (three cases; the remaining 13 with no delivery dates modeled well, using a procedure of assumed dates described below);

(b) buildings that appeared to have deliveries related to size of delivery truck, and thus with tank not filled (seven cases);

(c) data with at least one outlier (14 cases), as discussed below.

Note that many of the buildings fit into more than one problem category, so that the numbers add to more than 20. In particular, a number of the buildings with problem

(c) demonstrated problem (b) as well, confirming that the outlying data represent non-fill-ups for oil deliveries. Different types of improvements, described in the next section, were explored for each of these problem types.

Problem cases

Buildings with no meter reading (or delivery) dates

For the 16 buildings (out of the sample of 71) with undated consumption periods, including 11 with no oil delivery dates and five with no gas meter-reading dates, a date for each consumption period needed to be assumed for the PRISM run. For the results indicated by a pound (#) sign in Figure 4, the 15th of the month was assumed for cases with a maximum of one delivery per month, and the 30th of the month for cases with more than one delivery in one or more months (in which cases, multiple deliveries in a month were combined for a single consumption data point, as if there were a single large delivery at the end of the month). As indicated, 10 out of the 16 buildings in this category modeled well under these assumptions.

The next step was to vary the assumed date (using the 1st through the 31st of the month as alternatives), to see which assumption yielded the best model fit. In addition, multi-year data sets were broken down into single years, and in general this gave further improvements. This procedure was applied to all 16 buildings.

The validity of assumed dates was tested using a sample of 21 buildings that had actual delivery or meter reading dates; 15 of them were heated by oil, and six by gas. A data set was created that masked the actual dates for these buildings, and a PRISM analysis was run assuming that oil deliveries were made (or, for gas, meters were read) on the same day of each month. In all, 31 PRISM analyses were run on each building corresponding to the possible days of the month. The results of the 31 runs were then compared against the "correct" PRISM runs using the actual dates to determine suitable criteria for selecting the best set of assumed dates.

Although the problem of missing dates affects oil-heated buildings much more than gas-heated buildings, we included both types of buildings in the analyses. In particular, inclusion of gas-heated buildings with uniformly spaced monthly meter readings gave us an opportunity for a precise comparison of the correct date with the "best" assumed date for which the PRISM R^2 was highest. The objective was to answer whether R^2 and/or CV(NAC), the two most important PRISM reliability statistics, represent useful criteria for determining which is the best assumed date, and to assess to what extent the NAC from assumed dates closely matches the "real" NAC.

Table 1 gives a detailed snapshot of the PRISM results. "Actual" PRISM estimates, from consumption data with known reading or delivery dates, are compared with "masked" results, in which assumed dates are used. From the 31 runs for each building, the highest R^2 and lowest R^2 values are shown. For each building, the set of dates for which R^2 is highest is adopted as the best results from assumed dates; these are labeled "masked" in the table, to compare with "actual" representing those obtained from the actual meter reading or delivery dates. Also given in the table are the "worst" results, corresponding to the dates for which R^2 was lowest. In all cases, the best R^2 was far superior to the worst R^2; the results are clearly sensitive to choice of date.

In a number of cases, when reliable PRISM results are obtainable from the actual data, the results from the best set of assumed (masked) dates are reliable as well. Comparing the reliability of NAC, CV(NAC) from masked dates is less than 20% different than CV(NAC) from actual dates for all six gas-heated buildings but for only five of the 15 oil-heated buildings. That the reliability is similar for gas data is not surprising, since gas meter readings are generally spaced about one month apart. As oil deliveries are irregular and do not always correspond to a fill up, masking dates in several cases usually lowers the reliability, except for a few cases where it increases.

Table 1. Comparison of PRISM results from "actual" and "masked" (assumed) meter reading or delivery dates.

HOUSE ID	FUEL	RSQUARED (actual)	RSQUARED (masked)	RSQUARED (worst)	NAC (actual)	NAC (masked)	(a) se(NAC) (actual)	se(NAC) (masked)	CV(NAC) (actual)	CV(NAC) (masked)	(b) NAC(act) -NAC(mas)	(b)/ NAC(act)	(b)/(a) RATIO	DEL'Vs MULT 5002	N (act)	N (mas)
122	GAS	0.817	0.764	0.346	6506	6876	591	589	.091	.086	-370	-5.7%	-.626	N	7	7
280	GAS	0.980	0.968	0.852	4824	4833	95	125	.020	.026	-9	-.2%	-.095	N	30	30
281	GAS	0.970	0.965	0.830	3072	3089	64	71	.021	.023	-17	-.6%	-.266	N	34	34
287	GAS	0.988	0.992	0.750	3571	3592	100	85	.028	.024	-21	-.6%	-.210	N	11	11
357	GAS	0.965	0.969	0.719	24357	24352	1322	1214	.054	.050	5	.0%	.004	N	8	8
358	GAS	0.728	0.749	0.651	7123	7144	538	483	.076	.068	-21	-.3%	-.039	N	13	13
165	OIL	0.898	0.958	0.817	18849	17722	1321	824	.070	.046	1127	6.0%	.853-	N	15	10
238	OIL	0.829	0.835	0.538	17441	16896	1599	1724	.092	.102	545	3.1%	.341	Y	14	9
239	OIL	0.701	0.758	0.577	23966	23334	3089	3147	.129	.135	632	2.6%	.205	Y	12	9
275	OIL	0.898	0.797	0.734	2857	2890	142	229	.050	.079	-33	-1.2%	-.232	N	33	18
278	OIL	0.723	0.480	0.358	46595	41863	3420	4797	.073	.115	4732	10.2%	1.384	Y	27	19
290	OIL	0.932	0.790	0.543	9747	9380	316	702	.032	.075	367	3.8%	1.161	N	32	18
323	OIL	0.832	0.949	0.733	98068	91805	11037	5987	.113	.065	6263	6.4%	.567	Y	15	9
341	OIL	0.914	0.565	0.397	22966	23272	872	2314	.038	.099	-306	-1.3%	-.351	Y	22	18
342	OIL	0.909	0.576	0.435	21774	22164	966	2446	.044	.110	-390	-1.8%	-.404	Y	22	18
343	OIL	0.833	0.850	0.597	46076	46704	3381	3469	.073	.074	-628	-1.4%	-.186	Y	29	19
347	OIL	0.923	0.702	0.525	17694	17881	927	2596	.052	.145	-187	-1.1%	-.202	Y	22	16
362	OIL	0.707	0.517	0.201	46534	38913	4078	6733	.088	.173	7621	16.4%	1.869	Y	15	10
453	OIL	0.913	0.734	0.555	10840	11171	1007	1456	.093	.130	-331	-3.1%	-.329	N	14	13
695	OIL	0.745	0.747	0.542	46117	44361	2508	2899	.054	.065	1756	3.8%	.700	N	46	22
777	OIL	0.753	0.812	0.618	41217	39363	3953	3304	.096	.084	1854	4.5%	.469	Y	37	19

See text for explanation.

Building 323, which shows an improvement of CV(NAC) from 11.3% for the actual run to 6.5% for the masked run, is an interesting example. Some of the deliveries are in multiples of 500, suggesting that the delivery amount was determined by the truck's tank size rather than by the building's tank size. The NACs from the two PRISM runs differ by 6.4%, and thus by about one standard error of NAC. In this case, masking the "delivery" dates caused an improvement in the consumption data, similar to that seen from combining gas or electricity data for two consecutive periods when an estimated meter reading between the two periods is indicated. (Note that the R^2 improvement from masking this building's data is due, at least in part, to the decreased number of data points in the masked run.)

As shown in Figure 5, the NAC estimates from the actual vs. masked data are quite close. For the gas data, only one out of five differ by more than 1%, and that one differs by 6%, which is less than its CV(NAC) of 9%. For the oil data, only two of the 15 have NACs that differ by more than 6%. The building (#362) with the largest difference in NAC was also a building with reported deliveries in multiples of 500; in this case masking the dates caused a deterioration in the reliability of NAC (with an increase of CV(NAC) from 9% to 17%).

Selection of the highest R^2 appears to be a useful criterion, particularly since the R^2 values from different sets of assumed dates vary considerably. For the 15 oil-heated buildings, the median R^2 for the "best" assumed dates is 0.76, which is fairly close to the median R^2 of 0.83 from the runs on the actual dates, whereas the median R^2 for the "worst" assumed dates is only 0.54. In general, for the 31 sets of assumed dates, the lowest R^2 is considerably below the highest R^2.

These analyses indicate that, by assuming delivery dates, useful PRISM results can be obtained for buildings with monthly but undated delivery data. Finding the date of the month that gives the highest R^2 in most cases produces an NAC estimate

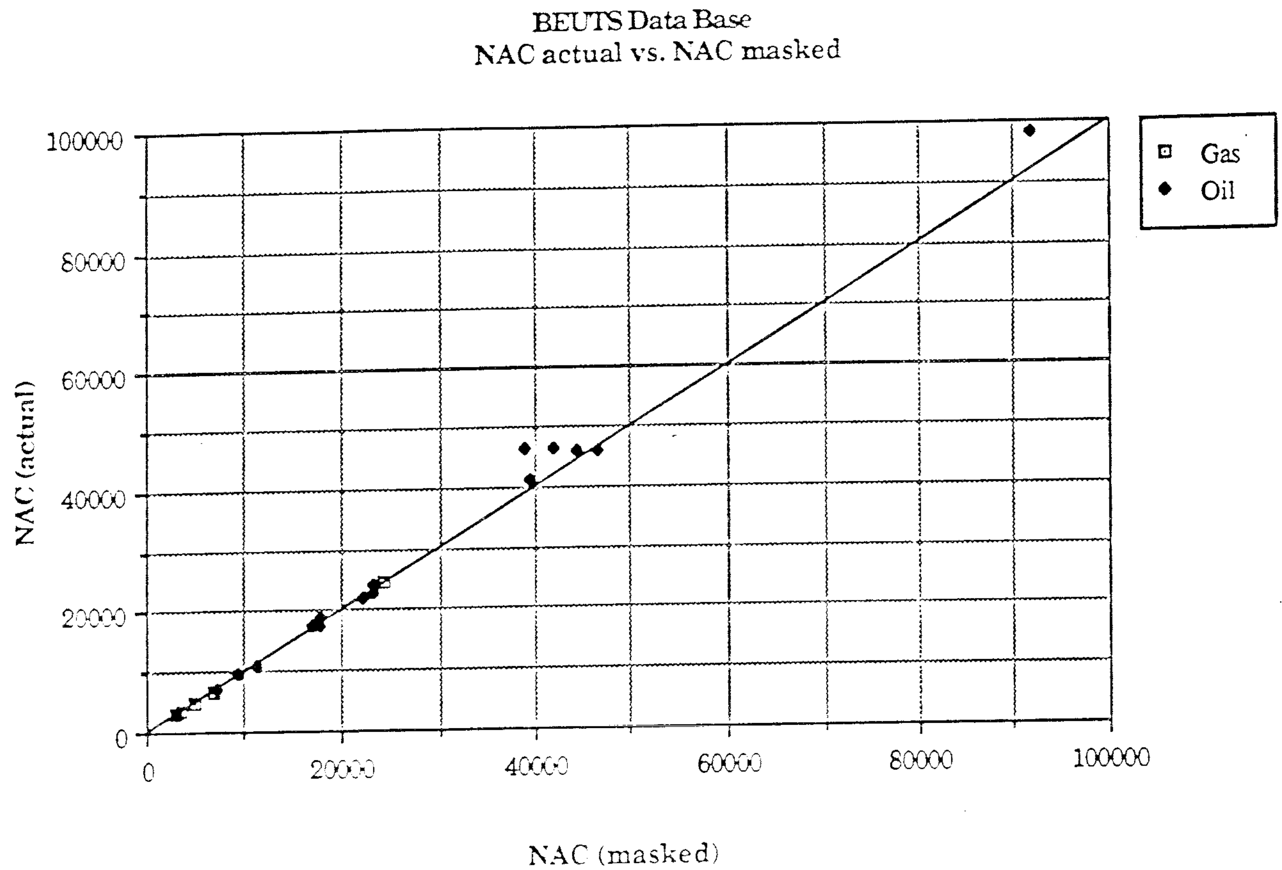


Figure 5. NAC estimates for "actual" PRISM runs vs. NAC for the best "masked" run (based on R^2).

close to the best value (i.e., to the value that would be obtained if the actual dates were known), and often maintains the reliability of NAC. This generally appears to be a promising procedure for retrieving useful information from monthly delivery data for which delivery dates are missing and could greatly expand the usefulness of the entire BEUTS data base, for which delivery dates are missing for a large fraction of the oil-heated buildings.

Buildings with deliveries appearing to be related to truck size

A number of the buildings had "consumption" data that included repeated values, e.g., deliveries of 2000 gallons of oil in several months. Possible explanations for this are numerous. The fuel oil deliverer's truck may be smaller than the oil tank in the building, so that the deliveries represent the size of the truck tank rather than a complete fill for the building's tank. The repeated data could be a request by the building owner or manager to put in only a fixed, budgeted amount of fuel. Or, the building owner might monitor the oil level in the tank and request a fill-up when the tank is near empty, in which case the delivery would represent consumption. Each of these explanations would manifest itself differently in the extent to which the data can be modeled. Therefore, one would not necessarily expect reliable PRISM fits of such data. Nevertheless, reasonable results could often be obtained, in some cases by identifying and treating the non-fill delivery as an outlier. The example, Building #360, used in the discussion of the next problem type illustrates this well.

Buildings with at least one outlying data point

For a number of the buildings, a single outlier in the consumption data was evident. When we looked more closely at the original data, a phenomenon unique to oil data appeared, affecting eight out of the 14 cases with outliers. A single, high outlying point covered a delivery of a very short period: four days or less, or as short as one day in some cases. The original data shown in Figure 6 for Building #360 illustrates a one-day delivery, on February 2, 1987, of an even 3,000 gallons. Apparently, a partial fill was done on the day before and the oil truck returned the following day to complete or add to the tank fill, giving an anomalously high outlier. In a few other buildings, anomalously low outliers were also seen.

The first PRISM run on these data attributed all of the oil consumption in the one-day delivery to a one-day period. As seen in the plot of consumption vs. degree-days in Figure 7a, this produces a huge outlier that has a disproportionately large effect on the PRISM fit. Combining this one-day period with the previous period gives a more reasonable consumption vs. degree-day plot (Figure 7b), and a shift in the PRISM fit from highly unreliable to highly reliable (Table 2). This example illustrates the striking improvement that can result from a careful examination of the original data and, in particular, from a judicious combination of two consecutive data points.

Table 2. Comparison of PRISM results before and after data combination: Building #360 as example of single outlier

	R^2	CV(NAC)	Ref. temp. τ ($\pm$se)
Before Combining	0.14	0.54	53($\pm$22)°F
After Combining	0.83	0.07	66($\pm$ 6)°F.

Testing for outliers

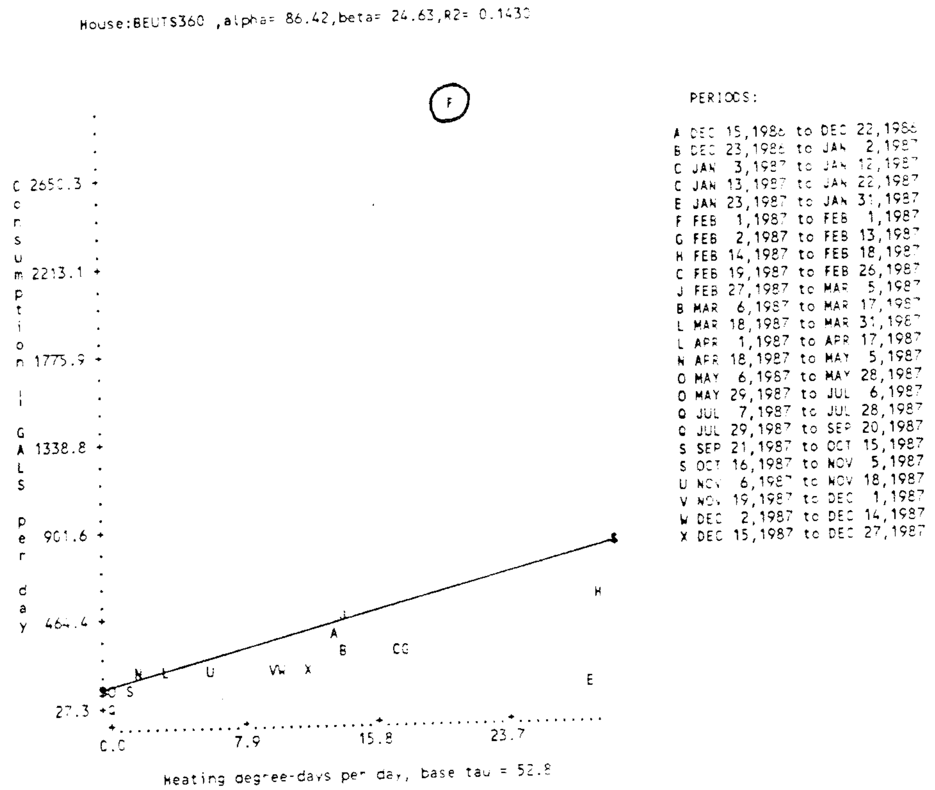
PRISM analyses that lead to poor results may cause higher-than-necessary attrition of the sample size if the analyst does not have the time or the facilities to take a closer look at the cause of these poor results. In the past, the only way to

DELIVERY DATE	# of Gallons/ Therms/Mlbs
YEAR 1	
12/15/86	2,829—
12/23/86	3,000—
CONT. OF YEAR 2	
9/21/87	3,048—
10/16/87	3,079—
11/6/87	3,100—
11/19/87	3,000—
12/2/87	3,140—
12/14/87	3,140—
12/28/87	3,100—

DELIVERY DATE	# of Gallons/ Therms/Mlbs
YEAR 2	
1/3/87	2,900—
1/13/87	3,000—
1/23/87	2,707—
2/1/87	1,000—
2/2/87	3,000—
2/11/87	3,006—
2/19/87	2,725—
2/27/87	2,538—
3/6/87	3,000—
3/18/87	3,001—
4/1/87	3,100—
4/18/87	3,000—
5/6/87	3,150—
5/29/87	3,020—
7/7/87	2,980—
7/29/87	600—

Figure 6. Sample of original consumption data for BEUTS Building #360 (illustrating a one-day delivery on February 2, 1987).

a)



b)

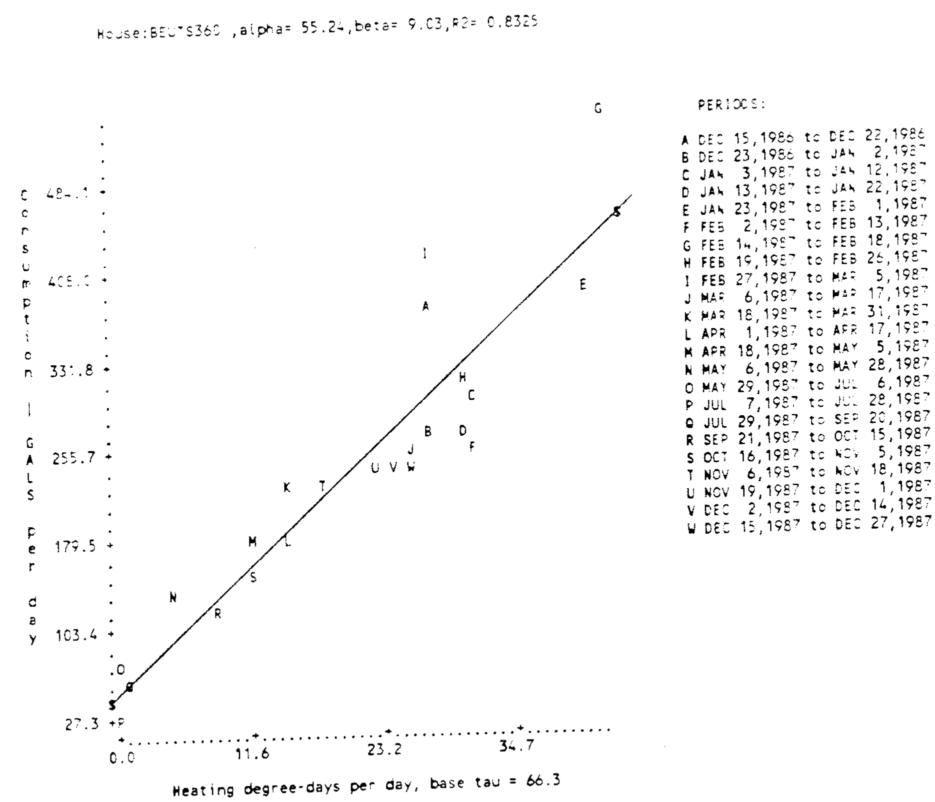


Figure 7. a) PRISM plot of consumption vs. heating degree-days for BEUTS Building #360, demonstrating one-day outlier (point F); b) consumption vs. heating degree-days for same building with outlier (point F) combined with previous day's delivery (point E).

determine the cause of a problem with data from the PRISM analyses was to look at the individual-building plots of consumption vs. period or heating degree-days. In earlier work, dramatic PRISM improvements from combination of two consecutive data points have been reported for gas-heated houses, in which an unspecified estimated reading entered as an actual reading can lead to a high outlier adjacent (in time) to a low outlier (Fels and Reynolds, 1990; Reynolds et al., 1990). For oil data, unevenly spaced deliveries lead to more complicated, and more prevalent, outlier effects, partly because of the possibility of very short consumption periods and also because of the potentially important connection between outliers and incomplete tank fills.

Although an outlier in consumption data can be "real", reflecting anomalous behavior in that consumption period, the possible undue influence of a single outlier and the likelihood that it results from a data problem has led us to explore more quantitative ways of detecting data errors and outliers. The use of studentized residuals, one such method, is briefly described here and explained in detail in the Appendix. The method is promising, both in terms of our long-range goal to develop automated procedures of data screening and error checking, and for the immediate objective of improving the PRISM analyses of oil-heated buildings.

The externally studentized residual, or "R-Student statistic", is a special rescaled version of the raw residual (with mean 0 and variance 1), wherein residuals are in effect computed relative to the model's fit of the data in the absence of the suspected outlier. Consider, for example, a data set with a single outlier, and, after removal of that outlier, a linear regression of the remaining (N-1) data points. In the original PRISM fit of the N data points, the single outlier can have an unduly large influence, pulling the PRISM fit near or through that point. The resulting residual for the outlier may be no larger than the other residuals in the data set, making the original set of residuals not useful for outlier detection. On the other hand, relative to a "corrected" linear regression of N-1 data points (with the outlier excluded), the outlying data point's residual, i.e., the difference between actual and predicted values, may be expected to be very large relative to the residuals of the other data points. This is the rationale behind the studentized-residual procedure.

Roughly speaking, a data point may be considered an outlier if its studentized residual lies outside the 95% confidence interval, which for a PRISM fit of 12 data points corresponds to a studentized residual of magnitude greater than about 2.0.* To detect meter reading errors in natural gas data, one simply looks for consecutive data points with studentized residuals that are of opposite sign and of magnitude greater than 2, i.e., outliers in opposite directions. For oil data, the studentized residual may be useful not only for identifying high/low consecutive data points, but also for detecting "spikes", resulting from oil deliveries made one or a few days apart.

In the preceding example, as shown in Figure 7a and Table 2, the studentized residual was 27.8 for the outlier in the original data, compared with a maximum magnitude of 1.4 for all others for that building, and only 1.8 in magnitude for the outlier when combined with the previous data point. This clearly illustrates the unambiguous outlier detection provided by the studentized residual and the reliability improvement in the PRISM fit that can result from data combination for that outlier (Table 2). Studentized residuals have been computed for the problem cases studied thus far, and their usefulness as a detector of consumption data outliers for PRISM -- being explored in this study for the first time -- looks encouraging. Other examples of the usefulness of studentized residuals will be presented throughout this report.

---

*In some cases, a cutoff of 2.0 may be too stringent for detecting high/low readings. See the Appendix for guidelines on selection of cutoff criteria.

Summary of data improvements

Using these tools for outlier detection, improvements to the BEUTS data set were made. These tools were not only applied to the 20 unreliable cases, but also to selected buildings that met the reliability criteria but that nevertheless showed the possibility of further improvement.

The results of these analyses are shown in the plot of CV(NAC) vs. R^2, called "Run B", in Figure 8. Comparison of this plot with Run A in Figure 4 and comparison of the quartiles indicated in each plot show clearly the substantial improvement in PRISM results. The median R^2 improves from 0.75 in Run A to 0.84 in Run B and the median CV(NAC) improves from 0.09 to 0.07. Using the same cutoff criteria of $R^2 \geq 0.6$ and CV(NAC) ≤ 0.15, the following breakdown of reliable vs. unreliable fits results:

- 61 reliable ($R^2 \geq 0.6$ and CV(NAC) ≤ 0.15);
- 10 unreliable ($R^2 < 0.6$ and/or (CV(NAC) > 0.15).

Therefore, the percent of reliable cases has increased considerably, from 72 to 86%, and a number of them shifted from very unreliable to very reliable cases. The improvement is much more pronounced for the oil data than for the gas data: the median R^2 improved from 0.74 to 0.82 for the 57 oil-heated buildings, and from 0.95 to 0.98 for the 14 gas-heated buildings. Apparently the additional analysis for oil data is worthwhile.

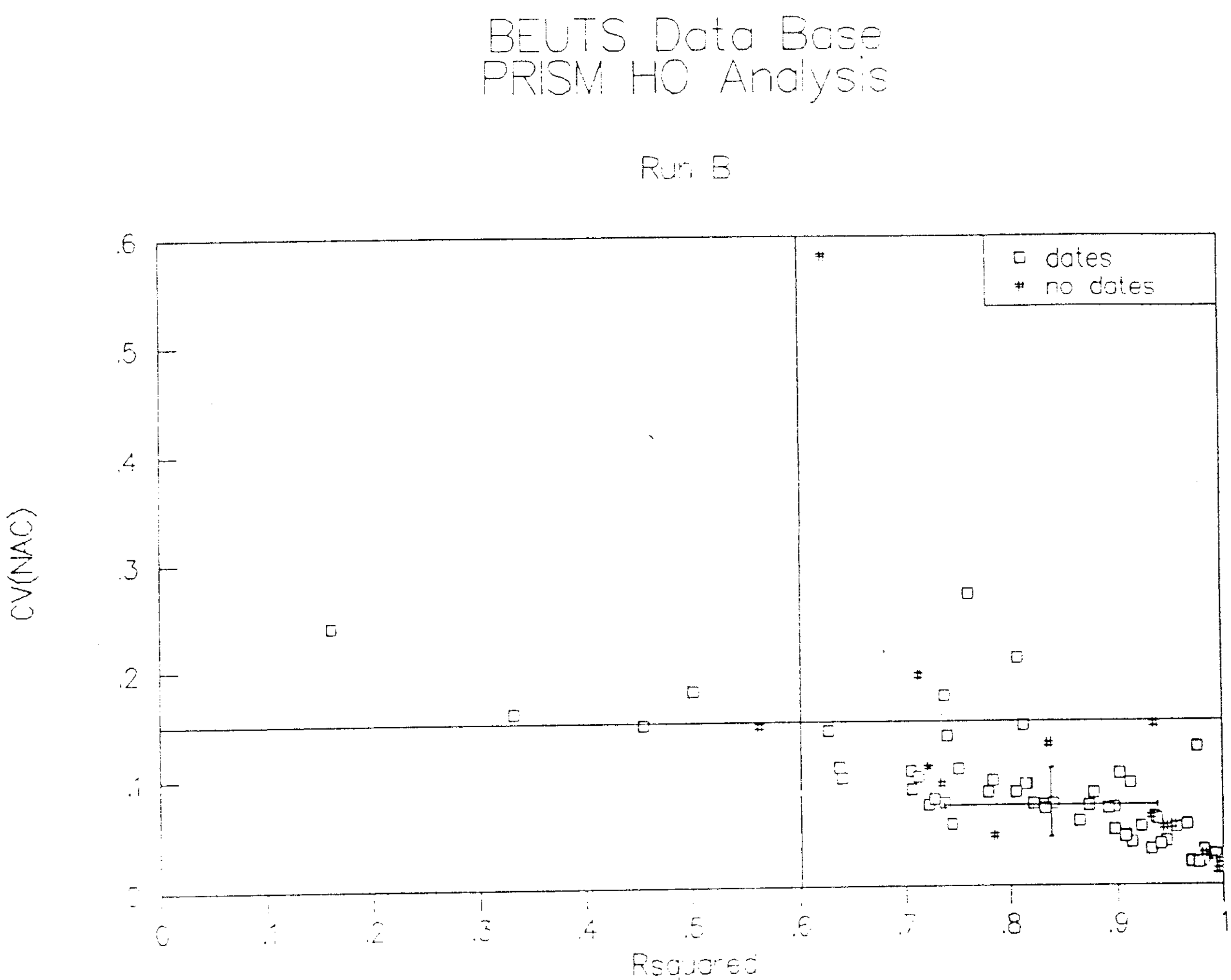


Figure 8. CV(NAC) vs. R^2 for BEUTS sample of 71 buildings -- Run B -- to show improvement in PRISM estimates as a result of improvements to data quality. See Figure 4.

3. BUILDING MONITORING DATA PROJECT

Description of data set

The second data set for this project was from the thirty oil-heated multifamily buildings participating in the Building Monitoring Data (BMD) project, an Energy Authority project generating data currently being compiled by Fred Goldner of Energy Management Associates. This data set was ideally suited to the needs of this study, both because of an expected high level of data quality (the buildings are well managed under one company and operate on similar energy management systems), and because of the availability of extremely detailed monitoring data in addition to the building-level consumption data analogous to our BEUTS data set (Goldner, 1991). Specifically, there were three parts to the data set:

- oil delivery data (amounts and dates of delivery) for the 30 buildings, received initially for the one-year period from May 1990 through April 1991, and later for a multi-year period spanning most of 1986 through 1991 (computer printout of the data was received from the management company);
- detailed daily and sometimes hourly energy consumption, furnace runtime, and temperature data measured and recorded by the energy management systems installed in nine of the buildings (compiled by Fred Goldner for a nine-building subset of the 30 buildings)*; and
- vacancy-rate data for nine of the buildings (the same as the previous subset) compiled from detailed rental records from the management company.

Analysis of each of these data sets is described below.

PRISM analysis of BMD data

A detailed PRISM analysis was performed on the most recent year of oil delivery data provided, May 1990 through April 1991, for all 30 buildings. As in the BEUTS

*The detailed data set is described in Progress Report #3 (for June-July, 1991) for this project. The consumption data are emphasized in this report.

analyses, a plot of CV(NAC) (percent standard error of NAC) vs. R^2 was used to determine quality of fit of the 30 buildings using the PRISM model. Figure 9a shows the preliminary results (Run A).

A closer look at the raw data for those buildings that did not model well, particularly those with low R^2 and/or high CV(NAC) values, revealed that there appear to be many instances with obvious outliers in the oil data, as there were in the BEUTS analyses. One source of these outliers appears to be one-day deliveries, or instances in which a partial delivery was made on one day and the deliverer returned the following day to fill the rest of the oil tank. These outliers are treated by combining the values with the previous day's consumption. Other paired outliers, e.g., from deliveries spaced two or three days apart, were treated in a similar manner, and PRISM was run on the combined data. The resulting plot of reliability criteria appears in Figure 9b (Run B). One can see a substantial improvement in the model fits as the results move closer to the bottom right corner of the plot, corresponding to high R^2 and low CV(NAC).

The improvement is evident in the example shown in Table 3 which gives summary PRISM results for Run A and Run B for building BMD13. Figures 10ab show the corresponding consumption vs. heating degree-day plots. The results once again demonstrate the stability of the NAC estimate even though the other parameters remain less reliable. In Figure 10a, there is an obvious high/low pair of deliveries, points G and H; the studentized residuals for this high/low pair were -2.3 and 2.5, respectively. (The other studentized residuals for these data were 0.6 or lower in magnitude.)* When these two data points are combined, the improvement in R^2, from 0.42 to 0.90, and the increased reliability of NAC, from a CV(NAC) of 0.22 to 0.07, are striking. This confirms that the studentized residuals analysis provides quantitative

*The Appendix on studentized residuals contains additional discussion of this example.

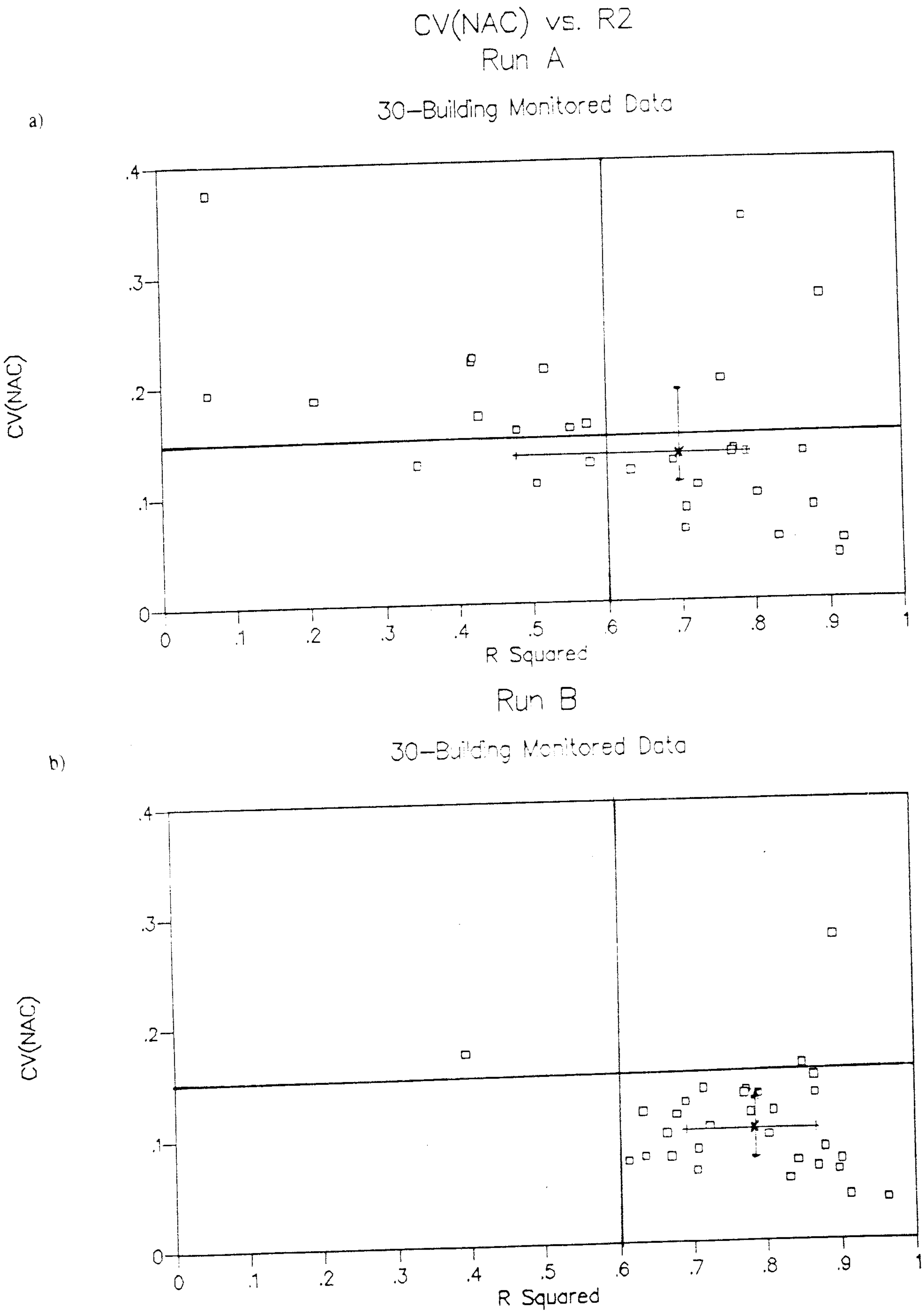
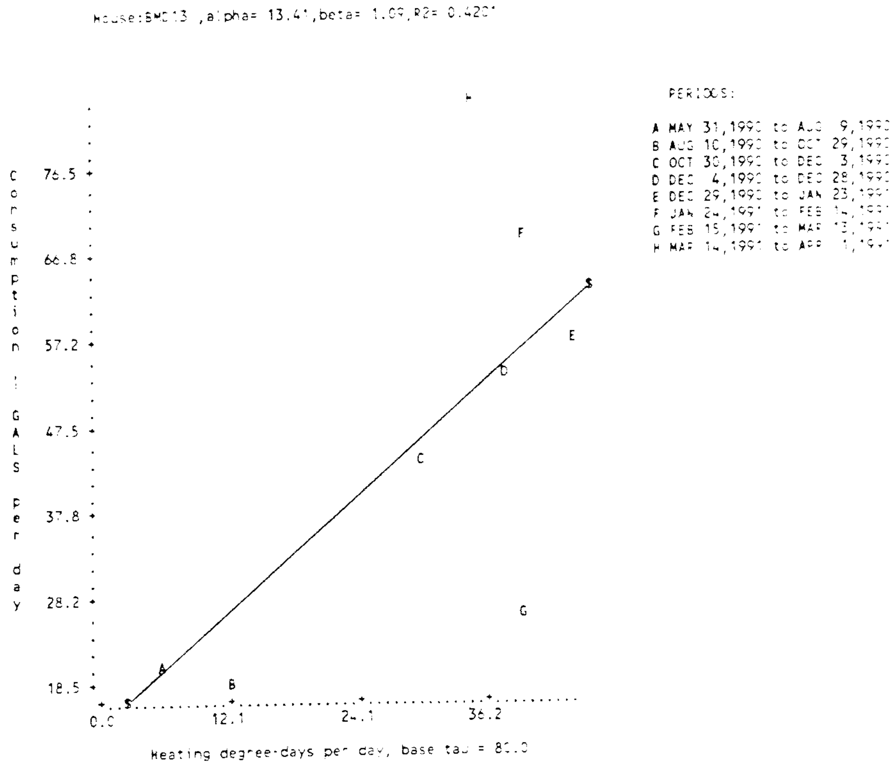


Figure 9. Plot of CV(NAC) vs. R^2 for 30 buildings from Building Monitoring Data (BMD) project for a) Run A (original data); b) Run B (after data improvements).

a)



b)

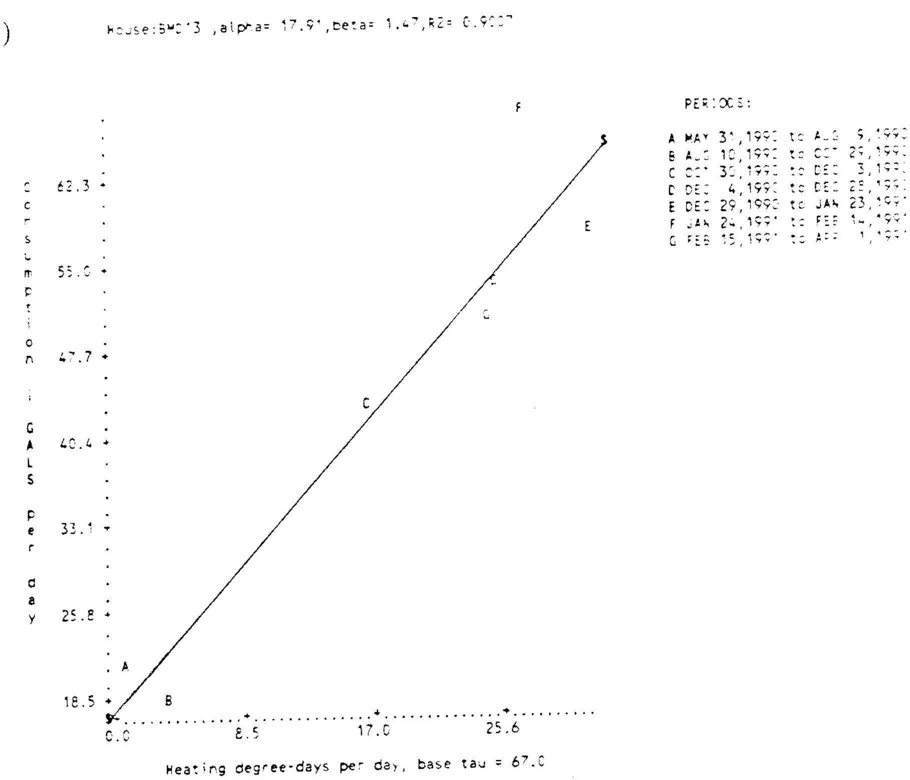


Figure 10 PRISM plot of consumption vs. heating degree-days (Building BMD13) for a) original data; b) after combining high/low pair (points G and H).

guidance for major improvements in the PRISM fits.

Table 3. Building BMD13: Example of improvements in PRISM results due to combination of high/low outliers.

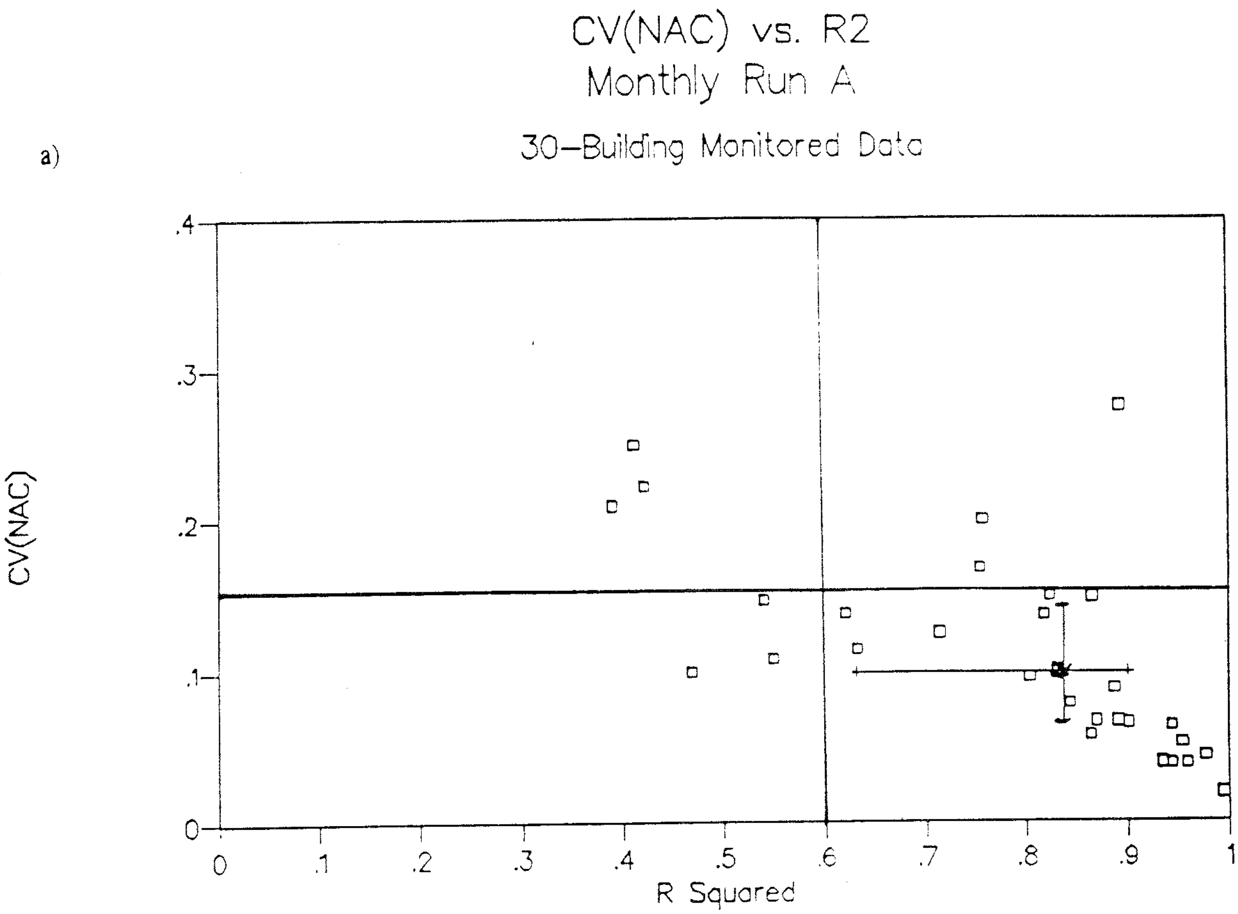
Run #	N	R^2	$\tau(\pm se)$ (°F)	NAC(±se) (gallons/year)	CV(NAC)
Run A	8	0.420	80(258)	14958(3279)	0.22
Run B	7	0.901	67(14)	14411(1040)	0.07

Monthly aggregation

Since many of the buildings had frequent deliveries, this data set gave us the opportunity to explore whether additional improvements in the PRISM fits resulted from monthly aggregation of the data. It also provided a more direct comparison of PRISM fits of oil data with those of gas and electricity utility billing data, which are generally in fairly even monthly increments. To do this, we combined the delivery data for the 30 buildings into approximate monthly sums, using the actual date for the latest delivery in the month. Again, a plot of CV(NAC) vs. R^2 was produced (Monthly Run A; Figure 11a). By combining outlying data points and rerunning the PRISM analysis, improved fits were evident (Monthly Run B; Figure 11b).

The median values of R^2 and CV(NAC) in Table 4, and the quartile distributions superimposed on the plots in Figures 9 and 11, indicate the similarity between results from the second (outlier-corrected) run of the more frequent data (Run B) and those from the first run of the combined monthly data (Monthly Run A). Median R^2 is 0.78 for Run B vs. 0.84 for Monthly Run A, and median CV(NAC) is 0.10 for Run B vs. 0.09 for Monthly Run A. This similarity is reassuring since, in many cases, the combination of outlying points for the original data set should be automatically accomplished when oil delivery data within a month are combined. Strikingly similar, not only in median

a)



b)

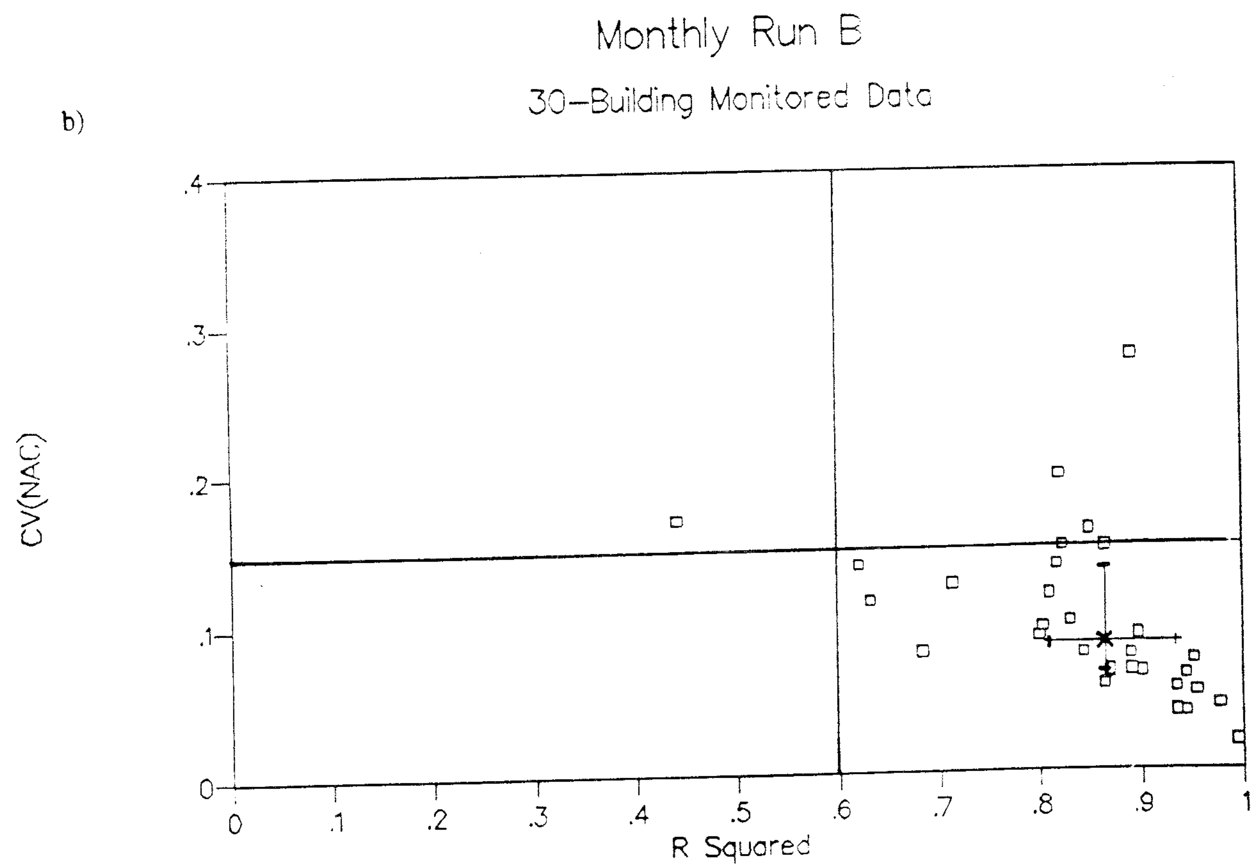


Figure 11. BMD data plot of CV(NAC) vs. R^2 for a) original data aggregated into monthly intervals (Monthly Run A); b) improved monthly aggregated data (Monthly Run B).

values but also in the distribution of results, are the two outlier-corrected runs (Run B and Monthly Run B). Apparently, monthly aggregation is not needed once the outliers have been corrected.

Table 4. Median statistics for PRISM analyses of oil delivery data from BMD data base.*

	R^2	CV(NAC)
Run A	0.698	0.13
Run B	0.784	0.10
Monthly Run A	0.837	0.09
Monthly Run B	0.864	0.08

Run A = original data set
Run B = original data set with outliers combined
Monthly Run A = original data set combined into monthly increments
Monthly Run B = monthly data set with outliers combined

*See Figures 9 and 11 for corresponding quartile plots.

Note that two extreme cases, the lowest-R^2 case and the highest CV case in Figure 9b, remain similarly extreme cases even after monthly aggregation (Figure 11b), whether or not outliers are corrected. Therefore, their unreliability seems to result from some source other than identifiable outliers. Investigation of the reason for this unreliability, to determine, for example, whether there is a physical reason that consumption in these buildings does not follow outside temperature, would be worthwhile in a follow-on study.

In general, it appears that aggregating the data into monthly sums does not improve the fits substantially over simple treatment (correction) of outliers in the original, more frequent delivery data. This is a satisfying result in the sense that straightforward outlier correction minimizes the reduction of data points, whereas monthly aggregation often goes beyond that to unwarranted loss of information. (Note that care must be taken when comparing the R^2 values for these two runs since they reflect statistics based on a different number of degrees of freedom; the median number of data points for Run B is 10 vs. seven for the combined monthly analysis.)

The delivery data from the 30-building monitoring project complement well the BEUTS data in that the former represents the highest quality data one can reasonably expect for oil-heated multifamily buildings. An example of the PRISM results is given in Table 5 for nine of the 30 buildings for which we received detailed data. Even these "good" data are not immune to problems. Nevertheless, as seen in the BEUTS analyses, careful but simple treatment of outliers can yield substantial improvements in the PRISM results.

PRISM Sliding Analysis of oil delivery data

In order to examine the ability of PRISM to monitor consumption changes over time in oil-heated buildings, historical oil delivery data were entered for nine of the 30 buildings in the BMD data base. (These nine are the set of buildings for which detailed monitored data have been compiled.) Seven of the nine buildings had data spanning 1986-1991, and the other two had only the most recent two years of data.

The PRISM Sliding Analysis is a tool used often in cases where several years of consumption data are available. PRISM is run on yearly subsets of the data, moving the analysis forward approximately one month, or one consumption period, at a time. The resulting plot of NAC vs. time provides a clear picture of changing consumption over time, and the onset of energy savings as a result of conservation measures in the building. This approach has been successfully applied to monthly metered data, for gas or electricity, but not previously to oil delivery data with uneven spacing.

Figure 12 shows an example of the sliding analysis applied to the oil delivery data for building BMD7. The first data point represents NAC and its error bars from

Table 5. PRISM "Compare" file for the nine-building subset of the BMD data base.

***** ******* PRISM-Heating Only (HO) **

UNIT ID	PRE OR POST	SAMP TYPE	TIME PERIOD	# PDS	# DAYS	RAW CONS X	RXR	TREF	BASE LEVEL X PER DAY	HEAT SLOPE X PER HDD	HEATING PART X PER YEAR	NAC X PER YEAR
BMD1		AP	06/14/90-04/04/91	4	294	8827G	0.893	48.0(10.0)	12.59(14.15)	4.876(4.113)	8247.(2893.)	12847.(3534.)
BMD2		AP	05/24/90-04/10/91	9	321	19083G	0.842	67.6(16.0)	31.51(16.21)	2.070(0.810)	11386.(4884.)	22894.(1628.)
BMD3		AP	07/13/90-04/16/91	8	277	16201G	0.789	49.7(7.7)	35.17(12.83)	4.900(2.445)	9469.(2525.)	22315.(2951.)
BMD5		AP	05/24/90-04/29/91	19	340	41722G	0.722	58.3(7.1)	53.87(31.71)	9.972(2.905)	34047.(7678.)	53724.(5559.)
BMD6		AP	05/18/90-04/05/91	10	322	25411G	0.769	52.9(7.7)	40.46(19.67)	7.342(3.428)	17829.(4079.)	32605.(4333.)
BMD7		AP	07/02/90-04/19/91	8	291	31727G	0.866	53.4(6.7)	45.25(29.12)	10.634(3.801)	26699.(6015.)	43226.(5740.)
BMD8		AP	05/01/90-04/25/91	20	359	44612G	0.632	63.4(9.9)	52.00(44.36)	7.838(2.354)	35176.(*****)	54170.(6388.)
BMD9		AP	05/09/90-04/25/91	12	351	22188G	0.394	75.0(41.4)	23.83(73.24)	2.529(1.580)	19149.(*****)	27854.(4824.)
BMD10		AP	05/21/90-04/23/91	11	337	50695G	0.897	73.2(13.6)	24.85(63.79)	7.622(1.723)	53527.(*****)	62605.(3945.)

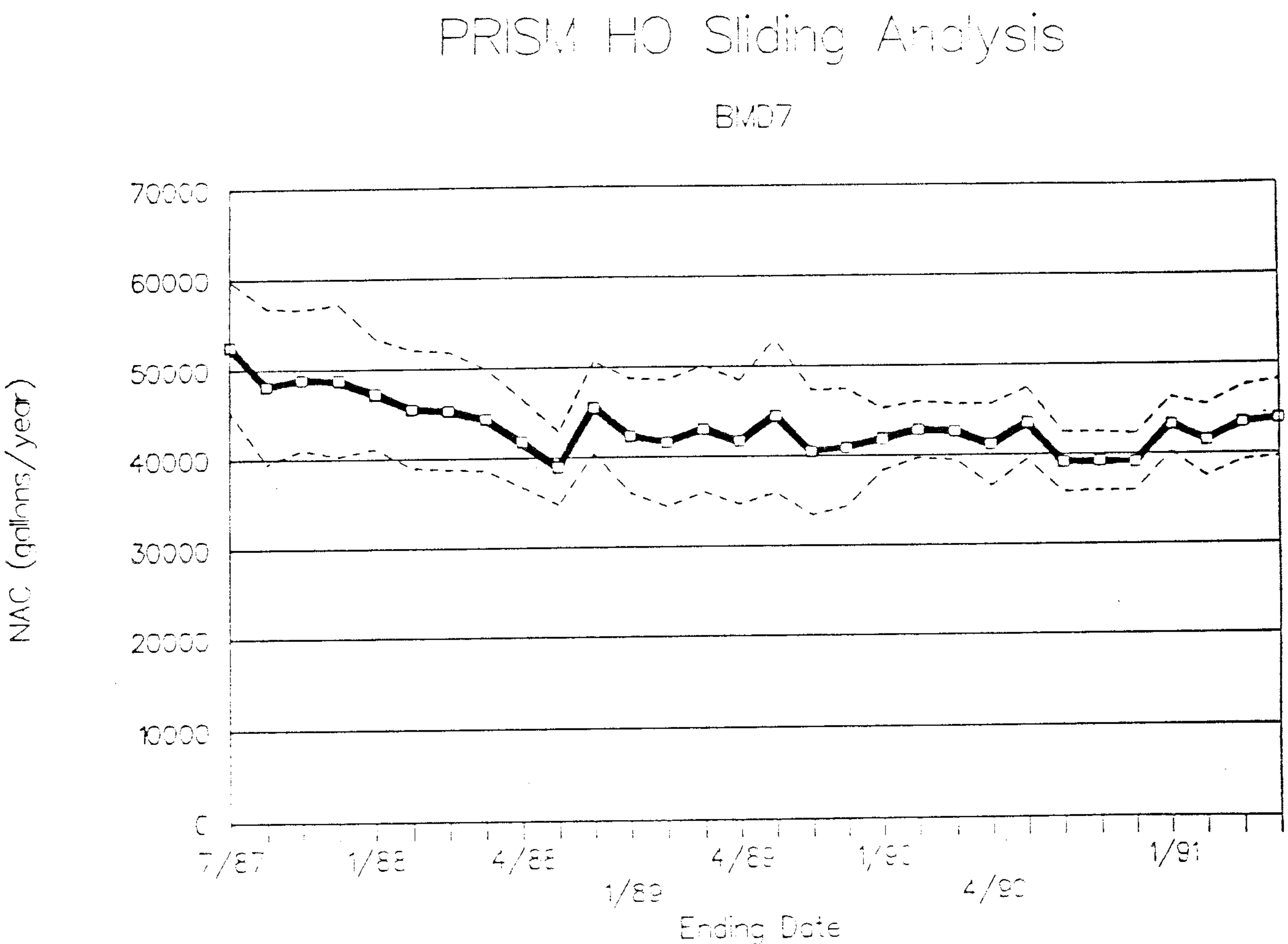


Figure 12. PRISM Sliding Analysis of oil delivery data for Building BMD7. Each point represents the NAC from a PRISM analysis run on approximately one year of data. The dashed lines represent the standard errors of the estimates. See Table 6 for exact values and dates covered by each point.

a PRISM analysis spanning June 1986 through July 1987, the second spanning November 1986 through November 1987, etc. Table 6 summarizes the PRISM results obtained from each set of data points. Note that, for oil delivery data, moving forward one month at a time is not always possible since deliveries are erratically spaced. In this case, we used a set of consumption periods that spanned as close to one year as possible; Table 6 shows the exact dates represented by each PRISM run.

Evident in the NAC plot for building BMD7 is a steady decline in consumption in the first year, followed by a leveling off in the most recent four years, with some minor fluctuations. As seen in Table 6, the PRISM analysis is fairly well determined: 29 of the 30 yearly analyses gave an R^2 of 0.60 or higher and 21 gave a CV(NAC) less than or equal to 0.15. As indicated by the BEUTS data base and the single-year analyses of BMD data, generally reliable PRISM results are obtainable from delivery data for oil-heated buildings, even though NAC seems to be less well determined than for gas-heated buildings with available metered data.

Figure 13 shows another sliding PRISM analysis for a second building from this data set, BMD8. Table 7 provides the dates and estimates of the PRISM analyses. Again, consumption trends are clearly visible: a decrease in consumption until January 1989 when it began to level off, followed by an increase to recent periods.

Study of vacancy rates

A question often asked about energy use in multifamily buildings is whether vacancy rates are a strong determinant of energy consumption. In terms more specific to our PRISM study objectives: should vacancy rates be explicitly included in a weather adjustment of a building's energy consumption? An unexpected addition to this study has been the availability of vacancy-rate data for the buildings in the BMD data base, which allows exploration of the vacancy-rate question.

Table 6. Summary of PRISM sliding analysis results for Building BMD7.

Date	Rsq	Reference Temperature(°F)	NAC (±std. err) (gallons/year)
06/25/86-07/30/87	0.699	87.0(-9.0)*	52598.(7269.)
11/05/86-11/04/87	0.742	67.0(18.5)	48077.(8690.)
12/02/86-12/03/87	0.788	70.4(22.7)	48875.(7848.)
12/31/86-12/28/87	0.792	69.3(21.7)	48729.(8433.)
01/10/87-01/05/88	0.873	68.8(15.0)	47224.(6225.)
01/24/87-01/23/88	0.847	74.0(28.8)	45480.(6522.)
02/06/87-02/10/88	0.836	74.0(29.8)	45301.(6577.)
02/28/87-03/01/88	0.843	76.6(36.7)	44349.(5714.)
04/17/87-04/01/88	0.917	57.2(7.4)	41676.(4946.)
07/30/87-09/06/88	0.932	61.0(6.0)	38770.(4112.)
11/04/87-11/21/88	0.850	55.1(6.2)	45547.(5148.)
01/05/88-01/03/89	0.725	50.2(11.7)	42324.(6367.)
02/10/88-02/18/89	0.602	54.2(13.2)	41524.(7095.)
03/01/88-03/10/89	0.631	46.0(17.7)	43011.(7104.)
04/01/88-04/11/89	0.607	59.0(14.9)	41568.(6880.)
09/06/88-08/18/89	0.574	50.0(18.2)	44467.(8608.)
10/11/88-10/06/89	0.683	57.8(16.6)	40378.(6947.)
11/21/88-11/21/89	0.679	59.5(22.0)	40859.(6532.)
01/24/89-01/17/90	0.889	86.0(-9.0)	41773.(3509.)
02/18/89-02/15/90	0.908	86.0(-9.0)	42799.(3164.)
03/10/89-03/10/90	0.907	86.0(-9.0)	42552.(3079.)
04/11/89-04/17/90	0.863	70.1(18.6)	41053.(4581.)
08/18/89-07/02/90	0.877	62.0(11.7)	43440.(3982.)
10/06/89-10/10/90	0.923	70.5(11.2)	39093.(3336.)
11/21/89-11/21/90	0.923	70.1(12.5)	39074.(3274.)
12/12/89-12/18/90	0.915	70.5(12.2)	39045.(3241.)
01/17/90-01/24/91	0.940	56.0(4.2)	43256.(3073.)
02/15/90-02/20/91	0.892	55.4(6.5)	41420.(4001.)
03/10/90-03/14/91	0.892	54.5(5.8)	43341.(4159.)
04/17/90-03/14/91	0.901	57.3(8.1)	43808.(4252.)

* A standard error of the reference temperature equal to "-9.0" indicates that the value is indeterminate.

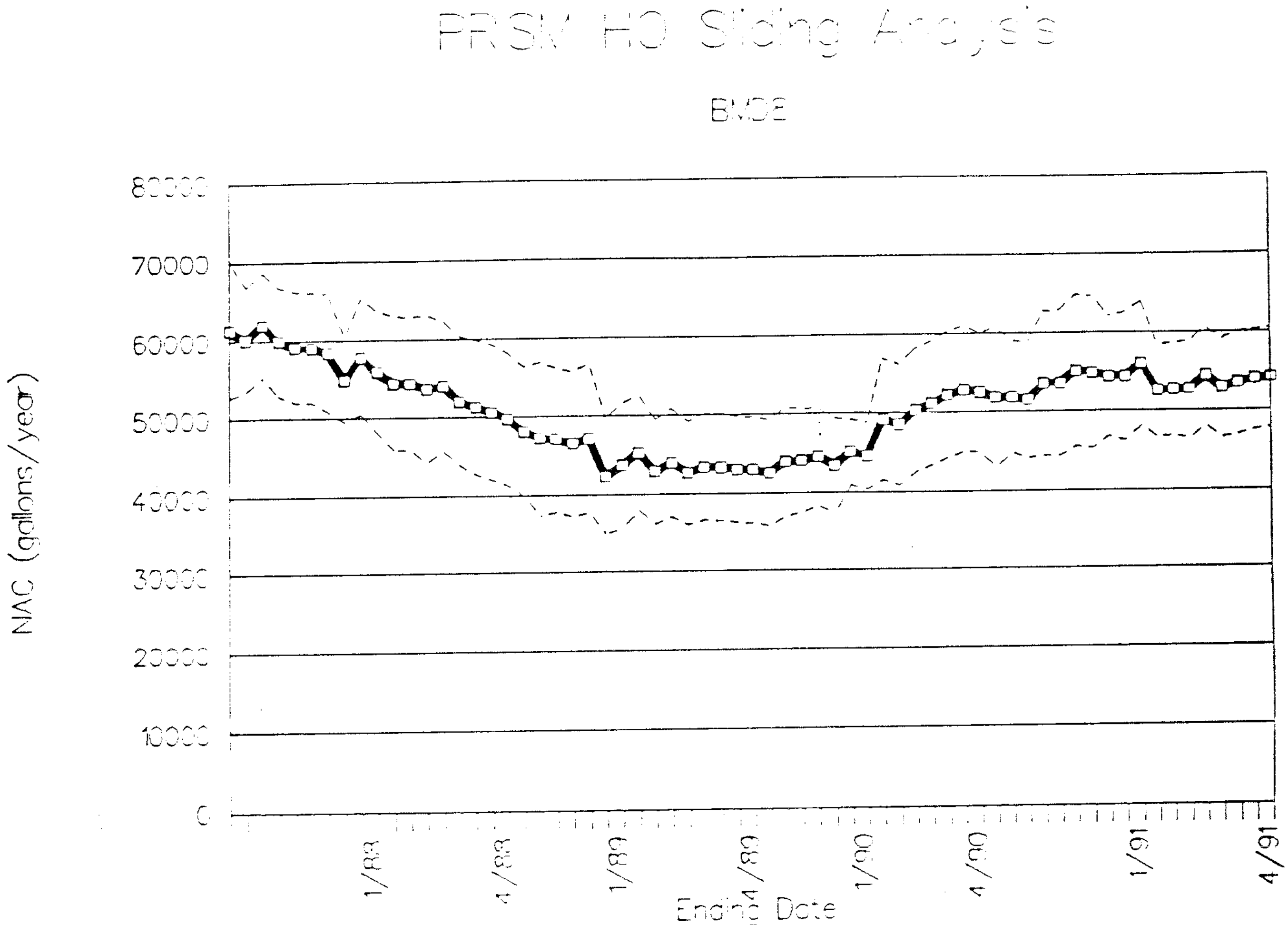


Figure 13. PRISM Sliding Analysis of oil delivery data for Building BMD8. Each point represents the NAC from a PRISM analysis run on approximately one year of data. The dashed lines represent the standard errors of the estimates. See Table 7 for exact values and dates covered by each point.

Table 7. Summary of PRISM sliding analysis results for Building BMD8.

Date	Rsq	Reference Temperature(°F)	NAC (±std. err) (gallons/year)
05/30/86-05/05/87	0.565	85.0(550.3)*	61281.(8728.)
07/07/86-08/28/87	0.635	86.0(-9.0)	60148.(6682.)
10/13/86-09/30/87	0.606	86.0(-9.0)	62007.(6686.)
11/01/86-10/29/87	0.602	86.0(-9.0)	59870.(6977.)
11/15/86-11/14/87	0.599	86.0(-9.0)	59116.(7058.)
12/01/86-11/28/87	0.600	86.0(-9.0)	59066.(7041.)
12/14/86-12/18/87	0.593	83.1(113.4)	58421.(7541.)
12/27/86-12/30/87	0.754	85.0(449.7)	54890.(5474.)
01/08/87-01/05/88	0.686	67.0(11.2)	57786.(7369.)
01/20/87-01/14/88	0.654	75.5(24.8)	55825.(7776.)
01/29/87-01/29/88	0.561	77.9(37.0)	54449.(8571.)
02/07/87-02/06/88	0.559	78.4(46.3)	54333.(8442.)
02/16/87-02/17/88	0.517	85.0(783.9)	53717.(9423.)
02/28/87-02/27/88	0.517	83.0(137.7)	53977.(8343.)
03/10/87-03/16/88	0.512	80.3(74.8)	52052.(8230.)
03/20/87-03/31/88	0.497	79.2(64.7)	51285.(8566.)
04/16/87-04/20/88	0.553	60.5(11.4)	50640.(8621.)
05/05/87-05/18/88	0.580	58.4(10.5)	49726.(8473.)
06/22/87-07/18/88	0.595	60.0(10.2)	48138.(8292.)
09/30/87-09/14/88	0.579	60.7(11.6)	47157.(9777.)
10/29/87-10/18/88	0.587	61.2(12.1)	47036.(9226.)
11/14/87-11/09/88	0.593	60.0(11.4)	46523.(9257.)
11/28/87-11/22/88	0.580	61.7(12.0)	47135.(9467.)
01/05/88-01/06/89	0.508	71.5(23.8)	42270.(7406.)
01/14/88-01/14/89	0.453	71.1(25.6)	43587.(7988.)
01/29/88-01/26/89	0.555	66.1(14.9)	45278.(7366.)
02/06/88-02/08/89	0.535	68.3(17.3)	42845.(6639.)
02/17/88-02/18/89	0.559	64.9(13.8)	43977.(6906.)
02/27/88-03/03/89	0.544	66.6(15.5)	42606.(6672.)
03/16/88-03/11/89	0.589	65.0(13.3)	43393.(6778.)
03/31/88-03/25/89	0.589	65.0(13.5)	43286.(6776.)
04/20/88-04/13/89	0.596	63.4(12.7)	43025.(6690.)
05/18/88-05/11/89	0.596	62.5(12.6)	42884.(6837.)
07/18/88-07/10/89	0.601	63.1(12.6)	42503.(6725.)
09/14/88-08/15/89	0.579	61.0(11.8)	43970.(6953.)
10/18/88-10/06/89	0.591	61.0(11.3)	43950.(6588.)
11/09/88-11/04/89	0.603	63.6(13.4)	44555.(6257.)

Date	Rsq	Reference Temperature(°F)	NAC (±std. err) (gallons/year)
11/22/88-11/22/89	0.644	55.9(11.8)	43396.(6056.)
01/14/89-01/13/90	0.857	58.9(5.9)	45113.(4157.)
01/26/89-01/29/90	0.852	57.2(5.5)	44559.(4281.)
02/08/89-02/03/90	0.626	65.0(12.7)	49194.(7702.)
02/18/89-02/16/90	0.613	65.0(12.7)	48454.(7672.)
03/03/89-03/05/90	0.592	67.0(14.3)	50328.(7814.)
03/11/89-03/13/90	0.584	67.4(14.4)	51076.(7891.)
03/25/89-03/27/90	0.584	69.4(16.8)	52225.(8045.)
04/13/89-04/06/90	0.579	70.7(19.1)	52882.(8042.)
05/11/89-05/31/90	0.599	70.7(17.6)	52551.(7642.)
08/15/89-07/24/90	0.569	73.1(25.8)	51784.(8899.)
08/15/89-09/20/90	0.611	72.7(17.3)	51884.(7081.)
10/06/89-10/30/90	0.604	71.9(16.2)	51425.(7413.)
11/04/89-11/10/90	0.570	80.0(107.8)	53519.(9194.)
11/22/89-11/27/90	0.566	80.0(108.8)	53545.(9216.)
12/09/89-12/07/90	0.558	80.0(109.3)	55150.(9625.)
12/16/89-12/19/90	0.543	80.0(109.2)	54904.(9654.)
12/31/89-12/31/90	0.472	74.1(24.4)	54443.(7709.)
01/13/90-01/15/91	0.471	73.9(23.3)	54359.(8054.)
01/29/90-01/24/91	0.521	73.0(20.9)	56022.(7820.)
02/03/90-02/02/91	0.622	73.0(15.3)	52567.(5914.)
02/16/90-02/15/91	0.627	72.7(15.0)	52677.(5870.)
02/24/90-02/21/91	0.637	70.3(12.8)	52749.(6235.)
03/05/90-03/05/91	0.661	67.0(11.1)	54394.(6256.)
03/13/90-03/16/91	0.637	69.3(11.9)	52955.(6297.)
03/27/90-03/28/91	0.641	63.1(9.7)	53518.(6506.)
04/06/90-04/10/91	0.638	63.0(9.9)	53924.(6384.)
05/01/90-04/25/91	0.632	63.4(9.9)	54170.(6388.)

* A standard error of the reference temperature equal to "-9.0" indicates that the value is indeterminate.

Monthly vacancy-rate data were compiled from detailed occupancy records by apartment for each building, for the same nine buildings used in the sliding analyses.* Data for the one-year period from July 1990 to June 1991 were obtained. For the PRISM analyses, these data were combined with delivery data for July 1990 to June 1991. The vacancy-rate data used in the analyses are summarized in Table 8. Most buildings have almost constant vacancy rates. The similarity of the median and mean vacancies shows the lack of outlying high and low vacancies. For most of the buildings, peak vacancy was about 8% of the building's total units.

One approach to the vacancy-rate question is to explore whether the monthly residuals from the linear regression in the PRISM model, i.e., the part of each month's actual consumption not explained by PRISM, correlate with the monthly vacancy rates. Residuals reflect how much the actual consumption is above or below the consumption predicted by the PRISM fit during that period. If there were an above-average number of vacancies in one period, the total consumption of the building during that period would be lower than predicted by PRISM; therefore, the residual for that period would be negative. Similarly, a below-average vacancy should result in a positive residual. Therefore, a negative correlation between residuals and vacancies would be expected. These hypotheses assume that vacancy is a significant factor in the determination of the residuals, and that a vacant apartment will use less fuel than an occupied apartment.

For each of the nine buildings, regressions of PRISM residuals against vacancy were run. The numbers used for vacancy in each delivery period were weighted averages of the vacancies in each month spanned by the delivery period. The results are shown in Table 9.

---

*Data were obtained with Fred Goldner of Energy Management Associates, who is using the vacancy data in his study of the same buildings (Goldner, 1991).

Table 8. Summary of vacancy data for nine-building subset of BMD data base.

Building	Total Units	Median Vacancy	Mean Vacancy	Mean % Vacancy	Peak Vacancy	Peak % Vacancy	Peak Month
BMD1	23	1	.83	3.62%	2	8.70%	6/91
BMD2	36	0	.67	1.85%	3	8.33%	8/90
BMD3	37	2	2.33	6.31%	5	13.51%	4/91
BMD5	25	1	1.25	5.00%	2	8.00%	4/91-6/91
BMD6	49	0	1.08	2.21%	4	8.16%	5/91
BMD7	61	2	2.08	3.42%	3	4.92%	8/90-11/90
BMD8	41	3	3.08	7.52%	4	9.76%	7/90-8/90
BMD9	80	5	5.17	6.46%	9	11.25%	5/91
BMD10	102	3	4.33	4.25%	11	10.78%	7/90

Table 9. PRISM estimates for the nine-building subset of BMD data base and corresponding residuals vs. vacancy regression results.

	PRISM run					Residuals vs. Vacancy				
Building	N	R2	CV(NAC)	Tau	se(Tau)	r2	Intercpt(A)	se(int)	Slope(B)	se(slope)
BMD1	4	.893	.28	48.0	10.0	.555	-20.8	6.38	640.2	405.65
BMD2	10	.718	.08	54.8	9.4	.032	.89	11.63	-298.4	576.28
BMD3	9	.817	.11	50.0	5.9	.000	-.59	15.36	9.2	192.15
BMD5	21	.658	.09	53.8	6.3	.004	-10.57	56.68	215.44	758.73
BMD6	11	.815	.08	32.2	1.9	.013	-1.94	23.49	89.68	261.1
BMD7	9	.892	.10	54.4	5.4	.023	-12.71	26.61	379.25	929.8
BMD8	20	.629	.11	63.0	10.3	.002	34.69	60.25	-474.52	2571.13
BMD9	12	.374	.54	78.0	226.0	.01	15.44	44.66	-242.82	749.72
BMD10	12	.88	.07	67.0	9.6	.043	27.51	26.77	-937.19	1402.58

In general, the correlation between residuals and vacancy (see "Residuals vs. Vacancy r^2" column*) is extremely low: eight out of the nine runs have r^2 less than 0.05, and the one high value (0.56, for building BMD1) was from only four data points. Apparently, at a monthly level, vacancy does not have a significant effect on consumption above or below the value estimated by PRISM. Furthermore, although four out of nine runs had, as expected, negative values of the coefficient B (slope), most values of A (intercept) and B were poorly determined, with error bars larger than the estimates themselves.**

The poor residuals vs. vacancy correlations in Table 9 result from analyzing several months of data at a time. Suspicious that this may mask an effect that might be evident in individual monthly data, we also examined maximum and minimum vacancy data for each building to see if they corresponded to an extremum in the opposite direction for the residuals. No pattern was seen in individual monthly data, but an intriguing pattern was seen when the extrema were examined. For this, the number of buildings (out of nine) that had residuals classified as either high or low were tallied for each month. A residual was classified as high if it was higher in value than the residuals on either side of it, i.e., if it was a local maximum, and if it was larger than approximately 33% of the maximum value of the residuals. Analogous criteria for low were applied.

The number of buildings with a high or low residual in each month was then compared with the vacancy rate in each month averaged across the nine buildings.

---

*Note about notation: r^2 is used for the square of the correlation for residuals vs. vacancy, and other correlations reported here, to distinguish it from the PRISM model's R^2 statistic.

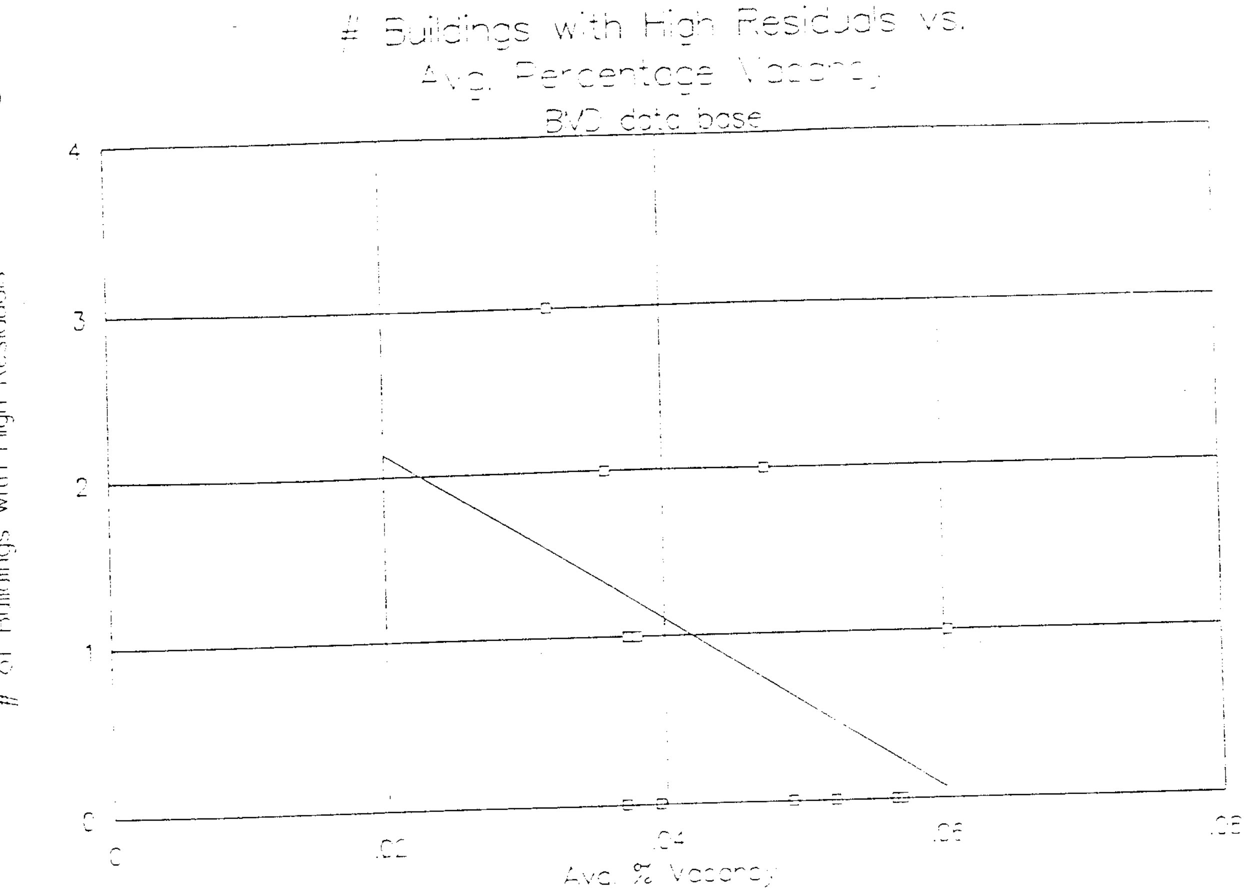
**Earlier work done on this vacancy study, which analyzed an incomplete data set excluding summer data, showed the importance of the presence of summer data and the effect on PRISM results (Progress Report #4, August-September 1991). This point will be addressed in a later section (see ECC Case #1).

The resulting frequency of high residuals and average vacancy for each month were then examined. The r^2 for the resulting scatterplot of high-residual frequency vs. vacancy, shown in Figure 14a, is 0.24, and the r^2 from the analogous low residuals plot in Figure 14b is 0.35. When all buildings are grouped together, the occurrence of extreme residuals from PRISM correlate with vacancy rate, albeit weakly but to a greater extent than the monthly residuals from individual buildings seem to. Whether or not this trend is important needs to be examined with data from a larger, more varied set of buildings. This method of summarizing the effect of vacancy on PRISM results should be useful in future studies, when such data bases become available.

An additional approach to exploring the effect of vacancy rates is to run PRISM on consumption data that have been adjusted by vacancy rates.* To adjust the data for vacancy, the amount of oil delivered during each period was divided by the weighted average of the number of occupied units in the building during the same period. PRISM was then run on these adjusted data, producing NAC in terms of gallons/unit; the results are summarized in Table 10. The quality of the PRISM fits are strikingly similar between adjusted and unadjusted runs, as seen by similar values for R^2, CV(NAC), the estimate of reference temperature, and its standard error. In some cases the adjusted run is a little worse than the unadjusted run, and, overall, adjusting the delivery data for vacancy does not seem to improve the quality of the PRISM results. More important, the vacancy-adjusted NAC (in gallons/unit) corresponds closely with the unadjusted NAC (in gallons/building) divided by the average number of occupied units. This average is weighted by the number of days in each month, and also accounts for turnovers and vacancies part of the way into the month. The percentage difference of these two values of NAC is less than 1% for seven of the nine buildings. This percentage difference is less than the CV(NAC) of both the adjusted and

*This was done in an earlier study (Goldman and Ritschard, 1986).

a)



b)

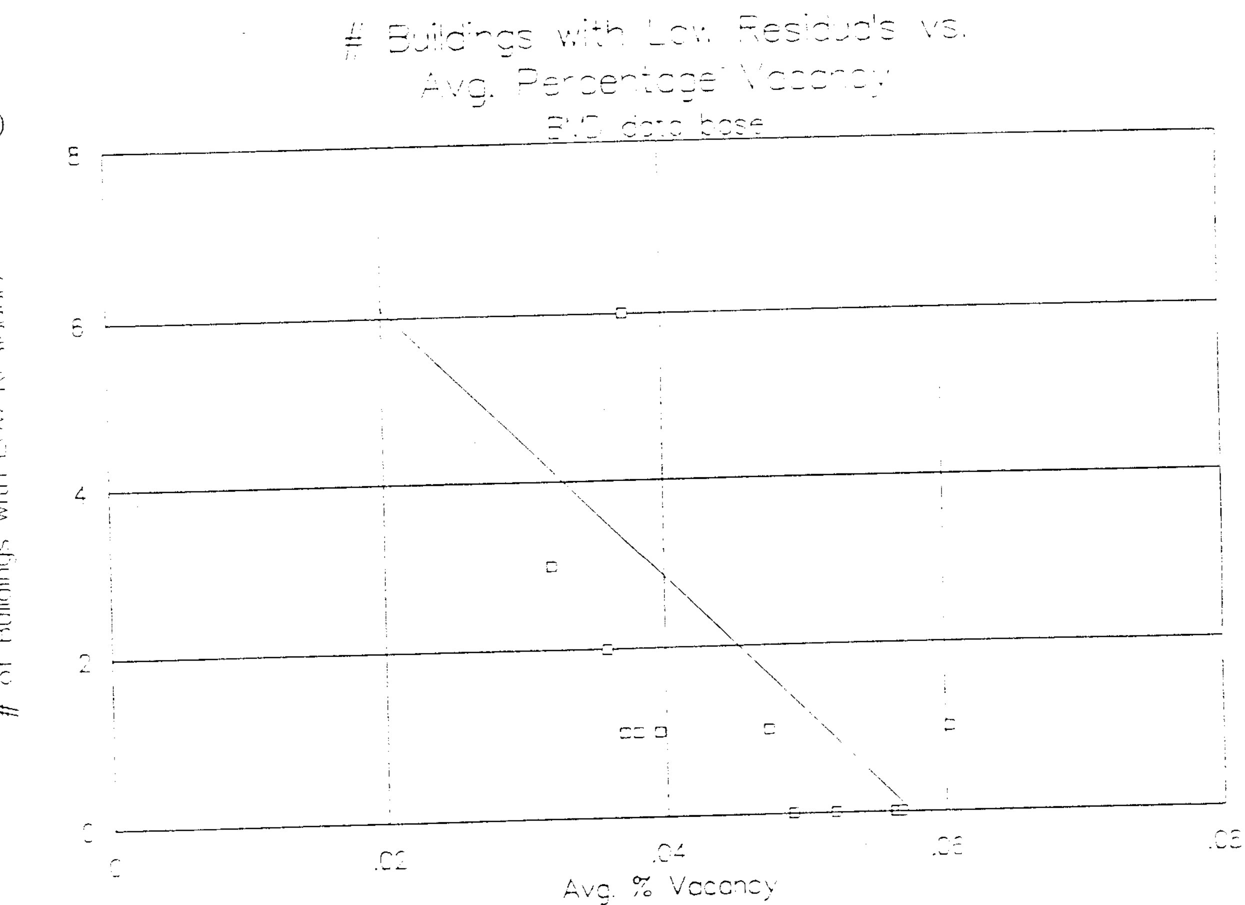


Figure 14. Frequency of PRISM residuals vs. average vacancy in each month averaged across the nine-building subset of the BMD data base for a) high residuals; b) low residuals.

Table 10. Comparison of PRISM results (with vacancy rates not included) with PRISM results based on each period's consumption divided by corresponding occupancy rate.

	N	R2	CV(NAC)	Tau (°F)	se(tau)	NAC (gallons/year)
BMD1	4	0.893	0.28	48.0	10.0	576
(adj)	4	0.886	0.28	48.0	10.4	579
BMD2	10	0.718	0.08	54.8	9.4	698
(adj)	10	0.714	0.08	54.4	9.2	699
BMD3	9	0.817	0.11	50.0	5.9	628
(adj)	9	0.811	0.11	50.0	6.0	628
BMD5	21	0.658	0.09	53.8	6.3	2508
(adj)	21	0.639	0.09	63.5	8.5	2406
BMD6	11	0.815	0.08	32.2	1.9	895
(adj)	11	0.716	0.12	52.5	8.2	727
BMD7	9	0.892	0.10	54.4	5.4	733
(adj)	9	0.891	0.10	54.4	5.5	733
BMD8	20	0.629	0.11	63.0	10.3	1441
(adj)	20	0.629	0.12	63.0	10.3	1440
BMD9	12	0.374	0.54	78.0	226.0	376
(adj)	12	0.371	0.53	78.0	227.0	376
BMD10	12	0.88	0.07	67.0	9.6	683
(adj)	12	0.88	0.06	66.5	9.4	684

unadjusted runs, in every case, meaning that the two values of NAC per occupied unit are statistically equivalent.*

The analyses reported here suggest that, at least for buildings where vacancy does not vary greatly from month to month, the consumption per occupied unit can be reliably and simply estimated by dividing NAC from the PRISM run of building-level oil-delivery data by the average number of occupied units over the PRISM run. The extent to which this simple approach is generalizable to multifamily buildings should be tested on other data sets.

Analysis of detailed data set

Detailed daily monitored data, received from Fred Goldner of Energy Management Associates for nine of the 30 buildings, span the summer of 1990 and include such variables as furnace runtime, metered oil consumption, high and low outdoor temperatures, makeup water, and consumption of domestic hot water. In addition, temperatures for individual apartments and water flows on an hourly basis are included for two of the nine buildings. Since the boiler is used for domestic water heating as well as space heating, the furnace runtime in the summer reflects oil use for water heating. See Goldner (1991) for a detailed study of these data.

Data for heating as well as non-heating months, not available for this report, will be available at a later date for this set of buildings. Such detailed data for a year-long period would be ideal for PRISM validation studies, for example, to assess whether

*The PRISM runs for building BMD6 reported in Table 10 show a large difference between the original and vacancy-adjusted NAC -- the only such case. These results are from the original unimproved data, because they met the reliability criteria for R^2 and CV(NAC). Therefore, this NAC sensitivity is puzzling. A closer look at the original data indicated two consecutive outliers, the combination of which gave an NAC estimate much closer to the NAC from the vacancy-adjusted data, and a much more reasonable reference temperature, i.e., 66°F, vs. 32°F from the original data.

PRISM's base-level estimate for the building is consistent with the data on domestic hot-water usage. For this project, the summer data were used as a snapshot of the entire data set to formulate questions that we might eventually want to ask of the longer time series of data.

One such question is whether the data on oil consumption and furnace runtime are linearly related. Since in theory they should be related by a constant ratio, we are in effect asking to what extent this theoretical relationship holds in the world of imperfect data and imperfect equipment. This may be a particularly useful question for oil-heated buildings, since, as is evident throughout the analyses in this report, oil deliveries may be too infrequent for an accurate determination of their weather dependence, and in view of the possibility of non-filled tanks, may be an inaccurate representation of consumption. Installation of runtime meters might provide a practical and greatly improved data source, e.g., for PRISM analyses, but only if the furnace runtime data accurately represent consumption.

Simple plots of daily oil consumption and furnace runtime data in Figures 15a-c show their ratio to be approximately constant for most but not all of the buildings. Whether the scatter is due to measurement error or physical phenomena needs to be explored. In general, it appears that either one of these variables may provide adequate information on oil consumption in the building, assuming that the factor to convert furnace runtime to oil consumption is known. If it turns out that better results from PRISM are obtainable from (frequent) runtime data than from (infrequent) oil delivery data, then use of runtime meters in conservation projects involving oil-heated buildings may improve the quality of savings estimates considerably.

An excellent opportunity for a direct comparison of PRISM applied to furnace runtime vs. oil delivery data was provided by a building in the third data set obtained for this study, ECC Case #1. In the next section, results clearly show the possible improvements in data accuracy offered by furnace runtime data.

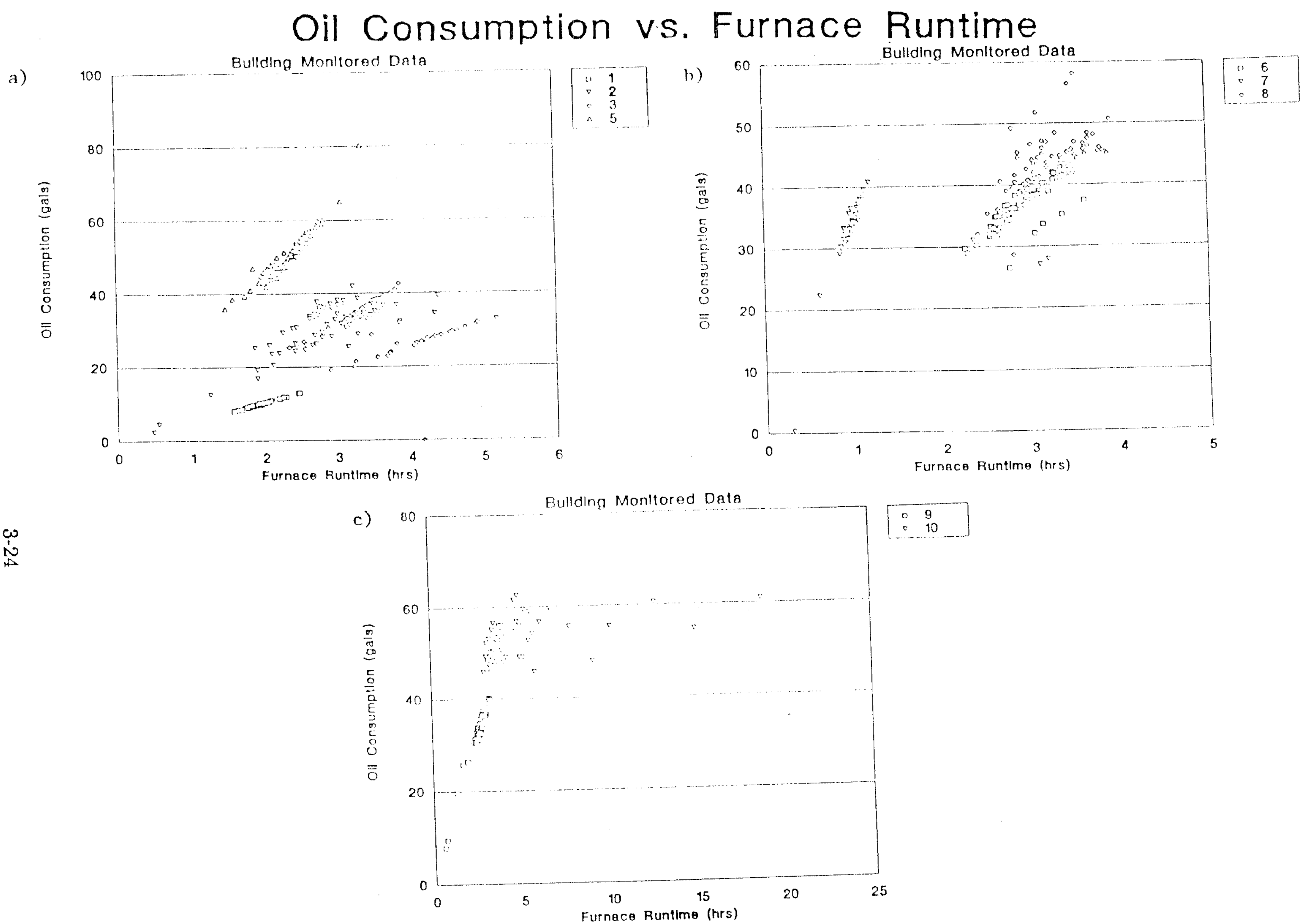


Figure 15. Oil consumption vs. furnace runtime for each of the nine buildings from the BMD data base. Plots are shown for a) buildings 1, 2, 3 and 5; b) buildings 6 through 8; and c) buildings 9 and 10.

4. ENERGY CONSERVATION CASES

Description of data set

In order to test further some of the findings obtained from the BEUTS and BMD data sets, a third data base was obtained for a set of buildings that had participated in a New York City program of Energy Conservation Cases (ECC; Judd et al., 1989). The two owner/managers of the seven buildings* for which we collected data kept careful fuel delivery records, as well as records of energy conservation measures, over a number of years. Analysis of their data provided an opportunity to test the multifamily-building PRISM approach under optimal data circumstances, i.e., a long time series of high-quality data for large and small buildings in which conservation effects may be evident. In addition, conversations with interested and conscientious building managers gave us an opportunity to begin to assess the role of this type of analysis in an interactive setting where PRISM-based monitoring of conservation actions could provide valuable feedback to multifamily building owners or managers and useful input to the owners' next conservation investments.

The six buildings, all oil-heated, consist of the following:

- ECC Case #1: a six-unit building in Woodside (Manager #1), NY, with oil consumption data from an elapsed time meter on the oil burner as well as oil delivery data spanning April 1978 to May 1991;
- ECC Case #2: a 315-unit building pair in the Bronx (Manager #2), with oil delivery data spanning January 1986 to November 1991; and
- ECC Case #3: a set of four buildings containing a total of 536 units in Staten Island (also Manager #2), with oil delivery data for each building spanning January 1986 to October 1991.

Results from the analysis of each of these three ECC data sets are presented below.

*Data from a third owner are expected soon. Because of the unusual personal management style of that building, we hope to have the opportunity to analyze these data at a later date.

The results reflect an extensive analysis of Case #1, and a less complete analysis of Cases #2 and #3.

ECC Case #1: Building in Woodside, New York

The data set for the Woodside building, owned and operated by Manager #1, is exceptional. The owner had kept oil delivery records dating back to 1979, when he had taken possession of the property. He provided CEES not only with the delivery data, but also with the records he kept on fuel consumption based on burner runtime. Therefore, this data set allowed us to make a direct comparison between oil delivery data and actual oil consumption for the same consumption intervals.

Sliding analysis of delivery data

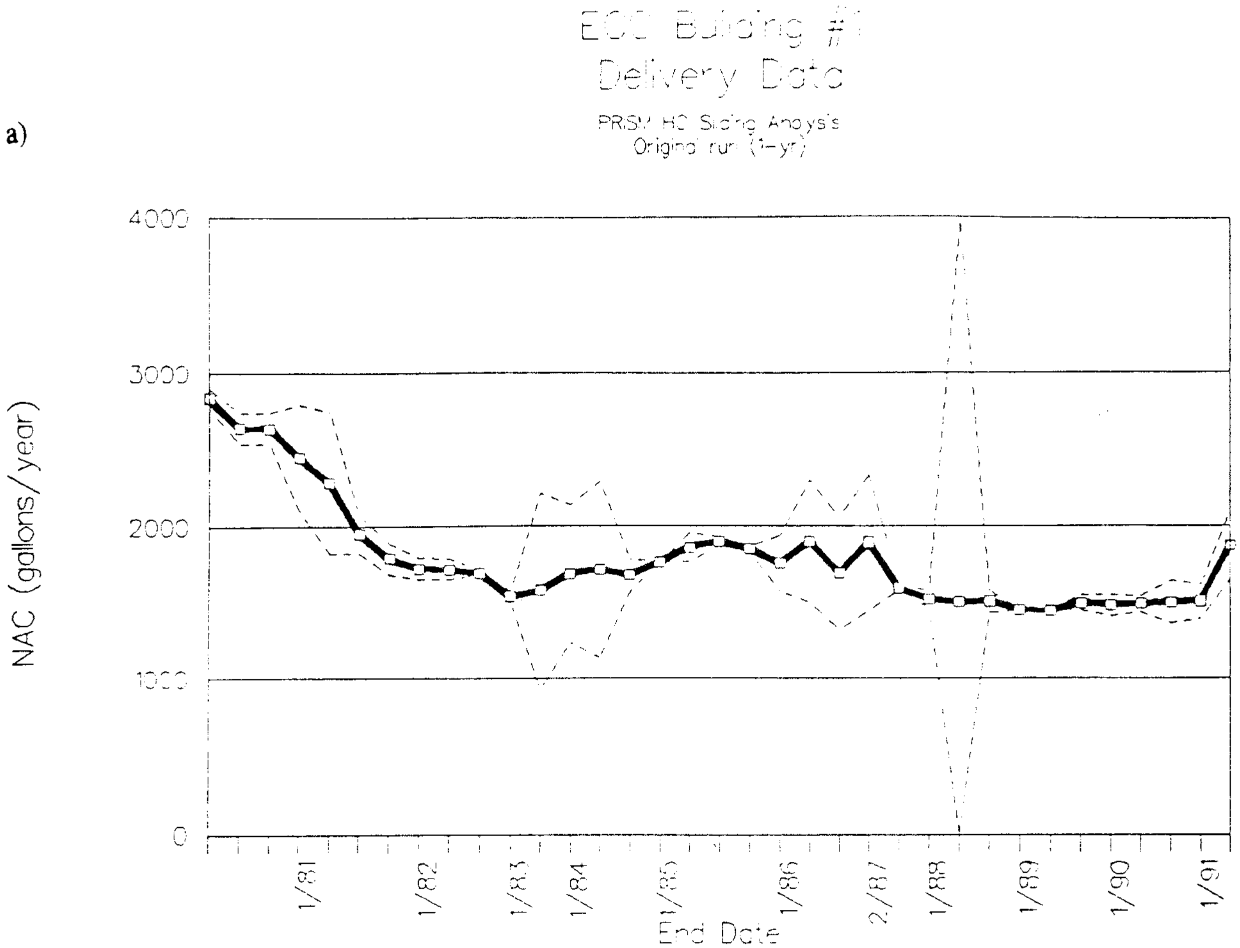
The first step was a sliding PRISM analysis of the delivery data, that were grouped and analyzed in approximately one-year periods. The NAC results and corresponding error bars are plotted in Figure 16a. The most evident change appears to be a marked decrease in consumption near the beginning of the data series in 1979, i.e., shortly after the current owner took possession of the building; this was followed by periods of fairly consistent energy consumption. In general, NAC is extremely well determined, but less stable NAC estimates, with anomalously large error bars, are evident for certain short periods, particularly the periods ending during the winters of 1981 and 1983, the summer of 1986, and the winter of 1988, the most recent period.*

Since many of the one-year periods included only four or five data points and since, after 1981, consumption appeared to be fairly smooth, a sliding analysis was also performed using two-year periods to determine whether the low number of data points

---

*Note that the period with ending date of February 1988 gives a very large standard error of NAC (the PRISM result showed it to be indeterminate). We will discuss this particular period at length later.

a)



b)

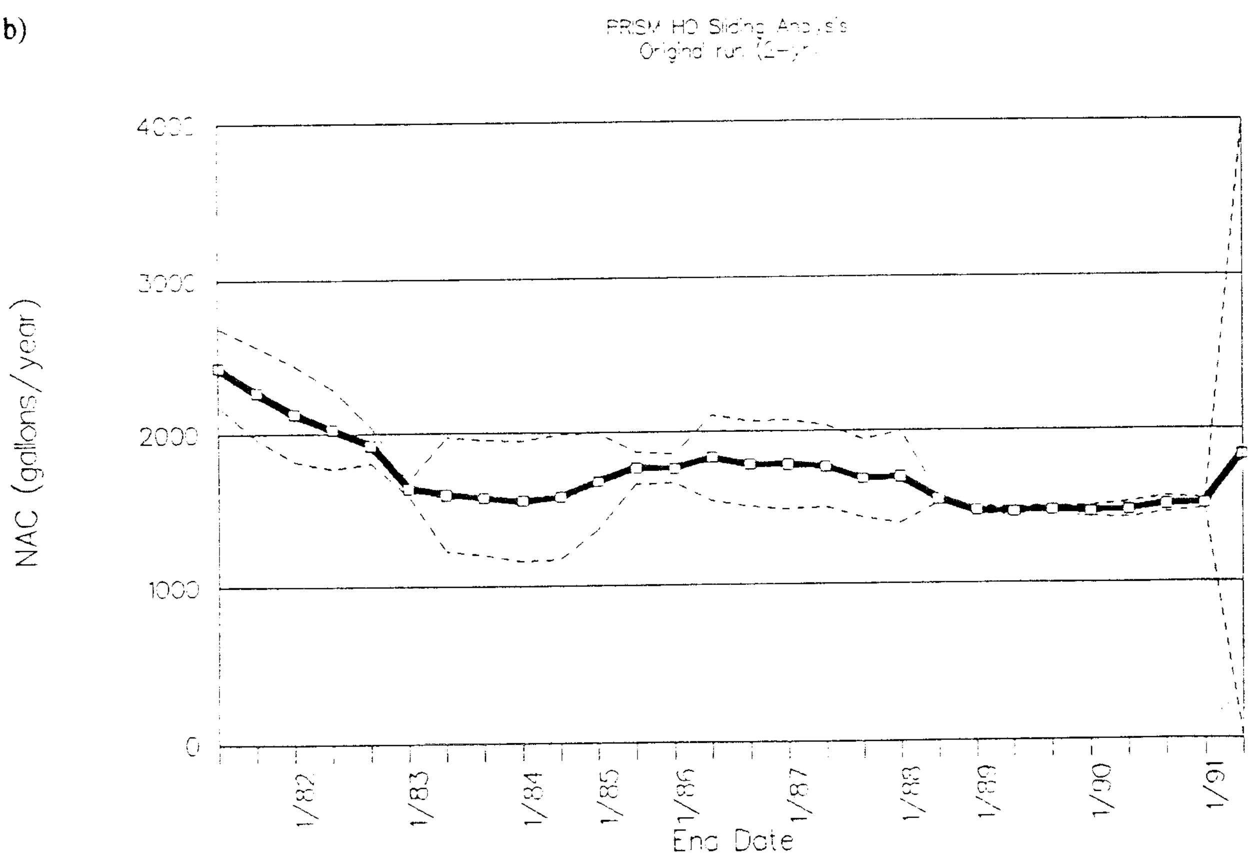


Figure 16. PRISM Sliding Analysis of original oil delivery data from ECC Case #1 (from Energy Conservation Cases). The dashed lines represents standard errors of the NAC estimates. Plot a) shows one-year analyses, while b) shows two-year analyses.

was the source of the instability in the estimates. The resulting plot is shown in Figure 16b. Although use of two-year periods appears to produce some smoothing, areas of instability are still evident. From other studies (Fels, ed., 1986), we know that the error bars for NAC and for all PRISM parameters are expected to increase substantially, and then settle to smaller values as the analysis moves through a period of change. This possibility is explored for the first period of instability (winter 1981). Since NAC is flat during the other periods of instability, other sources of instability such as data anomalies need to be explored.

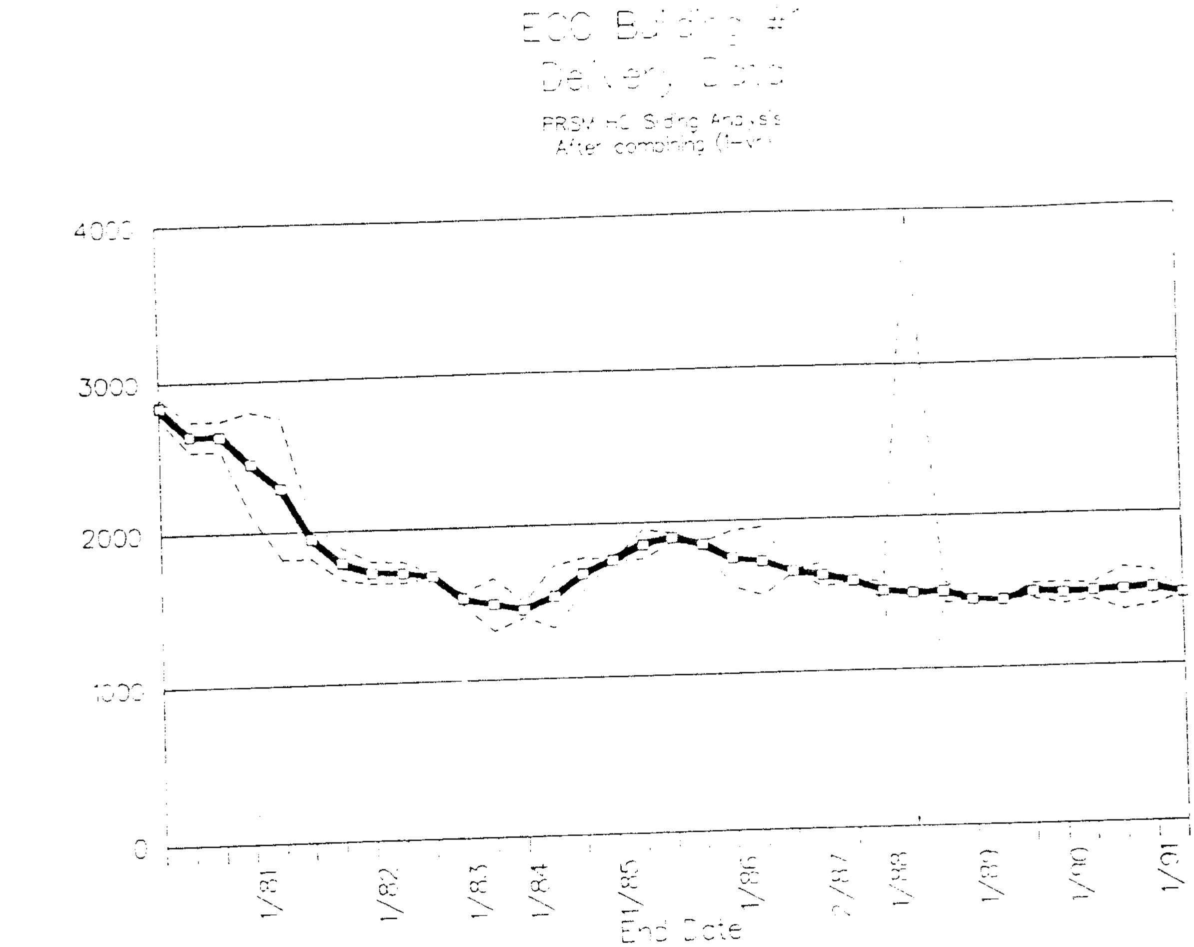
Applying previous methods developed for identifying anomalous data points, we were able to make reasonable combinations of data points during these unstable periods. The results of another set of sliding analyses applied to the improved raw data are shown in Figures 17a for one-year analyses and 17b for two-year analyses, to compare respectively with Figures 16ab for unimproved data.

A discussion of our conclusions on the improvement of the data set follows. The steps taken to explore possible data problems are described in detail, as examples of lessons that might help PRISM users analyze similar data sets of multifamily buildings.

Winter 1981: changing consumption

For the one-year period ending in winter 1981, no data problem, nor any need to combine data, was indicated. As can be seen from both Figures 16 and 17, the NAC estimates and error bars in this time period are unaffected by improvement of the data set, i.e., these results appear to be correct. The dramatic conservation effect, leading to the drop in NAC for this time period, is clearly seen in the consumption vs. degree-day plot in Figure 18. The data points for time period April 10, 1979, through October 30, 1980 (labeled A through G in the plot) can be distinctly separated from those for the

a)



b)

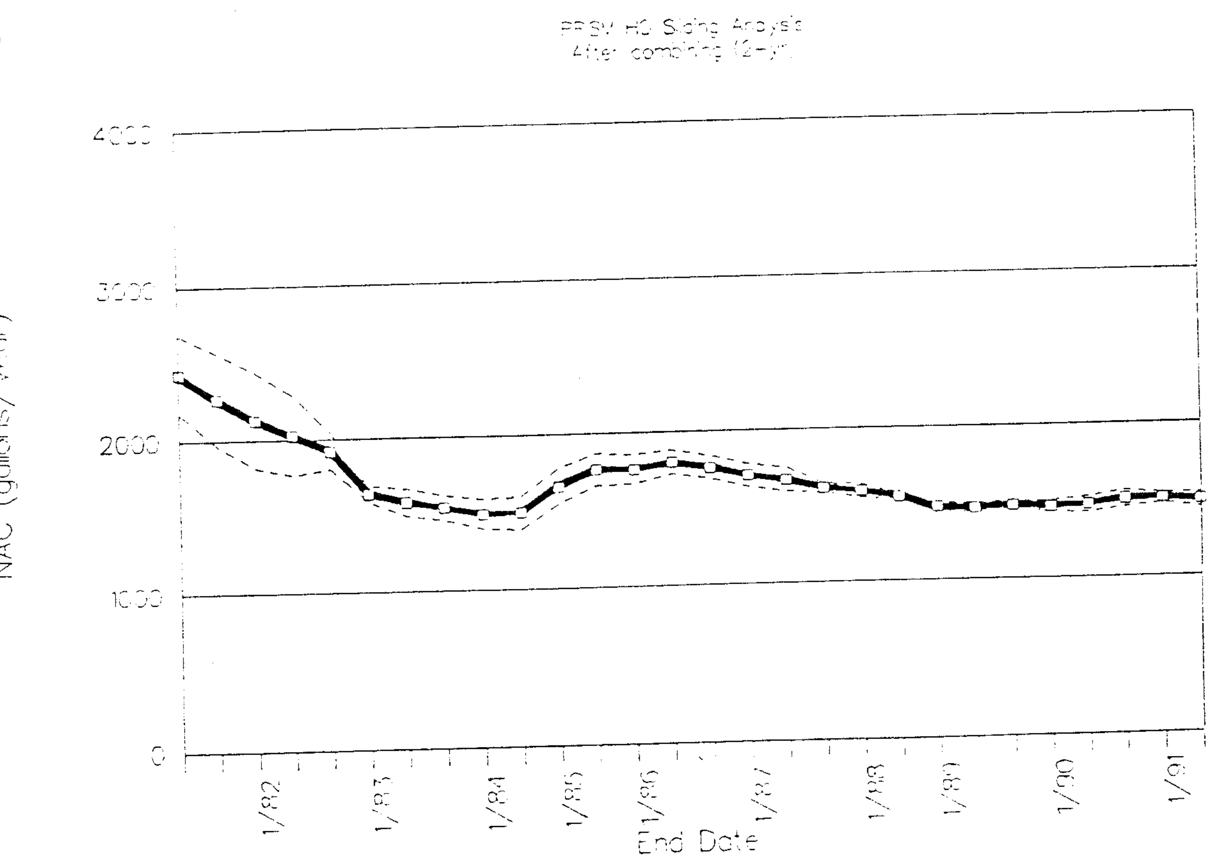


Figure 17. PRISM Sliding Analysis of improved oil delivery data from Building ECC Case #1. The dashed lines represent standard errors of the NAC estimates. Plot a) shows one-year analyses, while b) shows the two-year analyses.

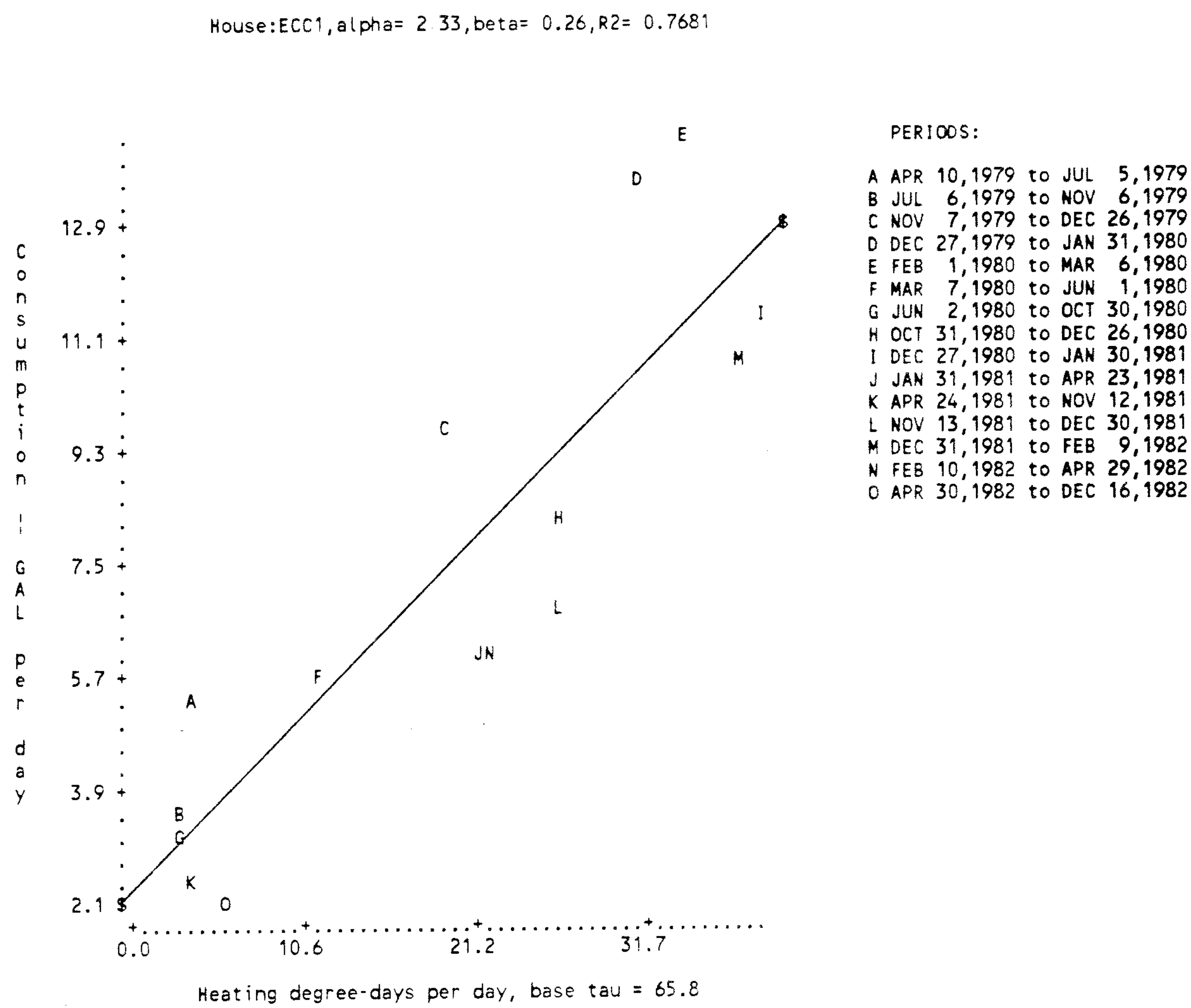


Figure 18. PRISM plot of consumption vs. heating degree-days (ECC Case #1) for unstable period encompassing Winter 1981 data. Conservation is indicated by separation of points A through G from points H through O.

period October 31, 1980, through December 16, 1982 (labeled H through O). The PRISM results for these two ("before and after") periods are distinct and highly reliable: R^2 = 0.979 and CV(NAC) = 0.04 for the first period, and R^2 = 0.986 and CV(NAC) = 0.05 for the second period. The corresponding NAC decreased by 1083 gallons/year, or 38%, from a level of 2741 gallons/year, and this savings estimate is well determined, with a standard error of 139 gallons/year, or 3.8%.

These results show that, instead of a data anomaly, there appears to be a change in the building that is influencing its energy consumption. PRISM is behaving exactly as it should, with results that are stable and reliable when pre-retrofit consumption data are well separated from post-retrofit data, and unstable when data from the two periods are mixed. As confirmation of this hypothesis, the ECC report on this building indicates that the owner implemented several conservation measures after purchasing the building.

Winter 1983: combination of two data points due to non-fill delivery

A PRISM analysis of the two-year period from April 1981 through May 1983 shows very unstable results: R^2 = 0.547, CV(NAC) = 0.23 and a reference temperature of 88.0°F with an indeterminate standard error. A plot of the original consumption data vs. heating degree-days (Figure 19) indicates an obvious need for combination of two data points: points F and G. Use of studentized residuals confirmed this hypothesis. The studentized residuals for points F and G were, respectively, -3.5 and 1.3; the latter is not as large as expected for a large outlier of the extent indicated in the plot, apparently because the studentized residual for point C, at 1.15, was similar to point G. Combination of the two data points F and G produced much improved results: R^2 = 0.983, CV(NAC) = 0.06, with a reference temperature of 53.0 ($\pm$3) °F. The substantial improvement in the stability of the PRISM estimates from the combination

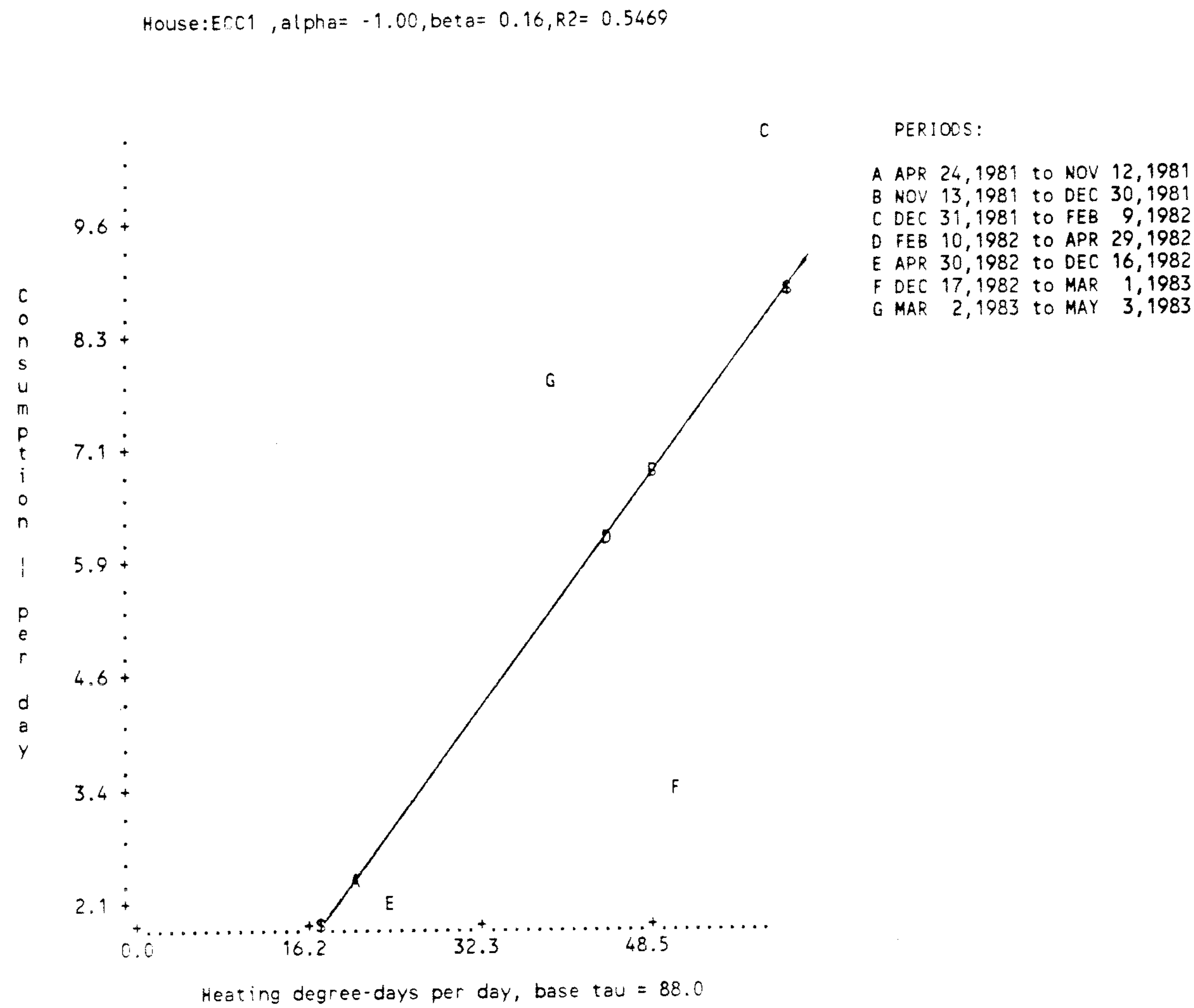


Figure 19. PRISM plot of consumption vs. heating degree-days (ECC Case #1) for unstable period during Winter 1983. Combination of outlying low/high points F and G is indicated.

of these two points can be seen in the sliding analysis plots (Figures 17ab).

Returning to the raw data file sent by the owner, a handwritten note indicates that one delivery, on March 2, 1983, was for one tank only since the other was being cleaned. This would explain the low reading shown for that time period, point F: a non-fill for period F followed by a filling of both tanks in period G leads to delivery data that underestimate consumption in period F, overestimate it in period G, and when combined accurately represent consumption. This hypothesis is corroborated by the analysis of the runtime-metered consumption data that follows.

Summer 1986: combination of three data points

The period encompassing summer 1986, April 1985 through July 1987, gave PRISM results of $R^2 = 0.528$, CV(NAC) = 0.15 and, again, a high reference temperature of 88.0°F, with an indeterminate standard error. The plot of delivery data for this period (Figure 20) shows an unusual consumption pattern: points C and E appear as large outliers, while point D lies very near the regression line. Results from the studentized residuals, -2.2 for point C, 2.9 for point E, and the magnitude of all others smaller than 0.5, support the identification of outliers. Although generally we look for pairs of high-low outliers, we tried combining these three data points. The PRISM estimates improved substantially: $R^2 = 0.996$, CV(NAC) = 0.02, with a reasonable reference temperature of 58.5 ($\pm$2)°F. Again, these more reliable results are clearly shown in the sliding analysis plots.

Going back to the raw data provided no evidence of any data anomaly. Based on evidence in the previous case of a low outlying data point being attributable to fill-up of only one of the two tanks, we hypothesized the following: that the first two of these three deliveries were fills of one tank only, while the third was a fill-up of both tanks. Furthermore, unlike the first and third, the second delivery in fact represented

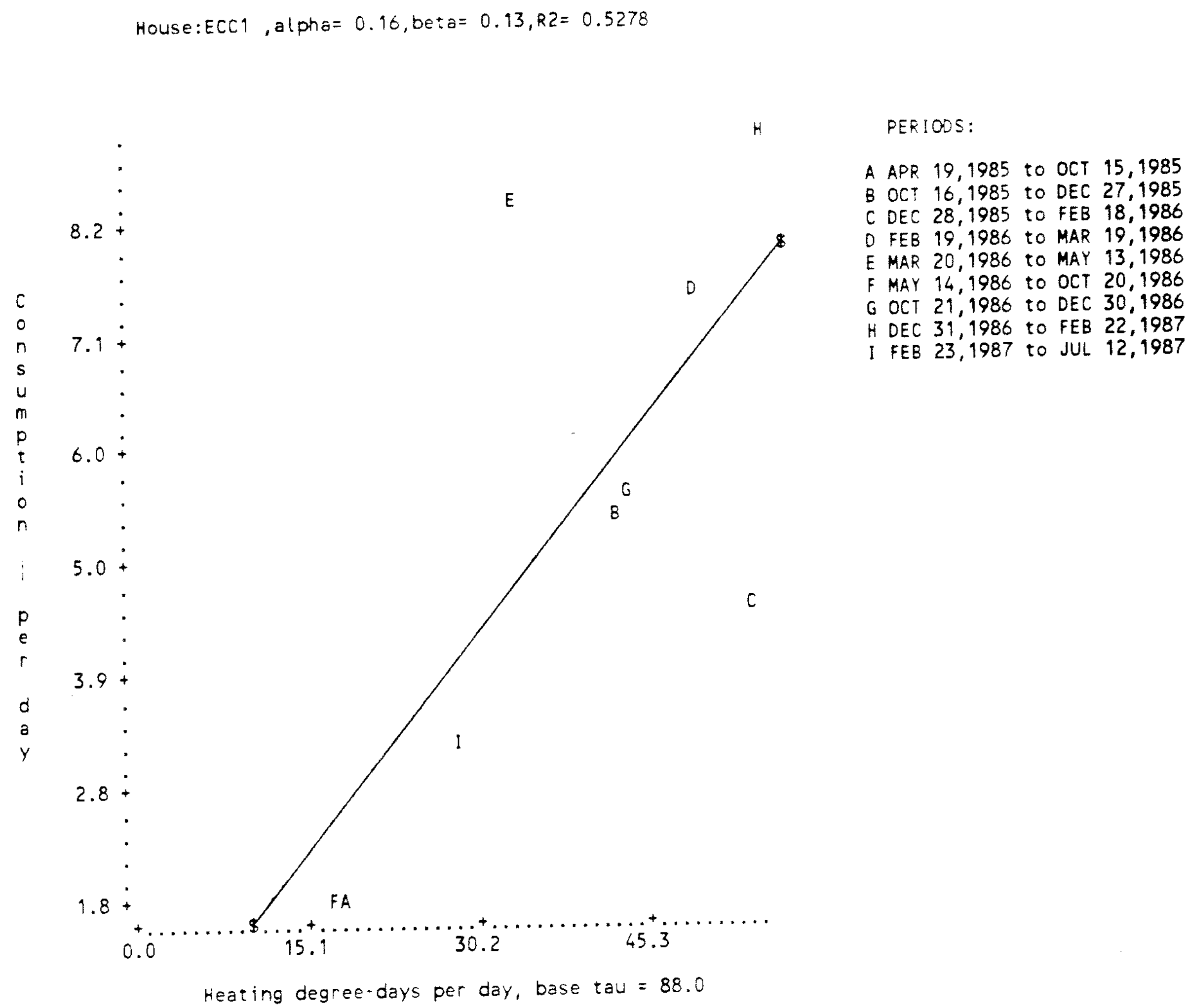


Figure 20. PRISM plot of consumption vs. heating degree-days (ECC Case #1) for unstable period during Summer 1986. Points C and E appear to be outliers and are combined with point D.

consumption because less than one tank of oil was consumed during that shorter-than-usual period (interval D was one month long, vs. one-and-a-half months for interval C and two months for interval E). The consumption data for these same periods support this hypothesis, as will be shown.

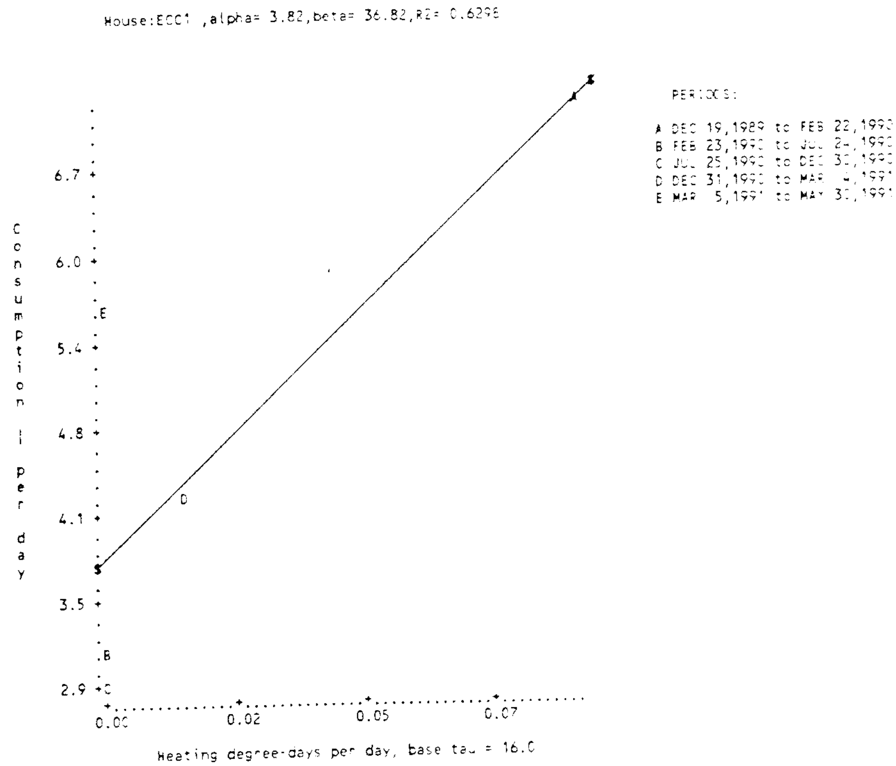
Most recent data: low-τ problem

Analysis of the most recent two-year period, December 1989 through May 1991, gave PRISM results of R^2 = 0.630, CV(NAC) = 0.12 and a low reference temperature of 16°F, with indeterminate standard error. The plot (Figure 21a) does not provide much insight into the data problems: because of the extremely low reference temperature, all but two of the data points lie on the vertical axis.

In order to view the information in a more useful form, we reran the PRISM analysis fixing the reference temperature at a more reasonable value, 65°F. The resulting plot, shown in Figure 21b, indicates two consecutive outliers, points D and E, as well as an additional point, A, having a large effect on the regression line. Note the similarity between this plot and the one shown earlier in Figure 19 for the winter 1983 case. As before, we combine points D and E and obtain vastly improved results: R^2 = 0.9998, CV(NAC) = 0.01, and a reference temperature of 75($\pm$4)°F. (Note that the very high R^2 should be interpreted in view of the small number of data points; after data combination there are only four data points, the minimum required by PRISM.)

In many cases throughout this report as well as in earlier studies, combination of outlying data points has led to major improvements in R^2 and CV(NAC), with only a minimal effect on NAC. This low-τ case with few data points is an important exception: the data improvement changes the NAC enormously, by 22%, bringing it into line with the previous values of NAC shown in the sliding analyses. Comparison of the last NAC value in the sliding analyses in Figure 17b vs. Figure 16b, and Figure 17a vs. Figure 16a, shows this difference clearly. This is not a totally unexpected result,

a)



b)

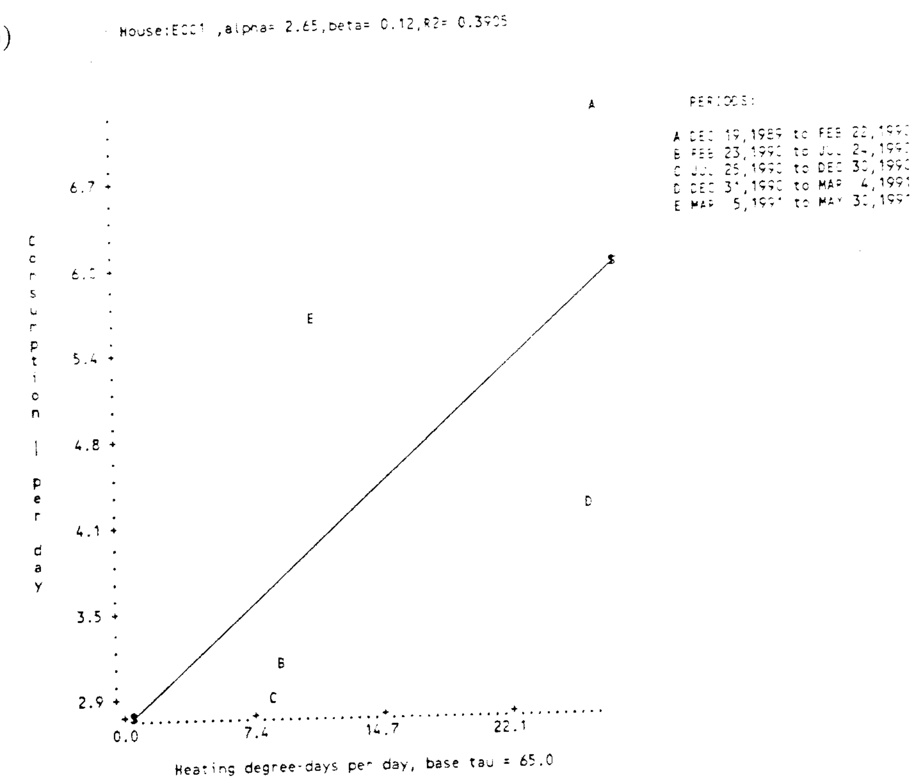


Figure 21. PRISM plot of consumption vs. heating degree-days (ECC Case #1) for period resulting in a very low reference temperature (16°F): a) for original data; b) for original data but running the PRISM analysis with reference temperature fixed at 65°F.

since the standard error of the NAC estimate from the low-τ PRISM run of the original data was indeterminate.

The fact that this is a low-τ case deserves further discussion. Since the reference-temperature estimate from the original data, 16°F, was only two degrees above the minimum daily average temperature observed in the time period December 1989 through May 1991, covered by the PRISM run, there are too few heating days in the time period of analysis (which had only four days with temperatures below 16°F), and in the 12-year normalization period as well, to give anything but erratic results.* (A very different situation occurs for high-τ cases, in which almost the entire analysis period becomes heating days; although the individual parameters are badly determined in such cases, the NAC nevertheless is likely to be highly reliable.)

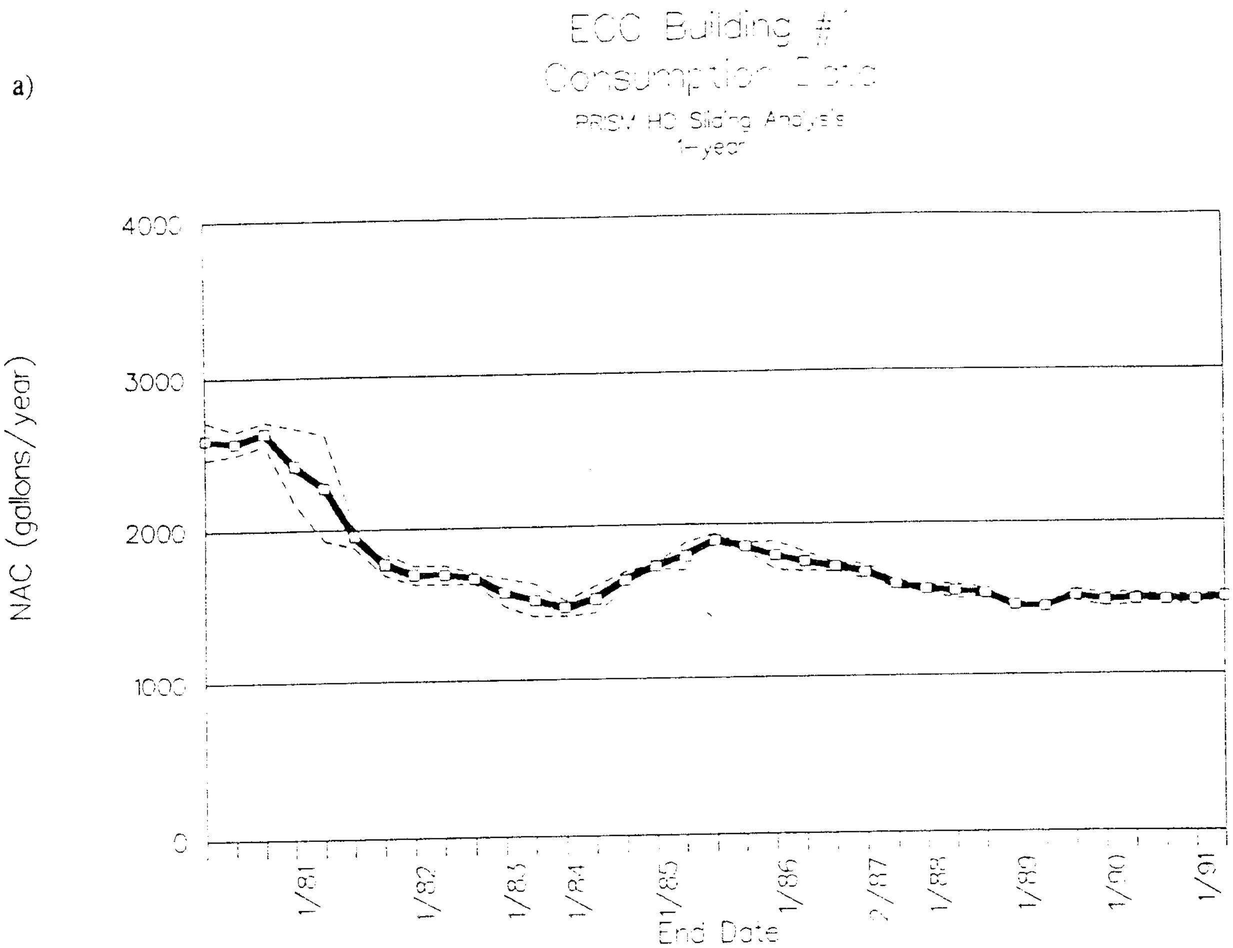
It is clear from this example that any low-τ case, within about 3°F of the minimum temperature, always should be examined for data problems. Although low-τ cases from the PRISM Heating-Only model are rare occurrences, the resulting instability of NAC is sufficiently troublesome to warrant further study of the problem. It is reassuring that, at least for this example, the problem was corrected by data combination.

Consumption (runtime-metered) data

In addition to delivery data for this building, actual consumption, based on furnace runtime meter readings, and converted to gallons of oil consumed, was provided for the same time periods. Figures 22ab show the sliding PRISM estimates of NAC using one-year and two-year periods, respectively. The resulting fit is extremely good for all periods, with the exception of the early periods during which conservation

*This low-τ case showed an interesting variation of R^2 with τ-value: $R^2 = 0.630$ (maximum value) at $\tau = 16$°F, declining to a minimum $R^2 = 0.28$ at $\tau = 40°$, and rising smoothly to $R^2 = 0.41$ at $\tau \geq 80°$ F. The NAC estimated at τ-values between 65°F and 88°F were within 5% of the apparently correct value obtained after data improvement, i.e., as shown in Figure 17b.

a)



b)

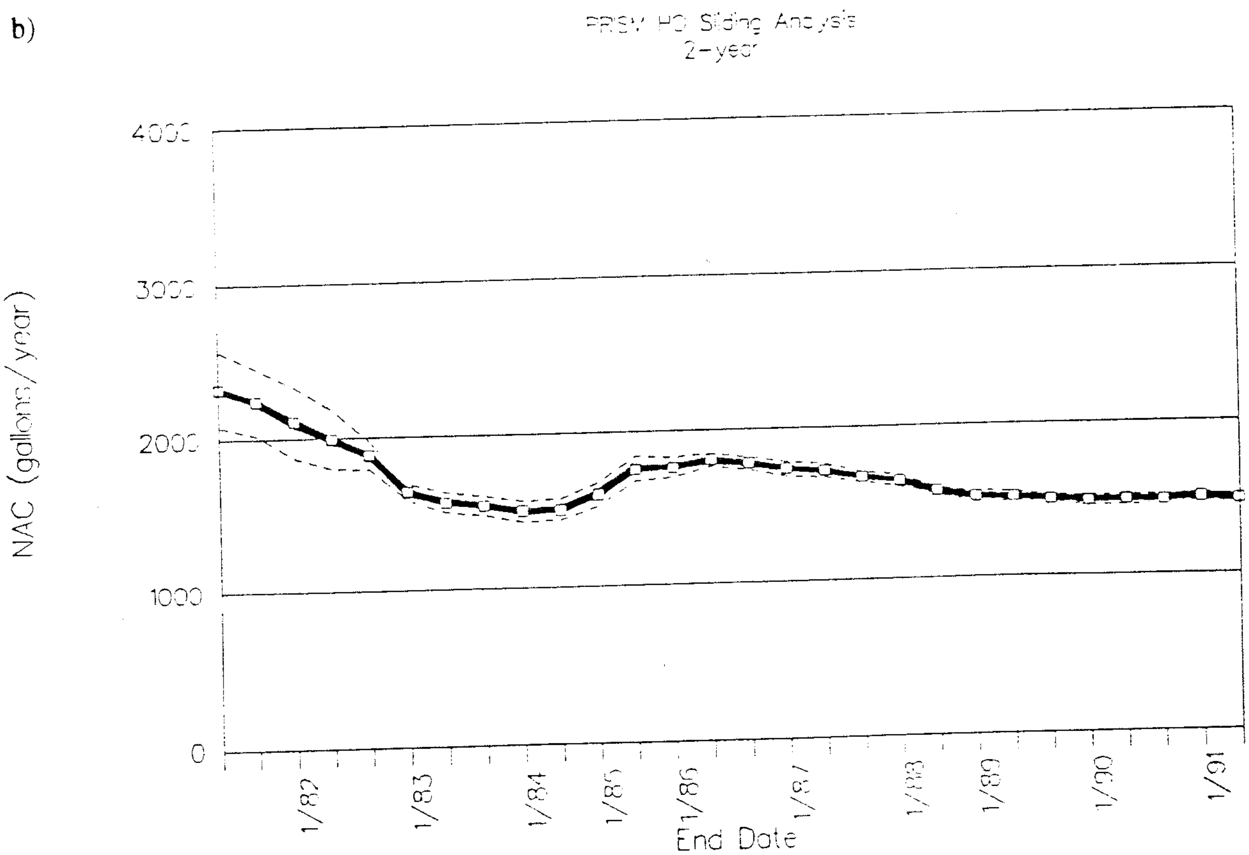


Figure 22. PRISM Sliding Analysis of furnace runtime measure of oil consumption for Building ECC Case #1 plotted for a) one-year analyses and for b) two-year analyses.

measures were implemented, as previously discussed. The extremely stable and reliable NAC results suggest that no data anomalies exist in the consumption (runtime) data. Additional evidence is in the extremely high R^2 values and low CV(NAC) over the entire time period: $R^2 \geq 0.86$ and CV(NAC) ≤ 0.15 for all runs in the one-year sliding analyses, and $R^2 \geq 0.97$ and CV(NAC) ≤ 0.07 following the winter 1981 period of consumption change. This is in contrast to periods of instability in the analogous plots (Figures 16ab) from the delivery data. Apparently, the anomalies in the delivery data are from deliveries that do not accurately represent consumption, e.g., from tank non-fills, rather than from periods of anomalous consumption. The good PRISM fits of the runtime meter data, however, lead to the conclusion that, when no major changes in the building are taking place, month-to-month fluctuations in this building's consumption are almost entirely explained by outside temperature.

We are fortunate to have a set of accurate consumption data for verifying the procedure for identifying data outliers and for justifying the combination of consecutive data points in the delivery data. Table 11 illustrates this for the winter 1983 case. The first delivery is anomalously low, apparently because the second tank was being cleaned, but the second delivery compensates so that the summed delivery accurately represents consumption over the two periods. This kind of result can be compared to a high-low reading more typical of gas and electricity data, where a meter-reading error may occur one month with the correct reading the following month that compensates for the original error. This oil consumption data set verifies that data combination in oil delivery data is physically justified, as it has been in earlier analyses of utility meter readings.

The consumption data give us the opportunity to investigate the underlying reason for the one unstable PRISM run not explained by conservation measures or data outliers: the PRISM run with ending date of February 1988 (see Figure 16a). Whereas analysis of the delivery data gave an R^2 of 0.999 and an NAC of 1491 gallons/year, with an indeterminate CV(NAC), the comparable analysis of consumption data gave $R^2 = 0.998$ and NAC = 1547 gallons/year, with CV(NAC) = 0.03. It is surprising that the delivery data do not show any evidence of a data anomaly. The consumption data, in fact, are extremely close in value to the delivery data.

Table 11. Sample comparison of oil delivery and consumption (runtime-metered) data.

Period	Oil Delivery (gallons)	Oil Consumption (gallons)
Dec 18-Mar 2	247	499
Mar 3-May 4	498	231
Total (Dec 18-May 4)	745	730

The only obvious difference between the delivery and consumption data is one additional data point in the consumption data during the summer of 1987: a summer-only, July 1 through July 13, data point. Apparently, the owner reads the furnace runtime meter every June 30 regardless of oil delivery schedule. A sensitivity study of the influence of the addition of a summer-only data point on the quality of PRISM results was performed. Table 12 shows the following results:

- PRISM analysis of original delivery data;
- PRISM analysis of original consumption data;
- PRISM analysis of delivery data using best "τ" from run #2 (58.3°F) as a fixed reference temperature; and
- PRISM analysis of consumption data combining the two summer data points in June and July for a direct comparison with delivery data.

These results indicate that the presence of the additional summer-only information, which aids PRISM in pinning down the base-level consumption, can greatly impact the stability of all the PRISM estimates.

Table 12. Comparison of PRISM results for ECC Case #1, demonstrating importance of summer data.

		N	R^2	$\tau(\pm se)$ (°F)	NAC (gallons/year)	CV(NAC)
(1)	Delivery data (original)	4	0.999	79(ind*)	1491	ind
(2)	Consumption data (original)	5	0.998	58.3(±3.0)	1547	0.030
(3)	Delivery data (fixed tau)	4	0.998	58.3(fixed)	1523	0.029
(4)	Consumption data (combined Jun/Jul)	4	0.999	79(ind)	1505	ind

* "ind" = indeterminate standard error

To test the hypothesis that it is the presence of summer-only data and not simply the presence of an additional data point that leads to the improved results, several different combinations of consumption data were run through PRISM. Results with comparable reliability to those in row (2) of Table 12 were obtained when the summer-only data point was included, whether four or five data points were used, but less reliable individual parameters resulted when the summer-only data point was combined with the data point on either side of it. Similar results were obtained from analogous data that included the summer of 1990.

This case study verifies the importance of distinct summer information for a reliable determination of PRISM estimates.* It has been well established from previous studies of houses that summer data are essential for a reliable determination of base-level consumption, and thus for reliable PRISM runs in general (Rachlin et al., 1986). These analyses clearly show that this result extends to multifamily buildings as well. Consumption periods that include summer months are essential for reliable PRISM runs, and summer-only information is likely to increase the reliability considerably. The lack of distinct summer information is explored further in the set of Staten Island buildings (ECC Case #3).

ECC Case #2: Pair of buildings in the Bronx

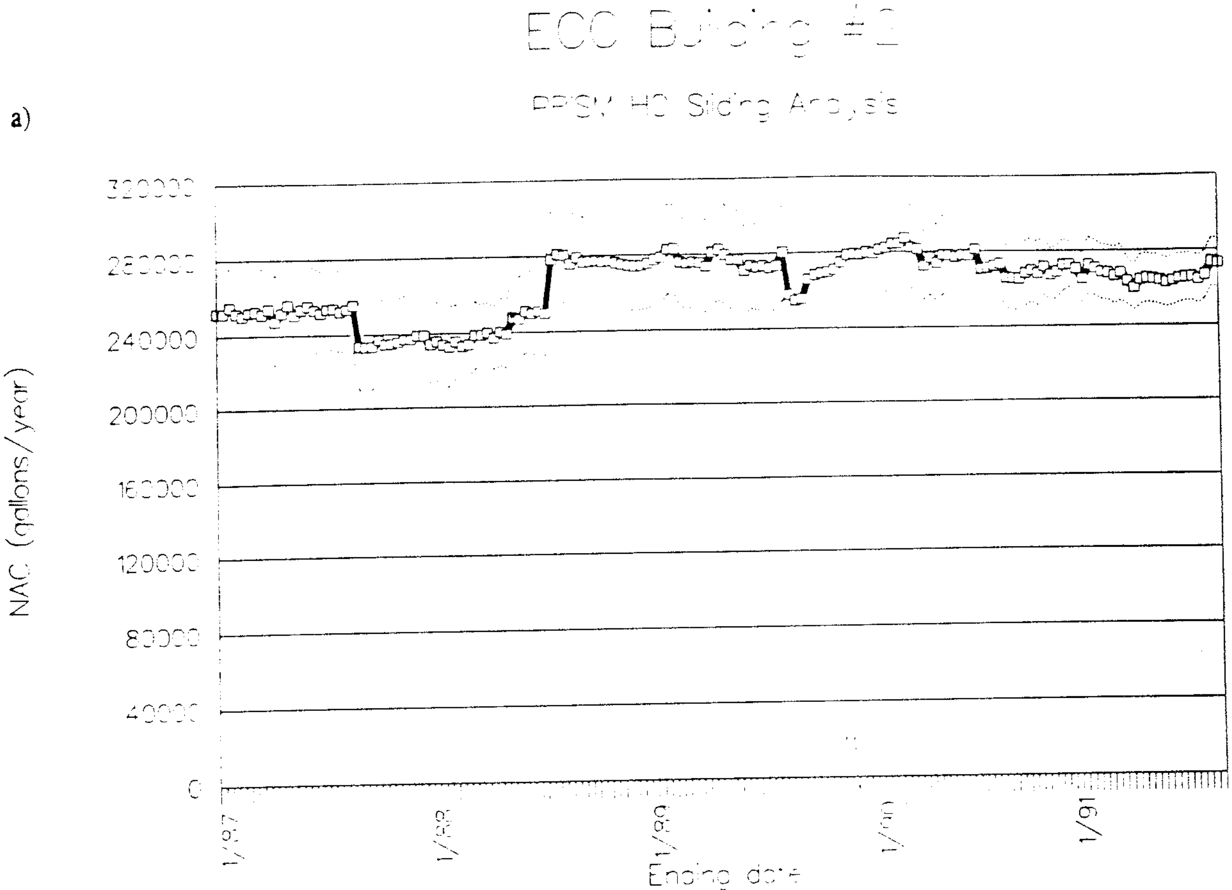
This data set corresponds to oil deliveries for a pair of buildings with 17 floors each and a total floor area of 379,000 square feet. The delivery data, supplied by Manager #2, span the period January 1986 through September 1991. This pair of buildings had as many as seven oil deliveries in one month, and provided the opportunity to study the benefits of different delivery data aggregations, especially monthly.

Figure 23a shows NAC from a PRISM Sliding Analysis of the frequent data. The average (mean) CV(NAC) for the entire period is 0.078, with the largest value at 0.101, and the smallest 0.044. The most outstanding features of this plot are the sudden consumption changes occurring in periods ending August 1987, June 1988, and

---

*In a preliminary analysis of the vacancy rate data from the BMD data set, Progress Report #4 for August through September 1991, the lack of complete overlap between the delivery data and vacancy-rate data led to an interesting comparison of 10-month, without summer, vs. 12-month results. The resulting CV(NAC) was markedly larger for several of the buildings when no summer data were included, even when the number of delivery data points was reduced by only one. For example, for building BMD10, CV(NAC) increased from 0.06 with 11 data points, including summer, to 0.24 with 10 data points, excluding summer. The standard error of the reference temperature increased as well. Thus, this study confirmed the importance of summer data in PRISM estimation.

a)



b)

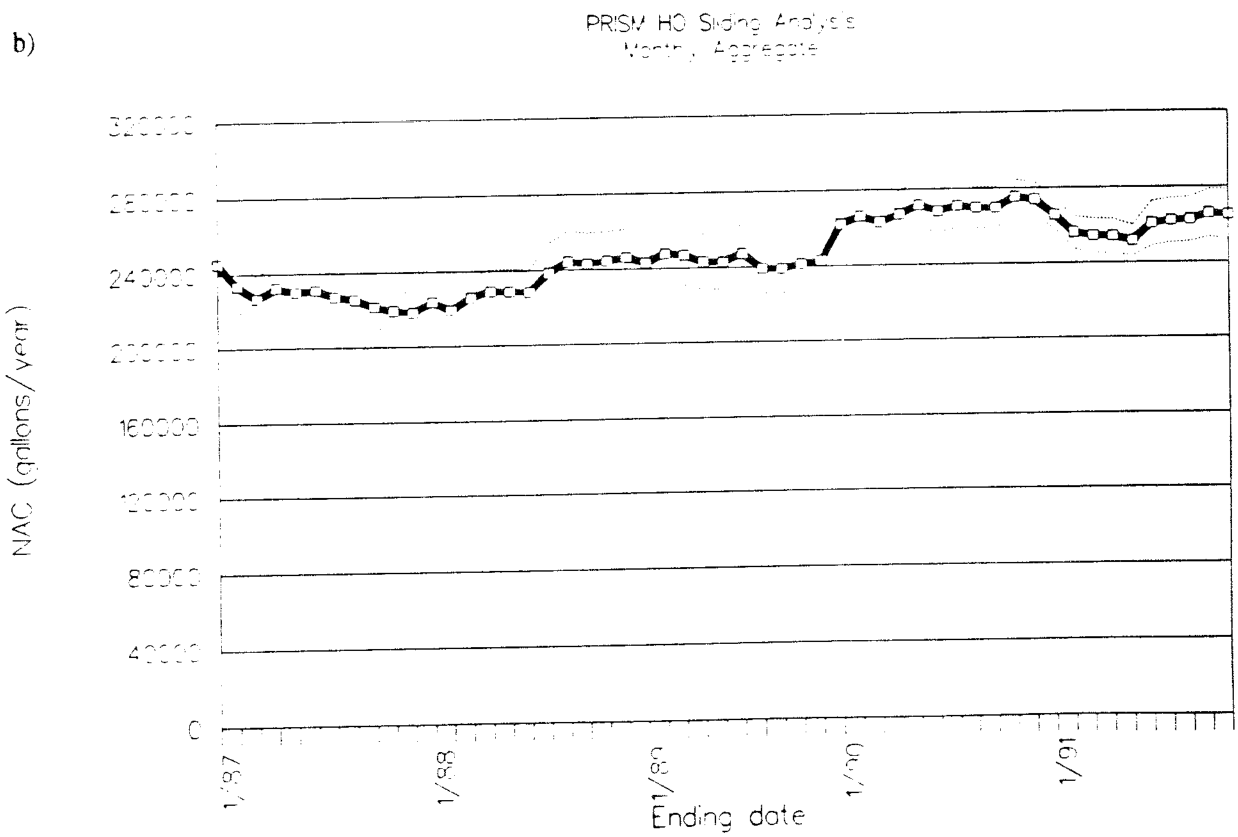


Figure 23. PRISM Sliding Analysis of oil delivery data from ECC Case #2 using a) original (more frequent) data and b) frequent data aggregated into monthly sums.

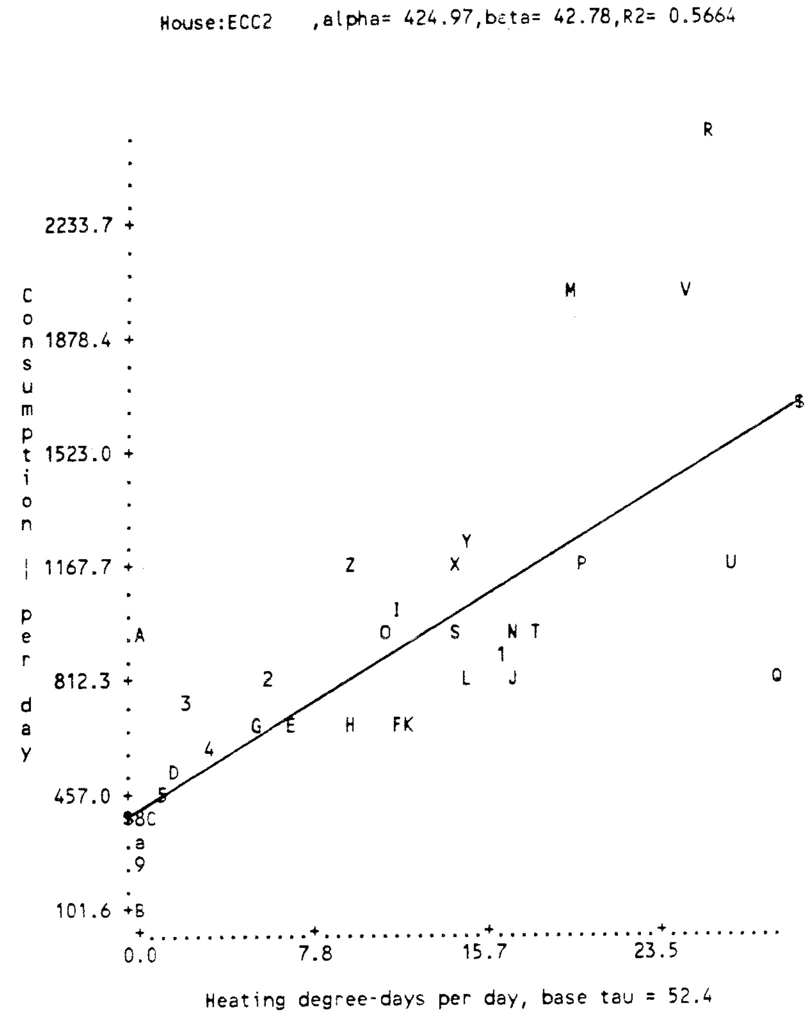
June 1989. A closer look at the individual PRISM analyses for these time periods reveals that in each case there is a single summer data point that is an outlier. The presence of this outlying data point results in a substantial increase in the estimate of the base-level component. For example, Figure 24 shows the consumption vs. heating degree-day plot for the time period August 22, 1986 through August 9, 1987. At the point in Figure 23a where the first large decrease occurs, the delivery data for the time period August 22, 1986 through August 28, 1986 is dropped, point A on the plot. The effect of not including this point in the second time period, August 29, 1986 through August 31, 1987, is a substantial decrease in base level, from 425 gallons/day to 265 gallons/day, which apparently has a substantial impact on NAC.

Figure 23b shows a different sliding analysis of the period from 1986 through 1991, with the delivery data aggregated into monthly sums. To compare with the more frequent data, the average CV(NAC) is 0.051, the largest is 0.074, and the smallest is 0.028. It is evident that the data are smoothed, especially in the three time periods that had data anomalies in the more frequent data. As a result, overall trends in energy consumption are seen more clearly. There appears to be an increase in consumption in more recent years: the highest NAC, in 1991, is 22% higher than the lowest NAC, in 1987. This rise in consumption, with relative standard error of 5.6%, appears significant.

An important lesson learned from this case study is the strong influence of a single summer data point that can be smoothed by using data aggregation. For oil data such as these, the variability caused by the "noisiness" of the more frequent deliveries seems to indicate the usefulness of monthly aggregation.

ECC Case #3: Set of four buildings on Staten Island

This set of oil delivery data for four separate buildings, also obtained from



PERIODS:

A AUG 22,1986 to AUG 28,1986
B AUG 29,1986 to OCT 9,1986
C OCT 10,1986 to OCT 22,1986
D OCT 23,1986 to NOV 3,1986
E NOV 4,1986 to NOV 13,1986
F NOV 14,1986 to NOV 20,1986
G NOV 21,1986 to NOV 27,1986
H NOV 28,1986 to DEC 4,1986
I DEC 5,1986 to DEC 10,1986
J DEC 11,1986 to DEC 17,1986
K DEC 18,1986 to DEC 26,1986
L DEC 27,1986 to JAN 1,1987
M JAN 2,1987 to JAN 4,1987
N JAN 5,1987 to JAN 9,1987
O JAN 10,1987 to JAN 15,1987
P JAN 16,1987 to JAN 20,1987
Q JAN 21,1987 to JAN 26,1987
R JAN 27,1987 to JAN 30,1987
S JAN 31,1987 to FEB 4,1987
T FEB 5,1987 to FEB 10,1987
U FEB 11,1987 to FEB 15,1987
V FEB 16,1987 to FEB 18,1987
W FEB 19,1987 to FEB 24,1987
X FEB 25,1987 to MAR 1,1987
Y MAR 2,1987 to MAR 5,1987
Z MAR 6,1987 to MAR 10,1987
1 MAR 11,1987 to MAR 17,1987
2 MAR 18,1987 to MAR 24,1987
3 MAR 25,1987 to APR 1,1987
4 APR 2,1987 to APR 9,1987
5 APR 10,1987 to APR 21,1987
D APR 22,1987 to APR 30,1987
5 MAY 1,1987 to MAY 12,1987
8 MAY 13,1987 to MAY 26,1987
9 MAY 27,1987 to JUN 24,1987
9 JUN 25,1987 to JUL 20,1987
a JUL 21,1987 to AUG 9,1987

Figure 24. PRISM plot of consumption vs. heating degree-days for ECC Case #2 for time period August 22, 1986 - August 9, 1987. This plot shows the effect of a single outlying summer data point (point A) on the PRISM estimates.

Manager #2, span the period January 1986 through October 1991. Each building has 13 floors for a total area of 519,000 square feet. There are typically no more than three deliveries during a single month in the heating season.

In the ECC report for these buildings, installation of a separate gas-fired hot water heater in 1983 was noted. The anticipated effect on a PRISM analysis of the oil data for these buildings would be an estimate of base-level consumption near zero.

A PRISM Sliding Analysis was performed on each of the four buildings. In every year of data, for each building, at least one data combination was clearly indicated, generally by an anomalously high one-day delivery preceded by a low delivery; a substantial improvement in R^2 and CV(NAC) resulted from the data combinations. In building #4, for example, combination of points A and B in Figure 25 dramatically transformed a terrible fit into a very good fit: R^2 increased from 0.04 to 0.87, and CV(NAC) decreased by 90%, from 0.59 to 0.07. This improvement is also evident in the sliding analysis plots for this building where the second large increase in consumption (Figure 26a) is smoothed after combining these two data points (Figure 26b).

After data combination in the one-year analyses, all CV(NAC) values for building #1 were 0.08 or below, and R^2 values were above 0.75. For all four buildings, the worst CV(NAC) was 0.15, and the lowest R^2 was 0.64.

In spite of the generally reliable NAC estimates, two of the individual PRISM parameters, base level and reference temperature, were generally not well determined. As is evident from the one-year results in Table 13, the reference temperature estimate is often unreasonably high, at or near the maximum daily temperature of the analysis period (the resulting standard error is indeterminate).* The concomitant base-level

*When the reference temperature has an indeterminate standard error, the standard error of NAC is somewhat underestimated.

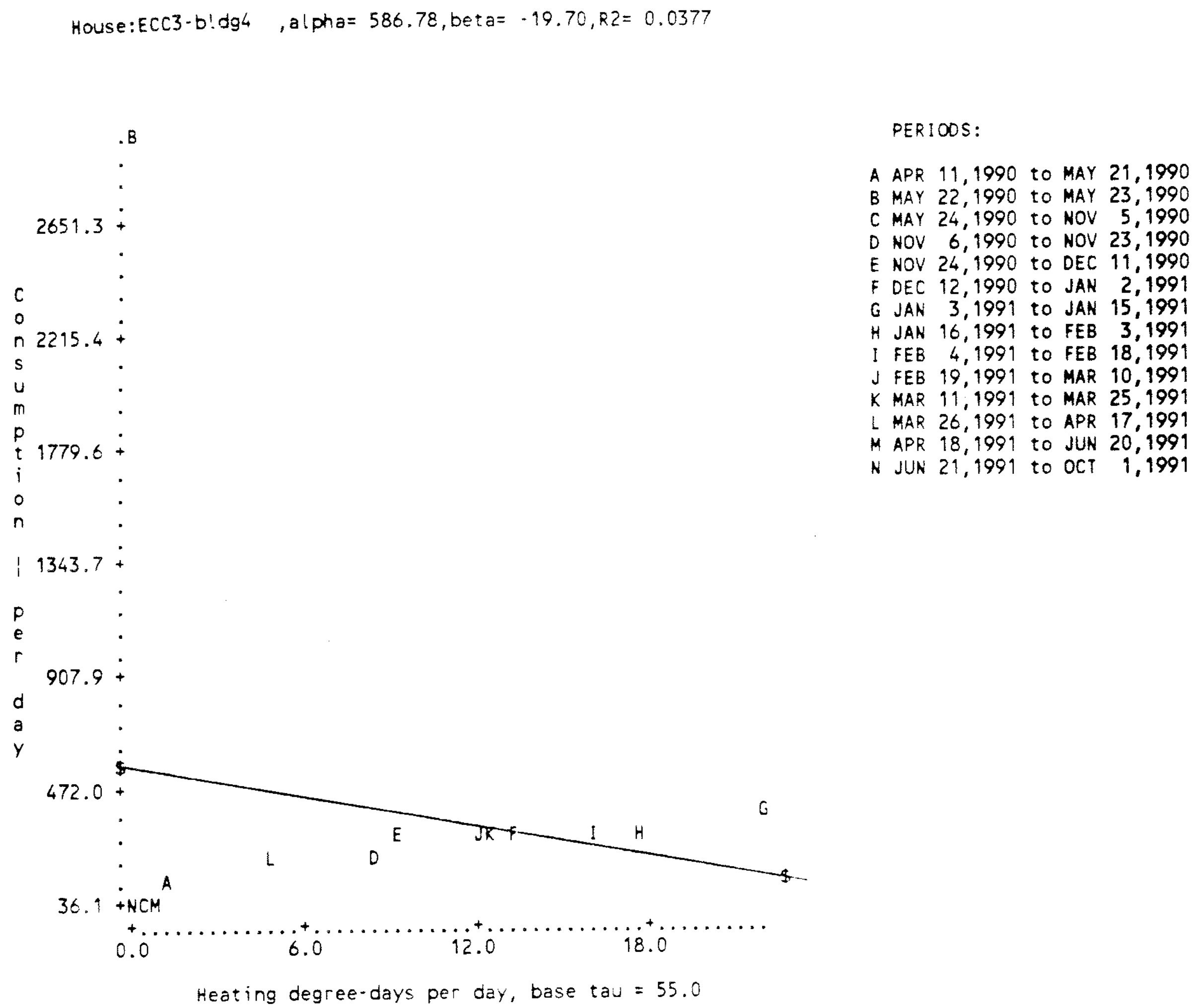
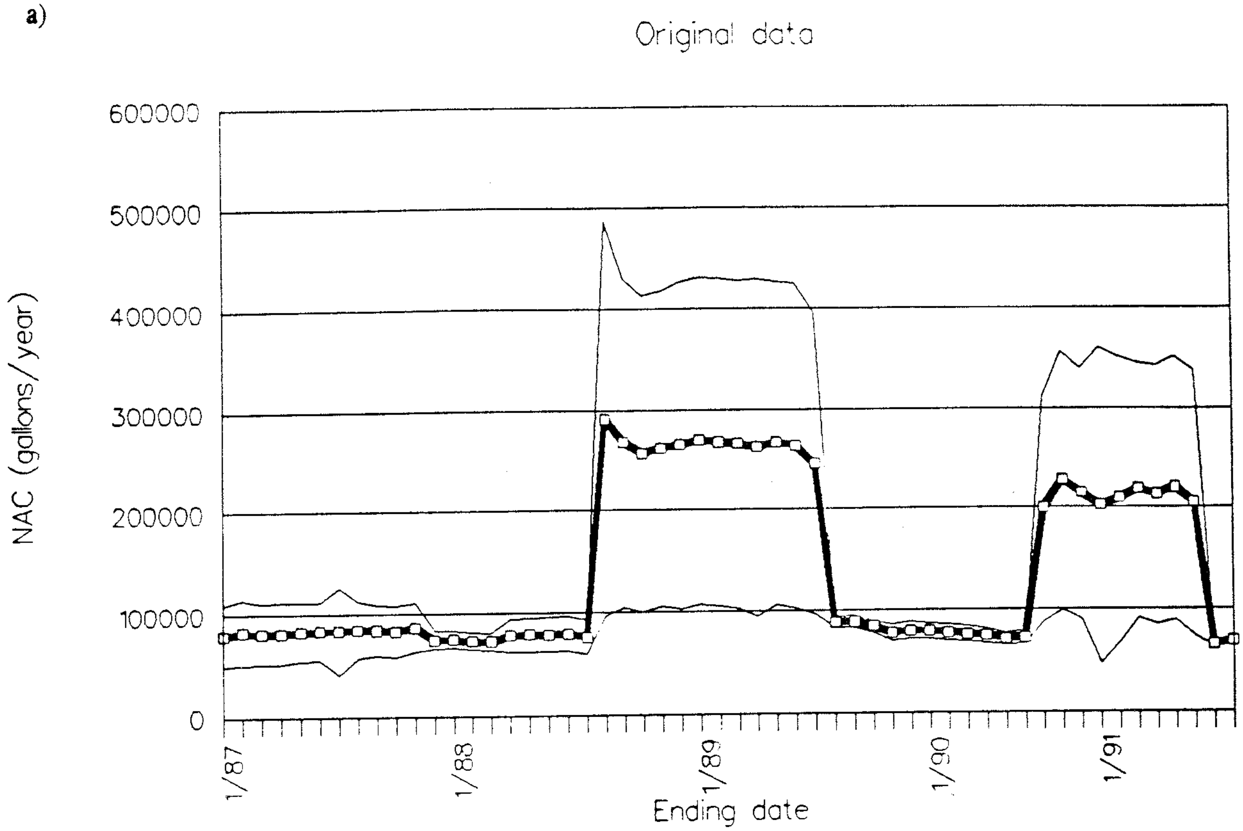


Figure 25. PRISM plot of consumption vs. heating degree-days for building #4 from ECC Case #3. Combination of outlying points A and B is indicated.

Staten Island Building #4
PRISM HO Sliding Analysis

a)



b)

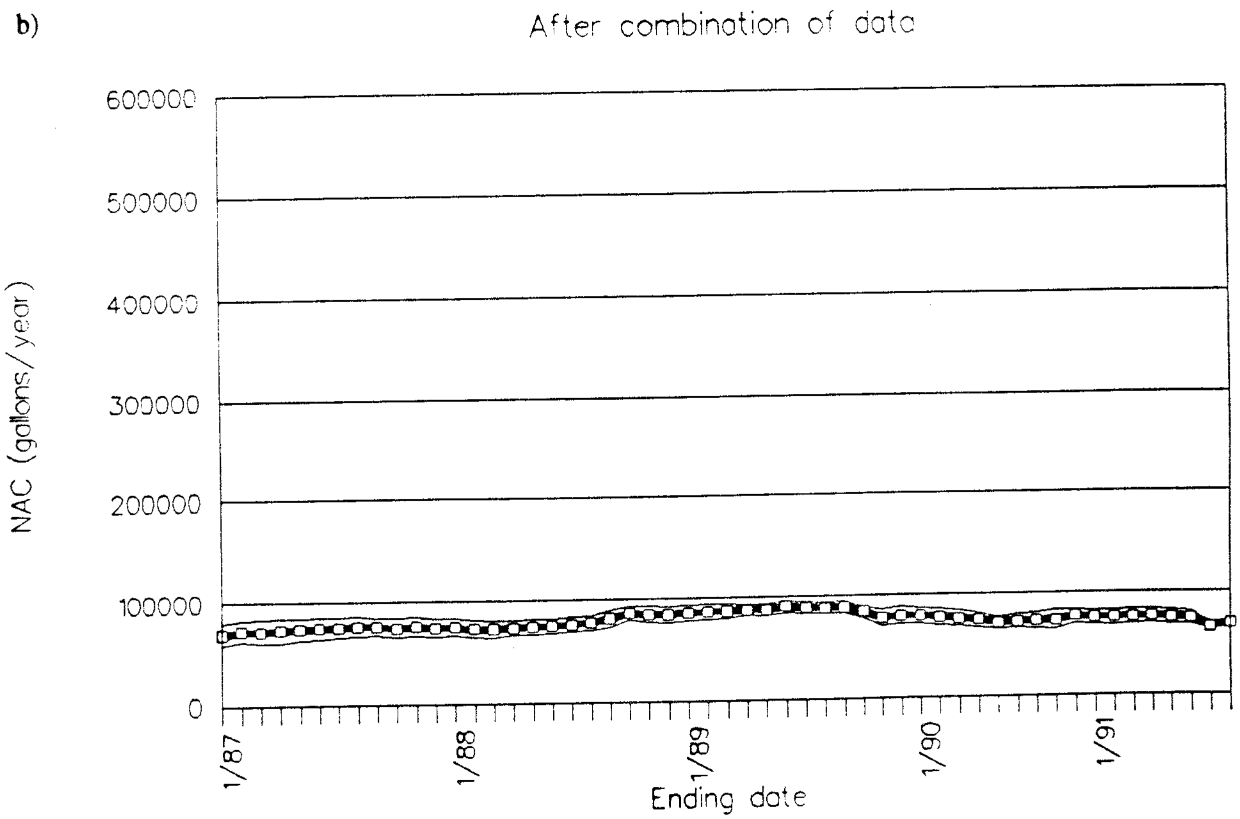


Figure 26. PRISM Sliding Analysis of building #4 from ECC Case#3 of a) original data and b) after data improvements.

Table 13. Summary of one-year PRISM analyses for ECC Case #3.

	N	R^2	Tau(se)* (°F)	NAC (gallons/year)	CV(NAC)
Bldg #1					
1986-1987	11	0.975	87(ind)	72142	0.03
1987-1988	13	0.824	86(ind)	80753	0.08
1988-1989	16	0.753	90(ind)	86741	0.07
1989-1990	10	0.904	86(ind)	69584	0.07
1990-1991	14	0.916	68(9)	77454	0.06
Bldg #2					
1986-1987	14	0.817	73(18)	81208	0.11
1987-1988	13	0.637	86(ind)	79400	0.12
1988-1989	16	0.824	90(ind)	98115	0.07
1989-1990	10	0.784	80(72)	64898	0.14
1990-1991	15	0.655	66(15)	95251	0.09
Bldg #3					
1986-1987	11	0.909	87(ind)	74232	0.07
1987-1988	15	0.839	74(17)	92553	0.09
1988-1989	17	0.779	90(ind)	94091	0.06
1989-1990	11	0.734	86(ind)	70149	0.12
1990-1991	13	0.915	75(13)	79652	0.06
Bldg #4					
1986-1987	13	0.736	76(41)	74466	0.15
1987-1988	14	0.836	79(39)	76058	0.10
1988-1989	16	0.810	90(ind)	89049	0.06
1989-1990	11	0.797	86(ind)	69557	0.09
1990-1991	13	0.874	85(149)	71990	0.07

---

*"ind" indicates indeterminate standard error.

estimates were also badly determined, with either negative or positive values considerably lower than the standard errors, indicating that the expected near-zero value was within the error bars for the base-level estimate.

As seen in earlier studies (Fels et al., 1986), the strong correlation between the base-level and the reference-temperature estimates, where a decrease in base level may be accompanied by an increase in reference temperature, contributes to the strong stability of the NAC estimate. Therefore, although individual parameter estimates may be unstable, NAC is well determined; thus, savings estimates calculated from the NAC estimates still would be reliable. Without the data combination from identification of outliers, this reliability would not be possible.

5. GUIDELINES FOR FUTURE PRISM APPLICATIONS TO MULTIFAMILY BUILDINGS

The three data sets for this study offer a rich range of opportunities for assessing the quality of consumption data in large multifamily buildings, and confirm the usefulness of PRISM for analyzing these data. The lessons have been many. We have learned about techniques for detecting and correcting outliers, and have realized vastly improved results because of these techniques. Moreover, we have learned about the benefits of careful data analysis, particularly that highly reliable consumption indices may be retrievable from data for buildings that, under initial assessment, do not lend themselves readily to simple energy analysis. This finding has practical implications for evaluations of conservation programs, where sample attrition due to buildings with bad data can be a major problem.

In previous studies of single-family homes and small-commercial buildings (Reynolds and Fels, 1988; Reynolds et al., 1990), the value of careful data screening was evident, and even dramatic in a fraction of cases. This study shows that in large multifamily buildings, particularly those with oil heating, careful screening of original data often leads to dramatic improvements in the ability to monitor building consumption. In fact, data improvements were part of the PRISM approach for such a large number of the buildings in this study that integration of data screening into the PRISM procedure seems a worthwhile goal.

In this section, we present a succinct summary of the lessons learned thus far, for future energy analysts of multifamily buildings who want to use PRISM. The lessons are specific to multifamily buildings, and build on more general guidelines for PRISM analyses published elsewhere (Fels, ed., 1986; Fels and Reynolds, 1990). The approaches suggested here draw explicitly on examples from the three data bases presented earlier (Sections 2-4).

The following summary of available tools and problem types likely to be encountered is given in a style intended for easy reference by PRISM users, with brief descriptions of procedures and relevant examples from this study. The summary is intended to be used in conjunction with the PRISM documentation supplied with the PRISM software as well as with earlier sections of this report.

Basic tools and techniques to supplement PRISM analysis

Tool #1: Careful examination of original data

Probably the most important tool for maximizing the reliability of PRISM applications is careful examination of the original data: checking the original data entry if a possible error is indicated, and using good common sense to question whether an out-of-line consumption point is physically reasonable or a misrepresentation of consumption (due to a meter-reading error, estimated reading, non-fill oil delivery, transcription error, etc.). After PRISM has been run on the original consumption data, and after data corrections as well, examination of the consumption vs. degree-day plots (optional PRISM output) can lead to clear identification of data anomalies and outliers (e.g., that should be tested for high studentized residuals; see Tool #3). If the reference temperature estimated by PRISM is unreasonably high or low, near the maximum or minimum observed daily temperature in the time period of analysis, a consumption vs. degree-day plot with the reference temperature fixed at a mid-range value such as 60°F may be more informative (see Problem Type #6).

Procedure: Check data entry carefully; examine consumption vs. degree-day plot (optional PRISM output) from a) PRISM run on original data, and b) after data corrections or justified data combinations, from PRISM run on corrected data. Fixing the reference temperature at about 60°F may provide more useful plots if the PRISM-estimated reference temperature is anomalously high or low.

Examples:

- BEUTS Building 343 (Figures 2ab for original data and Figure 3 for consumption plot);

- BMD Building 13 (Figures 10ab for consumption vs. degree-day plots of original and corrected data); and

- ECC Case #1 (Figure 18) showing change in consumption, apparently before and after an energy conservation measure.

Tool #2: CV(NAC) vs. R^2 plots

To examine the reliability of PRISM runs for a set of buildings, e.g., for each year of data that is available, a crossplot of CV(NAC) vs. R^2 provides a useful snapshot of the reliability of that data set's PRISM results. Reliable results appear in the lower right-hand corner of plot, as defined by pre-established reliability cutoffs, and other cases needing further attention are easily identified.

Procedure: Run PRISM on data spanning a similar time period for each of a set of buildings (e.g., one year of data). Plot CV(NAC) vs. R^2 from each PRISM run. Decide on preliminary reliability criteria; for multifamily buildings, we recommend initial criteria of $R^2 \geq 0.6$ and CV(NAC) ≤ 0.15. Addition of corresponding lines to plot to show cutoffs may be useful. As time allows, examine consumption data for those buildings failing criteria (i.e., not falling in the lower right-hand corner of plot as defined by cutoff lines) to determine possible data improvements (see section on data improvement techniques).

Examples:

- Sample of buildings from BEUTS data set (Figure 4 from Run A of original data and Figure 8 from Run B of improved data); and

- Sample of buildings from BMD project (Figures 9ab from original vs. improved data, and Figures 11ab from monthly aggregations of data).

Tool #3: Studentized Residuals

As a result of this study, studentized residuals have been found to be a useful procedure for detection of outliers in PRISM analysis. The method of computation is described in the Appendix.

Procedure: Run PRISM on original consumption data and compute studentized residuals. If one or more large values are found, e.g., of magnitude greater than 2.0 for 12 data points, confirm presence of outlier(s) by examining consumption vs. degree-day plot, and follow procedure outlined in Problem Types #1 and 2.

Examples:

- BEUTS Building #360 (Figure 7a) showing huge studentized residual for a one-day delivery;

- BMD Building 13 (Figures 10ab); and

- ECC Case #1 (Figure 19 showing high/low outlier and Figure 20 showing triple outlier).

Tool #4: PRISM Sliding Analysis

When more than one year of consumption data are available, PRISM Sliding Analysis is a useful technique for viewing consumption trends over time. In addition, data anomalies are evident in the error bars of NAC.

Procedure: Run PRISM on yearly subsets of the consumption data, moving forward approximately one month, or one consumption period, at a time. For example, for a multiyear data set that starts with January of year one, i.e., Jan YR1, Jan YR1-Dec YR1 would correspond to run #1, Feb YR1-Jan YR2 to run #2, Mar YR1-Feb YR2 to run #3, etc. Generate plots of R^2 vs. time and NAC (± standard error) vs. time. Examine these plots for possibility of data errors or anomalies: increases in the error bars of NAC will occur naturally as consumption goes through change, due, for example, to energy conservation measures, but erratic increases in the error of NAC in the absence of consumption change are likely indications of data outliers or problems.

Examples:

- BMD Buildings 7 (Figure 12) and 8 (Figure 13); see discussion of PRISM Sliding Analysis in Section 3;

- ECC Case #1 (Figures 16ab from original data and 17ab from improved oil delivery data, and Figures 22ab from consumption (runtime-metered) data; see winter 1981 case for effect of conservation changes, and winter 1983 and summer 1986 for effects of data improvements);

- ECC Case #2 (Figure 23a from original data and 23b from monthly aggregate of data); and

- ECC Case #3 (Figure 26a from original data and 26b from improved data).

Data improvement techniques

Problem type #1: High/low pairs of outliers

A higher-than-expected consumption point adjacent in time to a lower-than-expected consumption point can be an indication of an estimated reading (for gas or electricity), a non-fill delivery (for oil), or a data error (any fuel). Combination of the two consecutive data points can lead to enormous improvements in the reliability of the PRISM estimates.

Procedure: Identify high/low outlier pair from examination of consumption vs. degree-day plots and from studentized residuals obtained from PRISM run on the original data. Assess whether combination of a high/low pair of consumption data points is justified, on the basis of whether the two data points are substantially distanced from the PRISM regression line relative to the other data points (see examples) and/or the studentized residual is +2 or higher for the high outlier and -2 or lower for the low outlier. (These values are guidelines for PRISM run on 12 data points. As described in text, and in the Appendix, this should not be considered a hard-and-fast rule in that the presence of less extreme outliers can affect the value of the studentized residual for a more extreme outlier.) Combine high/low outliers, if justified. Recheck the revised consumption data and revised PRISM results (using consumption vs. degree-day plots) both for the improvements and the presence of additional outliers.

Examples:

- BMD Building #13 (Figures 10ab) for consumption plots (before and after data combination) and Figure 9 and Table 4 for resulting summary showing improvement of Run B over Run A for BMD data base;
- ECC Case #1 (Figure 19 for high/low outlier and Figure 20 for triple outlier, validated by runtime-metered consumption data; see Table 12 and discussion in Section 4); and
- ECC Case #3 (Figure 25).

Problem type #2: Single outlier

In oil data, much more than for other fuel types, the presence of isolated outliers seems to be fairly common, and generally seems to correspond to one-day deliveries or non-fill deliveries spanning a short time period.

Procedure: Identify single outlier using consumption vs. degree-day plots and studentized residuals, as described for Problem Type #1. If studentized residual is larger in magnitude than 2, and/or if the plot indicates the data point as an obvious outlier, try combining outlying consumption data point with consumption for previous period. Recheck for presence of outlier.

Examples:

- BEUTS Building #360 (Figure 6, Figures 7ab and Table 2 showing improvement from data combination); see also "Buildings with at least one outlying data point," in Section 2; and

- ECC Case #2 (Figure 24) for single outlying summer data point.

Problem type #3: Oil delivery data without delivery dates

For buildings for which oil delivery data are designated by month but not by day of month, useful PRISM information may be retrievable by monthly aggregation.

Procedure: Aggregate multiple deliveries in each month, and enter data in usual PRISM format, specifying month and year but not day of month. Then assume delivery day for each month (e.g., '1' for first day in the month) and run PRISM; repeat for all possible days, i.e., obtain 31 PRISM runs corresponding to 31 assumed dates. Choose the best assumed date as the one for which the corresponding PRISM run has the highest R^2, and use those results.

For a simpler, though probably less reliable, procedure, use 15 as the assumed delivery date for single-delivery months and 30 as the assumed date for months with multiple deliveries, in which case, the sum of the multiple deliveries becomes the consumption data point for that month.

Examples:

- Results of experiment to study assumed dates (Figure 5 and Table 1); see also "Buildings with no meter reading (or delivery) dates" in Section 2 ;

- BEUTS Building #323 (in Table 1) showing improvement from masked dates; and

- BEUTS buildings with no delivery dates (indicated in Figure 4).

Problem type #4: Buildings with infrequent deliveries showing repeated values

Repeated values in oil delivery data (e.g., many deliveries equal to or close to 2000 gallons) may, but possibly may not, indicate non-fill deliveries and thus inaccurate representation of consumption. Of course, repeated values could represent consumption, if, for example, the oil dealer is summoned when the tank is near empty, and the resulting delivery corresponds to the capacity of the oil tank. Therefore, whether they lead to single outliers from short-period deliveries or to high/low pairs of outliers, or to consumption data modeled well by PRISM, needs to be carefully examined.

Procedure: Using procedures outlined for single and paired outliers (Problem Types #1 and 2), check consumption data for outliers. Combine consecutive pairs of consumption data, if indicated, and recheck.

Examples:

- BEUTS Building #360 (Figures 6 and 7ab); see "Buildings with deliveries appearing to be related to truck size," in Section 2.

Problem type #5: Buildings with frequent deliveries showing repeated values or other scatter in the data

Occasionally, the data are sufficiently frequent that the scatter can be removed by aggregation of data points into fairly even (e.g., monthly) time periods.

Procedure: Aggregate consumption data by approximately monthly periods. For example, for the PRISM data entry, assign latest delivery date in month to sum of the deliveries in that month. Check for presence of outliers in the first or last delivery in month; moving it to an earlier or later time period might remove the outlier effect.

Examples:

- BEUTS Building #323 (in Table 1) showing improved results from masking delivery dates. See discussion on buildings appearing to be related to truck size, in Section 2, and Run A and Run B on original data vs. Monthly Run A and Monthly Run B of BMD data (Figures 9ab vs. Figures 11ab, and discussion in text), in Section 3; and

- ECC Case #2 (Figure 23a from original data and Figure 23b from monthly aggregation of data).

Problem type #6: Anomalously low or high estimate of reference temperature ("Low-τ" or "High-τ" cases)

An anomalously low or high value of reference temperature (low-τ or high-τ cases, respectively) can result from a PRISM analysis, particularly when there are insufficient summer data for an accurate determination of base-level consumption. This problem is more likely to be seen in unevenly spaced and infrequent oil data than in gas or electricity data from monthly meter readings. Generally, when the PRISM Heating-Only model is used, the NAC from high-τ cases is well determined and usable, but the individual parameters are not. Probably the only situation in which NAC is suspect is in extremely low-τ cases, which seems to be a rare occurrence, probably remedied by data combination for outlier correction, or (as a last resort) by setting the reference temperature to a mid-range value such as 60°F. (In low-τ cases, the reference temperature is so near the minimum observed daily temperature that there are almost no days with heating degree-days, so that an accurate statistical model of consumption is impossible.)

Procedure: If the PRISM-estimated reference temperature (τ) is very near the maximum or minimum observed daily temperature (in which case its standard error will be huge or indeterminate), rerun PRISM with τ fixed at a mid-range value of 60°F and check the consumption vs. degree-day plot for presence of outliers. Identify and correct outliers, if indicated, using procedure for Problem Types #1 and 2. If no data improvement from outlier appears possible, vary the consumption period slightly to check the stability of NAC (see section on PRISM Sliding Analysis). Although individual parameters from high-τ or low-τ cases will not be reliable (as clearly indicated by their large standard errors), NAC for high-τ cases is likely to be reliable. The reliability of NAC for low-τ cases needs to be assessed.

Note: Further research is needed to ascertain the best treatment of low-τ cases, and, in particular, whether and in what situation a reliable estimate of NAC may be obtained. Note also that, when the standard error of τ is indeterminate, as it will be when it is at or very near the maximum or minimum observed temperature in the period of analysis, or if the PRISM run is based on fixed-τ, the standard error of NAC and corresponding CV(NAC) appearing in the PRISM output will be underestimated. This is because, in such cases, the estimation of se(NAC) does not encompass any uncertainty of τ.

Examples:

- ECC Case #1: (Figures 21ab showing low-τ case); see also "Most recent data: low-τ problem," Section 4;

- ECC Case #1: Table 12 showing importance of summer data, and high τ resulting from insufficient summer data; and

- ECC Case #3: Table 13 showing high-τ cases with apparently reliable NAC estimates in spite of badly determined individual parameters.

6. SUMMARY AND FUTURE DIRECTIONS

Starting with oil delivery or gas metered data provided by building managers or owners, this study has explored the usefulness of the original data and the benefit of painstaking and informed analysis in improving the data's usefulness. PRISM has been confirmed as a valuable tool both for data screening and for producing reliable indices of weather-adjusted consumption in multifamily buildings. Specific prescriptions for identifying data anomalies and for determining possible data corrections or improvements, developed as a result of this study, look promising for future analysts of energy consumption in large multifamily buildings.

Even after data improvements, the reliability of PRISM estimates for oil-heated multifamily buildings on average remains somewhat lower than has been seen in numerous studies of houses and multifamily buildings with gas heating, as well as of houses with oil heating. Previously recommended criteria of reliability, in terms of cutoff values for CV(NAC) and R^2, may need to be relaxed somewhat for large oil-heated buildings in order to retain a large enough fraction of the buildings as modelable. Although it is clear that occasional lack of summer data and non-fill deliveries interfere with PRISM model reliability, more work is needed to understand the extent to which the lower reliability seen in oil data is due to data timing, i.e., less frequent and unevenly spaced deliveries, as distinct from physically based problems such as non-fill deliveries. In spite of the lower reliability overall, the successful application of PRISM to a large fraction of the oil-heated buildings analyzed in this study, with reliable NAC estimates and high model R^2 statistics, is an encouraging indication that readily available consumption data may be sufficient for meaningful monitoring of energy conservation in large multifamily buildings.

Lessons derived from our analyses of the three data sets comprising this study have been summarized in the form of specific procedures for use in PRISM applications.

In a real-world evaluation, the analyst may have neither the time nor the facilities to review individual cases to the extent done here. On the other hand, the sample sizes of multifamily buildings participating in energy conservation programs are typically very small, especially in comparison with the usual large samples of single-family houses, so that extra time spent in analysis of each building may well be feasible. Numerous examples in this study have provided evidence that painstaking analysis can transform a seemingly not useful data set into one that can yield reliable and useful consumption indices, making the extra work for data improvement worthwhile.

The procedures developed here are intended to provide preliminary guidelines for multifamily building analysis. They will no doubt need to be revised after testing on additional data sets of multifamily buildings; their trial application by PRISM analysts will provide essential input to future refinements. A long-range research objective is to turn these preliminary procedures into unambiguous prescriptions and criteria for model improvements, e.g., in terms of minimum values of studentized residuals or required increase in R^2 values, to determine under what conditions data improvements are warranted. The appropriate set of criteria could then be established prior to data analysis, to ensure that the scorekeeping of energy savings be done with consistent rules and objective procedures.

REFERENCES

Belsky, D., et al. (1980) Regression Diagnostics: Identifying Influential Data and Sources of Colinearity, John Wiley, New York, NY, pp. 6-21, 42-51, 64-69.

Beschen, D. and M. Brown (1990) "Evaluation Plan for the Weatherization Assistance Program (Draft)," Oak Ridge National Laboratory, Oak Ridge, TN.

DeCicco, J., et al. (1986) "PRISM Applied to a Multifamily Building: the Lumley Homes Case Study," Energy and Buildings, Vol 9, #1-2, pp. 77-88.

Fels, M., ed. (1986) "Measuring Energy Savings: The Scorekeeping Approach," Special PRISM Issue of Energy and Buildings, Vol. 9, #1-2.

Fels, M. (1986) "PRISM: An Introduction," in Special PRISM Issue of Energy and Buildings, Vol. 9, #1-2, pp. 5-18.

Fels, M., et al. (1986) "Exploratory Scorekeeping for Oil-heated Houses," Energy and Buildings, Vol. 9, #1-2, pp. 127-136.

Fels, M. and C. Reynolds (1990) "Now That I've Run PRISM, What Do I Do With the Results?", Home Energy, September/October.

Gerardi, R., et al. (1991) "Short-term Test Methods for Predicting the Thermal Performance of Buildings", New York State Energy and Research Development Authority Report #1129-EEED-AEP-89, Albany, NY.

Goldman, C. and R. Ritschard (1986) "Energy Conservation in Public Housing: a Case Study of the San Francisco Housing Authority," Energy and Buildings, Vol. 9, #1-2, pp. 89-98.

Goldner, F. (1991) "Summer 1990 Operational Use Profile," Draft Report of Energy Use and DHW Consumption Research Project, Energy Management Associates, New York, NY.

Hewett, M., panel leader (1988) Proceedings from the Panel on Multifamily Building Technologies, Summer Study on Energy Efficiency in Buildings, American Council for an Energy Efficient Economy, Washington, D.C.

REFERENCES (continued)

Hirst, E. (1987) "Cooperation and Community Conservation," Final Report of Hood River Conservation Project, DOE/BP-11287-18, Bonneville Power Administration, Portland, OR.

Judd, P., et al. (1989) "Building Energy Use Tracking System: Final Project Report," New York City Department of Housing, Preservation and Development, New York, NY.

Myers, R. (1986) Classical and Modern Regression with Applications, Duxbury Press, Boston, MA., pp.84-87, 144-155.

National Oceanic and Atmospheric Administration (NOAA; 1970-1991) Local Climatological Data - Monthly Summary [New York City Central Park], National Climatic Data Center, Asheville, NC, monthly publication.

Rachlin, J., et al. (1986) "Seasonality of Non-heating Consumption and Its Effect on PRISM Results," Energy and Buildings, Vol. 9, #1-2, pp. 139-148.

Reynolds, C., et al. (1990) "Using Monthly Billing Data to Find Energy Efficiency Opportunities in Small Commercial Buildings," Proceedings of the ACEEE 1990 Summer Study on Energy Efficiency in Buildings, Vol. 10, pp. 10.221-10.232, Washington, D.C.

Reynolds, C. and M. Fels (1988) "Reliability Criteria for Weather Adjustment of Energy Billing Data," Proceedings of the ACEEE 1988 Summer Study on Energy Efficiency in Buildings, Vol. 10, pp. 10.236-10.251, American Council for an Energy-Efficient Economy, Washington, D.C.

Appendix. Studentized Residual Analysis*

This Appendix provides the formulation for the computation of studentized residuals (also called R-student statistics), which have been found to be useful for outlier detection in PRISM analysis. (For a detailed discussion of this topic see Belsky et al. (1980) and Myers (1986).)

The studentized residual (or externally studentized residual), t_i is given by:

$$t_i = \frac{e_i}{S_{(-i)}(1-h_{ii})^{1/2}}$$

where e_i = $y_i - \hat{y}_i$, the i^{th} residual from the PRISM fit of N data points

$S_{(-i)}$ = the square root of the residual sum of squares (RSS) if the i^{th} point (the suspected outlier) were deleted and the regression recalculated

h_{ii} = the i^{th} diagonal element of the hat matrix H, where $\hat{y} = Hy$. h_{ii} takes values between $1/n \leq h_{ii} \leq 1$. The larger the value of h_{ii}, the more influence point i has on the slope of the regression line. In energy data, a point will have a large value of h_{ii} if it has an extremely large or small number of average heating degree-days per day compared to other points in the data set.

From the formula, we can see that t_i will be large when:

a) the residual at point i is large (i.e., point i is an outlier);

b) point i has a large influence on the regression line slope (h_{ii} is high);

c) $S_{(-i)}$ is small, which occurs when the other points aside from point i are nearly colinear.

*The work reported in this Appendix was summarized by Wayne Rosen as part of his graduate research project, which was supervised by Professor Richard DeVeaux, Department of Civil Engineering and Operations Research, Princeton University.

Statistical theory states that given n data points, and assuming the PRISM model is the correct model, the studentized residual for a single data point has the t-distribution on n-3 degrees of freedom (since the PRISM model uses three parameters). Thus we can test

H_0: $t_i = 0$ (the studentized residual at point i is not statistically different from 0)

H_1: $t_i \neq 0$ (the studentized residual is statistically different from 0; point i may be an outlier or point of high influence).

Thus, t_i can be used to test whether a data point is an outlier or point of particularly high influence. This is extremely useful because it offers a numerical means for outlier detection to complement the visual method of inspecting PRISM's residual plots (and could possibly be automated in the future). A second benefit is that it may detect points that are not necessarily outliers, but nevertheless have a high influence on the slope of the PRISM regression analysis (β). (An example of this is demonstrated in the section on ECC Building #1 entitled "Most recent data". In this example, because PRISM chose such a low value for the reference temperature, one data point was determining the slope of the regression line. See text for a more thorough discussion of this analysis.)

In order to use the studentized residuals effectively, an appropriate level for α (the probability that given H_0 to be true, $|t_i| \geq t_{\alpha,n-3}$) must be chosen. (Note that the α described here is not to be confused with the PRISM parameter α, or base-level consumption.) For single outliers, it appears that $\alpha = 0.10$ can be used to detect a weak outlier and $\alpha = 0.05$ to detect an obvious outlier. However, for detecting consecutive high/low combinations, we can make α less strict. For example, if $\alpha = 0.20$, and given that H_0 is true for both high/low points, the chance that two consecutive points would have studentized residuals with opposite sign and whose absolute values are larger than $t_{\alpha, n-3}$ is (0.20)x(0.20)x(1/2) = 0.02. Thus, criteria for high/low detection will be less strict than for single outlier detection. This is important because we have found that a high/low combination may go undetected in the presence of a third outlier in the same data set if α is too small. The third outlier sharply increases $S_{(-i)}$ for both high/low points, resulting in smaller values for their studentized residuals.

Interpretation of studentized residuals must also be in conjunction with the results of the PRISM regression analysis. In situations where the model fits very well, t_i is often too sensitive, resulting in a large value for a particular studentized residual for a small aberration. Therefore, we recommend relying on the studentized residual for PRISM fits with $R^2 < 0.90$. An example is the after-combination PRISM fit for

BMD13 in Fig. 10b, for which a not-extreme outlier (point F) has a very high studentized residual (almost 4.0). As can be seen in this plot, the fit of the data is very good (R^2 = 0.90). Thus the small deviation of this single point in this regression line results in an oversensitivity in studentized residual results.

One final word of caution on these residuals: if all points fall on a perfect line, with the exception of point i, t_i will be infinite since $S_{(-i)} = 0$. This could occur not only for a set of "perfect" data, but also in the case where tau is found to be low and one point is determining the slope of the PRISM regression line (thus all other points lie directly on the y-axis at zero heating degree-days). If an infinite studentized residual does occur, we recommend a close look at the plot of consumption vs. heating degree-days to determine what the PRISM fit looks like.