Enhanced Billing Analysis in Multifamily Buildings: Making Good Use of Available Data

Margaret F. Fels and Cathy L. Reynolds
Princeton University

Available energy consumption data from 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 a large number of centrally heated multifamily buildings in New York City are used, and the results of PRISM applied to the data are analyzed. Emphasis is on the problems encountered in oil delivery data and possible data improvements that can result from careful data screening and application of statistical techniques for outlier detection. By combining PRISM with these techniques, reliable indices of weather-adjusted consumption appear generally feasible for oil-heated as well as gas-heated multifamily buildings.

Introduction

Good energy information from buildings is needed both to identify energy conservation opportunities and to measure the actual savings achieved by whatever energy conservation actions are taken. In single-family houses, energy bills have proved to be a very 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 to single-family houses is far better understood than it is to multifamily buildings (Fels, ed. 1986). This study uses PRISM to explore the usefulness of available data in large multifamily buildings.

In multifamily housing, particularly low-income housing, acquisition of the needed 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 through 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 clearly warranted simply on the basis of costs. An added motivation for improved energy analysis techniques in large multifamily buildings is the vast energy conservation potential represented by that sector (Hewett 1988).

A two-fold objective of this study is to assess the quality, accuracy, and usefulness of available energy consumption data in multifamily buildings, and to determine the usefulness of PRISM for monitoring consumption and measuring energy savings in these buildings, particularly those with oil heating. A related objective is 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 (for example, when applied to a year of consumption data for a building), the PRISM results serve as verification of high-quality data. Furthermore, when PRISM does not work well, the method becomes a useful pre-processor of the data, an identifier of the 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.

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. Through studies of multifamily buildings in New York City sponsored by the New York State Energy

Research and Development Authority (NYSERDA), several rich data sets became available for testing this hypothesis.[1]

BEUTS Data Base

The first data base, obtained from the New York City Department of Housing, Preservation and Development (HPD), consisted of a subset of the 765 multifamily buildings that participated in the Building Energy Use Tracking System (BEUTS; see Judd et al. 1989). Each building's BEUTS report contained data on the building characteristics and reported fuel consumption. The subsample for this study was designed to include a sufficiently 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. The resulting subsample of 71 buildings included gas-heated (N=14) as well as oil-heated (N=57) buildings. Buildings with less than adequate data (e.g., with no oil delivery dates) were intentionally included as well. For PRISM analysis of the fuel consumption data, daily temperature data from the National Weather Service station in New York City were used (NOAA 1970-91).

As an example, the consumption data received for one oil-heated building (building #343), in PRISM format, are shown in Figure 1a. Running these data through the PRISM Heating-Only (HO) model gives the plots and the results shown in Figure 1b. 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) gal/year, i.e., the relative standard error of NAC, or CV(NAC), is only 7.3% of the estimate.[2] 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.

Note that the oil deliveries for this building are generally close to 3,000 gal, but they are spaced unevenly. The wide variation in per-day consumption (Figure 1b) illustrates this well. This is the opposite of what we see for gas-heated buildings: 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 was run on the data for all 71 buildings, for an initial assessment of the quality of the data and the reliability of PRISM applied to the data set. Two indicators of the goodness of fit from PRISM are CV(NAC), which ideally is very small, and the model's R^2-statistic, which ideally is close to 1.0 (its maximum value). Figure 2a shows a plot of CV(NAC) vs. R^2 for the initial PRISM runs on the 71 buildings (called "Run A").

The lower right-hand corner of Figure 2a, 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 cutoffs 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. A total of 51 (72%) out of the 71 buildings meet these reliability criteria when the original data as received were used.

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. Several types of data problems became evident:

(a) buildings with no meter reading (or, more commonly, oil delivery) dates specified (6 cases; the remaining 10 with no delivery dates modeled well, using a procedure of assumed dates described below);

(b) buildings which appeared to have deliveries related to size of delivery truck (e.g., in multiples of 500), and thus with tank not filled (7 cases);

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

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.

a)

```
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
```

b)

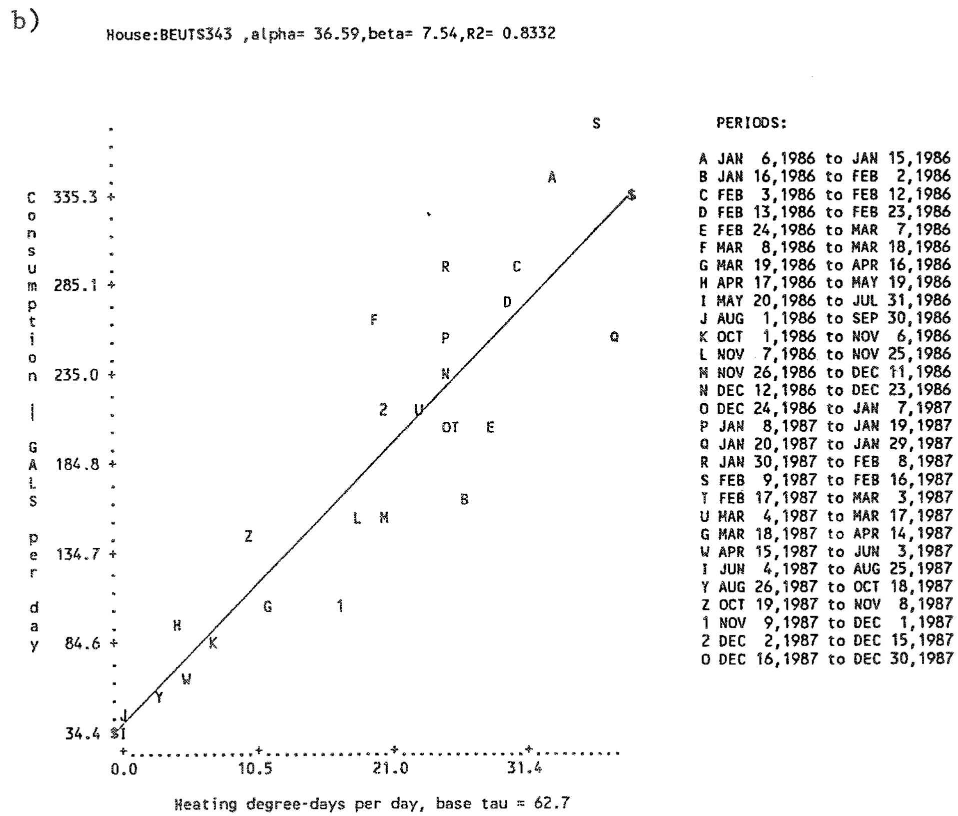

Figure 1. Sample of Consumption Data: a) Data for Building #343 from the BEUTS data base, in PRISM format; b) PRISM plot of consumption vs. heating degree-days, with PRISM fit of the data indicated by the line

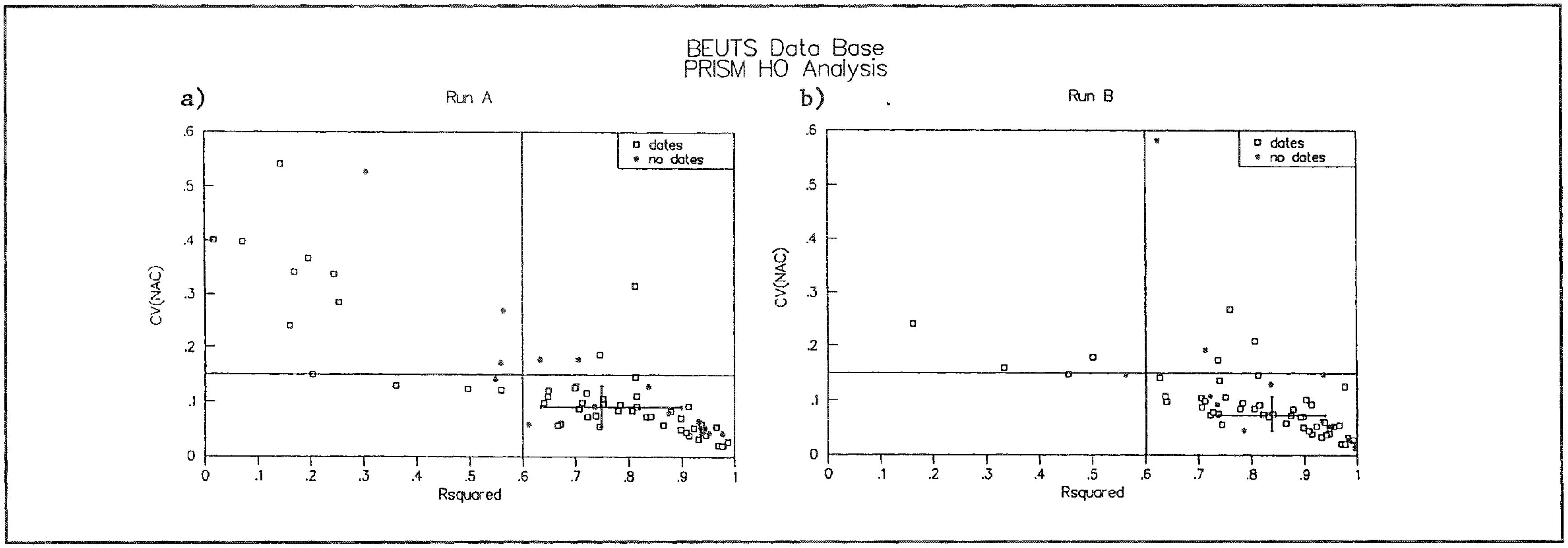

Figure 2. CV(NAC) vs R^2 for BEUTS Sample of 71 Buildings: a) Run A using original data; b) Run B after data improvements. Symbols distinguish buildings with meter reading or delivery dates from those without. (For the latter, assumed dates were used.) Superimposed on each plot are the quartile distributions of CV(NAC) and R^2.

Different types of improvements, described below, were explored for each of these problem types. The resulting improvements constitute "Run B", and are summarized in Figure 2b.

Problems Cases from the BEUTS Data Base

Buildings with No Meter Reading (or Delivery) Dates. For the 16 buildings in the subsample with undated consumption periods (11 with no oil delivery dates, and five with no gas meter-reading dates specified), a date for each consumption period needs to be assumed for the PRISM run. For the results indicated by a pound (#) sign in Figure 2a, 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 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. This procedure was applied to all 16 buildings.

The validity of assumed dates was tested using a sample of 21 buildings (15 oil-heated and 6 gas-heated) from the data set that had actual delivery or meter reading dates. 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, in order to determine suitable criteria for selecting the "best" set of assumed dates.

An examination of the PRISM results indicates that the assumed set of dates giving the highest R^2 value from the PRISM run in general produces an NAC close in value and in reliability to the NAC estimated from actual dates (see longer version of this report[1]). For the gas data, the NAC estimates from the "actual" vs. "best assumed" dates differ by more than 1% for only one out of five cases, and that one, which differs by 6%, corresponds to the case with the highest CV(NAC). For the oil data, only two of the 15 have NACs that differ by more than 6%. The building with the largest difference in NAC is also a building with reported deliveries in multiples of 500. Selection of the highest R^2 appears to be a useful criterion, 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.

Comparing the reliability of NAC, CV(NAC) from the best assumed dates is less than 20% different (on a relative scale) 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. For oil data, for which spacing of deliveries is very irregular, and for which deliveries do not always correspond to a fill up, use of assumed dates in several cases lowers the reliability, but in a few cases it increases the reliability.

In one such case, CV(NAC) improved from 11.3% for the actual dates to 6.5% for the best assumed dates, apparently because some of the deliveries were 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. 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.

With this procedure, useful PRISM results appear obtainable for buildings with monthly but undated delivery data. Finding the date of the month which gives the highest R^2 in most cases gives an NAC estimate close to the best value (i.e., to the value that would be obtained if the actual dates were known), often maintains the reliability of NAC, and in general appears to be a promising procedure for retrieving useful information from monthly delivery data for which delivery dates are missing. This, for example, could greatly expand the usefulness of energy consumption data such as that found in the BEUTS data base, for which delivery dates are missing for a large fraction of the oil-heated buildings. A comparison of Figures 2a and 2b shows the improvement resulting from this procedure applied to the 16 buildings with undated consumption data in the BEUTS subsample.

Buildings with Deliveries Appearing to be Related to Truck Size. A number of the buildings had "consumption" data which 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 amount of fuel because of budgeting. 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 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 8 out of the 14 cases with outliers): a single high outlying point covered a delivery of a very short period (four days or less), as short as one day in some cases. The original data 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 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 in Figure 3a, 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 3b), and a shift in the PRISM fit from highly unreliable to highly reliable (Table 1). 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.

Testing for Outliers

PRISM analyses which 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 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 (i.e., one 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" (e.g., from 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" is one such method that appears promising, both as a possible procedure for automating data screening and error checking, and for the immediate objective of improving the PRISM analyses of oil-heated buildings.[3]

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, and thus the original set of residuals is 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 value greater than about 2.0 in magnitude.[4] This provides a convenient method for detecting 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. 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 an oil delivery made shortly (usually 1-4 days) after another delivery.

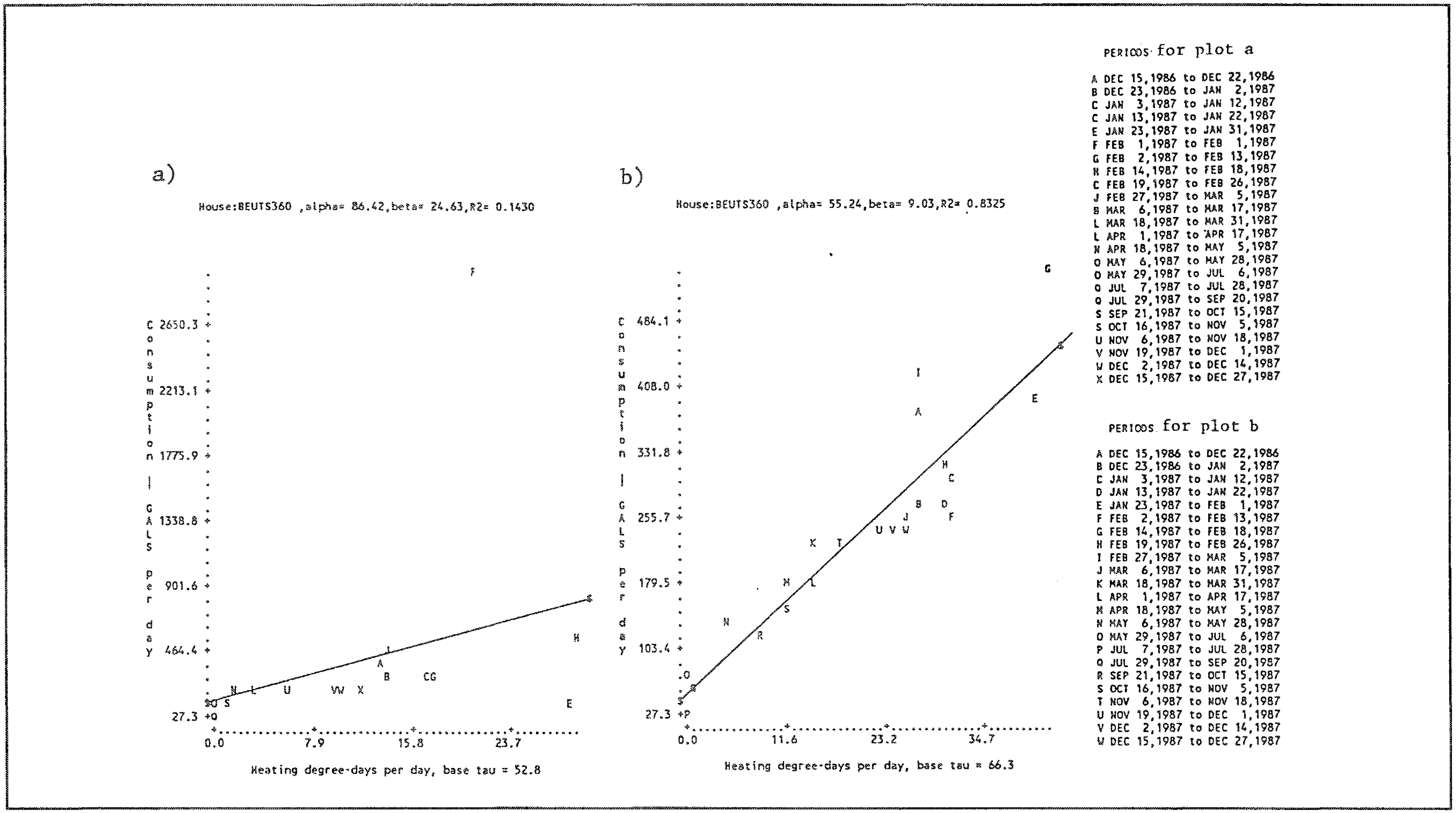

Figure 3. Example of Improvement from Outlier Correction: a) PRISM plot of consumption vs. heating degree-days for BEUTS Building #360, demonstrating One-Day Outlier (point F); b) consumption vs. heating degree-days for same building with outlier combined with previous day's delivery (point E). Results are in Table 1.

Table 1. Comparison of PRISM Results before and after Improvement from Outlier Correction; Oil-heated Building #360 (see also Figure 3)

	R^2	CV(NAC)	Ref. Temp τ (+se)
Before combining	0.14	0.54	54($\pm$22)°F
After combining	0.83	0.07	66($\pm$6)°F

In the preceding example (Figure 3 and Table 1), 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 example is a clear illustration of unambiguous outlier detection provided by the studentized residual, and the major improvement in PRISM fit (from unreliable to reliable) that can result from data combination for that outlier. We have computed studentized residuals 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.

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 possible outliers in the data.

The results of these analyses are shown in the plot of CV(NAC) vs. R^2 in Figure 2b (Run B). Comparison of this plot with Run A in Figure 2a and comparison of the quartiles indicated in each plot show clearly the substantial improvement in PRISM results. Using the same cutoff criteria, the number of reliable cases increases considerably, from 51 (72%) to 61 (86%) out of the 71 cases. Furthermore, several of them shift 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.

Building Monitoring Data Project

The second data set for this project was from thirty oil-heated multifamily buildings participating in a detailed monitoring project. This data set was ideally suited to the needs of this study, both because of an expected high level of data quality (the set of buildings is extremely well managed--all by the same company, and all operating on similar energy management systems), and because of the availability of very detailed monitoring data in addition to building-level consumption data (Goldner 1991).

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 analyses, a plot of CV(NAC) vs. R^2 was used to determine quality of the PRISM fit of these 30 buildings. Figure 4a shows the preliminary results (Run A).

Taking a closer look at the raw data for those buildings that did not model well (in particular, those with low R^2 and/or high CV(NAC) values), we found many instances with obvious outliers in the oil data, including outliers from "one-day deliveries", similar to those encountered in the BEUTS data base. Others show outliers in high/low pairs, as from a partial-fill followed by a complete-fill delivery. For example, for building BMD13, the studentized residuals for an obvious high/low outlier pair are -2.3 and 2.5 (vs. magnitudes of 0.6 or lower for all other points). 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.

After outliers were identified, PRISM was run on the combined data for all 30 buildings. The resulting plot of reliability criteria appears in Figure 4b (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)).

Since many of the buildings had very 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. This 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. The results, included in the longer version of this report, indicate that

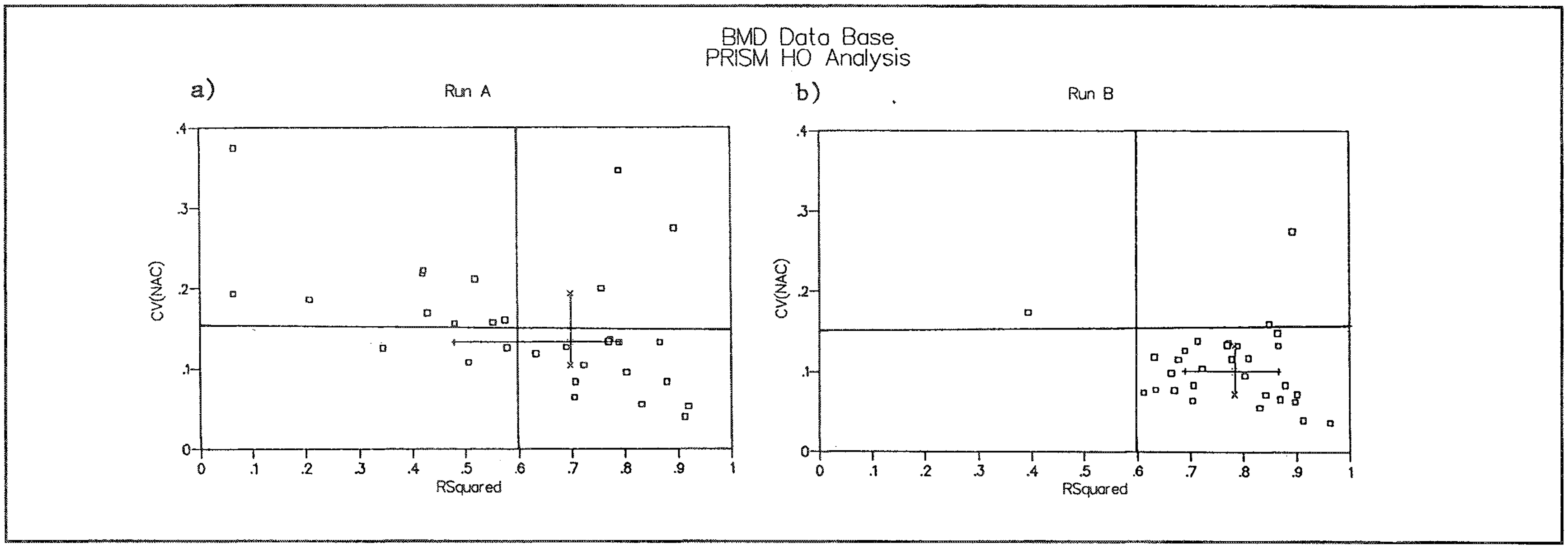

Figure 4. CV(NAC) vs. R^2 for 30 Buildings from Building Monitoring Data (BMD) Project: a) Run A (original data); b) Run B (after data improvements). Superimposed on each plot are the quartile distributions of CV(NAC) and R^2.

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.[5] 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.

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. Even these "good" data are not immune to data problems. Nevertheless, it is clear from these analyses that careful but simple treatment of outliers can yield substantial improvements in the PRISM results.

Energy Conservation Cases

A third data base, from 23 buildings that had participated in a mid-1980's HPD program of Energy Conservation Cases (ECC), included buildings with careful fuel delivery records, as well as records of energy conservation measures, over a number of years (Judd et al. 1989). Analysis of these data provided an opportunity to test the multifamily-building PRISM approach under optimal data circumstances, namely, a long time series of high-quality data, for large and small buildings, in which conservation effects may be evident.

Included here is a snapshot of the results for one of the buildings (ECC Case #1), whose data set was particularly interesting. The building's owner, who had kept oil delivery records dating back to 1979, provided CEES not only with the delivery data, but also with the records he kept on fuel consumption based on burner runtime, allowing a direct comparison between oil delivery data and actual oil consumption for the same consumption intervals.

The NAC results from a sliding PRISM analysis of the delivery data are plotted in Figure 5a. Although one-year PRISM analyses are generally preferred, these data are grouped in approximately two-year periods because of infrequent (less than monthly) deliveries.[6] A marked decrease in consumption near the beginning of the data series (1979), shortly after the current owner took possession of the building, is evident. In general, NAC is extremely well determined, but less stable NAC estimates are evident for certain short periods. From other studies (Fels, ed. 1986), we know that the error bars for NAC (and for all PRISM parameters) may be expected to increase substantially, and then settle to smaller values, as the analysis moves through a period of change: the first period of instability (Winter 1981) may be thus explained. Since NAC is so flat during the other periods of instability, other sources of instability such as data anomalies are suspected.

Identification of anomalous data points led to reasonable combinations of data points during these unstable periods. The results of the sliding analysis applied to the improved raw data are shown in Figure 5b. The effects of data smoothing are evident.

The analyses thus far have assumed that in general oil deliveries represent consumption (i.e., fill-ups of the oil tanks). The availability of furnace runtime metered data provides a direct check of this assumption. The results of running PRISM on the runtime-meter data (converted to gallons of oil consumed) are summarized in Figure 5c, which shows the sliding PRISM estimates of NAC using two-year periods. The resulting fit is extremely good for all periods, with the exception of the early periods during which conservation measures were implemented, as discussed above. 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: $R^2 \geq 0.82$ and CV(NAC) ≤ 0.11 for all runs, and $R^2 \geq 0.95$ and CV(NAC) ≤ 0.05 following the Winter 1981 period of consumption change. This is in contrast to the periods of instability in the analogous plots (Figures 5a-b) from the delivery data.

This analysis indicates that the anomalies in the delivery data are from deliveries that do not accurately represent consumption (e.g., from non-fills of the tanks), rather than from periods of anomalous consumption. This was verified for specific data points, corresponding to only one of two tanks being filled, leading to a low/high pair of delivery data points, and consumption (runtime) data that fit the PRISM model very well (see detailed report). Furthermore, an additional summer data point in several years of the runtime data gave a greatly improved PRISM fit, validating the importance of summer consumption information in PRISM analysis. The good PRISM fits of the runtime meter data overall lead to the conclusion that, in periods 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.

Summary and Future Directions

Starting from 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

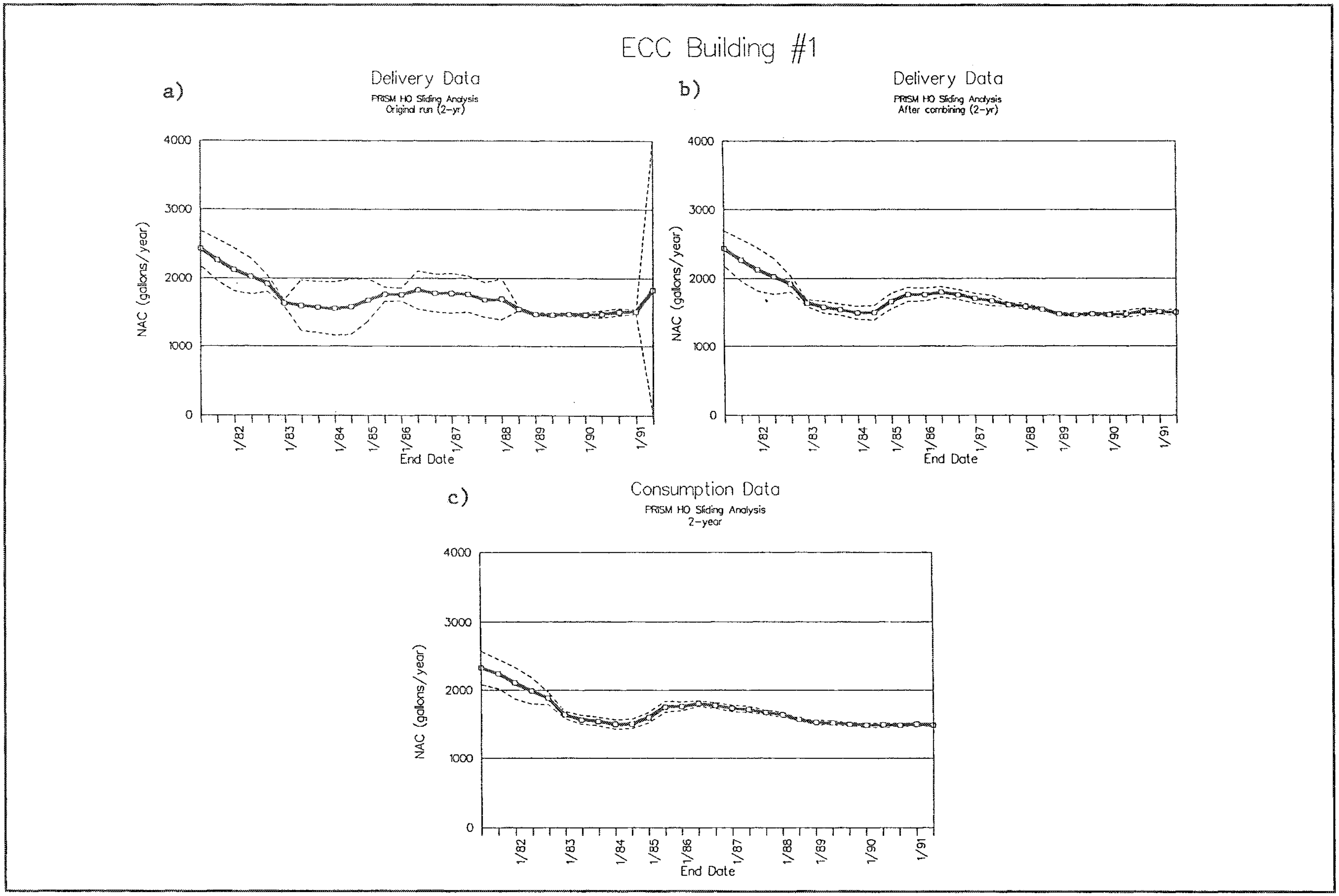

Figure 5. Results of PRISM Sliding Analyses of Data from Energy Conservation Case (ECC) #1: a) from original oil delivery date; b) from improved oil delivery data; and c) from furnace runtime measure of oil consumption. Two-year periods are used. The dashed lines represent standard errors of the NAC estimates.

usefulness. Augmented by new procedures for identifying data anomalies and for determining possible data corrections or improvements, PRISM has been shown to be a valuable tool both for data screening and for producing reliable indices of weather-adjusted consumption in multifamily buildings. These procedures should be useful for increasing the reliability and thus the sample size in single-family studies as well.

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. 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, and occasional lack of summer data, as distinct from physically based problems such as non-fill deliveries. Nevertheless, the successful application of PRISM to a large fraction of the oil-heated buildings analyzed in this study is an encouraging indication that readily available consumption data may be sufficient for meaningful monitoring of energy conservation in large multifamily buildings.

In a real-world evaluation, the analyst may not have the resources 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 typically large samples of single-family houses) so that extra time spent in analysis of each building may well be feasible. Future research is needed to refine these tools and to develop unambiguous criteria for deciding under what conditions data improvements are warranted (e.g., in terms of minimum values of studentized residuals, or minimum increase in R^2 values, etc.). 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.

Acknowledgments

This study was funded by the New York State Energy Research and Development Authority (NYSERDA). The authors would like to thank Norine Karins at NYSERDA for her valuable input and helpful comments on the progress of this research, and Fred Goldner of Energy Management Associates for providing much of the data and for answering numerous questions. We would also like to thank Mark Engel of Ralph Langsam Associates; Barbara Tillman of Grenadier Realty and George Birman for providing additional data; Wayne Rosen and Erik Nelson for painstaking data analysis; and Richard DeVeaux for valuable statistical input.

Endnotes

1. This paper is a summary of a more detailed report (Fels and Reynolds 1992).

2. Coefficient of variation CV(NAC) = [standard error(se) of NAC]/NAC, or, equivalently, the relative standard error of NAC, which is written as a ratio (e.g., 0.065) or as a percent (6.5%).

3. For statistical background on studentized residuals, see Belsky et al. 1980, and Myers 1986.

4. In some cases, a cutoff of 2.0 may be too stringent for detecting high/low readings. See the longer version of this report for guidelines on selection of cutoff criteria and for detailed results.

5. Note: Comparison of the PRISM R^2 values from data before and after aggregation should consider the fact that they reflect statistics based on different numbers of degrees of freedom.

6. See longer version for one-year results, and a detailed resolution of data anomalies for this building, as well as detailed results for six additional buildings.

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.

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.

Fels, M. and C. Reynolds. 1992. "Energy Analysis in New York City Multifamily Buildings: Making Good Use of Available Data." Draft Final Report to NYSERDA, PU/CEES, Princeton, NJ (longer version of this report).

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. "Multifamily Building Technologies: Panel 2." *Proceedings of the ACEEE 1988 Summer Study on Energy Efficiency in Buildings*, Volume 2, American Council for an Energy-Efficient Economy, Washington, D.C.

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. 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*, Volume 10, pp. 10.236 - 10.251. American Council for an Energy-Efficient Economy, Washington, D.C.

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*, Volume 10, pp. 10.221 - 10.232. American Council for an Energy-Efficient Economy, Washington, D.C.