PRISM: An Introduction

MARGARET F. FELS

Center for Energy and Environmental Studies, Princeton University, Princeton, NJ 08544 (U.S.A.)

(Received January 1986)

WHY KEEP SCORE?

In the past, programs designed to induce energy conservation in housing have nearly all been casual about their measurement of energy savings. In many cases, savings are unashamedly asserted without being measured: the monitors keep score with a yardstick scaled by the number of participants or number of dollars spent rather than the actual amount of energy saved, or they rely on engineering models which, lacking calibration to real-world experience, notoriously overestimate the actual savings. This is particularly distressing given that the single most important objective of these programs, the saving of energy, is intrinsically quantifiable and relatively accessible by means of energy consumption data recorded systematically for another purpose — billing. Furthermore, weather adjustments are easily made from readily available temperature data, so that effects of conservation need not be obscured by differences in weather from one year to the next.

The need for reliable scorekeeping in energy conservation is increasing. Many utilities in the U.S. have undertaken extensive retrofit assistance programs for their customers, not only because of the federal Residential Conservation Service (RCS) [1], which mandates nearly free energy audits for customers, but also because of a growing commitment to energy conservation as a utility investment strategy. RCS audits have reached millions of homes. The Low-Income Weatherization Program, federally funded but managed at the community level, is bringing to many more homes an extensive, often costly, set of retrofits in addition to the energy audit. New conservation strategies, involving monetary rewards for conservation actually achieved (for example, payments from the utilities serving the retrofitted houses to the conservation company carrying out the retrofits [2]) or shared-savings arrangements between building owner and energy service company, require the savings estimates to be both accurate and unambiguous. Finally, the homeowner, whether participating in these programs or acting independently, needs feedback on the effectiveness of his or her conservation investments.

Companies which offer conservation services invariably need help in informing the customer about how much energy — and money — his or her purchase is likely to save. With records of actual savings achieved, companies could understand, quantitatively, the value of what they sell. The resulting picture could become one of satisfied, savvy customers dealing with a company able to convey accurately the value of its own services.

IT'S EASIER THAN IT MAY SEEM

Perhaps surprisingly, keeping accurate scores on the actual amount of energy saved is straightforward, and the required data, whole-house meter readings and average outdoor temperatures, are readily available. The PRInceton Scorekeeping Method, PRISM, uses utility meter readings from before and after the retrofit installation, together with average daily temperatures from a nearby weather station for the same periods, to determine a weather-adjusted index of consumption, Normalized Annual Consumption or NAC, for each period. The procedure is depicted in Fig. 1(a). Analogous to (and, based on field measurements, clearly more accurate than) the U.S. federally mandated miles-per-gallon rating, the NAC index provides a measure of what energy con-

INPUTS: OUTPUTS:

Fig. 1. Schematic diagram showing the data requirements for the Princeton Scorekeeping Method (PRISM) and the estimates that result from it: (a) the basic procedure for one house; and (b) the procedure for calculating control-adjusted savings for a group of treated houses.

sumption would be during a year under typical weather conditions. Total energy savings are derived as the difference between NAC in the pre- and post-periods. A 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 in another.

PRISM is certainly not the first method to include weather normalization. In fact, the simple relationship between a house's energy consumption for space heating and outside temperature was recognized, in the published literature, at least 80 years ago [3]. Even before natural gas pipelines were available, weather information was crucial input to gas dispatching and production decisions [4]. In the current literature, there is a variety of methods that have features similar to PRISM's [5].

The origins of PRISM date back to Princeton University's earliest energy analyses of buildings, in the 1970s [6, 7]. In its current form, PRISM differs from other approaches in several important ways: in its *physical foundation*, which allows a physically meaningful interpretation of the results; in its emphasis on *reliability*, particularly of the NAC index, which in general is extremely well determined; in its *standardized output*, which facilitates comparisons across programs, and its accurate *error diagnostics* attached to all the estimates it produces; in its *availability*, to a wide variety of potential scorekeepers; and, finally, in its objective of *generality*, to all fuel types and to a wide range of building types and climates.

We define the word "scorekeeping" to mean the measurement of actual energy savings. PRISM is thus a particular scorekeeping method. Its purpose is to describe the past rather than predict the future. A static model, PRISM is not appropriate, as some dynamic models are, for the management of a building — to schedule thermostat setbacks, for example. On the other hand, these dynamic models are overly complicated for scorekeeping, which requires only long-term averages of consumption, i.e., data that

are readily available for large numbers of houses, or buildings in general. PRISM is designed to be a scorekeeping tool that makes the best possible use of such data.

TOWARD A STANDARDIZED APPROACH

Until recently, the haphazard array of approaches used to evaluate retrofit programs has made it impossible to compare savings from one program to another, or to aggregate the effects across programs. When the first "scores" came in from selected RCS and weatherization programs, many of them were disappointing [8]; yet, the lack of a coordinated approach has made it impossible to learn from mistakes or to plan for more effective programs in the future.

The progression of recent conferences testifies to the increased commitment to scorekeeping based on real data [9 - 14]. At an evaluation conference in Columbus, Ohio, held in 1982 (the first of its kind), many participants disputed the merits of billing data and argued the success of their programs either on the basis of number of participants or from engineering estimates of the energy saved, rather than from knowledge of actual savings achieved [9]. Since then, especially at biennial summer studies in Santa Cruz, California, the discussion of evaluation has shifted from "whether" to "how" to use real data [11, 12]. Now, as seen at the 1985 evaluation conference in Chicago, there is agreement on the importance of a standardized approach for measuring energy actually saved, and, among many, consensus on PRISM as the method of choice [14]. (Two of the several PRISM-based evaluations reported at the conference are summarized in refs. 15b and 15c, in this issue.) The seemingly inevitable, and occasionally embarrassing, shortfall of actual savings relative to engineering estimates is now part of the common experience. Further, the availability of a well tried method for measuring these savings is allowing the concern to shift to broader issues: how to choose the control group, how to use the savings estimates to evaluate a program's cost-effectiveness, and what conservation lessons can ultimately be learned from comparisons across programs.

Adjustment for the performance of a group of untreated, "control" houses can be an important part of scorekeeping, when it is desirable to decouple the savings induced by the measures of interest from the savings that would otherwise have occurred due to external events (such as increased energy prices). Evidence of extensive and continuing conservation over the decade since the Arab oil embargo, in the population at large, confirms the importance of adjusting the savings by a control group (see ref. 15p, in this issue). PRISM applied to both treatment and control houses, as shown in Fig. 1(b), gives a measure of control-adjusted as well as weather-adjusted savings for the treatment group. The analysis can then be updated for succeeding years, to track the durability of the savings (see ref. 15c, in this issue).

Invariably, an evaluation of a conservation program ought to go beyond the PRISM analysis, to determine the cost-effectiveness of various tried approaches to conservation, for example, or to clarify the reasons why some households saved more than others. The savings estimates, along with other PRISM outputs, provide reliable input to such analyses. Thus the PRISM analysis depicted in Fig. 1 may be thought of as standardized scorekeeping, representing stage one of the evaluation, while subsequent analyses, limited by available data and shaped by the specific needs of the project being evaluated, constitute stage two.

In this special double issue, PRISM is presented as a standardized, easy-to-use approach which utilities, communities, researchers and entrepreneurs throughout the country can adopt for measuring energy savings. Fifteen applications of PRISM are reported, ranging from specific studies of the interpretation of PRISM parameters to full-scale evaluations of retrofit programs [15].

A brief outline of these papers is given at the end of this introductory paper. We present now a description of the method: its physical rationale, the statistical procedure underlying it, and a sample savings analysis to illustrate its use. Our description here is fairly detailed; it is intended to serve as a reference for the other papers. The occasional variation on the method presented here is identified in the relevant paper.

SUMMARY OF THE METHOD

The Princeton Scorekeeping Method (PRISM) is a statistical procedure for calculating changes in energy consumption over time. For each house (or building) being analyzed, the procedure requires meter readings (or, for fuel oil, delivery records) for approximately one year in each period of interest. The consumption data are then corrected for the effects of weather, which of course is never the same for two different years, and also for differences in the time spanned by the different periods.

PRISM differs from other weather-normalization procedures in that the house's break-even temperature is treated as a variable, rather than a constant such as 18.3 °C (65 °F). Three physical parameters result from the model applied to the billing data for the heating fuel* of an individual house: *base-level consumption*, as a measure of appliance usage in the house; *reference temperature*, as a reflection of interior-temperature settings; and *heating slope*, as a measure of the lossiness of the house. Derived from these parameters, the NAC index is the reliable estimate of the consumption which would occur in a year of typical weather.

The physical basis for the model

Generally, whether for natural gas, oil or electricity, a house's heating system is first required when the outdoor temperature (T_{out}) drops below a certain level (the heating reference temperature τ), and for each additional degree drop in temperature a constant amount of heating fuel (the heating slope β) is required. Thus, the required heating fuel is linearly proportional to ($\tau - T_{out}$), and the proportional constant β represents the house's effective heat-loss rate. In addition, the house may use a fixed amount of the heating fuel per day (the base level α) in an amount independent of T_{out}. Formally, the expected fuel consumption per day, f, as illustrated in Fig. 2 for an idealized house, is given by

$$f = \alpha + \beta(\tau - T_{out})_+ \qquad (1)$$

*We use the word "fuel" to mean electricity as well as natural gas, fuel oil, or any external energy source.
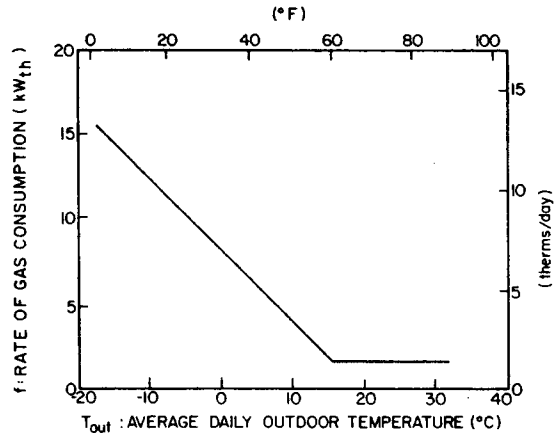
Fig. 2. Daily gas consumption (f) as a function of outdoor temperature (T_{out}), for a single idealized house. The house's energy signature is defined by a base level of 2.0 kW_{th} (equivalent to 1.7 therms/day in natural-gas units), a heating slope of 480 W_{th}/°C (0.22 therms/°F-day), and a reference temperature of 15.6 °C (60 °F).

where the term in parentheses is the heating degree-days h to base τ, i.e., $h(\tau)$, and the "+" indicates zero if the term is negative. This relationship is derived in Appendix 1.

The derivation of eqn. (1) leads to a simple physical interpretation for each of the three parameters. The reference temperature τ, which will vary from house to house, is likely to be influenced primarily by the indoor temperature T_{in} (which may be regulated by a thermostat setting)* and, in addition, an offsetting contribution from intrinsic gains (i.e., heat generated by appliances, occupants, and the sun). The heat-loss rate β depends on the conductive and infiltration heat losses, and, inversely, on the furnace efficiency, while the base level α represents the fuel requirements of appliances (including lights, for electricity, and the water heater if fueled by the heating fuel).

If τ is not accurately determined, or if it changes significantly over the time periods studied, the error or change in τ will inversely affect α, and β as well. Figure 3 illustrates this for the idealized house by plotting f vs. $h(\tau)$ for one correct and two incorrect values of τ. A straight-line fit through each set of points will have a different slope and

*In a large centrally heated building, the main boiler may be directly controlled by the outdoor temperature rather than a thermostat, thus replicating the constant-τ assumption of PRISM. Such a building is the subject of a paper in this issue [15f].
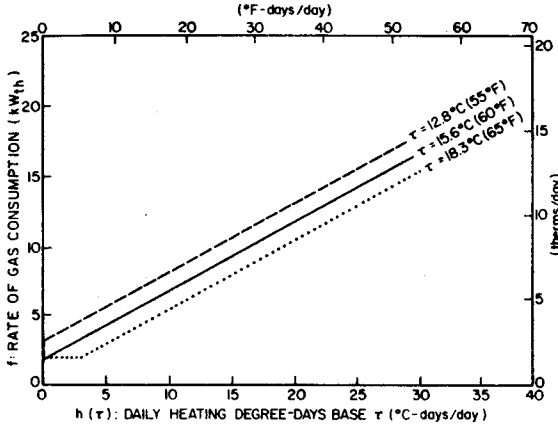
Fig. 3. Daily gas consumption (f) as a function of degree-days base τ [$h(\tau)$], for a single idealized house. The three curves correspond to the same consumption and temperature data, with degree-days calculated to different bases τ.

intercept. Therefore, an assumed (incorrect) reference temperature, such as the value of 18.3 °C so commonly used, is likely to lead to less physically meaningful values of the base level and the heat-loss rate.

Individual-house analysis

Based on this physical interpretation, the two data requirements for the analysis are actual meter readings (approximately monthly), from which consumption is calculated, and local average outdoor temperatures (daily), from which heating degree-days to different reference temperatures are computed in exact correspondence to the consumption periods. Generally, the most reliable results are obtained from a year's data [15n]. The input to the procedure is then F_i and H_i where:

F_i = average daily consumption in time interval i

$H_i(\tau)$ = heating degree-days per day computed to reference temperature τ in time interval i.

Here F_i is computed as the consumption in interval i divided by N_i, the number of days in that interval, and $H_i(\tau)$ is computed from T_{ij}, the average daily outdoor temperature for the jth day of interval i, over N_i days, i.e.,

$$H_i(\tau) = \sum_{j=1}^{N_i} (\tau - T_{ij})_+ / N_i \qquad (2)$$

The set of data points $\{F_i\}$ and $\{H_i\}$ for an approximately year-long period are then fit to a linear model:

$$F_i = \alpha + \beta H_i(\tau) + \epsilon_i \qquad (3)$$

where ϵ_i is the random error term. For a guessed value of reference temperature τ, the base-level and heating-slope parameters α and β are found by standard statistical techniques (ordinary least-squares linear regression). Using an iterative procedure based on Newton's method [16], "best τ" is found as the value of τ for which a plot of F_i vs. $H_i(\tau)$ is most nearly a straight line. Formally, τ is determined as the value for which the mean-squared error is minimized, or equivalently for which the R^2 statistic is highest. The corresponding values of α and β are the best estimates of base level and heating slope*.

The application of PRISM to real data is illustrated in Fig. 4, for a gas-heated house. The gas consumption data, F_i, plotted against time in Fig. 4(a), fall into a very straight line in Fig. 4(b) when plotted against heating degree-days H_i computed to best τ, the reference temperature determined by the model. The complete PRISM results for this house and this period are given in Table 1. At 0.985, the R^2 statistic indicates a very good straight-line fit, corresponding to the line drawn in Fig. 4(b).

The house's index of consumption for its heating fuel, NAC (Normalized Annual Consumption), is obtained from the model parameters, α, β and τ, applied to a long-term (say, ten-year) annual average of heating degree-days. NAC is calculated as follows:

$$\text{NAC} = 365\alpha + \beta H_o(\tau) \qquad (4)$$

where $H_o(\tau)$ is the heating degree-days (base τ) in a "typical" year. Once a normalization period is established, the values of H_o over the range of possible τ require a one-time cal-

---

*The SI units we recommend for PRISM parameters are: kW for α, W/°C for β, °C for τ, and GJ/year for NAC and other annual consumption estimates. Fuel-resource energy (for natural gas and oil) and site electrical energy are differentiated by the subscripts "th" for thermal and "elec" for electrical, respectively. The corresponding imperial units (therms for natural gas, etc.) are given in the list of conversion factors in the Foreword to this issue.

HOUSE : T 120 PRE
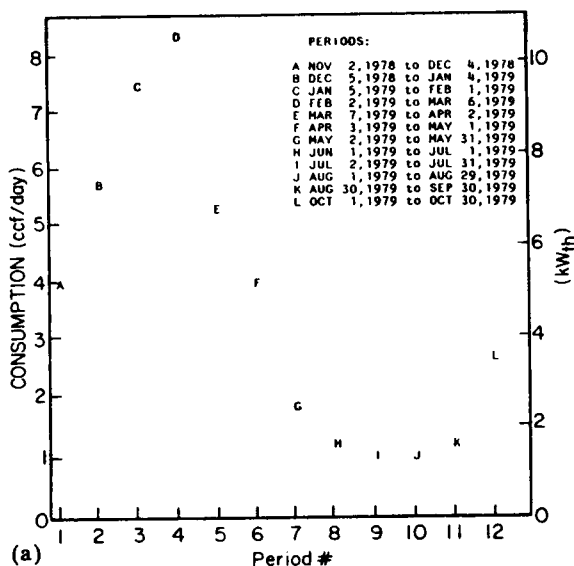
(a)
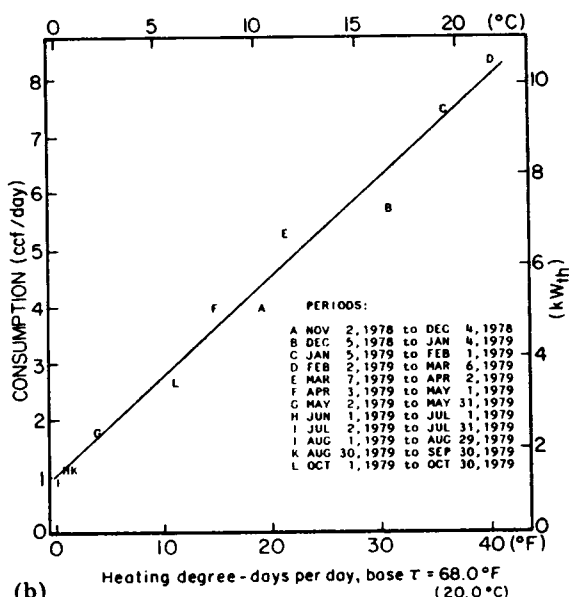
(b)

Fig. 4. Plots of consumption data (F_i) for sample gas-heated house in New Jersey. In (a) F_i is plotted against i, the (monthly) time period; in (b) F_i is plotted against $H_i(\tau)$, i.e., heating degree-days computed to best τ determined by PRISM. The straight line is the least-squares fit of the data, giving the results shown in Table 1.

culation only for each weather station used*. For standardization, we recommend a consistent normalization period; we use the twelve-year period from 1970 through 1981.

*Daily temperature data to compute H_i for each period and H_o for the normalization period, to any integer value of τ, are available from the National Weather Service for the appropriate weather station [17]. Degree-days to non-integer τ are found by linear interpolation.

TABLE 1

Sample PRISM results for a gas-heated house*

τ	= 20.0 (±1.5) °C
	[68.0 (±2.8) °F]
α	= 1.12 (±0.33) kW$_{th}$
	[0.90 (±0.26) ccf/day]
β	= 400 (±30) W$_{th}$/°C
	[0.18 (±0.01) ccf/°F-day]
βH_o	= 107.1 (±9.1) GJ$_{th}$/year
	[996 (±84) ccf/year], or 75% of NAC
NAC	= 142.5 (±4.0) GJ$_{th}$/year
	[1324 (±37) ccf/year]
R^2	= 0.985

*The sample gas-heated house is house T120 from the Modular Retrofit Experiment [15a]. The estimates are derived from PRISM applied to the pre-retrofit consumption data shown in Fig. 4. Each number in parentheses is the standard error of the estimate.

Reliability of the estimates

In general, the NAC estimate provides a reliable consumption index from which energy savings and conservation trends may be accurately estimated. The small standard error of NAC for our sample house, at 3% of the estimate, is typical of PRISM results. On the other hand, the three parameters, α, β and τ, which define a house's energy signature, are less well determined, as is confirmed by the standard errors for the sample house (Table 1) as well as by other studies in this issue (see Table 2 of ref. 15k). As a result, the parameters' changes over time are often difficult to interpret due to the interference of physical and statistical effects.

The stability of the NAC index is evident in Fig. 5, which shows, for the sample house, the progression of NAC as the estimation year is slid forward one month at a time. The drop in consumption after the retrofit is evident. Note the larger standard errors of NAC for the periods falling between the pre- and post-retrofit periods. In general, NAC is quite insensitive to exactly which months are included. (The gap in the plot reflects an inevitable characteristic of real-world data sets, namely, estimated or missing readings.) The analogous plots for the individual parameters demonstrate the temporary instability as the estimation window passes through the retrofit period, and thus the importance of excluding the retrofit period from the estimation periods used for scorekeeping (see examples in ref. 15f, in this issue).
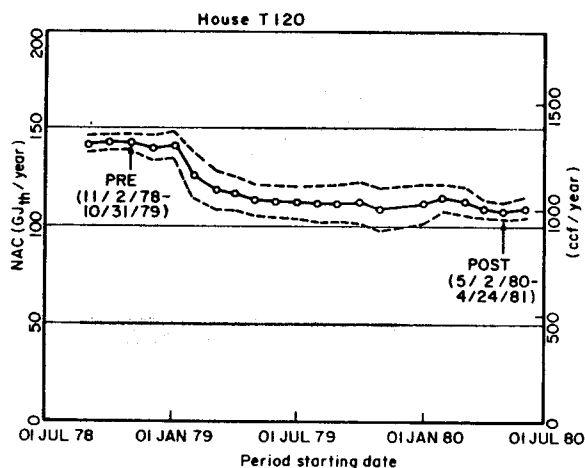
Fig. 5. Illustration of sliding PRISM, in which the one-year estimation period, starting on the date indicated, is moved forward one month at a time. NAC estimates are shown, for the sample gas-heated house used for Fig. 4 and Table 1. Dashed lines give standard errors of the estimate. Pre- and post-retrofit periods used for scorekeeping are indicated.

When a continuous series of consumption data is available, a sliding analysis such as this one can be a powerful tool not only for selecting the final estimation periods, but also for flagging anomalies in the data, and, more generally, for monitoring gradual changes in consumption. This technique has been used in a systematic study of the stability of PRISM parameters (see ref. 15n, in this issue).

While it is tempting to attribute a change in α to water-heater wrap or more efficient appliances, for example, or a drop in $\beta H_o(\tau)$ to added ceiling insulation or other measures to tighten the structure (through β) or to lowered thermostat settings (through τ), such simple physical inferences from individual-house results are usually not valid (see Appendix 1). For one thing, events affecting only one parameter are not likely to occur in isolation. For another, PRISM's assumptions that α, β and τ are constant on average from month to month over a year-long estimation period are not likely to hold perfectly in any real house. The change in the temperature-dependent component of NAC, $\beta H_o(\tau)$, is usually better determined than the change in β or τ individually, just as NAC is better determined than either that component or the base-level component, 365α. Nevertheless, the effect of the seasonal variability of appliance and water-heater usage, investigated

elsewhere in this issue [15m], interferes with the interpretation of these components.

As is well demonstrated by the studies reported in this special issue, NAC is a reliable and stable *index* of consumption. At best, the other PRISM parameters provide physically meaningful *indicators*, whose changes may not be statistically significant but whose behavior can often suggest the reason for a consumption change. The need for careful interpretation of these indicators is an important theme of this issue. Accurate standard errors for all the parameters [α, β, τ, $\beta H_o(\tau)$, and NAC] are part of the standard PRISM output. Developed for this model, the "composite" method for estimating the errors includes the uncertainty in the estimation of τ as well as the estimation error from fitting eqn. (3) [16]. It turns out that α and β are much more sensitive to variations in τ than is NAC (see Fig. 2 of ref. 15k). Even in extreme cases when one or more of the parameters is poorly determined, the standard error of NAC is usually only 2 - 4% of the estimate*. This stability of NAC is PRISM's most important feature.

Estimation of group savings

The NAC estimate provides the basic index for measuring energy savings, in groups of houses from one to thousands. Computed as the change in NAC between two periods of interest, the savings estimates are weather-adjusted, and thus are independent of changes in the weather between the two periods.

When adjustment by a control group is needed, an ideal control group is one constructed by random selection of participants from a larger set, where some or all of those not selected for treatment become the controls. (This approach is used in ref. 15a.) Often such advanced planning is not possible. A less ideal though generally adequate pro-

---

*NAC can be reliable even in the event of an extreme anomaly, for example, when best τ is established at the highest value of daily temperature for the estimation period. (The associated standard error of τ is infinite.) Only two such anomalies occurred in the data set of 276 cases from which the example in Fig. 4 was taken [15a]. For each of the two cases, only six (bimonthly) data points were available. In both cases, NAC was well determined: the standard errors of NAC were 3.0% and 6.0% of the estimate, and the corresponding R^2 values were 0.99 and 0.94, respectively.

cedure is to match non-participants to participants after the fact, so that the control and treatment groups have similar profiles, defined, for example, by energy consumption (i.e., pre-retrofit NAC), energy prices, household size and income, and house area. Another possible, and less cumbersome, alternative is to make the aggregate of the utility serving the retrofitted houses into a surrogate control group (see Table 7 of ref. 15a, and ref. 15p).

The scorekeeping procedure presented here includes both weather and control adjustments. Using billing and weather data for approximately year-long periods before and after (and not including) the period during which the retrofits were performed, PRISM is applied to each control and treatment house included in the program. From the resulting NAC_{pre} and NAC_{post} estimates, representing respectively a house's NAC for the pre- and post-retrofit periods, the raw, weather-adjusted savings for each house is then computed as:

(absolute) $S_{raw} = NAC_{pre} - NAC_{post}$ (5a)

(percent) $S_{raw, \%} = (1 - NAC_{post}/NAC_{pre})$
$$\times 100 \qquad (5b)$$

From the individual-house estimates, average values (medians or means*) are calculated for each group: $NAC_{pre}(T)$, $NAC_{post}(T)$, $S_{raw}(T)$ and $S_{raw, \%}(T)$ for the treatment group, and $NAC_{pre}(C)$, $NAC_{post}(C)$, $S_{raw}(C)$ and $S_{raw, \%}(C)$ for the control group. The savings for the treatment group (or for an individual treated house) may then be adjusted by the control houses, as described in Appendix 2, to give $S_{adj}(T)$ and $S_{adj, \%}(T)$.

It is important to know the errors associated with the various savings estimates. For a group of houses, the standard error of the median provides a robust measure of whether the savings in the treatment group(s) are distinguishable from the savings in the control group. For individual houses, the standard error of each savings estimate is readily computed from the standard error of NAC for each house, as given in Appendix 3.

---

*Although either median or mean values of NAC may be used, we recommend the median as the more 'robust' (i.e., insensitive to outliers) measure of the center of a group's distribution, and the standard error of the median as the measure of its accuracy (see Appendix 3).

SAMPLE SCOREKEEPING ANALYSIS

Our sample house (Table 1 and Figs. 4 and 5) is one of the 58 "house doctor" houses in the Modular Retrofit Experiment (MRE), a collaborative project between Princeton University and the natural gas utilities in the New Jersey area (see ref. 15a). The control group consisted of 40 additional houses. To illustrate the scorekeeping approach, we start from the savings estimated for the single house and continue through the computation of control-adjusted savings for the entire house-doctor group.

As indicated in Fig. 5, for the pre- and post-retrofit periods indicated, NAC for the sample house dropped from $NAC_{pre} = 142$ (± 4) GJ_{th}/year to $NAC_{post} = 107$ (± 4) GJ_{th}/year. The resulting raw savings were*:

$S_{raw} = 35(\pm 6)GJ_{th}$/year

or, relative to NAC_{pre}

$S_{raw, \%} = 25(\pm 4)\%$

The small standard errors in the savings indicate that the savings were significant.

This house saved more than the average house in the house-doctor group, for which the median savings were*:

$S_{raw}(T) = 21(\pm 3)GJ_{th}$/year

or, relative to NAC_{pre}

$S_{raw, \%}(T) = 15(\pm 2)\%$

The median savings in the control group were considerably lower, though far from negligible:

$S_{raw}(C) = 14(\pm 3)GJ_{th}$/year

or,

$S_{raw, \%}(C) = 10(\pm 1)\%$

The median control-adjusted savings for the house-doctor group, computed by the procedure described in Appendix 2, were:

$S_{adj}(T) = 9(\pm 2)GJ_{th}$/year

or,

$S_{adj, \%}(T) = 8(\pm 1)\%$

---

*The number in parentheses for the individual-house savings is the standard error of the corresponding estimate. The number in parentheses for the group savings (T or C) is the standard error of the corresponding sample median (see Appendix 3).

As has been the case in other retrofit programs, the control adjustment substantially deflates this experiment's raw savings. Nevertheless, the savings in the house-doctor group relative to the control group were highly statistically significant (see ref. 15a).

OTHER APPLICATIONS

The accuracy of the estimates from our sample analysis is typical of other applications of PRISM, both to gas-heated and oil-heated houses and to electrically heated houses without cooling. Summaries of model performance are found elsewhere in this issue (for example, refs. 15a and 15b for natural gas, ref. 15k for oil, and refs. 15c and 15h for electricity). For houses in heating-dominated climates, we have found R^2-values to average 0.97, and standard errors for NAC to average 3 - 4% of the NAC estimate. Even in the face of some anomalies in the individual-parameter results, NAC and the corresponding savings estimates are usually stable and reliable.

The model used in the above example, and in almost all papers in this issue, is the "heating only" PRISM model for individual houses. Two adaptations of this model have also been developed, for individual houses with electric cooling as well as heating [15h], and for large aggregates of gas-heated houses for which only total utility sales data are available [15p]. For the former, cooling analogues of β and τ are added to the model in eqn. (3). For the latter, a variation of $H_i(\tau)$ in the same equation is used to account for the billing lag. For both adaptations, NAC is on average as well determined as it is in individual-house, heating-only applications.

OVERVIEW OF THE ISSUE

PRISM was first developed for our own buildings research program. Its 1982 application to the Modular Retrofit Experiment is presented in this special issue as a prototype PRISM-based evaluation [15a]. Since then, PRISM has been widely applied to other groups of single-family houses. The Statistics Laboratory at the University of Wisconsin has used PRISM for the evaluation of Wisconsin's low-income weatherization program involving 1000 houses [15b]. Researchers at Oak Ridge National Laboratory are using PRISM as stage one of a two-stage approach, to evaluate RCS and other utility conservation programs, such as Bonneville Power Administration's Residential Weatherization Pilot Program [15c]. The Center for the Biology of Natural Systems at Queens College has applied PRISM to a smaller sample of houses in New York City, for a detailed comparison of two approaches to low-income weatherization [15d]. The method is being used extensively in Minnesota to monitor the success of a variety of city and state programs; using PRISM, the Minneapolis Energy Agency has carried out a definitive comparison of predicted vs. actual savings from RCS retrofits [15e].

Recently, researchers have begun to recognize the almost untapped resource of energy savings in multifamily buildings. In apartments in New Jersey, for example, the average energy usage per unit floor area may be double what it is in single-family houses, in spite of potential benefits from common walls in apartments*. As part of its shift in research emphasis from single-family to multifamily buildings, the Center for Energy and Environmental Studies at Princeton has extensively instrumented a 60-unit gas-heated apartment building in Asbury Park, New Jersey [15f]. High interior temperatures coupled with an unusual boiler configuration challenge the interpretability of the PRISM estimates; an engineering analysis of additional data provides an improved understanding of the results. Lawrence Berkeley Laboratory's study of a complex of apartment buildings in the San Francisco Housing Authority offers another test of the applicability of PRISM to large multifamily buildings [15g].

---

*For example, the per-unit NAC in a 126-unit gas-heated apartment complex in New Jersey gave 3.5 GJ_{th}/m^2 before a major retrofit and 1.7 GJ_{th}/m^2 after it [18], vs. 0.9 GJ_{th}/m^2 [15p] for the average gas-heated customer in the state. (The comparison assumes an average area of 150 m^2 per house, vs. 65 m^2 measured for the apartment complex.) The 60-unit building studied in this issue showed a similarly high average NAC per unit area [15f].

Whereas the methodology development initially emphasized gas-heated houses, special problems relating to other fuels have been the focus of recent research. Analysis of electrically heated houses can be confounded by electric cooling, even in a heating-dominated climate [15h], or by the presence of a heat-pump system which, to some extent, violates the assumptions underlying PRISM [15i]. Otherwise, gas and electricity have much in common: the data bases of monthly (or bimonthly) meter readings are equally accessible, and the seasonal dependence of non-heating consumption has a similar effect on the PRISM parameters for both fuels [15m]. Further, the effect of supplemental heating by wood on a PRISM analysis of the consumption of a conventional heating fuel is likely to be independent of whether the fuel is gas or electricity; the effect is explored here for electrically heated houses in the Portland, Oregon, region [15j]. Oil heating poses a new set of problems, not the least of which is infrequent, unevenly spaced deliveries [15k]. For any fuel, sufficent data over a year or more are needed for PRISM to work reliably; a systematic study of the stability of the model parameters provides some guidelines concerning PRISM data requirements [15n].

Often, anomalies that occur for individual houses are no longer evident in aggregated PRISM results. One short-cut aggregate approach is to apply a variation of PRISM to total utility sales data for residential heating customers (gas or electricity); extensive analysis of gas-heating customers in New Jersey has yielded promising results [15p]. Taking the analysis a step further, for the same data set and for another group of houses, the relative roles of two possible sources of conservation — shell tightening and lower thermostat settings — are inferred from the PRISM analyses [15q].

In all of these studies, NAC emerges as an extremely reliable index of consumption. The other PRISM parameters provide useful indicators of the components of NAC, but they require a sensitive interpretation with a careful consideration of their errors.

The fifteen papers presented in this special double issue provide convincing evidence that a simple method applied to whole-building billing data can become a powerful consumption monitoring, or scorekeeping, tool. The papers report progress on a particular method, PRISM. It would be naive to expect all houses, and especially the people occupying them, to obey the simple principles embodied by this method. Nevertheless, the truth told by actual meter readings, the basis of PRISM, cannot be ignored. The success of the studies thus far confirms that PRISM, though not without room for improvement, is a particularly useful way of extracting scorekeeping information from billing data.

FUTURE DIRECTIONS

The papers in this special issue emphasize applications of PRISM to conventional housing in heating-dominated climates; for climates in which the energy used for cooling rather than heating dominates, and for houses with a large solar component in their design, more research is needed. The studies in this issue focus on the fuel (gas, electricity, oil) used for space heating; in future work, the method should allow for the interaction between fuels when more than one fuel is used in a house by its furnace and appliances ("total energy scorekeeping"). The statistical procedure used for PRISM analyses is based on least-squares regression; more robust techniques, under development, would reduce the influence of anomalous data and improve the reliability of the estimates. The data bases for the analyses are primarily energy bills; the extent to which the value of billing data will be enhanced by additional data available from instrumentation, such as submeters and temperature sensors, needs to be explored. Thus far, the studies demonstrate the applicability of PRISM at two levels of analysis, to individual-house data for large samples of houses, and to utility aggregate data representing large fractions of the population; in between these two extremes, there may be additional strategies for dealing with large numbers of houses, such as clever statistical sampling of the houses being monitored, or substation or trunk-line metering to represent community-level consumption.

The primary objective of our current scorekeeping research is to realize the full potential of billing data for monitoring consumption in all climates and building types. The most

productive approach, for addressing these scorekeeping concerns and ultimately for learning about the effectiveness of conservation measures, will be studies of actual consumption data. We anticipate that the best research laboratory for these studies will continue to be real-world conservation projects.

APPENDIX 1

Derivation of physical model underlying PRISM

The space heating energy, E_h, required to maintain a house at temperature T_{in}, is proportional to the difference $(T_{in} - T_{out})$, where T_{out} is the outdoor temperature. The proportionality constant L represents the lossiness of the house. Thus, when $T_{out} < T_{in}$,

$$E_h = L(T_{in} - T_{out}) \qquad (A1)$$

The lossiness has two contributions, from air infiltration losses L_i and from transmission losses L_t, i.e., $L = L_i + L_t$*. Some of this heating is supplied by the house's intrinsic gains Q, representing heat gains from appliances, occupants and the sun**, and the rest by an amount of fuel f_h burned at efficiency η, i.e.,

$$E_h = \eta f_h + Q \qquad (A2)$$

Therefore, the required external fuel for space heating is

$$f_h = L(T_{in} - T_{out})/\eta - Q/\eta \qquad (A3)$$

which may be rewritten:

$$f_h = \beta(\tau - T_{out}) \qquad (A4)$$

where

$$\beta = L/\eta \qquad (A5)$$

---

*The transmission lossiness $L_t = \Sigma U_j A_j$, where A_j is the area of each exposed surface, and U_j is the corresponding transmission coefficient. To a good approximation, the infiltration lossiness $L_i = V\rho C_p$, where V is the volume flow rate of outdoor air entering the building, ρ is the density of air and C_p is the heat capacity of air; this ignores moisture-related heat loss (due to latent heat to evaporate water inside the house). See Chapter II of ref. 6, as well as the discussion in ref. 15f, in this issue.

**Our definition of intrinsic gains adds solar gains to the comprehensive list compiled by Shurcliff [19].

and

$$\tau = T_{in} - Q/L \qquad (A6)$$

Thus the house's reference temperature τ (the outdoor temperature below which external fuel is required for heating) is below T_{in}, by an amount proportional to the house's intrinsic gains.

If the heating fuel is also used for other purposes such as water heating, appliances, and (for electricity) lighting, at a rate α, then the rate at which heating fuel is consumed per day is given by:

$$f = \alpha + \beta(\tau - T_{out}) \qquad (A7)$$

for $T_{out} < \tau$. This is the relationship shown in Fig. 2, and corresponds with eqn. (1) in the text (see also refs. 6 and 7).

In a single-family house, the usual control system is a thermostat, which regulates the indoor temperature T_{in}. In this case, the constant-τ assumption of PRISM requires that several factors be constant from month to month: average indoor temperature T_{in}, average internal gains Q, and average house lossiness $L = \eta\beta$ (see eqn. (A6)). The constant-β assumption requires that L and η do not vary on average from month to month (see eqn. (A5)). The constant-α assumption requires non-varying energy usage for appliances, etc., fueled by the heating fuel.

Given these assumptions, several classes of interventions will induce predictable changes in α, β, and τ. Reduction of monthly average thermostat settings will decrease τ. Structural retrofits will affect β and τ: in an ideal house (seen through PRISM), a decrease in L will decrease both β and τ. An improved furnace efficiency η will also decrease β. A shift to more efficient appliances will lower α. However, by decreasing internal gains Q, this shift will increase τ (leaving β unchanged), and thus lead to an increased requirement for heating fuel that will partially offset the benefits from more efficient appliances. Any change in Q will affect τ: the addition of a household member might lower τ, for example, whereas the shift to a more efficient appliance (fueled by the heating or a non-heating fuel) will raise τ.

These theoretical expectations are valid for the ideal house such as was used in Fig. 2. When PRISM is run on real data, for which α, β and τ are not truly constant over any

estimation period, statistical covariance among the three parameters often interferes with simple associations between known interventions and the observed trends in the parameters. The problem is particularly acute when the periods of estimation include major changes. Pre-retrofit and post-retrofit periods should therefore be selected to exclude interventions wherever possible.

APPENDIX 2

Computation of group savings estimates

We let the notation $[X]|_T$ and $[X]|_C$ represent the median (or mean) of the set of values of the quantity X for the treatment (T) or control (C) group, respectively.

In analogy with the individual-house savings in eqns. (5a) and (5b), the raw, weather-adjusted savings for the treatment group is given, in absolute terms, by:

$$S_{raw}(T) = [NAC_{pre} - NAC_{post}]|_T \qquad (A8a)$$

and, in percent terms relative to NAC_{pre}, by

$$S_{raw,\%}(T)/100 = [1 - NAC_{post}/NAC_{pre}]|_T \qquad (A8b)$$

Using similar pre- and post-periods, raw savings for the control group are analogously given by:

$$S_{raw}(C) = [NAC_{pre} - NAC_{post}]|_C \qquad (A9a)$$

and

$$S_{raw,\%}(C)/100 = [1 - NAC_{post}/NAC_{pre}]|_C \qquad (A9b)$$

To adjust the savings in the treatment group by the control, we define a control-adjustment factor:

$$C_{adj} = [NAC_{post}/NAC_{pre}]|_C \qquad (A10)$$

Then the control-adjusted savings of the treatment group are obtained by the following:

$$S_{adj}(T) = [C_{adj} \times NAC_{pre} - NAC_{post}]|_T \qquad (A11a)$$

and

$$S_{adj,\%}(T)/100 = [C_{adj} - NAC_{post}/NAC_{pre}]|_T \qquad (A11b)$$

which can be simplified for a single treatment and control group as

$$S_{adj,\%}(T) = S_{raw,\%}(T) - S_{raw,\%}(C) \qquad (A11c)$$

These formulae apply to an individual treated house (i.e., to a treatment group of one) as well as to the entire group. (For the MRE results presented earlier and in ref. 15a, eqn. (A11b) was applied individually to each location (module), for which a separate C_{adj} was calculated.)

APPENDIX 3

Standard errors of savings estimates

The standard errors of the savings estimates are obtained from the standard errors of NAC_{pre} and NAC_{post}, i.e., $se(NAC_{pre})$ and $se(NAC_{post})$, which are computed by a method developed for PRISM [16] and are included in the standard output for each house analyzed. For an individual house:

$$se(S_{raw}) = [se^2(NAC_{pre}) + se^2(NAC_{post})]^{1/2} \qquad (A12)$$

$$se(S_{raw,\%})/100 = \{(NAC_{post})^2[se^2(NAC_{pre})]$$
$$/(NAC_{pre})^4 + [se^2(NAC_{post})]/(NAC_{pre})^2\}^{1/2} \qquad (A13)$$

where S_{raw} and $S_{raw,\%}$ are computed from eqns. (5a) and (5b), respectively.

When a group of houses is analyzed, the center of the distribution of the quantity X may be represented by either the mean or median value of X, i.e., by mean(X) or median(X). For each measure, there are corresponding measures of the width of the distribution of X: the standard deviation, sd(X), is generally used with mean(X), and the interquartile range, IQR(X), i.e., the length of the interval containing the middle 50%, with median(X).

The standard error of the sample mean, se[mean(X)], gives a measure of the variability of the sample mean. For a group of N houses, this is computed from sd(X) as follows:

$$se[mean(X)] = sd(X)/\sqrt{N} \qquad (A14)$$

In direct analogy with eqn. (A14), the standard error of the sample median may be computed from IQR(X):

$$se[median(X)] = IQR(X)/\sqrt{N} \qquad (A15)$$

This provides a measure of the variability of the sample median*. For a given quantity X, eqn. (A14) or (A15), respectively, tells how accurately the mean or median has been estimated for the larger group of houses from which the study group was drawn. Thus two alternative representations for the center of the distribution of X for a set of houses may be written:

$$\text{mean}(X) \pm \text{se}[\text{mean}(X)]$$

and

$$\text{median}(X) \pm \text{se}[\text{median}(X)].$$

In that they are more insensitive to outliers, the median measures (median value, interquartile range, and standard error of the sample median) are robust alternatives to the mean measures (mean value, standard deviation, and standard error of the sample mean). Since outliers may strongly influence the mean value, in an amount that may substantially distort the resulting representation of a group's savings (for $X = \text{NAC}_{\text{pre}} - \text{NAC}_{\text{post}}$), the median measures are usually more meaningful. On the other hand, mean measures are occasionally more convenient, since classical t-tests of significance are readily available for them. In addition, a comparison of the mean and median values is often useful, for obtaining a sense of the skewness of the distribution.

For scorekeeping, we recommend computation of both sets of measures for the quantities of interest. In general, we rely on the median measures, after they have been compared with the mean measures.

For the sample scorekeeping analysis presented in the text, the standard errors of the savings for the single gas-heated house were computed from eqns. (A12) and (A13); these represent measurement errors. The standard error of the sample median of the house-doctor group's savings, from eqn. (A15), was

used to represent a measure of the variation across houses. To some extent, the latter includes the effect of the measurement error for each house (eqns. (A12) and (A13)), which is generally, for all PRISM parameters, much smaller than the corresponding estimate's variation from house to house.

---

*The standard error of the median is a new quantity inferred from other work [20, 21]. Use of eqns. (A14) and (A15) for the standard error of the mean or of the median implicitly assumes that the errors in estimating a given quantity in different houses are uncorrelated. If this assumption is invalid, these formulae understate the uncertainty of their respective mean and median estimates, though not by much if (as is usually the case) the measurement errors are small relative to the house-to-house variation of the estimate.

REFERENCES

1 U. S. Department of Energy, Residential Conservation Service Program: Final Rule, *Federal Register, 47* (123) (June 25, 1982) 27752 - 27803; see also *National Energy Conservation Policy Act, Public Law 95 - 619*, Part 1 of Title II, 1978.

2 The Bradley Plan: Home energy conservation program, *Congressional Record, 125* (107) (July 31, 1979) 10950 - 53, Washington, DC; see also *Energy Security Act, Public Law 96-294*, Subtitle C of Title V, 1980.

3 Examples of early papers are: Highest economy in furnace heating: Proper temperatures, ventilation and coal consumption for different outside temperatures, *The Metal Worker, Plumber and Steam Fitter, 66* (Nov. 10) (1906) 47 - 49 ("For any given outside temperature there is a corresponding amount of heat that must be supplied in order to offset the heat losses through the walls and windows", p. 47); and R. P. Bolton, Heating requirements in New York weather, *Heating and Ventilation Magazine, VIII* (March) (1911) 1 - 10 ("An analysis of the weather conditions actually prevailing during the heating season in a city like New York may be made a very useful guide in determining the heating requirements, both for the whole period and for any given portion of the period", p. 1).

4 J. Wright (private communication), Gas Planning Department, Public Service Electric and Gas Company, Newark, NJ, 1985.

5 M. Fels, R. Socolow, J. Rachlin and D. Stram, *PRISM: A Conservation Scorekeeping Method Applied to Electrically Heated Houses*, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, *EPRI Report No. EM-4358*, Palo Alto, CA, December 1985, Appendix 1.

6 T. Schrader, *A Two-Parameter Model for Assessing the Determinants of Residential Space Heating, MS Thesis*, Department of Aerospace and Mechanical Engineering, Princeton University, *Report No. 69*, Center for Energy and Environmental Studies, Princeton, NJ, June 1978.

7 F. Sinden, A two-thirds reduction in the space heat requirement; and L. Mayer and Y. Benjamini, Modeling residential demand for natural gas as a function of the coldness of a month, in R. Socolow (ed.), *Saving Energy in the Home*, Ballinger Publishing Co., Cambridge, MA, 1978.

8 E. Hirst, Household energy conservation: a review of the federal Residential Conservation Service, *Public Administration Review, 44* (5) (1984) 421 - 430.

9 Battelle, Columbus Laboratories, *Workshop Proceedings: Measuring the Effects of Utility Conservation Programs, February 1982, Columbus, OH, EPRI Report No. EA-2496*, Palo Alto, CA, July 1982.

10 Synergic Resources Corporation, *Proc. Conference on Utility Conservation Programs: Planning, Analysis, and Implementation, October 1983, New Orleans, LA, EPRI Report No. EA-3530*, Palo Alto, CA, May 1984.

11 J. Harris and C. Blumstein (eds.), *What Works: Documenting Energy Conservation in Buildings*, Summary of the 1982 ACEEE Summer Study on Energy-efficient Buildings held August 1982 in Santa Cruz, CA, American Council for an Energy-Efficient Economy, Washington, DC, 1984.

12 *Doing Better: Setting an Agenda for the Second Decade, Proc. 1984 ACEEE Summer Study on Energy-efficient Buildings, August 1984, Santa Cruz, CA*, American Council for an Energy-Efficient Economy, Washington, DC, 1984.

13 *Proc. Weatherization 1985: Midwest Technical Conference, Chicago, IL, February 1985*, Argonne National Laboratory, Argonne, IL, 1985.

14 *Proc. Second National Conference, Energy Conservation Program Evaluation: Practical Methods, Useful Results, Chicago, IL, August 1985*, Vol. I (Session 4: PRISM Workshop, Session 5: Weatherization Program Evaluation, and Session 8: Using PRISM), Argonne National Laboratory, Argonne, IL, 1985.

15 *Energy and Buildings, 9* (1 and 2) this issue; papers in the order of their appearance:

a G. Dutt, M. Lavine, B. Levi and R. Socolow, The Modular Retrofit Experiment: design, scorekeeping and evaluation, pp. 21 - 33.

b M. Goldberg, A Midwest low-income weatherization program seen through PRISM, pp. 37 - 44.

c E. Hirst, Electricity savings one, two, and three years after participation in the BPA Residential Weatherization Pilot Program, pp. 45 - 53.

d L. Rodberg, Energy conservation in low-income homes in New York City: the effectiveness of house doctoring, pp. 55 - 64.

e M. Hewett, T. Dunsworth, T. Miller and M. Koehler, Measured vs. predicted savings from single retrofits: a sample study, pp. 65 - 73.

f J. DeCicco, G. Dutt, D. Harrje and R. Socolow, PRISM applied to a multifamily building: the Lumley Homes case study, pp. 77 - 88.

g C. Goldman and R. Ritschard, Energy conservation in public housing: a case study of the San Francisco Housing Authority, pp. 89 -98.

h D. Stram and M. Fels, The applicability of PRISM to electric heating and cooling, pp. 101 - 110.

i M. Fels and D. Stram, Does PRISM distort the energy signature of heat-pump houses?, pp. 111 - 118.

j M. Fels and D. Stram, The effect of burning wood on saving electricity, pp. 119 - 126.

k M. Fels, M. Goldberg and M. Lavine, Exploratory scorekeeping for oil-heated houses, pp. 127 - 136.

m M. Fels, J. Rachlin and R. Socolow, Seasonality of non-heating consumption and its effect on PRISM results, pp. 139 - 148.

n J. Rachlin, M. Fels and R. Socolow, The stability of PRISM estimates, pp. 149 - 157.

p M. Fels and M. Goldberg, Using the scorekeeping approach to monitor aggregate energy conservation, pp. 161 - 168.

q M. Goldberg and M. Fels, Refraction of PRISM results into components of saved energy, pp. 169 - 180.

16 M. Goldberg, *A Geometrical Approach to Nondifferentiable Regression Models as Related to Methods for Assessing Residential Energy Conservation, Ph.D. Thesis*, Department of Statistics, Princeton University, Report No. 142, Center for Energy and Environmental Studies, Princeton, NJ, 1982.

17 National Oceanic and Atmospheric Administration (NOAA), *Local Climatological Data — Monthly Summary [name of weather station]*, National Climatic Data Center, Asheville, NC, monthly publication.

18 R. Socolow, *Field Studies of Energy Savings in Buildings: a Tour of a Fourteen-Year Research Program at Princeton University, Report No. 191*, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, July 1985; presented at the Soviet-American Symposium on Energy Conservation, Moscow, June 6 - 12, 1985; to be published in *Energy*.

19 W. Shurcliff, *Superinsulated Houses and Double-envelope Houses: A Survey of Principles and Practice*, Brick Publishing Co., Andover, MA, 1981, p. 9.

20 E. Parzen, A density-quantile function perspective on robust estimation, in R. Launer and G. Wilkinson (eds.), *Robustness in Statistics*, Academic Press, New York, NY, 1979.

21 C. Hurvich, Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, private communication, November 1985.